sanofi aventis

Because health matters

Date 21-August-2006

Via fax and UPS

0402 6 AUG 21 PM 2:55

Division of Dockets Management (HFA-305)
Food and Drug Administration
5630 Fishers Lane, Room 1061
Rockville, MD 20852

Re: Docket No. 2006D-0191

Draft Guidance for Industry and Food and Drug Administration Staff; Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

Dear Sir/Madam:

Sanofi-Synthelabo Inc. and Aventis Pharmaceuticals, members of the sanofi-aventis Group, appreciate the opportunity to comment on the above-referenced guidance, "Use of Bayesian Statistics in Medical Device Clinical Trials; Availability."

This document provides guidance on statistical aspects of the design and analysis of clinical trials for medical devices that use Bayesian statistical methods.

The purpose of this guidance is to discuss important statistical issues in Bayesian clinical trials for medical devices and not to describe the content of a medical device submission. Further, while this document provides guidance on many of the statistical issues that arise in Bayesian clinical trials, it is not intended to be all-inclusive. The statistical literature is rich with books and papers on Bayesian theory and methods; a selected bibliography has been included for further discussion of specific topics.

This is the first FDA guidance dedicated to the Bayesian approach to the design and analysis of clinical trials. Although written for devices and not drugs, it has potential implications in terms of relevance for drug development, representing a breakthrough in terms of statistical methods, and opens the way to further research and progress.

GENERAL COMMENTS

- A clarification is needed on the use of "endpoints" and "parameters," as they are used interchangeably in the document.
- A section on how to perform multiplicity adjustment on multiple endpoints or multiple treatment comparisons could be added. In particular, it should be addressed whether the adjustment should be based on the individual posterior distributions of individual parameters of individual variables or the joint posterior distribution of all the parameters of all variables.

200 Crossing Boulevard, PO Box 6890, Bridgewater, NJ 08807-0890
Tel: (908) 304-7000 · www.sanofi-aventis.com

2006D-0191 C4

- It would be helpful to address whether the Bayesian approach can be used for certain endpoints and the Frequentist approach be used for other endpoints. For example, a Bayesian approach could be used for the primary efficacy endpoints and the Frequentist approach used for secondary efficacy endpoints and safety endpoints. Otherwise, if the Bayesian approach has to be used for all endpoints, the computations would be too complicated.

  The general guideline for Frequentist approach is to include all stratification factors in the analysis model, for example, center effects when randomization is conducted within centers. For Bayesian analysis, prior distribution for center effects may not be easy to specify and computation will increase substantially with these center effects in the model. Is it acceptable **not** to include some stratification factors in the analysis model and is there a general guidance that addresses this?

SPECIFIC COMMENTS

Introduction: Page 4

"This document provides guidance on statistical aspects of the design and analysis of clinical trials for medical devices that use Bayesian statistical methods.

The purpose of this guidance is to discuss important statistical issues in Bayesian clinical trials for medical devices and not to describe the content of a medical device submission. Further, while this document provides guidance on many of the statistical issues that arise in Bayesian clinical trials, it is not intended to be all-inclusive. The statistical literature is rich with books and papers on Bayesian theory and methods; a selected bibliography has been included for further discussion of specific topics.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word should in Agency guidances means that something is suggested or recommended, but not required."

Comment: The first sentence "This document provides… that use Bayesian statistical methods" is very clear and accurate. Why then use (twice, and in the title of §5) the ambiguous expression "Bayesian clinical trial" in the second paragraph: this erroneously suggests that the clinical aspects are also specific when the Bayesian approach for statistics is selected in a clinical trial? Section 5, §5.1 to §5.4 of the draft document appropriately develops the idea that the non-statistical aspects of clinical trial protocols are similar, irrespective of whether the design and analysis of the trial is Frequentist or Bayesian.

In conclusion, it seems preferable that the expression "Bayesian clinical trial", or even "Bayesian trial" (which appears to be shorthand for the former, in §3.4 or §3.9) not be used.

