

Title: Uncertainty and decision making in clinical development: the impact of an interim analysis based on the posterior predictive power on depression trials

Authors: Gijs Santen(1), Meindert Danhof(1), Oscar Della Pasqua(1,2)*

Institutions: 1 Division of Pharmacology, LACDR, Leiden University, Leiden, The Netherlands
2 Clinical Pharmacology & Discovery Medicine, GlaxoSmithKline, Greenford, United Kingdom

Objectives: Clinical trials for marketed anti-depressant have a failure rate of almost 50%¹. This failure rate may be attributed to several factors, among which (1) the heterogeneity of the patient population, (2) the insensitivity of the endpoint (Hamilton Depression Rating Scale, HAMD)² to detect drug effect, (3) the nature of the disease, which leads to a high placebo effect and (4) inadequately powered studies. An appropriate interim analysis may lead to early detection and stopping of failing clinical trials, and may prevent clinical trials to accrue patients long after sufficient evidence of statistical significance has been achieved. In previous work we have suggested the use of a Bayesian hierarchical model with composite random effects to analyse depression data. The current investigation explores the feasibility of applying a Bayesian approach for interim analysis, thereby improving the accuracy of estimates for futility and efficacy criteria.

Methods: A linear longitudinal mixed effects model is proposed that incorporates a random slope and intercept to model discrepancies of individual patients from the population mean. The model is used to fit the fraction of the population associated with each treatment arm in the trial at the time of the interim analysis. The posterior distributions of all parameters are used subsequently in simulations to generate new studies with the projected sample size for the actual trial. The percentage of significant study results is reported as the posterior predictive power (PPP). Ultimately, the PPP and corresponding credible intervals are used to make decisions on stopping for futility or efficacy. In addition to the improvement in parameter estimation, the method allows detection of the appropriate timing for an interim analysis. For this purpose, data can be simulated taking into account the enrolment rate of the actual trial. Based on pre-specified futility and efficacy criteria one can determine at any given time point whether the interim analysis should be performed or postponed. R was used for data manipulation and graphical summaries. Model fitting and simulation were performed in WinBUGS 1.413.

We demonstrate the relevance of this approach using three historical datasets obtained from GlaxoSmithKline's clinical trial database, which allowed evaluation of treatment effect in 5 different treatment arms (paroxetine & fluoxetine dose titration (n=290), paroxetine controlled release 12.5 and 25 mg fixed dose (n=150) and dose-titrated lamotrigine (n=150)). As decision criteria, cut-off values for the PPP were chosen so that a PPP lower than 50% resulted in early termination for futility and a PPP higher than 90% resulted in early termination for efficacy. The timing for the interim analysis was determined for each individual study using the simulation procedures described above.

Results: Early termination was achieved for 2 out of the 5 treatment arms; one for efficacy (25 mg paroxetine controlled release) and one for futility (lamotrigine). In fact, in both studies the active treatment was not statistically different from placebo. On the other hand, the treatment effects for the other arms for which stopping criteria were not met were all close to statistical significance (as determined using conventional methods).

Conclusions: Our results strongly suggest that the use of a Bayesian approach based on a longitudinal mixed effects model leads to adequate stopping decision in clinical trials in depression. Furthermore, the proposed methodology offers pragmatic alternative to the evaluation of the cut-off values for futility/efficacy criteria. These findings also show the relevance of incorporating uncertainty in model parameters when using modelling and simulation for the purposes of decision making.

References:

- [1] Khan A, Brodhead AE, Kolts RL, Brown WA. Severity of depressive symptoms and response to antidepressants and placebo in antidepressant trials. *J Psychiatr Res.*39:145-150.
- [2] Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23:56-62.
- [3] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing.* 2000;10:325-337.