The failure to know what isn't known: negative publication bias with lamotrigine and a glimpse inside peer review

S N Ghaemi

"There is a type of interaction between human beings which proceeds not from knowledge, or even lack of knowledge, but from failure to know what isn't known." John Kenneth Galbraith¹

The medical literature meets Galbraith's description. Some things we know, and know that we know. Other things we do not know, and know that we do not know. But perhaps the largest class involves those things we do not know, and do not realise that we do not know.

This latter state of affairs is exemplified by the problem of negative studies. It has become increasingly clear that the medical literature is biased toward positive studies; negative studies are less frequently published.² Sometimes this may reflect loss of passion, as disappointed researchers file away their negative results. Sometimes it may be systematic, as pharmaceutical sponsors may actively suppress negative data which would adversely impact their marketplace sales. And journals may also systematically reject negative studies—which will generate fewer readers, fewer citations and lower impact factors for the journal—more frequently than positive ones.

CLINICAL RELEVANCE OF NEGATIVE STUDIES

Negative studies may provide important information. "Negative" does not mean unimportant, and proving something ineffective is as important, perhaps more so, than proving it effective. To be a good physician, knowing when *not* to use which medications, as Pinel famously said, is even more important than knowing when and how to use them.³ Drugs that are ineffective might be somewhat harmful; almost all drugs are harmful in some way; thus the risk-benefit calculation

is always hurtful for an ineffective drug. Furthermore, drugs which are effective are often not used, since the ineffective drugs are seen as effective, and thus the truly effective drugs, which sometimes have more side effects, are pushed aside. It is as if the princess at a ball, hidden in servant clothes, is ignored in favour of a servant girl, dressed up like a princess.

In effect, unpublished negative studies are needed to put positive studies in perspective. The clearest example in psychiatry is with antidepressants. The published literature is almost entirely positive, but when unpublished negative studies at the Food and Drug Administration (FDA) are added, 51% of the actual studies are positive and 49% negative.⁴ Had clinicians known for the past two decades that the actual research evidence base has a score of 51 vs 49, not 90 vs 10, perhaps they would have been more judicious in their use of antidepressants.

Another major example is the use of gabapentin for bipolar disorder. (One might call it "gabapentin syndrome," this failure to recognise the hidden princess, like lithium, in favour of an attractive impostor.) Negative studies of gabapentin in mania, some of which were eventually published,⁵ revealed the impostor for who she was. And gabapentin use has declined but not until after 5 years or more of extensive use and happy profits. Now generic, and with no further defenders, the truth about the drug is easy to state. But the truth should be the same whether a drug is under patent or generic.

THE CASE OF LAMOTRIGINE

The issue of lamotrigine and its negative studies is more complex because lamotrigine has been proven effective (unlike gabapentin) for bipolar disorder in some circumstances, namely maintenance treatment. It was also proven ineffective in other situations (everything else: see below). I was a principal investigator in the maintenance studies with lamotrigine

in the mid 1990s, and on the Lamictal national advisory board for its manufacturer, GlaxoSmithKline (GSK), as well as a recipient of honoraria from GSK for lectures until early 2008. Thus I was surprised to discover the existence of several negative lamotrigine studies, available on the GSK website to the author in 2006. In 2004, a paroxetine related settlement with New York State Attorney General's office forced GSK to post a registry which would include much more information about pre-trial and clinical drug study results. Of the nine lamotrigine related bipolar disorder studies posted on the website (see table 1), two were positive and published supporting the FDA approved indication for delay of relapse in the long term treatment for bipolar disorder patients.^{6,7} A negative study in rapid cycling bipolar disorder and another in acute bipolar depression were published but both emphasised positive secondary outcomes as opposed to the negative primary outcomes. Five other negative studies involving rapid cycling bipolar disorder, acute bipolar depression and acute mania have not been published and are only available on the GSK website.

Failure to adequately publish these negative studies led to the creation of a clinical impression that lamotrigine is an "antidepressant," a view innocently expressed to me as recently as last week by an academic colleague. This mistaken impression occurred partly because the *prophylactic* benefits of lamotrigine for depressive episodes were confused with a presumed *acute* benefit. Partly it was due to the publication of one apparently positive study, and the non-publication of several negative studies.

The clinical relevance of the lamotrigine studies is notable: taking the negative outcomes into account, as of now, one might say that this agent is reasonably effective in maintenance treatment of bipolar disorder, particularly in prevention of depression. It is proven ineffective in acute mania, rapid cycling disorder and acute bipolar depression.