Page 5 Section 3.1 What is Bayesian statistics?

"Bayesian statistics is a statistical theory and approach to data analysis that provides a coherent method for learning from evidence as it accumulates. Traditional (Frequentist) statistical methods formally use prior information only in the design of a clinical trial. In the data analysis stage, prior information is considered only informally, as a complement to, but not part of the analysis. In contrast, the Bayesian approach uses a consistent, mathematically formal method called Bayes' Theorem for combining prior information with current information on a quantity of interest. This is done throughout both the design and analysis stages of a trial."

Comments: The one feature of Bayesian statistics, which the draft guidance selected in this paragraph in view of roughly and quickly characterizing this approach, is the most familiar to any statistician, as it stems from Bayes' theorem, but is not the most fundamental.

In addition Frequentist ideas like conditional power are also aimed at "learning from experience as it accumulates"; the empirical cumulative distribution function is another straightforward example of pooling all the available experience in the analysis.

Conversely, Bayesian non-informative approaches tend to underweight prior experience, and in the limit to ignore it.

The notions developed later in §4.1 (and also in the last paragraph of §4.8) are certainly more fundamental and comprehensive in characterizing Bayesian statistics in contrast to the Frequentist approach. Furthermore, Bayesian notions in §4.1 are necessary to §3.2 and §3.9.

In conclusion §3.1 should integrate the content of §4.1 and §4.8.

Page 5 Section 3.2: Why use Bayesian statistics for medical devices?

"When good prior information on clinical use of a device exists, the Bayesian approach may enable FDA to reach the same decision on a device with a smaller-sized or shorter-duration pivotal trial.

The Bayesian approach may also be useful in the absence of informative prior information. First, the approach can provide flexible methods for handling interim analyses and other modifications to trials in midcourse (e.g., changes to the sample size or changes in the randomization scheme). Second, the Bayesian approach can be useful in complex modeling situations where a frequentist analysis is difficult to implement or does not exist.

Good prior information is often available for a medical device; for example, from earlier studies on previous generations of the device or from studies overseas. These studies can often be used as prior information because the mechanism of action of medical devices is

typically physical, making the effects local and not systemic. Local effects are often predictable from prior information when modifications to a device are minor. Bayesian methods may be controversial when the prior information is based mainly on personal opinion (often derived by elicitation methods). The methods are often not controversial when the prior information is based on empirical evidence such as prior clinical trials. Since sample sizes are typically small for device trials, good prior information can have greater impact on the analysis of the trial and thus on the FDA decision process.

The FDA Modernization Act of 1997 mandates that FDA shall consider the least burdensome means of demonstrating effectiveness or substantial equivalence of a device (Section 513(a)(3)(D)(ii) and Section 513(i)(1)(D)). The Bayesian approach, when correctly employed, may be less burdensome than a frequentist approach."

Comments:
2nd paragraph: "absence of informative prior information" is redundant; "informative" should be removed

4th paragraph: although the guidance understandably supports data-based prior hyper-parameters, and dispels subjective prior hyper-parameters (again in §5.5), "skeptical" and "enthusiastic" prior hyper-parameters may be useful to assess the robustness of the conclusions, and this interesting use of priors could be mentioned here.

Page 8 Section 3.9; what are the potential difficulties in the Bayesian approach?

"Planning the design, conduct, and analysis of any trial are always important from a regulatory perspective, but they are crucial for the Bayesian approach. This is because decisions are based on:

- prior information
- information obtained from the trial, and
- the mathematical model used to combine the two.

Different choices of prior information or different choices of model can produce different decisions. As a result, in the regulatory setting, the design of a Bayesian clinical trial involves pre-specification of (and agreement on) both the prior information and the model. This includes clinical agreement on the appropriateness of the prior information and statistical agreement on the mathematical model to be used. Since reaching this agreement is often an iterative process, we recommend you meet with FDA early on to discuss and agree upon the basic aspects of the trial design.

