The Willis Lecture 2003

Evaluating Treatments for Stroke Patients Too Slowly
Time to Get Out of Second Gear

Charles Warlow, MD, FRCP, FMedSci

Thomas Willis was born in 1621, received his classical education and then medical degree from Oxford in 1646, became a highly successful and prosperous physician, and, in 1660, was appointed Professor of Natural Philosophy at the University of Oxford. Although at that time there were no neurologists as such, he coined the word neurology in his *Cerebri Anatome*:

“We should deliver an exact neurology, or doctrine of the nerves” and “without the perfect knowledge of the nerves the doctrine of the brain and its appendix would be left wholly lame and imperfect.”

Of course, his pharmacopoeia was primitive in the extreme and he had no reliable way of proving that any treatment interventions worked, unless they had a very obvious and immediate effect. But these days we have the most remarkable tool to evaluate our interventions, the randomized controlled trial, although it is not yet 60 years since the first trial in the modern era was published in the *BMJ* on October 30, 1948, of streptomycin for pulmonary tuberculosis (Figure 1). It was conducted in the UK just after World War II when the supply of the antibiotic was so limited that Austin Bradford Hill, the statistician, was able to persuade clinicians to use randomization between the new treatment and control as a form of fair rationing and to construct 2 groups of patients who were so similar in their prognosis at baseline that any definite difference in their outcome must be because of the new treatment.

And in the past 3 decades we have had another tool, meta-analysis, largely developed in Oxford by Richard Peto the statistician and Iain Chalmers who devised the Cochrane Library, and in the US by the late Tom Chalmers. Also, we now use the familiar forest plots to display the data in such an elegant and parsimonious way. The first such plot was probably published by John Lewis (Figure 2a). Each trial was then represented by a rectangle whose horizontal length was proportional to the confidence interval around the treatment effect shown as a short vertical line in the middle, and the pooled effect of all the trials was represented at the bottom as another rectangle. Steff Lewis, daughter of John, who is the statistician in our stroke research group, has produced an up-to-date and so more recognizable forest plot of that first attempt by her father (Figure 2b). The strength of the studies, essentially the number of outcome events, is proportional to the size of the black boxes, an idea popularized by Richard Peto, who thought that the horizontal lines representing the confidence intervals would confuse clinicians who might think, wrongly, that the longer the line, the more informative the study. The point estimate to the left of the vertical line of no effect generally represents treatment better than control, and the pooled effect derived from all the trials is shown by a diamond at the bottom, with the confidence interval represented by its horizontal length. To make matters convenient for the busy clinician, all the information from all the trials ever performed is being gradually gathered together electronically in the Cochrane Library, which is updated every 3 months. For just stroke, this library now contains 6755 references to 3160 trials and 72 systematic reviews produced by the Cochrane Stroke Review Group, coordinated by Peter Sandercock, and their number is increasing all the time (www.dcn.ed.ac.uk/csrg/).

Therefore, our ability to evaluate therapeutic interventions flowing from the work of basic scientists and elsewhere has been transformed beyond all recognition. What might work can be put to the test; does it actually work in real life? No longer are theory and anecdote good enough, however persuasive the scientific hypotheses; facts are needed from randomized trials. An apt text comes from Thomas Huxley in 1893, “The great tragedy of science—the slaying of a beautiful hypothesis by an ugly fact.”

Progress in Therapeutic Interventions for Stroke
Therapeutic progress in stroke has been astonishing since I was a medical student 40 years ago. The antiplatelet effects of aspirin were not described until the late 1960s. Anticoagu-
The Carotid Surgery Story

The introduction of cerebral angiography in Portugal in the 1920s by Egas Moniz\textsuperscript{11} and the rediscovery of arterial disease of the carotid artery in the neck rather than "cerebral thrombosis" as a common cause of ischemic stroke by people like Miller Fisher in the US\textsuperscript{12} encouraged surgeons to bypass or remove atheromatous lesions of the carotid artery in the neck. After early failures in China\textsuperscript{13} and Argentina,\textsuperscript{14} the first successful carotid endarterectomy was performed in the US in 1953 by Michael De Bakey.\textsuperscript{15}

Nine years later in 1962, Bill Fields and others in the US started the first randomized trial of carotid endarterectomy.\textsuperscript{16} This was an amazing achievement, considering randomized trials had only recently been developed. But the trial was too small to provide any definite evidence of efficacy and had absolutely no effect on American surgeons and many other surgeons who just continued operating. The origin of the internal carotid artery was just too accessible to be left alone!

