

BMJ

Methods in health service research: An introduction to bayesian methods in health technology assessment

David J Spiegelhalter, Jonathan P Myles, David R Jones and Keith R Abrams

BMJ 1999;319:508-512

Updated information and services can be found at:
<http://bmj.com/cgi/content/full/319/7208/508>

These include:

References

This article cites 18 articles, 5 of which can be accessed free at:
<http://bmj.com/cgi/content/full/319/7208/508#BIBL>

14 online articles that cite this article can be accessed at:
<http://bmj.com/cgi/content/full/319/7208/508#otherarticles>

Rapid responses

One rapid response has been posted to this article, which you can access for free at:

<http://bmj.com/cgi/content/full/319/7208/508#responses>

You can respond to this article at:

<http://bmj.com/cgi/eletter-submit/319/7208/508>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Topic collections

Articles on similar topics can be found in the following collections

[Other Statistics and Research Methods: descriptions](#) (600 articles)

[Bayesian statistics: descriptions](#) (19 articles)

Notes

To order reprints follow the "Request Permissions" link in the navigation box

To subscribe to *BMJ* go to:

<http://resources.bmj.com/bmj/subscribers>

takes an interdisciplinary approach to stemming the epidemic, building links with other international organisations, non-governmental organisations, the private sector, and the research community.

Communicating reality and vision

Much remains to be done to raise awareness and concern about cancer in the developing world. The yawning gap between poor and rich countries persists, and cheap effective technologies such as hepatitis B vaccine are not applied. There is a pressing need to deal pragmatically with today's problems by setting realistic priorities. Yet health professionals also have a responsibility to expand what is feasible. As Article 27 of the Universal Declaration of Human Rights states, "Everyone has the right . . . to share in scientific advancement and its benefits." This vision can be communicated through persuasive and practical arguments for placing cancer in developing countries squarely in context—and firmly on the agenda.

I thank N Muñoz and R Sankaranarayanan for discussions and advice, and P Kleihues, DM Parkin, K Sikora, A Narinesingh, E

LeGresley, and Y Daikh for their comments. Special thanks go to KJ Hughes.

Competing interests: None declared.

- 1 UNAIDS/WHO. *Epidemiological fact sheet on HIV/AIDS and sexually transmitted diseases*. Geneva: UNAIDS/WHO, 1998.
- 2 World Health Organisation Programme on Cancer Control. *Cancer strategies for the new millennium*. Proceedings of a conference held at the Royal College of Physicians, London, 1998. (www.who-pcc.iarc.fr/Publications/Publications.html; accessed 2 March 1999.)
- 3 Murray CJL, Lopez AD, eds. *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard University Press, 1996.
- 4 Pisani P, Parkin DM, Muñoz N, Ferlay J. Cancer and infection: estimates of the attributable fraction in 1995. *Cancer Epidemiol Biomarkers Prev* 1997;6:387-400.
- 5 Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int J Cancer* 1999;80:827-41.
- 6 Pisani P, Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 major cancers in 1990. *Int J Cancer* 1999 (in press).
- 7 Sankaranarayanan R, Black RJ, Parkin MD, eds. *Cancer survival in developing countries*. Lyons: International Agency for Research on Cancer, 1998.
- 8 Walboomers JMM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999 (in press).
- 9 WHO Expanded Programme on Immunisation. *Hepatitis B vaccine—making global progress*. Geneva: WHO, 1997.

(Accepted 9 April 1999)

Methods in health service research

An introduction to bayesian methods in health technology assessment

David J Spiegelhalter, Jonathan P Myles, David R Jones, Keith R Abrams

This is the third of four articles

MRC Biostatistics Unit, Institute of Public Health, Cambridge CB2 2SR

David J Spiegelhalter, senior statistician
Jonathan P Myles, research assistant

Department of Epidemiology and Public Health, University of Leicester, Leicester LE1 6TP

Keith R Abrams, senior lecturer in medical statistics
David R Jones, professor of medical statistics

Correspondence to: Dr Spiegelhalter david.spiegelhalter@mrcc-bsu.cam.ac.uk
Series editor: Nick Black

BMJ 1999;319:508-12

Bayes's theorem arose from a posthumous publication in 1763 by Thomas Bayes, a non-conformist minister from Tunbridge Wells. Although it gives a simple and uncontroversial result in probability theory, specific uses of the theorem have been the subject of considerable controversy for more than two centuries. In recent years a more balanced and pragmatic perspective has emerged, and in this paper we review current thinking on the value of the Bayesian approach to health technology assessment.

