

# PEOPLE OF CONSEQUENCE

## Sir Austin Bradford Hill and the progress of medical science

Richard Doll

Sir Austin Bradford Hill, honorary fellow of the London School of Hygiene and Tropical Medicine, dean of the school from 1955 to 1957, and professor of medical statistics from 1946 to 1961, died on 18 April 1991. With his death the country lost a man who has been described as the greatest medical statistician of the twentieth century, despite the fact that he held no degree in either medicine or statistics.

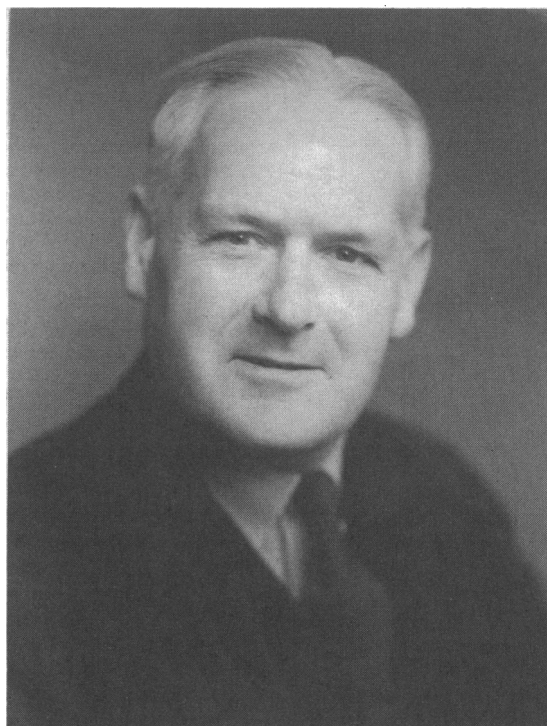
### Sir Austin's career

Sir Austin, or Tony as he was always known (causing many of his admirers to assume that the famous A Bradford Hill was christened Anthony) was born on 8 July 1897 into a family of which he was justly proud. Each of the four preceding generations of Hills had had one member or more mentioned in the *Dictionary of National Biography*, including an educationalist, a criminal law reformer, a mechanical inventor, a Johnsonian scholar, and two fellows of the Royal Society. These last were Sir Rowland Hill, Tony's great great uncle, who invented the rotary press and reformed the postal service, introducing the penny post in 1840 against strong opposition; and his father, Sir Leonard Hill, who made important contributions to understanding the cerebral circulation and the hazard of rapid decompression of divers.

At school Tony had determined to study medicine, but the first world war intervened and he enlisted at the first opportunity, in 1916, opting for a commission in the Royal Naval Air Service. Soon after completing his training he was posted to the Greek islands to support the attack on the Dardanelles. By November 1917, however, he had developed pulmonary tuberculosis and was invalided out to die at home. After an artificial pneumothorax the progress of the disease changed, and by 1919 he was sufficiently recovered to begin to think again about his future. Medicine was now out of the question, and on the advice of Major Greenwood, who had been a demonstrator in his father's department and had remained a close friend of the family, he decided to study economics as an external student of London University. By reading in bed while convalescing and with the aid of a correspondence college he obtained a BSc(Econ) with second class honours in 1922, having attended the University only twice to take examinations.

Tony enjoyed economics, but he had no desire to make it his career, and with Greenwood's support he obtained a grant from the Medical Research Council to investigate the reason for the high mortality in country districts at ages 15-30. The excess, he concluded, was most likely to be due to selective migration of the fittest to work in towns.<sup>1</sup> During this period he had an opportunity to extend his knowledge of statistics by attending the BSc course in statistics at University College London, where he was inspired more by Karl Pearson's ideas, philosophy, and enthusiasm than by the mathematical niceties of his statistical methods.

Thereafter Tony worked in association with Green-



GODFREY ARGENT

*Sir Austin Bradford Hill, described as the greatest medical statistician of the twentieth century. He held no degree in either medicine or statistics*

wood until 1939, first as a member of the scientific staff of the MRC's industrial health research board and then as reader in epidemiology and vital statistics at the London School of Hygiene and Tropical Medicine, where Greenwood had become the first professor of medical statistics in 1928. With the outbreak of war Tony was seconded first to the Ministry of Home Security, concerned with the effects of bombing on the civilian population, and later to the Air Ministry's medical directorate. In 1946 he succeeded to Greenwood's chair, becoming at the same time honorary director of the MRC's statistical research unit and adding part time deanship of the school between 1955 and 1957. In 1961 he chose to retire a little earlier than he had to, believing firmly that one should not stay in research and teaching late in life.

