in December. "We tend to overreact in this country, and try to turn off the spigot as soon as there's a problem." And even those who agree that there is an oversupply may doubt that the APS can—or should—do anything about it. "That there is a problem isn't challenged by anybody," says Donald Langenberg, APS president and chancellor of the University of Maryland system. "But that there is something the APS can do to solve the problem isn't clear."

That's where the election results could strike some sparks. YSN members have been demanding action, and the group's new voices in the APS hierarchy may try to push the idea of shrinking the physics graduate student pool further than most of the rest of the APS leadership is willing to go. For example, many YSN members have loudly complained about the number of physics Ph.D.s going to foreign citizens, which has grown from 200 to 600 over the past decade, while the number of physics Ph.D.s awarded to U.S. citizens and foreign students holding permanent visas has held steady at about 800. The YSN recently lobbied the Department of Labor to drop proposed plans to relax the restrictions against hiring foreign Ph.D.s. "Granting Ph.D.s to foreign students is fine," says Aylesworth. "But most of the students who come here have promised to return home after their degrees, and it doesn't always happen that way."

YSN members have also advocated supplying prospective graduate students with information packets detailing the gloomy job prospects, and pressuring funding agencies to divert resources away from graduate students and towards postdoctoral researchers looking for a job. "It doesn't cost much to train a physicist," says Levine. "Supporting a working physicist takes a lot more money."

But many other physicists are leery of the idea of squeezing funds, or of any other attempts to actively discourage students or force universities to reduce the number of physics graduate students. "It doesn't make sense for graduate departments to constrict the number of students," says Fermilab's Leon Lederman. "If the student knows the situation, it's caveat emptor." Warns David Balamuth, chairman of physics at the University of Pennsylvania and an APS fellow: "Putting more money into postdocs and less into graduate students could drive things the wrong way, and we might end up with a shortage in 15 years." As for the foreign student issue, even physicists who agree it's a problem shy away from the notion of placing new limits on foreigners' access to degrees or jobs. "I can't seriously imagine that's good public policy, or even morally defensible," says Langenberg.

Aylesworth and Levine tend to dismiss such criticism. "Our ideas are annoying to those people in the power structure whose job it is to produce more Ph.D.s," insists Levine. But he and Aylesworth also say they recognize there may be limits to what they can reasonably expect the APS to do; both are now downplaying the need for action on foreign students, for example, and they say that over the past year, the APS has responded to many of their complaints. "I could be the first elected official who kept his campaign promises before taking office," says Aylesworth.

For now, at least, the YSN firebrands seem prepared to take office at the end of the year in a spirit of compromise. And the APS leadership appears ready to welcome the new faces. "We need a few rabble-rousers in the APS," says Turner.

–David H. Freedman

David H. Freedman is a free-lance science writer in Brookline, Massachusetts.

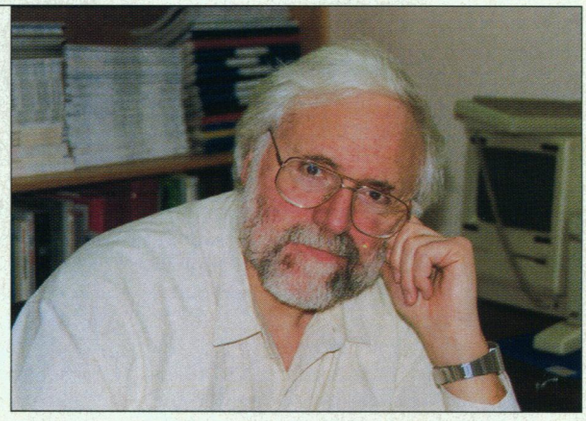
MEETING BRIEFS

Peer Review Goes Under The Microscope

Drummond Rennie, West Coast editor of the *Journal of the American Medical Association* (JAMA), says that his epiphany came in 1985: It struck him that as a scientific editor he was devoting most of his life to the peer-review system, and yet virtually no research had ever been conducted to test it. The result was the First International Congress on Peer Review in Biomedical Publications, organized by Rennie in 1989. Fifty abstracts arrived almost at the last moment, covering issues ranging from how much time reviewers spend on reviews to whether blinding the reviewer to the author of the paper helps to suppress bias. "They seemed to come out of nowhere," says Rennie. His relief faded, however, when he read the abstracts and realized that "a great deal of them were poor at best."