A concise definition of bayesian methods in health technology assessment has not been established, but we suggest the following: the explicit quantitative use of external evidence in the design, monitoring, analysis, interpretation, and reporting of a health technology assessment. This approach acknowledges that judgments about the benefits of a new technology will rarely be based solely on the results of a single study but should synthesise evidence from multiple sources—for example, pilot studies, trials of similar interventions, and even subjective judgments about the generalisability of the study's results.

A bayesian perspective leads to an approach to clinical trials that is claimed to be more flexible and ethical than traditional methods,¹ and to elegant ways of handling multiple substudies—for example, when simultaneously estimating the effects of a treatment on many subgroups.² Proponents have also argued that a bayesian approach allows conclusions to be provided in a form that is most suitable for decisions specific to patients and decisions affecting public policy.³

Summary points

Bayesian methods interpret data from a study in the light of external evidence and judgment, and the form in which conclusions are drawn contributes naturally to decision making

Prior plausibility of hypotheses is taken into account, just as when interpreting the results of a diagnostic test

Scepticism about large treatment effects can be formally expressed and used in cautious interpretation of results that seem "too good to be true"

Multiple subanalyses can be brought together by formally expressing a belief that their conclusions should be broadly similar

Use of bayesian methods in health technology assessment should be pursued cautiously; guidelines, software, and critically evaluated case studies are needed

Many questions remain: notably, to what extent the scientific community or regulatory authorities will allow the explicit consideration of evidence that is not totally derived from observed data. In this article we

outline the available literature, discuss the main techniques that are being suggested, and provide some recommendations for future work.

Nature of the evidence

A “bayesian” approach can be applied to many scientific issues, and a search for this term in the Institute for Scientific Information’s database yielded nearly 4000 papers over the period 1990–8. About 200 of these were relevant to health technology assessment. Using these as a source for forward and backward searches, and searching other databases (Embase and Medline) and sources, we identified about 300 papers, including about 30 reports of studies taking a fully bayesian perspective. A considerable further number of studies have taken a so called “empirical Bayes” approach, which uses elements of bayesian modelling without giving a bayesian interpretation to the conclusions; these are further mentioned below.

The published studies are dispersed throughout the literature and, apart from one recent collection of papers,⁴ the only textbook which might be considered to be on bayesian methods in health technology assessment focuses on the confidence profile approach.⁵ Published studies are mainly demonstrations of the approach rather than complete assessments, and though many articles advocate bayesian methods, practical take-up seems low.

Findings

Philosophy of the bayesian approach

Bayes’s theorem is a formula that shows how existing beliefs, formally expressed as probability distributions, are modified by new information. Diagnostic testing is a familiar situation to which the theorem can be applied; a doctor’s prior belief about whether a patient has a particular disease (based on knowledge of the prevalence of the disease in the community and the patient’s symptoms) will be modified by the result of the test.⁶

The unknown piece of information may, however, be a somewhat more intangible quantity than an individual’s true diagnosis—for example, the average survival benefit of drug A over drug B in a particular group of patients. Such quantities are not directly observable in any reasonably sized experiment and are considered to be unknown variables. Just as the full evaluation of a diagnostic test requires the prevalence of the disease to be specified, a bayesian analyst is prepared to make the bold step of specifying a probability distribution expressing the relative plausibility for this unknown quantity, before taking into account any evidence from a study. This “prior” distribution can then be combined with evidence from the study to form a “posterior” (formally proportional to the product of the prior and the likelihood function). The box shows an example.

The posterior distribution provides probabilities of events of clinical interest and so one could say, for example, that under specified assumptions “the chance is 15% that drug A improves average survival by at least three months over drug B.” This type of statement is impossible to make within the traditional statistical framework, in which the interpretation of P values and confidence intervals depends on rather convoluted

