

Guarding the Guardians:

Research on Peer Review

Chicago, IL

May 10 - 12, 1989

The First
International
Congress
on
Peer Review
in Biomedical
Publication

### Welcome!

Welcome to the First International Congress on Peer Review in Biomedical Publication! I believe that publication is central to science and that the peer review process is central to scientific publication. Your presence here in Chicago suggests that you agree.

When George Lundberg and I, in early 1986, committed ourselves to the idea of holding a conference to stimulate research in editorial peer review, we realized we were taking a risk. We might end up with no papers, and even if the papers eventually came in, we might end up with no audience to hear them presented. We need not have worried. As this book of abstracts makes clear, we have plenty of work to discuss, and plenty of discussants. I want to thank all of you who have contributed abstracts and all who have come to hear them presented.

The Congress could not have taken place without the enthusiastic work of many of my colleagues. Martha Carney, previously Director of the Department of Editorial Services and Administration for the AMA Scientific Information Group, helped with the early planning, and Elaine Williams, who now holds that position, shouldered a heavy administrative responsibility. Michele Bacuros, Administrative Assistant, has been responsible for, and good-humored about, the innumerable physical details of the conference, and Sharon Kremkau and the staff of AMA Meeting Management have provided much-needed support. JAMA editorial assistants Sharon Iverson, Steve Sarang, and Anne Frecka have cheerfully taken on and expertly handled a large, extra responsibility. We have repeatedly benefited from the expert advice and assistance of Roxanne Young, JAMA Associate Editor. I am also grateful to Jack Baker, AMA Vice President of Publishing. and John Sayban, JAMA Production Supervisor, for the

help they have given with advertising. Arnold Relman, MD, of *The New England Journal of Medicine*, Edward Huth, MD, of the *Annals of Internal Medicine*, Daniel Koshland, PhD, of *Science*, and Stevan Harnad, MA, of *Behavioral and Brain Sciences*, have been generous with advertising space to let the world know about the Congress.

Much of the early planning of the program as well as the detailed day-to-day management was done with the help of Elizabeth Knoll, PhD, now with the University of California Press, but until July 1988, Assistant to the Editor of *JAMA*. Her intellectual and practical assistance, whether given at *JAMA* or from afar as a member of the Advisory Board, has been immense. Annette Flanagin, RN, MA, took over as Assistant to the Editor in September 1988 and has been superb in pushing the process forward and handling every contingency with imagination and flair. Without Elizabeth and Annette, the Congress could not have taken place.

The members of the Advisory Board, particularly John Bailar, Stevan Harnad, Brian Haynes, Stephen Lock, and Pat Woolf have been an enormous resource (and excellent peer reviewers of the abstracts on peer review).

Putting this conference together has been both hard work and a lot of fun. I hope that all of you will enjoy it as much as I.

Drummond Rennie, MD Director, Peer Review Congress Deputy Editor (West), *JAMA* Adjunct Professor of Medicine, Institute for Health Policy Studies, University of California at San Francisco

### **CONGRESS ADVISORY BOARD**

Marcia Angell, MD Executive Editor, The New England Journal of Medicine

John C. Bailar III, MD, PhD Statistical Editor, The New England Journal of Medicine

Renee C. Fox, PhD Annenberg Professor of the Social Sciences, University of Pennsylvania

Vincent A. Fulginiti, MD Editor, American Journal of Diseases of Children

Stevan Harnad, MA Editor, Behavioral and Brain Sciences

R. Brian Haynes, MD, PhD Professor of Clinical Epidemiology and Biostatistics, McMaster University Edward J. Huth, MD Editor, Annals of Internal Medicine

Donald Kennedy, PhD President, Stanford University

Elizabeth Knoll, PhD Science Editor, University of California Press

Daniel E. Koshland, Jr, PhD Editor, Science

Stephen P. Lock, MD, FRCP Editor, British Medical Journal

George D. Lundberg, MD Editor, The Journal of the American Medical Association

Jose F. Patiño, MD Co-Editor, World Journal of Surgery Drummond Rennie, MD (Director) Deputy Editor (West), The Journal of the American Medical Association

M. Roy Schwarz, MD Assistant Executive Vice President, American Medical Association

David W. Sharp, MA Joint Deputy Editor, *The Lancet* 

Thomas P. Stossel, MD American Cancer Society Clinical Research Professor, Harvard Medical School

Kenneth S. Warren, MD Director for Science, Maxwell Communication Corporation

Patricia K. Woolf, PhD The Woodrow Wilson School for Public and International Affairs, Princeton University

Alfred Yankauer, MD, MPH Editor, American Journal of Public Health

### **Table of Contents**

		Page
Registration Information		2
Program		3 - 8
Continuing Medical Education Information		8
A Historical Perspective on Peer Review From the 18th Century	David A. Kronick	9
The Evolution of Editorial Peer Review	John C. Burnham	9
The Present Status of Peer Review	Stephen P. Lock	9
The Communities of Scientists	Elizabeth Knoll	10
Who Are the Journal Reviewers and How Much Do They Review?	Alfred Yankauer	10
The Community of Referees	Jane M. Smith	10
What Do Peer Reviewers Do?	Stephen P. Lock	11
Editors' Use of Editorial Peer Review in Different Categories of Indexed US Medical Journals	Ann C. Weller	11
Statistical Refereeing in the British Medical Journal	Martin J. Gardner	12
Duplicate Publication in Otolaryngology–Head and Neck Surgery	Byron J. Bailey	12
Publication Bias: What Is Its Magnitude and Nature?	Stevan Harnad	13
The Existence of Publication Bias and Risk Factors for Its Occurrence	Kay Dickersin	13
What Can and Should Be Done to Reduce Publication Bias? The Perspective of the Author and Reviewer	Thomas C. Chalmers	14
What Can and Should Be Done to Reduce Publication Bias? The Perspective of the Editor	David W. Sharp	14
Should All Controlled Trials Be Registered at Inception? A Model From Perinatal Medicine	Iain Chalmers	15
Re-Review of Accepted Manuscripts	Joseph M. Garfunkel	15
Comparative Review Before and After Acceptance of a Manuscript	Vincent A. Fulginiti	15
ndividual and Collective Appraisal of Manuscripts	Frans J. Meijman	16
The Effects of Blinding on the Peer Review of Manuscripts	Robert A. McNutt	16
Use and Evaluations of Peer Reviews by Authors	Joseph M. Garfunkel	17
A Cohort Study of Controlled Trials Initially Reported as Abstracts	Iain Chalmers	17
The Political Realities of Peer Review	Peter Budetti	18
Correcting the Literature Following Fraudulent Publication	Paul J. Friedman	18
The Interface Between Research Institutions and Journals	David Korn	19
Biomedical Information, Peer Reviews, and Conflict of Interest as They Influence Public Health	Erdem I. Cantekin	19

### Table of Contents (continued)

		Page
Comparison of Research Quality Guidelines in Academic and Nonacademic Environments	Joel J. Nobel	20
Steps in the Right Direction	Arthur H. Rubenstein	20
The Consequences of Wrong Decisions	Edward J. Huth	21
The Validation of Clinical Trials Data by the FDA: Relevance for Peer Review in Biomedical Publications	Stuart L. Nightingale	21
Regulatory Agency Use of Peer Review	Sheila Jasanoff	22
A Comparison of Peer Review Methods for Grant Applications at the National Institutes of Health	Mary Ann Scheirer	22
The Impact of Scientific Fraud	Eugene Garfield	23
Online Identification of Published Errata Notices	Lois Ann Colaianni	23
The Philosophical Basis of Peer Review	David Horrobin	24
Editorial Practices of National Dental Journals	Beverly A. Entwistle	24
Surgical Editorial Board Membership: Is Peer Review Possible More Than Once?	James T. Evans	25
Reviewer Agreement on Recommendation and Manuscript Attributes	Joan Ferrante	25
Reviewing the Editors: Articles Chosen by the Editors of the International Editions of $JAMA$	Erica Frank	25
Correlates and Consequences of the Major Peer Review Systems Used by US Scientific Journals	Lowell L. Hargens	25
A New Approach to Referees' Assessments of Manuscripts	Lowell L. Hargens	26
Some Virtues and Defects of the Peer Review System: A Study Based on Reviews Received by Two Groups of Dutch PhD Students	Sheila M. McNab	26
Quotation and Reference Accuracy in Surgical Journals	Howard I. Nadjari	27
Does the Quality of Manuscript Preparation Affect Editorial Decisions About Publication?	Leif I. Solberg	27
Borrowing, Generating, and Distributing Credit Through Research Papers: What Is an Optimal Linkage of Scientific Knowledge Claims?	Paul Ulbrich	28
Status of Peer Review of Papers by Biomedical Journals in India	B. L. Verma	28

### Registration Information

The Registration/Information Desk is located in the Great Hall Foyer.

#### Registration/Information Desk Hours

, , , , , , , , , , , , , , , , , , , ,	
Tuesday, May 9	4:00 PM - 9:00 PM
Wednesday, May 10	7:30 AM - 6:00 PM
Thursday, May 11	7:30 AM - 6:00 PM
Friday, May 12	7:30 AM - 5:00 PM

### Wednesday, May 10

All plenary sessions will be held in the Great Hall . Locations for all other activities are indicated next to function.

7:30 - 8:30 AM REGISTRATION AND CONTINENTAL BREAKFAST (GREAT HALL FOYER)

8:30 AM Welcome

James H. Sammons, MD, Executive Vice President, American Medical Association

**8:40 AM** George D. Lundberg, MD, Editor, *JAMA* 

8:50 AM Introduction

Drummond Rennie, MD, Deputy Editor (West), JAMA, and

Adjunct Professor of Medicine, Institute for Health Policy Studies,

University of California-San Francisco

9:00 AM MORNING SESSION

"THE PAST AND PRESENT OF PEER REVIEW"

Moderator: Drummond Rennie, MD

9:00 - 9:20 AM "A Historical Perspective on Peer Review From the

18th Century"

David A. Kronick, PhD, Medical Bibliographer, Briscoe Library,

University of Texas at San Antonio

Discussion

9:30 - 10:00 AM "The Evolution of Editorial Peer Review"

John C. Burnham, PhD, Professor of History and Lecturer in Psychiatry,

Ohio State University

Discussion

10:15 - 10:45 AM BREAK

10:45 - 11:15 AM "The Present Status of Peer Review"

Stephen P. Lock, MD, FRCP, Editor, British Medical Journal

Discussion

11:30 - 11:45 AM "The Communities of Scientists"

Elizabeth Knoll, PhD, Science Editor, University of California Press

Discussion

**12:00 - 1:45 PM LUNCHEON** (GOLD ROOM)

1:45 PM AFTERNOON SESSION

"THE STUDY OF PEER REVIEW I: JOURNAL PRACTICES"

Moderator: John C. Bailar III, MD, PhD, Statistical Editor.

The New England Journal of Medicine

1:50 - 2:05 PM "Who Are the Journal Reviewers and How Much Do They

Review?"

Alfred Yankauer, MD, MPH, Editor, American Journal of Public Health

Discussion

"The Community of Referees" 2:20 - 2:35 PM Jane M. Smith, MA, Assistant Editor, British Medical Journal "What Do Peer Reviewers Do?" 2:35 - 2:50 PM Stephen P. Lock, MD, FRCP, Editor, British Medical Journal Discussion (for both Smith's and Lock's presentations) 3:10 - 3:30 PM BREAK "Editors' Use of Editorial Peer Review in Different Categories 3:30 - 3:45 PM of Indexed US Medical Journals" Ann C. Weller, MA, Deputy Librarian for the Health Sciences, University of Illinois at Chicago Discussion "Statistical Refereeing in the British Medical Journal" 4:00 - 4:15 PM Martin J. Gardner, BSc, PhD, Professor of Medical Statistics, University of Southampton Discussion "Duplicate Publication in Otolaryngology-Head and Neck 4:30 - 4:45 PM Surgery" Byron J. Bailey, MD, Editor, Archives of Otolaryngology-Head and Neck Surgery Discussion

### Thursday, May 11

EVENING RECEPTION FOR ALL PARTICIPANTS (GOLD ROOM)

7:30 - 8:30 AM	CONTINENTAL BREAKFAST (GREAT HALL FOYER)
8:45 AM	MORNING SESSION  "PUBLICATION BIAS"  Moderator: R. Brian Haynes, MD, PhD, Professor of Clinical Epidemiology and Biostatistics, McMaster University
8:45 - 9:00 AM	"Publication Bias: What Is Its Magnitude and Nature?" Stevan Harnad, MA, Editor, Behavioral and Brain Sciences Discussion
9:15 - 9:30 AM	"The Existence of Publication Bias and Risk Factors for Its Occurrence"  Kay Dickersin, PhD, Assistant Professor, University of Maryland School of Medicine

Discussion

6:00 - 7:30 PM

#### 9:45 - 10:00 AM

### "What Can and Should Be Done to Reduce Publication Bias? The Perspective of the Author and Reviewer"

Thomas C. Chalmers, MD, Lecturer, Harvard School of Public Health. and Professor and Dean Emeritus, Mount Sinai School of Medicine, Boston Discussion