A change in the prior information or the model at a later stage of the trial may imperil the scientific validity of the trial results. For this reason, formal agreement meetings may be appropriate when using a Bayesian approach."

Specifically, the identification of the prior information may be an appropriate topic of an agreement meeting.

Extensive model building

The Bayesian approach often involves extensive mathematical modeling of a clinical trial, including:

- the probability distributions chosen to reflect the prior information
- the influence of covariates on patient outcomes or missing data
- the relationships between various sources of prior information used in the model.

We recommend you determine modeling choices through close collaboration and agreement with FDA's and your statistical and clinical experts.

Specific statistical and computational expertise

The Bayesian approach often involves specific statistical expertise. Computer-intensive calculations are often used to:

- analyze trial data
- check model assumptions
- assess prior probabilities at the design stage
- perform simulations to assess probabilities of various outcomes
- estimate sample size.

The technical and statistical costs for the above are often offset by the savings of a shorter trial or a more flexible analysis.

Choices regarding prior information

An FDA advisory panel may question prior information you and FDA agreed upon beforehand. We recommend you be prepared to clinically and statistically justify choices of prior information. In some cases, we recommend you perform sensitivity analyses to check robustness of models and priors.

Device labeling

Results from a Bayesian trial may be expressed differently from the way trial results are usually described in device labels. We recommend you ensure trial results reported on the device label are easy to understand.

Checking calculation

The flexibility of Bayesian models and the complexity of the computational techniques for Bayesian analyses create greater possibility for errors and misunderstandings. FDA, therefore, will carry out a detailed statistical review of a Bayesian submission.

Since the software used in Bayesian analysis is relatively new, FDA will often verify results using alternate software. FDA recommends you submit your data and any programs used for Bayesian statistical analyses electronically.

Bayesian and traditional analyses approaches may differ

Two investigators, each with the same data and a different preplanned analysis (one frequentist and one Bayesian), could conceivably reach different conclusions that are both scientifically valid. While the Bayesian approach is often favorable to the investigator with good prior information, the approach can be more conservative than a frequentist approach."

Comment: Planning the design and analysis accurately appears as a condition for the scientific validity of the conclusions. However, Bayesian methods resort more strongly to (especially distributional) modeling than frequentist methods, and frequentist methods are often based on asymptotic properties. There is nothing specific to Bayesian methods here. In particular, discrepant conclusions between two frequentist methods (models) may be as likely as between frequentist and Bayesian approaches.

Reference is made to the Bayesian approach being sometimes more conservative than the frequentist approach, with a general citation of Section 5. This comment should be expanded to be more understandable, including clarifying in what sense "conservative" is used.

Page 10 Section 4.1 Bayesian Statistics, Introduction

"The fundamental idea in Bayesian statistics is that one's uncertainty about an unknown quantity of interest is represented by probabilities for possible values of that quantity. For instance, unknown quantities of interest in device trials might be:

- clinical safety and effectiveness endpoints
- a patient's outcome to be observed in the future
- a missing observation on a patient.

Prior distribution and non-informative prior distribution

Before a trial begins and data are obtained, the investigator assigns prior probabilities to the possible values of the unknown quantity, known as the prior distribution. In principle, the prior can be based on the investigator's personal knowledge of the quantities of interest or on another expert's opinion, etc. If absolutely nothing is known about that quantity, something called a non-informative prior distribution may be specified. In trials undergoing regulatory review, however, the prior distribution is usually based on data from relevant previous trials.

Bayes' theorem and posterior probabilities

After data are gathered and information becomes available, the prior probabilities are mathematically updated according to a statistical result called Bayes' theorem. The updated probabilities, known as posterior probabilities, are probabilities for values of the unknown quantity after data are observed. This approach is a scientifically valid way of combining previous information (the prior probabilities) with current data. The approach adjusts to changing levels of evidence: today's posterior probabilities become tomorrow's prior probabilities.

The Bayesian paradigm

The Bayesian paradigm states that probability is the only measure of one's uncertainty about an unknown quantity. In a Bayesian clinical trial, uncertainty about an endpoint (also called parameter) is quantified according to probabilities, which are updated as information is gathered from the trial.

