

**EXPRESS YOUR
OPINION**

Editor's Page

Dear Colleagues,

The upcoming year 2001 applied symposium and the 5th international conference of 2001 are sponsored by the ICSA. We'd like to notify you ahead of time and encourage your involvement by contacting the chairs of the corresponding events or by sharing your research findings with the members. Your suggestions / comments are welcomed. More details inside

IN THIS ISSUE ...

Messages	2
Minutes	3
Reports	7
Professor PL Hsu	14
Special topic–Survey/Poll	.18
Controversial Statistical Issue	
– Bayesian approach in clinical trials	38
Announcements	53
Statistics' Delight / 統計趣聞	59
Members/Activities	61
Advertisement	69
Financial Reports	70
Memberships Form/Info ..	73

**GET INVOLVED IN
NEXT ISSUE**

Special topic – Finance
Statistics' Delight/ 統計趣聞
Controversial statistical
Issue – active controlled
clinical trials

**Get involved in the next
issue by email your articles
to the Editorial Board at
WANGS@CDER.FDA.GOV**

The More The Better ?

I once chatted with people about a report that men have, on average, 2 billion more brain cells than women have. A young man rushed to comment, “So men are smarter than women.” He simply meant, the more the better. A business lady countered, “I think women’s brain cells function more effectively and efficiently.” A doctor took a moment’s breath and said, “Well, I wouldn’t be surprised, because men need more brain cells to coordinate both their larger bodies and extra physical activities; who knows, quantity of brain cells may have nothing to do with intelligence.” A statistician inquired, “Maybe we should define what makes up smartness, a priori.”

In presidential balloting, definition also plays a key role. For instance, the elections in Taiwan earlier this year encountered such a situation. Although Chen Shui-bian won the election by a margin of 2.5% (he received 39.3% of the popular votes; James Soong, his nearest competitor, received 36.8%), he did not receive a majority of the votes. For President Chen and his new government, interpretation of these numbers was more than science and art, it was also reality and compromise. But what is the appropriate way of defining victory? Is it “The more the better?” This question will also prove its importance in the coming November U.S. presidential elections. Despite leading Gore by 14% six months ago, Bush heads by only 2% now. Can Gore turn the numbers around in the end as President Chen did? In this issue, sampling methods used in polls are summarized from the Gallup. An unprecedented phenomenon of Abandon/Secure, statistical method of forecasting the impacts on the undecided voters and prediction modeling are discussed.

Frequentist theory of hypothesis testing has been the norm of standard statistical practice in well-controlled clinical trials. While probability equations may be straightforward, scientific debates that impact the public are not. With Bayesian’s spirit, debates over defining prior distributions scientifically could easily fall into dispute. Should Bayesian approach outweigh the need to protect conventional wisdom of experiment-wise false positive rate? Will higher prediction probability always translate into “the more the better?” In this issue, several insightful articles, pro or skeptical on Bayesian’s view, contribute to this newly introduced column of Controversial Statistical Issue.

With members’ contributions and continuing interest on our Bulletin, amusing articles in each issue become our veins of support. As always, the Editorial Board welcomes your submissions.

Sue-Jane Wang
Editor-in-chief

!!! Controversial !!!
Statistical
Issue

Extracted from ICSA Bulletin July 2000 issue, pp 38-52.
Editor: Sue-Jane Wang, Ph.D.

Bayesian Approach in Clinical Trial

Bayesian Methods in Health Technology Assessment – the Case of Clinical Trials

Dr. I.J. Lauder

**Department of Statistics and Actuarial Science
The University of Hong Kong**

Introduction

Standard statistical practice for the design, monitoring and reporting of controlled clinical trials are founded on the Neyman and Pearson frequentist theory of hypothesis testing (36).

Kadane (29) argues that there are distinct advantageous aspects of the Bayesian approach with specific reference to the area of clinical trials, and that advances in Bayesian technology (20, 22) have made Bayesian inference a practical inferential tool.

In a key read paper Spiegelhalter et al (42) argue that that historical and continuing debate on the frequentist vs the Bayesian approach to clinical trials (4, 13, 46) is largely ideological and that a pragmatic approach is realistic to the conduct of trials. They concentrate on the areas of monitoring and reporting of clinical trials as this is where the

differences between the two methods is greatest.

The discussion on this paper is wide-ranging and illuminating and testifies to the interest and perceived relevance of Bayesian methodology in clinical trials.

Another important point of reference is the special issue of Statistics in Medicine (44) dedicated to papers on the pros and cons of the Bayesian approach to clinical trials.

In this short review, we attempt to summarize the current status of the Bayesian approach to clinical trials. To be concise, the methodology is discussed thinking mainly of two parallel groups in a fixed sample or sequential setting. This is not to say that Bayesian methods are not relevant to other types of design such as cross-over, equivalence, N of 1 and factorial trials and Phase IV safety monitoring (24, 25, 27, 40, 48). Neither is there space for discussion on

the decision theoretic approach using loss functions (3, 6, 7, 8, 32, 46), and the relevance to Phase I, II and III trials.

It is hoped that the references cited will provide a guide for interested researchers to assess the possible advantages of the Bayesian approach in the clinical trials setting. Again, to be concise, technical details will be kept to a minimum.

Main Issues

The main differences between the frequentist and Bayesian approaches occur in the monitoring and reporting of clinical trials. We summarize these issues as they are addressed in (42).

