

Conference Report

Steaming up windows and refereeing medical papers

RICHARD SMITH

Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?
T S ELIOT *The Rock* 1934

The little research that has been done on refereeing, asking one scientist for an opinion on whether another scientist's paper is worth publishing, suggests that if two medical referees are asked for separate opinions on a paper they agree on whether the paper should be published only slightly more often than could be expected by chance. So when the *BMJ* gathered together about 60 of its most respected referees we could hardly expect a cosy consensus. As it turned out, the temperature of the debate was so hot that after those of us returning to London were dragged still talking from the hall and shoved into Professor Tony Mitchell's minibus we managed with our hot breath to steam up its windows in seconds. That steamed up window was my image of the day, but when next morning I picked my way through the rubble of my notes I found that a few messages could be salvaged. In the tradition of autocratic editors I now give my (unrefereed) version of the day's events.

What are journals for?

Why do people write for medical journals, Professor Mitchell asked in his introduction. Some people write because they have something to say, but we must recognise that many write to advance their careers, beef up their applications for research grants, or simply for self-aggrandisement. Then why do people choose to write something rather than say it at a symposium? How often, too, are the results of a paper known long before they are ever published? All day long speakers kept invoking the mysterious "invisible colleges," which may spread information more quickly and often more clearly than published papers. When a few cardiologists gather together in a bar in Acapulco they may find out all they need to know about the latest research. Some journals, especially those which take a long time to publish a paper, may thus be reduced to "school magazines," telling people what they already know and making sure that all the members of the school are mentioned at least once so that all the parents will continue to buy the magazine.

So why are so many journals produced? Many, Professor Mitchell suggested, are produced for unholy reasons like making money for the publisher and increasing the status of a group or a specialty. Many doctors receive journals not because they choose to subscribe but as part of membership of a professional society or trade union. So if they are not choosing to receive the journals they may not be choosing to read them. This idea that perhaps all of those at the conference were frantically "transmitting" when few were "receiving" worried Professor Mitchell,

and he was also not convinced that there was much evidence that doctors changed their behaviour after reading scientific papers. These larger, more philosophical worries did not seem to concern many members of the audience, however, and we moved on to discuss the process of selecting articles.

Selecting articles for publication

Because the *BMJ* had organised the conference, and with five *BMJ* editors present, the conference unavoidably had a *BMJ* bias, but many of those present worked for other journals in one way or another. Different journals and editors, it emerged, work in very different ways, but what I write inevitably has a *BMJ* bias.

We receive some 5000 papers a year at the *BMJ*, and we publish about 600. We referee about half of our papers, but this means that we reject about half without refereeing them. The *BMJ* editors were not unduly worried by this responsibility because they thought that most of the papers rejected at this stage were so clearly unoriginal, overspecialised, incomprehensible, or scientifically flawed that the decision was not difficult. Some referees were worried, however, that good papers might be rejected at this stage, and they suggested that the *BMJ* ought to do some kind of study of these papers. Indeed, the main theme of the day, and the theme of the Editor's leading article (p 1224), is how little studied and "unscientific" is the process of selecting scientific articles for publication.

Once a paper has been selected for refereeing it is sent off with a crisp note asking the referee to comment on the originality, scientific reliability, clinical importance, and suitability for a general rather than specialist journal. Other journals give their referees longer and more explicit instructions, and Dr Bill Whimster from King's College Hospital, London, who described what it was like to receive one of the *BMJ*'s letters, suggested that it should contain more information. (There was little discussion of how referees were selected, but a couple of editors explained how they might sometimes select a "tough" referee for a paper they were not keen on publishing and a "soft" referee for one they liked.)

Dr Whimster suggested that the *BMJ* letter should be expanded to include: when we want the article back by; why we have chosen the person as a referee; whether we have chosen another referee and, if so, who; advice on the etiquette of refereeing papers; whether the statistics will be looked at by somebody else; how much the authors will be told; whether there will be any feedback; whether the referee is expected to correct the English; and how much we pay. Many of the referees at the conference did not want all this and were quite happy with the cryptic note they now receive. As many of them confirmed, established referees tend to continue to do what they have always done regardless of what the letter says, but the *BMJ* has recently increased the number of its referees considerably, and the question of what to ask referees to do has now become more important. Others agreed with Dr

Whimster that many of his suggestions would be very useful for new referees, and his guidelines also provided a guide to subsequent discussion.

The issue of confidentiality was one of the first to come up. Dr Whimster said that he photocopied all the papers he was asked to referee so that he could scrawl comments on to the photocopy. He also admitted that he kept the photocopy. Some of those at the conference thought that this was dreadful, particularly in this age of "piracy, plagiarism, and forgery." The next issue was whether the referee could discuss the paper with colleagues. The advantages are that he can then get specialist help on more esoteric points and also, by using his junior staff, he can help in training new referees. Again, some were worried that this might be a breach of confidentiality, but the consensus was that if the editor was told the name of the colleague who had helped then that would make consulting a colleague acceptable. Others suggested that if the referee did not think he was expert enough to referee the paper himself, instead of showing a colleague, he should send it back to the editor—perhaps with a suggestion on who might be expert enough.