By the late 1970s, the situation was completely out of control. Carotid surgery had become epidemic, at least in the US. Trivial lesions were being removed from patients who had no symptoms at all, neurologists were pointing out the risks involved and that these risks were often down-played by surgeons, and there was huge variation in practice.\textsuperscript{17} I spent many days of my life in the 1970s reviewing innumerable case series that added nothing to the evidence, only fuel to the argument. My boredom was only temporarily relieved when I read in the \textit{Archives of Surgery} that after the operation, "the patient stated that he could feel more blood going to his brain."\textsuperscript{18}

So the stage was set for our large European randomized trial, which began in 1981, and Henry Barnett’s equally large North American trial, which began in 1987. The much smaller Veterans Affairs Trial\textsuperscript{19} was stopped early when in 1991 the 2 large trials first reported that surgery did indeed benefit symptomatic patients, if the carotid stenosis was severe.\textsuperscript{20,21} During the same period, trials in asymptomatic stenosis patients showed there was benefit from surgery but, at least over a matter of a few years, the risk of stroke without surgery was not high enough to make the risk of surgery itself worthwhile.\textsuperscript{22} Evidence of any benefit over a longer period will eventually be provided by long-term follow-up in the Asymptomatic Carotid Surgery Trial.\textsuperscript{23}

This balance of early risk of surgery versus longer-term risk without surgery has become the key issue in selecting patients for treatment, not just those who \textit{might} have a stroke without surgery, but the smaller number who \textit{will}. After all, not everyone with even severe stenosis has a stroke, maybe 4 in 5 do not, either because they do not survive long enough or because in some way their carotid lesion is "safe" and unlikely to lead to more emboli to the brain. We must reduce the number needed to operate" to prevent 1 stroke, not only to make the operation more cost-effective in economic terms, but also to avoid exposing patients who do not actually need the operation or the risk, discomfort, and anxieties of surgery. After all, unlike drugs that can always be stopped if adverse effects occur, an operation that is done is done and cannot be undone.

Working out just who will benefit from carotid surgery is the aim of the ongoing individual patient data meta-analysis
of the 3 symptomatic trials being organized by Peter Rothwell in Oxford and Michael Eliasziw in Calgary. The data from the trials have been made compatible with each other, particularly how carotid stenosis was measured and strokes defined, and it is now clear that the trials actually have remarkably similar results (Figure 3). Furthermore, we can now better-examine the treatment effect in subgroups, checking that hypotheses generated in 1 trial are confirmed in the others. For example, the extraordinary importance of early surgery for maximum absolute benefit was suggested in the European trial and confirmed in the others.

All this makes a nice story, but it is more than 50 years since it started and it still has not finished. I have illustrated it, and the others stories, in Figure 4, using a traffic light analogy. From the mid 1940s to the early 1960s, the color of the carotid surgery row is red because at that time, there was no evidence of benefit but the operation had been suggested and was being performed; so, red means “stop,” because it was not yet routine clinical practice. The randomized trials started in the 1960s, so we were “amber” (yellow) (prepare to go like the traffic lights), meaning we should wait before crossing to routine clinical practice. The trials were inconclusive and stopped in the 1970s (so we were “red” again because there was no evidence of benefit) and restarted in the 1980s (amber). The trials did not report until the 1990s, when at last we became “green” (meaning to go), and the operation became securely embedded in clinical practice.

So why did all this take so long? There are several reasons. First, the operation was suggested when randomized trials had only just begun, and this methodology was embraced by
physicians for evaluating drugs before many surgeons were willing to put their operations to the test. And as Dave Sackett memorably remarked, “Therapeutic reports with controls tend to have no enthusiasm, and reports with enthusiasm tend to have no controls.”

