analysis in severe acute bipolar depression. Although this positive subgroup analysis may be valid, or perhaps not,[13] one is still faced with the inability to simply state the overall negative outcome as negative. Furthermore, the meta-analysis did not include the other negative studies in acute mania, rapid cycling or acute unipolar depression. Those studies remain unpublished in any form.

## CURRENT STATUS

The recent summary paper[11] includes five acute bipolar depression studies while my reading of the www.gsk.com website found only three when I accessed that site on 1 March 2008. In seeking to update the table with those new studies, I revisited the GSK website (accessed 7 May 2009) but I was unable to find the previous data registry of clinical trials at all; if it still exists, it is certainly hard to find. A visit to the NIMH www.clinicaltrials.gov website (accessed 7 May 2009) has 40 studies with lamotrigine in bipolar disorder but the majority are not sponsored by GSK and none of the studies listed in table 1 can be found at the NIMH website.

Thus despite claims that the pharmaceutical industry is taking this matter seriously, access to scientific research results seems to be taking a continual back seat to the patent protection claim of proprietary data; public health appears to be second in priority to capitalism.

## CONCLUSIONS

Evidence based medicine—or, more simply put, the science of medicine—cannot be taken seriously, and is certainly not valid, if the evidence base is only partial. The scientific literature currently is like an under cooked meal which we think is ready to eat. We never know whether what we see in the evidence is correct or biased in one direction or the other. Meta-analyses of large published datasets are not as meaningful as they seem when unpublished data languish elsewhere. Statistical tests for publication bias can only provide some sense of the problem; the real solution is to ask the question first, to not presume that our evidence base is anywhere near complete and, in contrast with our experience, to publish critical reviews of unpublished negative studies, rather than setting such a high bar on such reviews that they inevitably fail to make it to print.

**Competing interests:** In the past year, SNG has received a research grant from Pfizer and honoraria from GlaxoSmithKline, AstraZeneca, Pfizer and Bristol-Myers Squibb.

## REFERENCES

1. **Galbraith J.** *The Great Crash of 1929*. New York: Mariner Books, 1997.
2. **Dickersin K.** The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;**263**:1385–9.
3. **Pinel P.** *A treatise on insanity*. New York: Hafner Publishing, 1962.
4. **Turner EH,** Matthews AM, Linardatos E, *et al*. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;**358**:252–60.
5. **Pande AC,** Crockatt JG, Janney CA, *et al*. Gabapentin in bipolar disorder: a placebo-controlled trial of adjunctive therapy. Gabapentin Bipolar Disorder Study Group. *Bipolar Disord* 2000;**2**:249–55.
6. **Calabrese J,** Bowden C, Sachs G, *et al*. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder. *J Clin Psychiatry* 2003;**64**:1013–24.
7. **Bowden C,** Calabrese J, Sachs G, *et al*. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Arch Gen Psychiatry* 2003;**60**:392–400.
8. **Shorter E.** *Before Prozac: The troubled history of mood disorders in psychiatry*. New York: Oxford University Press, 2009.
9. **Ghaemi SN.** All mixed up: on the absence of diagnostic guidelines for mixed states in the ISBD Diagnostic Guidelines Task Force Report. *Bipolar Disord* 2008;**10**:129–30.
10. **Bowden CL,** Asnis GM, Ginsberg LD, *et al*. Safety and tolerability of lamotrigine for bipolar disorder. *Drug Saf* 2004;**27**:173–84.
11. **Calabrese JR,** Huffman RF, White RL, *et al*. Lamotrigine in the acute treatment of bipolar depression: results of five double-blind, placebo-controlled clinical trials. *Bipolar Disord* 2008;**10**:323–33.
12. **Feinstein AR.** Meta-analysis: statistical alchemy for the 21st century. *J Clin Epidemiol* 1995;**48**:71–9.
13. **Pocock SJ,** Assmann SE, Enos LE, *et al*. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med* 2002;**21**:2917–30.

