
Article:
Stevens, J.W. and O'Hagan, A. (2002) Incorporation of genuine prior information in cost-effectiveness analysis of clinical trial data. International Journal of Technology Assessment in Health Care, 18 (4). pp. 782-790. ISSN 0266-4623

https://doi.org/10.1017/S0266462302000594

International Journal of Technology Assessment in Health Care, **18:4** (2002), 782–790.

INCORPORATION OF GENUINE PRIOR INFORMATION IN COST-EFFECTIVENESS ANALYSIS OF CLINICAL TRIAL DATA

John W. Stevens
AstraZeneca R&D Charnwood

Anthony O'Hagan
University of Sheffield

Abstract

The Bayesian approach to statistics has been growing rapidly in popularity as an alternative to the frequentist approach in the appraisal of heathcare technologies in clinical trials. Bayesian methods have significant advantages over classical frequentist statistical methods and the presentation of evidence to decision makers. A fundamental feature of a Bayesian analysis is the use of prior information as well as the clinical trial data in the final analysis. However, the incorporation of prior information remains a controversial subject that provides a potential barrier to the acceptance of practical uses of Bayesian methods. The purpose of this paper is to stimulate a debate on the use of prior information in evidence submitted to decision makers. We discuss the advantages of incorporating *genuine* prior information in cost-effectiveness analyses of clinical trial data and explore mechanisms to safeguard scientific rigor in the use of such prior information.

Keywords: Bayesian statistics, Cost-effectiveness, Clinical trials, Prior information

The analytical appraisal of healthcare technologies is becoming increasingly important as a means of assessing the clinical and cost-effectiveness of clinical interventions. In the analysis of data that has been collected in a clinical trial, it is common to use frequentist statistical methods. The frequentist approach (also known as the classical approach) is the familiar one based on p-values, confidence intervals, and unbiased estimators. Standard frequentist methods in cost-effectiveness analysis are based on asymptotic normality of sample means or the nonparametric bootstrap, used to derive estimates and confidence intervals for cost-effectiveness ratios. However, it is questionable whether the information that has been produced has been of substantial value to decision makers in choosing between healthcare technologies. Indeed, there is an increasing view that this may in part be a direct result of the application of frequentist statistical methods (15), and that Bayesian methods provide a more natural framework for quantifying uncertainty and for decision making (2).

The authors gratefully acknowledge an anonymous referee for useful comments on an earlier version of this manuscript.

The arguments made in favor of the Bayesian approach are that they offer more intuitive and meaningful inferences, that they give the ability to tackle more complex problems, and that they allow the use of prior information in addition to the clinical trial data. These arguments are as sound and persuasive in cost-effectiveness analysis as they are in other fields (21).

A fundamental aspect of a Bayesian analysis is the use of prior information as well as clinical trial data in the final analysis. In general, advocates of frequentist statistical methods argue that prior information is intrinsically subjective and therefore has no place in science. They point out that an unscrupulous analyst can concoct any desired result by the creative specification of prior distributions for the parameters in the model. However, the potential for manipulation is not unique to Bayesian statistics. The scientific community and regulatory agencies have developed sophisticated safeguards and guidance to avoid conscious or unconscious biases. An example is the insistence on double-blind randomized clinical trials for the rigorous comparison of interventions. This—and similar requirements for the statistical analysis protocol to be established before the trial begins (10)—is necessary to obviate the potential for manipulation or bias that already exists in the use of frequentist statistics.

In our opinion, Bayesian methods have significant advantages over classical frequentist statistical methods in the appraisal of healthcare technologies and the presentation of evidence to decision makers. In this context, the purpose of this paper is to discuss the advantages of incorporating *genuine* prior information in cost-effectiveness analyses of clinical trial data and to explore mechanisms to safeguard scientific rigor in the use of prior information.

