

David Egilman <degilman@egilman.com>
12/11/2003 12:10:43 PM

Record Type: Record

To: David Egilman <degilman@egilman.com>, Mabel E. Echols OMB_Peer_Review/OMB/EOP@EOP
cc: "Barry Castleman (bcastleman@earthlink.net)" <bcastleman@earthlink.net>, "David Michaels (eohdmm@gwumc.edu)" <eohdmm@gwumc.edu>, "Jennifer Sass (jsass@nrdc.org)" <jsass@nrdc.org>
Subject: RE: OMB peer review Bulletin Comments from a public member who consults with industry and for injured workers

Sorry attachment omitted from previous email.

David Egilman MD, MPH
degilman@egilman.com
508-472-2809 cell
425-699-7033 fax
8 North Main Street Suite 404
Attleboro, Ma 02703
www.egilman.com

- peer review review.doc
- omb_letter.doc

b) Peer review

Some courts have demanded that an expert's opinions be based on data that has been published in peer review journals to determine whether she can testify before a jury. In addition courts have restricted experts opinions to those expressly stated by the authors of the published work. The courts have established two presumptions: (1) anything that is published in a peer review journal is "good science" and therefore admissible; and, (2) unless a theory, fact, and/or analysis has been previously published in a peer review journal such arguments and data cannot be "good science" and cannot form the basis for admissible expert testimony.

The court in *Daubert* did not cite a single scientific authority in support of their rulings on these points.¹ The sole authority mentioned is Peter Huber who is not known to be a scientist or a legal scholar. His book, cited twice by the Supreme Court, was never peer-reviewed and was funded by a political interest group. This ruling runs counter to the prevailing views in the scientific community, which is largely skeptical of the peer review process. Thus, "[t]here is mounting evidence that peer review in the United States is not functioning well, and there is growing concern among scientists and policy makers about the soundness of the peer review system".²

Drawing upon the *Daubert* rulings, federal judges were eager to consign threshold standards to questions of who can testify and what they can say to the self-selected editors of private journals, whose decisions are inscrutable and not subject to audit or appeal.

The misuse of the peer review system as an acid test for scientific credibility is outlined in the following spurious interpretations of peer-reviewed evidence. First, the peer review system is not intended to yield "the truth, the whole truth, and nothing but the truth." Second, peer review journals do not replicate and verify the experiments, research and analytical techniques, or data reported in the papers submitted for publication. Third, peer review journals do not warrant that the ideas and information contained in the articles they publish are accurate, valid, certain, reliable, or true, or otherwise amount to "good science." Fourth, the mere fact of publication does not mean that the ideas and information reported in an article are "generally accepted" by, or represent the "consensus" views of the relevant academic community. Fifth, conversely, the fact that ideas and information have not been published in a peer review journal does not mean that they are not generally accepted, or that they are "generally rejected," or that they cannot represent good science. Finally, the peer review system should not be regarded as more rigorous and reliable than the jury system's use of cross-examination. For these reasons, we believe the courts should not rely exclusively on peer review literature, but should consider peer review literature in conjunction with other materials and with expert testimony.

i) Purposes of the peer review system

The purposes and design of the peer review system are not to produce the truth, the whole truth, and nothing but the truth. Although witnesses are obliged to swear to "tell the truth, the whole truth, and nothing but the truth," the litigation process overall does not seek absolute truth but, instead, a much more modest, yet still quite challenging goal: providing justice to the parties to a case.³ Similarly, although the overarching goal of the peer review system may be to add to the total sum of scientific knowledge and to the grand scheme of scientific understanding, the practical and immediate purposes of the peer review system are much more modest. Briefly defined, peer review is an organized method for evaluating scientific work which is used by scientists to certify the correctness of procedures, establish the plausibility of results, and allocate scarce resources (such as journal space, research funds, recognition, and special honor).⁴ The primary goal of the peer review system is to provide a relatively permanent and readily accessible documentary record of ongoing scientific research and analysis, and thus to facilitate the development and exchange of ideas. The peer review system is designed to compel authors of such reports to comply with certain minimal formalistic and stylistic standards in order to weed out essays that

¹ *Daubert v. Merrell Dow Pharms., Inc.*, 509 US 113 S.Ct.

² Daryl E. Chubin and Edward J. Hackett *Peerless Science: Peer Review and U.S. Science Policy*, 1-5 (1990). at 1. See also: *The Proceedings of the First International Conference on Peer Review in Biomedical Research May 10-12, 1989*, and *J. Am. Med. Assoc.*, n. 10, 1990.

