Medical journals and the shaping of medical knowledge

Jan P Vandenbroucke

I shall begin by asking: “How is medical knowledge being shaped?”—a process that is much more complicated than we previously thought, and the answers may even prove unsettling. I will then discuss what medical journal editors do in the shaping of knowledge.

Fact versus opinion
The shaping of medical knowledge is to me the relative role of fact versus opinion, or of empiricism versus theory.

In the 1830s, medicine still believed in blood letting, cupping, and similar nonsense; worse, the indications for the interventions were guided by the position of the stars and their interplay with Galenic humours. Then came the triumph of numerical observation by Pierre Charles Alexandre Louis (figure 1), the founder of Médecine d’Observation in Paris. He showed that blood letting does not work for acute pneumonia. The members of the Médecine d’Observation movement were gloriously right, and heroes of medicine.

Let me now jump 160 years. One of the more fashionable theories of the 1990s has been about the immunological mechanism of septic shock. Gram-negative septic shock was ascribed to circulating endotoxin produced by the bacteria. Research in animals showed that gram-negative septic shock could be prevented if the blood was cleared of circulating endotoxin; this was done with antibodies—antiendotoxins that were tailor-made by the most modern molecular methods. The first randomised controlled trial was positive, but more than a dozen trials and thousands of patients later, we now know that these antibodies do not work, and may even be detrimental in gram-positive sepsis.

Have we made so little progress? Is the immunological theory about septic shock like that of the stars, and is it only Médecine d’Observation or its descendant, evidence-based medicine, that brings us the necessary empirical correction, so that we can distinguish right from wrong?

Louis revisited
A revisit to the original Médecine d’Observation argument is illuminating. We think that Louis showed that blood letting did not work, but if you read the original description, as Alfredo Morabia, an epidemiologist from Geneva, has done, you will become uncertain. 1 Louis did not compare patients who were or were not bled; the comparison was between those who came to hospital early in the course of their disease, and were bled early, and those who came late and were bled late. More of the patients who were bled early died. This comparison between two non-comparable prognostic groups would be completely unacceptable today. Also, the tables in the original text contain some disturbing arithmetic mistakes.

In retrospect, the celebrated numerical argument was poor. Nevertheless, we still accept as a “fact” that it was proven by the numerical method that blood letting does not work. So, why did people accept the argument, and why do we still accept it?

Could the reason be that we are glad we got rid of blood letting and the theories behind it—that we like to accept the fact because of our opinion?

Side-effect controversies
The easiest way to see the role of opinion in modern medicine is to look at side-effect controversies. The greatest side-effect story of the 20th century concerns lung cancer and cigarette smoking. The discovery that lung cancer was a side-effect of cigarettes was made almost 50 years ago. Today, we remain in the midst of discussions about the regulation of this product.

Earlier this year, a survey was published in the Journal of the American Medical Association about review articles on passive smoking and its ill-effects.2 Many of the reviews were more or less formal meta-analyses. The survey showed that the conclusions depended critically on the source of funding of the authors—that is, on their prior opinion—since certain authors with certain opinions seek certain sources of funding. We have seen a similar interplay between sources of funding and opinion of authors with reviews on the side-effects of calcium-channel blockers, β-agonists, salicylates and Reye’s syndrome, and countless others. This finding, you may say, is only apparent because these are merely reviews, and even if they are more or less formal meta-analyses, there is a large subjective element to them. If we could only get the data, subjectivity would play no part.

Another side-effect controversy concerns the safety of the most modern type of oral contraceptives, which made news headlines and even led to questions in parliament in several countries. Do the most modern third-generation...
contraceptives lead to a higher incidence of deep venous thrombosis than did previous types of contraceptive?

Three major case-control or cohort studies showed a relative risk of 2-0 or more, when venous thrombosis among women using the newer contraceptives was compared with that in women using slightly older brands.3-5 Opposed to these are three studies that have shown a relative risk of 1-5 or less.6 The authors of the latter studies have concluded that there is no increased risk. How are we going to make sense of this bewildering situation? Are we helped by the information in figure 2 that all of the latter studies showing no truly increased risk were directly funded by the contraceptive-pill producing pharmaceutical industry? “Not really”; as one colleague told me (a pharmacologist involved in early-phase studies for the industry), “this might very well mean that industry-sponsored studies are the better ones”.

This, of course, is observational epidemiology, mostly case-control studies, and we all know about the hazards of case-control studies. If only we had properly randomised controlled trials, this controversy would be avoided.