The frequentist approach uses type I and type II error for trial design and p -values and confidence intervals for analysis. The Bayesian approach specifies a prior

probability function on the parameters and moves to the posterior probability function through the combination of the prior with the likelihood in Bayes's formula. Treatments can then be compared by computing probabilities for clinical equivalence or difference, based on the posterior probability function and associated credible intervals (i.e. direct probability interpretation of the results.)

Interim analyses can be made easily at any time using the current posterior probability function and are not affected in their Bayesian form by the number of interim analyses. Interim Bayesian prediction can be made by use of predictive probability. These are an aid to deciding on the continuation or stopping of a trial.

The power of the design can be expressed in Bayesian terms by pretrial predictions based on the prior. This is known as predictive power.

Table 1 - Summary comparison of frequentist and Bayesian methods in clinical trials

Issue	Frequentist	Bayesian
External information	Informally used in design	Used formally by specifying a prior probability distribution
Parameter	A fixed state of nature	An unknown quantity which can have a probability distribution
Basic question	"How likely are these data given a particular value of the parameter?"	"How likely is a particular value of the parameter given these data?"
Reporting statistical results	Likelihood functions, p -values, confidence intervals	Plots of posterior distributions of the parameter, calculation of specific posterior probabilities of interest, and use of the posterior distribution in formal decision analysis.
Interim analyses	# of analyses dictates overall and nominal significance levels	Probability and credible interval calculations not affected by the

	and repeated confidence intervals.	number or timing of interim analyses.
Interim predictions	Conditional power analyses	Predictive probability of getting a firm conclusion.
Dealing with subsets in trials	Adjusted p -values (e.g. Bonferroni)	Subset effects shrunk towards zero by a “skeptical” prior

Table 1 gives a summary comparison of the frequentist and Bayesian approaches in relation to clinical trials.

The main criticism of the Bayesian approach is how to choose the prior. This issue is also addressed in (42). The use of different forms of prior such as reference priors, clinical priors, skeptical priors and enthusiastic priors is presented. The motivation is that the interpretation of the current results through the use of the respectively induced posterior probability functions can be used as a basis for discussion of the current status of the trial and as a form of sensitivity analysis.

Applications

We cite references where Bayesian methods have been applied in the RCT setting (1, 2, 5, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 23, 26, 28, 30, 31, 33, 34, 35, 37, 38, 39, 41, 43, 45, 47). Reference (47) comes from the authors own experience where Bayes predictive methods were used as an aid in a frequentist group-sequential design.

References

[1] K. Abrams, D. Ashby, and D. Errington. Simple Bayesian analysis in clinical trials – a tutorial. *Controlled Clinical Trials*, 15:349-359, 1994.
 [2] K. Abrams, D. Ashby, J. Houghton, and D. Riley. Assessing drug interactions:

Discussion

It has not been possible to summarize the evolving role of Bayesian methods in all areas of Health Technology Assessment. Other relevant areas where advances are occurring and should be occurring are:-

Evidence Synthesis, Observational Studies and Strategy, decision and policy making in health research and health care.

As we move into the 21st century, it is clear that the old positions of frequentist vs. Bayesian, at least in the pragmatic approach to clinical trials, is becoming less relevant. The argument here is not “Bayes is best”, but simply to alert the reader to “Bayes is a fact of life” is becoming more realistic in clinical trials practice. The constructive interplay of the two philosophies can only help to serve the primary outcome – the patient under study.

Acknowledgement

Table 1 is based on material in personal communication form from David Spiegelhalter as are the references for Bayesian RCT’s. Personal thanks are due.

Tamoxifen and cyclophosphamide. In Berry and Stangl [59], chapter 1, pages 3-66.
 [3] J.A. Bather. On the allocation of treatments in sequential medical trials. *International Statistical Review*, 53:1-13, 1985.