Now, 4 years later, comes the sequel, the Second International Congress, or Peer Review II, held in Chicago from 9 to 11 September. This time, the conference attracted some 270 researchers, most of them journal editors, and 110 scientific abstracts. Rennie says he still would not boast about the standard of the science of peer review, but it was "higher than last time."

Less satisfied was Marcia Angell, executive editor of the *New England Journal of Medicine*, who likened the study of peer review to the study of art, where quality is more easily recognized than quantified. "Most of the things you can measure aren't interesting," Angell suggested, "and most of what's interesting you can't measure." But Angell's dictum isn't absolute, as the following results go to show. Caveat emptor: These studies have not been peer reviewed.



Peer-review reviewer. Drummond Rennie.

Significant Statistics

As any medical researcher knows, one of the best ways to evaluate the efficacy and safety of a particular drug or treatment is a randomized, double-blind trial. The therapy is given to subjects at random, with everyone else in the trial getting another treatment or a placebo, and nobody involved—neither the physician, nor the technician giving the treatment, nor the patient—knowing which treatment is which. "Randomization is the only reliable way of avoiding selection biases," says Iain Chalmers, director of the United Kingdom's Cochrane Centre, which is part of an international collaboration that studies health care. "If it is not done properly, biases won't be avoided" and the studies will become meaningless.

You would expect journals to be especially vigilant in ensuring that they publish clinical results only from properly randomized, double-blind trials. But, as Chalmers and several other speakers told the meeting, that's not always the case.

Chalmers and his colleagues analyzed 206 studies of drugs and procedures published in four leading obstetrics and gynecology journals in 1990 and 1991, looking at the adequacy of the randomization and double-blind procedures. The results were sobering, said Chalmers.

Only a third of the studies reported hav-

ing generated random numbers by accepted methods such as lists of random numbers or computer algorithms; for one journal, the figure was a lowly 15%. And only 25% of the research groups used adequate methods to ensure that the identity of the treatments, subjects, and controls were concealed from all parties. For the same low-ranking journal, the figure was an extraordinary 5%. Unless the authors were misreporting their own procedures, says Chalmers, the trials themselves may be suspect.

Journal editors and reviewers should catch the kinds of lapses Chalmers' study identified, but the two reports that followed suggested they may not be equipped to do so. For one thing, noted statistician Douglas Altman of the Medical Statistics Laboratory of the Imperial Cancer Research Fund in London, the challenge of understanding and evaluating statistical methods keeps getting tougher. His analysis of the statistics papers cited in clinical studies showed that medical researchers are rapidly adopting new, computer-intensive techniques for analyzing data. "The changes," said Altman, "have serious implications for statistical refereeing." Increasingly, suggested Altman, journals and reviewers may be unable to assess whether these high-tech statistical methods are applied appropriately and correctly.

That's partly the fault of the journals, said Stephen George, a biostatistician at the Duke University Medical Center. George reported that only half of the 50 biomedical journals he and his collaborators surveyed this year had a biostatistician on board to review manuscripts that rely heavily on statistical techniques. No wonder, he remarked, that "a substantial proportion of papers published in the medical journals contain obvious flaws in statistical analysis, procedures, and execution."

But Stan Glantz, a professor of medicine at the University of California, San Francisco, and a statistical consultant for the *Journal of the American College of Cardiology*, thinks that catching many statistical lapses shouldn't require a biostatistician because most of the lapses are so elementary. "You don't need a Ph.D.," he said. "Just not having slept through the first 2 weeks of your statistics course."

