Jerome Cornfield's Bayesian approach to assessing interim results in clinical trials

James J Schlesselman
Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA 15261, USA
Corresponding author: James J Schlesselman. Email: jjs66@pitt.edu

Introduction

Jerome Cornfield (1912–1979) was a man of philosophical bent, engaging wit, deep thought, great mathematical talent, and formidable skill in written and oral debate.[1–7] He had extensive influence on biostatistics and medical research in the USA in the middle of the 20th century.

Cornfield was never awarded a university degree beyond a Bachelor of Science. Indeed, in the words of one distinguished colleague, 'he represented a mockery of excessive adherence to traditional qualifications'.[8] Nevertheless, Cornfield was elected president of the American Statistical Association, the American Epidemiologic Society and the International Biometric Society's Eastern North American Region, and he was a fellow of the Institute of Mathematical Statistics, the American Statistical Association and the American Association for the Advancement of Science.[9,10]

While at the National Cancer Institute in the 1950s, Cornfield developed statistical methods for laboratory studies and epidemiologic investigations. His writing on the nature of causation and on evidence that may be used to buttress an inference of cause-and-effect arose from his involvement in one of the great public health controversies of his time: lung cancer in relation to cigarette smoking.[1,11]

Among his many contributions to epidemiology, Cornfield defended case–control studies as an appropriate method to assess the potential effects of an exposure on the risk of disease; developed the odds ratio, based on a case–control study, as an approximation to the corresponding estimate of relative risk based on a cohort study; and developed the rationale for the use of relative risk, as opposed to absolute risk, in studies of disease etiology. He also gave a persuasive rationale for the use of observational studies in scientific inference.[2]

Shortly after his death, the scope of Cornfield's wide-ranging research activities was surveyed in a series of articles written by persons who knew him well. They discussed his contributions to laboratory research,[12] epidemiologic studies,[13] statistical theory[14] and clinical trials.[15,16]

Cornfield and clinical trials

> Despite the ambiguities involved in design, in decision making and in conclusion reaching, it is undeniable that the clinical trial has constituted an important contribution to medicine.... [this] surely can be explained as simply a further triumph of experimental method as applied to clinical medicine. As such, it would not have surprised even Claude Bernard, and only the statistical participation would have puzzled him.[6(p 420)]

The last major expression of Cornfield's thinking about clinical trials appeared in his 1976 article, 'Recent methodological contributions to clinical trials'. It covered six areas: decision-making; appraisal of uncertainty; likelihood ratios; the use of prior opinion in statistical analyses (Bayesian methods); patient subgroups (multiple comparisons); and randomisation. Cornfield began his discussion with three broad assertions: statistical methods could never provide unique, unequivocal answers to problems of design and analysis; the process of inference and decision-making in clinical trials is loosely structured, because that is the nature of an intrinsically complex enterprise; and analyses of data from clinical trials may inevitably lead to ambiguous answers.

Cornfield's skepticism about the role of statistical methods in decision-making and scientific inference had been reinforced by his recent experience with the interim monitoring of several large-scale, multicentre clinical trials. His reservations, however, had been voiced much earlier in a more general context in 1959, in his carefully reasoned, thought-provoking overview of research methodology: 'Principles of research'.[2]

Cornfield had long advocated that clinical trials be randomised whenever feasible and ethical. He also urged that individuals be randomised if possible, rather than larger units, such as clinics or hospitals. He had stated the frequentist rationale for randomisation in 1959:

> The device of randomization did two things. First, it controlled the probability that the treated and the control group differed by more than a calculable amount in their exposure to disease, in their immune history, or with respect to any other variable, known or unknown to the experimenters, which might have a bearing on the outcome of their trial. Furthermore, as the size of the two groups being compared increased, it assured that the probability that they differ by more than this amount approached zero.
>
> The second thing that randomization made possible was an objective answer to the question that must be asked at the conclusion of any trial: In how many experiments could a difference of this magnitude have arisen by chance alone if the treatment truly has no effect?[2(p 245)]

In his 1976 paper, Cornfield wrote:

> One of the finest fruits of the Fisherian revolution was the idea of randomization, and statisticians who agree on few other things have at least agreed on this. But despite this agreement and despite the widespread use of randomized allocation procedures in clinical and in other forms of experimentation, its logical status, i.e. the exact function it performs, is still obscure. Does it provide the only basis by which a valid comparison can be achieved or is it simply an ad hoc device to achieve comparability between treatment groups?[6(p 418)]

Although randomisation's 'logical status' and 'the exact function it performs' may have become obscure upon deep reflection, Cornfield nonetheless answered both of the questions posed: 'no', to the first, and essentially 'yes,' to the second. However, although Cornfield was a longstanding advocate of randomised clinical trials, he believed that observational studies could also provide a secure foundation for scientific inference and decisions. His defence of observational studies in 1959, accompanied by his earlier writing on this topic, are reminiscent of Bradford Hill's 1953 article, *Observation and Experiment*, and Hill's celebrated exposition of guidelines for assessing cause-and-effect from epidemiologic studies.[17,18]