### Sir Austin's influence on medical science

It is no exaggeration to say that Tony Bradford Hill had more influence on the past 50 years of medical science than many winners of the Nobel prize for medicine. This is not because of the importance of particular pieces of research but because of the effect his teaching had on the way medical research has developed. Three contributions stand out: his articles on medical statistics, which appeared serially in the *Lancet* in 1937 and were reprinted in book form as *Principles of Medical Statistics*<sup>2</sup>; his development of

ICRF Cancer Studies Unit,  
Radcliffe Infirmary, Oxford  
OX2 6BR

Sir Richard Doll, *emeritus*  
professor of medicine,  
University of Oxford

BMJ 1992;305:1521-6

epidemiological methods for investigating the causes of non-infectious disease with guidelines for determining whether an observed association reflected cause and effect; and his introduction of randomisation for the conduct of clinical trials.

### Principles of medical statistics

In 1992 it may be difficult to appreciate how great an impact *Principles of Medical Statistics* had on the practice and interpretation of research results by what was in 1937 an essentially innumerate medical profession. The position was accurately described in an editorial in the *Lancet*:

For most of us figures impinge on an educational blind spot. . . . This is a misfortune, because simple statistical methods concern us far more closely than many of the things we are forced to learn. . . . Many of our problems are statistical; and there is no other way of dealing with them. In preventive medicine this is so obvious that it has acquired general recognition. In laboratory work, although it has come more slowly, it is now widely recognized that it is very unsafe to base conclusions on statistically inadequate data. In clinical medicine, it is coming more slowly still; so slowly that many avoidable errors, and a sad waste of material, still hinder progress.<sup>3</sup>

It was Tony's genius that led him to respond to this situation not by pressing the need for deferring to a statistical consultant but by insisting that workers in medical problems, in clinical as well as preventive medicine, must themselves know something of statistical techniques, both in experimental arrangements and in the interpretation of figures. Only in this way would medical workers genuinely come to appreciate the contribution of statistics and learn to accept statisticians as partners in research, who on their side had to steep themselves in the realities of medical life.

Tony's 17 articles in the *Lancet* took the reader by the hand from procedures for presenting statistics and the use of the mean, median, and mode; through the variability of samples, the calculation of the standard deviation, and differences between proportions and the frequency of their occurrence by chance; to the use of the  $\chi^2$  test, the correlation coefficient, and life tables; ending with three lucid chapters explaining common fallacies and misunderstandings.<sup>4</sup> Most importantly the statistical procedures were all described in straightforward English, eschewing symbols and almost all formulas, except occasionally to summarise what had already been fully explained.

This presentation of the aim of the statistical method and of its practical application to the sort of problems that physicians regularly encountered met a need so successfully that when the articles were republished in book form a third edition had to be reprinted twice and a fourth, expanded edition produced within ten years. Various editions were subsequently translated into Spanish, Korean, Indonesian, Polish, and Russian, and an enlarged 12th edition appeared shortly after he died, with his son I D Hill as a joint author.<sup>5</sup>

Statistical analysis is now an integral part of almost every medical publication, due to the work of many gifted statisticians throughout the world. That the medical profession awoke to its need in the middle of the century was due primarily to its exposition by Tony Bradford Hill.

### Epidemiology of non-infectious disease

Through his investigation into the reasons for the increasing mortality attributed to lung cancer Tony developed model case control and cohort studies and set out guidelines for determining whether an observed association could be interpreted as indicating cause and effect.

### CASE-CONTROL STUDY

When the investigation of lung cancer began, the idea that epidemiological techniques could be used to discover the causes of non-infectious disease appealed to very few research workers and the few case-control studies of cancer and heart disease that had been carried out were relatively crude. In Bradford Hill's study of lung cancer 20 London hospitals were requested to notify patients who had been admitted with a diagnosis of cancer of the lung, stomach, or large bowel, without specifying which, in the hope that the social workers who interviewed them would be blind to the diagnosis at the time of interview; the patients were classified after discharge from hospital on the basis of their discharge diagnoses according to the confidence with which the diagnoses had been made.<sup>6</sup> Controls were chosen by the interviewers in a specified way from patients of the same sex within the same five-year age groups and in the same hospital at about the same time, excluding only patients suspected of having cancer of some other type.