**Randomised controlled trials**

Let us start with randomised trials of homoeopathy before we go on to allopathic medicine. There is a misconception about homoeopathy in that some people still like to give it the benefit of the doubt because its mechanism of action is not understood. At first, we did not understand the mechanism of action of foxglove or of willow bark either, but we accepted these folk medicines into our scientific medicine. And have we completely understood the mechanism of action of digitals or salicylates today? In homoeopathy, the situation is different: there is an infinite dilution that precludes any possibility of chemical activity.

Imagine a randomised trial about homoeopathy: you compare “solvent only” with an “infinite dilution”. That is the same as giving solvent in both arms. Such a trial is a game of chance between two placebos. What should be the outcome? Zero effect, except for some random variation that is always present. This means that if there are several randomised trials on homoeopathy versus placebo, you expect to get a nice funnel plot that is centred around “no effect”, as in figure 3.

On the x-axis you see all possible trial results: in the middle you see “no effect”. On the y-axis you see the size of the trials: smaller study sizes are situated in the lower half, and larger ones in the upper half of the plot. Each dot or number is the result of one or more trials. You expect small trials show quite a bit of scatter about the true “no effect”, whereas the larger the trial gets, the closer it comes to the truth. That is the funnel plot you expect, and, of course, this one is a nice computer simulation.

Reality is different. In 1997, The Lancet published a meta-analysis of randomised trials of homoeopathy, and the authors included a plot (figure 4). The x-axis is a logarithmic scale of the odds ratio. The y-axis gives the inverse of the variance, a measure is equivalent to study size. “No effect” is a log odds ratio of zero. There are a few largish trials, but the bulk of the evidence is nicely centred close to the log odds ratio of unity, corresponding to a relative risk of about 2-0, meaning that homoeopathic dilutions have almost twice the power to cure compared with homoeopathic solvent. You also see a bunch of outliers. Is there publication bias? Perhaps some of the smaller negative trials have not been published. Not so: the authors of this particular meta-analysis calculated that more than 4000 well-conducted average-sized randomised controlled trials would have to have been left deliberately unpublished to explain the results by publication bias only.

You and I do not like to give up yet, of course. Perhaps the quality of the trials was very poor. Again the authors have their response ready: the subset of high-quality trials published in the best journals still showed an effect of homoeopathy, at the highest dilutions.

Let us now make a dangerous thought experiment. Let us imagine that this funnel plot was not about homoeopathy, but about something that we believe in—our favourite disease and our favourite drug which we are certain works. If the funnel plot of all the randomised trials looked exactly like the one in figure 3, what would we say? We would look at the funnel plot and we would admit: “Yes, we know that the literature is always somewhat optimistic, but even if you leave out the outliers, a sizeable effect remains, doesn’t it? And if you look especially at the high-quality trials, the effect is still there. And what of the few largish trials with no effect? These were done by people who do such large but sloppy trials, they did not use the right dosage, or the right timing, and they failed to exclude properly patients with contraindications. And, now that we think of it, we can explain those outliers: this must be a subpopulation of patients with more severe disease, and that is why they are apart. The drug is more effective among them; you can see it. They have their own little funnel plot”. Does this sound familiar?

To some people that scenario is too much like invoking auxiliary hypotheses. Some people state that they will only accept evidence from meta-analyses with symmetrical funnel plots. If you take that position, you will certainly throw away good evidence from several trials, because some trialists have produced rubbish. Remember also that asymmetry is dependent on scale. A funnel plot that is symmetrical on a relative-risk scale might become asymmetrical on a risk-difference scale.

I warned you about this thought experiment. We seem to have reversed everything. We started out by stating that theories—stars in heaven, molecular immunology—were
significantly smaller than the placebo group. Apparently, published trials, the treatment group was repeatedly and advances than are trials without sponsors. Sponsors are more likely to report statistically significant blinded assessment and intention-to-treat analysis. It has been shown that randomised trials with outside products would also have better preventive anti-

coagulation—that is, fewer cases of venous thrombosis, as well as fewer side-effects (ie, lower risk of bleeding). In 1992, I took part in a meta-analysis of randomised trials of low-molecular-weight heparin versus classic heparin. Yes, we found those effects: more effective prevention and, if anything, an equal or lower risk of bleeding. The funnel plot, by the way, was symmetrical. Being inveterate sceptics, we divided the trials into two groups: those with a somewhat “higher” quality score and those with a somewhat “lower” quality score. Completely arbitrary, of course.

