

- [ScienceWatch Home](#)
- [Inside This Month...](#)
- [Interviews](#)

- Featured Interviews
- Author Commentaries
- Institutional Interviews
- Journal Interviews
- Podcasts

Analyses

- Featured Analyses
- What's Hot In...
- Special Topics

Data & Rankings

- Sci-Bytes
- Fast Breaking Papers
- New Hot Papers
- Emerging Research Fronts
- Fast Moving Fronts
- Corporate Research Fronts
- Research Front Maps
- Current Classics
- Top Topics
- Rising Stars
- New Entrants
- Country Profiles

About Science Watch

- Methodology
- Archives
- Contact Us
- RSS Feeds

scienceWATCH[®].com

TRACKING TRENDS & PERFORMANCE IN BASIC RESEARCH

[Interviews](#)
[Analyses](#)
[Data & Rankings](#)

2009 : May 2009 - Author Commentaries : Erick Turner

AUTHOR COMMENTARIES - 2009

May 2009



[+enlarge](#)

Erick Turner

Featured Scientist from *Essential Science Indicators*SM

According to *Essential Science Indicators* from *Thomson Reuters*, the paper "Selective publication of antidepressant trials and its influence on apparent efficacy," (Turner EH, et al., *N. Engl. J. Med.* 358[3]: 252-60, 17 January 2008) has been cited 79 times from the time it was published to December 31, 2008. It has been named as both a *Highly Cited Paper* and a *New Hot Paper* in the field of Clinical Medicine. As of the April 18, 2009 update of the *Web of Science*[®], this paper shows 114 total citations.

Lead author Dr. Erick Turner is Assistant Professor in the Department of Psychiatry and the Department of Pharmacology at Oregon Health and Science University (OHSU) in Portland. He is also Senior Scholar with OHSU's Center for Ethics in Health Care and Staff Psychiatrist with the Portland Veterans Affairs Medical Center. He has previously held positions at the US Food & Drug Administration and the National Institute of Mental Health.

In this interview, ScienceWatch.com correspondent Gary Taubes talks with Dr. Turner about this paper and the buzz it has created in the research and pharmaceutical industries.

SW: How did you get interested in the issue of selective publication of clinical trial results?

Before coming here to Oregon Health and Science University and the Portland VA, I was a medical officer at the US Food & Drug Administration (FDA), doing clinical reviews of new drug applications. Prior to that, I had been at the National Institute of Mental Health (NIMH), doing a fellowship. I thought that NIMH was a mecca and that I had access to the very best information. But when I left NIMH and went to the FDA, I realized I had been in the dark, basically clueless.

At the FDA, once I started reviewing new drug applications, I became aware of a large number of negative trials. I had never seen a negative trial before in a peer-reviewed journal. My impression had always been, and the teaching I'd always received, was that the peer-reviewed literature was the Holy Grail, the most authoritative source of medical information there was. As I came to see it, this just wasn't the case.

Seeing these negative trials at the FDA, I felt that I had this for-your-eyes-only kind of access to secret information. If I hadn't worked at the FDA, I wouldn't have known that these negative trials existed. Later, when I came to academia, I saw a naiveté among my colleagues and medical professionals, this belief that if something is published in a journal, it must be so. I felt it was wrong that medical professionals, on the front lines treating patients, should be getting a sanitized view of drug efficacy—they need to know

the whole story.

SW: Did this actually affect research or just prescribing practices?

Both. For research, it was an uphill battle getting approval to do clinical trials involving placebo. Our institutional review board (IRB) took the position that placebos were unnecessary and that all you had to do to prove a drug worked was to compare it with a so-called known effective drug.

The reasoning is this: Let's say that drug A is your known effective drug, and you "know" it's effective because it's been approved by the FDA. Now say that a new drug, call it drug B, performs as well as drug A in a clinical trial. Then, voilà, by the law of transitivity, you have demonstrated that the new drug is effective, too. The flaw in this logic is that it rests on the premise that drug A is superior to placebo in every clinical trial. I knew otherwise from my time at the FDA, but the IRB wasn't going to just take my word for it because they had seen only positive trials in the literature.