Decision rules

The pre-market evaluation of medical devices aims to demonstrate the safety and effectiveness of a new device. This demonstration is most commonly achieved through statistical hypothesis testing. For Bayesian trials, hypotheses are tested with decision rules. One common type of decision rule considers that a hypothesis has been demonstrated (beyond a reasonable doubt) if its posterior probability is large enough (e.g., 95 or 99 percent).

The Bayesian approach encompasses a number of key concepts, some of which are not part of the traditional statistical approach. Below, we briefly discuss these concepts and contrast the Bayesian and frequentist approaches."

> **Comment:** As a rule, in a Bayesian framework, the "unknown quantities (of interest)" are called "parameters (of interest)" (of the statistical and analytical parts of some model). Why not define and use the term "parameter" throughout, consistent with usual practice, and to clearly distinguish model parameters from observed data?

Page 12 Section 4.2; what is a prior distribution?

"Suppose that **x** is an endpoint (parameter) of interest in a clinical trial. The initial uncertainty about **x** should be described by a probability distribution for **x**, called the prior distribution and denoted by P (**x**).

As an example, suppose **x** is the rate of a serious adverse event. Its possible values will lie between 0 and 1. One prior distribution is the uniform distribution indicating no preference for any value of **x**. So the probability that **x** lies between 0.1 and 0.2 is the same as the probability that **x** lies between 0.4 and 0.5, or between 0.65 and 0.75, or in any interval of length 0.1.

Alternatively, the prior distribution might give preference to lower values of **x**. For example, the probability that **x** lies between 0.2 and 0.3 can be larger than the probability that **x** lies between 0.7 and 0.8."

Comment: Bayesian literature generally uses Greek characters to designate a parameter, and Latin characters (especially "x") for the data. The notations here and later should comply with this traditional use.

Page 13 Section 4.7; what is the likelihood principle?

The likelihood principle is important in all of statistics, but it is especially central to the Bayesian approach. The principle states that all information about the endpoint of interest, **x**, obtained from a clinical trial, is contained in the likelihood function. In the Bayesian approach, the prior distribution for **x** is updated using the information provided by the trial through the likelihood function, and nothing else. Bayesian analysts base all inferences about **x** solely on the posterior distribution produced in this manner.

Comment: The likelihood principle is almost universally accepted among Bayesian statisticians, and its application is especially useful, whereas frequentist statisticians need to "adjust" p-values and central estimates, especially in sequential design and analyses. This stands clearly as one of the most critical differences between Bayesian and frequentist statisticians in clinical research.

Page 14 Section 4.8; how do the Bayesian and frequentist approaches differ

"As outlined above, Bayesian analysts base all inferences on the posterior distribution, which (in adherence to the likelihood principle) is the product only of the prior and the likelihood function. Although the frequentist approach makes extensive use of the likelihood function, it does not always strictly adhere to the likelihood principle. For example, the interpretation of a frequentist p-value is based on outcomes that might have occurred but were not actually observed in the trial; that is, on something external to the likelihood.

Another way of saying this is that Bayesian inferences are based on the "parameter space" (the posterior distribution), while frequentist inferences are based on the "sample space" (the set of possible outcomes of a trial)."

Comment: The foundational difference between Bayesian and frequentist methods is never clearly expressed. Frequentist methods view the unknown parameters of interest as fixed constants of nature, whereas Bayesian methods view the unknown parameters as random objects for which probability statements can be made

Page 21 Section 5.7; determining the sample size

"The sample size in a clinical trial depends on:

- prior information
- mathematical model used in analysis
- distributions of parameters in the analytic model
- specific decision criteria
- variability of the sample.

If the population of patients is highly variable, the sample size increases. If there is no variability (i.e., everyone in the population has the same value for the measurement of interest), a single observation is sufficient. The purpose of sizing a trial is to gather enough information to make a decision while not wasting resources or putting patients at unnecessary risk.

