

Planning a Bayesian Trial: Prior Distributions and Hierarchical Models

Gene Pennello, Ph.D.

Division of Biostatistics

Office of Surveillance and Biometrics

Center for Devices and Radiological Health

Food and Drug Administration

Public Meeting on Draft Bayesian Guidance

July 27, 2006

Outline

- Section 5.5: The Prior Distribution
 - Informative Prior Distributions
 - Non-Informative Prior Distributions
- Section 5.6: Hierarchical Models
- Section 6: Analyzing a Bayesian Trial
 - Posterior Distribution Summaries
 - Hypothesis Testing, Interval Estimation
- Section 7: Bayesian Post-Surveillance

5.5 Initial Information About Endpoints: The Prior Distribution

- Prior to any collection of data in a study, initial uncertainty about endpoints (parameters) of interest is quantified by specification of probabilities for their possible values.
- This initial probability distribution is called the *prior distribution*.
- The study design for every Bayesian trial includes specification of the prior distribution.

I. Informative Prior Distributions

- *Informative prior distributions*
 - are formulated from prior information
 - can decrease the sample size required for a trial.
- Possible sources of prior information:
 - clinical trials conducted overseas
 - patient registries
 - clinical data on very similar products
 - pilot studies
 - historical controls (prior for concurrent controls)

Proposing Prior Information to FDA

- FDA should agree with your choice for an informative prior distribution. We recommend that sponsors
 - Make sure prior information is legal to use.
 - Hold a pre-IDE meeting with FDA to come to agreement on what prior information is scientifically valid for the Bayesian trial.
 - Submit prior information as part of the IDE (when an IDE is required).

Data-Based Prior Information

- Prior information based on data from prior studies.
- To be appropriate, prior studies usually need to be similar to the current study.
- Ideally, prior studies & current study are similar in
 - devices used
 - endpoints
 - patient population
 - patient management
 - investigational sites
 - time frame
 - objectives
 - protocol
 - physician training
 - physician experience
 - physician technique

Data-Based Priors (*cont.*)

- Scientists, clinicians, engineers from sponsor and FDA usually play the key role in determining if the prior studies are similar to the current study, and therefore in determining if they are appropriate as prior information.
- Ideally, studies used as prior information are representative of the possible sources of prior information that could have been selected.
- You should avoid systematically excluding non-favorable studies to avoid introducing bias into the Bayesian statistical analysis.

Opinion-Based Priors

- Problematic in the regulatory setting
- Approval of a device could be delayed or jeopardized if FDA advisory panel members or other clinical evaluators or do not agree with the opinions used to generate the prior.

Prior Information for Studies Having a Control

- Prior information can be for
 - the new device
 - the control device
 - both devices
- Prior information for the concurrent control:
 - Historical control data are a common source.
 - Can decrease the sample size for the concurrent control enabling a greater proportion of patients to be allocated to the new device, thereby increasing the experience with it.

Very Informative Prior Distributions

- Prior information can be too informative for a study.
 - E.g., prior information based on many more patients than are to be enrolled in the study
- A prior distribution is too informative if the prior probability that the pivotal study is a success, i.e., demonstrates the proposed claim(s), is excessively high.
- Various remedies are available:
 - modify or discount the prior distribution in some way
 - increase the stringency of the success criterion

II. Non-Informative Prior Distributions

- If prior information on study parameters is unavailable or does not exist for a device, then initial uncertainty about the parameters can be quantified with a *non-informative prior* distribution.
- Possible non-informative prior: uniform probabilities over the range from minimum to maximum possible parameter value.

5.6 Borrowing Strength From Other Studies: Hierarchical Models

- In a Bayesian analysis, the prior distribution is combined with the current study data to produce the posterior distribution, upon which all inference on the study parameters are based.
- The posterior distribution is produced according to an assumed Bayesian statistical model.

5.6 Borrowing Strength From Other Studies: Hierarchical Models

- *Hierarchical models* are often employed because of their flexibility in combining prior information with current study data.
- A hierarchical model assumes a hierarchical structure in the specification of the statistical distributions for the data and the parameters.

5.6 Borrowing Strength From Other Studies: Hierarchical Models

- To estimate current study parameters, hierarchical models "*borrow strength*" from the prior information.
- Borrowing strength translates to additional sample size.