INCORPORATION OF PRIOR INFORMATION

A Bayesian analysis synthesizes two sources of information about the unknown parameters of interest. The first of these is the sample data, expressed formally by the likelihood function. The second is the prior distribution, which represents other information available to the investigator. Whereas the likelihood function is fundamental also to frequentist inference, the prior distribution is used only in the Bayesian approach. If we represent the data by the symbol D and denote by θ the set of unknown parameters, then the likelihood is $f(D \mid \theta)$, the probability that data D will be observed conditional on the true value of the parameter θ. The prior distribution is then a probability density function $\pi(\theta)$, giving the probability that θ takes any particular value based on whatever other prior information might be available to the investigator. Bayes' theorem synthesizes these two sources of information through the equation:

$$p(\theta \mid D) \propto f(D \mid \theta)\pi(\theta). \quad \textbf{(1)}$$

As usual, the proportionality symbol '$\propto$' expresses the fact that the product of the likelihood and prior density on the right side of equation (1) must be scaled to integrate to one over the range of plausible θ values. The scaled product is then the posterior density $p(\theta \mid D)$, which expresses what is now known about θ based on both sample and prior information. Notice that if for some value of θ the likelihood in the right side of equation (1) is small, so that the data suggest that this value of θ is implausible, then the posterior distribution will also give small probability to this θ value. Similarly, if for some value of θ the prior distribution in the right side of equation (1) is small, so that the prior information suggests that this value of θ is implausible, then again, the posterior distribution will also give small probability to this θ value. In general, the posterior probability will be high for some θ only when *both* information sources support that value. The simple and intuitive nature of

Bayes' theorem as a mechanism for synthesizing information is an attractive feature of the Bayesian method.

In a cost-effectiveness analysis of a clinical trial, the data D will comprise observations that relate to the costs and efficacies, as experienced by individual patients under the two treatments. The parameter θ will comprise the unknown true mean costs and efficacies of the two treatments, averaged over the whole of the relevant patient population, or perhaps other unknown parameters from which these means may be constructed. The prior distribution represents information about the distributions of costs and efficacies under the two treatments that is available *prior* to (or, more generally, in addition to) observing the data D.

Using this framework, we are now able to make any inferences that we wish from the *posterior* distribution representing everything that we know about the unknown parameters in the model. If the amount of data that are available in an analysis is large, then the data will provide the majority of the evidence and will dominate the prior distribution, in the sense that the posterior distribution will be almost identical to the likelihood, and therefore the posterior conclusions will be almost entirely driven by the data, rather than the prior information. When sample sizes are much smaller and there is greater uncertainty in the trial data, the incorporation of prior information will play a much greater role, so that the appropriate specification of *genuine* prior information becomes much more critical. Genuine prior information will discriminate between different possible values of the parameters, since the information will suggest that some values or ranges of values are much more likely than others. Genuine prior information will be represented by an *informative* prior distribution that gives higher probability to some values than to others. This contrasts with a Bayesian analysis using a so-called noninformative prior, which is discussed in the next section. Genuine prior information would represent all available evidence that has been formally synthesized into probability distributions, and there is a growing body of literature on Bayesian methods for evidence synthesis (22).

In addition to prior information about means and variances, genuine prior information can be *structural*, where we assert, for instance, that two parameters should have similar values or that one should be larger than the other. In some situations, we may feel confident about expressing structural prior information while having relatively much weaker prior knowledge of the numerical values of the parameters. Such prior information can nevertheless play a crucial role in the analysis, as described later.

NONSUBJECTIVE PRIOR INFORMATION

We have highlighted the main objections of subjectivity and the attendant risks of bias, prejudice, and manipulation that have been made against the use of prior information in a Bayesian analysis. Some authors suggest that a compromise between a fully Bayesian analysis and a frequentist approach that does not use prior information is to use a noninformative prior distribution representing a state of prior ignorance (4). In so doing, we would obtain some of the benefits of a Bayesian approach, particularly that the results are presented in the intuitive way in which one would like to make inferences, without including any element of subjectivity in the analysis.

Another suggestion, in the context where an experimental treatment is being compared with a current or standard treatment, is to specify skeptical prior distributions for the parameters in the model (9;12). In this context, the prior distribution is specified in such a way that it automatically favors the standard treatment. Again, it is proposed that skeptical prior distributions should be expressed in a formalized, nonsubjective way.