Decision-making and appraisal of uncertainty in monitoring clinical trials

Cornfield's early work in statistics reflected his training in frequentist methods, for which 'probability' is a measure of the relative frequency with which events occur in the natural world. This interpretation differs from that of Bayesians, for whom 'probability' is a measure of a person's degree of belief in a proposition, such as 'the risk of lung cancer is increased 20-fold in persons who smoke 2 packs of cigarettes per day'. Both conceptions of probability have a long history of use and application,[19,20] and both conform to the same axioms and theorems of probability theory. In addition to what is meant by probability, the major difference in the Bayesian approach is that it imposes two requirements: first, one must specify a probability distribution (of belief) for values of the parameters involved in a statistical model to be used for an analysis; and second, when the study data become available, one must apply Bayes' theorem to update the prior probability distribution. This second requirement, called *Bayes rule*, is a mathematical consequence of Bayesians' insistence that any rational expression of degree of belief (personal probability) must be internally consistent.[21(pp 36–37, 64–66)]

In the mid-1960s, Cornfield repudiated frequentist methods for monitoring clinical trials and expressed his conversion to a then radical point of view, namely, that tests of statistical significance and *p*-values, due to RA Fisher,[22] tests of statistical hypotheses and confidence intervals, due to Jerzy Neyman and Egon Pearson,[23,24] and sequential analysis, due to Abraham Wald and George Barnard[25,26] – virtually everything that frequentists held dear – were seriously misguided.[27–31] Such thoughts had essentially been stated earlier by Jimmy Savage among others,[32,33] but it was Francis Anscombe's review of Peter Armitage's influential text, *Sequential Medical Trials*, which really garnered attention.[34,35]

Anscombe's carefully reasoned, trenchant criticism of sequential analysis and Neyman-Pearson theory of hypothesis tests was so important that senior statisticians at the US National Institutes of Health held 'an informal seminar' about the issue in June 1965.[36] On that occasion, Cornfield argued for a Bayesian approach to monitoring clinical trials. Although he subsequently continued to use the frequentist concepts of 'power' and 'level of statistical significance' to plan clinical trials,[37] Cornfield objected to their relevance in the interim analyses of a study's emerging data, and by implication in the final analysis as well.

Cornfield and the University Group Diabetes Program

The University Group Diabetes Program saga

> The UGDP was meant to be a model of clinical investigation. Planned and managed by a team of statisticians and clinicians, the study addressed a series of long-standing controversies over the clinical management of diabetes. Intended to demonstrate how a properly designed, randomized, controlled trial could resolve differences of clinical opinion, the UGDP instead became a symbol of all that was wrong with the statistical enterprise in medicine. Few recent controversies in medicine are comparable in length and rancor to that over the UGDP.[38(p 198)]

In his 1976 article, Cornfield referred to three clinical trials as examples of decisions that had to be made in the face of major, unexpected problems: the University Group Diabetes Program;[39,40] the Coronary Drug Project;[41,42] and the Diabetic Retinopathy Study.[43]

Because of the importance of the University Group Diabetes Program to the development of modern-day clinical trials,[38,44,45] I use that study to explain Cornfield's Bayesian method – relative betting odds – which was used as one of three statistical techniques to support the decision to discontinue tolbutamide. The University Group Diabetes Program also provides an instructive instance in which Cornfield's formidable skills in rhetoric, statistical analysis and logical thinking are clearly displayed in undermining the arguments of a study's critics.[4] It additionally provides a striking example in which Cornfield's Bayesian rationale for avoiding the use of p-values and tests of statistical significance were contradicted by his practice.

Cornfield was not involved in the University Group Diabetes Program's design, and he was not part of the research team initially. Nearly six years after the first patient had been enrolled and the study was fully in progress, his advice was sought by Christian Klimt, the head of the University Group Diabetes Program's Statistical Center. At that time, Klimt was vexed by a major problem: informal, interim analyses indicated that more cardiovascular deaths were occurring in patients treated with tolbutamide than with placebo – a completely unexpected, inexplicable finding. Bayesian analyses by Cornfield and computer simulations for frequentist-based analyses were therefore developed to assess the unfavourable trend, which also indicated that overall mortality on tolbutamide was no better than that on placebo and was possibly worse. On the basis of these analyses, Cornfield advocated that treatment with tolbutamide be stopped, and he remained the major defender of the statistical basis of that recommendation.[4,46,47]

With Bayesian and frequentist analyses in hand, a two-day meeting of the University Group Diabetes Program's Executive Steering Committee voted in June 1969 to drop tolbutamide as a study treatment, and duly informed the US Food and Drug Administration of their decision. Despite skepticism by some reviewers in the agency and by its external advisors, the US Food and Drug Administration announced its intention in May 1970 to place a warning of increased cardiovascular hazard on the label for tolbutamide and all chemically related agents (sulfonylurea drugs). This was done before any University Group Diabetes Program publication, and it set in motion a long-ensuing clinical and scientific controversy.