10:15 - 10:45 AM

**BREAK** 

#### 10:45 - 11:00 AM

### "What Can and Should Be Done to Reduce Publication Bias? The Perspective of the Editor"

David W. Sharp, MA, Joint Deputy Editor, The Lancet

Discussion

#### 11:15 - 11:30 AM

### "Should All Controlled Trials Be Registered at Inception? A Model From Perinatal Medicine"

Iain Chalmers, MB,BS, FRCOG, National Perinatal Epidemiology Unit, Oxford, England

Discussion

11:45 AM -1:45 PM LUNCHEON (GOLD ROOM)

#### 12:45 - 1:45 PM

### POSTER SESSION PRESENTATIONS (GREAT HALL FOYER)

#### "Editorial Practices of National Dental Journals"

Beverly A. Entwistle, MPH, Associate Professor, Department of Applied Dentistry, University of Colorado School of Dentistry

### "Surgical Editorial Board Membership: Is Peer Review Possible More Than Once?"

James T. Evans, MD, Department of Surgery, Mercer University School of Medicine

### "Reviewer Agreement on Recommendation and Manuscript Attributes"

Joan Ferrante, PhD, Managing Editor, Journal of Health and Social Behavior

### "Reviewing the Editors: Articles Chosen by the Editors of the International Editions of JAMA"

Erica Frank, MD, Resident, Cleveland Clinic

### "Correlates and Consequences of the Major Peer Review Systems Used by US Scientific Journals"

Lowell L. Hargens, PhD, Department of Sociology, University of Illinois

"A New Approach to Referees' Assessments of Manuscripts" Lowell L. Hargens, PhD, Department of Sociology, University of Illinois

### "Some Virtues and Defects of the Peer Review System: A Study Based on Reviews Received by Two Groups of Dutch PhD Students"

Sheila M. McNab, Buys Laboratory, University of Utrecht, The Netherlands

### "Quotation and Reference Accuracy in Surgical Journals"

Howard I. Nadjari, MD, Department of Surgery, Mercer University School of Medicine

# "Does the Quality of Manuscript Preparation Affect Editorial Decisions About Publication?"

Leif I. Solberg, MD, Department of Family Practice and Community Health, University of Minnesota

### "Borrowing, Generating, and Distributing Credit Through Research Papers: What Is an Optimal Linkage of Scientific Knowledge Claims?"

Paul Ulbrich, Sociology Department, Columbia University

## "Status of Peer Review of Papers by Biomedical Journals in India"

B. L. Verma, PhD, M. L. B. Medical College, Jhansi, India

#### 1:45 PM AFTERNOON SESSION

### "THE STUDY OF PEER REVIEW II: TESTING PEER REVIEW"

Moderator: John C. Bailar III, MD, PhD

Statistical Editor, The New England Journal of Medicine

#### 1:50 - 2:05 PM "Re-Review of Accepted Manuscripts"

Joseph M. Garfunkel, MD, Editor, Journal of Pediatrics

## 2:05 - 2:20 PM "Comparative Review Before and After Acceptance of a Manuscript"

Alan Bedrick, MD, Associate Professor of Pediatrics,

Arizona Health Sciences Center

Discussion (for both Garfunkel's and Bedrick's presentations)

#### 2:35 - 2:50 PM "Individual and Collective Appraisal of Manuscripts"

Frans J. Meijman, MD, Editor, Netherlands Journal of Family Medicine

Discussion

#### 3:05 - 3:30 PM BREAK

### 3:30 - 3:45 PM "The Effects of Blinding on the Peer Review of Manuscripts"

Robert A. McNutt, MD, Associate Editor, Journal of General Internal Medicine

Discussion

### 4:00 - 4:15 PM "Use and Evaluations of Peer Reviews by Authors"

Joseph M. Garfunkel, MD, Editor, Journal of Pediatrics

Discussion

### 4:30 - 4:45 PM "A Cohort Study of Controlled Trials Initially Reported as Abstracts"

Iain Chalmers, MB,BS, FRCOG, National Perinatal Epidemiology Unit,

Oxford, England

Discussion

7:30 - 8:30 AM CONTINENTAL BREAKFAST (GREAT HALL FOYER)

8:30 AM MORNING SESSION

" QUALITY ASSURANCE IN BIOMEDICAL

RESEARCH AND PUBLICATION "

Moderator: Patricia K. Woolf, PhD, The Woodrow Wilson School for Public and

International Affairs, Princeton University

8:40 - 8:55 AM "The Political Realities of Peer Review"

Peter Budetti, MD, Congressional Subcommittee on Health and the Environment,

US House of Representatives

Discussion

9:10 - 9:25 AM "Correcting the Literature Following Fraudulent Publication"

Paul J. Friedman, MD, Associate Dean, University of California-San Diego

Discussion

9:40 - 9:55 AM "The Interface Between Research Institutions and Journals"

David Korn, MD, Vice President and Dean, Stanford University Medical School

Discussion

10:10 - 10:25 AM BREAK

10:25 - 10:40 AM "Biomedical Information, Peer Reviews, and Conflict of Interest

as They Influence Public Health"

Erdem I. Cantekin, PhD, Professor of Otolaryngology, University of Pittsburgh

School of Medicine

Discussion

11:00 - 11:15 AM "Comparison of Research Quality Guidelines in Academic and

Nonacademic Environments"

Joel J. Nobel, MD, President, Emergency Care Research Institute,

Plymouth Meeting, Pa

Discussion

11:30 - 11:40 AM "Steps in the Right Direction"

Arthur H. Rubenstein, MD, Chairman of Medicine, University of Chicago

Discussion

11:50 AM - 12:00 Summary

Patricia K. Woolf, PhD

12:00 - 1:45 PM LUNCHEON (GOLD ROOM)

1:45 PM AFTERNOON SESSION

"THE PRACTICAL CONSEQUENCES OF PEER REVIEW"

Moderator: Thomas P. Stossel, MD, American Cancer Society Clinical Research

Professor, Harvard Medical School

1:50 - 2:05 PM "The Consequences of Wrong Decisions"

Edward J. Huth, MD, Editor, Annals of Internal Medicine

Discussion

2:15 - 2:30 PM "The Validation of Clinical Trials Data by the FDA: Relevance

for Peer Review in Biomedical Publications"

Stuart L. Nightingale, MD, Commissioner for Health Affairs,

US Food and Drug Administration

Discussion

2:40 - 2:55 PM "Regulatory Agency Use of Peer Review"

Sheila Jasanoff, PhD, Director, Program on Science, Technology, and Society,

Cornell University

Discussion

3:05 - 3:20 PM "A Comparison of Peer Review Methods for Grant Applications

at the National Institutes of Health"

Mary Ann Scheirer, PhD, Special Expert for Program Evaluation,

National Institutes of Health

Discussion

3:30 - 3:45 PM BREAK

3:45 - 3:55 PM "The Impact of Scientific Fraud"

Eugene Garfield, PhD, President, Institute for Scientific Information, Philadelphia

Discussion

4:05 - 4:15 PM "Online Identification of Published Errata Notices"

Lois Ann Colaianni, MLS, Director for Library Operations,

National Library of Medicine

Discussion

4:25 - 4:35 PM "The Philosophical Basis of Peer Review"

David Horrobin, MD, Research Director, Efamol Research Institute.

Kentville, Nova Scotia

Discussion

4:45 PM Summary

Drummond Rennie, MD

### CONTINUING MEDICAL EDUCATION INFORMATION

The American Medical Association designates this continuing medical education activity for 18 credit hours in Category 1 of the Physician's Recognition Award of the American Medical Association. To obtain CME credit, be sure to sign the attendance form at the registration desk on Friday, May 12.



# The Past and Present of Peer Review

# A Historical Perspective on Peer Review From the 18th Century

DAVID A. KRONICK, PhD

The University of Texas Health Science Center at San Antonio, Briscoe Library, 7704 Floyd Curl Dr, San Antonio, TX 78284

Awareness of the proliferation of scientific journals and concern about editorial responsibilities and duties are not new phenomena. In the 18th century we find editors of new journals beginning the prefaces to the first issue, even as new journal editors sometimes do today, with some variant of the expression "what, another new journal?" Despite the differences in the academic, social, intellectual, and economic environments of 18th-century and today's science, we can find similarities that can be interesting and instructive in the way 18th-century editors defined and exercised their prerogatives and responsibilities. These concerns extend to ideas and practices that are closely related to what we call "peer review" or "refereeing" today.

Although the beginnings of peer review are frequently associated with the actions of the Royal Society of London when it took over official responsibility for the *Philosophical Transactions* in 1752, antecedents of peer review practices go back to the 17th century and the first organizations of scientific societies in Italy, France, Germany, and elsewhere. The procedures for maintaining control of the quality of the contents of independent journals were in many cases both similar to and different than those adopted by the societies. A unique feature of the societies was that many offered premiums for "prize essays" and developed procedures for making awards that are closely related to the peer review process.

An examination of the periodicals and society proceedings of the 17th and 18th centuries reveals that although there were significant differences in such aspects as attitudes toward preserving the anonymity of the author, that sensitivity toward such issues as priority and the mechanisms of dealing with them was in many ways similar to the sensitivity we encounter today.

In a period when we are beginning to question the traditional systems of communicating scientific information, it is instructive to examine them in their early stages of development, when perhaps investigators and practitioners had more direct control over the process.

#### The Evolution of Editorial Peer Review

JOHN C. BURNHAM, PhD

Departments of History and Psychiatry, Ohio State University, 230 W 17th Ave, Columbus, OH 43210

Practically no historical accounts of the evolution of peer review exist. Nor are there extensive or standard accounts of the history of medical journalism. This paper is therefore a pioneer contribution.

The practice of editorial peer reviewing became general some time after World War II. Contrary to common assumption, editorial peer review did not grow out of or interact with grant peer review. And editorial peer review procedures did not spread in an orderly way, developed from editorial boards and passed on from journal to journal. Casual referring-out of papers on an individual basis may have occurred at any time, beginning in the early to mid-19th century. Institutionalization of the process, however, took place mostly in the 20th century—either to handle new problems in the numbers of papers submitted or to meet the demands for expert authority and objectivity in an increasingly specialized world.

Biomedical journals appeared in the 19th century as personal organs, following the model of more general journalism. Journal editors viewed themselves primarily as educators, and their major problem was obtaining enough material to fill an issue, not choosing among an abundance of submissions. Only when the number of papers submitted provided the opportunity to choose on the basis of quality did circumstances lead first one journal and then another to adopt expert or technical refereeing procedures.

#### The Present Status of Peer Review

STEPHEN P. LOCK, MD, FRCP

British Medical Journal, BMA House, Tavistock Square, London WC1H 9JR, United Kingdom

Despite the progress made in studying peer review and in improving the process (with checklists for referrees and attention to the statistical aspects, for example) editors need to remember that this has gone unappreciated by outsiders. Or has it? Perhaps what is urgently needed now is a survey of scientists "out there" to tell us whether they believe that peer review is subject to the traditional accusations of cost, delay, and bias or whether they are satisfied that the system works well and is valuable for science.

In any case, I believe editors need to go much more public than they have – what they mean by peer review, for example, which parts of the journal are subject to this and which are not, and their willingness to engage in proper dialogue with authors. My personal belief is that in 30 years' time anonymous referees' reports will have been relegated along with other shibboleths such

as anonymous editorials and book reviews, but a lot more boring debate will have taken place before that occurs.

#### The Communities of Scientists

ELIZABETH KNOLL, PHD

The University of California Press, 10995 LeConte Ave, Los Angeles, CA 90024

The history of editorial peer review appears to be the history of an informal, irregular process of consultation at the editor's discretion between editors with strongly stated interests and goals, and colleagues whom they knew personally and whom they judged to have more knowledge about a particular subject than they did themselves. Nowadays the big biomedical journals, at least, keep computerized reviewer files containing thousands of names of self-declared experts on various subjects, most of whom are not known at all to the editors, and reviews are written in a standardized form. with little discussion between editor and reviewer. In many respects editorial peer review is now more a bureaucratic than a collegial process. The "community of scientists" must also be seen as many communities of specialized professionals seeking personal security and recognition as well as impersonal truth. Nonetheless, we continue to talk in the personal and informal terms of "colleagues," "community" (usually in the singular), and "peer" (always in the sense of "equal").

Of course editors reject the negative side of the close, even clubby world that such terms imply — the suspicion that peer review relies on and reinforces an old boys' network that unfairly keeps out the less powerful members of other communities. As a defense against possible subjective bias in a process on which authors are ever more dependent and against which they feel themselves to have little redress, a new picture of a more formal, systematic, and impersonal peer review process seems to have emerged. The precision and objectivity of science (or a somewhat idealized notion of science) is attributed to the peer review process itself, which is sometimes described as a sort of truth-producing machine.