³ *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311, 1316 (9th Cir.) (on remand)

⁴ Daryl E. Chubin and Edward J. Hackett *Peerless Science: Peer Review and U.S. Science Policy*, 1-5 (1990)

appear overtly implausible, indefensible, erroneous, fallacious, or fraudulent. The articles that survive the screening process are presented in a standardized format in journals that are made readily available to all scholars and practitioners in a particular discipline, who may then use them in their own work or may challenge them in their own writing, e.g., by attempting to replicate the experiments described or to reanalyze the results reported. Thus, *the peer reviewed journal system is designed to provide one (among many) common and convenient forums for scientific debate, and is not the final summation of existing scientific knowledge.* Peer-review is not a necessary intermediary step in the process of scientific exploration; it represents neither the beginning nor the end of the complicated morass of discovery, hypothesis, and revision that generates scientific propositions.

The editors of a peer review journal and the referees picked by such editors to review a submitted manuscript do not replicate and verify the experiments, techniques, analyses, or data reported in such manuscripts, in any way. The peer review system resembles the law review structure although there are some rather significant differences:

- (1) Articles submitted to law reviews are vetted by apprentice student-lawyers while manuscripts tendered to peer review journals are sifted by scientists who, though generally experienced, are largely figureheads on the editorial board who may have little specific knowledge of the topic at hand; and
- (2) teams of law review editors spend endless hours in painstaking efforts to check the accuracy of every cite and footnote, while peer review referees can afford to spend no more than a few scant hours in assessing the facial plausibility of a manuscript.

ii) Peer review and general acceptance

The fact that an article has been published in a peer review journal does not establish that the facts, research or analytical methodologies, or conclusions reported therein are “generally accepted” by or represent the “consensus” views of the relevant scientific community.

Although it may well be that an article published in the New England Journal of Medicine would have been accepted for publication by the Journal of the American Medical Association, and vice versa, the number of papers submitted to such prestigious journals is quite large and the number that is accepted is quite small. Although the publication acceptance rate varies from one discipline to another, as well from one journal to another, as a general rule, the more prestigious the journal, the greater are the number of papers submitted and the lower the percentage that are accepted.

In order to investigate disciplinary differences in how scientific journals evaluate submissions, one researcher collected data from the Astro-Physical Journal, Physiological Zoology, and the American Sociological Review. Referees' evaluations of submissions to these journals differed strikingly: nearly half of the referee reports for American Sociological Review recommended outright rejection, while the corresponding proportions for the other two journals were about one fourth and one tenth. Final dispositions show even greater variation, with the astrophysical journal accepting 91% of submissions, Physiological Zoology 59% and American Sociological Review 13%.⁵

Journals in fields like physics, which have high acceptance rates, the rule of thumb for referees and editors is, “When in doubt accept”. After a second or third submission, most papers are granted publication. This high acceptance rate hides the lack of reliability and validity in the process. In fields like psychology, which have acceptance rates of about 20%, the number of papers submitted far exceeds the number of pages available. The rule here is, “When in doubt reject”. Those who know the game best enjoy far higher acceptance rates. Through apt choices of topics and methods and good contacts within one’s research field, one can succeed. These wildly

⁵ See abstract of Lowell L. Hargens, *Variation In Journal Peer Review Systems. Possible Causes And Consequences*, 263 *J. Am. Med. Assoc.* 1348 (1990). See also Ann C. Weller. *Editorial Peer Review in U.S. Medical Journals*, *J. Am. Med. Assoc.* 1344, 1346 (1990) (acceptance rates for two broad categories of us medical journals range from 32.3% to 51.8%) Rick Crandall, *Peer Review Improving Editorial Procedures*, *Bioscience* 607 (Oct. 1956)

divergent standards across scientific disciplines should be, if nothing else, a cautionary tale for courts looking to apply peer-review as a boilerplate criterion.

Whether a paper is accepted or rejected by single journal may depend, in fact, more on the fame or anonymity of the authors (or the relative prestige of the universities or other institutions with which the authors are affiliated) rather than on the quality of the author's work. Caprice and bias may play a much larger role than most champions of the peer review system would wish to acknowledge.⁶ Thus, in one particularly insightful study, two researchers resubmitted a series of twelve previously published articles to the "highly regarded and widely read American psychology journals," journals that had published the same articles 13 to 36 months earlier. The only difference between the articles that had been published and the articles that were resubmitted was that the names and affiliations of the "new" authors had been downgraded. Only three (different) resubmissions were caught, thus allowing nine manuscripts to continue through to the end of the review process.