- [4] D.A. Berry. Interim analysis in clinical trials – the role of the likelihood principle. *Amer. Statistician*, 41:117-122, 1987.
- [5] D.A. Berry. Monitoring accumulating data in a clinical trial. *Biometrics*, 45:1197-1121, 1989.
- [6] D.A. Berry. Bayesian methods in phase III trials. *Drug Information J.*, 25:345-368, 1991.
- [7] D.A. Berry. A case for Bayesianism in clinical trials. *Statistics in Medicine*, 12:1377-1393, 1993.
- [8] D.A. Berry. Decision analysis and Bayesian methods in clinical trials. *Cancer Treatment Research*, 75:125-54, 1995.
- [9] D.A. Berry, M.C. Wolff, and D. Sack. Decision making during a phase III randomised controlled trial. *Controlled Clinical Trials*, 15:360-378, 1994.
- [10] J.M. Brophy and L. Joseph. Placing trials in context using Bayesian analysis – GUSTO revisited by Reverend Bayes. *J. Amer. Medical Association*, 273:871-875, 1995.
- [11] J.M. Brophy and L. Joseph. Bayesian interim statistical analysis of randomized trials. *Lancet*, 349:1166-68, 1997.
- [12] B.P. Carlin, K. Chaloner, T. Church, T.A. Louis, and J.P. Matts. Bayesian approaches for monitoring clinical trials with an application to toxoplasmic encephalitis prophylaxis. *Statistician*, 42:355-367, 1993.
- [13] J. Cornfield. Sequential trials, sequential analysis and the likelihood principle. *Am. Statistn.*, 20:18-23, 1966.
- [14] R. Dersimonian. Meta-analysis in the design and monitoring of clinical trials. *Statistics in Medicine*, 15:1237-1248, 1996.
- [15] J.J. Digman, J. Bryant, H.S. Wieand, B. Fisher, and N. Wolmark. Early stopping of a clinical trial when there is evidence of no treatment benefit: protocol B-14 of the national surgical adjuvant breast and bowel project. *Controlled Clinical Trials*, 19:575-588, 1998.
- [16] P.M. Fayers, D. Ashby, and M.K.B. Parmar. Bayesian data monitoring in clinical trials. *Statistics in Medicine*, 16:1413-1430, 1997.
- [17] A. Fletcher, D. Spiegelhalter, J. Staessen, L. Thijs, and C. Bulpitt. Implications for trials in progress of publication of positive results. *Lancet*, 342:653-657, 1993.
- [18] L.S. Freedman and D.J. Spiegelhalter. Application of Bayesian statistics to decision making during a clinical trial. *Statistics in Medicine*, 11:23-35, 1992.
- [19] L.S. Freedman, D.J. Spiegelhalter, and M.K.B. Parmar. The what, why and how of Bayesian clinical trials monitoring. *Statistics in Medicine*, 13:1371-1383, 1994.
- [20] A. Gelman and D.B. Rubin. Markov chain Monte Carlo methods in biostatistics. *Statistical Methods Medical Research*, 5:339-55, 1996.
- [21] S.L. George, C.C. Li, D.A. Berry, and M.R. Green. Stopping a clinical trial early – frequentist and Bayesian approaches applied to a CALGB trial in non-small-cell lung cancer. *Statistics in Medicine*, 13:1313-1327, 1994.
- [22] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo Methods in Practice*. Chapman and Hall, New York, 1996.
- [23] J.B. Greenhouse and L. Wasserman. Robust Bayesian methods for monitoring clinical trials. *Statistics in Medicine*, 14:1379-1391, 1995.
- [24] A.P. Grieve. Evaluation of bioequivalence studies. *European J. Clinical Pharmacology*, 40:201-202, 1991.
- [25] A.P. Grieve. Bayesian analyses of two-treatment crossover studies. *Statistical Methods in Medical Research*, 3:407-29, 1994.
- [26] P. Gustafson. Comments on ‘Investigating Therapies of potentially great benefit: ECMO’ by J H Ware. *Statistical Methods in Medical Research*, 4:310-317, 1996.
- [27] P.H. Hsu, A. Laddu, D.C. Jordan, R.W. Stoll, and P.A. Deverka. Use of Bayesian drug causality assessment of adverse

- events in a United States pharmaceutical environment. *Clinical Pharmacology and Therapeutics*, 51:124-124, 1992.
- [28] M.D. Hughes. Practical reporting of Bayesian analyses of clinical trials. *Drug Information J.*, 25:381-393, 1991.
- [29] J.B. Kadane. Prime Time for Bayes. *Controlled Clinical Trials*, 16:313-318, 1995.
- [30] J.B. Kadane. *Bayesian Methods and Ethics in a Clinical Trial Design*. Wiley, New York, 1996.
- [31] R.E. Kass and J.B. Greenhouse. Comments on 'Investigating Therapies of potentially great benefit: ECMO' by JH Ware. *Statistical Science*, 4:310-317, 1989.
- [32] G.G. Koch. Summary and discussion for 'Statistical issues in the pharmaceutical industry: analysis and reporting of phase III clinical trials including kinetic/dynamic analysis and Bayesian analysis'. *Drug Information J.*, 25:433-437, 1991.
- [33] R.J. Lewis. Bayesian hypothesis testing: interim analysis of a clinical trial evaluating pheny-toin for the prophylaxis of early post-traumatic seizures in children. In Berry and Stangl [59], chapter 10, pages 279-296.
- [34] R.J. Lilford and J. Jackson. Equipoise and the ethics of randomization. *J. Royal Soc. Medicine*, 88:552-559, 1995.
- [35] M.K.B. Parmar, D.J. Spiegelhalter, and L.S. Freedman. The CHART trials: Bayesian design and monitoring in practice. *Statistics in Medicine*, 13:1297-1312, 1994.
- [36] S. Pocock. *Clinical Trials: A Practical Approach*. Wiley, Chichester, 1983.
- [37] S.J. Pocock and M.D. Hughes. Practical problems in interim analyses, with particular regard to estimation. *Controlled Clinical Trials*, 10:209S-221S, 1989.
- [38] S.J. Pocock and D.J. Spiegelhalter. Domiciliary thrombolysis by general practitioners. *British Medical Journal*, 305:1015, 1992.
- [39] A.A. Sasahara, T.M. Cole, F. Ederer, J.A. Murray, N.K. Wenger, S. Sherry, and J.M. Stengle. Urokinase pulmonary embolism trial, a national cooperative study. *Circulation*, 47 (supplement 2):1-108, 1973.
- [40] R. Simon and L.S. Freedman. Bayesian design and analysis of two x two factorial clinical trials. *Biometrics*, 53:456-464, 1997.
- [41] D.J. Spiegelhalter, L.S. Freedman, and M.K.B. Parmar. Applying Bayesian ideas in drug development and clinical trials. *Statistics in Medicine*, 12:1501-1517, 1993.
- [42] D.J. Spiegelhalter, L.S. Freedman, and M.K.B. Parmar. Bayesian approaches to randomised trials (with discussion). *J. Royal Statistical Soc. Series A-Statistics in Society*, 157:357-387, 1994.
- [43] D.K. Stangl. Prediction and decision-making using Bayesian hierarchical-models. *Statistics in Medicine*, 14:2173-2190, 1995.
- [44] *Statistics in Medicine*. Methodological and Ethical Issues in Clinical Trials. *Statistics in Medicine*, Special Issue, 1993.
- [45] J. Ware. Investigating therapies of potentially great benefit: ECMO (with discussion). *Statistical Science*, 4:298-340, 1989.
- [46] J. Whitehead. *The Design and Analysis of Sequential Clinical Trials (Second Edition)*. J. Wiley and Sons, Chichester, 1997.
- [47] Y.L. Yu, C.R. Kumana, I.J. Lauder, Y.K. Cheung, F.L. Chan, M. Kou, C.M. Chang, R.T. Cheung, K.Y. Fong. Treatment of acute cerebral haemorrhage with intravenous glycerol: a double-blind placebo-controlled randomised trial. *Stroke*, 23:7, 967-971, 1992.
- [48] D.R. Zucker, C.H. Schmid, M.W. McIntosh, R.B. Agostino, H.P. Selker, and J. Lau. Combining single patient (n-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *J. Clinical Epidemiology*, 50:401-10, 1997.