# Understanding confounding and mediation

Michael A Babyak

In both experimental and observational studies, many researchers attempt, often implicitly, to identify causal relations among variables. In trying to understand the possible causal processes that might have generated their data, the concepts of *confounding* and *mediation* play a prominent role. The two phenomena are often confused, and indeed are not always readily distinguishable. In the present paper, I will present a brief, somewhat simplified, introduction to confounding and mediation. I will present basic defining criteria, how to distinguish the two and also the problem of cases in which the distinction is not clear, along with some final caveats.

**Correspondence to:** Dr Michael A Babyak, Duke University Medical Center, Box 3119 DUMC, Durham, NC 27710, USA; michael.babyak@duke.edu

## CONFOUNDING

The term *confound* arises from the Latin *confundere*, to pour together or mix.[1] The English word *confuse* arises from the same Latin root (http://www.merriam-webster.com/). In the context of empirical research, the term confounding is most often encountered in situations where some "predictor" of interest, let's call it *a*, is presumed to be associated causally with some outcome, say, *c*. However, there may an additional variable, *b*, that also is associated with the predictor of interest, *a*, and the outcome, *c*. In the broadest application of the term, the effects of *a* and *b* are said to be confounded—that is, mixed. (There are in fact several ways in which the term confounder or confounding are currently used in research methodology but for our purposes here we will focus on the most common one[2].) This, however, is only a very broad use of the term. In order to distinguish it from mediation, we will use the more specific definition offered by methodologists, which is as follows. Confounding is present when the following conditions occur: (1) both the predictor of interest and the potential confounder must be associated with the outcome (*a* and *b* are related to *c*); (2) the predictor of interest and the confounder must be associated (*a* and *b* are associated); (3) the confounder is not a presumed causal consequence of the predictor (*b* cannot be caused by *a*). When these conditions are met, the effects of the predictor and the confounder, *a* and *b*, on the outcome, *c*, are confounded. As a corollary to the above criteria, adjusting

for the effect of the confounder, either by partialling or stratification, will reduce or eliminate the effect of the predictor of interest. In a regression context, the parameter estimate for the effect of *a* on *c* will be attenuated when *b* is entered into the same equation.

### Example of confounding

Let us look at a real world example of confounding. We will use one of the examples used by Rubin in a paper on propensity scoring.[3] A large observational study (Rubin adapted these data from Cochran[4]) examining the relation between the type of tobacco use and rates of cancer mortality showed, contrary to what we might expect, that individuals who smoked either a pipe or cigar had higher cancer death rates (35.5%) than individuals who smoked cigarettes (20.5%). What's going on here? Especially with observational data, where there is no control via randomisation, a good scientist will immediately ask the questions implied by the definition of confounding. Is there a variable that might be associated with both tobacco type and cancer mortality? If so, is that variable NOT in the suspected causal sequence between tobacco type and cancer mortality? A different way to pose this question is "are there differences between those who smoke cigarettes and those who smoke pipes or cigars—beyond the fact that they smoke different forms of tobacco—that might account for the difference in mortality?" One very likely such difference is age. At least in the present era, pipe and cigar smokers tend to be, on average, older than cigarette smokers. Thus it may be that the relation between tobacco type and mortality is spurious, existing only by way of the fact that tobacco type is related to age. Indeed, in the Canadian study, the average age of the pipe smokers was nearly 70 years while the cigarette smokers were only about 51 years old on average. Let us walk through the confounding criteria. Tobacco type and age are apparently associated with mortality; age and tobacco type are associated; age is not in the causal sequence between tobacco type and cancer mortality—that is, the type of tobacco a person uses cannot possibly cause chronological age. So, it looks like tobacco type and age are confounded. As additional evidence, it turns out that after adjusting for age, the magnitude of the relation between the original predictor, tobacco type, and the outcome, death, is diminished. In fact, after adjustment for age (Rubin uses Cochran's[4] stratification method rather than partialling in a regression model to make this adjustment), the difference in mortality rates between the two tobacco types vanishes. In fact, the difference is reversed, with the age adjusted mortality rate of almost 30% in the cigarette group compared with just under 20% in the pipe/cigar group. Thus accounting for the confounding effect of age reveals what is likely the "real" effect of tobacco type on survival. The original, unadjusted relation between tobacco type and mortality was merely an artefact produced by the correlation between age and tobacco type.