Blowing Smoke With Symposia

Publishing an article in a journal supplement or in the report of a symposium has, in the long run, much of the impact of publishing in a first-rate, peer-reviewed journal. Even though these publications are often sponsored by pharmaceutical or tobacco companies and may not be peer reviewed, they are indexed in MEDLINE and Index Medicus, used in the large statistical surveys called meta-analyses, housed in medical libraries, and entered into the debate on issues of public policy. The catch, said researchers such as Lisa Bero of the University of California, San Francisco, and Paula Rochon of Ontario's Baycrest Centre for Geriatric Care, is that the papers that appear in these venues often fail to measure up to those in traditional journals.

True, says Rochon, some peer-reviewed symposia reports and supplements contain excellent science. But when she and her colleagues from Baycrest and Harvard compared randomized control trials of drug therapies published in supplements between January 1990 and June 1992 with similar studies published in mainstream journals, they found a clear difference in quality. The reviewers in the study—who were blinded to where any individual paper had been published—consistently rated the supplement papers as poorer on a standardized scoring system of methods and results. The study also found that drug trials published in the supplements were less likely to have called on a biostatistician, less likely to acknowledge government or foundation support, and considerably more likely to be sponsored by a pharmaceutical manufacturer. The findings, she says, "raise serious questions about the [scientific] role of supplements."

"These [symposia studies on tobacco] are poorly described, yet cited very frequently in attempts to influence policy."

—Lisa Bero

An extreme case, said Bero, are reports of symposia sponsored by tobacco companies. Working with Rennie, she examined the conclusions and quality of 114 research articles on the highly contentious issue of second-hand tobacco smoke and its health effects. Of the studies published in mainstream journals, 10% found no adverse health effects. In the symposium reports, that figure jumped to 63%.

To the tobacco industry, the disparity shows that mainstream journals are biased against studies that exonerate second-hand smoke. If the symposium and journal papers were of comparable quality, there might be some merit to this argument, said Bero, but she reported that the average quality of the symposia papers was questionable. Fewer than half (43%) of the symposia articles contained a "methods" section, compared to virtually all (94%) of the peer-reviewed articles. "These [symposia studies] are poorly described," Bero said, "yet cited very frequently in attempts to influence policy as if they were peer-reviewed literature."

To Bero the only bias seemed to be one in favor of negative results in the symposia. But Marcia Angell, executive editor of the *New England Journal of Medicine*, argued that Bero's data on quality may not fully explain the scarcity of negative results in the mainstream journals. Researchers have to be careful, said Angell, not to let "political correctness" drive their conclusions. "I hold no brief with the tobacco industry," she said later, "on the other hand, science is science."

Reviewing the Reviewer

One sure way to improve the peer-review system would be to use only high-quality reviewers. But only one study of what makes a good peer reviewer has ever been published—in 1985, by Thomas Stossel, an editor of the *Journal of Clinical Investigation*. And Stossel came to a counter-intuitive conclusion: The higher the academic status of the reviewer, the lower the quality of the review. Now Arthur Evans of the University of North Carolina has evidence that age as well as rank can be an impediment to a reviewer.

He and his colleagues set out to replicate and extend Stossel's study based on the performance of 200 reviewers at the journal Evans helps edit, the *Journal of General Internal Medicine*. The reviewers were asked to judge the quality of original manuscripts, and the editors, in turn, assessed the quality of the reviews, based on such factors as constructiveness and clear analysis of the paper's strengths and weaknesses.

Good academic qualifications were a plus, Evans told his audience at Peer Review II; better reviewers tended to have fellowship training in clinical research methods, as well as a doctorate or an advanced degree in public health, and were from the most prestigious academic institutions. But, as Stossel had found, high rank was a definite handicap; indeed, the lower the academic rank, the better the review. Reviewers with top administrative positions and prolific publishing records often turned in hasty, superficial reviews. In contrast, assistant professors on their way up tended to spend more time on their reviews, and it showed.

When Evans and his colleagues controlled for rank, one other factor remained: age. The best reviewers turned out to be young, under 40. "It's possible reviewers get worse by age because they get busier," Evans said, "or the review quality improves by year of birth, because reviewers' training and education improves." The answer to that question may have to wait 4 more years for Peer Review III—the sequel to the sequel.

–Gary Taubes