Conditional Power – A Bayesian Procedure

Gordon Lan,
Pfizer Inc.

In many long-term clinical trials, data are monitored periodically even if the design of the study is fixed. Before the group sequential methods were introduced in the 1970's, many NIH-sponsored trials used the concept of "conditional power (CP)" as a guideline for interim data analyses. To simplify the discussion, let us assume that a new treatment is tested to show beneficial effects. Based on the observed data, the conditional power is the probability that the final Z-value will fall into the rejection region, which leads to a significant beneficial conclusion for the new treatment. If the ultimate outcome of a trial can be predicted with a high probability, i.e., if the CP is very high or very low, perhaps the trial should be terminated early. Note that when data are monitored, the CP depends on the unknown value of the treatment effect, θ , which determines the distribution of future observations.

During the Data Safety and Monitoring Board (DSMB) meetings, the Statistical Center may present $\{CP(\theta)\}$ for a spectrum of values of θ . The DSMB members then use

this information to determine whether the study should be terminated early. A classical Bayesian approach would use the predictive probability – a weighted average of $CP(\theta)$ by the posterior distribution of θ - for consideration of early termination. In practice, I have found that this Bayesian approach was used in a more "informal" way. The DSMB members have different scientific backgrounds, and each member uses his/her own prior distribution which may not be the same for all the members. Second, these members may have different exposure to the new treatment. Note that many clinicians sitting on the DSMB are experts in the medical field being studied; these clinicians may have access to external information about the new treatment, which they are not allowed to share with other members of the DSMB. Finally, when a DSMB member pools all the information available to him/her, he/she may not specifically use the Bayes formula to derive a predictive power, but rather instinctively derive a value and the decision is reflected in their vote on early termination.

To control the α -level, CP should be evaluated at $\theta = 0$. If $CP(0)$ is high, then the use of conditional power to stop early for benefit will only inflate the α -level slightly. This special aspect is only a small part of the CP consideration during the discussion in the DSMB meetings. The "informal" Bayesian modification as described above and other medical considerations carry much more weight than the control of the α -level in the early stopping decision process.

References

(1) Lan KKG and Wittes J (1994). "Data Monitoring in Complex Clinical Trials: Which

treatment is better?" *J of Statistical Planning and Inference*, Vol. 42: 241-255.

(2) Spiegelhalter DJ, Freedman LS and Blackburn PR (1986). Monitoring Clinical Trials: conditional or predictive power?" *Controlled Clinical Trials*, Vol. 7: 8-17.

Bayesian statistics in clinical trials – Is there a future?

Professor Andrew P. Grieve

**Pfizer Central Research,
Ramsgate Road,
Sandwich, Kent,
CT13 9NJ, UK.**

I work for a pharmaceutical company and therefore it should not be surprising that when addressing the issues of using Bayesian methods in clinical trials I look at them through pharmaceutical eyes as well as Bayesian eyes.

First let us remind ourselves that Bayesian ideas are not new. Thomas Bayes' original paper appeared posthumously in 1763; Laplace in 1774 was using posterior distributions in practical applications; in 1898, when studying uncertainty in estimating the correlation coefficient, Karl Pearson used a Bayesian approach; Gosset (Student) developed the sampling distribution of the correlation coefficient in 1908 to simplify the calculations required to determine the appropriate posterior distribution.

Despite this early work there is little evidence to suggest that in the field of clinical trial methodology Bayesian ideas caught on – a notable exception being Cornfield (1966). Things are however changing. Both in the United States and Europe there has been significant activity in bringing to the attention of practising statisticians the advantages of a Bayesian perspective in clinical trials (Berry, 1993; Spiegelhalter et al, 1994).

In drug development Bayesian ideas have until very recently been notable by their absence. Why? Certainly Bayesian methods are now formally acceptable by the major world regulatory authorities. The 1998 ICH Guideline on Statistical Principles in Clinical Trials (E9) says:

“ The use of Bayesian and other approaches may be considered when the reasons for their use are clear and when the resulting conclusions are sufficiently robust”.

While I for one find this endorsement somewhat lukewarm the door is now open to the use of Bayesian techniques. So what other hindrances could prevent their use? I think there are four major hindrances.