Such proposals are tempting, particularly in a regulatory framework where rigorous standards and safeguards are demanded. However, both ideas suffer from serious objections, and both fail to exploit the full potential of the Bayesian approach.

A key objection is that there is no unique way to implement either idea, and hence subjectivity is not removed. There is a considerable body of work in the technical Bayesian statistics literature attempting to define prior distributions to represent ignorance, known variously as noninformative, weak, or reference prior distributions (13). For all but the simplest situations, there is no general agreement on which formulation to use, and answers depend on the (subjective) choice of parameterization. There can be even less agreement over what constitutes a skeptical prior distribution. In our opinion, there are a number of misconceptions underlying these proposals. In particular, it is simply naïve to suppose that the prior distribution in a Bayesian analysis is the only place where subjective judgments are in danger of entering the analysis. Any statistical model, whether formulated for a frequentist or Bayesian analysis, is a matter of subjective judgment, and it is commonplace that different statisticians make different choices. Furthermore, the choice of which estimator, significance test, or confidence interval to employ is a subjective matter in frequentist statistics. Such subjectivity is reduced but not removed by the requirement that the basic form of the analysis should be prespecified. There is no such choice to make in Bayesian statistics, since once the posterior distribution has been obtained, there is a unique (objective) answer to any properly specified question about the parameters. Indeed, there is a fundamental sense in which the frequentist choice of an estimator, test, or confidence interval is comparable to the choice of a prior distribution.

The idea of skeptical prior distributions confuses two reasons for caution in adopting a new treatment. The first is a natural skepticism of the value of new treatments and the possible vested interests of those proposing them. But the degree of skepticism that we feel is not an absolute and should depend on the context. A new treatment that is well based in accepted science will be accorded less skepticism than a new drug whose mode of action is not fully understood. This skepticism should in fact be a natural part of the investigator's prior knowledge, and anyone bringing forward an analysis using *genuine* prior information should demonstrate that due caution and skepticism have been exercised.

The other reason for caution, however, is the considerable cost to society that is entailed in switching to a new treatment. Introducing a treatment that is subsequently found to be less effective than the one it replaced will be seen as a serious error. For this reason, we may demand a relatively high standard of proof of the value of new treatments. This is a quite separate matter from the prior information, but concerns how we derive inferences or decisions from the posterior distribution. Analogously, in a frequentist analysis it might concern the choice of p-value but not the choice of test.

It is a strength of the Bayesian approach that it is possible to separate the natural skepticism that should be a part of formulating one's prior distribution from the caution that should be exercised as part of decision making.

Concerning the first proposal, it is questionable whether a state of complete prior ignorance would ever exist. In a clinical trial comparing the cost-effectiveness of two treatments, there must be some information available to the sponsor and the investigators that supports the decision for further investigation, and this information is available for use in a fully Bayesian analysis. Indeed, it should actually be regarded as wasteful not to do so.

We therefore argue that these proposals are misguided, and that where *genuine* prior information exists, it should be used to maximize the information from a cost-effectiveness analysis. The risks associated with subjectivity remain, however, and must be addressed. This is particularly true in matters of regulation and public policy. The specification of a prior distribution in such a context must be transparent and defensible.

The only meaningful and useful way to specify prior information is through the genuine elicitation of prior beliefs. Similar to the requirements of a frequentist analysis, what is required is a formal process of elicitation that is documented and justified *before* the study starts. Decision makers who want to exercise further caution can do so by requiring a

sufficiently high posterior probability that the new treatment is more cost-effective than the standard treatment.

ADVANTAGES OF USING PRIOR INFORMATION

Clinical trials are generally powered to detect clinically relevant treatment effects and often involve many centers across several countries. Different clinical practices across different countries may lead to very different patterns of resource usage so that the pooling of resource usage and cost data is not straightforward. A simple solution to this problem in cost-effectiveness analyses is to report data from each country separately, although the ability of the trial data to allow conclusions to be reached will now be much reduced. The ability to incorporate prior information in a cost-effectiveness analysis is a positive benefit of the Bayesian approach to statistical inference. It allows the analysis to make use of *all* available information in addition to the clinical trial data, and hence to reach stronger, or more generally, more realistic conclusions than would be available from a frequentist analysis.