Thus in carrying out any attempt to estimate the relation between a putative predictor and an outcome, one of the primary questions to ask is whether one or more variables exist that might meet the criteria for confounding. Hopefully, this question was asked at the design stage of the study and, equally important, the potential confounders were measured and available for analysis.

## MEDIATION

Mediation very closely resembles confounding. In the broadest sense of *confusing* or *mixing*, a mediator is identical to a confounder. Mathematically, there is literally no difference at all. In fact, with one important exception, the criteria for mediation can be described using the criteria for confounding and simply substituting the word *mediator* for *confounder*. The important exception occurs on step 3, which concerns the presumed causal relation among the variables. The criteria for mediation are: (1) both the predictor of interest and the potential mediator must be associated with the outcome (*a* and *b* are related to *c*); (2) the predictor of interest and the mediator must be associated (*a* and *b* are associated); (3) the mediator is a presumed causal consequence of the predictor (*a* causes *b*). Again, partialling, or adjusting for the effect of the mediator, will reduce or eliminate the effect of the predictor of interest. Note that the only difference between the confounder and mediator criteria is in step 3—the mediator must be a presumed causal consequence of the variable of interest. The mediator thus stands in the midst of the causal chain from *a* to *b* to *c*, as its Latin origin *medius*—the middle—indicates.[1]

### Example of mediation

A consulting project I worked on a few years ago offers what I think is a splendid example of mediation. The context was an experiment that examined the effect of diet on coat colour in a special breed of mice.[5] This was highly sophisticated work involving genotyping and a number of other technical details of which I have only a faint grasp, but the gist of the experiment was this. Female mice of a very special genotype were randomly assigned to two dietary conditions, and were mated with males while on the diet. One diet was presumed to alter the expression of a gene that controlled coat colour in offspring; the control diet was believed to have no influence on the expression of that gene. Thus the offspring of the mice in the special diet condition were expected to have a different coat colour than those in the control condition. The researchers also measured the presumed mechanism—a process called methylation—that was thought to be the causal link between the diet and coat colour. In our mediation framework, the dietary condition is the predictor of interest (*a*); methylation is the potential mediator (*b*) and coat colour, the outcome (*c*). Following the mediation criteria, the researchers first showed that indeed, the offspring of mice in the experimental diet condition did have different coat colour compared with the offspring in the control condition (see upper portion of fig 1). So, *a* was related to *c*. They also showed that the degree of methylation was related to both coat colour—that is, *b* was related to *c*; and they showed that the dietary condition was associated with the degree of methylation—that is, *a* was related to *b*. Finally, they showed (see lower part of fig 1) that when methylation was entered into a regression model in which diet had predicted coat colour, the relation between diet and coat colour almost entirely vanished. In other words, methylation almost entirely explained the relation between diet and coat colour. Because their earlier work had shown that methylation was in the causal chain between diet and coat colour, they

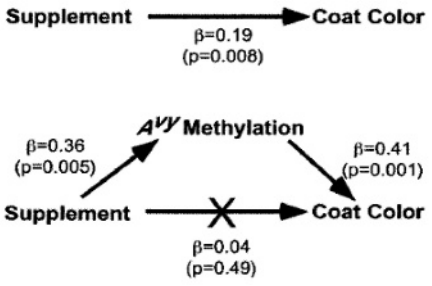


**Figure 1** Methylation as a mediator between difference in diet and coat colour of offspring. From Waterland and colleagues.[5]

concluded that the data were consistent with the belief that methylation accounted for, or mediated, the relation between diet and coat colour. In the parlance, diet had an *indirect effect* on coat colour by way of methylation.