When the system works well enough, this unrealistically exalted view of the peer review process is reassuring to members of academic promotions committees, authors whose careers and professional self-respect depend on publication in peer-reviewed journals, readers who lack the specialized training to be critical and thus must accept the journals' judgment of an investigation, and perhaps even journal editors themselves. When the system fails, the recrimination is sharp not because science is so severely damaged—science has survived many failed hypotheses and experimental errors—but because the professional costs to the individuals involved is so great.

Perhaps we can make the study of peer review systematic, even scientific. We can certainly make those who use it, control it, participate in it, and depend on it more self-aware and self-critical. Whether we can make the peer review process itself "scientific," in the unphilosophical sense of "error-free," and whether we should even want to do so, are other questions entirely.

# The Study of Peer Review I: Journal Practices

# Who Are the Journal Reviewers and How Much Do They Review?

ALFRED YANKAUER, MD, MPH

Departments of Family and Community Medicine and Pediatrics, University of Massachusetts Medical School, Worcester, MA 01655

To assess the nature and work load of reviewers for the American Journal of Public Health (AJPH), a sample of 264 reviewers was surveyed with virtually a 100% response. In 1987, respondents reviewed papers for 274 other journals, 81% of which were monitored by the Scientific Citation Index (SCI) or the Social Sciences Citation Index. Respondents reviewed most often for JAMA (27%), the American Journal of Epidemiology (26%), and The New England Journal of Medicine (23%). The median number of journals for which they reviewed was 3.6, and median of their estimated review time was 2.7 hours. Their weighted average review time was 2.4hours. The range was wide, and review time was inversely related to number of papers reviewed. At \$30 per hour for the 1,400 AJPH reviews in 1987, this represents over \$100,000 or close to the total cost of producing two issues of the journal. Only 31% of AJPH reviewers were not listed as author of a source publication in the 1987 SCI, and only 16% were not cited. The median number of SCI citations per reviewer was 13. The viability of the scientific establishment depends on the good will of its membership, and its moral sensibility depends on their honesty.

### The Community of Referees

JANE M. SMITH, MA, AND STEPHEN P. LOCK, MD, FRCP

British Medical Journal, BMA House, Tavistock Square, London WC1H 9JR, United Kingdom

We have long had an argument about the prevalence of refereeing among consultants in the United Kingdom. One of us (S.P.L.) thought that most consultants never reviewed papers; the other (J.M.S.) suspected that more might do so. We therefore performed a questionnaire survey to see how many consultants had refereed papers over the past 5 years (and how often).

We sent the questionnaire to a 1 in 48 sample of all UK consultants. It included questions on specialty, institution, years qualified, part of the country, the amount of refereeing, whether consultants wrote articles, and whether consultants were on the editorial board of a journal. If consultants did review papers we also asked how long they had been doing so, how they had been recruited, whether anyone had taught them, and why they did it.

Two hundred eighty-five questionnaires were returned in a usable form—a response rate of 73%. One of us (S.P.L.) was right: only 96 (34%) refereed papers. What was more surprising was the number claiming to write articles: 183 (64%). The other subgroup we identified was member of editorial board: there were 29 of them, nearly all of them both referees and writers. referees differed from the nonreferees in several respects. none of them surprising: most of the referees wrote articles and more of them came from London and university cities, worked in university hospitals, and were academics. On average the referees did not do very much refereeing. The median number of papers refereed in the past 2 years was 4 (interquartile range, 1 to 11). Most of the papers (1,029) were for journals in the referee's own specialty, with only 78 for general journals, 41 for journals in other specialties, and 6 for journals in other sciences. The editors, however, accounted for a disproportionately high number of those papers: the 16 who claimed to have included those papers they had refereed as part of their editorial responsibilities accounted for more than half the papers and had seen a median of 10 papers over the 2 years (interquartile range, 6 to 59). In general the amount of refereeing was linked to the amount of writing and editing. proportion of editors increased with the number of papers refereed, as did the proportion of writers.

#### What Do Peer Reviewers Do?

STEPHEN P. LOCK, MD, FRCP, AND JANE M. SMITH. MA

British Medical Journal, BMA House, Tavistock Square, London WC1H 9JR, United Kingdom

Despite the widespread use of referees, little information is available on individual workload and attitudes. We therefore undertook a prospective study of a 1 in 6 sample of all 1,264 of the British Medical Journal's (BMJ) active referees, asking them to keep records on all papers they refereed from January to September 1988 - whether for the BMJ or any other journal. To examine any difference between specialties we also included all the remaining psychiatrists (n = 65) and pediatricians (n = 67) in our files to see whether they behaved differently from each other. We hypothesized that pediatricians might review more papers for general journals than might psychiatrists. The records included questions on time taken, the type of journal, whether instructions from the editor were clear, whether the paper was in the referee's area of interest, and whether it

posed any conflict of interest. At the end of the 9 months we sent the referees a general questionnaire about their work place, position, and age and their attitudes toward refereeing.

From a total of 343 referees we received forms and/ or records from 301 referees. Two hundred four of these had refereed papers during the first 9 months of 1988. The referees in the three samples—the main sample and the pediatricians and psychiatrists—were predominantly from university cities, and almost half of them were academics. Among them they had reviewed 1,980 papers—a median of 4.0 among 1,248 papers for the main sample, 2.0 among 271 papers for the pediatricians, and 5.0 among 461 papers for the psychiatrists. Well over half the papers reviewed were for journals in the referees' own specialties—the work they did for the BMJ represented 13% of the papers. Our hypothesis that pediatricians might do more work for general journals than do psychiatrists was not borne out by the figures. In fact psychiatrists did significantly more refereeing for general journals (27%) than pediatricians (14%). The psychiatrists and the main sample took longer (1.44 hours), and among the main sample the average time taken was significantly longer for general journals (1.56 hours) than for specialty journals (1.29 hours). About half of the referees (91 main sample, 26 pediatricians, and 29 psychiatrists) performed some sort of editorial role, most of them for journals in their own specialties. For the vast majority of the papers refereeing seemed to have been a straightforward affair. in that few papers had not arrived without clear instructions, were not in the referee's area of interest, or had posed a conflict of interest.

# Editors' Use of Editorial Peer Review in Different Categories of Indexed US Medical Journals

ANN C. WELLER, MA

Library of the Health Sciences, University of Illinois at Chicago, 1750 W Polk St, Chicago, IL 60612

Most of the literature on editorial peer review assumes that there is a uniform system of editorial peer review. Most of the studies of editorial peer review are from editors of the well-known journals reporting on journals they edit. The objective of this study was to identify any characteristics of the editorial peer review process that differentiate two distinct categories of indexed US medical journals. The hypothesis states that these different categories of journals use different practices of editorial peer review.

All journals were indexed in *Index Medicus*. Journals in group 1 appeared on each of 3 lists of recommended journals, had a circulation of at least 10,000, and were cited at least 5,000 times per year. Sixteen journals met the criteria for group 1. Journals in group 2 were indexed but met none of the other criteria of group 1. Approximately half of the journals in *Index Medicus* qualify as

group 2 journals. When the journals are divided by broad subject areas, all group 1 journals are general medical journals or cover a major specialty of medicine, while 85% of group 2 are either a medical subspecialty, related discipline, or interdisciplinary journal.

One hundred twenty-four group 2 editors were mailed a questionnaire (69.4% returned). All 16 group 1 editors and a random sample of 16 group 2 editors who returned questionnaires were asked for an interview. Sixteen group 1 editors or managing editors (100%) were interviewed. Fifteen group 2 editors (93.75%) were interviewed. The following statistically significant (P < .05) (two-tailed t test) differences were found: Group 1 editors worked more hours per week (20.6 vs 12.6), received more manuscripts per year (1,617 vs 162), had a higher rejection rate (67.2% vs 48.2%), had more manuscripts with statistics (72.7% vs 54.1%), reviewed a smaller percentage of manuscripts (82.6% vs 93.1%), and solicited more editorials (86.9% vs 38.9%). Group 1 editors were more likely than group 2 editors to use associate editors to select reviewers (87.5% vs 62%), locate reviewers in a reviewer file (72.7% vs 40.7%), have a reviewer file that is searchable by subject (87.5% vs 53.8%) and computerized (81.3% vs 35.5%), use blind review (100% vs 32.9%), have a manuscript revised (88.7% vs 75.8%), share reviewers' reports with all reviewers (81.3% vs 35.7%), have a manuscript rereviewed after a complaint about rejection (62.5% vs 14.3%), and reverse a rejection decision (19.7% vs 7.6%). Group 2 was more likely to use editorial boards (55.4% vs 16.3%). Group 1 editors were less likely than group 2 to use the same review process for solicited manuscripts (40% vs 73.8%), give thorough in-house review (6.3% vs 32.4%), leave a signature on a signed reviewer's report (35.7% vs 85.9%), send revised manuscripts to the same reviewer (18.8% vs 63.1%), and decide the outcome when reviewers disagree (25% vs 62.7%).

Results are discussed in terms of what was learned during the interviews. Results indicate that there is not a uniform system of editorial peer review practiced by all editors. Group 1 and 2 editors use editorial peer review differently. Some differences can be attributed to journal size, but some represent a different approach to editorial peer review.

### Statistical Refereeing in the British Medical Journal

<sup>1</sup>MARTIN J. GARDNER, BSc, PhD, <sup>1</sup> AND JANE BOND, BSc, MSc<sup>2</sup>

<sup>1</sup>Medical Research Council Environmental Epidemiology Unit, Southampton General Hospital, Southampton, 509 4XY, United Kingdom, and <sup>2</sup>Department of Statistics, University of Southampton

Statistical refereeing of papers submitted to the *British Medical Journal (BMJ)* has been taking place for at least 10 years. The number of papers assessed has

increased regularly, and particularly in the last few years.

The study was undertaken to answer the question: "Are papers published in the *BMJ* improved by statistical refereeing?" A 6-month period (January to June 1988) was used as the study duration. The question was addressed by: (a) examining each published *BMJ* article that had been subjected to statistical refereeing, (b) against the originally submitted manuscript, and (c) in relation to the statistical referee's comments.

In particular, changes suggested in (c) were examined for any revisions from (b) to (a) to determine whether amendments were made appropriately.

Thirty general papers (January - March) and 25 clinical trials (January - June) were identified as eligible. For those where a detailed checklist\* was completed at submission, 3 of 25 general papers and 2 of 20 clinical trials papers only were considered of acceptable statistical standard for publication. Assessment of the same papers when published, by a statistical reviewer unconnected with the refereeing system, suggested that 21 of 25 and 17 of 20, respectively, were of acceptable statistical standard.

Information on detailed aspects indicated that improvements had taken place for the published versions in terms of both description of the design and conduct aspects of studies as well as the analysis and presentation. It is concluded that statistical refereeing largely improved the related contents of published papers. Ways of further improvement will be discussed.

\*See: Guidelines for writing papers.  $Br\ Med\ J$ . 1988;296:48-50.

# Duplicate Publication in Otolaryngology–Head and Neck Surgery

BYRON J. BAILEY, MD

Archives of Otolaryngology–Head and Neck Surgery, University of Texas Medical Branch, Substation #1, PO Box 103, Galveston, TX 77550

Editors, reviewers, and readers share a common desire for the publication of material that is accurate. timely, useful, and original. This study is designed to analyze the role of the peer review process as it relates to the issues of duplicate publication in the biomedical journals of the specialty of otolaryngology-head and neck surgery. We have analyzed the publications of every author and coauthor who published an original article in the Archives of Otolaryngology-Head and Neck Surgery between January 1980 through December 1987. During this 8-year period, 1,965 authors and coauthors published 1,060 original articles in our journal. A search of several medical literature databases permitted us to collect the titles of all publications by each author and coauthor for the period from January 1977 through December 1988.

The list of each author's publications was reviewed for titles that suggested duplicate publication, and each

of these sets of articles was flagged for comparison. This initial report will focus on the results of our analysis of the first 1,000 authors, selected at random. Of these 1,000 authors, 228 individuals appear to have engaged in the publication of as many as 938 duplicate articles.

Articles are analyzed for five levels of duplicate publication as follows: level 1—identical articles, or articles with identical paragraphs, illustrations, or conclusions; level 2—highly similar articles, or those in which the same data, patient series, or experiments are described with only superficial modifications; level 3—arbitrarily segmented articles in which "salami slicing" has produced several publications when one would have been appropriate; level 4—sequential articles in which there are reports concerning a growing series of patients, without changes in fundamental concepts or conclusions; and level 5—interdisciplinary articles in which the same information is modified slightly for presentation to more than one discipline.

The purpose of this study is to gather objective data and report our findings concerning the incidence and types of duplicate publication in a single specialty during a specific interval. This information will be valuable in the further education of authors, peer reviewers, editors, and readers regarding the issues of duplicate publication.

Our initial impressions are

- 1. Duplicate publication is not a trivial issue within our specialty's author population,
- 2. A relatively high percentage of authors engage in duplicate publication of their work,
- 3. Most authors who engage in duplicate publication do so only a few times, and
- 4. The peer review process is only partially effective as a screen to prevent duplicate publication.