Within a few months of the University Group Diabetes Program's first two publications,[39,40] several articles highly critical of the study's design, implementation and statistical analyses appeared in the medical literature.[48–51] Those articles were followed shortly thereafter by further criticism of the University Group Diabetes Program,[52,53] which continued throughout the 1970s.[54–56]

Cornfield was a representative of the University Group Diabetes Program in discussions with the US Food and Drug Administration, and he testified before a US Senate subcommittee in a hearing held over three days in September 1974 about the controversy.[46,57] Amazingly, on the last day of that hearing, Cornfield was allowed to cross-examine an articulate, forceful opponent of the University Group Diabetes Program, Holbrooke Seltzer, who was Professor of Internal Medicine at the University of Texas and a member of the Board of Directors of the American Diabetes Association.[47,58] Furthermore, at the request of Senator Gaylord Nelson, Chairman of the Subcommittee on Monopoly, Cornfield wrote replies to numerous interrogatories proffered by the attorney representing a group of diabetologists, including Seltzer, who opposed the actions of the US Food and Drug Administration and challenged the validity of the University Group Diabetes Program's findings.[59] Those challenges, which unsuccessfully sought to obtain the University Group Diabetes Program data, were ultimately decided by the US Supreme Court.[60]

The application of Cornfield's 'relative betting odds'

Although the University Group Diabetes Program was planned and monitored with frequentist methods, it became the first clinical trial in which a

Bayesian analysis was used, what Cornfield called 'relative betting odds', which is a Bayesian method for testing hypotheses.[27,29,33] The term 'betting odds' refers to an index of *personal belief* in a proposition or hypothesis (H_A), such as the following:

> H: the risk of cardiovascular death on tolbutamide is identical to that on placebo.

The strength of one's belief that H is true is explained by way of placing a bet. For someone to give odds of 10:1 in favour of H, before the University Group Diabetes Program data have been inspected, i.e. 'prior odds' that H is true, also called 'prior odds of H', means that one is prepared to bet 10 units that H is true in order to receive only 1 unit if H is shown to be false.

In Cornfield's application to the University Group Diabetes Program, betting odds on H were often expressed relative to a specific alternative hypothesis (H_A),

> H_A: compared to placebo, the risk of cardiovascular death on tolbutamide is increased by 25%.

Figure 1 shows that as data from the University Group Diabetes Program emerged, they began to indicate that H is false.

In the 204 patients on tolbutamide, there were 30 deaths in total by the eighth year of follow-up, 26 of which were attributed to cardiovascular causes. By comparison, there were only 21 deaths in the 205 patients on placebo, 10 of which were attributed to cardiovascular causes. The life-table estimate of the cumulative risk of cardiovascular death (intent-to-treat analysis) was 17.6 (standard error 2.5) per 100 on tolbutamide versus 6.0 (standard error 2.5) per 100 on placebo.

In view of such data, a person's 'posterior odds' in favour of H being true, i.e. one's belief that 'the risk of cardiovascular death on tolbutamide is identical to that on placebo', should be smaller than his 'prior odds'. The amount of decrease is given by Bayes theorem. That amount is what Cornfield called the 'relative betting odds'. It expresses the degree to which the acquisition of data should change one's 'prior odds' to 'posterior odds'. Expressed schematically:[61]

$$\text{posterior odds of H} = \text{relative betting odds} \times \text{prior odds of H}.$$

Values of relative betting odds less than 1 should diminish one's posterior odds that H is true, whereas values of relative betting odds greater than 1 should enlarge them.

Figure 2 displays some of the University Group Diabetes Program's many reported values of relative betting odds.[40] The alternative hypothesis (H_A) in Figure 2 is that 'the cumulative risk of cardiovascular death is increased by 25% over the cumulative risk on placebo'.

One sees from Figure 2 that the University Group Diabetes Program's accumulating data led to progressively decreasing values of the relative betting odds. By the 8th year of follow-up, the relative betting odds were 0.15. In other words, at that time a person's