Amazingly only one of the nine articles that had been published previously was accepted for publication the second time around. In each case, all reviewers agreed about the final disposition of the manuscript. But no one cited plausibility or lack of new results as reasons for rejection. In fact, most cited poor presentation or inadequacies in the methods as the rejection criteria.⁷ Similar criticisms were offered in the study by Garfunkel, et al.⁸ Steven Hamad, editor of Peer Commentary on Peer Review, concluded (referring to the work of Peters and Ceci) that: "...any reader who thought that science was a matter of consensus should emerge from [reading the work of Peters and Ceci] thoroughly disabused of that notion. And lest he be inclined to conclude that this outcome may pertain to particularly controversial material let him sample other [author's] treatments, nowhere will he find much evidence of univocality in important current work."⁹ In another test, four researchers sent the same articles submitted to the Journal of General Internal Medicine to different sets of reviewers. One member of each team was told the name and institutional affiliation of the author, while the second reviewer received the manuscript "blind" i.e. without the author's name and institutional affiliation. The researchers found that the "blind" reviews judged the article to be of a higher quality (and more deserving of publication) than did those reviewers who knew the authors name and affiliations.¹⁰ Others noted that the "prestige of author or [affiliated] institution" influenced the decisions of a sample of medical peer review journal to publish from 30% to 50% of the time.¹¹

Another investigator reviewed prior work on bias in peer review journals and concluded that the final disposition of a manuscript is influenced by the general stance of the conclusions (positive or negative), the prior publication records of the author(s) and other factors related weakly, if at all, to the quality of the research.¹² Additional evidence of bias on the basis of the size and magnitude of the reported effect and the size of the samples used can be found in many other studies.¹³ In a particularly interesting study, Davidson reviewed 107 clinical trials of new therapies published in 1984 and found a correlation between the source of the funding and the direction of the results: 39% of the studies funded by pharmaceutical companies favored the new treatment versus 64% for

⁶ See Daryl E. Chubin and Edward I. Hackett Peerless Science Peer Review and U.S. Science Policy, 94 (1990)

⁷ Douglas P. Peters and Stephen J. Ceci, Peer-review Practices of Psychological Journals: The Fate of Published Articles. Submitted Again *J. Behavioral and Brain Sci.*, 187 (1982).

⁸ Joseph M. Garfunkel Martin H. Ulshen, Harvey Hamrick Edward E. Lawson. Problems Identified by Secondary Review of Accepted Manuscripts 263 *J. Am. Med. Assoc.* 1369 (1990)

⁹ Stevan Harnard, ed., Peer Commentary on Peer Review: A Case Study in Scientific Quality Control, 2 (1932)

¹⁰ Robert A. McNutt, Arthur T. Evans, Robert H. Fletcher and Suzanne W. Fletcher, The Effects of Blinding on the Quality of Peer Review, 263 *J. Am. Med. Assoc.* 1371 (1990).

¹¹ Ann C. Weller, Editorial Peer Reviews in US Medical Journals, 263 *J. Am. Med. Assoc.* 1344, 1346 (1990)

¹² Kay Dickerson. The Existence of Publication Bias and Risk Factors for its Occurrence, 263 *J. Am. Med. Assoc.* 1385 (1990).

¹³ Michael Mahoney. Publication prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System, *J. Cogn. Ther. and Res.* 161 (1977); In: Spiegel-Rosing. Science Studies: Bibliometric and Content Analysis 7 *Soc. Stud. Sci.* 97 (1977); T.D. Sterling. Publication Decisions and their Possible Effects on Inferences Drawn from Tests of Significance - or Vice Versa, 541., *Am Stat. Assoc.* 34 (1959); T.C. Chalmers. R.S. Kofi. and O.F. Grady, A Note on Fatality in Serum Hepatitis. 49 *Gastroent* 22 (1965); B. Sommer. The File Drawer Effect and Publication in Menstrual Cycle Research. 11 *Psych. Wom. Q.* 233 (1987). Von Bakanic, Clark McPhail. and Rita Simon. The Manuscript Review and Decision-making Process. 52 *Am. Soc. R. Ev.* 631 (1987); and C.B. Begg, and J.A. Berlin. Publication Bias and Dissemination of Clinical Research, 81 *J. Natl. Cancer Inst.* 107 (1989).