The overall significance of duplicate publication is difficult to assess from a cost-benefit perspective and deserves thoughtful discussion and debate.

### **Publication Bias**

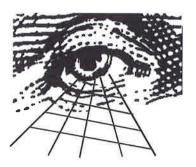
## Publication Bias: What Is Its Magnitude and Nature?

STEVAN HARNAD, MA

Department of Psychology, Princeton University, Princeton, NJ 08544

Quantitative comparisons between anonymous peer review and open peer commentary will be described and discussed. It has been suggested that we define "publication bias" as any factor that leads to a systematic deviation from the truth in publication. According to such a definition, the refusal to publish papers by women might not count as a publication bias (if it did not lead to a systematic deviation from the truth), whereas a zeal factor that multiplied all positive results

by 0.001% would. Obviously, there are different kinds of potential biases in publication, not all connected with ensuring the publication of the whole truth and nothing but the truth. The magnitude and significance of the deviation from the truth (and of the truth itself) are pertinent variables too. Systematicity is certainly not enough. It is quite possible, for example, that two of the standard candidates-author identity and direction of results—bias publication toward rather than away from the truth overall. Even error and fraud do not have face validity as deviators from the truth unless it can be shown how they could be built on systematically, in science's usual convergent, cumulative way, without being found out. Applied science may have more potential bias problems than basic science, but these may be more closely related to application criteria than to publication criteria. Systematic departure from the truth is also more likely to occur, but less likely to matter, in average, humdrum work that is driven by publish-or-perish prerogatives, than in the real growth regions of science, where the self-corrective factors of cumulativity and convergence prevail.



## The Existence of Publication Bias and Risk Factors for Its Occurrence

KAY DICKERSIN, PHD

The University of Maryland School of Medicine, Department of Ophthalmology, 22 S Greene St, Baltimore, MD 21201

Publication bias is the tendency on the parts of investigators, reviewers, and editors to submit or accept for publication study findings depending on the direction or strength of the results. The assumption has been made in the medical field that publication bias exists, although there are few data to support the notion.

Much of what has been learned comes from the social sciences, but less has been done in the field of medicine. In medicine, only three published studies have looked at the problem directly (Simes, 1986; Dickersin et al, 1987; Sommer, 1987). These three studies all provide evidence for publication bias, from the areas of cancer, clinical trials, and menstrual cycle research. Data from one of the studies (Dickersin et al, 1987) imply that unpublished trials may have more "negative" results because they include fewer patients. It also appears

that authors, not editors, are responsible for the decision not to publish "negative" results. Two additional studies have been conducted and the results are being prepared (Chalmers et al,<sup>4</sup> in preparation; Dickersin et al,<sup>5</sup> in preparation). Chalmers and his colleagues did not find that subsequent full publication of results initially published in abstracts was related to either study findings or quality.

Publication bias is important both from the scientific perspective (complete and accurate communication of knowledge) and from the perspective of those performing meta-analyses. If meta-analyses are to be a basis for making treatment decisions, then they must include all available data of an acceptable quality.

#### References

- 1. Simes RJ. J Clin Oncol. 1986;4:1529-1541.
- Dickersin K, Chan S, Chalmers TC, Sacks HS. Controlled Clin Trials. 1987;8:343-353.
- 3. Sommer. Psychol Women Q. 1987;11:233-242,
- 4. Chalmers T, Hetherington J, Tarnow-Mordi W, et al. In preparation.
- 5. Dickersin K, Meinert C, et al. In preparation.

# What Can and Should Be Done to Reduce Publication Bias? The Perspective of the Author and Reviewer

THOMAS C. CHALMERS, MD1; CYNTHIA S. FRANK1; AND DINAH REITMAN, HPS2

<sup>1</sup>The Technology Assessment Group, Harvard School of Public Health, 677 Huntington Ave, Boston, MA 02115, and <sup>2</sup>Clinical Trials Unit, Mount Sinai Medical Center

Publication bias is only one form of research bias in the field of clinical trials. Also of importance are prepublication and postpublication biases. These will be illustrated from personal experiences.

Prepublication biases: Anything that interferes with the truth being ascertained and reported in a meaningful way can be classified as a bias. Among those who perform clinical trials the common causes of bias are ignorance of how to perform an optimal study, sloth (the "quick and dirty study"), and greed (doing what brings in the money). An unbiased research project takes time, effort, and money spent on controlling bias as well as possible. The double standards of peer review and payment applied to clinical trials vs clinical practice exacerbate the problem.

Publication bias is commonly considered to be solely concerned with how "positive" or "negative" a study may be. Equally important are the reviewers who are seldom unbiased and have a profound effect on the papers they review.

Postpublication bias refers to the interpretation, review, and meta-analysis of published clinical trials. Data will be presented to show that whether or not a clinical trial is agreed with is highly correlated with the specialty training and practice of the reviewers. Postpublication bias has become more significant with the advent of meta-analysis, which, if not carried out with

care to avoid and measure bias, adds the imprimatur of quantitation to a potentially biased review.

The impact of these three categories of bias can be minimized (1) by eliminating the double standards applied to clinical research and practice and by designing and executing clinical trials that control bias as well as possible; (2) by registering and publishing all trials, which would overload the system and require the creation of electronic media, by disguising the source of articles sent out for review, and by requiring that all reviewers back their words with their signatures; and (3) by replacing the ordinary review articles with well-controlled meta-analyses and by requiring conflicts of interest to be clearly acknowledged.

# What Can and Should Be Done to Reduce Publication Bias? The Perspective of the Editor

DAVID W. SHARP, MA

The Lancet, 46 Bedford Sq, London WCIB 3SL, United Kingdom

"Publication bias," on the definition supplied for the Congress, has three facets, all of editorial concern—(A) the bias sometimes perceived by disappointed authors; (B) the "biases" (not all unhealthy) that a journal's policies, priorities, and procedures may introduce; and (C) the biases intrinsic in design and interpretation in what is submitted for publication. Preventing type C is one goal of peer review itself, but a study once executed cannot be redesigned. Journals publishing good, but not faultless, investigations can at least insist on discussion of outstanding confounding factors, and a case can be made for depositing hypotheses before studies are done. Biases against investigators or institutions personally or against certain lines of work (type A) may be complained of by disgruntled authors more often than they are valid. However, journals—especially general ones and those with low acceptance rates-will need to adopt policies that an outsider might perceive, mistakenly, as prejudiced (type B).

Not all bias is unhealthy. Bias is a secretive creature, and monitoring the paperwork on submissions to journals will not reveal many examples, and few serious. Journals must monitor their refereeing systems for efficiency, but should they be introducing policing systems as well—"blinding" referees (and, logically, themselves) to authorship, introducing tight codes of practice, and seeking solemn declarations from reviewers, for example-solely for the purpose of reducing publication bias? The editor-author-reviewer triangle will strain at its angles from time to time, and editors have an obvious duty to ensure that the review process is handled courteously and expeditiously and as fairly as possible. Yet absolute perfection is unattainable. A professional journal ought to be able to avoid the worst of bias without a devious bureaucracy more suited to the civil service of Byzantium.

### Should All Controlled Trials Be Registered at Inception? A Model From Perinatal Medicine

IAIN CHALMERS, MB, BS, FRCOG

National Perinatal Epidemiology Unit, Radcliffe Infirmary, Oxford OX2 6HE, England

Investigations in which statistically significant differences between treatment groups have not been observed are less likely than others to be reported in scientific journals. In clinical research, this selective suppression of "negative" results may lead to the adoption of ineffective or hazardous treatments. In an attempt to obtain information about unpublished trials in perinatal medicine, letters were sent to 42,200 obstetricians and pediatricians in 18 countries. As a result, 395 unpublished randomized trials were notified. Only 18 of the trials had been completed more than 2 years before the survey, a period during which at least 2,300 reports of perinatal trials had been published. One hundred twenty-five of the 395 unpublished trials had ceased recruitment within the 2 years prior to the survey, 193 were actively recruiting at the time of the survey, and 59 were about to begin recruitment. It seems unlikely that publication bias will be addressed successfully by attempts to obtain information about unpublished trials retrospectively. However, since the response rate to the request for details about ongoing and planned trials was good in this survey, prospective registration of trials at inception appears to be a feasible approach to reducing publication bias and its adverse consequences. An additional merit of prospective registration of clinical trials is that it should reduce unnecessary duplication (as opposed to necessary replication) in research and promote more effective collaboration. Arrangements are now being made to establish prospective registration of trials in perinatal medicine worldwide.

# The Study of Peer Review II: Testing Peer Review

### Re-Review of Accepted Manuscripts

JOSEPH M. GARFUNKEL, MD; HARVEY J. HAMRICK, MD; EDWARD E. LAWSON, MD; AND MARTIN H. ULSHEN, MD

Department of Pediatrics, University of North Carolina School of Medicine, Chapel Hill, NC 27599

To assess the adequacy of the initial review/editorial process, and whether substantive problems would be identified by further review, 25 papers accepted, after revision, for publication in the *Journal of Pediatrics* were sent for review to referees of similar background and experience as those of the original reviewers. The new

referees were unaware that the paper had been accepted. The hypothesis was that no important deficits would be identified on further review and that all manuscripts would be recommended for publication at relatively high priority. The two second reviews of each paper were then independently evaluated by two of three assistant editors with similar experience, who were asked to identify concerns that would have warranted further revision or rejection.

One manuscript was recommended for rejection by both new reviewers, and seven other manuscripts were recommeded for rejection by one new reviewer, three on the basis that another journal would be more appropriate. The average priority score was 2.5 on a scale of 1 (highest) to 5. Differences of opinion between new reviewers were most commonly identified for methods/study design (14 of 25) and discussion/interpretation of the data (15 of 25).

The only manuscript recommended for rejection by both reviewers was rejected by both editors. Of the seven papers that received one positive and one negative review, both editors would have accepted three, both editors would have rejected one, and the editors disagreed about three. There was no consistent pattern of differences regarding rejection or acceptance among the three editors, but one editor would have required more revisions than the other two. In this study, the widely used system of obtaining opinions from two reviewers, plus editorial review, successfully identified manuscripts that would have been recommended for publication by most reviewers. The editorial decision would have been the same at least 80% of the time. However, the system failed to identify all problems that warranted revision prior to publication.

# Comparative Review Before and After Acceptance of a Manuscript

VINCENT A. FULGINITI, MD<sup>1</sup>; JAMES J. CORRIGAN, MD<sup>2</sup>; HUGH D. ALLEN, MD<sup>3</sup>; ALAN BEDRICK, MD<sup>4</sup>; JOANN M. QUANE<sup>1</sup>; AND LAURA C. MARTIN, MS<sup>1</sup>

<sup>1</sup>American Journal of Diseases of Children, Tulane University School of Medicine, 1430 Tulane Ave, New Orleans, LA 70112; <sup>2</sup>University of Arizona; <sup>3</sup>Ohio State University; and <sup>4</sup>Department of Pediatrics, Arizona Health Sciences Center

We had all accepted manuscripts in 1988 submitted to a second set of reviews after the original manuscript had been reviewed, revised, and accepted. Second reviewers were matched by the editor, to the degree possible, by number, discipline, experience, and prior performance for the journal. Second reviewers were blind to the fact that the article had been accepted, although three second reviews had to be eliminated because the article had been published by the time the referee was asked to review the article. Informed consent was obtained from all authors for this process.

Thirty-five manuscripts were analyzed by October 25, 1988; the study is ongoing. Comparisons were made between the original scores made by the original referees, the scores assigned by at least two editorial board members, the scores assigned by the editor, and the scores assigned by the research referees. Comparisons were also made between the specific comments of the original referees and those of the second set of referees.

Second referees disagreed with the assessment of the original referees in 43% of instances. Since all manuscripts had been accepted, the disagreement was always in the direction of rejection. Reasons given for rejection by the second set of referees were (in descending order of frequency): information not new (28%), insufficient data (10%), questionable statistics (8%), excessive speculation (7%), confused presentation (6%), and miscellaneous (5%).

In 4 of the 35 of the manuscripts (11%), the second set of referees identified problem areas also identified by the original set of referees. In 3 of these 4 manuscripts the problem did not appear to have been satisfactorily addressed. There were 14 manuscripts (40%) that originally were in dispute by the editorial board members, ie, there was no unanimity for acceptance. Of these the second set of referees offered scores of rejection in eight (57%). Among the remaining 21 manuscripts with unanimous editorial acceptance, the second set of referees offered scores of rejection in seven (33%).

Our data reveal substantive disagreement among referees for the same issues in many manucripts; most were qualitative and judgmental, and a few were quantitative and objective. In-depth review by editorial board and editor appears to be essential to arrive at a reasoned opinion as to acceptability, since referees have widely divergent views on the value, substance, and acceptability of original research and observations. The fate of a given manuscript appears heavily dependent on the referee(s) selected and the weight given that opinion by the editorial board and editor.