THE ROLE OF THE FDA

It is worth mentioning that the FDA has encouraged this state of affairs, by viewing negative studies as uninformative, due to the possibility of being "failed" rather than truly negative (ie, the sample may have simply been unresponsive, or dosing might have been too low and so on). Thus drugs could have two positive studies, and 10 or so negative ones (as did a number of

Correspondence to: Dr S N Ghaemi, Tufts Medical Center, Department of Psychiatry, 800 Washington St, Box 1007, Boston, MA 02111, USA; nghaemi@tuftsmedicalcenter.org

Table 1 Published and unpublished randomised clinical trials of lamotrigine in bipolar disorder

Study	Diagnosis	Design	No	Duration (weeks)	Scale	Outcome	Result	Published
SCAA2008	Mania	LTG vs Li vs Pla	216	3	MRS	LTG -9.3 ± 10.9 Li -10.7 ± 11.6 Pla -9.5 ± 10.5	Failed*	No
SCAA2009	Mania	LTG vs Li vs Pla	229	6	MRS	LTG -11.6 ± 14.0 Li -15.6 ± 13.0 Pla -11.4 ± 12.3	Negative†	No
SCAA2010	Bipolar depression	LTG vs Pla	206	10	HDRS	LTG -10.6 ± 8.3 Pla -10.2 ± 8.1	Negative‡	No
SCA40910	Bipolar depression	LTG vs Pla	257	8	MADRS	LTG -12.2 Pla -11.2	Negative‡	No
SCAB2001	Bipolar depression	LTG vs Pla	195	7	HDRS	SDs not provided LTG (200 mg/day) -10.5 ± 8.1 ($p = 0.08$) LTG (50 mg/day) -9.3 ± 8.9 ($p = 0.24$) Pla -7.8 ± 7.9	Negative‡	Yes
SCAB2005	Rapid cycling	LTG vs Pla	137	32	TIME	Median survival time LTG 142 days Pla 133 days $p = 0.73$	Negative‡	No
SCAA2012	Rapid cycling	LTG vs Pla	182	34	TIME	Median survival time LTG 126 days Pla 79 days $p = 0.18$	Negative‡	Yes
SCAB2003	Prophylaxis	LTG vs Li vs Pla	463	72	TIME	Median survival time LTG 200 days Li 170 days Pla 93 days $p = 0.029$ for both Li and LTG	Positive	Yes
SCAB2006	Prophylaxis	LTG vs Li vs Pla	175	72	TIME	Median survival time LTG 141 days Li 292 days Pla 85 days $p = 0.02$ for LTG $p = 0.003$ for Li	Positive	Yes

Source for unpublished data: www.gsk.com (accessed 11 June 2009).

*Lamotrigine and lithium (active control) were equivalent to placebo, thus no information for or against lamotrigine efficacy can be concluded.

†Lamotrigine was equivalent to placebo but lithium was more effective than placebo.

‡Lamotrigine was equivalent to placebo in an adequately powered study to demonstrate modest effect sizes.

HDRS, Hamilton Depression Rating Scale; Li, lithium; LTG, lamotrigine; MADRS, Montgomery Asberg Depression Rating Scale; MRS, Mania Rating Scale from SADS-C; Pla, placebo; TIME, time to intervention with psychotropic medications due to full relapse or initiation of relapse into mood episodes.

Reprinted with permission from the *Medscape Journal of Medicine* 2008;10:211. Available at <http://www.medscape.com/viewarticle/579046> © 2008 Medscape.

selective serotonin reuptake inhibitors),⁸ and the FDA not only allowed approval but it did not require that the pharmaceutical industry publish its negative results. The pharmaceutical industry did the minimum necessary; the FDA set the minimum far below what should have been the acceptable scientific standard.

Furthermore, even though the FDA is supposed to allow access to information about all studies for drugs for which it provides an indication, in fact it still does not allow access to raw data from industry sponsored studies, which are viewed as confidential and proprietary. Rather, in my experience and those of

colleagues who have attempted to access such studies, Freedom of Information Act requests are met with abstracted summary results. While summary results are better than no results, full access to scientific data should be the standard at the FDA.

PEER REVIEW EXPERIENCE

It is rare for an author to have the opportunity to let readers see the winding course from writing to peer review to publication of a paper. Some papers are rejected repeatedly, as was ours, before publication, and thus the course can be rather serpentine. The chance to present

the process as it happened may be especially relevant in this case as it may highlight another obstacle in the way of adequate science: the systematic resistance of the peer review system to such critiques.

In the course of trying to investigate and publicise these negative findings, we ran into numerous roadblocks, as I discussed in an interview for the Carlat Report last year (<http://www.thecarlatreport.com/index.asp?menu=wp314200795443&page=wp3152007102735>). At first, I supplemented my analysis of the GSK website by going to other pharmaceutical company websites to look for data with drugs known

to be studied in bipolar disorder, my field of expertise. I found that only GSK had extensively reported negative data; in fact, the only other company with more than a few studies on its website for bipolar disorder was Eli Lilly, and all its studies were positive. I knew that negative studies had been conducted with olanzapine in bipolar disorder, studies which I had seen in poster format at conferences. I emailed Lilly employees to clarify where those studies were, and the response was that such data were in preparation for publication, and study results would not be posted on the web if they were being prepared for publication. To date, I have not seen negative studies published with olanzapine in bipolar disorder. Other companies are also notable for the absence of any negative studies in bipolar disorder on their websites. The FDA site (www.clinicaltrials.gov) also had no results for any of these agents.