26 The Bill Fields trial was well ahead of its time but the apparently negative result, a false-negative as it has turned out, annoyed the surgeons, and the physicians did not insist on the much larger trials that were needed for another 10 years. And when the physicians did, many surgeons were openly hostile and tried to prevent randomization. Some accused me of deliberately designing the European trial to make sure the result was negative. More money would have helped. The European trial was performed “on a shoe string” and we had to go to our Medical Research Council 3 times to keep it going. Research funding bodies so often expect results in 3 years, or 5 years maximum, which is the normal length of a research proposal, not the 15 to 20 years that may be needed. And stroke research is relatively underfunded, certainly compared with heart disease research. All these factors made the whole recruitment process far too long. It took 13 years in the European trial when it should have taken 2 or 3 years. Of course, the key to treatment success here lies in the long-term follow-up and we could not shorten that. The patients had to be followed-up for years, and even now the individual patient data meta-analysis is taking a long time because when data are owned by >1 party, there has to be considerable discussion about how analyses should be performed and presented. Maybe making trial data sets publicly available is a good idea, as is now the case for National Institutes of Health (NIH)-funded trials, but this would probably be impossible if commercial interests were at stake.

The Carotid Angioplasty Story

Despite all the lessons of the carotid surgery story, I fear history is repeating itself with angioplasty and stenting as an alternative to carotid surgery. This technique was developed in the 1960s (red in Figure 4) for the leg arteries and applied to the carotid artery from the early 1980s. Numerous case series have told us nothing about the relative safety and durability of the procedure compared with surgery. Trials started in the 1990s (and so to amber). The CAVATAS trial, published in 2001,28 was too small and therefore inconclusive, and other trials have been even smaller. Many of those performing angioplasty have continued on regardless, and the present trials are having difficulty recruiting. It has taken years to raise the funding for these new trials; some surgeons are once again trying to obstruct randomization, and so are some of those who perform angioplasty. Therefore, we are still in the amber section of progress and a long way from the “green light” of normal clinical practice. We do not seem to be learning very fast how to evaluate new interventions to prevent stroke, particularly perhaps when vested interests are at stake.

The Aspirin Story

Because it was known to cause bleeding, aspirin was first suggested as an antithrombotic drug in 1953 by Lawrence Craven, a general practitioner working in Glendale, a suburb of Los Angeles, but no one noticed because he published his thoughts in the now-defunct Mississippi Valley Medical Journal.29 The antiplatelet effects were not discovered until the late 1960s,8,9 after which several trials were started, from red to amber in Figure 4. For stroke prevention, the landmark 1978 Canadian trial of 585 patients was big by the standards of the day, but it was too small to be convincing on its own.30 The first Oxford meta-analysis of all the 25 antiplatelet drug trials then available appeared in 1988, and the medical community became reasonably convinced that aspirin should be used for secondary stroke prevention, so we were “green” on the slide.31 The second cycle of the meta-analysis published in 1994 was even more convincing,32 and the latest cycle involving ∼200 000 randomized patients in 287 trials

Figure 3. The overall results of the 3 main trials of carotid endarterectomy in symptomatic patients. From Figure 3 in Rothwell et al (2003)24 showing the absolute risk reductions (ARR) and relative risks (RR) in any stroke or operative death in the European (ECST), North American (NASCET), and Veterans Affairs (VA)309 Trials, by degree of carotid stenosis. p(het) indicates significance of differences between trials in relative risk reductions.
was published in 2002. So why has all this taken ~50 years?

First, aspirin was suggested not long after randomized trials were introduced; however, trials did get going fairly quickly once the antiplatelet properties were discovered. But it was not clear until the 1980s that when a treatment effect is rather modest, remarkably large sample sizes are needed to reveal it above the “background noise,” a lesson that Richard Peto and others have again and again emphasized to clinicians. Modest effects are worth having from the public health perspective if a disease is common, disabling, and expensive, as is the case for stroke, and the if treatment is not too expensive. Of course, from the individual patient perspective, the treatment must be acceptable, in other words easy to take and free of significant adverse effects, as is the case for aspirin. And “the number needed to treat” should not be so high that hundreds of people have to be treated, perhaps for years, to prevent 1 of them from having a poor outcome, in this case a stroke or other serious vascular event.