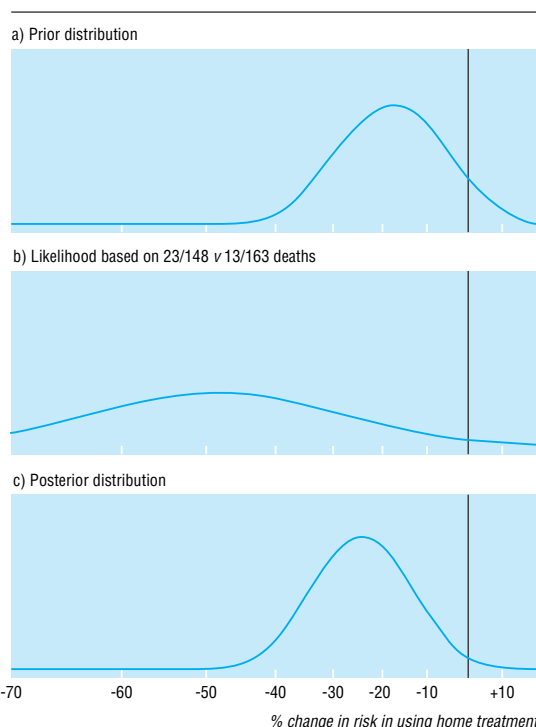


Fig 1 Prior (a), likelihood (b), and posterior (c) distributions arising from reanalysis by Pocock and Spiegelhalter⁷ of the GREAT trial of home thrombolysis.⁸ The prior distribution represents a summary of evidence external to the trial, the likelihood expresses evidence from the trial itself, and the posterior distribution pools these two sources by multiplying the two curves together

statements concerning the long run properties of statistical procedures under null hypotheses.

The table briefly summarises some major distinctions between the bayesian and the traditional

Bayes’s theorem after a randomised trial

Pocock and Spiegelhalter⁷ discuss a small trial of early thrombolytic treatment in preventing deaths from myocardial infarction, which had reported a remarkable 49% reduction in mortality.⁸ On the basis of both published and unpublished large trials, they argued that if treatment were provided two hours earlier “a 15–20% reduction in mortality is highly plausible, while the extremes of no benefit and a 40% reduction are both unlikely.” This opinion could be represented as a prior distribution as shown in figure 1(a), which expresses the relative plausibility arising from this external evidence.

Figure 1(b) shows the “likelihood” for the true risk reduction arising from the trial itself, which is simply proportional to the chance of observing the data (23/148 deaths in controls v 13/163 deaths with active treatment) for each hypothesised risk reduction. Bayes’s theorem states that the two sources of evidence can be combined by multiplying the prior and likelihood curves together and then making the total area under the resulting curve be equal to 1—this is the “posterior” distribution and is shown in figure 1(c). The evidence in the likelihood has been pulled back towards the prior opinion, thus formally representing the suspicion that the trial results were “too good to be true.”

The resulting distribution provides an easily interpretable summary of the total evidence, and posterior probabilities for hypotheses of interest can then be read from the graph. For example, the most likely benefit is a reduction in risk of around 24% (half that observed in the trial), the posterior probability that the risk is reduced by at least 50% is only 5%, and a 95% confidence interval is from 43% to 0% risk reduction. Subsequent experience has reinforced the conclusion of this analysis that it is very unlikely that home thrombolysis reduces mortality by 50%.

Brief comparison of bayesian and frequentist methods in randomised trials

Issue	Frequentist methods	Bayesian methods
Prior information other than that in the study being analysed	Informally used in design	Used formally by specifying a prior probability distribution
Interpretation of the parameter of interest	A fixed state of nature	An unknown quantity which can have a probability distribution
Basic question	"How likely is the data, given a particular value of the parameter?"	"How likely is a particular value of the parameter given the data?"
Presentation of 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	P values and estimates adjusted for the number of analyses	Inference not affected by the 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 (for example, Bonferroni)	Subset effects shrunk towards zero by a "sceptical" prior

approach. The latter is sometimes termed "frequentist" as it is based on the long run frequency properties of statistical procedures. There are many papers summarising the bayesian philosophy and its application to randomised trials: Cornfield's is a notable early example,⁹ and other authors have argued for the flexibility, coherence, and intuitiveness of the approach.^{1-3 10} Several authors have highlighted how the bayesian approach leads naturally into a formal decision theoretical approach to randomised trials.¹¹

Quantifying prior beliefs

The bayesian approach is most controversial when there is no hard evidence for the prior distribution and we have to rely on subjective judgment. This considerably broadens the area of potential application, although the reasonableness of the judgments will need to be justified. The traditional terms prior and posterior may also be misleading, giving the impression that the prior has to be fixed before the evidence is examined. It is more helpful to think of the prior as summarising all external evidence about the quantity of interest—for example, other published studies—which might arise during or after the study that is being considered.