## NOT SO FAST

Although all of the above may seem relatively straightforward, not surprisingly, there are potential complications. In the above mouse example, although the researchers were very confident that methylation was the active mechanism and had much biological evidence to support this view, there is always the possibility that methylation was not the sole or even the real causative mechanism. There still may have been some unmeasured mechanism, closely associated with methylation that changed the coat colour. Although unlikely, the possibility exists that methylation may have been a byproduct of or marker for some other unmeasured or as yet unknown mechanism. If this were the case, to the extent that methylation was strongly associated with the *real* (but unidentified) mechanism, it would behave statistically just like the real mediator. Such variables are always lurking outside our statistical models, more dangerously so in purely observational studies, but randomised experiments are not entirely immune to this problem.

A further complication is that it is not always clear whether a variable is a confounder or a mediator. In our examples above, age clearly could not have been caused by the type of tobacco used, and the mouse researchers were confident that methylation was caused by the diet, and in turn caused coat colour changes. However, in at least some instances there may be doubt or reasonable debate as to whether a variable is in the causal path or not. For example, shortly after completing graduate school, I was a co-author on a publication that examined the relation between a set of psychosocial characteristics and survival in men.[6] Briefly, 750 middle-aged men were interviewed in a standardised way, recorded on audio tape. The recordings were later scored along a number of dimensions of vocal style and verbal content, which were categorised as hostile, socially dominant, withdrawn and several more. The men were then followed for 22 years for survival status. Among the findings was that men who were classified as hostile tended to die over the follow-up period at a higher rate than men not so classified. Almost as a matter of reflex, to demonstrate that hostility was an "independent risk factor", we adjusted the survival model for a number of background variables, including age and traditional risk factors for the leading cause of death, heart disease. The traditional risk factors included serum cholesterol, cigarette smoking, diastolic blood pressure and age. The intent of the adjustment was to ensure that the observed association between the psychosocial variables and survival was not confounded with these additional background variables. With hindsight, let us consider each of the adjustment variables in the context of our confounding criteria. If age were related to hostility, and it often is, at least in other studies,[7] it would be a potential confounder and we would want to adjust for it. If the relation between hostility and survival were substantially diminished after adjusting for age, just as in our pipe, cigar and cigarette use example above, we would conclude that the original, unadjusted association between hostility and survival was an artefact produced by the non-causal association between hostility and age.

As for the traditional risk factors, cholesterol, blood pressure and cigarette use, however, the picture may not be quite as straightforward. Hostility has been shown to be related to all three of these risk factors, and each also have been shown to be related to survival.[8] However, it is possible that cholesterol, blood pressure and cigarette use *are causal consequences* of the personality trait of hostility. Compared with their more affable counterparts, hostile men may be more prone to ignoring or opposing medical advice about health habits and thus have worse risk profiles—higher cholesterol, higher blood pressure and heavier cigarette smoking. In this case, the three risk variables are part of the causal effect of hostility on survival. If the association between hostility and survival is diminished after adjusting for these three factors, we could not necessarily conclude that the unadjusted association between hostility and survival was an artefact. Instead, we might conclude that these risk factors were the mechanisms—the mediators—by which hostility was related to survival. Partialling out the contribution of the risk factors would actually underestimate the total effect of hostility on survival. In the mouse and coat colour experiment, this would be analogous to concluding that the diet manipulation was not related to coat colour because the association disappeared after adjustment for methylation. Methylation was not an artefact; it was *why* the diet was related to coat colour.