The most notable feature of the study was the analysis of the results, which set out for the first time the conditions that had to be met before it could be concluded that an association observed in a case-control study could be interpreted as indicating cause and effect: firstly, the elimination of bias in the selection of patients and controls, in the way patients reported their histories, and in the way interviewers recorded them; secondly, the existence of an appropriate time relation between exposure to the suspected agent and the development of the disease; finally, the lack of any other distinction between the affected patients and their controls that could account for the observed association or of any common factor that could lead both to the specific exposure and the development of the disease. This we would now call confounding. This paper set "a new standard for the case-control study. . . . It stands as a classic example for the investigation of a given outcome and an array of exposures. No previous research paper lays out the essentials of the case control method with such understanding and meticulous care."<sup>7</sup>

### EVIDENCE OF CAUSATION

The arguments that led to the conclusion that "smoking is a factor, and an important factor, in the production of carcinoma of the lung" were subsequently formalised in Bradford Hill's presidential address to the section of occupational medicine of the Royal Society of Medicine<sup>8</sup> and summarised under nine heads in later editions of his *Principles of Medical Statistics*.<sup>9</sup>

Tony's comment on two of these are still of special interest. Important though specificity is, he said, it must not be carried too far. Berkson, for example, argued that smoking could not be the cause of any of the diseases with which it was associated, because it was associated with so many.<sup>10</sup> This, Tony said, was nonsense. Smoke was a complex mixture (much more complex, as it turns out, than was dreamt of at the time) and its different constituents could cause different diseases in different ways, just as the consumption of milk, a carrier of infection, had been a cause of scarlet fever, diphtheria, tuberculosis, undulant fever, sore throat, dysentery, and typhoid fever. Biological plausibility, helpful though it is, is a feature we cannot demand. It was, Tony pointed out, the lack of biological knowledge in the nineteenth century that led a prize essayist writing on the value and fallacy of statistics to conclude, among other "absurd" associations, that "it could be no more ridiculous for the stranger who passed the night in the steerage of an emigrant ship to ascribe the typhus, which he there contracted, to the vermin with which bodies of the sick

### Inference of causality

Strength of association  
Consistency  
Specificity  
Relationship in time  
Biological gradient  
Biological plausibility  
Coherence of evidence  
Experiment  
Analogy

might be infected," while in the twentieth century there was no biological knowledge to support the evidence of the effects on the fetus of rubella in the pregnant mother.

Clearly, Tony concludes, "none of these nine viewpoints [see box] can bring indisputable evidence for or against a cause and effect hypothesis and equally none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us to answer the fundamental question—is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?"

#### COHORT STUDY

Tony was satisfied that the evidence of the case-control studies (combined with the ecological evidence of a relation between mortality from lung cancer and consumption of cigarettes) could lead to only one conclusion, but he found that it was dismissed by the many people, including many senior scientists, who had little understanding of the strengths and weaknesses of epidemiological methods. It was clear, therefore, that if further progress was to be made a new approach was needed, one that would avoid the difficulties associated with the interpretation of the case-control studies. It should "determine the frequency with which the disease appeared in the future among groups of persons whose smoking habits were already known." Such a study could be defined as prospective, in contrast to the retrospective case-control study, on the basis of definitions given in the *Oxford English Dictionary*.<sup>11</sup>

Doctors, Tony thought, would make a suitable population to study—they should not only be interested enough in the problem to be willing to report their smoking habits, and do so reasonably accurately, but they should also be easy to follow because of their need to remain on the medical register. Both beliefs proved correct. After a test on 200 doctors chosen at random from the medical directory, letters were sent, with the aid of the BMA, to all doctors known to be in the United Kingdom. Usable replies were received from two thirds.<sup>12</sup> Twenty years later Richard Peto and I were able to determine the vital state of 99.7% of those who had replied to the initial inquiry.<sup>13</sup>

The first results, obtained after only two and a half years, when 36 deaths attributable to lung cancer had been reported, showed a gradient in mortality from the disease with increasing amount smoked that was almost identical to that observed in the retrospective study (figure 1). Despite the small number of deaths, the trend with the amount smoked, standardised for age, was significant (two sided  $p=0.04$ ) and it was thought necessary "in view of the nature of the results to lay [them] before the survivors of the 40 000 men and women who made them possible."<sup>11</sup> Because of the short period of observation it had not been thought necessary to calculate person years at risk, which at the time had to be done clerically with the aid only of a Hollerith card sorter. It was not until more solid results were obtained after four and a half years, when a relation was also observed between smoking and coronary heart disease, that a precise analysis was carried out based on a count of person years at risk.<sup>14</sup>

Substantial numbers of deaths accumulated, and by the end of 1986—the cohort had been followed for 35 years—significant trends in mortality with increasing numbers of cigarettes smoked were observed for 25 of the 48 causes of death examined (table I, previously unpublished). With the exception of cancer of the rectum and hernia (the latter not examined elsewhere) similar trends have been observed in cohort studies in other countries,<sup>15</sup> the only cause of death not significantly associated with smoking in this series but found

TABLE 1—Causes of death associated with smoking in British doctors followed for 35 years

