

Uncertainty and decision making in the development of antidepressants: the impact of posterior predictive power on interim analysis

Gijs Santen¹, Meindert Danhof¹, Oscar Della Pasqua^{1,2}



¹Division of Pharmacology, Leiden/Amsterdam Center for Drug Research, The Netherlands
²GlaxoSmithKline, Clinical Pharmacology & Discovery Medicine, United Kingdom



INTRODUCTION

Recently, we have identified that enrolment rate is one of the key factors that determine the optimal timing of an interim analysis in the clinical development of antidepressants. The current investigation explores the feasibility of applying a Bayesian approach for interim analysis that takes enrolment rate into account.

Based on simulated data, we show how a utility function can be used to assess the optimal timing of an interim analysis to detect a potential treatment effect and/or stop a trial.

METHODS

Data

Historical data on HAMD scale¹ are used for illustration of the concepts. Data from two double-blind randomised placebo-controlled trials in depression were obtained from GlaxoSmithKline's clinical trial database.

- Study 1: paroxetine CR 12.5 mg 25 mg (n=150, 7 visits, 8 weeks)
- Study 2: lamotrigine (n=150, 8 visits, 7 weeks)

Model

The HAMD data was analysed using a hierarchical linear model with treatment-time and baseline-time as fixed effects. Simulations were performed based on parameter estimates from earlier analyses from studies on a similar patient population.

Utility function at interim analysis

The utility function of the interim analysis is defined as the difference in the frequency of stopping decisions between no significant effect and a desired treatment effect size.

Decision process

The posterior predictive power (PPP, fig. 1) is defined as the power to detect a treatment effect given all the evidence. The PPP is used as decision basis for estimation of futility and efficacy stopping criteria. In conjunction with the observed enrolment rate (figure 2), an evaluation of the utility function is made by exploring how different scenarios and timings contribute to its maximisation for simulated data.

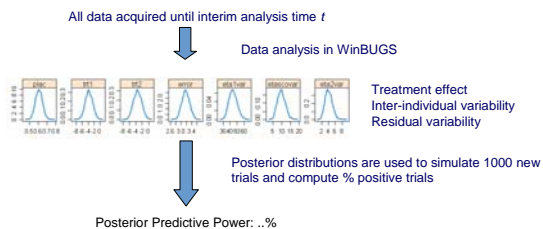


Figure 1. Computation of the posterior predictive power (PPP)

METHODS CONTINUED

Software

All modelling was performed using WinBUGS 1.4.1². Simulations, calculations and summaries were obtained using the language and environment for statistical computing, R (version 2.5.1)

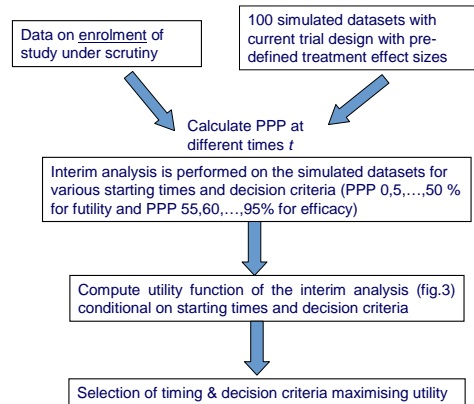


Figure 2. Method used to optimise decision criteria & timing of IA

The selected decision criteria are subsequently applied to the historical datasets and evaluated

RESULTS: DETERMINING DECISION CRITERIA

The results of the evaluation of the optimal decision criteria for an interim analysis of study 1 and 2 are shown in figure 3A and 3B, respectively. The utility curve is shown for criteria ensuring that type I error < 5% and type II error < 20%.

Based on the utility curves, an interim analysis for study 1 could be started at 80 days after start of enrolment. For study 2, an interim analysis can take place at 240 days after start of enrolment. In both cases cut-off criteria of 20% and 85% for futility and efficacy (red circles) were selected.

Application of the selected decision criteria to the original datasets led to a stopping decision of the 25 mg paroxetine arm in study 1 for efficacy. Study 2 was stopped for futility. The savings in patients were substantial (25-40%). Analysis of the complete datasets revealed that these decisions were in line with the trial results.

RESULTS: DETERMINING DECISION CRITERIA

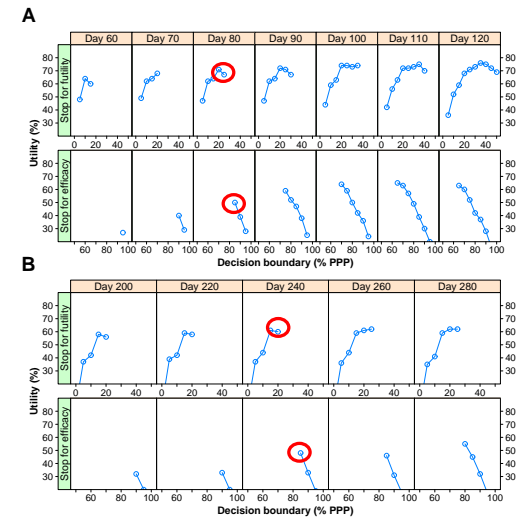


Figure 3. Utility of the interim analysis for study 1 (A) and study 2 (B) versus varying decision criteria. Top panels: stopping for futility. Bottom panels: stopping for efficacy. Each panel depicts a different starting time for the interim analysis.

CONCLUSION & PERSPECTIVES

- A methodology is proposed to determine the decision criteria for a Bayesian interim analysis that accounts for enrolment rate.
- The use utility curves provides a robust summary of the timing and decision criteria required to optimally predict effect size during an interim analysis.
- Validation of the method on historical datasets with anti depressants show that substantial reductions can be reached in population size.

References:

- ¹ Hamilton M: A rating scale for depression. J.Neural.Neurosurg.Psychiatry 1960; 23:56-62
- ² Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing. 2000;10:325-337.