Peer review plays a critical role in determining the nature and level of medical practice. As a consequence, the process has an important impact on patient outcome.

Not only does peer review directly influence the content of the medical literature, the process indirectly provides the basis of medical literature, of pre- and postgraduate medical education, of modifying the composition of bedside teaching, and of influencing the use or rejection of various medical innovations. The composition of medical school faculties and their promotion are strongly influenced. Frequent contributions to the peer-reviewed literature are regarded as an important or even mandatory prerequisite; failure to be published is regarded as strongly negative.

Each day, at thousands of morning reports, attending rounds, consultations, in hospitals and physicians' offices, the products of peer review are used as guides to the management of vast numbers of patients.

In the past several years, there has been a growing uneasiness with the quality of the process. There is widespread dissatisfaction with peer review in the humanities. Should these concerns be extended to medicine? Few answers have emerged. Such basic questions as "how good is the process?", "how valid?", "how fair?", "do the benefits outweigh the risks?", "what is the impact on patient outcome?" are there alternatives to the present process?" could the process be improved? largely have gone unanswered.

These questions are more pressing for the medical literature than for the strictly scientific literature. For the latter, the human consequences are largely indirect. In the case of the medical literature, the ultimate question is whether the present peer review process improves or is detrimental to patient outcome. There are no simple answers, but there is evidence that there is cause for concern.

Potential benefits include:

1. A reduction in the number of gross errors which appear in the literature.
2. An automatic mechanism for quality control.
3. A reduction in the number of published manuscripts to a manageable figure.
4. The enforcement of some set of standards or norms for practice.
5. The stimulation of efforts to produce better work and better writing.

Potential risks include:

1. Exclusion of new ideas or approaches which conflict with orthodoxy and, thus, retardation of medical progress.
2. Perpetuation of wrong ideas or practices which are part of orthodoxy.
3. Providing an unwarranted stamp of approval for work which is wrong or shoddy.
4. Needless delay in providing information which is helpful to patients, resulting in poorer patient outcome.

Evidence of Inadequacies in the Peer Review System

Some of this evidence is at the anecdotal level. The description of the Krebs' cycle by Hans Krebs was rejected by a famous British journal. The Turkish biochemist who reviewed the manuscript had studied cat pectoral muscle, which is notoriously deficient in Krebs' cycle intermediaries. He reasoned that if he couldn't find the intermediates, they must not exist. The harrassing reviews of several experts almost prevented the publication of the original paper by Berson and Yalow describing radioimmune assay. Gibbs, the father of one of the most important ideas in medicine, the concept of free energy, was driven to publish in the Connecticut State Journal of Chemistry. These are examples of the difficulties of publishing new and revolutionary ideas.

There are also numerous examples of the publication of medical fantasies. We must not forget that bronchial
asthma is caused by worms, also toxemia of pregnancy. Alzheimer's disease is ameliorated by intrathecal cholinergic drugs, and a substantial number of patients can be cured of a probably nonexistent disease by drastic surgery.

Are these florid examples aberrations, or is there more convincing evidence? A substantial number of papers published in the "best" medical journals use incorrect statistical methods and scientific design. What can be said about the conclusions drawn from these analyses?

The only "clinical" trial of peer review we know of was conducted in the field of psychology. Twelve papers, which had been accepted and published in the most prestigious psychology journals, were resubmitted to the various journals, unchanged except for the titles of the submitting institutions. Three of the resubmissions were recognized by the reviewers and editors and thus were not suitable for the experiment. Eight of the remaining nine resubmissions were rejected the second time around by a total of 36 reviewers. The most common reason for rejection was unacceptable methodology. This trial provides a number of lessons. Most of the resubmitted articles were not recognized the second time around, hardly the stuff of which immortal contributions are made. How could there be such disparity in expert opinion? The results are not a strong endorsement of expert judgment. And is there a coterie of experts who cannot recognize "unacceptable" methods in their areas of expertise? If the same pattern exists for medical articles, then, indeed, our patients are in trouble!