Thus the critical considerations in determining the presence of confounding or mediation, or distinguishing between the two, are the extra-statistical arguments regarding cause that are brought to bear on the question at hand. Without such knowledge, it is impossible to distinguish between mediation and confounding. Even when extra-statistical arguments can be made, however, we are not always guaranteed to be correct about our conclusions. With respect to confounding, unmeasured or uncontrolled variables are always a threat. In a randomised experiment, the threat is relatively low, but it is not impossible for some hidden but important imbalance between treatment conditions to explain the apparent success of a treatment—in other words, the treatment condition may be confounded with the variables that are not reasonably balanced across treatment conditions. The threat is much more significant with observational data studies, and volumes have been devoted to the best approaches to identifying confounding (see, for example, Harrell[9] and Steyerberg[10]). With respect to mediation, caution also must be duly exercised in drawing firm conclusions. Even in a well controlled experiment, it is possible that the putative mediator is simply a proxy for some other mechanism. In observational studies, particularly those which attempt to draw inferences about mediation with cross sectional data, this threat is quite severe. Even when the criteria for mediation are met and the regression model produces the desired pattern of results, the most we can conclude is that the data are consistent with—or perhaps stated more precisely, the data are not inconsistent with—mediation. Our putative mediator may well be confounded with some unmeasured variable or variables. And of course, we may simply have the causal model all wrong. Again, in cross sectional observational designs in particular, we often cannot know which direction cause is occurring, or whether there is an unmeasured variable out there that is a cause of one or even all of the measured variables.

As a final note, the enterprise of understanding causes among variables measured in a variety of experimental and non-experimental designs has become extremely sophisticated in recent years but these advanced ideas are far beyond the scope of this short paper. For those interested in digging deeper, Pearl's work on graph theory, although at times technically daunting, has been especially important in this area.[11] For mediational

analysis, although still by far the most frequently used approach, the original method recommended by Baron and Kenny[12] of simply adjusting for the mediator and watching the change in regression estimates also has evolved considerably. As a starting point, the interested reader might examine the work of MacKinnon and colleagues[13 14] (also see http://www.public.asu.edu/~davidpm/ripl/mediate.htm) and also Kraemer and colleagues[15] to learn more.

**Competing interests:** None.

## REFERENCES


1. **Glare PGW.** *Oxford Latin dictionary*. Oxford: Oxford University Press, 1982.
2. **Greenland S.** Confounding. In: Boslaugh S, ed. *Encyclopedia of epidemiology*. Thousand Oaks, CA: Sage, 2007:227–32.
3. **Rubin DB.** Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 1997;**127**:757–63.
4. **Cochran WG.** The effectiveness of subclassification in removing bias in observational studies. *Biometrics* 1968;**24**:295–313.
5. **Waterland RA,** Jirtle RL, Waterland RA, *et al*. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 2003;**23**:5293–300.
6. **Houston BK,** Babyak MA, Chesney MA, *et al*. Social dominance and 22-year all-cause mortality in men. *Psychosom Med* 1997;**59**:5–12.
7. **Barefoot JC,** Peterson BL, Dahlstrom WG, *et al*. Hostility patterns and health implications: Correlates of Cook–Medley Hostility Scale scores in a national survey. *Health Psychol* 1991;**10**:18–24.
8. **Siegler IC,** Peterson BL, Barefoot JC, *et al*. Hostility during late adolescence predicts coronary risk factors at mid-life. *Am J Epidemiol* 1992;**136**:146–54.
9. **Harrell FE.** *Regression modeling strategies: With applications to linear modeling, logistic regression, and survival analysis*. New York: Springer, 2001.
10. **Steyerberg EW.** *Clinical prediction models*. New York: Springer, 2009.
11. **Pearl J.** *Causality*. Cambridge: Cambridge University Press, 2000.
12. **Baron RM,** Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**:1173–82.
13. **MacKinnon DP,** Dwyer JH. Estimating mediated effects in prevention studies. *Eval Rev* 1993;**17**:158.
14. **MacKinnon DP,** In AC, Beatty LA. *Analysis of mediating variables in prevention intervention studies. Scientific methods for prevention intervention research*. Washington, DC: DHHS Pub, 1994:127–53.
15. **Kraemer HC,** Wilson GT, Fairburn CG, *et al*. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry* 2002;**59**:877–83.


# *Et al*

The articles we select for *Evidence-Based Mental Health* must pass two stages: first they must pass our basic validity criteria and then the editors assess each article for clinical relevance. A number of articles meet the inclusion criteria but are not abstracted due to lack of space. We will highlight the most interesting of these here and list the rest.