One source of a prior distribution is the pooled subjective opinion of informed experts, which can be elicited interactively by using computer programs¹² or questionnaire methods.¹³ Such opinions should rely on extensive experience: for example, Peto and Baigent state that "it is generally unrealistic to hope for large treatment effects" but that "it might be reasonable to hope that a new treatment for acute stroke or acute myocardial infarction could reduce recurrent stroke or death in hospital from 10% to 9% or 8% ... but not to hope that it could halve in-hospital mortality."¹⁴ This closely mimics the prior opinion used in the box above to illustrate how extreme results based on small studies should not be taken at face value. Another source of prior opinions is, of course, meta-analyses of previous similar studies.

One important use of a prior distribution is in planning the sample size of a randomised trial. Instead of using a single (possibly optimistic) alternative hypothesis as the basis for the power calculation, the prior distribution can be used to produce an "expected power," taking into account reasonable uncertainty about the true treatment effect.¹³

There has been an increasing move towards "off the shelf" priors—for example, those intended to represent the opinions of an archetypal "sceptic" and those of an "enthusiast"¹⁵; these can be used to represent extreme opinions in sensitivity analyses and

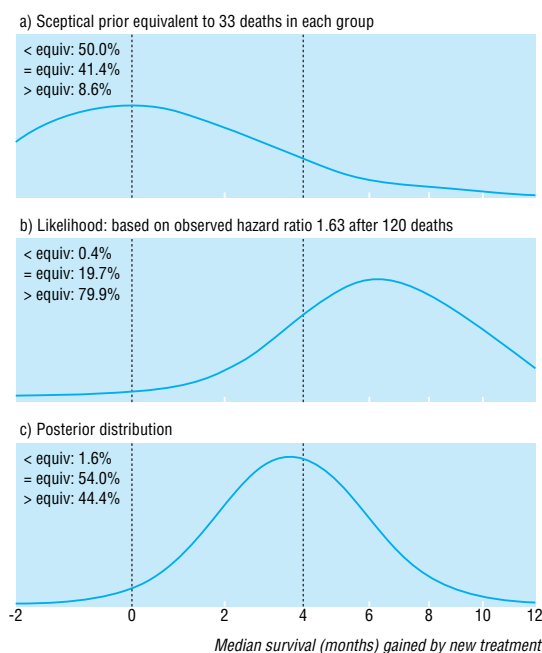


Fig 2 Prior, likelihood, and posterior distributions arising from Cancer and Leukaemia Group B trial of standard radiotherapy versus additional chemotherapy in advanced lung cancer.¹⁵ Dashed lines give boundaries of range of clinical equivalence, taken to be 0 and 4 months median improvement in survival. Numbers by each graph show probabilities of lying below, within, and above the range of equivalence

in sequential monitoring of trials (see below). One published example concerns the use of sceptical priors in determining whether there is sufficient evidence for a treatment to be generally recommended (box).

Applications in monitoring randomised trials

In the traditional frequentist approach, randomised trials are designed to have a fixed chance (usually 5%) of incorrectly rejecting the null hypothesis, and various techniques have been developed for adjusting the apparent significance level of a result to allow for the fact that the data have been analysed more than once. The bayesian approach sees no need for this and instead monitors the trial on the basis of the current posterior distribution, providing an updated summary of the evidence about the treatment effect at the time of any analysis. Several monitoring schemes have been suggested, some of which are based on decision theory.¹¹ The most frequently illustrated technique is simply based on the "tail" areas of the posterior distribution—

for example, stop the trial if the chance that the treatment is more effective than control is greater than 99%.¹⁷ If desired, the probability of the treatment effect being greater than some clinically important difference may be used, or, in the case of equivalence studies, that the treatment difference is less than, say, 10%.

A sceptical prior may be thought of as a handicap that the trial data must overcome in order to provide convincing evidence of benefit. In the light of early positive results, the approach shows a degree of conservatism which can be remarkably similar to that of frequentist stopping rules.¹⁸ The use of sceptical priors has been described in a tutorial and in meta-analyses,^{19 20} and a senior statistician with the US Food and Drug Administration has said that he “would like to see [sceptical priors] applied in more routine fashion to provide insight into our decision making.”²¹

The table also considers predictions made at an interim stage in a randomised trial. Whereas the frequentist conditional power calculations are based on a hypothesised value of the true treatment effect, a bayesian approach can answer a crucial question: if we continue the study, what is the chance we will get a significant result?