Among the trials of surgical patients that had the higher-quality scores, generally published in journals with somewhat higher impact factors, the picture was less rosy: the preventive benefit was still there, though much smaller, but it was now countered by a clear tendency for an increased risk of bleeding. This seems logical, as far as anticoagulation goes: the more preventive efficacy, the more bleeding. By contrast, the trials of lesser quality showed a very large preventive benefit as well as many fewer cases of bleeding.

There have been more trials and more meta-analyses, but in general, their conclusions have shifted to a position of caution. They now state that indeed the newer heparins are a great advance, especially for their ease of administration and because they do not necessitate a stay in hospital, but on the whole their therapeutic and side-effect profiles are being judged more and more often as equivalent to those of the older heparins. Duncan Thomas has described how the original biochemical insights into the mode of action of low-molecular-weight heparins had raised the hope that they would be at least as effective as the old heparins but at the same time show less risk of bleeding.

Subsequent biochemical studies, however, revealed that the mechanism of action was, after all, closer to that of classic heparin, so a therapeutic benefit was not that likely. Apparently, the earlier and methodologically weaker trials were much more subject than the stronger and subsequent trials to some “guidance” by the original biochemical insights.

Facts and explanations

A hopeful pattern seems to be emerging: could it be that when empirical observations, say from randomised trials, align themselves perfectly with basic science theory, and vice versa, we are on safe ground?

Let us examine closer the correspondence between facts and explanations, going back to cigarette smoking and lung cancer. The empirical evidence from observational epidemiology is overwhelming. For almost half a century,
we have lived with the conviction that tobacco causes lung cancer, a notion based almost solely on epidemiological evidence (the animal experiments with tobacco smoke were somewhat artificial both in exposure and in outcome). What we lacked was a biological mechanism. Some 2 years ago, a great “hurray” greeted the pages of *Science*: in bronchial epithelial tissue, certain compounds of tobacco smoke, benzo[a]pyrenes, form DNA adducts at mutational hot-spots of the p53 oncogene.15

At last, we had an explanation! When you look at the paper, it is an associative type of explanation. We are told that in certain carcinomas of the lung, p53 has often undergone mutation, though not always the same mutation, and that the putative mutagens from tobacco smoke bind preferentially, though not always, to these hot-spots. They do so, not only in bronchial epithelial cells, but also in other cells.

Suppose that we did not know anything about the epidemiological evidence of tobacco causing lung cancer. What would we make of these observations? First, they would most certainly not have been made. Nobody would have regarded this as a research topic of interest. Second, they would most certainly not have been published by *Science*. They might not have been published at all—because they would have no meaning. On the contrary: the finding that putative mutagens could bind to p53 hotspots, where they might cause mutations, would be counter to the prevailing knowledge, since it would not be known that tobacco, from where these compounds come, causes cancer. These observations would most likely be discarded as probable laboratory error, or as something quite aspecific occurring in all kinds of cells, and as yet unexplained.

This thought experiment has become quite real in the side-effect controversy about third-generation contraceptives. One coagulation test on acquired activated protein C resistance has been devised that seems to separate the newer contraceptives from the older ones.16 Some of us have already cried “hurray”: now we have not only the epidemiology but also the beginning of a molecular explanation. Those researchers who think that the epidemiological evidence shows there is nothing wrong, however, have allied with other coagulation scientists to say that the results are most likely to be a laboratory error, and even if confirmed, they will most probably be meaningless epiphenomena.

In the end, in the tobacco example as well as in the oral-contraceptive example, it is our belief in the epidemiology that colours our perception of the basic science. Side-effect controversies are very nice examples for making us wonder about the shaping of medical knowledge.

I must confess that whenever I run along this train of thought, I get slightly nervous. Don’t we come dangerously close to a totally relativistic position about truth? On the one hand, we use a single chemical argument, about the impossibility of chemical action at infinite dilutions, to discard even the very best randomised trials on homoeopathy. On the other hand, we gladly permit randomised trials to wipe away the immunological theory about the prevention of gram-negative septic shock.

**The interplay**

What is the interplay between fact and theory? In conversation, Henrik Wulff from Denmark told me that in his view completely empirical knowledge could not exist: it would be a collection of facts like a pile of grains of sand, without any structure or purpose. We need the theory to give it structure and purpose.

The theory will vary over time. If 50 years ago, you asked a doctor why he used salicylates for fever, he would have answered, “because it has antipyretic effect”, which is saying in Latin and Greek that it lowers fever, a trick at which medicine is so apt. If you had asked me 25 years ago, when I received my medical training, I would have answered that salicylates reset the body’s temperature clock in the hypothalamus. Today, students will tell you about endorphins, prostaglandins, and, especially, cytokines. The theory changes and becomes more and more detailed, which is not important. The important observation is that we need a framework of explanatory stories to order the facts.