studies not funded by pharmaceutical companies.¹⁴

iii) The inadequacies of peer review

Some observers have noted that many journals rather routinely ignore or tolerate, but in any event reproduce, inaccurate statistics. Thus, as early as 1966, an article in the *Journal of The American Medical Association* disclosed that an analysis of the quality of statistics reports that had appeared in ten different medical peer review journals “suggested that from a statistical viewpoint” only 41 (28%) of 149 published studies were acceptable, 7 (5%) should have been rejected, and 101 (68%) should have been revised before publication, that is, only that submitted manuscripts contained significant errors but that these errors went undetected (or at least uncorrected) by the editors and referees of those ten medical journals.¹⁵ Gardner and Bond’s own and more recent study of peer review articles published in the distinguished *British Medical Journal* revealed equally alarming results - only 11% of incoming manuscripts contained accurate statistical reports; that figure improved to 84% after publication.¹⁶ To make matters worse, many if not most journals do not do what the *British Medical Journal* routinely does — independently assess and correct the statistics cited in articles selected for publication.

Thus, a recent survey revealed that although statistical analyses were prominent in 54% to 73% of the articles published in a sample of American biomedical journals (depending on the classification of the journals by the researchers) the percentage of journals that ‘always or usually use a statistical review’ ranged from 30% to 37%.¹⁷

Another recent study involved sending 23 manuscripts that had been accepted for publication, after review and revision, by the *Journal of Pediatrics* (but not published at the time) to two additional referees for evaluation.¹⁸ The researchers discovered that, although there was substantial agreement between the first and second set of reviewers, “the initial two-part review failed to identify a significant number of concerns that probably should have been corrected, or at least addressed, before publication.”¹⁹ Still another study checked 150 randomly selected citations and quotations taken from articles published in three medical journals. Forty-eight percent (72) of the citations had some error, while nine percent (13) had major citation errors.²⁰ Moreover, of the 137 citations with no major errors, 23% (32) had major quotation errors, sometimes saying that a source supported a particular proposition, when it did not, as for example: An article by Lowe is cited to support an increased risk of esophageal cancer with alcohol consumption; the article is, in fact, about treatment and contains absolutely no mention of etiology.²¹

The failure of peer review journals to adequately screen out false, substandard, or fraudulent reports might be excusable if the journals did a creditable job of subsequently informing their readers about the existence of the published mistakes (and the dangers of relying on reports containing such flaws) once the errors were noticed. Unfortunately, the journals do not do much better after the fact than they do before.

Inasmuch as the peer review system is conducted by human beings, there is no reason to expect it to be perfect, and it is not. This imperfection would not matter as much if major errors were caught quickly and corrected completely. Even after a mistake is discovered, however, it can take decades to correct completely. Indeed peer

¹⁴ R.A. Davidson. Sources of Funding and Outcome of Clinical Trials, I. *J. Gen. Intern. Med.* 155 (1986). See generally. Daryl E. Chubin. A Soap Opera for Science. 37 *Bioscience* 259 (April 1987); Rick Crandall. Peer Review: Improving Editorial Procedures. *Bioscience* 607 (Oct. 1986); Chubin and Hackett, *supra.* at 8, 11-12. 94-103.

¹⁵ Martin J. Gardner and Jane Bond, An Exploratory Study of Statistical Assessment of Papers Published in the *British Medical Journal*, 263 *J. Am. Med. Assoc.* 1355 (1990), citing S. Schor and I. Karten, Statistical Evaluation of Medical Journal Manuscripts, 195 *J. Am. Med. Assoc.* 145-50 (1966).

¹⁶ *Ibid.*

¹⁷ Ann C. Weller. Editorial Peer Review in US Medical Journals. 263, *J. Am. Med. Assoc.* 1344 (1990).

¹⁸ Joseph M. Garfunkel Martin H. Ulshen, Harvey Hamrick Edward E. Lawson. Problems Identified by Secondary Review of Accepted Manuscripts 263 *J. Am. Med. Assoc.* 1369 (1990)

¹⁹ *Ibid.* at 1370

²⁰ James T. Evans, Howard I. Nadjari, and Sherry A. Burchell, Quotational and Reference Accuracy in Surgical Journals 263 *J. Am. Med. Assoc.* 1353 (1990).

²¹ *Ibid.* at 1354

review journals seem unusually reluctant to acknowledge mistakes and to correct those mistakes that are conceded. For example, one contemporary probe recounts the saga of a paper that reported “unique cell cultures” that were the cause of much celebration when they were discovered a generation ago and its use and reliance by many scientists since that time.²² By the early 1970’s it was learned, though, that the putatively “unique” cultures were anything but. Nevertheless, those portions of the scientific community that were most responsible for the problems (and the ones best situated to rectify the errors) were decidedly reluctant to issue the necessary retractions. For this reason, it is believed that many scientists may still have not received the word and may yet be unaware of the contamination of the cell cultures.²³

More generally, according to a recent report by two investigators, “Little is known about the ultimate scientific fate of retracted, invalid literature. We identified 32 completely retracted articles ... and measured their subsequent use in the scientific literature.... After retraction, these studies were cited in support of scientific concepts 733 times or roughly nine times apiece. Comparison with a control group revealed that retraction reduced subsequent citation by approximately 35%.”²⁴ As a result invalid work is not being effectively purged from the scientific literature.²⁵

iv) Other venues for the presentation of scientific information

Just as the mere fact of publication does not signify “general acceptance,” the fact of non-publication does not indicate that ideas and information have “generally rejected” or that they are not and cannot be “generally accepted”.