Multiplicity—estimating the prior

We often wish simultaneously to carry out a set of related analyses—for example, meta-analysis of individual trial results—allowing for between centre variability in the analysis of a multicentre trial or analysing subsets of cases in a single trial. We call these subanalyses. The traditional frequentist approach tries to maintain a constant probability of wrongly rejecting the null hypothesis (type I error) by some adjustment—for example, a Bonferroni method for multiple comparisons.

The bayesian approach integrates subanalyses by assuming that the unknown quantities (for example, the treatment effects specific to subsets) have a common prior distribution, with the important difference that this prior distribution has unknown parameters that need to be estimated. Such models are known as hierarchical and can, in theory, have any number of levels, although three is generally enough. Non-bayesian versions (multi-level, random effects and random coefficient models) use either likelihood or “empirical Bayes” approaches to estimate the model parameters.

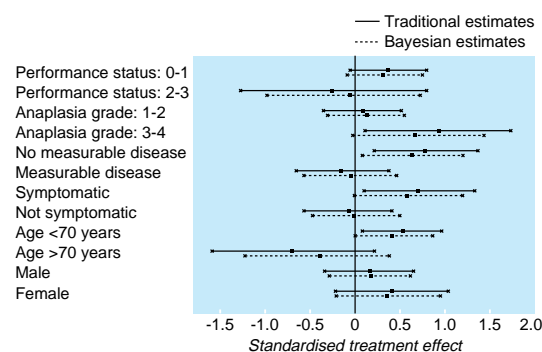


Fig 3 Traditional and bayesian estimates of standardised treatment effects in a randomised trial of treatments for cancer. The bayesian estimates are pulled towards the overall treatment effect by a degree determined by the empirical heterogeneity of the subset results

Is a confirmatory trial necessary?

Parma et al illustrate the use of a sceptical prior distribution in deciding whether or not to perform a confirmatory randomised trial.¹⁶ They discuss a Cancer and Leukaemia Group B trial of radiotherapy and chemotherapy versus standard radiotherapy in patients with locally advanced stage III non-small cell lung cancer. This trial showed an adjusted median improvement in survival of 6.3 months (95% confidence interval 1.4 to 13.3 months) in favour of the new treatment, which has a two sided P value of 0.008. They give two reasons why this might not lead to an immediate recommendation for radiotherapy and chemotherapy as standard treatment. Firstly, the toxicity of chemotherapy might mean a minimum worthwhile improvement is demanded; the authors suggest a figure of around four months. Secondly, a natural scepticism exists about new cancer treatments, derived from long experience of failed innovations.

These two aspects can be formalised within the bayesian framework. Firstly, one can report the probability that the new treatment not only provides a positive improvement but that this exceeds a minimum clinically worthwhile improvement. Secondly, scepticism is expressed by a prior distribution that is centred on zero improvement and shows a 5% chance that the true improvement is greater than the alternative hypothesis in this study—namely, that the true improvement is five months.

Figure 2 shows this sceptical prior distribution, which is equivalent evidence to that of an “imaginary” trial in which 33 patients taking each treatment died. The dashed vertical lines indicate the null hypothesis of no improvement and the minimum clinically worthwhile improvement of four months. Between these lie what can be termed the range of equivalence, and the figure shows that the sceptical prior expresses a probability of 41% that the true benefit lies in the range of equivalence and only 9% that the new treatment is clinically superior.

The likelihood function shows the inferences to be made from the data alone, assuming a “uniform” prior on the range of possible improvements; Parma et al call this an enthusiastic prior. The probability that the new treatment is actually inferior is 0.4% (equivalent to the one sided P value of $0.008 \div 2$). The probability of clinical superiority is 80%, which might be considered sufficient to change treatment policy.

The posterior distribution shows the impact of the sceptical prior, in that the chance of clinical superiority is reduced to 44%, hardly sufficient to change practice. In fact, Parma et al report that the National Cancer Institute intergroup trial investigators were unconvinced by the Cancer and Leukaemia Group B trial due to their previous negative experience, and so carried out a further study. They found a significant median improvement, but of only 2.4 months, suggesting that the sceptical approach might have given a more reasonable estimate.