Cause of death	Death rate/100 000 men/year (standardised for age)			
	Cigarettes smoked/day			
	0	1-14	15-24	≥25
<b>Cancers:</b>				
Lung	14	100	182	327
Larynx		3	5	14
Mouth or pharynx	2	6	8	25
Oesophagus	2	14	26	41
Stomach	25	35	45	38
Rectum	6	10	16	45
Pancreas	13	27	25	47
<b>Vascular diseases:</b>				
Pulmonary heart disease		5	7	22
Ischaemic heart disease	526	752	825	956
Myocardial degeneration	65	122	111	175
Hypertension	34	28	49	63
Arteriosclerosis	22	31	33	67
Aortic aneurysm	12	31	65	71
Cerebral thrombosis	91	94	151	141
Cerebral haemorrhage	61	77	80	90
Subarachnoid haemorrhage	8	9	12	22
Other cerebrovascular diseases	66	122	96	126
<b>Other causes:</b>				
Pulmonary tuberculosis	4	8	8	20
Pneumonia	68	110	141	149
Chronic bronchitis	9	77	93	180
Peptic ulcer	7	11	32	30
Cirrhosis of liver*	6	12	22	65
Hernia		4	2	7
Suicide*	25	29	32	60
Poisoning*	8	15	16	30

\*Attributed to confounding with consumption of alcohol or personality, or both.

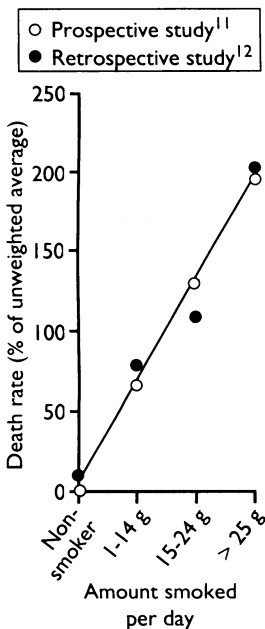
to be associated in the others is cancer of the bladder. Three associations reflect confounding with other factors (namely, cirrhosis of the liver, suicide, and poisoning) but other associations (excluding that with rectal cancer and including that with bladder cancer) are now thought to imply a causal relation. In consequence, Peto and Lopez calculated that smoking is responsible for more than 1.5m deaths a year in the 30 so called developed countries, constituting nearly 16% of all deaths over 35 years of age, with a loss of expectation of life of over 20 years for the 700 000 who die aged under 70.<sup>16</sup>

#### SUBSEQUENT DEVELOPMENTS

In the past 40 years the retrospective and prospective studies designed by Bradford Hill have become the basic tools of epidemiology throughout the world. Their names have been changed to case-control and cohort studies, and the techniques of statistical analysis have been elaborated beyond recognition so that the effects of multiple factors can be taken into account. The case-control study nested within a cohort has been developed; this can increase the power of a simple cohort study, and epidemiologists have learnt to collaborate with laboratory workers in seeking objective measures of exposure to risk. Mistakes in interpretation, however, continue to be made, less commonly from lack of specialised statistical knowledge than from, in Bradford Hill's words, "the tendency of workers to accept figures at their face value without asking themselves at every turn 'what is at the back of these figures? What factors may be responsible for this value? In what possible ways could these differences have arisen?'"<sup>17</sup>

#### Randomised controlled trial

Equally important was Tony's introduction of the randomised controlled trial. In 1946 he persuaded two MRC committees to adopt it: firstly, to test the value of a pertussis vaccine,<sup>17</sup> and secondly, a few months later, to test the efficacy of streptomycin in the treatment of pulmonary tuberculosis.<sup>18</sup> The results of the second



Relation between amount smoked by continuing smokers and mortality from lung cancer obtained in case-control<sup>12</sup> and cohort<sup>11</sup> studies: mortality in men aged 45-74 years, standardised for age, expressed as percentage of unweighted average in different smoking categories and amount smoked per day in grams of tobacco (all methods of smoking combined)

trial were published first and it has undeservedly received the accolade of the first randomised trial.

Randomisation was not, of course, a new idea. According to Peter Armitage it had been proposed in a limited way 300 years before by van Helmont, a medicinal chemist, when he challenged the academics of his day to compare the efficacy of their treatment with his. "Let us take out of the hospitals, out of the camps, or from elsewhere, 200 or 500 poor People that have Fevers, Pleurisies, etc. Let us divide them into halves, let us cast lots, that one half of them may fall to my share, and the other to yours . . . we shall see how many funerals both of us shall have."<sup>19</sup> Alas, the challenge was not taken up, and van Helmont failed to get the chance of the 300 florins he was prepared to wager on the result. The idea was not developed in a more practical way until the 1920s, when Fisher adopted it as a basic principle of experimental design in agriculture.<sup>20,21</sup> He showed how it could be used not only to ensure the avoidance of bias in the selection of subjects for treatment but also to provide an optimal estimate of the likelihood that the differences observed between the treatment groups could have arisen by chance if the treatments were in reality equally effective or ineffective.