In traditional frequentist clinical trial design, the sample size is determined in advance. Instead of specifying a particular sample size, the Bayesian approach (and some modern frequentist methods) may specify a particular criterion to stop the trial. Appropriate stopping criteria may be based on a specific amount of information about the parameter (e.g., a sufficiently narrow credible interval, defined in **Section 6: Analyzing a Bayesian Clinical Trial**) or an appropriately high probability for a pre-specified hypothesis.

At any point before or during a Bayesian clinical trial, you can obtain the posterior distribution for the sample size. Therefore, at any point in the trial, you can compute the expected additional number of observations needed to meet the stopping criterion. In other words, the sample size distribution is continuously updated as the trial goes on. Because the sample size is not explicitly part of the stopping criterion, the trial can be ended at the precise point where enough information has been gathered to answer the important questions.

Special considerations when sizing a Bayesian trial

When sizing a Bayesian trial, FDA recommends you decide in advance on the minimum sample size according to safety and effectiveness endpoints because safety endpoints may lead to a larger sample size. FDA also recommends you include a minimum level of information from the current trial to enable verification of model assumptions and appropriateness of prior information used. This practice also enables the clinical community to gain experience with the device.

When hierarchical models are used, we recommend you provide a minimum sample size for determining the amount of information that will be "borrowed" from other studies.

We recommend the maximum sample size be defined according to economical, ethical, and regulatory considerations.

Various approaches to sizing a Bayesian trial are described in Inoue et al. (2005), Katsis & Toman (1999), Rubin & Stern (1998), Lindley (1997), and Joseph et al. (1995a,b)."

Comments: The guidance should start by acknowledging that it is impossible to compute the sample size with the Bayesian approach, as the posterior distribution is conditional on the unique, but still unobserved, data set. In addition, as no sampling distribution is used in the Bayesian method, it is impossible to rely on simulated (sampled) data.

However, it is necessary to get some rough idea of the sample size. And this is easily done with the frequentist approach, including diverse degrees of uncertainty about the prior information. Why not acknowledge this fact, and use the frequentist approach prospectively to compute a rough approximation of the sample size required with a Bayesian design and analysis?

Page 22 Section 5.8; assessing the operating characteristics of a Bayesian design

Because of the inherent flexibility in the design of a Bayesian clinical trial, a thorough evaluation of the operating characteristics should be part of the trial design. This includes evaluation of:

- probability of erroneously approving an ineffective or unsafe device (type I error)
- probability of erroneously disapproving a safe and effective device (type II error)
- power (the converse of type II error: the probability of appropriately approving a safe and effective device)
- sample size distribution (and expected sample size)
- prior probability of claims for the device
- if applicable, probability of stopping at each interim look.

Comments: This paragraph makes explicit mention of purely frequentist notions, such as type I and II error rates, and power. It is acknowledged here that the frequentist approach is very well adapted to prospective thinking about a design, as it relies on parameterized sampling distributions. Can this fact be acknowledged in the text?

Page 23 Section 6.3; Interval estimation

"Bayesian interval estimates are based on the posterior distribution and are called credible intervals. If the posterior probability that an endpoint lies in an interval is 0.95, then this interval is called a 95 percent credible interval.

For construction of credible intervals, see Chen & Shao (1999) and Irony (1992). Other types of Bayesian statistical intervals include highest posterior density (HPD) intervals (Lee, 1997) and central posterior intervals. "

1.Sanofi-aventis comments
Docket 2006D-0191

Comment: The section title is misleading, as it is the probability of a given interval, which is estimated. Interval estimation suggests that the boundaries of the interval are random variables, as in the frequentist approach.

On behalf of Sanofi-Synthelabo Inc, and Aventis Pharmaceuticals, members of the sanofi-aventis Group, we appreciate the opportunity to comment on the "*Draft Guidance for Industry and Food and drug Administration Staff; Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials*" and are much obliged for your consideration.

Sincerely,

Richard Gural
Vice President
Regulatory Development