Observers of medicine have always understood that we need both, as we can read in these captions from an introduction to medicine at the end of the 18th century. In his introduction to the teaching of physic in Edinburgh between 1768 and 1789, William Cullen stated: “. . . for two thousand years past there have been two plans proposed . . . the dogmatic and the empiric . . . and that, in the present state of science, either of them is by itself insufficient”.17 It is intriguing to recall that very little of the medicine he taught, either the theory or the empirical observations, was effective in our modern eyes; nevertheless, somehow he understood the distinction. He also understood that they constantly correct each other.

Facts and theories remain inextricably linked. They are linked at two levels. At a general level, facts cannot meaningfully exist without theory, even if that theory is a mere tautological description, such as salicylates being called antipyretic. At a deeper level, more disturbing, and leading to some of the bitterness in the controversies, is that the very generation of facts, the production of facts, is loaded with theory, with opinion, or, depending on your point of view, with prejudice—as in several of the examples I have given.

That dual link explains why at the cutting edge of scientific progress, where new ideas develop, we will never escape subjectivity: at a single point in time, we will all make our choices differently because of different insights and therefore different interpretations. Only the future will tell who was right.

**The role of journals**

Where do we obtain our facts as well as our theories? Both are being published daily, in the medical journals that we read. Medical journals help to shape our medical knowledge by supporting a theory, by challenging a theory with facts, or by deliberately pointing to new theories. Who decides what we read? The editors. Which brings me to the second part of my lecture.

There is a frequent misunderstanding by medical scientists, those who produce the papers, that editors are there to publish what is scientifically worthwhile. There is an equally frequent misunderstanding by practising physicians that medical journals provide continuing education, much like a textbook in weekly deliveries. Both should read Sir Theodore (“Robbie”) Fox’s account on the role of medical journals.13 Sir Theodore worked for *The Lancet* for nearly 40 years, and was editor for about half that time. Close to retirement in 1963, he gave three lectures on medical journals. His account opened with one of my favourite quotes about the role of medical publishing in the care of the patient: “Whenever a doctor meets a new
patient or a new problem, he ought to have at his disposal every known fact that may help him."

This quote ought to be a motto for all libraries of medicine, or even better perhaps the logo of an internet site. But, we have our problem again: the "facts" are mentioned. They are already qualified as all the facts that are helpful, which is a subset of all available facts. Sir Theodore was too clever an editor to leave it at that. He makes short shrift of the notion that medical published work might be a mere collection of facts brought in by authors. He made an important distinction between "recorder journals" and "newspaper journals". Recorder journals might come close to the desired function of a journal from the point of view of the author: they are not really meant to be read, but only to deposit science for future reference. They are important as I never cease to tell junior residents: your science does not exist until published. General journals—newspaper journals—by contrast, have other aims: they are active shapers of knowledge and of opinions. In the words of Sir Theodore: newspaper journals "inform, interpret, criticize, and stimulate." 34

This means, and Sir Theodore was quite explicit in this, that they take sides and choose to print whatever they think is necessary for the enlightenment of their readers. That is true at the several levels at which they publish. Newspaper medical journals not only publish digested news or opinion, but also retain some recorder function, at least for very important facts. Even at that level, editors make choices. I like to quote the delightful little statistic from Ben-Shloma and Davey Smith about the publications on the Barker hypothesis in the British Medical Journal and The Lancet. 19 This is a much debated hypothesis about prenatal and early-life influences in the development of chronic disease later in life. In the BMJ almost all published papers supported the Barker hypothesis, whereas The Lancet was at most neutral but preferred to be critical—the difference was statistically significant.

Editorial campaigns

In the past 50 years, the BMJ and JAMA have strongly waged campaigns against smoking. The Lancet supported the UK National Health Service. The New England Journal of Medicine has waged campaigns against gun ownership, as has the Annals of Internal Medicine generally for a more enlightened outlook of internal medicine. More recently, the BMJ, Annals of Internal Medicine, and JAMA have greatly championed evidence-based medicine, whereas The Lancet and the NEJM have taken a little distance.