#### A REVOLUTIONARY IDEA

It is difficult now to appreciate how revolutionary Tony's suggestion seemed when he first recommended randomisation in the clinical field. Until then new treatments had mostly been introduced on the basis of the results reported by senior members of the medical profession who had tried them out on a series of patients (or in the case of preventive measures on a series of healthy people) and concluded that the outcome was better than that reported in the past. This was fine when the new treatment had a dramatic and quickly detectable effect, as occurred with insulin, vitamin B-12, and the sulphonamides; but new treatments as efficacious as these were few and far between, and the system led to many false claims for treatments that were useless and sometimes actually harmful. It also hindered the recognition of small advances that could, however, be important if the condition treated was common. The snag was variability, both in the type of patient who presented to different practitioners at different times and in the type selected by the practitioner for treatment from among those who presented.

By the end of the nineteenth century it had been recognised that this variability could be compensated for by the simple expedient of having a concurrent series of controls and imposing a deliberate rather than a haphazard system of treatment allocation. A simple solution was to treat alternate patients with the new remedy, and this system was adopted by Fibiger when diphtheria antiserum became available in the 1890s.<sup>22</sup> It was also recommended a few years later by Pearson when he was asked to assess the value of the various forms of immunisation that had been used in the British army.<sup>23</sup> Obvious though it seems, this simple technique was adopted only slowly, and it was still the exception rather than the rule in 1937 when Tony recommended it in his series of articles in the *Lancet*.<sup>24</sup> The method was, however, far from ideal, as practice proved that bias could enter into the doctors' decisions to include patients in their experimental series if they knew which treatment the patient might be expected to receive before the decision was made about the patient's suitability for inclusion. Tony appreciated this and explained shortly before he died that in his 1937 articles "I deliberately left out the words 'randomization' and 'random sampling numbers' . . . because I was trying to persuade doctors to come into controlled trials in the very simplest form and I might have scared

## Forty five years on

In the four and a half decades that followed the streptomycin trial many refinements have been introduced that have improved the power of the technique. Two modifications to the design that were suitable for special situations were introduced by junior members of the department with Tony's advice. One was the cross over design, in which each patient serves as his or her own control. This could be used when the object was to relieve symptoms without modifying the course of the disease. The other modification was a factorial design, in which several treatments are assessed at once by giving them or their corresponding placebos randomly in four possible combinations of two treatments or eight combinations of three treatments. This can be used when the treatments are likely to act by different mechanisms.

Other developments were in the statistical techniques used to assess the significance of the results. Sequential analysis was introduced when, for ethical reasons, it was desirable to watch the sum of the results as the outcome for each new patient was added,<sup>33</sup> and others improved the statistical methods for combining the results in subgroups of patients with different characteristics and for taking account of the duration of illness or survival, as well as the fact of recovery or death.<sup>34,38</sup> Finally Peto and Peto introduced the log rank test, which allows the results in subgroups to be combined and the duration of illness to be taken into account in an optimally efficient way, and they gave, in conjunction with other leading statisticians, such a clear account of its use<sup>39,41</sup> that it has become a standard procedure.

All, however, was not well and by the mid-1970s some people again began questioning the use of randomised trials, urging instead a return to the discredited system of historical controls.<sup>42</sup> The problem was that randomised trials, even with all their technical improvements, often led to inconclusive and even contradictory results. The frustrating frequency with which inconclusive results occurred was due to trials being conducted on too small a scale. Numbers did not need to be large when effects were large, as in the case of streptomycin for tuberculosis, or when the response could be measured quantitatively, as in the trials that permitted measurement of the change in the size of a gastric ulcer; but most new treatments could be expected to have only moderate effects, and these could be difficult to recognise if the outcome assessed was an all or none event that occurred relatively rarely, such as death or stroke. These effects could, however, be important if the disease treated was common.

them off. I think the concepts of 'randomization' and 'random sampling numbers' are slightly odd to the layman or, for that matter, to the lay doctor, when it comes to statistics. I thought it would be better to get doctors to walk first, before I tried to get them to run."<sup>24</sup>

#### THE STREPTOMYCIN TRIAL

The second randomised trial that Tony persuaded an MRC committee to undertake had two features other than randomisation that made it a landmark in the history of medical research. One was the emphasis that was laid on recording events in such a way that the record would be unbiased by knowledge of the treatment that the patient had received. This could be achieved in three ways: by counting as end points only indisputable events like death; by blinding the investigator as well as the subject to the nature of the subject's treatment; and by arranging for events that involved an element of subjective judgment, such as the degree of change in an x ray shadow, to be determined independently by someone ignorant of the treatment schedule. After consideration the committee, on Tony's advice, adopted the first and third.