1. Philosophical Objections.

Many statisticians find it hard to accept the use of Bayesian methods because they are based on a view of probability that sees it as a subjective measure of belief rather than the more traditional definition based on repeated sampling. This is certainly a difficulty. However, I believe that any Bayesian Statistician working for a pharmaceutical company needs to remember that he/she is working in a regulatory environment that is “frequentist”. A

consequence is that such a statistician needs to ensure that their Bayesian methods, even if based on a subjective concept of probability, are calibrated within this “frequentist world”. This can be done by simulating from known cases and examining the operating characteristics of the proposed method.

2. Trust

There is a feeling that the use of subjective priors may allow unscrupulous companies and/or their statisticians to try to dupe, or pull the wool over regulators’ eyes. I think this is very unlikely. First, as I said above, Bayesian statisticians need to calibrate their methods and I cannot imagine that a method which it can be shown inflates the false positive rates would be acceptable. Secondly, Professor Stephen Senn has pointed out that “nowhere is the discipline of statistics conducted with greater discipline than in the pharmaceutical industry”. I believe that this statement will be equally true of Bayesian statistics. Documentation is a creed of pharmaceutical statistics and it will be no different with Bayesian methods. The prior distributions will need to be specified in the protocol, as will utility functions if required, they will need to be justified, and they will not be able to be changed. In my mind it is unlikely that undocumented, subjective priors will be allowed.

3. Conservatism

Pharmaceutical companies tend to be conservative as far as statistics are concerned. The argument is often made that we do not need new methods because we have successfully registered drugs with the existing methods. That is true, but it does not preclude the development of alternative, more efficient, methodology. As statisticians we need to learn

to sell technical advancement not solely to fellow statisticians but also to budget holders. The impact of statistics on the bottom line is a very convincing argument for changing methods!

4. Lack of Tools

Many statisticians will tell you that they would have used Bayesian methods had the necessary tools been available. They are now. The development over the last 10 years of Monte Carlo Markov Chain methods has made the use of Bayesian methods a practical proposition for the applied statistician.

These objections, although substantial, are not insurmountable. I believe that there is a good prospect that more statisticians in the pharmaceutical industry will in the future use Bayesian ideas. Their use may start in early phases of drug development where decision making tends to be internal to the company. But with familiarity their use will widen even if the prospect of a Bayesian 21st century, as predicted by Dennis Lindley, is not yet certain.

References

- Berry, D.A. (1993). A case for Bayesianism in clinical trials. *Statistics in Medicine*, 12, 1377-1404.
- Cornfield, J (1966). A Bayesian test of some classical hypotheses – with applications to sequential clinical trials. *Journal of the American Statistical Association*, 61, 577-594.
- Spiegelhalter, D.J., Freedman, L. S. and Parmar, M.K.B. (1994). Bayesian approaches to randomised trials. *Journal of the Royal Statistical Society Series A*, 157, 357-387.

On A Bayesian Subject Allocation Method in Clinical Trials

Daniel Y.T. Fong

Clinical Trials Centre, Faculty of Medicine,
The University of Hong Kong

Email: dytfong@hku.hk

1. Subject Allocation in Controlled Clinical Trials

Clinical trials are mandatory before a regulatory authority can approve use of a medicinal product. Before the start of any clinical trial, the trial protocol must be approved by an Institutional Review Board to ensure that the trial design is ethical and that the rights of all study subjects are secured. Specifically, all participating subjects must sign a proper informed consent form to ensure they are well aware of the study details and their rights to withdraw at any time without penalty.¹

Since R.A. Fisher introduced the concept of randomization in 1926, randomized controlled trials (RCTs) have been recognized as the “gold standard” for the evaluation of treatment efficacy and safety by including active and placebo controls.² Randomisation

is a vital element in design of experiments to minimise bias from uncontrolled systematic factors and to facilitate valid statistical inference. However, a common challenge to RCTs is the allocation of study subjects to a sub-optimal treatment (e.g., placebo/active control) although it is unclear which treatment will be optimal until all data are gathered and analysed. This leads to tension between individual ethical concerns to provide the best-known treatment for individual patients and collective ethical concerns to advance current knowledge for the benefit of future patients.^{3,4}

A Bayesian approach for subject allocation to study treatments, termed the Kadane-Sedransk-Seidenfeld (KSS) method, has been previously proposed.³ This aims to optimize the individual and collective ethics simultaneously. Before recruitment starts, a group of medical experts is formed and their opinions on the efficacy of the study treatments at different levels of certain prognostic factors are solicited. A statistical model, called Prior, is then developed to imitate the expert opinions. Given the prognoses of a newly recruited subject, predicted outcomes when the subject was assigned with different treatments are computed using the Prior model. The subject is then allocated accordingly to the “optimal” treatment with the best-predicted outcome. The expert Prior model is continually updated as more efficacy data are accrued through an automatic computer program. Periodic calibration of the Prior model may also be performed by further advice from the expert panel when presented with some hypothetical efficacy data.

In general, such a Bayesian allocation method is less appealing for trials that aim to study long-term efficacy, as most of the subjects are likely to be recruited and assigned a treatment before any efficacy data are

accrued to update the Prior model. On the other hand, there may be over-emphasis on the veracity of the initial Prior model for trials that focus on short-term efficacy. This is because only a few elicitation exercises can be performed for calibration. In the following discussion, we will focus on situations where efficacy data are steadily accrued, with sufficient calibration of the Prior model well before subject recruitment is completed.

