

THE USE OF BAYESIAN METHODS IN EARLY CLINICAL STUDY DESIGN

Susie Collins, Mark Whitlock

Early Clinical Development, Pfizer, Cambridge, UK

ABSTRACT

In early clinical research it is important to use all of the information available when designing our early-phase clinical studies. Bayesian methodology is widely employed within Pfizer to formally incorporate historical data and information in order to run smaller, cheaper, and more ethical studies and to make more reliable decisions on whether to progress an asset into later-stage development.

Techniques for incorporating historical data and information are discussed including the use of elicitation to set an informative prior for the between study variance component in cases where there are few historical studies.

Two examples are presented where Bayesian methods have been employed for the design of an Alzheimer's psychosis study; the first showing how estimates of variability from historical studies can be combined across studies to provide an estimate of variability for input into sample sizing calculations for the future study; the second showing how placebo responses from the same historical studies can be combined into a placebo response prior which can then be incorporated into the statistical analysis of the future study, reducing the number of patients required in the placebo arm.

How estimates of variability from historical studies can be combined across studies to provide an estimate of variability for input into sample sizing calculations for the future study

When designing clinical studies the one question every statistician is asked is "What sample size do we need?". In order to answer this question, it is necessary to have an estimate of the variability for the endpoint of interest. This can be estimated by means of data from historical studies under the assumption that these studies are relevant or 'exchangeable', i.e. similar to the future, planned study.

If data are only available from a single historical study then this single estimate of variability can be used. Or, to be more conservative, this estimate can be inflated using a 1-sided upper 80% confidence limit for the true standard deviation based on the chi-squared distribution (although this does not account for the additional study-to-study variation).

Ideally, however, there are data available from several historical studies and in this scenario, in early clinical research at Pfizer, we generally adopt a meta-analysis approach, or more specifically, a Bayesian meta-analysis approach combining the standard deviation estimates across the studies to give a predicted standard deviation for a future study.

The meta-analysis approach fits a mixed effects model to the historical SD estimates with study as a random term (like PROC MIXED in SAS). In the Bayesian paradigm, priors on the random effects are additionally incorporated, i.e.

$$y_i = \mu + s_i + e_i, \text{ where}$$

y_i is an individual standard deviation from a study s_i
 $s_i \sim N(0, \omega^2)$ is the study random effect
 $e_i \sim N(0, \sigma_i^2)$ is the within study error, based on the standard error of each SD

A predicted standard deviation (y') for a future study is then given by $y' = \mu + s'$. To be conservative, we usually take the 80th percentile of the distribution of these predicted values as the estimate of the standard deviation for sample sizing calculations.

BugsXLA, Winbugs or Openbugs packages can all be used to perform this Bayesian meta analysis.

Flat priors can be used for μ and ω , i.e.

μ : Normal, centred on the mean of the data
 ω : Half Normal, positive.

It is sensible to assess the sensitivity of the model to the prior distribution by using a range of different priors and checking how much influence this has on the results.

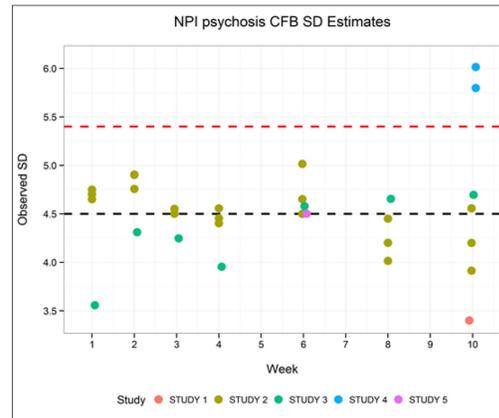
If the number of historical studies is limited, it can be difficult to estimate the between-study variability (ω). This can be overcome in the Bayesian paradigm by incorporating an informative prior distribution for this between-study variation.

If an informative prior is needed, elicitation can be used, similar to the approach described in Spiegelhalter et al (2004)[†] p.168. The elicited value for the maximum credible difference between the effects of two randomly chosen studies (Δ) is used to determine the scale parameter for a Half-Normal prior distribution. Simulation work by Woodward (2011)^{††} p. 56 has shown that the Half-Normal scale parameter needs to be set to approximately $\Delta/3$ in order to induce a prior belief of about 95% probability that the difference between the effects is Δ .

