

## ORIGINAL ARTICLE

# Clinical trials and haemophilia: does the Bayesian approach make the ideal and desirable good friends?

A. IORIO and M. MARCUCCI

*Internal and Vascular Medicine – Haemophilia Centre, University of Perugia, Perugia, Italy*

**Summary.** When the disease is rare and/or the outcome is uncommon the trial design does not warrant precise and unbiased estimates due to a lack of power or the expected length of recruitment and observation periods. Is there any reliable method to control for bias and consequently achieve an advantage from estimates generated by different study designs? An interesting statistical approach suitable to solve this problem has been theorized by Thomas Bayes. A Bayesian analysis is aimed at answering the question ‘How this trial will modify our belief about that treatment effect?’ In summary, the Bayesian approach can be defined as the explicit and quantitative use of any kind of external evidence in the design, analysis, and interpretation of an experimental trial. The results of a Bayesian analysis is the 95% credible interval in which we believe the

estimate to lie with a probability of 95%, or the estimate of the probability that the quantity of interest is less than a specific value. The principal advantages of the Bayesian approach are that it allows to directly make probability statements about quantities of interest; it allows to easily make predictive statements, conditional on the current state of knowledge; it enables evidence from a variety of sources to be taken into account within a coherent modelling framework; it requires the investigator to explicit prior beliefs and demands. Exemplifications of the advantages of the Bayesian approach will be given discussing some papers published in *Haemophilia*.

**Keywords:** Bayesian, clinical trial, methodology, rare bleeding disorders, systematic reviews

## Introduction

The usual presentation for the simplest clinical question, both in research and clinical practice, is the occurrence of two competing treatments with no definitive evidence of superiority of one over the other: a condition called clinical equipoise [1].

To escape the impasse, the researcher has to set up and run the experimental trial: he first decides on the worthwhile treatment effect to be proved (i.e. 30% less bleedings with A as compared to B) and then he measures the *relative effect* to the *required precision* with the *required power*. This is all a researcher should care about, here is our answer. The clinician now easily prescribes the winning treatment, and there is not any residual equipoise left.

When disease and outcome are both common, it is plausible to get a result by applying this standard approach, i.e. the randomized controlled trial (RCT). The RCT is the best application to clinical research of the frequentistic approach. In simple words, what you demonstrate with the help of statistics is only that treatment X produces effects unlikely to be due to chance. Unfortunately (or may be fortunately), in most of the cases where a clinician has to make a choice, no frequentistic proven evidence is there to help. A typical example is when the disease is rare and/or the outcome uncommon. In this setting, in fact, the trial design does not warrant precise and unbiased estimates due to lack of power or expected length of recruitment and observation periods [2].

Here is, finally, the key question: do we have any gain from potentially biased estimates? Is there any reliable method to control for bias and consequently get advantage from estimates generated by different study designs? Is there any method to control for bias other than doing an RCT? We are going to speculate on potential answer to these questions, taking the cue

Correspondence: A. Iorio, Internal and Vascular Medicine – Haemophilia Centre, University of Perugia, Via G. Dottori 1, 06129, Perugia, Italy.

Tel.: +39 349 120 4919; fax: +39 075 578 2309;

e-mail: iorioa@unipg.it

Accepted after revision 26 February 2009

from the publication of a Bayesian paper in this issue of *Haemophilia* [3].

Our main aim is to introduce the reader to the advantages of the Bayesian approach, taking into special account haemophilia and rare bleeding disorders. To reach our goal, we also are going to briefly review two additional papers published on this Journal [4,5] as practical examples of the Bayesian method application.

### The Bayesian approach to health care research

Luckily indeed, the frequentistic approach is not the only one we can rely upon to make reasonable choices, and, moreover, it may not be the best one. A different statistical approach to the 'natural' clinical reasoning has been theorized by Thomas Bayes [6] and proposed as an alternative to the frequentistic method in health research.

The novelty in the idea of the Bayesian approach can be told through discussion of how we can think about the result of clinical trials, for example a generic *treatment effect*. A typical frequentistic analysis would give a *P value*, an *estimate* and a *confidence interval* as a summary of the contribution of that trial to the knowledge about that treatment effect. A Bayesian analysis goes beyond this by answering the question 'How this trial will modify our beliefs about that treatment effect?' This viewpoint prompts the analyst to explicitly state (i) a reasonable opinion concerning the treatment effect 'without' the trial evidence (that is the *prior distribution*); (ii) the information about the treatment effect based 'only' on data coming from the trial (that is the *likelihood*); and to *combine* these two pieces of evidence to produce; (iii) a final opinion about the treatment effect (that is the *posterior distribution*). This final combination is statistically worked out using Bayes's theorem. In summary, the Bayesian approach can be defined as the explicit and quantitative use of any kind of *external* evidence in the design, analysis and interpretation of an experimental trial.