The other feature was the attention paid to ethical considerations. This was before any medical body had issued lines of guidance, before ethical committees to

vet proposals had been thought of, and long before such considerations had attracted the attention of academic lawyers. The first issue that had to be faced was whether it was ethical to withhold the drug from any patient. Streptomycin had been discovered two years previously; it had been shown to have a powerful effect *in vitro* and in guinea pigs, and the published results of its clinical use were distinctly encouraging. Only a small amount was available in Britain, and foreign exchange was not available to buy more. It was agreed, therefore, to use the limited supplies to treat patients with miliary tuberculosis and tuberculous meningitis, conditions that had been invariably fatal. The amount left over was insufficient to treat more than a tiny proportion of the many desperately ill people with other types of tuberculosis and the committee concluded that "it would . . . have been unethical *not* to have seized the opportunity to design a strictly controlled trial which could speedily and effectively reveal the value of the treatment."<sup>25</sup>

Two other questions that the committee considered were whether the physician in charge of patients could modify the schedule and whether the control patients should be given placebos. The doctor, it was agreed, must always "do for his patient whatever he really believes to be essential for that patient to return him to health."<sup>24</sup> This meant that if any patient seemed likely to benefit by collapse therapy—the only specific treatment available for pulmonary tuberculosis before the introduction of streptomycin—the treatment must be given, even if it upset the balance of the streptomycin and the control groups. As for the use of a placebo, this was ruled out in the patients' interest, as it would have required an intramuscular injection four times a day for four months. The response to treatment could be assessed objectively without it, psychological factors

could have little impact on such a serious disease; and there was, as Tony said, "no need in the search for precision to throw common sense out of the window."<sup>25</sup>

The committee did not discuss the need to obtain "informed consent" as it had not at the time become a subject of concern. The patient's agreement to participate was mentioned in the reports of most of the trials in which Tony was involved, but the method by which it was obtained was not described. To Tony the overriding issue was the welfare of the patient, and he inveighed strongly against the requirement to obtain formal consent if this required a frightening account of the risks associated with the condition from which the patient suffered. "Does the doctor," he wrote, "invariably seek the patient's consent before using a new drug alleged to be efficacious and safe? If the answer is No, then what process, one may ask, makes it needful for him to do so if he chooses to test the drug in such a way that he can compare its effects with those of the previous orthodox treatment?"<sup>27</sup> The question he asked might be answered differently today, but the principle with which he was concerned, that there should not be "double standards" for the ethics of therapeutic trials and of routine medical care, is still valid.<sup>26, 27</sup>

In the event, the ethics of the trial were not challenged, there was no problem of compliance, and the effect of the drug proved to be so great that even though supplies were sufficient for the treatment of only about 50 patients a clear reduction in fatality emerged after the patients had been observed for six months (table II,  $p < 0.01$ ).<sup>24</sup> Streptomycin was accepted as the standard treatment for tuberculosis and the MRC's committee was able to turn its attention to finding out, by further randomised trials, how to overcome the development of bacterial resistance.<sup>18</sup>

## Two logical developments

Both the production of overviews and the expansion of trials to include many more patients are natural developments of Tony's randomised trial, and both are now being widely adopted.

Compiling all the evidence and reaching a conclusion on the basis of the totality of the results would seem such an obvious approach that it is difficult to see how anyone could object to it—and, in fact, no one did when published work was examined by "experts" who then wrote "authoritative reviews" of the material. The trouble was that the opinions of the experts were largely subjective and dependent on the school to which they belonged; often they disagreed. The novelty of the modern overview is that the review should be undertaken according to strict statistical criteria. Reviews of this sort have met with opposition from some clinicians who are reluctant to surrender their clinical judgment to what they consider to be a mindless juggling with figures. Carried out crudely without an understanding of the nature of the individual trials, overviews might justify such a description, but the scientific overview that has been pioneered by Peto is a different procedure. It requires intensive inquiry to ensure not only that all properly conducted trials have been included, some of which may not have been published, but also that all the trials have been conducted and the results reported in appropriate ways. As the criteria adopted for admitting patients to each trial need not be identical—each trial is treated as an independent entity—no claim is made about the precise quantitative effect of treatment in any subsequently selected group. The method shows whether the totality of results can be attributed to chance, and if not whether the overall effect is beneficial or harmful. The quantitative result provides an indication only of the sort of effect that might reasonably be anticipated with broadly similar groups of patients.

The method has been used extensively in the past few years and has recorded some brilliant successes. It has, for example, shown the value of prescribing tamoxifen

following the surgical treatment of breast cancer.<sup>43</sup> Altogether 29 700 women have been studied in 41 randomised trials. Half received tamoxifen, the groups being otherwise treated similarly by surgery, radiotherapy, or cytotoxic chemotherapy. Few of the individual trials gave significant results, but the overview showed clear benefit, irrespective of the addition of cytotoxic therapy. Improvement was greater with longer treatment: the odds of death for women treated for less than two years were reduced by 11%; for an average of two years, by 19%; for longer than two years, by 22%. The reduction, moreover, has increased with the passage of time up to at least 10 years from the start of treatment, and only an ultrapessimist would not expect the benefit to continue and be somewhat greater still if treatment was more prolonged.