Other Antiplatelet Drugs
For dipyridamole, although its antiplatelet properties were discovered at the same time as aspirin, there is far less randomized evidence. In combination with aspirin, it may be more effective than aspirin alone in preventing strokes but not, it seems, in coronary events. I think more trials are needed, and I suspect that Boehringer-Ingelheim, who make dipyridamole, like many other pharmaceutical companies, have not been persuaded that really large trials were and are necessary, and maybe did not want to or were unable to fund them even if they were persuaded. We are still “on amber” for dipyridamole with aspirin (Figure 4). Although Sanofi were bolder in funding a 20 000-patient trial, they were unlucky that the thienopyridine clopidogrel in the CAPRIE trial did not seem to be very much better than aspirin, but at least thienopyridines are an alternative for those few who are aspirin-intolerant; so, we are now “at green” (Figure 4).

Of course, it is very difficult to get adequate funding for the very large trials that are required. Quite recently even the combined resources of the UK Medical Research Council, the Veterans Affairs Cooperative Studies Program, and the Canadian Institutes of Health Research were insufficient to fund a proposed trial large enough to test, in both de novo patients and so-called aspirin failures, aspirin versus the combination of aspirin with dipyridamole versus the combination of aspirin with clopidogrel.

Aspirin for Acute Stroke
Antiplatelet drug trials for acute stroke were very delayed. The International Stroke Trial and the Chinese Acute Stroke
Trial showed a modest benefit for aspirin, but not until 1997.36,37 I think this delay occurred because just performing acute stroke trials was much more difficult than secondary stroke prevention trials. Acute stroke patients were distributed over the medical wards and therefore difficult to study in large numbers in an organized way, and we were all daunted by the large sample sizes that the cardiologists had shown us were required in their acute coronary trials. And, of course, aspirin had no commercial potential and so the trials had to be funded almost exclusively by government at the time when many, including even the UK Medical Research Council (MRC), had been seduced by the theoretical attractions of neuroprotection. The situation was worsened by the increasing competition for patients and the willingness of clinicians to accept the extremely generous per-patient funding provided in the industry-led neuroprotection trials, which takes us to the next story.

The Neuroprotection Story
Neuroprotection for acute stroke has, so far, been a failure.38 It was first suggested in the early 1980s. Trials were started fairly quickly, some would say far too quickly before the basics were sorted out, and we are still “at amber,” waiting for a convincingly positive result. Part of the problem is that once again, industry is expecting too much of the compounds and rushing to get to market before the patents expire. The drugs were not all that promising at the preclinical level, and the trials were far too small to demonstrate a modest but still clinically worthwhile effect. However, a small effect may not be commercially worthwhile and therefore not worth looking for.

The Anticoagulant Story
Anticoagulants go back even further than aspirin. What came to be called heparin was discovered in 191639 and used in patients from the 1930s, whereas oral anticoagulants were discovered in the 1930s and introduced into clinical practice in the 1940s.40 By the 1950s, both heparin and oral anticoagulants were being used by enthusiasts to treat and prevent strokes. Although the early randomized trials were very small and unconvincing, some physicians proved themselves as tenacious as some surgeons in clinging to their treatments despite the lack of evidence. For acute stroke, the true believers have only recently let go.41 However, it has turned out that long-term anticoagulation for stroke prevention works, at least in atrial fibrillation patients; the enthusiasts were right, as of course they often are. In 1993, the Dutch-led European Stroke Prevention Trial showed that warfarin was really excellent for secondary prevention, and better than aspirin.42 But for those in sinus rhythm, the enthusiasts were wrong, as of course they often are, and so far there is no trial evidence indicating that warfarin is better than aspirin.43

Although the delays here were perhaps caused by the problem of funding trials of drugs of no interest to industry, I suspect the problem was mostly the climate of acceptance of a long-established treatment, albeit with no supporting randomized evidence.