The most pressing reason to believe that there are real problems is the growing proliferation of iatroepidemics, systematic errors in medicine which lead to harm or death in masses of patients. A partial list of some current iatroepidemics in chest and heart disease might include: barbiturates for global cerebral ischemia, routine chest x-ray examinations, V/Q lung scanning for the specific diagnosis of pulmonary embolism, the vast overuse of Swan-Ganz catheters, electronic fetal monitoring, the vast overuse of chronic cardiac pacemakers for the sick sinus syndrome, sodium deprivation for the treatment of eclampsia, extracranial-intracranial bypass for strokes, methyldprednisolone for septic shock, and nonspecific lung biopsy.

Even though there may be controversy about some of the individual listings, these originated in and were perpetuated by peer-reviewed journals.

To be fair, the causes of iatroepidemics are multifactorial, but peer review does play an important role.

One of the most important problems with the present system is that its operation may be largely indifferent to the consequences of a given paper for patient welfare. Some papers do have enormous patient consequences.

For example, the clinical trial which demonstrated that lumpectomy and radiation were as effective as radical mastectomy for the treatment of small cancers of the breast ran into serious review difficulties which delayed publication for many months. In fact, the lay press interceded and perhaps hastened publication. The fundamental basis of review difficulties can be guessed. A number of surgical experts believed in radical mastectomy. Undoubtedly, from the best of motives, a series of objections to the details of the paper were generated which delayed publication. Another mechanism (see below) could have expedited publication. The major point is that the delay resulted in a substantial number of women undergoing an unnecessary psychologic and physically mutilating procedure which could have been (perhaps should have been) avoided.

Another example involves the use of barbiturates in the treatment of global cerebral ischemia. A paper which was flawed by methodologic and interpretive problems, seemed to show that monkeys with ischemic brains, administered massive doses of barbiturates, did better than control animals. This study led to the use of barbiturates in innumerable patients. There was the usual flurry of anecdotal reports which either supported or interdicted the practice. To the credit of the innovators of the method, an international, well-designed clinical trial was organized. In brief, barbiturates don't work, and the risk-benefit ratio of its use is unfavorable. The paper which summarized the results of the clinical trial languished in the review process for months or years. During this period, patients with global brain ischemia continued to be treated with a useless and potentially dangerous form of therapy.

Conventional Defense of the Present System

There is almost no one intimately involved in the peer review process who does not sense some limitations. However, the usual evaluation paraphrases Winston Churchill's description of democracy, ie, that peer review is not perfect but is the best system which has been devised.

This is actually a rather gratuitous conclusion. Given lack of an evaluative attempt and failure to generate modifications or alternatives, such a rationalization is self-serving.

Subjective Elements in Review

It has not been emphasized sufficiently that peer review involves substantial (at times overwhelming) subjective judgments. This is indicated by various phenomena. Two expert reviews of a given paper are widely disparate. This could only arise if the judgment of at least one expert is wrong. This problem is usually
resolved by bringing in a third expert and the decision to publish is based on two of three. For some journals, even one unfavorable review may diminish the priority and result in disapproval. These are often regarded as the "best" journals.

There is the phenomenon of "the new boy on the block." Contributions by individuals not in the field may have harder going than from well established (less threatening) experts. This despite the theoretic advantages of bringing input from outside disciplines. There is a parallel phenomenon, "the old boy on the block." A well-established worker in a given field may be able to publish an article which would have been rejected had it been submitted by a lesser known figure.

Interestingly, purely objective judgments, such as the adequacy of the statistical analysis and, more importantly, whether the material to be published has been subjected to an appropriate rigorous clinical trial, are not involved in most judgments.

Are There Ways the System Could Be Modified?

It is not difficult to come up with ways to modify the present system.

1. The role of the reviewer in acceptance or rejection could be modified. Almost every expert reviewer has a conflict of interest; he or she represents their discipline as it now exists and unconsciously tends to defend it. The input of the expert should be largely technical. Most journals leave the final decision of publication to their editorial board, but are heavily influenced by the expert opinion on the suitability of publication. This influence is expressed by providing two sheets to the reviewer. One sheet contains the information to be transmitted to the author. The other expresses the opinion of the reviewer as to suitability for publication and priority of publication. This second sheet should be eliminated. This action would tend to downgrade the influence of reviewers in determining the outcome of the review. Such a downgrading is appropriate in our opinion. Parenthetically, one journal (Lancet) has traditionally used this approach. It is a superb journal.

2. Journals should reserve space for particular articles which are of great interest, but which receive poor reviews. At present, this function is subserved by letters, opinion articles or by editorials, but these seldom have the impact of an original paper. To "protect" the readers, publication of these could be accompanied by the reviews (if desired, unsigned) opposing the conclusions of the paper, or by an editorial.