2. The Use of Prior

Subjective bias introduced by the use of Prior has been a criticism of Bayesian methods. We discuss here a slight elaboration regarding the use of Prior for subject allocation.

Prior can be treated as an adaptive allocation device, governed by both soliciting expert opinions (qualitatively) and by accruing efficacy data with some hypothesized statistical assumptions (quantitatively). It aims to imitate expert opinion for the best current treatment to patients having different prognostic risks. The use of subject prognoses to decide an optimal treatment is attractive but debatable. Specifically, there might be an imbalance between treatment groups with respect to certain prognostic factors resulting from the Bayesian method. For instance, severely ill patients may tend to be assigned or not assigned to an investigating treatment by the experts. Some prognostic factors might thus be confounded with the treatment effect and thus the outcome is not collectively ethical.

In addition, the Prior model is updated in accordance with certain statistical distribution assumptions that are often hypothesized for the sake of mathematical convenience, and no justification can be made. To a certain extent, calibration by the expert

panel can help to warrant the veracity of the continually updated Prior. Frequent calibration is, however, not feasible, as the resulting Prior relies too much on the medical judgement of the expert panel, with few statistically valid grounds for such decisions. Nevertheless, when there is a significant discrepancy between the updating process and reality, it will again be unethical on the collective level, though the outcome may not be too serious at the individual level when compared with standard randomization procedures.

Moreover, the Prior will be extremely biased when data are missing owing to inferiority of a particular study treatment when those recorded tend to be responsive to the treatment. The Prior will then continually assign subjects to that study treatment not because it is “optimal”, but because its inferiority is not recognized.

3. Logistic Difficulty

To build up a good Prior, the expert panel should have a good span of knowledge and hold distinctive views in the area of the study treatment. However, when there is limited practical knowledge of the investigating treatment, gathering expert opinion may present great difficulty.⁴

In addition, clinical trials data are first recorded on case report forms (CRFs) by a study nurse at the study site. After certain data monitoring procedures, all data are entered twice independently into two separate databases that are compared and have all discrepancies removed before analysis is undertaken. Therefore, there will be a time lag before the efficacy data of a subject on the CRFs are transferred into a database and checked for data quality. By the time the Prior is updated, there may have been a

number of additional subjects recruited who need to have their treatments decided, and more efficacy data accrued. The advance of electronic data capture technology might help to mitigate this issue, but further technological development is needed.

4. Conclusions

The motive of the Bayesian method that aims to assign the best treatment to patients by imitating expert opinion is no doubt superior. Only a few personal viewpoints are briefly outlined, while other issues such as how the sample size is computed would also be worth discussing. A broad conclusion is that reservations are yet to be resolved before the Bayesian method for subject allocation can be practically employed in most comparative trials. Randomized controlled trials remain the gold standard design for evaluation of treatment efficacy and safety, provided a proper monitoring system and good clinical practice (including proper administration of informed consent) is secured.

Acknowledgement

Financial support from the Committee on Research and Conference Grants is gratefully acknowledged.

References

1. Karlberg J and Tsang K. Introduction to Clinical Trials. Clinical Trials Centre, Hong Kong (1998).
2. Box JF. R.A. Fisher and the design of experiments, 1922-1926. *American Statistician* 34:1-7, 1980.

3. Kadane JB. Bayesian Methods and Ethics in a Clinical Trial Design. John Wiley & Sons, New York (1996).
4. Palmer CR and Rosenberger WF. Ethics and Practice: Alternative Designs for Phase III Randomized Clinical Trials. *Controlled Clinical Trials* 20:172-186, 1998.

Bayesian vs. Frequentist

Tie-Hua Ng
U. S. Food and Drug Administration

Summary. A simple experiment of tossing a coin is used to illustrate the basic concept of Bayesian approach and highlight the basic differences between the Bayesian and the frequentist approaches.

Bayesian approach. Let's consider a parameter of interest, say, θ . The basic assumptions of the Bayesian approach are (1) θ is an unknown random variable, and (2) there is a prior distribution for θ . We then compute the posterior distribution of θ given the data as follow:

$$\Pr[\theta | \text{data}] \propto \Pr[\text{data} | \theta] \Pr[\theta],$$

$\Pr[\text{data}|\theta]$ is the distribution of the data given θ , and $\Pr[\theta]$ is the prior distribution of θ .

Coin Experiment. Let's consider a simple experiment of tossing a coin. The parameter of interest θ is the probability of a head. We say that it is a fair coin, if $\theta = 0.5$, and that it is an unfair coin, otherwise. Suppose that we make n tosses. Based on the outcomes of these n tosses, we are interested in answering the question "Is it an unfair coin?" This may be formulated, in the frequentist framework, as testing the null hypothesis that $\theta = 0.5$ against the alternative hypothesis that $\theta \neq 0.5$.

Prior Distribution. Suppose that there are 30 coins in the box --- 6 fair coins labeled as blue, 21 of the unfair coins labeled as red with $\theta=0.6$ and 3 unfair coins labeled as green with $\theta=0.7$. For a given coin, θ is an unknown constant. On the other hand, if the coin is selected at random so that each coin has an equal chance of being selected, then the parameter of interest, denoted by Θ , is a random variable with the following distribution.

$$\Pr[\Theta = 0.5] = 0.2,$$

$$\Pr[\Theta = 0.6] = 0.7,$$

$$\Pr[\Theta = 0.7] = 0.1.$$

This is known as the prior distribution of Θ .