Overviews have certainly come to stay; but, valuable though they are, they are not the ideal solution; it may take years before sufficient well conducted trials have been carried out for a definite conclusion to be reached, and the alternative of a trial that is large enough to provide a clear answer on its own is certainly preferable. For a common condition even moderate improvements in outcome may save thousands of premature deaths a year, but a trial to show such an effect may need 10 000 or 20 000 patients if a convincing demonstration of benefit is to be obtained. Tony could never have envisaged the conduct of such a trial, as in his time it was only with the greatest difficulty that a dozen specialists could be persuaded to randomise patients and adopt a common treatment schedule. Trials like ISIS-2, which recruited over 17 000 patients and showed that streptokinase and aspirin could each reduce deaths from myocardial infarction by a fifth—and in combination reduce it by two-fifths—have been made possible only by building on the precepts that Tony taught: namely, to keep the trial simple, cause the participating clinician as little extra work as possible, and allow the clinician maximum freedom in the use of treatments other than those under test.

TABLE II—Controlled trial of streptomycin in the treatment of pulmonary tuberculosis.<sup>28</sup> Values are numbers (percentages) of patients followed for six months

Treatment	Condition on radiological examination				
	Dead	Deteriorated	Unchanged	Slightly improved	Considerably improved
Bed rest (n=52)	14 (27)	18 (34)	3 (6)	13 (25)	4 (8)
Bed rest plus 2 g streptomycin 4 hourly (n=55)	4 (7)	10 (18)	3 (6)	10 (18)	28 (51)

#### SUBSEQUENT DEVELOPMENTS

The success of the trials of treatment for pulmonary tuberculosis and of the subsequent trials of antihistamines for the prevention and treatment of the common cold,<sup>29</sup> of cortisone and aspirin for the treatment of rheumatoid arthritis,<sup>30,31</sup> and of adrenocorticotrophic hormone, cortisone, and aspirin in the Anglo-American-Canadian trial of the treatment of acute rheumatic fever,<sup>32</sup> all of which followed in a few years under Tony's guidance, established randomisation as the method of choice for testing the efficacy of treatments. Its use spread throughout English speaking countries, the Netherlands, and Scandinavia. Individual clinicians protested that controlled trials failed to reproduce the benefits that they had recorded in their own series because, they claimed, the trialists treated a wide range of patients with a particular disease instead of selecting those who would respond to the treatment. It gradually sank in, however, that this objection required the protesters to characterise the responsive patients with sufficient clarity for them to be put in a separate stratum before randomisation. The protests gradually faded, and randomisation spread throughout the developed world.

#### Sir Austin, the man

That Tony's teaching bore fruit and the methods he introduced were adopted so quickly was due not only to their intrinsic worth but also to the clarity of his exposition and the temperate way in which he advanced his ideas, for Tony was a quiet, unassuming, private person who sought to lead but not to drive. In committee he expressed his opinion firmly and cogently but never sought to impose it. He was consequently listened to with respect, and his advice was nearly always accepted. In public he sought to avoid controversy. Though distressed by Sir Ronald Fisher's attacks on his interpretation of the association between smoking and the incidence of lung cancer—which on one occasion went so far as accusing Tony of scientific dishonesty—he preferred to let the facts speak for themselves rather than embark on public dispute. He took immense trouble over his lectures, which he gave without the use of visual aids, and rehearsed and read so well that his audience often thought that he was speaking without a text. Even those who were not interested in the subject matter were, moreover, kept attentive by the occasional witty aside which, it must be said, might sometimes be considered indelicate by the prudish.

As a department head he had his door always open to any junior who sought his advice, and he saw his job as being to provide the conditions under which his university and research staff could be most productive. The consequence was that no one who worked in his department ever wanted to leave and it was only with the greatest difficulty that they could be prised away to take up professorships elsewhere. I can say without the slightest hesitation that those of us who inherited his empire—Donald Reid who so sadly died shortly afterwards, Peter Armitage, and myself—were all saddened when we learnt of his decision to retire a year early. Our

admiration of Tony as a scientist and teacher was equalled by our fondness for him as a friend.

This article has been adapted from the memorial lecture for Sir Austin Bradford Hill at the London School of Hygiene and Tropical Medicine, 7 May 1992.