3. Expedited publication of papers with important patient outcome implications might be useful, even at the expense of a truncated review process. Journals do have a social and medical responsibility for patient outcome and this would be one way of expressing it.

4. Journals might have a stamp which conveys the following information for articles which involve management modalities:

The material in this article has not been subjected to an appropriate prospective, randomized, controlled clinical trial and, thus, its value in patient management has not been rigorously established.

Aside from its accuracy (most published management modalities have not been tested in this way), such a statement would provide an important educational experience for the readership. This proposal would not slow the publication of important clinical observations. It would provide an important caveat.

There are at least ten papers on Swan-Ganz catheters which give lip service to the necessity for an appropriate clinical trial and then provide extensive descriptions of the virtues of the catheter without providing these data. This plays an important role in perpetuating the cult.

5. Editorial boards or their equivalents might develop retrospective or ongoing analyses of the quality, relevance and outcome of papers which are published. This would provide data which are sorely needed.

Even if none of these ideas is feasible or useful, then other alternatives should be considered.

In this respect, the American Medical Association is to be commended for its sponsorship of an upcoming conference on peer review which will undoubtedly take up many of the issues which have been discussed in this column.

What Individual Readers Could Do

No structural or organizational alterations can provide absolute protection against the risks of peer review. The most important element for improving the results of peer review involves the individual readers of journals.

Medical students should be trained to be more critical and skeptical of the printed word. Physicians who read journals should learn to recognize the limitations of individual papers, as well as the limitations of the scientific method as applied to the discipline of medicine.

A tough minded, critical readership would minimize the risks of peer review, especially as these risks impinge upon their patients.

Apology

We have deliberately refrained from emphasizing the benefits of the review process. The voluntary participation of reviewers is laudable. In the absence of acceptable data, it is possible that the present system is "the best one that can be devised." Our opinions, being mere opinion, may simply be incorrect. That should not stand in the way of obtaining data and attempting reforms.
ACKNOWLEDGMENT: We thank Dr. Alfred Soffer and Dr. Drummond Rennie for bringing some of these issues to our attention and providing insight.

REFERENCES
3 Brown GW. What can we believe? (editorial). Arch Dermatol 1985; 121:1391-93
10 Badgley RF. An assessment of research methods reported in 103 scientific articles from two Canadian medical journals. Can Med Assoc J 1961; 85:246-50
15 Robin ED. Saltem plus boni quam mali efficere consaire: At least try to do more good than harm. Pharras (in press)

Proponent View

Alfred Soffer, M.D., F.C.C.P.

Doctors Robin and Burke cite six potential benefits inherent in current peer review procedures used by medical journals. Five of these benefits relate to the judgmental or "watch dog" characteristics of peer review. Only one considers the function of the referee as teacher. I am disturbed by the philosophy confined to me by the editor of a distinguished medical journal:

The comments of our editorial board members and out-of-office reviewers are for our use primarily and only secondarily for the author's consideration. Sometimes we return detailed comments, but often we choose to send only a letter of rejection.

Such an attitude places far too much emphasis upon the reviewer's role in recommending acceptance or rejection and far too little on his or her contributions as constructive critic. The submission of a manuscript can be an invaluable educational experience only if the referee is encouraged to be a teacher as well as arbiter. An ancient Islamic text contains this statement, “Blade sharpens blade and scholar sharpens scholar.” When peer review is used merely as justification for rejecting manuscripts, then the “blade” is merely a weapon in authoritarian hands.

Under potential risks of peer review, Drs. Robin and Burke list “an unwarranted stamp of approval for work which is wrong or shoddy.” Mistakes will occur in any system of critical review, but will elimination of peer review prevent publication of slipshod research? A colleague who served, prior to 1951, as the editor of one of the largest medical society journals once told me, “I never used referees. I knew enough about most of the subjects in clinical medicine to make up my own mind about the value of submitted manuscripts.” It may have been true then, but we can never return to those days. The current referee system may fail us from time to time, but it represents today’s single greatest protection of scientific integrity and excellence.

Drs. Robin and Burke are concerned about needless delay in publication. Delays are critical in daily newspapers or weekly publications, but can this be a realistic concern when a periodical appears monthly?