Figure 1. Cumulative mortality rates per 100 persons at risk, by year of follow-up.[40] TOLB: tolbutamide; PLBO: placebo.
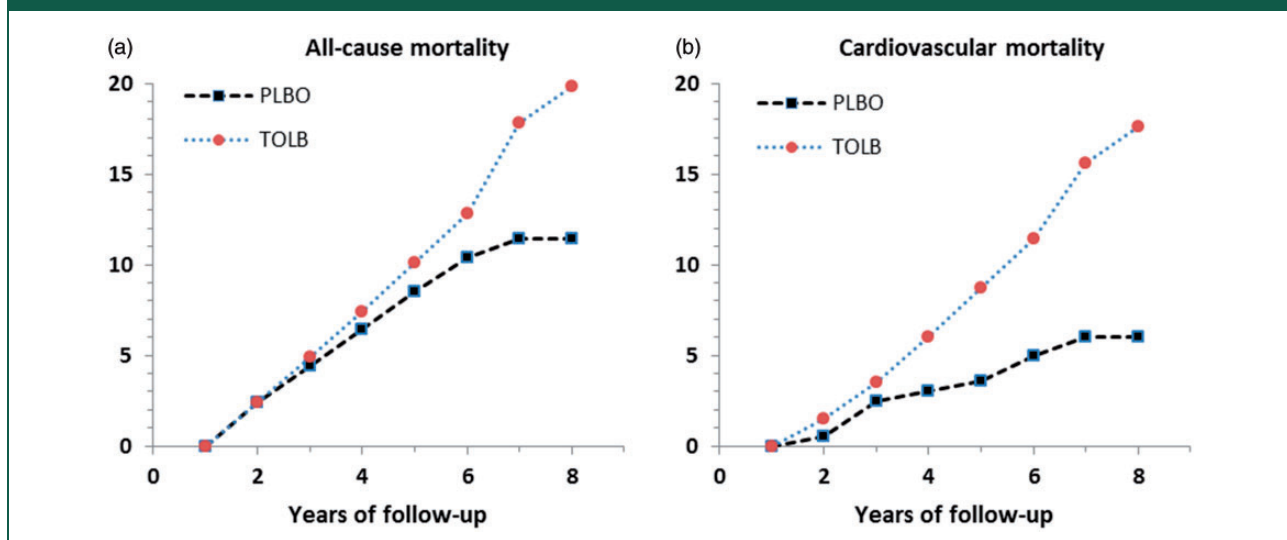
Figure 2. Relative betting odds for the difference in cumulative cardiovascular mortality (tolbutamide − placebo), by year of follow-up.

	1	2	3	4	5	6	7	8
RBO	1.00	1.00	0.91	0.87	0.52	0.39	0.21	0.15

prior odds in favour of H, i.e. that 'the risk of cardiovascular death on tolbutamide is identical to that on placebo', should have been diminished by 85%.

Relative betting odds depend not only on the data but also on the degree of belief in the alternative hypotheses against which the null hypothesis is tested.[27,62,63] Further complications arise in calculating relative betting odds for 'composite hypotheses', such as 'the risk of cardiovascular death on tolbutamide is within ±5% of the risk on placebo', or 'the risk of cardiovascular death on tolbutamide is at least 25% greater than the risk on placebo'.[29,33]

Cornfield's defence of the University Group Diabetes Program

> . . . an investigation originally designed to produce new knowledge, suddenly found itself involved in a difficult and unwanted task of decision making. From the purely formal hypothesis testing point of view that dominated the early thinking in clinical trials, what had happened was that the same body of data had been used to formulate a hypothesis and to test it. From that point of view the University Group Diabetes Program results should have been treated as suggesting a hypothesis to be tested in a new and independent trial. But to the investigators this was inappropriate. They had to decide for themselves and their patients whether the evidence available to them justified the future exposure of anyone to these agents . . . [6(p 409)]

At the outset of his rebuttal of the University Group Diabetes Program's critics, Cornfield stated that the study's prudent decision and moderately worded conclusion had been:

> . . . received by some critics with a hostility which has no discernible scientific basis. . . . The subsequent analysis is undertaken to illuminate these alternatives [i.e., independent repetition of the UGDP vs. acceptance of its findings] and not to defend the UGDP. Its

concentration on the strength of the evidence against tolbutamide should of course not be permitted to obscure the more general UGDP finding that lowering of blood glucose level did not appreciably lower the eight-year mortality from cardiovascular disease as compared with patients on diet alone.[4(p 1676)]

Cornfield's claim that his purpose was 'not to defend the UGDP' is at odds with what he did, namely, rebut every statistical criticism that had been levelled at the study. Furthermore, his remark that 'the more general UGDP finding that lowering of blood glucose level did not appreciably lower the eight-year mortality from cardiovascular disease as compared with patients on diet alone' was based on a serendipitous result: the University Group Diabetes Program was not designed to address this issue. Nonetheless, by insightful analysis using standard statistical methods and ruthless logic, Cornfield demonstrated that no factual basis supported the critics' concerns that randomisation had somehow broken down and produced major baseline inequalities which nullified the results; that excess mortality on tolbutamide was confined to a small number of clinics, and because of this the University Group Diabetes Program's findings could not be generalised to medical practice; and that dropouts, non-adherence to treatment, and the lower-than-expected mortality in the placebo group undermined the University Group Diabetes Program's findings concerning tolbutamide.