How much should referees subedit or rewrite was the next question. For the *BMJ* this question is easily answered because we have full-time subeditors and we do not want our referees to bother themselves much with the English. Other journals who have no full-time staff are grateful, however, for help with subediting. But where does subediting end and rewriting begin, asked Professor Mitchell, and if a referee rewrites a paper should he become an author and share in the glory or opprobrium that may come to that paper. Certainly, one function of refereeing is to advise authors on how their paper might be improved.

How should a referee go about his job? What is the main thing that an editor wants from him? Both editors and referees were agreed that the main function of the referee was to say whether a paper was scientifically reliable and true. With this in mind, the first thing that most referees looked at were the methods and results sections. If they did not think that they would be able to repeat the study after reading the methods or if they thought that the methods were flawed or the data did not support the conclusions then there was an end to refereeing. A few paragraphs telling the editor why the paper was flawed that he might or might not pass on to the author would be enough. Important in this context were references to papers in press or personal communications: some referees thought that if these were not to hand and were crucial to understanding and repeating the method then the paper should be rejected. Dr Stephen Lock, Editor of the *BMJ*, pointed out that under the agreement of the Vancouver Group (the International Committee of Medical Journal Editors)¹ authors were expected to supply copies of duplicate material, papers in press, and conference proceedings.

Eventually the conference discussed the difficult question of statistical refereeing. As Douglas Altman, a statistician, wryly observed, it was symbolic of the status given to statistics that the meeting had discussed references and punctuation before statistics. Professor Chris Booth, director of the Clinical Research Centre, thought that all editorial boards should include a statistician. The poor quality of the statistics in medical papers was, he said, a disgrace. Others thought that Professor Booth was getting carried away and forgetting that statistics is no more of an exact science than clinical medicine: statisticians disagree over whether a paper is statistically sound almost as much as doctors disagree over whether the paper is scientifically sound. The consensus was, however, that the standard of statistical refereeing did need to be improved, and that perhaps those referees incapable of commenting on the statistics in a paper should say so to the editor.

Another thing that editors wanted to know from referees was whether a paper was "believable." Sometimes the results seem just too perfect (one referee mentioned refereeing a paper on hypertension in which most of the readings taken with a random-zero sphygmomanometer ended in 0), and a suspicion is raised that the results have been "cooked." This seems to be happening more and more, and astute referees are one

of the few defences an editor has against such fraud. Most were agreed, however, that clever fraud is hard to spot.

The last question that the *BMJ* asks of its referees is whether they think a paper should be published in a general journal (like the *BMJ*) or in a specialist journal. This was the question that many referees found the most difficult. Deciding what belongs in a medical journal and what in a specialist journal is rather like deciding where the Midlands end and the North begins and begs some knowledge of who reads the journals. In the end the meeting seemed to decide that this was a question best left to the editors with the proviso that sometimes a specialist would be able to identify a part of his specialty that was deserving of a general audience.

At such a gentlemanly meeting payment was barely discussed. The *BMJ* pays its referees £5 for an opinion, but most specialist journals do not because they cannot afford it. Many referees do the job for the good of the scientific community, and many would be far happier with "feedback" from the *BMJ* rather than payment. At the moment the only feedback that referees get is to see whether the paper appears or not, but with the forthcoming computerisation this practice will change. To receive some feedback might not only please referees but, some thought, help to train them and raise standards.

The hanging committee

Once we receive papers back from our referees we consider them at our editorial committee (or "hanging committee" as we call it after the Royal Academy's hanging committee) at which we decide which papers we will publish, and in Nottingham we held one of these meetings in front of three observers. At the hanging committee, which usually consists of two or three editors and two outside consultants, we consider only the papers that referees have advised us are scientifically reliable, although we occasionally spot flaws ourselves at this stage. In the last analysis, we, the editors, publish what we like on the perhaps rather arrogant assumption that what we like our readers will like.

Various new points emerged in this session. Firstly, a couple of the papers we considered had in a way been submitted too early. If the authors had waited a little longer and collected more subjects then they could have come up with clearer answers, but we and the observers recognised that people publish not only to get across a message but also to lay claim to a new technique or to lengthen their bibliographies for their next job applications. We accept this and decided to publish both of those papers.

Then what do we do with papers that are resubmitted? Well, we always look at them again and sometimes we rerefered them. At our observed hanging committee we considered two papers that had been rerefered and where the two referees had come to different conclusions. We argued long and hard, but whether our observers thought that the committee was a good way of deciding what to publish was never quite clear. Perhaps, however, there is no other way.