Let's suppose that we select a coin at random from the box without knowing the color of the coin (so Θ is not observed). We then toss the coin 10 times and count number of heads. Let X be the number of heads. Then we have a bivariate distribution (Θ, X) . Although Θ is fixed once a coin is selected, if we do not know its value, Θ can still be considered random because we know nothing more than what we already knew before the coin was selected.

Clearly, given $\Theta = \theta$, X follows a Binomial distribution with parameters $(10, \theta)$, as shown in Figure 1. If we multiply the distribution of X given Θ by the prior distribution, we get the joint distribution of Θ and X as given in Figure 2. Therefore, the prior distribution acts like weights. If we do not know the proportions of the three types of coins, then we do not know the priors. In that situation, non-informative priors are often used. That is we assume that there are 10 coins of each type. We see that a non-informative prior means equal weights and that it is the "best guess" of the prior distribution of Θ , if there is no prior data.

Posterior Distribution. The posterior distribution of Θ is the conditional distribution

of Θ given X . It is computed by normalizing the joint distribution of Θ and X (Figure 2) for each X . In other words, for each column in Figure 2, divide the joint probabilities by the column sum. As an example, $\Pr[\Theta = 0.6 | X = 7]$ of 0.75 is computed by dividing 0.15 by 0.2 which is the column sum at $X = 7$. The posterior distribution of Θ is shown in Fig. 3.

Is it an unfair coin? Let's suppose that we observe $X = 9$, then the posterior probability of an unfair coin is 0.954 (see Figure 3) which highly indicates that it is an unfair coin. So, it is logical to conclude that it is an unfair coin. In that case, the error probability of concluding an unfair coin is 0.046. On the other hand, if we observe $X = 3$, then the posterior probability of an unfair coin is 0.564. If we conclude that it is an unfair coin, then the error probability of concluding an unfair coin is 0.434. Therefore, if we want to limit the error probability of concluding an unfair coin to 0.05, then we would conclude that it is an unfair coin if the posterior probability of an unfair coin given X exceeds 0.95. In that case, we would conclude that the coin is unfair, if $X = 9$ or 10.

Bayesian versus Frequentist. Frequentist approach is derived from the distribution under the null, that is $\Theta = 0.5$. At 0.05 significance level, we would conclude that the coin is unfair if $X = 9$ or 10 (see Figure 1). In this particular example, both approaches lead to the same conclusion, that is to conclude that the coin is unfair, if $X = 9$ or 10. This might not be the case in general.

The frequentist approach controls the error probability given that it is a fair coin ($\Theta = 0.5$) regardless of the prior distribution of Θ . In contrast, the Bayesian approach controls the error probability conditional on X , which depends on the prior distribution. Clearly, if we know the prior distribution, Bayesian approach has a big advantage over the

frequentist approach. On the other hand, if we do not know the prior distribution and use a non-informative prior, depending upon the true prior and observed X , the error probability of concluding an unfair coin may not be controlled.

Reference

Ng, T-H. "Can Bayesian approaches be used in clinical trials?" Presented at PhRMA Workshop, Baltimore, MD, September 7-10, 1997, and ICOSA Applied Statistics Symposium, June 18-20, 1999.

Figure 1-3 in the process of loading.

Bayesian Statistics in a Regulatory Environment

Gregory Campbell

Director, Division of Biostatistics
Center for Device and Radiological Health
U.S. Food and Drug Administration

The Center for Devices and Radiological Health (CDRH) in the U.S. Food and Drug Administration (FDA) has recently embarked on an initiative to investigate how Bayesian design and analysis can be used effectively by medical device companies in their pre-market submissions to CDRH.

While Bayesian statistics has been around for quite a while, a natural question is why would CDRH embark on this path at this time. Prior information plays a key role in the regulation of medical devices. In contrast to pharmaceutical products, the mechanism of action of many devices is physical as opposed to pharmacokinetic, local as opposed to systemic. There is often an abundance of information for the same device or a very similar one. This is so because the nature of medical device development is a series of (often gradual) changes in the design and construction. Further, there are sometimes data on trials conducted abroad or from pilots and data registries. If the control is also a device, its performance may be already very well characterized. The remarkable change in the past few years in Bayesian statistics, a change that has catapulted the subject from the

mathematical treatment of idealized conjugate priors to very realistic modeling, has been the arrival of fast computation combined with clever algorithms for Bayesian calculations. This advance now allows one to compute the posterior distribution (and any function thereof) for virtually any prior distribution.

In 1997 CDRH launched an effort to investigate whether the application of this methodology would be of use in the pre-market review of applications by medical device companies to the FDA. To this end CDRH enlisted the help of a Bayesian expert (D. Malec) for a period of over a year to explore the various issues that might arise in the use of Bayesian statistics in a regulatory environment. More recently, the Division of Biostatistics has hired two statisticians who are well-trained in Bayesian methods. In November, 1998, CDRH held a workshop (jointly sponsored by the Health Industry Manufacturers Association) that was attended by over 200 participants and very well-received. It featured presentations by FDA personnel, by academicians including D. Berry, W. Strawderman, and M. Escobar, and real case studies by representatives of four device manufacturers.

In the past year there have been two Pre-Market Approval applications (PMAs) that have been approved by the FDA with Bayesian analysis. A description of the analysis for these is in their Summaries of Safety and Effectiveness at the FDA website (www.fda.gov/cdrh/pdf/p970015b.pdf, www.fda.gov/cdrh/pdf/p970033b.pdf).