# Individual and Collective Appraisal of Manuscripts

FRANS J. MEIJMAN, MD

Huisarts en Wetenschap (The Netherlands Journal of Family Medicine), PO Box 14015, 3508 SB Utrecht, The Netherlands

In what ways do individual and collective appraisals of manuscripts differ? Is the final appraisal of a manuscript substantially affected by a collective consensus procedure of an editorial board, or can one rely on the (draft) resolution of an individual member of the board (or referee)? A study was made of discrepancies between appraisals made by an individual member of the editorial board and collective appraisals made by the same board, covering manuscripts submitted to the *Netherlands Journal of Family Medicine* over two periods, July 1985 through January 1987 (147 manuscripts) and

February 1987 through September 1988 (164 manuscripts). Discrepancies arose with regard to 40 manuscripts (27%) in the first period and 46 manuscripts (28%) in the second period. For 22 manuscripts (15%) in the first period and 30 manuscripts (18%) in the second period, there was a crucial discrepancy between unconditional refusal and (conditional) acceptance. Considerable interindividual variation within a period and intraindividual variation between both periods in the degree and type of discrepancy were observed. If we can assume that a decision made by more than one person is "better" than a decision by one person, it may be concluded that the extra time and effort involved in collective selection of manuscripts by a full editorial board can markedly improve the quality of editorial decision-making.

# The Effects of Blinding on the Peer Review of Manuscripts

ROBERT A. MCNUTT, MD; ARTHUR T. EVANS, MD; SUZANNE W. FLETCHER, MD; AND ROBERT H. FLETCHER, MD

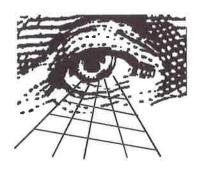
Editorial Office, *Journal of General Internal Medicine*, and Department of Medicine, The University of North Carolina, Chapel Hill, NC 27599

In most biomedical journals, decisions to publish original research depend in part on recommendations of external peer reviewers. But the process by which peer review should be conducted is controversial. Editors have suggested ways to avoid bias in peer review, including blinding of reviewers to authors. However, the effects of this procedure have not been evaluated by means of formal, published research. We are conducting a randomized controlled trial to test the hypothesis that blinding reviewers to authors and their institutions improves the peer review process.

Each manuscript sent to the Journal of General Internal Medicine (JGIM) that reports the results of original research is sent to two external reviewers. For the manuscript sent to one of each pair of reviewers, randomly selected, identification of authors and their institutions is removed; the other manuscript is sent out as received. Reviewers are encouraged (but not required) to sign their reviews. The editors, blinded to the authors, institutions, and reviewers, grade the manuscripts (according to importance of the question, originality, methods, presentation, and appropriateness for JGIM) as well as the quality of the reviews from both the editors' and authors' points of view. Authors also grade the reviews. The process is then unblinded and decisions regarding acceptance of the manuscript are made in the usual way, with all the information available and after discussion with the editorial staff. The study is planned to include 270 reviews (135 manuscripts) to detect a 10% difference in outcomes with  $\alpha$ = .05 and  $\beta = .10$ .

As of October 1988, seventy manuscripts have been entered in the study and 30 reviews are complete. At the rate manuscripts are being received, we expect to achieve the preselected sample size and be able to present the results at the International Congress on Peer Review in Biomedical Publication in May 1989. Blinding requires about 10 minutes and has been successful in 58% of reviews. So far, 45% of reviewers have signed their reviews.

These data will allow us to answer, for *JGIM*, the main research question: does blinding reviewers to authors and their institutions improve the quality of reviewers' recommendations, from either editors' or authors' perspectives, and does it change their recommendations? Other questions these data will be used to answer are (1) which characteristics of manuscripts are most strongly related to acceptance, (2) which aspects of the reviews authors and editors value most, (3) which characteristics of reviewers best predict quality reviews, and (4) how reviewers' decisions to sign reviews affect their reviews.



# Use and Evaluations of Peer Reviews by Authors

JOSEPH M. GARFUNKEL, MD; HARVEY J. HAMRICK, MD; EDWARD E. LAWSON, MD; AND MARTIN H. ULSHEN, MD

Department of Pediatrics, The University of North Carolina School of Medicine, Chapel Hill, NC 27599

A questionnaire survey was conducted among 60 authors whose papers had been rejected by the Journal of Pediatrics to determine (1) whether the authors had used the opinions of the reviewers to modify the paper before submitting it to another journal; (2) the author's evaluation of the quality of the reviews; and (3) the author's evaluation of the editor's communication. As controls for the latter two objectives, the same questionnaire was sent to 30 authors of accepted papers. Over a 6-month period, the questionnaire was mailed to randomly selected authors 2 months after the notification of the editorial decision. Responses were received from 67% of authors whose papers were rejected and 90% of those whose papers were accepted, and the responses of the latter were more complete. Only one of 40 authors indicated no intention to submit the rejected paper elsewhere, and only five authors had sent or planned to send their papers to another journal without revision. Two papers had been accepted by other journals without revision. One of these had already been rejected by still another journal.

For all aspects of reviewer comments and editorial suggestions, authors of rejected papers consistently rated the opinions to be less helpful and constructive than did those of accepted papers. Mean ratings for representative evaluations of reviews were as follows:

Content	Accepted*	Rejected*
Quality of writing	2.84	3.13
Methods or study design	2.41	2.51
Statistical analysis	2.76	3.17
Discussion/		
interpretation of data	1.97	2.05
Evaluation		
Helped strengthen paper	2.07	2.37
Comprehensive	2.03	2.14
Constructive	1.88	2.32

\*Scale: 1 = highest, 4 = lowest.

We conclude that most authors try to use the criticisms of reviewers to improve papers before submitting them to other journals and that authors of rejected papers evaluate the same review/editorial process less positively than authors of papers that are accepted for publication.

# A Cohort Study of Controlled Trials Initially Reported as Abstracts

IAIN CHALMERS,<sup>1</sup> CURTIS L. MEINERT,<sup>2</sup> THOMAS C. CHALMERS,<sup>3</sup> JIMI HETHERINGTON,<sup>1</sup> WILLIAM TARNOW-MORDI,<sup>1</sup> SUZANNE TONASCIA,<sup>2</sup> KAY DICKERSIN,<sup>2</sup> AND MIRIAM E. ADAMS<sup>3</sup>

<sup>1</sup>The National Perinatal Epidemiology Unit, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE England; <sup>2</sup>the School of Hygiene and Public Health, The Johns Hopkins University; and <sup>3</sup>the School of Public Health, Harvard University

The aims of this study were (1) to document (a) the frequency with which controlled trials initially reported as abstracts are subsequently reported in full and (b) the interval between publication of abstracts and publication of full reports; and (2) to test the hypotheses that controlled trials initially reported as abstracts are more likely to be followed by publication in full if (a) they have been judged to be of relatively good methodological quality on the basis of the information provided in the abstract and (b) the investigators have observed differences that they interpret as favoring the newer of the treatments compared.

The Oxford Database of Perinatal Trials was used to identify controlled trials in obstetrics, neonatal pediatrics, and anesthesiology, initially reported as abstracts. The database was also used to identify subsequently

published full reports of these trials. Assessors unaware of the eventual publication status categorized each abstract after assessing the methodological quality of the trial and whether or not clinically and/or statistically significant differences had been reported.

One hundred eighty-two controlled trials initially reported as abstracts were identified. Thirty-six percent of these were subsequently published in full, and 83% of these full reports had been published within the 2 calendar years following the calendar year during which the abstract had been published. Trials subsequently published in full were no more likely than others to have been judged methodologically superior ( $x^2 = 0.16$ ,  $2 \, df$ , P = .92), nor were the initial abstracts more likely to have reported differences favoring the test treatment ( $x^2 = 1.71$ ,  $2 \, df$ , P = .42).

Clinical investigators and journal editors should be concerned by these results for two reasons. First, they suggest that about two of three controlled trials initially reported as abstracts are never reported in sufficient detail to allow a proper consideration of the likely validity of their conclusions. The second reason for concern is that the results provide no reassurance that, among controlled trials initially reported as abstracts, methodological superiority influences the likelihood of subsequent publication in a full report. The results are reassuring, however, in suggesting that abstracts reporting differences favoring a new treatment are no more likely to be followed by full reports than abstracts in which the new treatment is judged to be similar or worse than the standard (control) therapy against which it has been compared.

# Quality Assurance in Biomedical Research and Publication

#### The Political Realities of Peer Review

PETER BUDETTI, MD

Subcommittee on Health and the Environment, Committee on Energy and Commerce, US House of Representatives, Annex I, Room 512, Washington, DC 20515

Responsible stewardship of the public trust requires that publicly funded activities be held accountable to public officials. For the most part, accountability is in the form of direct oversight of the content, budget, and operation of programs by executive agencies and congressional committees. Because of the special expertise involved in science, accountability with respect to scientific activities is in a sense largely delegated by public officials to scientists through peer review. Scientists are permitted to judge the merit and relative value to society of investing public funds in each other's work. As a consequence, peer review must be conducted with total propriety and must itself be accountable to public

entities. This "proxy accountability" is essential to maintaining the public confidence needed for continuing support for scientific research. It is also essential for continuing confidence in peer review as an appropriate guardian of public resources. Similarly, the entire scientific enterprise must be above reproach. This requires that it be viewed as self-policing rather than selfprotective. Should public confidence in peer review be eroded, the alternative is either reduced willingness to fund scientific research or more direct and intrusive methods of accountability. At the extreme, such direct accountability is sometimes associated with terms such as "pork barreling." Peer review must be conducted responsibly not only with respect to funding decisions, but also in the routine oversight of scientific colleagues and in determining what research will be published in the literature. Cases of fraud or other misconduct call into question the adequacy of refereed journals, institutional oversight, and study sections as reliable restraints on the temptations to misconduct, and raise the specter of more direct forms of accountability.

### Correcting the Literature Following Fraudulent Publication

PAUL J. FRIEDMAN, MD

Office of the Dean, University of California, San Diego, School of Medicine, La Jolla, CA 92093

After concluding that Robert Slutsky had submitted three fraudulent papers, the faculty ad hoc committee of the UCSD School of Medicine recommended a review of his entire bibliography to clear the record. This study is a follow-up describing the results of the attempt to inform the medical journals and their readers of our conclusions.

Following a year-long evaluation of the publications, a second faculty committee reported to 30 journals which of their share of the 135 papers were "valid," "questionable," or "fraudulent" and requested publication of the criteria used and the findings. Journals responded slowly to this request, half requiring additional letters over a 2-year period to elicit a reply. Only 3 of 7 journals notified of fraudulent papers responded with full retraction. Of the 13 journals that had only "valid" papers, 5 printed the statement to that effect. Statements concerning 39 of 60 nonvalid papers were published in 13 other journals.

Only seven notices covering 15 papers were found searching under the *Index Medicus* subject heading "Retraction of Publication." Searching Slutsky's bibliography retrieved 18 retracted papers. We have learned that only explicit statements of "retraction" will be properly cross-indexed.

It is concluded that many journals lacked policies and procedures for handling allegations of research fraud, which is analogous to the position in which universities have been until recently. Consultation with attorneys appeared to be more common than with editorial board members. Policies followed in some

cases were not consonant with those subsequently adopted by the International Committee of Medical Journal Editors.

Journals should accept institutionally authorized retractions or corrections when coauthors fail to act and respond more cooperatively to such requests. Statements of retraction or validation should be more consistently identified to facilitate indexing. A consensus definition of "retraction" would be useful to coordinate policies with the National Library of Medicine.

# The Interface Between Research Institutions and Journals

DAVID KORN, MD

Office of the Vice President and Dean, Stanford University School of Medicine, Stanford, CA 94305

In recent years there has been a crescendo of public concern about complex issues of scientific misconduct and fraud. Since much of the attention and many of the examples that have come to public notice have involved real or alleged transgressions in biomedical research, the clamor has been directed principally at biomedical research institutions, especially academic medical centers: the National Institutes of Health, the principal sponsor of such research in the United States today, and the Department of Health and Human Services. Although there is little substantiation, there is a general perception that instances of scientific misconduct are disturbingly prevalent; that reward systems as they have currently evolved in biomedical science and academic medicine are at the very least tolerant, if not encouraging, of sharp practices that can too often and too easily cross the line of flagrant dishonesty; and that academic medical centers are neither willing nor capable of responding to these problems forthrightly. candidly, and effectively. At Stanford University School of Medicine, policies and procedures for investigation of allegations of scientific misconduct have been in place —and used—for more than 6 years. In that time we have handled a modest caseload of challenging problems that in different ways represented violations of professional and scientific norms, and in the process, our policies and procedures have developed and been refined, much in the way of the evolution of a body of case law.

Although these processes have proved generally effective, our experience has revealed a number of substantive procedural difficulties that serve as serious obstacles to the goals of prompt resolution and equitable disposition of cases of scientific transgression. The difficulties lie mainly athwart what I choose to call the "interfaces" between those processes and authorities that fall within the purview of the School of Medicine and those that lie in venues external to the University, including the various private and public sponsoring agencies, the editors of scientific journals, the scientific community, and the general public. At each interface, a common problem arises from the tension between the

university's traditional practice of dealing with allegations of faculty misconduct (and with disciplinary actions more generally) in strictest confidence and considerations of appropriate or required notification and extent of disclosure to each of the several external constituencies who have, or assert the claim of, a "right to know."