Flexibility of Hierarchical Models

- The extent of the borrowing is adaptive:
 - If prior information is consistent with current study results, a large degree of strength is borrowed from the prior information.
 - For very different sets of results, the strength borrowed can be negligible.

Appeal Hierarchical Models to Regulators

- Hierarchical models reward sponsors for having prior information consistent with actual device performance.
- Protect against over-reliance on prior information that turns out to be overly optimistic or pessimistic for current study parameters.

Example of a Hierarchical Model

- Registry of an approved device used as prior information for an investigational device.
- 1st level of the hierarchy (patient level):
 - patients in the current study are exchangeable
 - patients in the registry are exchangeable
- 2nd level of the hierarchy (study level):
 - Current study and registry are exchangeable

Example of a Hierarchical Model

- Suppose a success rate is the primary endpoint.
- Under the 2nd level of the hierarchy, the success rates for the registry and the current study are
 - considered to be different (no two studies alike)
 - related because they are assumed exchangeable.
- Registry
 - provides information about success rate in current study,
 - not as much as if the patients in the registry and the current study were pooled (considered exchangeable).

Make Exchangeability Assumption Tenable: Calibrate Studies

- If the prior and current studies are known to differ in important covariates (e.g., prognostic variables), the exchangeability assumption is likely not to hold.
- Prior and current studies might still be regarded as exchangeable after statistical adjustment for all of the known important covariates.

Make Exchangeability Assumption Tenable: Calibrate Studies

- The statistical adjustment for covariates in effect calibrates the prior studies to the current study, so that the exchangeability assumption is more tenable.
- For proper calibration, covariate information at the patient level is ordinarily needed in both the prior studies and the current study.
- For some prior studies (e.g., literature), patient level information may not be available.

Applications of Hierarchical Models

- Combine data across a priori exchangeable studies.
- Combine data across a priori exchangeable centers in a multi-center trial.
- Adjustment for multiple testing of a priori exchangeable subgroups.

6. Analyzing a Bayesian Clinical Trial

- All inferences on the study parameters are based on their posterior distribution.
- Recall the posterior distribution is the combination (update) of the prior distribution with the current study data.

6.1 Summaries of the Posterior Distribution

- Inferences on a parameter are usually facilitated by summaries of the posterior distribution, especially
 - graphic representation of the posterior distribution for the parameter
 - the posterior mean
 - the posterior standard deviation
 - the posterior probability of a claim
 - Bayesian hypothesis test and credible interval

6.2 Hypothesis testing

- FDA often bases approval on demonstrating claims via hypothesis tests.
- A Bayesian hypothesis test (decision rule) is based on the posterior probability for the hypothesis.

6.2 Hypothesis testing

- Frequentist hypothesis tests control type I and II error rates.
- Regardless of whether a trial is Bayesian or frequentist, FDA is keenly interested in the type I and type II error rates associated with its decision to approve or not approve the device under review.
- Therefore, FDA recommends that you investigate the type I and II error rates of your proposed Bayesian hypothesis test (decision rule).
- Simulation is often used to compute these error rates.

6.3 Interval estimation

- In frequentist statistics, a *confidence interval* is an interval estimate for a parameter that indicates likely values for it.
- The Bayesian counterpart is the *credible interval*. If the posterior probability that a parameter lies in an interval is 0.95, then this interval is called a *95% credible interval*.
- Special types of credible intervals:
 - *central posterior interval*
 - *highest posterior density (HPD) interval*

7. Post-Market Surveillance

- The pre-market clinical study can serve as prior information, bridging the two distinct phases in the total product life cycle.
- If the pre-market study was Bayesian, the posterior distribution (or some modification of it) can serve as the prior distribution for the post-market.

7. Post-Market Surveillance

- Millions of potential signals: One class of devices may have, say, 100 possible adverse patient events, 50 possible device events (e.g., malfunction), 100 manufacturers, 10 devices per manufacture, 2 models per device.
- By chance, many spurious signals are expected.
- Bayesian hierarchical models can be employed on the adverse event/device combinations to reduce the number of spurious signals, thereby illuminating the true signals.
- Spurious signals can get pulled back to normal levels because strength is borrowed from adverse event rates for other combinations.
- CDRH is currently developing Bayesian hierarchical modeling tools specifically designed to datamine post-market medical device reports.