The philosophical underpinning for the success of this effort among the medical device community is that there needs to be an obvious potential benefit for the manufacturer for undertaking a clinical trial that is designed and analyzed with Bayesian methodology. One obvious area is in the use of valid

scientific evidence outside the scope of the current trial. In contrast to the traditional frequentist trial, wherein one allows no formal incorporation of any prior information whatsoever, there is often good quantitative prior data for device studies that could be utilized. If such prior quantitative data were predictive of the current trial, it could help to short-circuit any argument about the validity of the prior and its use could dramatically reduce the size of the current study. For example, in a hierarchical modeling situation, the question for prior quantitative data sources is whether the studies are what is called exchangeable with each other and with the current trial. (This is to be distinguished from exchangeability of subjects across trials.) This determination is not merely a statistical one but requires clinical expertise as well. Is it the same or a very similar device? Is the patient population and the inclusion /exclusion criteria the same or very close?

Other ways in which Bayesian trials can differ are as follows: 1) The rule to declare success for a clinical trial can be very different in a Bayesian study. Rather than rely on the size of some P-value that measures how deviant the results are under the null hypothesis, the decision for the Bayesian trial would be based on the posterior distribution in some fashion. 2) Bayesian predictive methods can be helpful in assessing the validity of a surrogate endpoint or in monitoring an on-going trial. 3) Trials for ethically sensitive devices can be designed (and analyzed) that allow for adaptive assignment of treatments to patients. 4) Bayesian methods can be employed to impute missing data.

An underlying philosophy is that there is no one unique methodology (be it frequentist or Bayesian) for a particular device

to plan to show it safe and effective. It is up to the company to propose the design of the trial and its planned analysis and then meet with the FDA for a discussion of it. Medical device companies are encouraged to meet early with CDRH representatives to discuss the proposed design and analysis; for Bayesian trials, this is especially important.

Many interesting research questions have arisen in this effort to investigate the use of Bayesian statistics in the regulatory environment. For example, what is the Bayesian analog of Type I error that protects the American public from devices that are unsafe or ineffective? What is a realistic way to anticipate an expected sample size for a Bayesian trial? What are the implications of the Likelihood Principle? What are frequentist properties of Bayesian stopping rules? How sensitive are hierarchical Bayesian analyses to non-informative priors on hyperparameters (parameters of the prior distribution), especially when there is only a single previous quantitative data set? How informative is the *de facto* prior that is generated from a hierarchical model on multiple previous data sets, when non-informative priors are assumed for hyperparameters? Other interesting questions relate to Bayesian subgroup analysis and to the implementation of a decision analysis approach.

Statisticians who are thinking about future training for applied work should give serious consideration to becoming familiar with Bayesian methodology. It is expected that this methodology will continue to play an ever increasingly important role in applied statistics in general and in the regulation of medical devices in the U.S. in particular.

MESSAGE FROM THE PRESIDENT

Dear ICSA Members:

Since the founding of ICSA, our association has been growing steadily. This is reflected by the increasing number of members and their interest in participating in symposiums and conferences. To wit, over 200 members attended the 2000 Applied Statistics Symposium in Piscataway, New Jersey. Although most participants were from North America, many of them were from China, Hong Kong, Taiwan and some from Australia, Jordan and other countries. The symposium was a big success. Our sincere thanks go to the Applied Statistics Symposium Committee.

To continue the growth of our association, we must build bridges between ICSA and the statistics community, including industry, government and academia worldwide. Of course, the bridge builders are the members. With your hard work, you can make ICSA serve your organization. When you attend an ICSA symposium or conference, you are benefited by acquiring new knowledge and making new friends, which in turn benefit your organization. Further your bosses and colleagues will know more about ICSA and the bridge will be strengthened.

We need to recruit new members so that we have more bridge builders. The Membership Committee is currently compiling a master list of potential members, who will be invited later this year to join ICSA. You can also participate in the membership drive by asking your friends to join us if they are not members of ICSA or by putting them on the master list by sending their names to me or to any member of the Membership Committee.

The finance of our association is in excellent condition (please see the Treasurer's report). We are constantly trying to use the resources to better serve our members, as ICSA is your association. Any suggestion to enhance our service to you is highly appreciated. On the other hand we must strengthen and secure our future financial situation. The Board of Directors has passed a motion to create a committee to study the long range financial planning of our association. This committee will undoubtedly steer us toward financial success.

As you know, the term of our Executive Director, Dr. Naitee Ting, will end this December. A new Executive Director will be elected. I would like to take this opportunity to thank Dr. Ting for his outstanding service to ICSA. He has spent endless hours to help me and my predecessors to fulfill our duties as presidents. Thank you, Naitee.

Our next membership meeting will be on August 16 in Indianapolis at the Joint Statistical Meetings. I hope that you are planning to attend JSM, and that I will see you there.

Chien-Pai Han,

President

**Special Thanks
from The
Editorial
Board**

The Editorial Board would like to thank the following volunteers during the preparation of this issue: Dr. Timothy Chen, Dr. Yi Tsong, Dr. H.M. James Hung, and Dr. Yu Liang. The Board recognizes them the guest members for the Editorial Working Committee.

If you have a new idea and are interested in joining us, please send your C.V. including your plan to the Editorial Board WANGS@CDER.FDA.GOV for consideration.

We encourage your active involvement in the ICSA Bulletin. Every effort counts.

□□□□□