Outcome

Dr Lock asked the last session of the conference, just as he asks in his leading article, whether all the enormous amount of work that refereeing demands is worth it. Do we end up with better papers? Do we throw out the duds? We do not know. Not surprisingly those at the conference (a vested-interest group if ever there was one) believed that refereeing probably did improve papers. But Dr Alex Paton, a postgraduate dean who has had a long association with the *BMJ*, cried out almost in despair that he thought that if we began to take refereeing too seriously doctors would spend so much time refereeing that they would have no time to write any new papers.

We need some proper studies. But are such studies possible? One that was suggested was to look at the bias of the referees by introducing blind refereeing. But can a paper be blinded?

Many thought that it would be impossible—especially in small specialties—to blind a paper so that the referee could not spot where the paper came from. Others thought it would, while Douglas Altman, the statistician, pointed out that there was still a big difference between knowing where a paper came from and being 80% certain that you did.

So amid all this disagreement do we and other editors need to change what we do? I think we do a little. We need to improve the letter we send out to referees, and we need more referees. We ought, too, to provide our referees with better feedback. Finally, we must study in as scientific manner as we can manage

the results of what we do. Who will referee the paper once we have written it or whether we should submit it to the *Lancet*, the *New England Journal of Medicine*, or the *International Journal of Refereeing, Communication Engineering, and Process and Outcome Analysis* are questions that remain unanswered.

Reference

- ¹ International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Br Med J* 1982;284:1725-814.

Letter from . . . Chicago

On Oak Street Beach

GEORGE DUNEA

On a warm summer's day in Chicago the fashionable models soak up the sun on Oak Street Beach while earnest scientists lying in the sand work hard to confirm the recent theories of a group of Wisconsin investigators. Ignoring the slender beauties in their scant bikinis, these scientists ogle their plumper sisters and can hardly be restrained from applying calipers to determine their waist to hip ratios. About 0.7 is said to be normal, and everything below that characterises the lower body obesity type, with excess fat being distributed mainly about the hips and thighs. In this condition, which is apparently acquired by over-eating during adolescence, the fat cells are increased in number but normal in size, which makes weight reduction difficult. Upper body obesity, by contrast, is often caused by overeating during adulthood. It is characterised by a normal number of fat cells that are enlarged, often tend to lose their insulin receptors, and present an eightfold increased risk of developing diabetes. Testosterone levels are often higher and may play a part in causing the fat to accumulate in the upper part of the body. Weight reduction is relatively easier, because the fat cells need to be shrunk rather than reduced in number. About 25% of all obese women belong to one of these two types, the rest being in between.¹

According to the study, an unbelievable 40% of all American women are obese.¹ This is not borne out by studies on Oak Street, but could require looking at a more representative sample on a less fashionable beach. Perhaps the study could be cautiously expanded to include psychological data, because psychologists have long been interested in the morbidly obese and especially in the effect of surgical correction of such obesity. Earlier reports suggested that surgery for morbid obesity almost uniformly brought about marital discord. A recent study,² however, found that on the whole marriages were improved by correcting obesity. In this study, however, the morbidly obese had more

unhappy marriages to begin with, and among this group there was a high postoperative divorce rate, thought to reflect the liberation of previously unhappy spouses who had put up with the marriage because of physical or psychological handicaps.²

For those selecting non-surgical ways of losing weight, life remains difficult, to say the least. "So difficult, such suffering, so horrible" groaned Soviet maestro Mstislav Rostropovich after slimming down 45 pounds (20 kg) in four months to 185 pounds (82 kg). "For breakfast I eat one hard boiled egg and half a grapefruit. I just cry." To prevent such suffering more than 100 manufacturers have recently marketed "starch blocker" pills made from an amylase inhibitor derived from raw red kidney beans and claimed to prevent the digestion of bread, pasta, rice, and potatoes. In Chicago thousands of bottles were sold each week of this pill, whose safety was established in a four-week study of 36 Indiana women. There was concern, however, about the presence in beans of various noxious substances, including lectins, which may cause haemagglutination. After several patients developed digestive disturbances and had to enter hospital, the Food and Drug Administration moved to ban further sales of the product—which is too bad, because one pill was claimed to prevent the absorption of at least 400 calories of starch, corresponding to three cups of spaghetti, three slices of pizza, or 50 French fries.

Another consequence of undue preoccupation with food is bulimia (ox hunger) or the "binge and bust" syndrome. This often occurs in young well-educated and well-to-do women, not overweight, who on an average of 11 times a week indulge in episodes of ravenous eating, consuming at one sitting anywhere from 1000 to 20 000 calories, and then induce vomiting to avoid gaining weight. This is regarded as a form of relieving tensions akin to abusing alcohol or drugs but less immoral, though physically harmful in that it may lead to rupture of the stomach, oesophagitis, and teeth decay from the acid. Some women are seemingly so obsessed by food that they spend their whole time doing little else but working, sleeping, and going through the eat-purge cycle—a form of behaviour unacceptable in our civilisation though perhaps not unknown to the Romans, whose lavish villas were often equipped with a vomitorium.