By assuming a common prior distribution for each subanalysis we are expressing scepticism about large differences in their outcomes, although the precise degree of similarity is generally considered unknown and estimated from the data—for example, by measuring the between trial variability in a meta-analysis. Full bayesian and empirical Bayes approaches can lead to similar conservatism (box).²²

Non-randomised studies and synthesis of evidence

Most authors have concentrated on the application of bayesian methods when designing randomised trials or pooling results from published trials, but a small number of papers have considered applying these methods to data collected from non-randomised studies. For example, in a paper analysing data from two case-control studies (one being very small) and a cohort study, the authors show the results of using different sources of information for the prior and likelihood.²⁴ Other authors have discussed the integration of evidence from several types of non-randomised studies²⁵ and the integration of findings from both randomised and non-randomised studies within a bayesian framework.²⁶

Bayes's theorem for subset analysis

Dixon and Simon describe a bayesian approach to dealing with subset analysis in a randomised trial in advanced colorectal cancer.²⁵ The solid horizontal lines in figure 3 show the standardised treatment effects within a range of subgroups, using traditional methods for estimating treatment by subgroup interactions. Four of the 12 intervals exclude zero; because multiple hypotheses are being tested, however, an adjustment technique such as Bonferroni might be used to decrease the apparent statistical significance of these findings.

The bayesian approach is to assume that deviations from the overall treatment effect that are specific to subgroups have a prior distribution centred at zero but with an unknown variability; this variability is then given its own prior distribution. Since the degree of scepticism is governed by the variance of the prior distribution, the observed heterogeneity of treatment effects between subgroups will influence the degree of scepticism being imposed.

The resulting bayesian estimates are shown as dashed lines in figure 3. They tend to be pulled towards each other, owing to the prior scepticism about substantial interaction effects between subgroups and treatments. Only one 95% confidence interval now excludes zero, that for the subgroup with no measurable metastatic disease. Dixon and Simon mention that this was the conclusion of the original trial; the bayesian analysis has the advantage of not relying on somewhat arbitrary adjustment techniques as it can be generalised to any number of subsets, and it provides a unified means of both providing estimates and tests of hypotheses.

Decision making

Another important feature of a bayesian approach is the way in which the resulting posterior probability distribution can be combined with quantitative measures of utility as part of a formal decision analysis. As with the elicitation of beliefs regarding probabilities, the elicitation and quantification of utilities is challenging, and this is one of the least developed areas of bayesian analysis. Such formal uses of decision theory have been applied in health technology assessments in various settings, including the development of clinical recommendations for prevention of stroke,²⁷ monitoring and analysis in randomised trials,¹¹ and assessment of environmental contamination on public health.²⁸

Recommendations

Bayesian analysis is widely used in a variety of non-medical fields, including engineering, image processing, expert systems, decision analysis, gene sequencing, financial predictions, and neural networks, and increasingly in complex epidemiological models. Health technology assessment has been slow to adopt bayesian methods; this could be due to a reluctance to use prior opinions, unfamiliarity, mathematical complexity, lack of software, or conservatism of the health care establishment and, in particular, the regulatory authorities.

There are strong philosophical reasons for using a bayesian approach, but the current literature emphasises the practical advantages in handling complex interrelated problems and in making explicit and accountable what is usually implicit and hidden, thereby clarifying discussions and disagreements. Perhaps the most persuasive reason is that the analysis tells us what we want to know: how should this piece of evidence change what we currently believe?

The perceived problems with the bayesian approach largely concern the source of the prior and the interpretations of the conclusions. There are also practical difficulties in implementation and software. Current international guidelines for statistical sub-

missions to drug regulatory authorities state that "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,"²⁹ and it seems sensible that experience should be gained in the use of bayesian approaches in health technology assessment in parallel with traditional approaches, with careful consideration of the sensitivity of results to prior distributions.

For future practical and methodological developments, we recommend:

- An extended set of case studies showing practical aspects of the bayesian approach, in particular for prediction and handling multiple substudies, in which mathematical details are minimised;
- The development of standards for the performance and reporting of bayesian analyses;
- The development and dissemination of software for bayesian analysis, preferably as part of existing programs.

This article is adapted from *Health Services Research Methods: A Guide to Best Practice*, edited by Nick Black, John Brazier, Ray Fitzpatrick, and Barnaby Reeves, published by BMJ Books.

Competing interests: None declared.