### Practical advantages of the Bayesian method

The posterior distribution provides both point estimates and 'credibility intervals'. The credibility interval is another plus value of the Bayesian approach. Credibility intervals are similar to the confidence intervals of the frequentistic approach, but they truly have the more intuitive interpretation that many of us wrongly ascribe to confidence intervals [7]. Indeed, in the frequentistic framework

the percentage defining the confidence interval (i.e. 95%) refers to the long-term frequency with which the interval holds down the result of repeated experiments; on the contrary in the Bayesian approach, the 95% credible interval effectively is that interval where we believe the estimate lies with a probability of 95% [7,8]. The posterior distribution may also make us able to calculate the probability that the quantity of interest (i.e. odds ratio, relative risk, hazard ratio) is less than a specific value. Thus the Bayesian approach allows to elaborate a meaningful cut off which may indicate, in a therapy contest, the probability to have some beneficial effects from an intervention [7].

Furthermore, the Bayesian approach is naturally suitable to build up hierarchical models, very useful to pool together the information provided by case reports, case series, retrospective cohorts, prospective observations. In doing this, the method allows to take into account and adjust for the different level of bias embedded in the sundry study designs. Particularly in this field, seminal frequentistic methods to pool data from observational studies are available, but they are unanimously considered prone to the effect of any bias embedded in the source data [9].

In summary, the advantages of the Bayesian approach are that: (i) it allows to directly make probability statements about quantities of interest (i.e. 75% probability that the risk of inhibitor of a given patient is below 10%); (ii) it allows to easily make predictive statements, conditional on the current state of knowledge, including the amount of uncertainty (i.e. there is a 80% probability that an additional dose of a given factor concentrate will stop bleeding in 50% of patients refractory to first dose); (iii) it enables evidence from a variety of sources, regarding a specific topic, to be taken into account within a coherent modelling framework (i.e. observational and randomized evidence about prophylaxis effectiveness pooled according to a definite system of probabilistic weights); (iv) it requires to explicit considerations about the rationale for all the phases of the study, since the investigator has to elicit prior beliefs (i.e. about the size of treatment effect of a bypassing agent) and demands (i.e. about the minimum bleeding rate reduction that would be considered clinically worthwhile).

### Limitations of the Bayesian method

Let's now try to discuss how the commonly acknowledged limits of the Bayesian approach are of little relevance in the framework of rare disorders. The

main criticism to the Bayesian approach is the introduction of a degree of 'subjectivity' in the method. Actually, researchers and clinicians in the rare disorders field are more familiar than others with the idea of subjectivity. Subjectivity lays principally in the dependence of the Bayesian construct on a prior distribution, whose elicitation is non-trivial and, at present, not standardized: similarly, the lack of agreed guidelines is a usual situation for researcher in the rare disorders field. Anyway, from a technical viewpoint, whichever prior distribution is used ('non-informative', 'vague', uniform), it is important that a sensitivity analysis is performed to test the effect of variation in the prior to the posterior.

Another relative disadvantage of the Bayesian analysis is its complexity to computational implementation, which constrains to the use of a few very specific softwares, and requires very specific skill in Bayesian statistics. If this can refrain from using the Bayesian theory where frequentistic approach are also available, it is often easier and less time and resource demanding than performing RCTs in the field of rare bleeding disorders.

### Examples of application of the Bayesian method to hemorrhagic disorders

Coming back to our initial argumentation about science progress in the field of rare bleeding disorders, the Bayesian approach makes it easy to incorporate all the 'imperfect' evidence springing out from registries and observational studies. This observational evidence can be used as very informative prior to gain the maximum advantage from the results of the scarce randomized data generated in the field.

Seeking for practical examples in the haemophilia field, we searched the archives of *Haemophilia* looking for the occurrence of the term 'Bayesian'. We found two interesting papers.

#### *Pharmacokinetics*

The use of Bayesian probability to predict of blood levels of factor concentrates has been discussed by Morfini [4]. Actually, the Bayesian method has been widely used to predict blood levels of drug, and it is particularly suitable to predict the infusion amount required to get a desired FVIII/IX level. The *prior* here are (possibly large) population pharmacokinetic data. If the prior distribution is relatively uninformative (few points and/or high variability), the posterior probability is likely to be less precise. The ideal method is three-staged; first, individual's

data are fitted separately to the pharmacokinetic model; secondly, the population parameters are estimated by combining all individual parameters; finally, between-subject variability is minimized by selecting the best random distribution model. Notwithstanding the theory, Morfini concludes that, even in the best conditions, the values predicted by the Bayesian compromise were reported to be up to 30% apart from the values actually measured in the real subject. It has to be noticed that the limited power of calculation available at time of the mentioned analysis could have allowed only an incomplete probability modelling, and full 'multi-parametric' or *joint* Bayesian models are certainly worthy to be tested on the powerful WinBUGS engine.