- Hill AB. Internal migration and its effects upon the death rates: with special reference to the county of Essex. *Medical Research Council Speco*, 1925. (Report series No 95.)
- Hill AB. *Principles of medical statistics*. London: The Lancet, 1937.
- Mathematics and medicine [editorial]. *Lancet* 1937;ii:31.
- Mathematics and medicine. *Lancet* 1937;ii:31.
- Hill AB, Hill ID. *Bradford Hill's principles of medical statistics*. 12th ed. London: Edward Arnold, 1991.
- Doll R, Hill AB. Smoking and carcinoma of the lung. *BMJ* 1950;ii:739.
- Susser M. Epidemiology in the United States after World War II: the evolution of technique. *Epidemiol Rev* 1985;7:147-77.
- Hill AB. The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* 1965;58:295-300.
- Hill AB. Reflections on the controlled trial. *Ann Rheum Dis* 1966;25:107-13.
- Berkson J. The statistical investigation of smoking and cancer of the lung. *Proceedings of the Mayo Clinic* 1959;34:206.
- Doll R, Hill AB. The mortality of doctors in relation to their smoking habits. A preliminary report. *BMJ* 1954;i:1451-5.
- Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. *BMJ* 1952;ii:1271-86.
- Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *BMJ* 1976;ii:1525-36.
- Doll R, Hill AB. Lung cancer and other causes of death in relation to smoking. *BMJ* 1956;ii:1071-6.
- Smoking and health: a report of the Surgeon General*. Washington DC: US Government Printing Office, 1979.
- Peto R, Lopez AD. Worldwide mortality from current smoking patterns: WHO consultative group on statistical aspects of tobacco-related mortality. In: Durston B, Jamrozik R, eds. *Tobacco and health 1990: the global war*. Perth: Organising Committee of the Seventh World Conference on Tobacco and Health, 1992.
- Medical Research Council Whooping-cough Immunization Committee. The prevention of whooping-cough by vaccination. *BMJ* 1950;ii:1463-71.
- Medical Research Council Tuberculosis Chemotherapy Trials Committee. The treatment of pulmonary tuberculosis with isoniazid. *BMJ* 1952;iii:735-46.
- Armitage P. The role of randomisation in clinical trials. *Stat Med* 1982;1:345-52.
- Fisher RA. The arrangement of field experiments. *Journal of the Ministry of Agriculture* 1926;33:503-13.
- Fisher RA. *The design of experiments*. Edinburgh: Oliver and Boyd, 1935.
- Fibiger J. Om serumbehandling af Difteri. *Hospitalstidende* 1898;6:309-25, 337-50.
- Pearson K. Report on certain enteric fever inoculation statistics. *BMJ* 1904;iii:1243-6.
- Hill AB. Memories of the British Streptomycin Trial in Tuberculosis. *Controlled Clinical Trials* 1990;11:77-9.
- Hill AB. Medical ethics and controlled trials. *BMJ* 1963;i:1043.
- Medical Research Council. Responsibility in investigations on human subjects. In: *Report of the Medical Research Council for the year 1962-3*. London: HMSO, 1964:21-5.
- Collins R, Doll R, Peto R. Ethics of controlled trials. In: Williams CJ, ed. *Introducing new treatments for cancer: practical, ethical, and legal problems*. Chichester: Wiley, 1992:49-65.
- Medical Research Council Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment for pulmonary tuberculosis. *BMJ* 1948;ii:769-82.
- Medical Research Council. Clinical trials of antihistamine drugs in the prevention and treatment of the common cold. *BMJ* 1950;ii:425-9.
- Medical Research Council and Nuffield Foundation. A comparison of cortisone and aspirin in the treatment of early cases of rheumatoid arthritis. *BMJ* 1954;i:1223-7.
- Medical Research Council and Nuffield Foundation. A comparison of cortisone and aspirin in the treatment of early cases of rheumatoid arthritis (2). *BMJ* 1955;iii:695-700.
- Rheumatic Fever Working Party. Treatment of acute rheumatic fever in children: a cooperative clinical trial of ACTH, cortisone, and aspirin. *BMJ* 1955;ii:555-74.
- Armitage P. *Sequential medical trials*. Oxford: Blackwell Scientific, 1960.
- Cochran WG. Some methods for strengthening the common  $\chi^2$  tests. *Biometrics* 1954;10:417-51.
- Cox DR. Some procedures connected with the quantitative response curve. In: David FN, ed. *Festschrift for J. Neyman, Research Papers in Statistics*. New York: Wiley, 1966:55-71.
- Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society* 1972;B34:187-219.
- Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports* 1966;50:163-70.
- Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719-48.
- Peto R, Peto J. Asymptomatically efficient rank invariant test procedures (with discussion). *Journal of the Royal Statistical Society* 1972;A135:185-207.
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976;34:585-612.
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer* 1977;35:1-39.
- Gehan EA, Freireich EJ. Non-randomized controls in cancer clinical trials. *N Engl J Med* 1974;290:198-203.
- Early Breast Cancer Trialists Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy: 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet* 1992;339:1-85.
- ISIS-2 Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction. *Lancet* 1988;ii:349-60.