Cornfield also rebutted a number of clinical concerns, including those expressed about the eligibility criteria for study patients, the use of a fixed dose of tolbutamide, the determination of the principal cause of death, the definitions of baseline risk factors, the failure to obtain data on patients' smoking histories, and the decision to stop treatment with tolbutamide before a more conclusive demonstration was available for its apparently harmful effect.

Despite Cornfield's advocacy of Bayesian methods in clinical trials and the use of his relative betting odds to support the University Group Diabetes

Program's decision to terminate treatment with tolbutamide, his rebuttal relied extensively on the very frequentist methods that he was arguing against: p-values and tests of statistical significance as a basis of support for or against hypotheses. Bayesian methods were conspicuous by their absence.

What can one say about the discrepancy between Cornfield's preaching against p-values and his practice? One answer is that he chose to use well-accepted frequentist methodology in defending the University Group Diabetes Program because it suited his purpose, was not misleading, and because framing a rebuttal through a series of Bayesian analyses would have been incomprehensible to and likely rejected by the vast majority of readers. A related answer is that a Bayesian rebuttal would have required Cornfield to specify prior distributions of belief in the issues he was discussing. One specification could have reflected his personal beliefs, or that of others defending the University Group Diabetes Program, but such prior distributions had never been expressed before the University Group Diabetes Program data were examined. Another tactic could have involved the application of a 'neutral prior', that is, a prior distribution of belief that represented weakly held opinion about the issues involved. This would have led to posterior distributions that approximate the likelihood function,[64] with results similar to those based on p-values. None of the analyses, however, would have satisfied the University Group Diabetes Program's fiercest critics, who were alleging that the study was a seriously flawed investigation beyond the remediation of statistics. The only choice of prior distribution that would have satisfied them would have been one that led to accepting their position.

As emphasised by Lindley,[65(p 313)] Bayesian methods are deficient in situations involving conflict, a circumstance that embroiled the University Group Diabetes Program both internally (some of the study investigators opposed the decision to stop treatment with tolbutamide) and externally (for various reasons many clinicians not involved in the University Group Diabetes Program did not believe its results). Although the University Group Diabetes Program was challenged by some critics because the study had stopped treatment with tolbutamide 'too soon', i.e. before more data on mortality were at hand, it is doubtful in retrospect whether more data from the same study would have ever settled the controversy.

Cornfield's legacy

It is sometimes claimed, or at least implied, by frequentists that, in contrast to Bayesians, they seek rules which minimize the long-run frequency of errors of inference or decision. If Bayesian procedures really lacked this property, then it would be difficult to accept them or to make them the basis of any scientifically defensible system of data analysis or behavior. We shall here argue that no such contradiction exists.[30(p 15)]

Cornfield's retirement from the US Civil Service in 1967 left the NIH with no influential advocate of Bayesian theory, which had been subject to some highly contentious argument.[32,66–71] Only two other clinical trials, the Coronary Drug Project[41,42,72] and the Urokinase Pulmonary Embolism Trial,[62,73] used Cornfield's relative betting odds (RBO) methodology, which is now largely abandoned for interim monitoring.[63(p 340)] Bayesian methods currently use the posterior distribution of the parameter of interest, such as the relative risk or the difference in risk. To limit the number of different estimates that can arise from a given set of data, some recommend that Bayesian calculations be made under three broadly different assumptions: a neutral prior, a skeptical prior (against the alternative hypothesis of interest) and a prior that favours the alternative hypothesis,[74,75] which is similar to a proposal made by Cornfield and Greenhouse.[31(pp 823–825)]

Only a few of Cornfield's articles on clinical trials are now cited in articles and texts. His evocative phrase 'RBO', which so effectively conjures the image of a casino, has been supplanted sadly by the non-descript 'Bayes factor', a term coined by IJ Good.[76–78] Despite Cornfield's key contribution to the interpretation of the University Group Diabetes Program, the investigators did not use the term 'RBO' in their publications, perhaps because of its connotation and the controversy in which they were involved. The University Group Diabetes Program, however, did use 'RBO', which was said to be an abbreviation for 'likelihood ratio', and they cited Cornfield's article,[29] but without mentioning that relative betting odd was his acronym for 'RBO'.

Even if one subscribes to the use of Bayes' rule for modifying prior belief, in practical applications such as the University Group Diabetes Program, making decisions and taking action will involve additional considerations: assessing the design of a study versus its implementation, and weighing costs, benefits and ethical concerns, all of which can be subject to widely differing opinions.[79] With regard to ethics, Anscombe's[35] proposed Bayesian method to decide whether to continue recruitment to a clinical trial when accumulating data suggest that one of the treatments being compared is superior failed to account for physicians' patient-centred

perspective,[79,80] a viewpoint which Cornfield evidently shared (see quotation at the start of 'Cornfield's defence of the University Group Diabetes Program' section).