There are at least four other ways in which scientists can introduce new ideas and information without publishing in peer review journals, ways that may be at least as effective, at least as impressive, and at least as well-established, legitimate, and well-regarded as publishing articles in peer review journals. First, scientists frequently are invited to give addresses at seminars, conferences, and colloquia; the printed versions of these speeches are usually reprised in the printed proceedings or reports of those conferences. Although the speeches are not formally peer reviewed either before or after they are delivered, the fact the authors were selected by their peers as speakers serves much the same quality control function. Second, “electronic bulletin boards” similarly offer instant dissemination by the authors — and, at times, instant questioning and critiques by those scientists who regularly peruse these electronic colloquia.

Third, reports and studies generated by and for federal, state, and local governmental agencies (and proposals submitted to such agencies by researchers seeking grants) likewise provide a way for scientists to exchange ideas and information. As noted in the National Research Council’s recent study on so-called “Gray Literature,” government agencies rely heavily, indeed, at times almost exclusively, on such reports, notwithstanding the fact that they are not published in peer review journals.²⁶ Finally, private industry, corporate research consortia, and independent think tanks correspondingly commission the same sort of reports and internal company studies from both staff scientists and independent consultants and rely on them just as heavily as government agencies, despite the fact that, once again, these papers are not formally published in peer review journals.

Some of the reports thus generated, whomever the sponsor and whatever the forum, are “pure or basic science”, in the sense that they have no immediate and specific application; these scholarly reports most closely approximate the pristine studies that are the ideal of peer review journals. Other studies and reports may lie closer to opposite extreme of “applied science,” that is, utilitarian accounts that are generated in response to particular technical questions or specific technological or biomedical problems — such as queries about the safety of a bridge design, evaluations of best location of a gas tank in a new car, predictions of how a new airport would affect highway traffic, or estimations of the side-effects of a new pharmaceuticals.

²² Michael Gold, *A Conspiracy of Cells* (1956).

²³ *Ibid.*

²⁴ Abstract of Mark. P. Pfeiffer and Gwendolyn L. Snodgrass, *The Continued Use of Retracted, Invalid Scientific Literature*. 263 *J. Am. Med. Assoc.* 1420 (1990).

²⁵ See also Paul F. Friedman, *Correcting the Literature following Fraudulent Publication* 263 *J. Am. Med. Assoc.* 1416 (1990).

²⁶ *Environmental Epidemiology, Health, and Hazards* (1991)

Perhaps the paradigmatic examples of high quality science that is almost never published are medical evaluations of particular patients with a specific disease. In such a situation, the questions are distinctively practical ones — whether, for example, to treat a malignancy with radiation or chemotherapy. Yet, whether the work undertaken yields a new, theoretical advance in basic science or results in new, practical applications of venerable principles, scientists will assess them on the basis of what they say, how they say it, and how useful they are, and not on where and whether they were peer reviewed ahead of time. This work may just have well appeared in surgical notes or may have been presented at a seminar. The fact remains that some of the best science is the most specific science. Most importantly, this sort of science may not be publishable not because it is wrong, or outside the mainstream of generally accepted views, but simply because a peer review journal editor might decide that it would not be of sufficient interest to enough of her subscribers.

There is also the problem of bias: (1) bias against some authors and scholars;²⁷ (2) bias in favored other authors and other institutions;²⁸ (3) bias against negative results;²⁹ (4) bias against or in favor of particular results;³⁰ (5) bias against truly innovative ideas;³¹ and (6) bias against nuts-and-bolts “clinical research” as opposed to more ethereal “theoretical research.” As to the last point, one researcher studied patterns associated with the publication of the results of clinical trials and concluded that the full reporting of the results of clinical trials is very irregular. “Substantial numbers of clinical trials are never reported in print and among those that are, many are not reported in sufficient detail to enable judgments to be made about the validity of their results.”³²

v) Non-peer reviewed data provides a reliable basis for causal relationships

The best illustration that scientific evidence can be reasonably relied upon by experts in the field, even if it has not appeared in a peer-reviewed journal, can be seen in the working of the federal agency created to test the safety of chemicals. The Secretary of the then Department of Health, Education and Welfare established the National Toxicology Program (NTP) in 1978 to test chemicals for their health effects and to develop and validate improved test methods. NTP routinely conducts experimental studies to determine whether substances are carcinogenic or have other adverse health effects. The results of these studies are issued directly by NTP in the form of a Technical Report.