With respect to the scientific literature, our investigations of cases of alleged scientific misconduct have occasionally revealed serious lapses from accepted standards of scientific performance and reporting. In such circumstances, we have concluded that correction or retraction of published work is a necessary and, indeed, an essential element in the final disposition of the case. Problems arise either because of insufficient agreement on policies and procedures among scientific publishers and editors on how to deal with these often disagreement matters or because of inadequate resolution and fortitude to act upon them. Of particular concern are instances in which requests for correction or retraction of publications arise from the findings of a formal investigation and are transmitted by the cognizant institution, whether university or research sponsor, with varying degrees of authorial and coauthorial cooperation, let alone enthusiasm. Despite considerable recent attention to this matter by such groups as the International Committee of Medical Journal Editors, the Institute of Medicine, the National Institutes of Health, and the American Association for the Advancement of Science/American Bar Association-sponsored National Conference of Lawyers and Scientists, it is my contention that serious problems remain unresolved. In this presentation, I intend to develop this theme in the time-honored manner of an illustrative case report.

# Biomedical Information, Peer Reviews, and Conflict of Interest as They Influence Public Health

ERDEM I. CANTEKIN, PhD,1 AND TIMOTHY W. MCGUIRE, PhD2

<sup>1</sup>Department of Otolaryngology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, and <sup>2</sup>The Graduate School of Industrial Administration, Carnegie Mellon University

The peer review process essentially controls the flow of information, as well as most other fields of intellectual inquiry. Peer review exerts limitations on both the collection and the dissemination of information through the federal funding decisions of research support and through its role as the ultimate arbitrator in the publication decisions of the scholarly journals. Information, which is gathered and disseminated under the control of peer review, eventually influences the public health as this knowledge is adopted in the practice of medicine. In an ideal world, the peer review process is not biased by the self-interest or conflict of interest of those individuals who participate as they perform their assumed

selfless duty for the advancement of knowledge for the betterment of mankind. However, in the real world this is not necessarily the case.

Two years ago, we submitted a manuscript reporting negative results from a clinical trial in which efficacy of an extensively prescribed medication was evaluated. As of today, the peer review process in leading medical journals has not yet allowed the publication of this research because of certain direct and indirect means of control exerted by individuals who have conflicts of interest with the dissemination of our negative findings. It is our opinion that this particular case demonstrates possible adverse effects of peer review that to some extent may have directly influenced the public health in this country.

In any social system, like the peer review system, self-regulation without built-in safeguards for oversight and accountability has the potential to diverge from the utopian original intent. The example we shall discuss shows that such divergence may be not in the best interest of the public, which the peer review system is intended to serve. We believe that today's built-in safeguards in peer review are not adequate to cope with the complexity of the modern-day biomedical information flow. We shall suggest some possible improvements to the peer review system.

# Comparison of Research Quality Guidelines in Academic and Nonacademic Environments

JOEL J. NOBEL, MD

Emergency Care Research Institute, 5200 Butler Pike, Plymouth Meeting, PA 19462

Our objective, in this ongoing study, is to examine safeguards intended to assure quality research and valid reporting and to learn if controls employed in nonacademic research environments offer useful models. Our inquiry was a broad one. We first determined the prevalence of institutional guidelines dealing with the quality of research undertaken in all 143 accredited medical schools in the United States and Canada by letter query and telephone follow-up. Nonacademic research institutions, because of their diversity, are being addressed selectively. By quality of research, we refer broadly to those those characteristics of a research project such as data capture, traceable calibration of equipment, recording and retention of data, examination of statistical validity, and independent verification of audit trails between articles submitted to journals and original data in laboratory notebooks and comput-

Although this is an ongoing study, work to date has led to the following preliminary impressions: (1) Few schools attempt to control the quality of research beyond legally mandated institutional review boards and animal research regulations. (2) The primary obstacle to acceptance of research guidelines related to quality of

research is the basic concept of academic freedom. (3) The academic research community is highly cloistered, invested in its own values, and has not examined the measures that nonacademic institutions use to assure quality. (4) While publication of fraudulent or shoddy research is typically attributed to causes such as "pressure to publish" or "individual lack of integrity," in many nonacademic institutions guidelines exist to prevent precisely these deficiencies and often succeed in doing so. Such guidelines are worthy of examination by academic institutions. (5) The concept of independent peer review assumes both honesty and lack of error between data acquisition and published articles. The host institution's failure to independently establish an audit trail between data and submitted articles is a serious gap that would not be tolerated in many nonacademic research settings. (6) Academic freedom and the related absence of institutional guidelines for quality of research may well represent one of the few remaining areas of nonaccountability in this country. One paradox is that the academic institution is legally constructed as a corporation. Yet, through the tradition of jealously guarded academic freedom, it often disclaims legal responsibility for the conduct of its employees or the quality of the mutual work product—research. There is no other type of institution in the United States that has successfully refuted that relationship in court and that, indeed, is a very serious prospect for the academic research community to contemplate.

This Congress on Peer Review should stimulate thoughtful introspection and willingness to examine lessons learned in other types of research environments, many of which have been quite successful in preventing error and fraud. Failure of the academic research community to reexamine its fundamental assumptions and to consider alternatives is likely to lead the public and its legislators to conclude that just as "war is too important to be left to the generals," "research is too important to be left to the researchers." Academic freedom is a trust, a compact agreed to by society, and if reform is not undertaken from within, it will, inevitably, be imposed from without.

### Steps in the Right Direction

ARTHUR H. RUBENSTEIN, MD

Department of Medicine, The University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637

Investigations of scientific fraud suggest that factors in the research environment may contribute to the occurrence of scientific misconduct even though they are not the direct causes of these occurrences. Examples include pressures to "publish or perish," an emphasis on competition and secrecy in research performance, and inadequate interaction of young researchers with their peers and mentors. There is concern that not only ethics but also the quality of scientific research in general may suffer in this environment.

These concerns have prompted research institutions,

professional organizations, government agencies, and congressional oversight committees to search for policies that will strengthen the integrity and quality of the research environment. As in the case of public concern over the research use of human and animal subjects, these policy discussions raise fundamental questions about the adequacy and effectiveness of the current self-regulatory system in assuring responsible research practices and preventing scientific misconduct.

In response to these issues, the National Institutes of Health requested the Institute of Medicine to impanel a Committee for the Study on the Responsible Conduct of Research. The Committee has recently published its report entitled "The Responsible Conduct of Research in the Health Sciences." Some of the conclusions of that report will be highlighted in this presentation. In developing recommendations, the Committee sought to define appropriate roles for government, universities, research institutions, professional organizations, and scientific journals that would stimulate local institutional and professional efforts without creating an unjustifiable regulatory burden on the research community. These recommendations represent the steps that the Committee believes are most appropriate for action at this time in seeking to promote integrity in health sciences research.

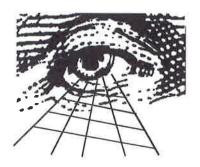
# The Practical Consequences of Peer Review

#### The Consequences of Wrong Decisions

EDWARD J. HUTH, MD

Annals of Internal Medicine, American College of Physicians, Independence Mall West, Sixth St at Race, Philadelphia, PA 19106-1572

Given the acceptable imperfections of peer review, a journal editor often faces the need to decide on what to publish and what not to publish without a clear consensus from external reviewers and editorial associates on the proper decision. This is the editorial version of Harry Truman's "the buck stops here." In this circumstance, the responsible editor must calculate what may result from the wrong decision. There are two main classes of wrong decisions: publishing a paper with invalid and potential damaging conclusions and failing to publish a paper with probably valid and potentially valuable conclusions. The essence of the problem lies in estimating who is likely to be injured by a wrong decision and the consequences of the injury. The parties at risk for injury are authors, the journal, the journal's publisher, physician readers, patients, particular institutions in society (such as pharmaceutical firms), and society at large. The decision will then necessarily hinge on the editor's ethical premises that determine in his or her view an acceptable degree of injury to the parties most at risk. The kinds of editorial analysis that follow from this argument are illustrated by some experiences in the editorial office of the *Annals of Internal Medicine*.



### The Validation of Clinical Drug Trials Data by the FDA: Relevance for Peer Review in Biomedical Publications

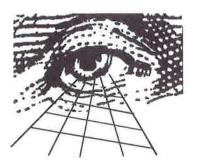
STUART L. NIGHTINGALE, MD

Office of Health Affairs, Food and Drug Administration, 5600 Fishers Ln, Rockville, MD 20857

This paper describes how the US Food and Drug Administration (FDA) reviews the analyses and data submitted to it by sponsors of clinical drug trials. The FDA physicians, pharmacologists, chemists, statisticians, and others review the submissions on clinical trials not only for evidence of safety and efficacy, but also to assess whether the data submitted can be considered reliable and valid. Another procedure by which the FDA assesses validity is the on-site inspection by the FDA staff of clinical trial records, including the raw data collected by clinical investigators. Invalid, "sloppy," or fraudulent data will not be acceptable to support approval of drug marketing applications. Sanctions for violations of the FDA regulations such as the repeated or deliberate submission of false data to the FDA can include the placing of certain restrictions on the investigator in conducting future clinical trials, the disqualification of a clinical investigator from future participation in clinical research on the FDA-regulated investigational products, and, in rare cases, prosecution. Information on violations of the FDA regulations by specific clinical investigators is generally available from the FDA under the Freedom of Information Act, and cumulative lists of clinical investigators who have been disqualified or have consented to specific limitations on their clinical research are available on request.

This paper describes techniques that the FDA uses to identify studies and clinical investigators for intensive review. The paper explores the potential utility to journal editors of various types of information that can be requested from the FDA. Examples of specific situations in which journal editors have benefited or could have benefited from information that was available from the FDA are given. Ethical concerns regarding withholding

information pertinent to the FDA decision-making from the Agency until it is published are explored. The appropriate role of authors, reviewers, and journal editors in this area is discussed.



### Regulatory Agency Use of Peer Review

SHEILA JASANOFF, PhD

Cornell University, Program on Science, Technology, and Society, 632 Clark Hall, Ithaca, NY 14853

This paper is based on a research project looking at peer review as it is used in the regulatory process. In the mid-1980s, peer review came to be seen as a promising way of validating the scientific and technical basis for public decisions relating to health, safety, and the environment. My research, which focused on the way two US federal agencies (the Environmental Protection Agency and the Food and Drug Administration) use peer review, sought to illuminate how a scientific review process can maintain its integrity in the highly contentious and political environment of regulatory decision-making.

The following questions were central to my project: How can regulatory peer review avoid the appearance of "capture" by political interests? Do peer reviewers restrict themselves to "doing science" or do they engage in policy-making under the guise of peer review? How does the public participate in regulatory peer review? What are the consequences of regulatory peer review (eg, "better" science, constraint on agency discretion, reduced conflict, or co-optation of policy)?

The results of my study indicate, first, that peer review in the regulatory setting is a very different process from peer review by scientific and medical journals. Review in the regulatory process is more public, more collective (reviewers act as committee), and more binding on the final decision-maker. Predictably, as well, the issues confronting peer reviewers in the regulatory environment are less scientific and more intertwined with policy than in the context of journal peer review.

How well does regulatory peer review work? First, we have to see the process for what it really is—not as a technique for separating scientific and policy considerations, but as an avenue for developing stable agreements over mixed issues of science policy. Seen in this light, regulatory peer review is most effective when it

permits simultaneous negotiation of differences over facts and values relevant to regulation, and least effective when it approximates the adversarial format of the courtroom. When peer review functions well, it can be an effective forum for the exchange of information between experts and the public (ie, risk communication)

As in journal peer review, there are numerous ways in which regulatory peer review can be biased, so that its conclusions seem politicized. Agencies should be particularly sensitive to (1) the problem of "interlocking directorates," (2) the timing and function of peer review, and (3) the opportunities for public participation.

### A Comparison of Peer Review Methods for Grant Applications at the National Institutes of Health

MARY ANN SCHEIRER, PHD

Building 31/4B25, Office of Science Policy and Legislation, Office of the Director, National Institutes of Health, Bethesda, MD 20892

This presentation describes selected results from a study recently conducted within the National Institutes of Health (NIH), which compared several methods for peer review of biomedical grant applications. Traditionally, most grant applications submitted to the NIH are reviewed by standing committees, which assess scientific merit and assign "priority scores" indicating a numerical rating for a funding recommendation. The study to be reported compared this committee method to review by mail, to examine whether a mail review could be as effective as the face-to-face committee process in making recommendations concerning scientific merit.

The study design involved dual reviews of 150 applications, within 8 review committees and with independent mail reviews by 5 reviewers for each application. Data were collected to compare the priority scores resulting from each method (including variability among reviewers) and to rate the summary statements that consolidate reviewers' comments from each method. Systematic group process data were also recorded at each committee meeting.