Cornfield became a staunch advocate of Bayesian methods, but he was never an ideologue on their behalf. His thinking and writing about 'the Bayesian outlook' from 1966 forward was accompanied by his contemporaneous use of frequentist techniques, *p*-values among them,[4] which he claimed to deplore in principle. The resultant inconsistency between this theorising and practice of a supremely logical man might be explained by Cornfield being not only a scientific pragmatist, i.e. using methods that were accepted by the large majority clinicians and scientists, but also a skeptic, which was expressed in his 1976 article, '*Recent methodological contributions to clinical trials*':[6] there can never be a completely reliable foundation of scientific inference or decision-making in the face of uncertainty, and statistical methods alone, Bayesian or otherwise, are incapable of solving this problem.[79]

If the above remarks are true, then why did Cornfield convert to and espouse Bayesian methods? One answer, offered by Cornfield, is that frequentist methods not only lead to logical conundrums, but they are also inflexible in practice, especially in the very instances where flexibility is scientifically most needed.[28–30,36(pp 862–866, 877–881)] In adopting that position, he was apparently willing to overlook problems with Bayesian theory,[64(pp 167–176),81] arguing that the advantage to be gained by improved flexibility was not accompanied by any important loss from abandoning frequentist techniques.

Cornfield's paper on recent methodological contributions to clinical trials was presented on 30 April 1976 at the Reed-Frost Symposium of the Johns Hopkins University. At the time Cornfield spoke, he did not realise that he would soon cross a threshold and pass into history. One year earlier, however, he had mused about the future:

> ...what is statistics and where is it going? When one tries, spider-like, to spin such a thread out of his viscera, he must say, first, according to Dean Acheson, 'What do I know, or think I know, from my own experience and not by literary osmosis?' An honest answer would be, 'Not much; and I am not too sure of most of it'.[5]

While Cornfield's contributions to clinical trials may be nearly forgotten, the Bayesian perspective that he advocated is becoming more widely accepted,[74,75,82,83] a development that would have pleased him.

References

1. Cornfield J. Statistical relationships and proof in medicine. *Am Stat* 1954; 8: 19–21.
2. Cornfield J. Principles of research. *Am J Ment Defic* 1959; 64: 240–252.
3. Cornfield J. Discussion by J. Cornfield, B.M. Jill, D.V. Lindley, S. Geisser, and C.M. Mallows. In: Meyer DL and Collier RO (eds) *Bayesian Statistics*. Itasca, IL: Peacock Publishers Inc., 1970, pp.85–125.
4. Cornfield J. The University Group Diabetes Program. A further statistical analysis of the mortality findings. *J Am Med Assoc* 1971; 217: 1676–1687.
5. Cornfield J. A statistician's apology. *J Am Stat Assoc* 1975; 70: 7–14.
6. Cornfield J. Recent methodological contributions to clinical trials. *Am J Epidemiol* 1976; 104: 408–421.
7. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol* 1978; 108: 100–102.
8. Frederickson DS. Remarks. *Biometrics* 1982; 38(Suppl): 7.
9. Greenhouse SW and Halperin M. Jerome Cornfield (1912–1979). *Am Stat* 1980; 34: 106–107.
10. Greenhouse SW. A tribute. *Biometrics* 1982; 38(Suppl): 3–6.
11. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB and Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst* 1959; 22: 173–203.
12. Mantel N. Jerome Cornfield and statistical applications to laboratory research: a personal reminiscence. *Biometrics* 1982; 38(Suppl): 17–23.
13. Greenhouse SW. Jerome Cornfield's contributions to epidemiology. *Biometrics* 1982; 38(Suppl): 33–45.
14. Zelen M. The contributions of Jerome Cornfield to the theory of statistics. *Biometrics* 1982; 38(Suppl): 11–15.
15. Ederer F. Jerome Cornfield's contributions to the conduct of clinical trials. *Biometrics* 1982; 38(Suppl): 25–32.
16. Green SB. A conversation with Fred Ederer. *Stat Sci* 1997; 12: 125–131.
17. Hill AB. Observation and experiment. *New Engl J Med* 1953; 248: 995–1001.
18. Hill AB. The environment and disease: association or causation? *Proc Royal Soc Med* 1965; 58: 295–300.