Example: Combining estimates of variability for sample sizing an Alzheimer's (AD) psychosis study

Inclusion criteria (randomised, placebo controlled, double-blind, parallel group studies run in a nursing home with patients diagnosed with AD and psychosis; sample size ≥ 30 ; Neuropsychiatric Inventory (NPI) psychosis endpoint) were used to perform a literature search for 'exchangeable' historical studies. This resulted in a handful of just 4 relevant studies, in addition to a current study which had just reported out its primary analysis.

A plot of the variability for the NPI psychosis change from baseline at each timepoint within each study revealed that the variability was reasonably consistent across time.



A Bayesian meta-analysis (as described earlier) was used to calculate a conservative estimate of the standard deviation at each timepoint.

Timepoint	SD
4 weeks	5.2
6 weeks	5.0
8 weeks	4.9
10 weeks	6.0

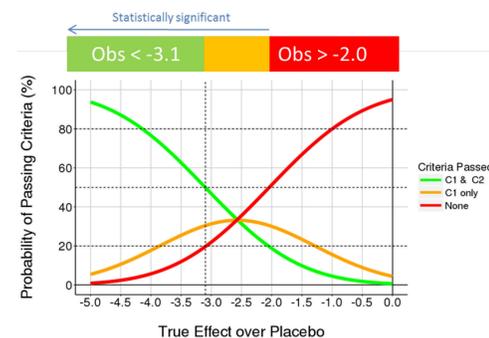
Additionally, since the length of study was undecided at this point, a more complex meta-analysis, with an additional source of variation (time), was used to calculate a conservative estimate of the SD for an 'average' timepoint (estimate=5.4 [red dashed line]).

These estimates were then plugged into sample sizing calculations for various scenarios of study length/variability and target value.

Example: SD = 5.4, target value of 3.1 (based on meta-analysis of effect sizes from similar historical studies)

Criterion 1: 95% confident effect < 0
 Criterion 2: 50% confident effect < -3.1

A sample size of 38 per group (parallel design) gave what was considered to be desirable operating characteristics.



How placebo responses from the same historical studies can be combined into a placebo response prior which can then be incorporated into the statistical analysis of the future study, reducing the number of patients required in the placebo arm

A similar meta-analysis approach (analysing the average placebo response rather than the standard deviation) can be utilised to estimate a placebo response prior distribution for incorporating into the statistical analysis of the future study. This allows a reduction in the number of patients that are required for enrolment into the study and so achieves savings in time and cost.

Example: Estimating a Placebo response prior for an Alzheimer's psychosis study

The results of the same literature search were also used to estimate a Placebo response prior for inclusion in the statistical analysis of the future AD psychosis study.

Study	CFB mean @ 6 weeks	CFB sem	n
STUDY 1	-4.83	-	129
STUDY 2	-5.52	0.43	117
STUDY 3	-4.59	0.43	121
STUDY 4	-1.94	0.47	90

As noted, there were only a few relevant studies with available data and so it was necessary to incorporate an informative prior for between-study variability in order to estimate this Placebo prior distribution.

This prior was obtained using elicitation (as described earlier). Relevant members of the study team were asked "What is the maximum difference in placebo change from baseline mean that we might expect to see between two studies?". In an ideal world, this should be done without seeing the data. For the available historical data, the observed maximum difference over the 10 weeks was approximately 2, but including the results of the recently read-out study increased this to 4 which was considered a plausible maximum between two studies.

This elicited value then determined the Half-Normal prior (as described earlier) used in the Bayesian meta-analysis to produce a posterior distribution to utilise as the placebo response prior in the analysis of the future study. [Note that as part of the analysis, we planned to check the concordance of the prior with the observed data with a view to down-weighting the prior if it was sufficiently discordant.]

The effective number of placebo subjects that the prior is equivalent to can be calculated and the sample size for the future study reduced accordingly. Assuming a Normal likelihood and prior (and hence posterior), the effective sample size can be calculated as the residual variance for the study being planned, divided by the variance of the prior distribution. This is explained in more detail in Spiegelhalter et al (2004) p.62-64. In this example, the prior at 6 weeks was equivalent to 7 subjects (9% of the total study sample size). Although this is seemingly small, it would result in a substantial time (~1-2 months) and cost (~\$0.5M) saving in such a hard to recruit early clinical trial.

Early Clinical Development



[†] Spiegelhalter et al (2004) Bayesian Approaches to Clinical Trials and Health-Care Evaluation
^{††} Woodward P. (2011) Bayesian Analysis Made Simple. An Excel GUI for WinBUGS