Initial findings from the study confirm the high scientific quality of review by both methods, but show a considerable divergence in priority scores awarded. Mail reviewers' scores averaged about 30 points better than committee scores (on the 100 to 500 scale). Mail scores showed a narrower range among applications, but a wider range of reviewers' scores within applications. These differences may reflect the NIH's recent instructions to committee reviewers to spread out their scores across the entire priority score range. Data are presented to relate the differences in priority scores to the observations of committee processes.

The two summary statements for each application were subsequently rated on quality dimensions by

external raters who were blind as to the review method used for each. They tended to rate summary statements from the committees as of higher quality overall. However, for a substantial minority of the applications, the mail summary statement received the higher ratings.

The findings suggest that priority scores fluctuate quite readily with changes in methods. Thus, the scores may not be a precise way of differentiating among applications, in spite of the apparent (numeric) precision of such scores.

#### The Impact of Scientific Fraud

EUGENE GARFIELD, PHD

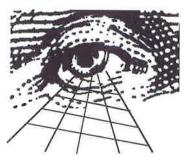
The Institute for Scientific Information, 3501 Market St, Philadelphia, PA 19104

The present study is a citation analysis of selected publications by Stephen E. Breuning, who in 1988 was convicted in federal court of scientific fraud. The goal of the study was to determine the impact of fraudulent research on the scientific literature. By examining this impact, journal editors can better decide how necessary it might be to implement various proposals for preventing fraudulent research from being published or, failing this, for alerting investigation to fraudulent research in the literature.

The study examined 23 publications by Breuning using the *Science* and *Social Sciences Citation Indexes*. It was shown that Breuning's publications received 218 citations from 91 publications during 1980-1988. Of these 218 citations, 83 (38%) were self-citations by Breuning and/or his coauthors. The tracking of these non–self-citations over time revealed a steady and rapid increase from 1981 (2) to a peak in 1985 (41), followed by a sharp decline in 1986 (30) and 1987 (7) that coincides with public disclosure of Breuning's fraud. These data indicate that the scientific literature effectively purges itself of fraudulent research. Authors "shun" work that is publicly exposed as fraudulent.

A citation context analysis was performed to determine how Breuning's research was used by citing authors. Excluding self-citations, a total of 170 citation contexts in 65 papers were examined. Of these citation contexts, 68 (40%) cited Breuning in combination with a series of other publications. Excluding these serial citation contexts, analysis of the 102 unique citation contexts revealed that 43 (42%) were "neutral": citations that simply declared a result, statistic, or statement reported in Breuning's papers without ascribing a value to it. Another 31 (30%) were "positive," indicating agreement or consensus with Breuning's findings. The remaining 28 citation contexts (27%) were "negative," indicating disagreement or variance with results, statistics, or statements reported by Breuning. A final analysis showed that 62 (61%) of the 102 unique citation contexts were "inconsequential." They did not influence the direction, design, discussion, or conclusion of the citing paper. The 40 (39%) that were found to be "material" were contained in 18 (28%) of the 65

papers that cited Breuning. Therefore, the impact of Breuning's research on the literature was meaningful, especially since it, and many of the citing papers, involved drug treatment of mentally retarded people. This particular instance of scientific fraud may warrant the measures currently proposed to detect fraud involving research on human subjects prior to publication or to alert authors to fraudulent research already in the literature.



## Online Identification of Published Errata Notices

LOIS ANN COLAIANNI, MLS; SHELDON KOTZIN, MLS; AND NANCY SELINGER, MLS

The National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

In 1987, the National Library of Medicine (NLM) implemented a procedure that identifies substantive errors to the text, abstract, or descriptive parts of an article introduced during the publication process that were corrected in subsequently published errata notices. The NLM, as the producer of the widely used biomedical database MEDLINE, believes that modern technology provides a way to inform online users when they retrieve a citation to an article for which an error has been noted. This is accomplished by amending the original citation with a reference to a published erratum notice. If the erroneous data are part of the information provided in MEDLINE, these data are corrected and the erratum notice is added to the title. When the error occurs in a portion of the article that is not included in the MEDLINE citation, only the reference to the erratum notice is added.

In 1987 nearly 2,500 substantive errata were noted by the NLM. Included in this number are some significant life-threatening dosage errors. Sometimes these errors were discovered by NLM indexers and were corrected, with the journal editor's approval, in MEDLINE before a published erratum notice appeared.

The NLM will report on the errata data for 1987 and 1988. While the NLM alerts users to the existence of errors, usually this occurs several months after the original citation appeared. Patterns in the intervals between the originally published article and its erratum notice, as well as between original publication and the availability of the updated information in MEDLINE, are also part of the report.

### The Philosophical Basis of Peer Review

DAVID HORROBIN, MD

The Efamol Research Institute, PO Box 818, Kentville, Nova Scotia B4N 4H8, Canada

Peer review can be properly judged only if its philosophical basis is kept in mind. Most people agree that the main purpose of peer review is quality control, to ensure the publication of quality reports of good research. Though important, this is inadequate as a mission statement for medical journals. A physician's duty is "to cure sometimes, to relieve often, to comfort always." The duty of a medical journal—and of the peer review process—is to render all possible assistance to the physician. The physician can best be helped by the advancement of knowledge. Knowledge may progress by the slow accretion of carefully collected information but also by radical and startling innovation. If an editor asks "Would it really matter for medicine if this research were never published?" for the great majority of papers, the answer must be no. The same type of research is proceeding in many centers and given time the same results will be obtained. For this type of paper peer review works well as a quality control measure. I do not think it works so effectively for those very rare papers that are radically innovative. These are the papers that might produce a revolution. If they are correct a major advance in patient care is likely. Even if incorrect the effort to prove them wrong is likely to lead to more rigorous conventional research.

It is by these revolutionary papers that peer review must be judged. It is not adequate to say that the system works well 99% of the time if the 1% of failures are the most innovative papers. Here definitions of the word "peer" become important. Who are the true peers of the

authors of these papers? Certainly not the average researchers to whom the papers are likely to be sent for review. Probably not even other innovative researchers whose work will be challenged by the new study; they have too much vested interest to review the radical concept fairly. It may not be possible to find true peers in terms of knowledge, perspective, and innovative ability. Here the concept of peer review breaks down. The editor must take exceptional care if he or she is not to be pilloried in a Nobel lecture 30 years later, or by a historian 100 years later who asks, "Why oh why was this not published?"

So how do editors handle this situation? Not very well, in my experience. I edit Medical Hypotheses, the only journal in medicine fully devoted to ideas. It has published approximatly 2,000 papers, most from wellqualified authors in mainstream institutions. A recent survey has shown that for many authors Medical Hypotheses was the journal of last resort. Their papers had been repeatedly turned down by the specialist journals that should have published them. Reviewers repeatedly described the work as too risky and too innovative, as a sort of pornography that should not be shown to the children of the research community lest they be corrupted. The problem lay not in the inadequacy of the papers but in the fact that they were almost never judged by true peers. The papers at this conference show that there is little risk of editors forgetting their quality control responsibilities. In contrast I think there is a major risk of editors forgetting their responsibility to patient care and thus the requirement to encourage high innovation. Innovative concepts, like new babies, require tolerance and nurturing. Patients are ill-served by a review system that penalizes innovation and encourages conformity. Editors must take exceptional care to ask the question "Who is this author's peer?" when confronted with those rare papers that may change the face of medical care.

#### **Poster Session Abstracts**

## **Editorial Practices of National Dental Journals**

BEVERLY A. ENTWISTLE, MPH

Department of Applied Dentistry, University of Colorado School of Dentistry, 4200 E Ninth Ave, Denver, CO 80262

The selection, roles, and expectations of editors of national dental journals have not been discussed to any degree in the literature. To examine these variables, written questionnaires were sent to editors of 22 national dental journals in April 1988. Results from the 17 responses (59% response rate) are reported using descriptive statistics. Editorial configurations of the journals vary considerably. Eight common roles of the chief dental editor emerged. Important characteristics for choosing editorial board members include (1) knowledge of a specific content area, (2) reputation, and (3) performance as an ad hoc reviewer. Formal course work

in editing, writing, or related areas is not required for board or editorial positions. Two to four members generally review each manuscript. A decision by the chief editor to reject a manuscript before review usually is based on four reasons: (1) subject not appropriate for the journal, (2) flawed research design, (3) unethical research, and (4) poorly written. Board members basically perform a substantive edit at a macro level for structure and at a micro level for content. Their role is strictly advisory in nature. Most reviewers are not formally oriented to the review process, except through some type of personal communication by the dental editor. Logistics and requirements for choosing dental editors appear to reflect a high degree of inbreeding based on dedication to the journal and demonstrated writing skills. How one gains actual editorial skills and the broad perspective and skills of orchestrating an entire journal are not clear. Discussion will focus on recommendations for improving the skills of peer reviewers and editors of dental journals.

### Surgical Editorial Board Membership: Is Peer Review Possible More Than Once?

JAMES T. EVANS, MD

Department of Surgery, Mercer University School of Medicine, 1550 College St, Macon, GA 31207

The current methods of editorial board selection result in significant overlap of membership. This reflects on the policies of peer review and reviewer selection. No written description of editorial board selection was published in any of seven surgical journals studied. The by-laws of several organizations contained references to editorial membership. The methods of editorial board selection identified either directly or indirectly include (1) membership or officer of organization, (2) current or former associate of chief editor, (3) membership of another editorial board, or (4) past office holder of an organization. Of seven journals studied, there were 260 total editorial board positions available. These were filled by 198 individuals. Of these, 34 individuals served on two or more boards, with three individuals serving on four boards each. An author selecting a journal for submission is reminded that due to significant overlap of editor and editorial board membership, peer review probably only occurs once. If overlap were eliminated, 62 new individuals would be brought into the peer review process and authors would be assured that second submissions would have a better chance of unbiased, fresh peer review. The degree of multiple membership by journal is as follows:

	Total		Organ-
	<u>Positions</u>	Overlap	<u>izations</u>
*Am J Surg	62	15	6
Am Surgeon	34	7	3
*Ann Surg	32	7	4
Arch Surg	34	7	5
Curr Surg	39	7	1
*Surgery	46	13	3
*Surg Gynecol Obstet	22	4	1

<sup>\*</sup>Journal with members on four boards.

## Reviewer Agreement on Recommendation and Manuscript Attributes

JOAN FERRANTE, PHD1, AND EUGENE B. GALLAGHER, PHD2

<sup>1</sup>Department of Sociology, Anthropology, and Philosophy, Northern Kentucky University, Highland Heights, KY 41076, and <sup>2</sup> *Journal of Health and Social Behavior*, Department of Behavioral Sciences, University of Kentucky Medical Center

We compared 310 pairs of reviewer recommendations and written critiques submitted to the *Journal of Health and Social Behavior* (n = 120 manuscripts). We concentrated on two areas: the extent to which reviewers agree (or disagree) on an overall recommendation and the extent to which reviewers agree (or disagree) on 8 manuscript attributes. The attributes include (1)

writing quality, (2) significance, (3) adequacy of literature review, (4) adequacy of sample, (5) conceptual quality, (6) appropriateness of statistical techniques, (7) appropriateness of interpretation, and (8) documentation and presentation. We also compared reviewer characteristics such as gender and length of time to return critiques to see if these affected reviewer agreement.

Our findings lend moderate support for the notion of objectivity in peer review. The reviewers agreed 79% of the time on the overall recommendation. Among the eight attributes, there was highest agreement (84%) about documentation and presentation and lowest agreement about adequacy of literature review (59%).

# Reviewing the Editors: Articles Chosen by the Editors of the International Editions of *JAMA*

ERICA FRANK, MD, MPH, AND GEORGE D. LUNDBERG, MD<sup>2</sup>

<sup>1</sup>Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44106, and <sup>2</sup>American Medical Association

This study reviews the article choices of the editors of the international editions of The Journal of the American Medical Association (JAMA). There are currently 10 JAMA editions (in France, Switzerland [French], Switzerland [German], Italy, China, Japan, Southeast Asia, Yugoslavia, India, and Turkey) published outside the United States, with a total circulation per issue of more than a quarter of a million copies (vs approximately 390,000 copies per weekly issue of the US JAMA). Their release schedule varies by country (biweekly to bimonthly) as do their contents (regional editorial boards are free to select any JAMA articles and to complement them with domestic commentary in selected instances). We examined the variation in lag time to publication (means ranging from 3.8 to 8.5 months), article type selection, concordance of article choices between countries, and effect of recommending an article (compliance frequency ranging between 11% and 64%) for each international edition for a 12-month period in 1987-1988.

#### Correlates and Consequences of the Major Peer Review Systems Used by US Scientific Journals

LOWELL L. HARGENS, PhD

Department of Sociology, University of Illinois, Urbana, IL 61801

Data on cohorts of manuscripts submitted to each of three US scientific journals (the *American Sociological Review*, the *Astrophysical Journal*, and *Physiological Zoology*) show striking differences in acceptance rates, average number of revisions before a paper is accepted, and time lags between first submission and final editorial disposition. These differences are consistent with previous research suggesting that disciplines differ substantially in various journal peer review outcomes and that, for US scientific journals at least, these differences have been very stable over the last 20 years.