SW: So what you're saying is that a drug can be effective sometimes and not others, and so you never know if one of the other times is the case in your particular clinical trial?

Yes. People often think of efficacy as an all-or-nothing phenomenon. Once a drug gets the stamp of efficacy from the FDA, it's as if the light switch of drug efficacy has been flipped into the "on" position. In reality, it works more like a dimmer switch, along a continuum. That's where the concept of effect size comes in, which we can talk more about later.

SW: Did you work on this before the work for the 2008 NEJM article?

Well, in 2005 I took part in a published debate and teamed with Martin Tramèr, a researcher in Switzerland. The opposing team took the position that placebos should almost never be used. We made the argument that you could make a serious public health gaffe by approving a drug that might have no advantage over placebo because you've tested it against a drug that beats placebo only some of the time. If you include a placebo arm in the trial, you can see whether the study drug and the active comparator separate from the placebo in that particular trial. Without the placebo arm, you're really just guessing.

In 2004 I wrote an essay in *PLoS Medicine* ("A taxpayer-funded clinical trials registry and results database," *PLoS Medicine* 1[3]: 180-2, December 2004). That was when there was a lot of talk about registries in the wake of the Vioxx scandal and the State of New York's lawsuit against Glaxo regarding Paxil for pediatric depression. My point was, let's not reinvent the wheel by creating a new system to combat publication bias. We already have a registry and results database—it's called the FDA. It's a gold mine of clinical trial data that goes back several decades, and we're ignoring it.

In that essay, I used a couple of anti-anxiety drugs as examples and contrasted what the journals said about them with what the FDA reviews said. After publishing this essay, I thought it would be interesting to expand upon that approach and look at the entire class of antidepressants.

SW: When did you start work on that?

I believe it was late 2005.

SW: Did you still have access to the FDA reviews to do the project?

Yes, and to some extent, you do, too. FDA reviews are posted online for drugs that have been approved since 1997. Unfortunately, very few people know about this resource, even though it's right there in the public domain for anyone to access.

Getting the FDA reviews on the newer drugs was relatively easy. But the challenge was for the older drugs, like Prozac, Zoloft, and Wellbutrin. All these had been approved prior to 1997, so those reviews are not posted on the FDA website. How did I get those reviews? I filed a Freedom of Information Act

Figure 1:

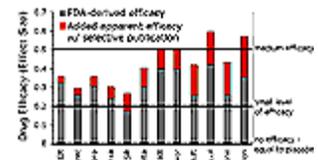
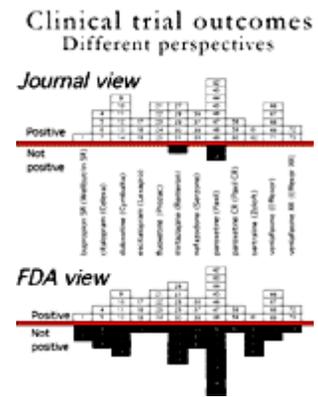


Figure 2:



[View/download figures & descriptions.](#)

(FOIA) request in 2005, but my request fell into a bureaucratic black hole. The FDA isn't known for its efficiency in handling FOIA requests, even when compared to other government agencies.

SW: So what did you do?

Well, I knew that a colleague of mine, Arifula Kahn, had written some papers based upon reviews he had gotten from the FDA. He runs a clinical trials center in the Seattle area, which is close to Portland. I went up to see him, and he graciously let me take his stuff to Kinko's—armloads of these notebooks containing FDA reviews—and I spent hours there copying. That was one source. Then another colleague, David Antonuccio at University of Nevada at Reno, kindly sent me some more FDA reviews.

SW: How did the reviewers at the *New England Journal* respond when you submitted the paper?