Initially, my colleagues and I tried to write up a review of all the companies, not singling out only GSK, although we put more emphasis on the GSK and Lilly studies. In initial informal peer review with a senior colleague, I found a great deal of scepticism about how the paper would be received and, although suggested revisions were given, my colleague demurred from co-authoring the paper. I then submitted the paper to the *Journal of the American Medical Association (JAMA)*, which had published numerous similar studies in medicine, although not psychiatry, and whose editors I knew were interested in this topic. The paper was immediately rejected by one of the editors in July 2006, without comment. I asked for specific feedback, and received a letter with numerous complaints, such as what follows: "There is a considerable literature on this specific topic, almost all of which you failed to cite". The editor goes on to note that some of these papers were co-written by the editor, which had "profound effects". He continued: "This failure on your part indicates a naivety [sic] or ignorance of the broader picture....You thus fail entirely to give the paper context". Much of the context had been taken out by the informal peer review of my senior colleague, who thought it was too politically sensitive. The editor continued, noting an editing mistake I made: "As a psychiatrist, you will appreciate the annoyance any *JAMA* editor might naturally feel when the manuscript he reads has an abstract written in the *New England Journal of Medicine* style. Unimportant, perhaps, but indicative of the fact that you cannot read *JAMA*, and that you did not bother to

read our instructions for authors". In fact, I read *JAMA* weekly but left it to my research assistants to check the instructions. He dismissed the abstract as "classic...pretty much useless," the methods as "wandering and discursive," the results as "incomplete, and so all you really had were the GSK trials. That's your paper. And the result is trivial, first, because the settlement has made GSK a special case, and secondly, because we already know very well that such trials are often, perhaps usually, unpublished, and the debate has long ago moved on to how to force compliance with registration." He concluded: "If you had sent us a crisp paper that clearly stated a hypothesis, and a credible way of investigating it; if you'd given us the context, clear methods and adequate statistical analysis; if you had provided the relevant citations, and if your hypothesis and investigation had been on something that hadn't been already documented by others: then we might have been interested. Unhappily, you did not. I hope this will help you in the future. Best wishes".

I was not entirely certain of the sincerity of those wishes but I did learn something from the *JAMA* editor about this paper: my initial goal was too broad; I could not address the whole pharmaceutical industry, even to make the simple point that almost all companies failed to disclose any negative data on their websites. I decided to emphasise the GSK database and the lamotrigine studies, and limit the discussion of the other companies and drugs. I then sent this shortened, focused paper—trying to apply all of the *JAMA* critiques—to the *American Journal of Psychiatry*. It was rejected without review. I then sent it to the *British Journal of Psychiatry* as an editorial; it was rejected without review. I finally decided to send it to the most cited journal that specialises in bipolar disorder, the journal *Bipolar Disorders*, where I also serve on the editorial board. It was rejected after two peer reviewers analysed it as follows: the main critique of the first peer reviewer was about why only websites were chosen for review; this reviewer wanted me to assess all scientific meetings to see what presentations had been made as posters, as well as review papers, chapters etc; this larger task was obviously not feasible. The reviewer thought the "conclusions are unwarranted" because these additional sources were not included. My point is that I was testing the companies' claim that they would provide their data on websites, which they did not. The second peer

reviewer, among other questions, was not convinced that we had done what we said: "The authors do not describe how they systematically compared data from different sources. How were they able to determine that negative data exist but that a specific company failed to make public? The authors need to provide specific examples where a negative study is reported in a specific public website but fails to be identified on the company's website".

In sum, our main critique from *JAMA* was that the problem was already well known, and our paper prepared shoddily, with too little information except what had already been known with GSK. The main critique from the world of bipolar research was the opposite: our claim was not to be believed unless we made extensive efforts to seek out other places where negative data might have been presented.

I have come to the conclusion that when a paper is rejected multiple times in peer review, for opposing reasons, then it is a good paper.⁹

Having run out of options, after 2 years of attempting publication, I had resigned myself to not being able to publish my paper on unpublished studies, until, after an interview for the *Carlat Report*, where I described this odyssey, the editors of the online journal *Medscape Journal of Medicine* invited me to send them my paper. Three peer reviewers approved it after some revision, and it was eventually published (http://www.medscape.com/viewarticle/579046_1).

RESPONSES TO THIS CRITIQUE

In response to this critique, some have noted that GSK published some of the data in combined form in one paper a few years ago,¹⁰ and more recently published a meta-analysis of the combined acute bipolar depression studies.¹¹ However, the initial partial publication of some of the negative studies included little data on the lack of affect. Furthermore, combining multiple papers in one location, and doing so only once, has much less influence than publishing each paper separately when the results are positive, or publishing several papers from the same single positive study, as is common practice. Needless to say, the positive papers also get published in journals with high impact factors and are commonly associated with press releases and related media coverage.

As for the meta-analysis, a kind of "statistical alchemy"¹² is performed whereby the overall negative results, although acknowledged, are leavened by the inclusion of a positive subgroup

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 paper¹¹ 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.

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 one².) 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

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