The Thrombolysis Story
And what of thrombolysis? What came to be known as streptokinase was described in 1933 and purified enough to be introduced into clinical practice in the 1950s44 (red in Figure 4). Thrombolytic treatment was described for stroke in 1958,45 and the first randomized trial was published in 1963.46 Of course, the early trials could not exclude intracerebral hemorrhage before computed tomography was introduced in the early 1970s and the patients fared very badly. It was not really until the 1980s that randomized trials were restarted, but, from Joanna Wardlaw’s Cochrane Review,47 I do not think the evidence is yet convincing enough to change routine practice worldwide. More and bigger trials are underway but are delayed, I fear, because of competition for patients from the far better-funded industry neuroprotection trials and hostility from some grant application referees who believe in thrombolysis and cannot tolerate the genuine uncertainty of others (there is, I believe, a general problem in refereeing large multicenter grant applications because there are few referees independent of the study and those who are may well be negative). Furthermore, neurology is traditionally an outpatient specialty, so most neurologists find it difficult to be available 24 hours per day and fully engaged in hyperacute medicine. Those who are often are overburdened with routine clinical work; acute stroke trials require serious commitments of time and energy. And, of course, there are not that many stroke specialists, so we should be collaborating more with acute emergency physicians and geriatricians.

So, 70 years after the start of the thrombolysis story and 30 years after the introduction of computed tomography we still do not quite know who to treat, even in well-organized and enthusiastic stroke units. We should learn from the cardiologists who, although they did not believe the 1985 Oxford meta-analysis,48 got themselves organized, randomized 60 000 patients, and as a result now treat acute myocardial infarction routinely with thrombolysis.49

The Blood Pressure-Lowering Story
Despite the vast amount of evidence for primary prevention of stroke by blood pressure-lowering,50 there was, until very recently, remarkably little evidence for secondary prevention. But in 2001 the Perindopril Protection Against Recurrent Stroke Study (PROGRESS) trial showed that lowering blood pressure in the long-term really does substantially reduce stroke recurrence risk, and probably the lower the pressure, the better.51 Servier were bold to fund PROGRESS, which was an investigator-led trial. They had no influence on the design, data collection, analysis, and reporting; it is a model of how trials should be.

But why were we so painfully slow with blood pressure-lowering? Maybe those who look after well people with high blood pressure performed numerous trials, but neurologists and stroke physicians did not talk much to them or to us. Or was it that neurologists did not pay much attention to treatment in general and of stroke in particular? For example, I have heard it said that it is not neurologists in the US who treat blood pressure but general internists. Or were we prepared to generalize from the primary prevention evidence without performing trials in stroke survivors? And maybe we who knew so much about autoregulation were too fearful of hypotensive-induced brain damage. Certainly, this fear has made trials in acute stroke of blood pressure-lowering very
controversial,52 and blood pressure-raising may even be a better option.53 We are still “at red” (on Figure 4), we just do not know what to do with the blood pressure in acute stroke. And once again, funding is a huge problem.

The Cholesterol-Lowering Story
Long-term cholesterol-lowering after stroke was not considered until the trials to prevent coronary events suggested that stroke occurrence risk was also reduced. But now we know that stroke recurrence is also reduced from the Heart Protection Study.54 We were misled by the observational epidemiology, which had not shown that increasing cholesterol was definitely a risk factor for stroke.55