19. Fienberg SE. A brief history of statistics in three and one-half chapters: a review essay. *Stat Sci* 1992; 7: 208–225.
20. Fienberg SE. When did Bayesian inference become "Bayesian"? *Bayesian Anal* 2006; 1: 1–40.
21. Lindley DV. *Understanding Uncertainty*. New York: Wiley, 2006.
22. Fisher RA. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, 1925.
23. Neyman J and Pearson ES. On the problem of the most efficient tests of statistical hypotheses. *Philos Trans Royal Soc Lond Ser A* 1933; 231: 289–337.
24. Neyman J and Pearson ES. The testing of statistical hypotheses in relation to probabilities a priori. *Proc Cambridge Philos Soc* 1933; 29: 492–510.
25. Wald A. Sequential tests of statistical hypotheses. *Ann Math Stat* 1945; 16: 117–186.
26. Barnard GA. Sequential tests in industrial statistics. *J Royal Stat Soc (Suppl)* 1946; 8: 1–21.
27. Cornfield J. A Bayesian test of some classical hypotheses, with applications to sequential clinical trials. *J Am Stat Assoc* 1966; 61: 577–594.
28. Cornfield J. Sequential trials, sequential analysis and the likelihood principle. *Am Stat* 1966; 20: 18–23.
29. Cornfield J. The Bayesian outlook and its applications (with discussion). *Biometrics* 1969; 25: 617–657.
30. Cornfield J. The frequency theory of probability, Bayes' theorem, and sequential clinical trials. In: Meyer DL and Collier RO (eds) *Bayesian Statistics*. Itasca, IL: Peacock Publishers Inc., 1970, pp.1–28.
31. Cornfield J and Greenhouse SW. On certain aspects of sequential clinical trials. In: Neyman J and LeCam LM (eds) *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4. Berkeley, CA: University of California Press, 1967, pp.813–829.
32. Savage LJ. The foundations of statistics reconsidered. In: Neyman J (ed.) *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1. Berkeley, CA: University of California Press, 1961, pp.575–586.
33. Edwards W, Lindman H and Savage LJ. Bayesian statistical inference for psychological research. *Psychol Rev* 1963; 70: 193–242.
34. Armitage P. *Sequential Medical Trials*. Oxford: Blackwell, 1960.
35. Anscombe FJ. Sequential medical trials. *J Am Stat Assoc* 1963; 58: 365–384.
36. Cutler SJ, Greenhouse SW, Cornfield J and Schneiderman MA. The role of hypothesis testing in clinical trials. *J Chron Dis* 1966; 19: 857–882.
37. Cornfield J. Fixed and floating sample size trials. In Engle RL, Jr. (Symposium Chairman). *Symposium on Statistical Aspects of Protocol Design*. Bethesda, MD: Clinical Investigation Review Committee, Clinical Investigations Branch, National Cancer Institute, National Institutes of Health, 1970:181–187, 197–204 (discussion).
38. Marks HM. *The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900–1990*. Edinburgh: Cambridge University Press, 1997.
39. University Group Diabetes Program Research Group. A study of the effects of hypoglycemic agents on vascular complications in patients with adult onset diabetes: I. Design, methods, and baseline characteristics. *Diabetes* 1970; 19(Suppl 2): 747–783.
40. University Group Diabetes Program Research Group. A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: II. Mortality results. *Diabetes* 1970; 19(Suppl 2): 785–830.
41. Coronary Drug Project Research Group. The Coronary Drug Project: initial findings leading to modifications of its research protocol. *JAMA* 1970; 214: 1303–1313.
42. Coronary Drug Project Research Group. The Coronary Drug Project: findings leading to discontinuation of the 2.5-mg/day estrogen group. *JAMA* 1973; 226: 652–657.
43. The Diabetic Retinopathy Study Research Group. Preliminary report on effects of photocoagulation therapy. *Am J Ophthalmol* 1976; 81: 383–396.
44. Meinert CL. *Clinical Trials: Design, Conduct, and Analysis*, 2nd ed. New York, NY: Oxford University Press, 2012.
45. Greene JA. *Prescribing by Numbers*. Baltimore, MD: The Johns Hopkins University Press, 2007.
46. Cornfield J. Statement of Dr. Jerome Cornfield, Chairman, Department of Statistics, The George Washington University, Washington, DC. In *Subcommittee on Monopoly* 1974: 10778–10794.
47. Cornfield J. Interrogation of Holbrooke S. Seltzer, M.D. In: *Subcommittee on Monopoly* 1974: 10889–10895.
48. Feinstein AR. Clinical biostatistics VIII. An analytic appraisal of the University Group Diabetes Program (UGDP) study. *Clin Pharmacol Ther* 1971; 12: 167–191.
49. Leibel B. An analysis of the University Group Diabetes Study Program: data results and conclusions. *Can Med Assoc J* 1971; 105: 292–294.
50. Salsburg DS. The UGDP study. *J Am Med Assoc* 1971; 218: 1704–1705.
51. Schor S. The University Group Diabetes Program. A statistician looks at the mortality results. *J Am Med Assoc* 1971; 217: 1671–1675.
52. Seltzer HS. A summary of criticisms of the findings and conclusions of the University Group Diabetes Program (UGDP). *Diabetes* 1972; 21: 976–979.
53. Schor SS. Statistical problems in clinical trials: the UGDP study revisited. *Am J Med* 1973; 55: 727–732.
54. Feinstein AR. Clinical biostatistics XXXV. The persistent clinical failures and fallacies of the UGDP study. *Clin Pharmacol Ther* 1976; 19: 78–93.
55. Feinstein AR. Clinical biostatistics XXXVI. The persistent biometric problems of the UGDP study. *Clin Pharmacol Ther* 1976; 19: 742–785.