When designing a cost-effectiveness trial, there will always be appreciable prior information, at least in regard to efficacy, because by the time the decision is made to conduct such a trial at Phase III or Phase IV, there will already be other clinical trial data available. Such data would usually be supportive of the value in continuing with the development of the drug, or else such investment would not be made.

The prior information would almost inevitably favor the new treatment as being more effective than the comparator and, if this information is to be utilized, a Bayesian framework is essential. This seems to imply the opposite of the idea of a skeptical prior, where we automatically favor the standard treatment rather than the new. The extent to which it can be legitimate to incorporate such prior information in a formal claim of clinical or cost-effectiveness needs to be explored. However, as we shall see later, very natural and defensible structural information can be more influential than numerical information.

ELICITATION OF PRIOR INFORMATION

A good Bayesian analysis will incorporate *genuine* prior information based on all that is known about the parameters in the statistical model. In the absence of an adequate process for the elicitation of prior information representing genuine prior beliefs, it is reasonable to be critical about the inclusion of prior information in an analysis.

At the present time, there is limited guidance or research on the elicitation of prior information. The literature on elicitation techniques is small but growing, both in the context of clinical trials (6;7;8;19) and more generally (11;17). However, there is still a need to provide clear methodology to assist practitioners in formulating the prior distributions that represent *genuine* prior information.

Given the relatively standardized process of drug development through Phases I to IV, there may be scope for equally standardized procedures to bring forward information from earlier phases into cost-effectiveness trials. In time, and with sound examples of Bayesian analysis incorporating *genuine* prospectively elicited prior beliefs, the true value of Bayesian methods can be demonstrated.

WHOSE PRIOR INFORMATION

The analysis of a clinical trial, whether to prove efficacy or to demonstrate cost-effectiveness, will be presented to regulatory authorities or decision makers external to the company. In the current environment, to present a Bayesian analysis in which the sponsor's own prior

beliefs are used to augment the trial data will, in general, not be acceptable to an external agency.

Drug regulatory authorities generally start from a position of skepticism when reviewing a submission for the approval of a new treatment based on safety and efficacy. The results of a confirmatory clinical trial are expected to be robust and "to estimate with due precision the size of effects attributable to the treatment of interest" (10). In most countries the decision to recommend a new treatment as being more cost-effective is usually the responsibility of agencies other than drug regulatory authorities, and is not generally the subject of formal regulation as with drug regulatory approval. However, there is a clear movement toward tight regulatory frameworks for cost-effectiveness that mirror those already in place for clinical safety and efficacy (1;3;5;16).

Simple acceptance of the sponsor's own unsupported prior opinions is clearly out of the question. What needs to be explored is the extent to which the sponsor's prior information, elicited prospectively in a rigorous and transparent way, might be acceptable.

No matter how well supported the sponsor's prior distribution may be, it seems likely that agencies will wish to see the results of analyses of the data using other prior distributions. So-called noninformative prior distributions have a role to play, but not for the purpose of dismissing genuine prior information. Their role will be to help the agency to gauge the impact of the prior information. Prior distributions that discount the sponsor's prior information or introduce further opinions from independent clinicians are other possibilities. It may be that emerging guidelines will require sponsors to elicit prior information from independent experts, who might even be nominated by the regulator.

We believe that, with the rapidly growing interest in Bayesian methods in health technology assessment, regulatory agencies need to address these questions in a dialogue with professional statisticians in industry and academia. Indeed, it is this belief that has primarily motivated our writing of this article.

AN EXAMPLE OF STRUCTURAL PRIOR INFORMATION

The authors have earlier presented (18) a Bayesian analysis of the U.K. data from the multicenter TACTIC trial (20), which compared two forms of inhaler—pressurized metered dose inhaler (pMDI) and Turbuhaler®—in the treatment of asthma.