The Good News Story: Clips Versus Coils
Sometimes an intervention can and is evaluated and introduced quickly into clinical practice. My last story is about detachable platinum coils. They were introduced in 1990, and within only 4 years patients were being randomized to coiling versus clipping in the International Subarachnoid Aneurysm Trial.56 In 2002, this trial showed a very clear advantage to coiling in anatomically suitable ruptured aneurysms, less death and dependency and less epilepsy, although longer follow-up will be needed to assess the durability of coils in preventing recurrent rupture. What is so remarkable is that this trial involved a head-to-head comparison of interventions by 2 separate groups of specialists, interventional neuroradiologists versus neurosurgeons. And yet within just more than a decade, we have a new intervention evaluated and into practice, and so are “at green” (Figure 4). So what was the secret of this success? Perhaps because the trial was led by 2 colleagues in Oxford, Andy Molyneux the neuroradiologist and Richard Kerr the neurosurgeon, perhaps because it was centered in Oxford where Richard Peto had established such a strong trials culture, perhaps because the UK MRC funded the trial reasonably generously, and perhaps because in some of the high randomizing centers the only way to get coils funded was in the trial, an echo back to the streptomycin story. Of course, it remains to be seen whether a trial not performed in the US is accepted by US neurosurgeons. I imagine there will be little resistance from US interventional neuroradiologists.

The Future
All the successful treatments in stroke have so far been supported by very large trials and meta-analyses, large enough to provide precise results and therefore to be convincing. And almost all were driven by academic investigators funded by government agencies and charities. When successful trials were funded by industry, they were performed completely independently, like PROGRESS51 and the Heart Protection Study.54 Trials funded, organized, and controlled by industry have not so far been particularly successful, and in the future, better partnership between academia and industry will be required.

I fear we have now done most of the easy bits of evaluation. But there is still a lot to do. What about evaluating the myriad of techniques that are used in rehabilitation, and in stroke nursing, and in influencing the abnormal physiology in acute stroke? What about evaluating the myriad of techniques being used in evaluating more mundane but still important issues such as feeding, like in the Feed or Ordinary Diet (FOOD) Trial being performed by Martin Dennis?57 Of course, these trials are difficult to design, but the challenge is well-worth taking. And what about sorting out intracranial vascular malformations in which there is not a single randomized trial of radiotherapy, surgery, or interventional radiology, all of which are in common use? Too difficult perhaps, but what a challenge to the younger generation of stroke researchers. And when all that is done, what about starting on the proper evaluation of diagnostic tests and screening?

I have mentioned a lot of the hurdles in getting treatments evaluated faster, changing hypotheses into facts, and how to jump over some of these hurdles. In summary:

- Some treatments came into routine clinical practice before or only just after randomized trials were developed. This is no longer an excuse. Randomized trials are the gold standard for evaluating all new treatments, not just drugs. And meta-analysis is the way to understand the results of several trials and plan future research when necessary.
- Randomization of a new treatment needs to be started early, before the intervention becomes epidemic.
- Trials need to be surprisingly large, particularly when the anticipated treatment effect is modest but still worthwhile, as it would be if the intervention was safe and inexpensive.
- Working with industry is helpful provided that the trialists retain their independence, particularly of design, data collection, analysis, reporting, and ownership.
- Government agencies and charities must not abdicate funding to the commercial sector. There are far too many noncommercial interventions that need to be tested, such as nonpatented drugs, surgery, devices, physiotherapy techniques, and so on.
- Noncommercial funders must better appreciate that they are in competition for patients with extremely rich commercial organizations.
- Everyone needs to understand that it can, legitimately, take a long time to evaluate interventions, particularly if long-term follow-up is required, so long-term funding is needed, too.
- People who referee grant applications should not willfully wreck them, and any competing academic as well as financial interests should be declared and allowed for. Gratuitous hostility is corrosive and just delays evidence becoming available and effective treatments being widely used.
- Enthusiasts must realize that they will never get their enthusiasms into practice unless the skeptics are allowed to test them in trials.
- Ethical obstructions are getting worse and must be tackled head-on, but this is another topic that I have not had time to discuss.
- More money for stroke research would be welcome.
But trials should get easier as stroke units and stroke specialists proliferate and become more organized; we need more of us to do the job.

I hope you will agree that it is time to get out of second and into third gear and evaluate promising interventions for stroke much faster. Who knows, we may even be able to get into fourth gear! This can be achieved in the future if we learn the lessons of the past.

References


Charles Warlow

*Stroke*. 2004;35:2211-2219; originally published online July 29, 2004;
doi: 10.1161/01.STR.0000138024.30064.71

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2004 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/35/9/2211