I think the typical number of reviewers for a manuscript is three. Our manuscript got eight. One of the reviewers later told me he had never heard of a paper being assigned so many reviewers.

The overall reception was positive, but with that many reviewers, some with a few comments and some with lots of comments, we had our work cut out for us. Later, in the second review cycle, one of the reviewers picked up on an error I made in labeling one of the figures—there was a "figure 2" should have been labeled "figure 3." He said, "You know I hate to put you through the wringer again, but it makes me wonder if you could have made other mistakes. I'm going to insist that you do double data extraction and entry."

This added months to the project, because in a sense we had to do the study all over again. I imposed on some people to repeat the extraction and entry of the data, to make sure we didn't make any mistakes and to guard against bias. There were a few minor changes, but it didn't change the overall result—in fact, some of the findings got a bit stronger.

SW: What did the study ultimately conclude? What was the message, in effect?

Let's back up a little bit. To do this study, first we identified all the FDA-registered trials on 12 antidepressants, and then we tracked those trials into the published literature to answer two questions. Number one was whether the trial was published, and number two, if it was published, was it done so in a way that agreed with the FDA, or was it spun?

So if you go by the journal articles, as doctors are taught to do, then 94% of the trials were positive, meaning that the drug was statistically superior to placebo with a P value of less than .05. On the other hand, if you go by the FDA reviews, you see that the true proportion of positive trials is 51%. That's quite different. Nearly 100% dropped to essentially 50-50. That was one message.

SW: 50-50? What does that mean? That it's a toss-up whether the drug works?

As we were talking about earlier, drug efficacy is not like a light switch, so efficacy isn't completely present half the time and completely absent the other half. It's not the case at the level of the clinical trial, and it's not the case at the level of the individual patient, either. What this 50-50 means is that the average response to the study drug was statistically better than the average response to placebo in about half of the trials. In the other half, the drug was usually still numerically better than placebo, but not enough to be statistically better. Having said that, we did find a few trials where the drug was actually numerically worse than placebo. Needless to say, those results weren't published.

SW: If the drug is only superior to the placebo in only half of clinical trials, does the drug actually work?

"We, as American taxpayers, have this gold mine of clinical trials data at the FDA that could be leveraged and made a

That's where meta-analysis comes in. Meta-analysis lets you combine trials that show a statistically significant difference between drug and placebo with trials that don't. Then you can see whether, for all trials combined, the drug-placebo difference is statistically significant. For each of the 12 drugs we looked at, it was.

Getting back to the point that drug efficacy is not an all-or-none phenomenon, meta-analysis also tells you how big that difference is. In other words, it allows you to put its efficacy somewhere along a continuum. The effect sizes we calculated for each drug based on the FDA data were substantially less than the effect sizes based on the published literature. Whether those effect sizes are clinically significant, in addition to statistically significant, is a matter of debate.

SW: Were you surprised at the response to the article?

I thought it would get some press response but it far exceeded my expectations. I wonder if it would have gotten so much attention if it hadn't been published in the *New*

lot more
accessible
to us."

England Journal of Medicine. If it had been published in a low-profile journal, it might have been ignored. But because it got so much press, the pharmaceutical companies apparently decided they couldn't ignore it and should respond to it.

SW: What was the industry response?

It reminded me of the movie title *The Empire Strikes Back*. One of the press articles about our article was in the *New York Times*. Ben Carey led off his article with a sentence containing the phrase "antidepressants like Prozac and Paxil." Lilly struck back with a press release the very next day, taking issue with the *New York Times* article and the implication that they weren't transparent, pointing out that our *NEJM* article showed that all the Prozac trials had been actually published.

Then they complained about our *NEJM* article. We had described a couple of Cymbalta trials as being unpublished. They said that was wrong, that those two trials had been published and that they had been presented in various meetings. In addition to their saying this in the press release, they repeated it in a "Dear healthcare provider" letter that probably went out to thousands of prescribers around the country.

