

Flaws are found in validating medical studies Many see need to overhaul standards for peer review

By Michael Kranish, Globe Staff | August 15, 2005

WASHINGTON – They are two of the most widely publicized pieces of medical research in recent years: Reports in prestigious journals declared that women who underwent hormone replacement therapy, and people who ingested large amounts of Vitamin E, had relatively low rates of heart disease.

Each study was vetted by peer review, the basic process for checking medical research, in which other researchers judge whether papers meet scientific standards.

But after research contradicted those studies – frustrating anyone who had followed their recommendations – some specialists began looking at whether peer review had failed to identify serious flaws in the research.

But the specialists found that it was almost impossible to discover what had happened in the vetting process, since peer reviewers are unpaid, anonymous, and unaccountable. Moreover, their reviews are kept confidential, making it impossible to know the parameters of the reviews

Now, after a study that sent reverberations through the medical profession by finding that almost one-third of top research articles have been either contradicted or seriously questioned, some specialists are calling for radical changes in the system.

In advance of a world congress on peer review next month in Chicago, these specialists are suggesting that reviewers drop their anonymity and allow comments to be published. Some are proposing that peer reviewers be paid to ensure a more even quality of review and analysis among all journals.

Dr. Drummond Rennie, who relies on review as deputy editor of *JAMA*, the *Journal of the American Medical Association*, said of the process, “The more we look into it, the harder it is to prove whether it does good or bad.”

Rennie has called for greater study of whether peer review improves research, and he has a personal policy of disclosing his name when he reviews articles.

“It would be lovely to start anew and to set up a trial of peer review against no peer review,” Rennie said. “But no journal is willing to risk it.”

Rennie’s journal published the study, which said that subsequent research had found that almost one-third of the top papers that appeared in top journals over a 13-year period from 1990 to 2003, had been either contradicted or found to have potentially exaggerated results. All the articles had undergone vigorous peer review, leading to questions about whether problems should have been caught by reviewers.

The author of that study, Dr. John Ioannidis, an adjunct professor at the Tufts University School of Medicine, said that flaws in the system were not solely responsible for the problems with the initial studies, but he said that they may be “part of the puzzle” that should be examined to improve research.

Ioannidis has proposed making peer reviews public so that “one could see whether someone said, ‘This is a great study, publish it,’ or whether there was constructive scientific thinking, comment and criticism.” He noted that he could not examine any peer reviews, including those for the hormone replacement and Vitamin E studies, because of the confidentiality surrounding peer review.

Under the system of peer review, a researcher submits findings to a journal for publication. Along with a review by editors, the article is sent to several specialists in the field.

These reviewers are not paid for their time, their names are usually not published, and their comments usually remain secret. They are usually not allowed to contact the researchers directly to ask questions, and they do not try to replicate the research.

The system has often had successes; many journal editors say peer review has saved countless prominent scientists from publishing seriously flawed work, and has spared the public from following mistaken medical advice.

But peer review also lacks consistent standards. Procedures vary among the world's 10,000 or so journals. A peer reviewer often spends about four hours reviewing research that may have taken months or years to complete, but the amount of time spent on a review and the expertise of the reviewer can differ greatly, especially at lesser-known journals.

"It has been bandied about as a sort of 'Good Housekeeping Seal of Approval,'" said Marcia Angell, former editor of the *New England Journal of Medicine*. "It is only as good as the peer reviewers and editors."

The increasing focus on peer review will be highlighted next month, when dozens of journal editors and specialists in peer review meet in Chicago. Dozens of papers will be presented on topics that include whether peer review adds value, and whether conflict-of-interest rules are working.

J. Scott Armstrong, a professor at the University of Pennsylvania who has spent years analyzing peer review, has voiced hope that the conference will lead to radical change in the way journals conduct peer review. The system, he said, is outdated and outmoded.

Pointing to a move by some journals to put their information on the Internet and to publish the

names of reviewers, he predicts that the current system of anonymous reviewers will be replaced by a version of Amazon.com, in which scientists from around the world contribute their thoughts to constantly updated research.

Change is not likely to come, however, at the upcoming Fifth International Congress on Peer Review and Biomedical Publication. That gathering is intended as a forum for discussion, rather than decision-making.

There is no governing body that defines what constitutes good peer review, or that demands that certain standards be followed.

Moreover, some of the editors at some of the large journals are not eager to change the system. Dr. Jeffrey Drazen, who is the editor of the *New England Journal of Medicine*, said he supports keeping the review system anonymous and unpaid.

"We don't think the system is broken and needs to be overhauled," Drazen said.

Drazen also said peer review is not necessarily at fault when a study is not replicated by subsequent research. "As a scientist, the things that give me the most joy is when someone is able to replicate something I published," Drazen said. "That means that you got it right. But sometimes people cannot replicate things. It is a mistake to view it as black and white . . . if you do a second study but can't replicate the primary findings, it doesn't necessarily mean the original research was wrong."

Ioannidis, the author of the study on flawed research, said he had examined articles from top journals published from 1990 to 2003, and had found that 16 percent of those studies were later contradicted, and that another 16 percent were not found to have had as strong a result in subsequent research.

Many factors led to the conflicting results, he said, including the fact that scientific research is often updated when larger or better-controlled trials are conducted. But flaws in the initial studies, including integrity and methodology, could not be ruled out.

Some journals are trying to improve the system by making themselves more open to the public. The Public Library of Science publishes a magazine called *PLOS Medicine*, which charges authors \$1,500 per article but which provides its journals online for free.

PLOS Medicine also encourages peer reviewers to reveal their identity, but it does not demand it. The journal's senior editor, Barbara Cohen, said some reviewers want anonymity out of concern about retribution, which she described as "you trashed my paper at *Nature*, now I'm trashing yours at *Science*," referring to two leading journals.

Cohen also said she is sympathetic to younger peer reviewers who fear that providing criticism of a senior person in the field will hurt their career. This is a common complaint among reviewers.

But given the high number of studies that end up either wrong or deeply flawed, much of the

medical profession is looking for new ways to examine research.

Armstrong, the professor who has read dozens of studies on peer review, cited numerous embarrassing incidents that he said had called the peer review process into question.

In one study, for example, researchers submitted a plagiarized paper to 110 journals, but only two publications recognized the problem.

In another study, researchers examined 18 papers that had been published in peer-reviewed journals by a person who later admitted scientific fraud; they found that 16 of the papers had an average of 12 errors each.

One such error was that "the father in one family had his first child at age eight and the next at age nine," Armstrong wrote.

Michael Kranish can be reached at kranish@globe.com.
© Copyright 2005 The New York Times company