General medical journals have been instrumental in making medicine and medical research accept medical ethics and medical ethics committees, if only by stating that they would no longer publish experiments on human beings unless they had been approved by an ethics committee. From time to time, they have of course extended the ethical theme somewhat, for example when deciding that duplicate publications or advance publications are almost as grossly unethical as downright fraud. Some observers could not help but wonder whether these extensions of their undisputed moral role did not, by accident, also serve their economies—economies that have for the "big five" rather become like those of any newspaper: they live by selling and they sell more if the readers think them interesting. Losing readers is the only way. Imagine someone writing that the editors of some medical journal have completely sold out their journal into the hands of "basic molecular scientists". Would that strike as an accusation? Not really—though, to some it might.

In the end, editors of general newspaper journals are much like politicians. They have a vision of a better world, and they want to impose that vision on their readership, their electorate. To be able to do so, they should continue to be read, just as politicians should continue to be elected.

An internet without editors?

Shall we ever escape these editorial politicians? Opposition looms. The first type of opposition dreams of a world without censoring by opinionated editors who are responsible to no-one. The internet will bring it. A world should be created in which all facts are put on the internet, for anyone to see, without any censoring and with open discussion appended to it. Imagine that from this day, we abolish all edited journals by decree, and we invite anyone who thinks that he or she has a relevant fact or opinion to put it somewhere on the internet. Tomorrow a mass outpouring of facts begins.

The first consequence would be that this mass of facts would cease to exist as medical knowledge, because it is irretrievably lost to the ordinary medical doctor. He or she does not have the time and the search capacities. And even if this were remedied by computerised search machines, the number of facts and opinions that are retrieved mechanically would be so overwhelming and at the same time so conflicting that the poor general practitioner would not know what to make of them, and cease to look at them.
Within a few days after the abolishment of edited journals, I guess that we will be receiving commercial e-mail messages from some young unknowns who tell us that they have the time as well as the technical resources to retrieve the information for us. They will add that, of course, since so much information on the internet is nonsensical, they also will obligingly sieve out the nonsense. They will sell us the compiled, and sometimes slightly abbreviated information at three levels: one about particular subjects, suited to active scientists in a subdomain; one that is more general, corresponding to a medical specialty, say, the circulation, the gut, or the brain; and finally they will bring a general selection with information and digested news that is of value to anyone who wants to keep abreast of medicine. Because they are not conversant with all aspects of medicine, they will invoke the help of some trusted friends to read the retrieved information and judge its quality. Within a week, my best guess, our edited journal system will be born again, inclusive of peer review.

Why would these young unknowns set up such a service for us? Well, to gain a little money on the side, perhaps, because it is more leisurely to sit behind a computer than to be on call, but also because they have a message. They will state that some area or some aspect of medicine is underserved, and they will emphasise it in their selections. That is what will make them tick, more than the money.

**Cochrane Collaboration**

Ideology is always behind editing. That much becomes clear when we look at the second type of opposition, which is already on the internet and quite successfully so. The Cochrane Collaboration is a structured attempt to make sense of all the existing printed and computerised medical information that pertains to medical practice in diagnosis and therapeutics. The actions of the Cochrane Collaboration are not basically different from those of an editor. They collect, select, compile, and bring edited messages: summary plots of risk ratios. They might select, by criteria other than our existing general newspaper editors, but they do select, and their selections are based on a very specific view of their role in medicine. They maintain that they might be better servants of medical practice, and some general medical journal editors make allowance that there might be some truth in those claims, and have started alliances. It remains a competition between different editorial systems with different ideologies. The Cochrane Collaboration does not do away with the editorial function; it is, if anything even fiercer, because it wants to remedy what the founders of the collaboration see as shortcomings of the ordinary general medical journal.

**First editorial message**

So, we will live with editorial censoring and editorial opinion, forever. It has never been different. In the first editorial of the first issue of *The Lancet* in 1823, the founder Thomas Wakley clearly stated his mission: “It had long been a subject of surprise and regret . . . that . . . there has not hitherto existed a work that would convey to the Public . . . reports of the Metropolitan Hospital Lectures . . . we shall exclude from our pages the semibarbarous phraseology of the Schools, and adopt as its substitute, plain English diction . . . we are aware that we shall be assailed by much interested opposition . . .”. His mission was the dissemination of information that was previously held jealously as a power base in the capital; he would do so in plain English, and most tellingly, against much interested opposition. The italics are in the original.

When recollecting some *Lancet* features of the past decades, the championing of the National Health Service, the opposition against some previous ironlike prime minister, the stand against the previous leadership of WHO, and most recently the warnings against several pharmaceutical industries and even against a single charity, wouldn’t you think that Wakley had never left the office?

I thank C. M. J. B. Vandebroucke-Grauls for help and support in the preparation of this lecture.

**References**