As of July 1992, NTP had issued four hundred sixteen Long Term Technical Reports and Short Term Toxicity Study Reports for hundreds of chemicals. These Technical Reports are widely relied upon by state and federal agencies for a host of decisions on regulating toxic substances, resolving disputes between conflicting studies, and similar purposes.

NTP employs an external scientific review process to insure the integrity of its research, but results are not published in a peer-reviewed journal and would not satisfy most interpretations given to the Daubert ruling by the lower courts. It would be inappropriate to publish the results of an NTP cancer study in a peer-reviewed journal for the following reasons:

First, if the NTP determines that a substance is carcinogenic, withholding the release of the study results pending acceptance in a peer-reviewed journal would delay the initiation of the appropriate regulatory process to consider how to protect the public. In fact, NTP anticipated this kind of situation and established a procedure whereby, if preliminary evidence suggests that a chemical is highly carcinogenic, federal regulatory agencies, manufacturers, trade associations, labor unions, public interest groups, and others are affirmatively notified even before the

²⁷ *ibid.*

²⁸ *ibid.*

²⁹ Dickerson, *supra*.

³⁰ Boyce Rensberger, “DNA Fingerprinting Is Disputed,” *Washington Post*, Dec. 20, 1991, at A3 and “Science Editor Denies Yielding to FBI Pressure,” *Science & Government Report*, Feb. 15, 1992 at 1. (e.g., The Journal of Occupational and Environmental Medicine has a reputed tendency to publish negative studies on potential environmental and workplace hazards and receives its financing and editorial support from industry and corporate medical directors.)

³¹ See, e.g., *Science Beyond the Pak*, 249 Science 14 (July 1990) and See generally, Thomas S. Kuhn, The Structure of Scientific Revolutions (2d ed. 1970)

³² Ian Chalmers, Underreporting Research is Scientific Misconduct, 263 J. Am. Med. Assoc. at 1405 (1990)

Technical Report is completed. To await the completion of a journal's peer review process could unnecessarily prolong the public's exposure to a harmful substance.

Second, the vast majority of the studies used to evaluate the health effects of chemicals are conducted using widely accepted and very explicit scientific protocols. Because these protocols have become so commonplace, the studies are not likely material for journals that necessarily must seek research on new scientific methods. This absence of peer review, however, has no bearing whatsoever on the scientific quality or probative value of NTP Technical Reports or their acceptance in the scientific community.

Moreover, data not published in peer-reviewed journals are commonly used for many other varieties of scientific and regulatory purposes. Under a broad array of federal statutes, including the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), The Federal Food, Drug and Cosmetic Act (FFDCA), the Toxic Substances Control Act (TSCA), and others, manufacturers are required to test their compounds for health and environmental effects as a condition for receiving government approval for sale and use. Generally, these studies, conducted in accordance with standardized test procedures in order to ensure their validity, are submitted to the appropriate federal government to determine how the compounds should be regulated.

For example, FIFRA requires that all pesticides sold or distributed in the United States must be registered by the U.S. Environmental Protection Agency.³³ Such data are used to determine whether the statutory standard for registration has been met, that a pesticide "when used in accordance with widespread and commonly recognized practice ... will not generally cause unreasonable adverse effects on the environment."³⁴

The EPA publishes detailed guidelines specifying the type of data necessary to support a pesticide registration.³⁵ The regulatory guidelines, and the underlying Pesticide Assessment Guidelines and Standard Laboratory Practices, explicate requirements for over a hundred different studies on the health and environmental effects of individual pesticides. Yet nowhere in the federal statute or the regulatory guidelines is it required that a study be published in a peer-reviewed journal as a precondition for consideration in the EPA's pesticide registration decision. Rather, in its general policy regarding acceptability of data, the Agency will:

- (1) Determine whether the data submitted to fulfill the data requirements specified in this part are acceptable. This determination will be based on the design and conduct of the experiment from which the data were derived, and an evaluation of whether the data fulfill the purpose(s) of the data requirement. In evaluating experimental design, the Agency will consider whether generally accepted methods were used, sufficient numbers of measurements were made to achieve statistical reliability, and sufficient controls were built into all phases of the experiment;
- (2) Evaluate the conduct of each experiment in terms of whether the study was conducted in accordance with good design and laboratory practices, and whether the results were reproducible.³⁶

To date, over six hundred pesticide active ingredients have been federally registered for use, largely on the basis of studies not published in peer-reviewed journals.