"A gigantic asylum is a gigantic evil, and figuratively speaking a manufactory of chronic insanity." An observation by John Arlidge on the large psychiatric hospitals being built in Victorian England but the sentiment could be equally applied to the overcrowded prisons of the 21st century. Approximately 25% of sentenced men in prison in the UK, and over 40% of sentenced women, have been estimated, using clinical criteria, to require psychiatric treatment or further assessment. The closure of inpatient beds, failure of community services and in the US mandatory sentencing for drug offences have all been implicated in the transformation of prisons to the new asylums. It is also an international issue (see HeadtoHead (http://blogs.bmj.com/ebmh-talk): Welcome to the Asylum! http://blogs.bmj.com/ebmh-talk/2009/04/30/welcome-to-the-asylum). Research on prison mental health struggles to get into the pre-eminent journals—try browsing Medline—but these three papers have managed to spark some editorial interest. For those who still need telling that diversion schemes are a good thing, this cohort study from Texas (*Am J Psychiatry* 2009;**166**:103–9), where they have over 100 "correctional facilities", finds that inmates with schizophrenia, bipolar disorder and major depression had as much as a threefold increase in risk of at least four incarcerations over the 6 year study period. We all know that suicide happens commonly in prison, with rates in the UK at least five times that of the general population. Paradoxically, this review of risk factors (*J Clin Psychiatry* 2008;**69**:1721–31) finds that being married and being employed are associated with suicide, possibly related to having more to lose. Untreated mental illness and alcohol problems, seemingly obvious but perhaps not so in the prison setting, are modifiable risk factors. So is being placed in a single cell, although given that most prisons are bursting at the seams I thought those had gone, along with "doing porridge". The increased risk of suicide is even greater in imprisoned adolescents, of which there are about 100 000 in the USA. This systematic review (*J Am Acad Child Adolesc Psychiatry* 2008;**47**:1010–19) comprising mainly US studies looked at the prevalence of mental disorder in juveniles in prison and found that 3% had a psychotic illness with approximately 11% of the boys and 29% of the girls having a major depressive disorder.

Everybody knows that if an antidepressant has not worked after a few weeks you should up the dose. Surely that is an evidence-based decision? Actually the evidence is pretty weak and this randomised trial (*J Clin Psychiatry* 2008;**69**:1383–92) comparing an increased or maintained dose of duloxetine in patients who have not improved after 6 weeks seems to confirm that it makes little difference what you do. Having said that, is anyone still prescribing duloxetine? After receiving unflattering reviews in both the *Drug and Therapeutics Bulletin* and *Prescrire International* it appears that we have an Anglo-French Accord stating that as an antidepressant duloxetine is singularly unimpressive.

A pharmacology question: which class of drug developed by Smith and Kline in the 19th century has proven popular with politicians including JFK and Adolf Hitler? Was commonly prescribed for asthma and hay fever? Gee whizz, if you haven't got it by now here is the giveaway: these days they are more likely to be prescribed for narcolepsy and attention deficit hyperactivity disorder (ADHD). Amphetamines of course, and despite them being around for over 100 years, a Cochrane review could not find any evidence for an effective drug or psychosocial treatment for dependence.[1] The authors of this Swedish study (*Am J Psychiatry* 2009;**166**:103–9) think they have the answer. Once you navigate your way past the rat brains, cocaine and positron emission tomography scanning in the introduction, you will find a small randomised controlled trial comparing the opioid receptor antagonist naltrexone to placebo with positive results. Naltrexone is already being used for heroin and alcohol dependence but some more trials with larger numbers are needed to demonstrate that it really does work for amphetamines. Ecstasy (MDMA) is perceived as a modern day derivative of amphetamines but has been around for almost as long—as anyone one who has read PG Wodehouse's short story *Bertie Wooster finds a Rave* can attest. The third most commonly used illegal drug in the UK, after cannabis and cocaine, there

EBMH

# Understanding confounding and mediation

Michael A Babyak

*Evid Based Mental Health* 2009 12: 68-71
doi: 10.1136/ebmh.12.3.68



## Notes