What they didn't mention was that these two articles where the trials were "published" had taken these two negative trials, bundled them with four positive trials, and reported them in two review articles. Each review article covered the same six trials and concluded that their drug had demonstrated efficacy in a majority of those trials.

SW: Doesn't that speak to the question of how you define "published?"

Yes. We defined a trial as "published" if it was published in its own stand-alone journal article. We explicitly excluded review articles, which meant we didn't count those two review articles reporting that duloxetine (Cymbalta) worked in a majority of trials. If those two negative trials had been published in stand-alone journal articles, the way all their positive trials were published, we certainly would have counted them as published. But maybe having two negative stand-alone articles wouldn't have been too good for marketing purposes.

In any case, yes, it begs the question, what does the word "published" mean exactly? If you have a negative trial and all you do is mention it ever so briefly in a review article, can you say it's published? Selective publication deals with whether a trial is published, but it also has to deal with how it's published. If you're picking and choosing—a positive trial gets its own journal article, but a negative trial doesn't and gets buried in a review article—that meets the definition for publication bias, which means that the publication fate depends on the trial's outcome.

Another complaint came from Wyeth in the form of a letter to the editor of the *NEJM*. They didn't try to argue that their unpublished Effexor trials were really published. Instead they took the position that those trials didn't deserve to be published because they were so-called "failed" trials which, they argued, are scientifically uninterpretable.

The logic of failed trials rests on a premise we talked about earlier, the dubious notion of an infallible gold standard drug. In those Effexor trials, they had used Paxil as their active comparator or gold standard. Paxil didn't beat placebo in those trials, so the argument goes: "Since our infallible gold standard failed, something must obviously be wrong with the trial. Therefore we shouldn't be concerned that Effexor didn't beat placebo in this trial, either. Because there's something wrong with the trial, let's just pretend it never happened."

In our response, we explained why we didn't buy into the dogma regarding failed trials. We pointed to the FDA data showing that Paxil had been unable to beat placebo in 9 of the 16 trials when it was going through its own clinical trials program. What sense does it make to assume that post-approval Paxil is somehow infallibly superior to placebo while pre-approval Paxil clearly wasn't? It's the same drug, isn't it? Rather than the trial not "working," maybe the trial worked fine but neither Paxil nor Effexor did. I'll wager that there are drug classes with larger effect sizes where there's no need for this notion of "failed trials."

Back to the letter—Wyeth, just like Lilly, asserted it was transparent with its clinical trials data. They said they advocated that all trials should be published regardless of outcome. Now, even if you adhere to the dogma of failed trials, a failed outcome is still an outcome, isn't it? Either you mean all trials or you don't. You can't have it both ways. If you're waiting to see how the trial turns out and then, with 20-20 hindsight, applying some sort of post-hoc litmus test to decide whether to publish it, that again meets the definition of publication bias. We suggested that they should simply publish failed trials, make the argument in the journal article that they should be ignored, and let the readers decide.

SW: Since you're arguing that the peer-review system doesn't help with this problem of selective

publication, is there anything you would do to fix the problem?

I believe there are some loopholes in the peer-review system that can be closed. Journals could routinely request protocols for clinical trials and other interventional studies. The reviewer should review the original protocol before looking at the methods and results in the manuscript. That's how I was taught to work at the FDA. When the results come in, you don't look at them; first you go back and re-review the original protocol to see what methods they specified *a priori*. Only then do you look at the results, and you make sure those results were obtained using the original methods. That way you eliminate the phenomenon known as HARKing, or hypothesizing after the results are known, and you can be very confident that the results haven't been spun.

Drug companies can't very well spin methods with the FDA because they know the FDA has and reviews the original protocol. They also tend not to withhold entire trials from the FDA. The FDA would say, "Hey, you registered ten trial protocols with us, but you're submitting results on only five of those trials. What are you trying to pull?"