The Agency not only relies on unpublished data to establish null associations, but uses the data to evaluate the hazards posed by the substances, both environmental and health-related. Section 6 of FIFRA authorizes the Administrator to cancel or suspend pesticide registration if the product does not comply with FIFRA or generally causes unreasonable adverse effects on the environment.³⁷ The Administrator's cancellation or suspension decision is not limited to consideration of studies published in peer-reviewed journals. In fact, the statute requires pesticide registrants at any time to submit to the EPA "additional factual information" regarding a pesticide's

³³ FIFRA § 3(a), 7 U.S.C. § 136(a)

³⁴ FIFRA § 3(c)(5)(D), 7 U.S.C., § 136a(c)(5)(D)

³⁵ FIFRA § 3(c)(2), 7 U.S.C. § 136a(CX2)

³⁶ 40 CFR § 158.80 (1988)

³⁷ FIFRA § 6, 7 U.S.C. § 136d

unreasonable adverse effects, a category clearly far broader than only information published in a peer-reviewed journal.

The EPA routinely cancels or suspends pesticides on the basis of studies not published in peer-reviewed journals. For example, on October 7, 1986, EPA issued an emergency suspension of a widely used pesticide ingredient, dinoseb. The suspension order was based upon scientific data EPA summarized as demonstrating that “dinoseb is a developmental toxicant in laboratory animals.”³⁸ These data were unpublished studies submitted to the EPA by dinoseb’s manufacturers but deemed by the Agency to be scientifically valid and sufficient indicators of cause for concern to warrant suspension of the pesticide.³⁹

Analogously, in accordance with the Toxic Substances Control Act, manufacturers must immediately disclose to the EPA any adverse effects caused by their products.⁴⁰ On September 15, 1992, for example, by letter IBM Corporation notified EPA under TSCA Section 8(e) that diethylene glycol dimethyl ether and ethylene glycol monoethyl ether acetate caused adverse reproductive effects, including increased spontaneous abortions. This disclosure was based on a preliminary, and yet unpublished, analysis of 1980-1989 data collected as part of an epidemiological study of IBM’s semiconductor manufacturing employees and their work environment at two sites.

c) The reliability of peer-review vs. cross-examination

Some observers have advanced the notion that the peer review system is much more rigorous and thereby more likely to reveal truths and uncover falsehoods than cross-examination.⁴¹ That the peer-review system is more reliable, strenuous, and exacting than cross-examination disregards the court’s institutional use of cross-examination as a potent vehicle for the truth.

The myth that peer review journals provide rigorous scrutiny of submitted papers is largely based on the underlying fable that peer review editors and referees duplicate the experiments and analyses discussed in those papers. As noted above, they do not. Indeed, the meager amount of time spent by unpaid volunteer referees in reviewing submitted manuscripts hardly allows for painstaking replication or other forms of rigorous scrutiny. It is a little-known but well established fact (ironically, well established in the peer review literature) that the average time spent on a review ranges between ‘less than’ 2.0 hours and 2.4 hours, depending on the study.⁴²

By contrast, cross-examination often subjects a witness to scrutiny that is longer lasting, more intensive in depth, and more extensive in the breadth of subjects examined. First, a witness must make herself available a sufficient length of time before she testifies to counsel for the opposing party. These lawyers may then depose the expert not only on every minute fact substantiating her testimony, but on any belief, trait, or past action that may provide a relevant or irrelevant characterization of her testimony. There are literally no limits to what a witness can be forced to answer or account for in her pre-trial testimony. There are rarely time limits placed on this testimony or on the cross-examination of a witness inside the courtroom. Second, once on the witness stand, a witness may be obliged to endure a line-by-line exegesis of his article and asked to explain and defend every point. Third, the questions asked on cross-examination are not mercifully limited to the specific parameters of the paper or report that is the foundation of the expert’s conclusion; instead, almost anything the expert has done, said, or written is fair game. For example, on cross-examination an expert witness can be subject to close questioning not only on the basis of the specific sources of her testimony, but can be impeached by anything she had said or written that is potentially inconsistent. Her motives and credentials can be scrutinized in ways unimaginable to peer review editors and referees. The peer review system has nothing comparable. All in all, whatever the pressures are generated by the peer review system, cross-examination is, quite literally, a far more trying experience.