- 1 Kadane JB. Prime time for Bayes. *Contr Clin Trials* 1995;16:313-8.
- 2 Breslow N. Biostatistics and Bayes. *Stat Sci* 1990;5:269-84.
- 3 Lilford RJ, Braunholtz D. The statistical basis of public policy—a paradigm shift is overdue. *BMJ* 1996;313:603-7.
- 4 Berry DA, Stangl DK. *Bayesian biostatistics*. New York: Dekker, 1996.
- 5 Eddy DM, Hasselblad V, Shachter R. *Meta-analysis by the confidence profile method: the statistical synthesis of evidence*. San Diego, CA: Academic, 1992.
- 6 Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*, 2nd ed. Boston: Little Brown, 1991.
- 7 Pocock S, Spiegelhalter DJ. Domiciliary thrombolysis by general practitioners. *BMJ* 1992;305:1015.
- 8 GREAT Group. Feasibility, safety and efficacy of domiciliary thrombolysis by general practitioners: Grampian region early anistreplase trial. *BMJ* 1992;305:548-53.
- 9 Cornfield J. Recent methodological contributions to clinical trials. *Am J Epidemiol* 1976;104:408-21.
- 10 Lewis RJ, Wears RL. An introduction to the Bayesian analysis of clinical trials. *Ann Emergency Med* 1993;22:1328-36.
- 11 Berry DA, Wolff MC, Sack D. Decision-making during a phase-III randomised controlled trial. *Contr Clin Trials* 1994;15:360-78.
- 12 Chaloner K, Church T, Louis TA, Matts JP. Graphical elicitation of a prior distribution for a clinical trial. *Statistician* 1993;42:341-53.
- 13 Parmar MKB, Spiegelhalter DJ, Freedman LS. The chart trials: bayesian design and monitoring in practice. *Stat Med* 1994;13:1297-312.
- 14 Peto R, Baigent C. Trials: the next 50 years. *BMJ* 1998;317:1170-1.
- 15 Spiegelhalter DJ, Freedman LS, Parmar MKB. Bayesian approaches to randomised trials. *J Roy Stat Soc Series A* 1994;157:357-87.
- 16 Parmar MKB, Ungerleider RS, Simon R. Assessing whether to perform a confirmatory randomised clinical trial. *J Natl Cancer Inst* 1996;88:1645-51.
- 17 Freedman LS, Spiegelhalter DJ, Parmar MKB. The what, why and how of bayesian clinical trials monitoring. *Stat Med* 1994;13:1371-83.
- 18 Grossman J, Parmar MKB, Spiegelhalter DJ, Freedman LS. A unified method for monitoring and analysing controlled trials. *Stat Med* 1994;13:1815-26.
- 19 Fayers PM, Ashby D, Parmar MKB. Bayesian data monitoring in clinical trials. *Stat Med* 1997;16:1413-30.
- 20 Dersimonian R. Meta-analysis in the design and monitoring of clinical trials. *Stat Med* 1996;15:1237-48.
- 21 O'Neill R. Early stopping rules workshop: conclusions. *Stat Med* 1994;13:1493-9.
- 22 Louis TA. Using empirical Bayes methods in biopharmaceutical research. *Stat Med* 1991;10:811-29.
- 23 Dixon DO, Simon R. Bayesian subset analysis in a colorectal cancer clinical trial. *Stat Med* 1992;11:13-22.
- 24 Ashby D, Hutton J, McGee M. Simple bayesian analyses for case-control studies in cancer epidemiology. *Statistician* 1993;42:385-97.
- 25 Eddy DM, Hasselblad V, Shachter R. An introduction to a bayesian method for meta-analysis—the confidence profile method. *Medical Decision Making* 1990;10:15-23.
- 26 Abrams KR, Jones DR. Meta-analysis and the synthesis of evidence. *IMA J Math Med Biol* 1995;12:297-313.
- 27 Parmigiani G, Ancukiewicz M, Matchar D. Decision models in clinical recommendations development: the stroke prevention policy model. In: Berry A, Stangl DK, eds. *Bayesian biostatistics*. New York: Dekker, 1996:207-33.
- 28 Wolfson LJ, Kadane JD, Small MJ. Expected utility as a policy-making tool: an environmental health example. In: Berry A, Stangl DK, eds. *Bayesian biostatistics*. New York: Dekker, 1996:261-77.
- 29 International Conference on Harmonisation. Statistical principles for clinical trials, 1998. www.ich.org/pdf/pma/e9.pdf; accessed May 1999.