In this analysis, we considered the simple binary efficacy outcome of whether a patient experienced one or more exacerbations during the trial period (negative outcome) or no exacerbations (positive outcome). Cost data were also available at the patient level. The cost data were typically highly skewed, and it was more reasonable to assume that they were log-normally distributed rather than normally distributed, an assumption that was supported by simple plots of the data.

In addition, two patients in the pMDI treatment group had very high costs, although neither patient had any exacerbations during the trial. These costs resulted in substantially higher variances for patients in the pMDI group who did not have exacerbations, even after a log transformation. Now, although there was no prior information about costs, we would not expect the variances of log-costs to be very different between the two treatments or for patients with or without exacerbations. In effect, we argued that the two extreme observations were indicative of the highly skewed nature of costs generally, and suggested that one might expect extreme costs in each treatment group irrespective of whether the patients experienced exacerbations. It is possible that extreme costs only occurred for patients in the pMDI group who did not have exacerbations purely because of random sampling, owing to the relatively small sample sizes, and that, with larger sample sizes, extreme costs may also have occurred in any one of the other categories. Consequently, based on this prior belief, we formulated a prior distribution that gave generally rather weak numerical prior

information about the parameters but incorporated the structural information that the ratios of variances of log-costs were likely to be reasonably close to unity.

The effect of the prior information was to shrink the posterior estimates of variance closer together. In fact, the largest shift was in the variance of log-costs for patients in the pMDI group who did not have exacerbations, which is where the extreme observations were. The detailed analysis found a probability of 0.7 that Turbuhaler® is cheaper than pMDI with weak prior information, whereas this probability moved to below 0.5 when we added prior information about the variances of log-costs. The sensitivity of this probability to the incorporation of the structural prior information is marked.

We conducted a further analysis using a noninformative prior distribution. In principle, this favors neither treatment and allows the trial data to determine the posterior distribution. In this example, however, that approach turned out to be more favorable to Turbuhaler® than our full prior specification. The structural prior information plays a key role in this analysis to moderate the random effect of where extreme costs fall within the data. Although in this case it has led to favoring pMDI, because it was in the pMDI group that the extreme costs occurred, the structural information is neutral and could just as easily have favored Turbuhaler® had the data turned out differently.

We believe that these data and the analyses using various prior distributions offer an important message regarding the legitimate use of prior information. Furthermore, we argue that the appropriate analysis of these data required the use of structural prior information, and so could *only* have been obtained through Bayesian methods. Similar approaches for combining different sources of evidence across studies are used in Bayesian meta-analysis (22).

We emphasize here that the data analyzed in this example are not the full set of data arising from the trial. The cost-effectiveness of Turbuhaler® was established clearly using the Canadian data (14), which comprised the majority of all cases, so that this analysis is for exploratory purposes only.

CONCLUSIONS

There is an increasing requirement to quantify the uncertainty associated with the cost-effectiveness of new or established healthcare technologies. The conventional approach is to use frequentist methods, although the Bayesian approach to statistical inference provides a more natural framework for presenting such information. Genuine prior information can provide more realistic conclusions, particularly where sample sizes are relatively small, as is often the case in cost-effectiveness analyses. As we have also seen, with the use of structural prior information we can also produce more meaningful inferences that do not suffer from the criticism of subjectivity that is often made of Bayesian analyses.

There is still a considerable amount of research to be done on the elicitation and specification of prior information in cost-effectiveness analyses. It is particularly important for guidance to be developed concerning the use of Bayesian methods and genuine prior information in the formal assessment of technologies by healthcare providers and agencies. The transition from an academic appreciation to genuine submissions of evidence using a Bayesian approach will only happen when there is an acceptance of the value of genuine numerical and structural prior information. We hope that this article will stimulate a debate on this topic and the development of formal procedures for the elicitation and incorporation of prior information in evidence submitted to decision makers.