³⁸ 51 Fed. Reg. 36636 (October 14, 1986)

³⁹ Love v. Thomas, 858 F.2d 1347 (9th Cir. 1988), cert denied 490 U.S. 1035 (1989)

⁴⁰ Toxic Substances Control Act, 15 U.S.C. § 2607(e)

⁴¹ Chubin, Hackett, Ozonoff, Clapp. Brief of Amici Curiae in Support of Petitioners in Daubert v. Merrell Dow Pharms., Inc., 509 US 113 S.Ct.

⁴² What do peer reviewers do?. 263 J. Am. Med. Assoc. 1341. 1342 (1990) (references for the British Medical Journal averaged less than two hours reviewing each paper submitted for publication)

Indeed, the Supreme Court has recognized that cross-examination is the greatest legal engine ever invented for the discovery of truth. “The opportunity for cross-examination ... is critical for ensuring the integrity of the fact-finding process,” and cross-examination is “the principal means by which the believability of a witness and the truth of his testimony are tested.”⁴³

In fact, in the Supreme Court’s Daubert decision, the Court grants that impeaching an expert with her own testimony is the most reliable way of exposing pseudoscientific assertions on the witness stand.

Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.⁴⁴

However, in the Court’s opinion, the admissibility of scientific evidence remains the ultimate safeguard against the use of spurious scientific evidence. For instance, through the Court concedes the testimony of experts can be “too particular, too new, or of too limited interest to be published” (e.g., in the reanalysis of data or in opinions on individual causation), the Court nonetheless validates the use of peer-review to screen out “substantive flaws in methodology”.⁴⁵ Moreover, the Daubert decision grants the courts the freedom to direct judgement on matters where the strength of evidence supporting causation fails to meet the preponderance of the evidence standard.

[I]n the event the trial court concludes that the scintilla of evidence presented supporting a position is insufficient to allow a reasonable juror to conclude that the position more likely than not is true, the court remains free to direct judgement.⁴⁶

⁴³ **Daubert v. Merrell Dow Pharms., Inc.**, 509 US 113 S.Ct.

⁴⁴ **ibid.**

⁴⁵ **ibid.** (The court cites Relman and Angel, How good is peer review?, *New. Eng. J. Med.* 827 (1989) in arguing that submission of materials to peer-review journals reliably screens out substantive flaws in methodology, and furthermore, should be used presumptively for meeting the reliability component of a Daubert challenge to expert qualification to testify.)

⁴⁶ **Ibid.**

December 11, 2003

Dr. Margo Schwab
Office of Information and Regulatory Affairs Office of Management and Budget
Executive Office of the President
725 17th St NW
New Executive Office Building
Room 10201
Washington, DC 20503

By email: OMB_peer_review@omb.eop.gov

Dear Dr. Schwab:

I am writing to you with my comments on OMB's proposed bulletin on peer review and information quality.

OMB's proposal is a solution to a problem which has yet to be defined. Ironically this proposal has not been peer reviewed. There is a body of literature which has examined the quality of peer review. (See attached) Most of this literature indicates that, even in the academic sphere, peer review is of poor quality. This is one of the reasons why peer-reviewed journals have repeatedly published fabricated data and studies.

More importantly, OMB's proposal will enhance corporate America's ability to influence regulations. As Adam Smith first noted, the capitalist system can only work efficiently if all costs are borne by the manufacturer or seller of a product. The major cost elements that are unaccounted for in the current system are the costs corporations inflict on workers, consumers and the environment when they sell products.

Corporations have the most to gain by limiting regulation, which is the only mechanism available to the government to force corporations to bear the true costs of production and sale of their products. Because of this, corporations have a great incentive to corrupt medical and scientific literature and they have done so during this century. This corruption includes peer-reviewed literature on topics ranging from tobacco to lead.

Your proposal will give corporations a fourth or fifth crack at limiting regulations to protect the public. It should not go forward.

However, since its passage is inevitable, I urge you to at least cover it with a fig leaf of honesty. Scientists who receive funding from corporations whose products will be subject to regulations must be barred from the peer review process. Disclosure of conflicts does not remove the conflict. There is no way to balance conflicts. Scientists who work for financially interested corporations cannot be “balanced” by honest scientists; they must be excluded from the process.

Sincerely yours,

David Egilman MD, MPH
Clinical Associate Professor Department of Community Health
Brown University
degilman@egilman.com
508-472-2809 cell
425-699-7033 fax
8 North Main Street Suite 404
Attleboro, Ma 02703
www.egilman.com