To determine the extent to which differences in journals' peer review systems can account for the differences in the outcomes of peer review evaluation, I developed a model encompassing four important characteristics of peer review systems: (1) the number of referees to whom a typical manuscript is initially sent, (2) the proportion of referee recommendations that favor publication, (3) the level of agreement between referees' recommendations, and (4) the degree to which an editor acts on referees' recommendations.

Data on these characteristics for each of the three journals in my study show how journals' peer review systems influence major outcomes of the peer review process. Data on disciplinary differences in typical peer review systems suggest that consensus on research priorities and techniques may play an important role in determining a field's typical journal peer review system, which in turn influences the kinds of peer review outcomes prominent in that field.

## A New Approach to Referees' Assessments of Manuscripts

LOWELL L. HARGENS, PhD,<sup>1</sup> AND JERALD R. HERTING, PhD<sup>2</sup>

<sup>1</sup>Department of Sociology, University of Illinois, Urbana, IL 61801, and <sup>2</sup>Department of Sociology, Stanford University

Studies of referees' assessments of manuscripts submitted to scientific journals assume that a merit or publishability dimension underlies referee assessments. To measure the level of agreement between referees' recommendations, researchers have used coefficients, such as the intraclass correlation coefficient and Cohen's kappa, that require researchers to assign arbitrary scores to recommendation categories (accept, revise and resubmit, reject, etc) or to distances between categories. Using data on referee evaluations of manuscripts submitted to five journals-American Psychologist, American Sociological Review, Law and Society Review, Physiological Zoology, and Personality and Social Psychology Bulletin—we show how an extension of recently developed methods for analyzing crosstabulations with ordered categories (L. A. Goodman, Annals of Statistics, 1985, pp 10-69) allows researchers (1) to test the assumption that a publishability dimension underlies referees' assessments and (2) to derive scale values for the recommendation categories. Our results suggest that a latent publishability dimension underlies referees' assessments for four of the five journals. The results also show that the greatest distance between adjacent recommendation categories is between the lowest and second lowest categories, suggesting that recommendations that a paper be rejected

are more reliable than more favorable recommendations. We show how these results can be used in attempts to measure the level of agreement between referees' assessments for a given scientific journal. Our results also point to analytic difficulties faced by researchers who wish to compare levels of referee agreement for different journals.

### Some Virtues and Defects of the Peer Review System: A Study Based on Reviews Received by Two Groups of Dutch PhD Students

SHEILA M. MCNAB

Buys Ballot Laboratory, University of Utrecht, Princetonplein 5, 3508 TA Utrecht, The Netherlands

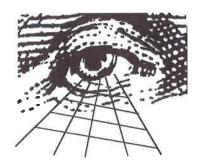
The study is based on reviews of 20 papers submitted to 9 European and 6 US biomedical journals over an 18month period. All the papers were written by Dutch graduate students preparing their PhD theses in the Departments of Medical and Physiological Physics and Pharmacy at Utrecht University. Dutch scientists often seek wider publicity for their work by publishing in English-language journals of standing based abroad. All the reviews considered are for articles that were scrutinized at source by a scientific expert and by a native speaker of English. Analyses of the reviews and discussion with the recipients have highlighted some of the virtues and defects of the peer review system. For the groups under study some shortcomings of the system caused major problems, whereas other defects were turned to advantage, becoming virtues.

Graduate students work under a supervisor and are still in a learning situation. They tend to write most of their papers toward the end of their 4-year contracts. Delays in publication due to an inadequately administered review process can thus be extremely frustrating. Another problem is that articles on the theme of a thesis are often interdependent. A reviewer may find a paper unacceptable because the companion paper is not yet in print. Since some work of the groups concerned is interdisciplinary, it can happen that the reviewer selected lacks the precise specialization required and a third reviewer has to be sought. The delay mechanism inherent in current refereeing practices is a major hazard for PhD students. Sometimes a very critical review can be turned to advantage. Instead of undermining a student's respect for his supervisor, it provokes frank discussion and can even enhance mutual esteem. The student seizes on the weak points and fights back vigorously. Students can benefit from the fact that supervisors, by reviewing papers, keep in touch with recent developments. By "farming out" a reviewing-task a professor may be breaching the confidentiality rule, but reviewing can be a useful educational activity.

Since the journals concerned are based in the United States and in Europe, this was an opportunity to look for differences in the respective review procedures. In the

samples some trends could be detected. The US review tends to be more formalized and detailed, the editor often adding a personal touch. But the excessive time allowed for correction and the resultant publication delay are not geared to the Dutch PhD system. In continental Europe reviewers tend to write less and be more haphazard in their criticism. This can be attributed partly to a lack of guidelines and partly to the language barrier. Reviewers in continental Europe often ask for their own nationals to be cited as references. In inter-European journals the editor is more of a figure-head, but is sometimes more tolerant of the longer article. In summary, the review process is colored by the traditions of the cultural area in which it originates.

Of the 20 articles, 19 were eventually accepted by the journal to which they had been originally submitted. This is a high acceptance rate but not unreasonable in view of the background circumstances. Although in some cases the review system delayed publication, it prevented it in only one case. The groups under study are reasonably satisfied with the review system, but propose the following improvements. Journals should shorten the time between receipt, final acceptance, and publication. Editors should take more care to find reviewers with precisely the right specialization. It would also be helpful if reviewers declined to review papers outside their special field.



## **Quotation and Reference Accuracy in Surgical Journals**

HOWARD I. NADJARI, MD; SHERRY ANN BURCHELL, MD; AND JAMES T. EVANS, MD

Department of Surgery, Mercer University School of Medicine, 1550 College St, Macon, GA 31207

The accuracy of quotations and references in three surgical journals was assessed to help address the question of responsibilities in peer review of papers selected for publication. Fifty references from each of three surgical journals in the same month of issue were selected randomly and evaluated for citation and quotation errors. The three journals serve as official organs for ten organizations. The results are displayed in the following table:

	Citation Errors		Quotatio	Quotation Errors	
	Minor	Major	Minor	Major	
Am J Surg	11	Š	3	17	
Surgery	8	4	0	10	
Surg Gynecol Obstet	22	4	0	10	

The overall error rates were 58%, 32%, and 54%, respectively. The major quotational errors were reevaluated for consensual agreement by at least two authors independently. The major quotational errors identified raise doubt that the reference was read. Responsibility for accurate citations and references has been assigned to authors by deLacey, Record, and Wade¹ and Eichorn and Yankauer.² On the other hand, Key and Roland³ favor assigning the task to editorial boards. We currently agree with Key and Roland, but conclude neither has been dramatically effective.

#### References

- 1. Br Med J. 1985;291:884-886.
- 2. Am J Public Health. 1987;77:1011-1012.
- 3. Arch Phys Med Rehabil. 1977;58:136-137.

# Does the Quality of Manuscript Preparation Affect Editorial Decisions About Publication?

LEIF I. SOLBERG, MD

Department of Family Practice and Community Health, University of Minnesota, Box 381 UMHC, Minneapolis, MN 55455

The editor of a small clinical research journal was concerned about the apparent frequency of preparation problems in manuscripts submitted for publication. Since there are no published data on these problems (format, writing style, content, and data presentation), he conducted a retrospective review of 65 consecutive research articles submitted to (1) provide feedback to potential authors on specific preparation expectations for a research article, (2) identify which areas of manuscript preparation were abused most frequently, and (3) determine whether the quality of preparation affected editorial decisions about publishability. Each article was judged on a specific 1-4 scale for each of 11 preparation areas, for the importance of the research question and the adequacy of the research design.

The preparation areas with the most frequent problems (scored 1 or 2) were (1) conclusion—61%, (2) discussion—54%, (3) methods—51%, (4) writing style—47%, and (5) question identification—43%. IMRAD Format Adherence had the least problems (11% 1 or 2 scores). Although all papers with initial acceptances had high preparation scores ( $\overline{m} > 3.25$ ), there was no clear relationship between preparation scores for those papers initially accorded other responses (revise =  $\overline{m}$  of 2.79, rewrite =  $\overline{m}$  of 2.62, and reject =  $\overline{m}$  of 2.68). Mean judgment about importance of the research question did not correlate with editorial decision. However, there was a clear relationship between research design adequacy and editorial judgment.

Although this small unblinded documentation of one editor's judgments has limited validity for the objectivity of editorial decision making, the specification and frequency of manuscript preparation problems could be used to improve author understanding of how to improve the clarity of submitted manuscripts.

### Borrowing, Generating, and Distributing Credit Through Research Papers: What Is an Optimal Linkage of Scientific Knowledge Claims?

#### PAUL ULBRICH

Sociology Department, Columbia University, New York, NY 10027

The extension of certified knowledge through peer review according to peer-determined criteria distinguishes science from other institutions. The distinction is becoming difficult to maintain as corporate interests and their representatives are increasingly welcomed into circles of scientific decision making. More precise indicators of a research project's scientific, rather than financial, success will be called for. The count of citations to publications resulting from the project is a commonly used though weak indicator. Several refinements in citation practices will be necessary before the citation count can measure the usefulness of research results to fellow scientists with as much validity as profit-from-patents, for example, measures the usefulness of research results to corporate investors.

Several suggested refinements include (1) asking multiple authors of a paper to assign the percentage of credit from citations due to each author; (2) weighting each citation received with the overall citation count of the citer; (3) allowing authors to distinguish citations given that represent a theoretical debt from those that represent an evaluation and of assigning a numerical value to each. (Evaluation is defined as prediction of citations and could be by all of all. Success or failure in predicting could be monitored and result in a gain or loss of professional credibility. Gate-keeper status would be clearly merited, and peer review members who compete with the author of a superior claim might increase their rank through a high evaluation.); and (4) calculating the constraint imposed by the specific publication on future citations received, eg, the rank of journal (tendency to be cited), number of reader-authors, and the visibility of the paper.

With such refinements implemented, researchers could negotiate with their public or private patrons with greater confidence. Self-interest would be unambiguous, globally acknowledged, and, where possible, standardized and pursued by technology, freeing more of the researcher's attention for problem-solving. Objective criteria would exist for ranking scientific knowledge claims and their authors by credibility. Problems and their peer review evaluators could be selected by their rank and position within the network of scientific knowledge claims. The Institute for Scientific Information in Philadelphia would find its mission expanded, as encapsulated by the term "The Institute for Scientific Credit." Bibliographic software would be redesigned to act as a credit bookkeeper. Technology for recognizing, achieving, and rewarding consensus would keep pace with the technology for understanding Nature.

From another perspective the exchange of professional credibility through peer-reviewed research pub-

lications would resemble an economy based on units of professional credit valid throughout the international system of science. Researchers could be paid according to the scientific usefulness of their work. National governments, acting as the ultimate peer review, could promote policy objectives by advancing professional credit to targeted peer groups at minimal additional cost to the tax payer.

### Status of Peer Review of Papers by Biomedical Journals in India

B. L. VERMA, PhD

Biostatistics Unit, Department of Social and Preventive Medicine, M. L. B. Medical College, Jhansi 284 128, India

Lack of or inadequate peer-review of papers are considered to be two of the major factors contributing to the poor standard of biomedical journals. A study, to look into the prevailing status of peer review of papers by biomedical journals, published from India, was undertaken in 1988. Ninety-five such journals that were currently subscribed by 3 medical libraries of Uttar Pradesh (India) were included for the purpose. A predesigned questionnaire, seeking information on editorial policies of the journals and editors' opinion toward some issues related to the peer review, was mailed to the editors. Those who did not respond at the first instance were reminded again after 3 months. In all, only 29 completed questionnaires were received. This report, however, is based on only 24 completed questionnaires (25.3%) as 5 had to be excluded owing to some deficiencies.

Although all journals had printed "instructions to contributors" in some form, only 4 (16.7%) had comprehensive instructions and only 3 (12.5%) had specific guidelines on statistics. Twenty journals (83.3%) had no policy for keeping a biostatistician on the editorial board and 6 (25.0%) were not using peer review at all. Nine journals (37.5%) did not have any criteria on what constitutes a paper, whereas 15 (62.5%) were without any rule to investigate whether the claimed work was actually done. Many journals (9, 25%) had no instruments to check whether the main study has been "sliced" into many pieces. Surprisingly, there were 11 journals (45.8%) without having any policy of statistical scrutiny of submitted papers. In response to our questions on editors' opinions about peer review practices, 6 editors (25%) did not agree for double-blind reviewing of biomedical papers. To check manipulations in data, the majority (15, 62.5%), however, agreed that raw data should be made available by the authors, if required by the journals. Some editors (4, 16.7%) totally disfavored peer review on the grounds of cost as well as man-hours involved vis-à-vis the benefits it produces in terms of quality of papers.



American Medical Association, 535 N Dearborn St, Chicago, IL 60610