POLICY IMPLICATIONS

Bayesian methods provide a more natural framework for quantifying uncertainty and for decision making, and there is a growing interest is using such methods in submissions of

cost-effectiveness. A fundamental aspect of a Bayesian analysis is the use of prior information, and decision makers should regard this as a useful contribution to the overall evidence available to them when choosing between healthcare interventions. Guidelines should be developed that provide recommendations for the elicitation process and the synthesis of such information into probability distributions. Submissions of evidence on the cost-effectiveness of new interventions using the Bayesian approach must include supporting documentation that demonstrate clearly that a formal process of elicitation has been followed if the prior information is to be accepted as credible. Analysts involved in cost-effectiveness submissions must develop their familiarity with the Bayesian approach to statistics as well as their skills in the elicitation of prior information.

REFERENCES

1. Academy of Managed Care Pharmacy. A format for submission of clinical and economic data in support of formulary consideration by managed health care systems in the United States. Available at: http://www.amcp.org/publications.
2. Ashby D, Smith AFM. Evidence-based medicine as Bayesian decision-making. *Stat Med.* 2000;19:3291-3202.
3. Australia Commonwealth Department of Human Services and Health. *Guidelines for the pharmaceutical industry on the preparation of submissions to the Pharmaceutical Benefits Advisory Committee: Including major submissions involving economic analysis*. Canberra: Australia Government Publishing Service; 1995.
4. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. *Health Econ.* 1999;8:257-261.
5. Canadian Coordinating Office for Health Technology Assessment. *Guidelines for economic evaluation of pharmaceuticals: Canada*. 2nd ed. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 1997.
6. Chaloner KM, Church T, Louis TA, et al. Graphical elicitation of a prior distribution for a clinical trial. *The Statistician.* 1993;42:341-353.
7. Chaloner KM, Rhame F. Quantifying and documenting prior beliefs in clinical trials. *Stat Med.* 2001;20:581-600.
8. Fayers PM, Cuschieri A, Fielding J, et al. Sample size calculation for clinical trials: The impact of clinician beliefs. *Br J Cancer.* 2000;82:213-219.
9. Heitjan DF. Bayesian interim analysis of Phase II cancer clinical trials. *Stat Med*. 1997;16:1791-1802.
10. ICH harmonised tripartite guideline: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Statistical principles for clinical trials 1998. ICH.E9. Available at: http://www.ifpma.org/ich1.html.
11. Kadane JB, Wolfson LJ. Experiences in elicitation. *The Statistician*. 1998;47:3-19 (discussion: 55–68).
12. Kass RE, Greenhouse JB. Comments on "Investigating therapies of potentially great benefit: ECMO" by JH Ware. *Stat Sci*. 1989;4:310-317.
13. Kass RE, Wasserman L. The selection of prior distributions by formal rules. *JAMA*. 1996;91:1343-1370.
14. Liljas B, Ståhl, Pauwels RA. Cost-effectiveness analysis of dry-powder inhaler (Turbuhaler®) versus a pressurized metered dose inhaler in patients with asthma. *PharmacoEconomics*. 1997;12:267-277.
15. Luce BR, Shih Y-CT, Claxton K. Introduction: Bayesian approaches to technology assessment and decision making. *Int Technol Assess Health Care*. 2001;17:1-5.
16. National Institute for Clinical Excellence. Technical guidance for manufacturers and sponsors. 2001. Available at: http://www.nice.org.uk/.
17. O'Hagan A. Eliciting expert beliefs in substantial practical applications. *The Statistician*. 1998;47:21-35 (discussion: 55–68).
18. O'Hagan A, Stevens JW. A framework for cost-effectiveness from clinical trial data. *Health Econ*. 2001;10:302-315.

19. Parmar MKB, Spiegelhalter DJ, Freedman LS. The CHART trials: Bayesian design and monitoring in practice. *Stat Med.* 1994;13:1297-1312.
20. Pauwels RA, Hargreave FE, Camus P, et al. A one-year comparison of Turbuhaler® versus pressurized metered-dose inhaler (pMDI) in asthmatic patients. *Chest.* 1996;110:53-57.
21. Sheingold SH. Can Bayesian methods make data and analyses more relevant to decision makers? A perspective from Medicare. *Int J Technol Assess Health Care.* 2001;17:114-122.
22. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res.* 2001;10:227-303.