Another thing that could be done to deal with selective publication is something we touched on earlier and covered in my 2004 *PLoS Medicine* essay. We, as American taxpayers, have this gold mine of clinical trials data at the FDA that could be leveraged and made a lot more accessible to us. Once the FDA approves a drug for a certain condition, the reviews are considered public information and accessible to us through the Freedom of Information Act (or FOIA). In practice, though, it's not as easy as it sounds. For example, if you want to see the FDA review on a drug like Prozac, you have to file a FOIA request, and you might have to wait years to receive it. Even though that review has undoubtedly been requested hundreds of times, when your request comes in, they go back through the same process all over again, digging out the review, going through it and redacting out so-called trade secrets (don't get me started on that). What a waste of resources! Why not just post it on the FDA website once and be done with it? As for the reviews that are posted on the FDA website, they could be made much more user-friendly.

"I felt it was wrong that medical professionals, on the front lines treating patients, should be getting a sanitized view of drug efficacy—they need to know the whole story."

SW: What advice would you give physicians when they're looking at the journal articles and deciding how to evaluate a particular drug?

I would say *caveat emptor*. When reading journal articles, consider the possibility that the FDA's version of what happened might be different. I would especially say *caveat emptor* regarding articles dealing with off-label uses. With FDA-approved drug-indication combinations, you have the FDA's conclusion that it meets their standards for efficacy and safety. That way you know that the FDA looked at the unspun data and found that the drug beat placebo in at least two clinical trials. You have no such assurance with off-label uses. For all you know, the positive trials you're reading about could be spun, and there could be other negative trials that have never seen the light of day.

In the case of off-label prescribing, ask yourself this: if this drug is so great for this indication, why hasn't the FDA approved it? There may be a very good reason for that. Perhaps a new use has been found for a drug, but the drug is old enough that it's gone off patent. Drug companies have little financial incentive to spend millions of dollars pursuing FDA approval, which would benefit all their generic competitors. But if the off-label use is for a new drug, one that still has plenty of patent time remaining, they should have plenty of financial incentive to get FDA approval. Maybe they've tried to get FDA approval and failed, or maybe they've chosen not to apply for FDA approval. Either way, you have to wonder why.

SW: Considering the unconventional nature of your research, have you found it easy to get grants to keep it going?

Not at all. First I should clarify that there was no grant money to fund our antidepressant study. But now that it's done, you might think the objective evidence of our paper's impact would translate into grant dollars to keep this kind of work going. Unfortunately, that hasn't happened, possibly because grant money seems to be siloed according to established fields and study areas. My work has elements from several of those silos, but it seems to fall into the gaps between them.

It seems ironic that there's so much grant money to generate new data—data that may or may not lead anywhere—but so little attention to the full and truthful dissemination of that data back to the scientific community. ■

Erick Turner, M.D.
Oregon Health and Science University

Portland, OR, USA

Erick Turner's current most-cited paper in *Essential Science Indicators*, with 79 cites:

Turner EH, *et al.*, "Selective publication of antidepressant trials and its influence on apparent efficacy," *N. Engl. J. Med.* 358(3): 252-60, 17 January 2008. Source: *Essential Science Indicators* from Thomson Reuters.

Additional Information:

The paper above has also been selected as the **New Hot Paper** in Clinical Medicine for **May 2009**.

KEYWORDS: ANTIDEPRESSANTS, EFFICACY, PUBLICATIONS, FDA, NEGATIVE STUDIES, PLACEBO, PRESCRIBING PRACTICES, PEER-REVIEW SYSTEM, INSTITUTIONAL REVIEW BOARDS, PHARMACEUTICAL INDUSTRY.

 PDF

[back to top](#) 

2009 : May 2009 - Author Commentaries : Erick Turner

[Scientific Home](#) | [About Scientific](#) | [Site Search](#) | [Site Map](#)

[Copyright Notices](#) | [Terms of Use](#) | [Privacy Statement](#)