#### *Assessment of inhibitor risk*

Lee and Roth [5] proposed a Bayesian approach to the evaluation of inhibitor risk for new factor concentrates to be approved for clinical use. They were prompted by the observation that the FDA proposed requirements (based on the frequentistic confidence interval approach, i.e. 95% upper confidence limit for inhibitor development below 6.8%) were so that, for a product to succeed, it must have an extremely low underlying risk of inhibitor development. Several commercially available FVIII products with an excellent safety record would not pass the proposed cut off. The authors proposed an alternative set of acceptance criteria, based on a Bayesian statistical paradigm, which was able to let all accepted product pass, but still blocked a known immunogenic product. They took advantage from the Bayesian determination of the probability that the product under evaluation has an inhibitor risk below the pre-set limit. In fact, the fundamental problem with applying the 6.8% cut off as a fixed metric for determining a safety endpoint threshold is that the upper bound of any confidence interval will vary according to the results of the sample obtained. What it is indeed to estimate, given the data collected, is the probability that the true inhibitor rate falls below 6.8%, a quantity that cannot be inferred from a frequentistic confidence intervals approach. Furthermore, the authors were maximally cautious in the tricky choice of an acceptable prior distribution. This was based on the published literature about inhibitor incidence rates in PTPs, with different weight given to clinical trials and population surveys. They also showed in a sensitivity analysis that assuming a prior uniform distribution did not change to a sensible extent the posterior probability.

### Comparison of treatments to stop bleeding in inhibitor patients

A third paper approaching another hot topic in haemophilia is published in the March issue of *Haemophilia*. Treur *et al.* [3] employed the Bayesian approach to compare the relative efficacy of APCC (FEIBA) and rFVIIa (Novoseven) to stop bleeding in inhibitor patients. The authors took advantage of the flexibility of the Bayesian approach in order to pool all the available evidence (very heterogeneous in nature, either observational, comparative and single arm, or randomized head to head comparisons). They propose a survival analysis model in which the hazard function (the probability to stop bleeding at a given time) was represented by a mathematical construct including parameters for baseline hazard and speed of hazard variation. The baseline hazard parameter incorporates medication type, dosage, method of efficacy rating and efficacy rater. The model estimates the pooled effect of these components basing on all the data of the studies included in the analysis. These estimates are then used to compare the predicted effect of the typically used regimens of FEIBA and Novoseven and their 95% credible intervals.

### Concluding remarks

The Bayesian approach is aimed to allow that evidence coming from different kind of sources is used in a unified model. This is what everyday treaters do to give their patients the better treatment. They naturally rely on their 'prior' assumptions. The same degree of 'subjectivity' is embedded in the Bayesian approach, but it is mathematically worked out and controlled. Further, it is relatively easy to understand what happens if you vary the weight (your 'confidence') given to – let's say – method of assessing efficacy or dosage administered. You may choose your own 'prior' and the method will build up for you the consequences.

Other researchers, interested in more common diseases, can concentrate on enrolling a huge number of patients in more and more RCTs. In the haemophilia field, we hope that more and more resources will be dedicated to the huge modelling and computational effort needed to gain the maximum advantage from the evidence we already have available.

### Disclosures

The authors stated that they had no interests which might be perceived as posing a conflict or bias.

### References

- 1 Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med* 1987; **317**: 141–5.
- 2 Lilford RJ, Thornton JG, Braunholtz D. Clinical trials and rare diseases: a way out of a conundrum. *BMJ* 1995; **311**: 1621–5.
- 3 Treur M, McCracken F, Heeg B *et al.* Efficacy of recombinant activated factor VII (rFVIIa) vs. activated prothrombin complex concentrate (aPCC) for patients suffering from haemophilia complicated with inhibitors: a Bayesian meta-regression. *Haemophilia* 2009; **15**: 420–36.
- 4 Morfini M. Pharmacokinetic studies: international guidelines for the conduct and interpretation of such studies. *Haemophilia* 2006; **12**: 6–11.
- 5 Lee M, Roth D. A Bayesian approach to the assessment of inhibitor risk in studies of factor VIII concentrates. *Haemophilia* 2005; **11**: 5–12.
- 6 Bayes T. An essay toward solving the doctrine of chances. *Philos Trans R Soc Lond* 1763; **53**: 370.
- 7 Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. *Health Technol Assess* 2000; **4**: 1–130.
- 8 Higgins J, Green S (eds) *Cochrane Handbook for Systematic Reviews of Interventions*. <http://www.cochrane-handbook.org> [Version 5.0.1], 518. 1 September 2008. The Cochrane Collaboration, 2008.
- 9 Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res* 2001; **10**: 277–303.