56. Feinstein AR. How good is the statistical evidence against oral hypoglycemic agents? *Adv Intern Med* 1979; 24: 71–95.

57. Subcommittee on Monopoly (1974). Hearings on the present status of competition in the pharmaceutical industry, Second Session, Part 25, Oral Hypoglycemic Drugs. Select Committee on Small Business, United States Senate: September 18, 19, and 20, 1974. See http://babel.hathitrust.org/cgi/pt?id=mdp.39015005 125193 (last checked 26 March 2015).

58. Seltzer HS. Statement of Holbrooke S. Seltzer, M.D., Chief of Metabolism at the Veterans' Administration Hospital and Professor of Internal Medicine at The University of Texas, Dallas, Tex. In *Subcommittee on Monopoly*. 1974: 10880–10889.

59. Cornfield J. Correspondence between Senator Gaylord Nelson and Neil L. Chayet, Dr. Jerome Cornfield, Dr. Christian R. Klimt, and Dr. Jeremiah Stamler. In *Subcommittee on Monopoly* 1974: 11507–11523.

60. US Supreme Court. Forsham v. Harris, 445 U.S. 169 (1980). Forsham v. Harris No. 78-1118. Argued October 31, 1979, Decided March 3, 1980, 445 U.S. 169. See http://supreme.justia.com/cases/federal/us/445/169/case.html (last checked 26 March 2015).

61. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. *Ann Intern Med* 1999; 130: 1005–1013.

62. Urokinase Pulmonary Embolism Trial Study Group. The urokinase pulmonary embolism trial. A national cooperative study. *Circulation* 1973; 47(Suppl 2): 1–108.

63. Jennison C and Turnbull BW. *Group Sequential Methods with Applications to Clinical Trials*. New York: Chapman & Hall/CRC, 1999, p.340.

64. Royall R. *Statistical Evidence: A Likelihood Paradigm*. New York, NY: Chapman & Hall, 1997, pp.169–176.

65. Lindley DV. The philosophy of statistics. *J R Stat Soc Ser D Stat* 2000; 49: 293–337.

66. Hartley HO. In Dr. Bayes' consulting room. *Am Stat* 1963; 17(1): 22–24.

67. Kempthorne O. Discussion of "The Bayesian outlook and its application". *Biometrics* 1969; 25: 647–654.

68. Bross IDJ. Applications of probability: science vs. pseudoscience. *J Am Stat Assoc* 1969; 64: 51–57.

69. Lindley DV. The future of statistics: a Bayesian 21st century. *Adv Appl Prob* 1975; 7(Suppl): 106–115.

70. Lindley DV. Comment on "Why isn't everyone a Bayesian?" *Am Stat* 1986; 40: 6–7.

71. Chernoff H. Comment on "Why isn't everyone a Bayesian?" *Am Stat* 1986; 40: 5–6.

72. Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *New Engl J Med* 1980; 303: 1038–1041.

73. Urokinase Pulmonary Embolism Trial Study Group. Urokinase pulmonary embolism trial. Phase 1 results: a cooperative study. *J Am Med Assoc* 1970; 214: 2163–2172.

74. Spiegelhalter DJ, Abrams KR and Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. New York, NY: Wiley, 2004.

75. Spiegelhalter DJ, Freedman LS and Parmar MKB. Bayesian approaches to randomized trials. *J R Stat Soc Ser A* 1994; 157: 357–416.

76. Good IJ. Significance tests in parallel and in series. *J Am Stat Assoc* 1958; 53: 799–813.

77. Good IJ. A Bayesian significance test for multinomial distributions. *J R Stat Soc Ser B* 1967; 29: 399–431.

78. Good IJ. Letter re Bayes factors: what they are and what they are not. *Am Stat* 1999; 55: 173–174.

79. Armitage P. The evolution of ways of deciding when clinical trials should stop recruiting. *JLL Bulletin: Commentaries on the History of Treatment Evaluation*, 2013. See http://www.jameslindlibrary.org/articles/the-evolution-of-ways-of-deciding-when-clinical-trials-should-stop-recruiting/ (last checked 18 November 2015).

80. Armitage P. Sequential medical trials: some comments on F.J. Anscombe's paper. *J Am Stat Assoc* 1963; 58: 384–387.

81. Efron B. Why isn't everyone a Bayesian? *Am Stat* 1986; 40: 1–5.

82. Berry DA. Bayesian clinical trials. *Nature Rev Drug Discov* 2006; 5: 27–36.

83. Food and Drug Administration. Guidance for industry and FDA staff. Guidance for the use of Bayesian statistics in medical device clinical trials, 2010. See http://www.fda.gov/RegulatoryInformation/Guidances/ucm071072.htm (last checked 26 March 2015).