

Title: Model-based and trial simulation approach to estimate antidepressant drug effect in multi-center clinical trials: relevance of placebo response in the different recruitment centers

Authors: Roberto Gomeni*

Institutions: CPK/Modelling & Simulation, Clinical Pharmacology Discovery Medicine, GlaxoSmithKline, Verona, Italy

Objectives: Develop a modeling and a trial simulation strategy to estimate the magnitude of antidepressant treatment effect by factoring the overall signal of clinical efficacy into the contribution of the signals generated in each recruitment center. Investigate the relationship between the detectable magnitude of treatment effect and the distribution of placebo response in the recruitment centers. Propose a methodology to characterize the performance of each recruitment center (level of placebo response) and to evaluate and compare treatment effect in different trials in presence of heterogeneous distribution of the placebo response.

Methods: Data were derived from GSK clinical database on Paroxetine (GlaxoSmithKline clinical trial register [<http://ctr.gsk.co.uk/medicinelist.asp>]). For the present investigation, we considered 3 randomized, double-blind, placebo controlled, parallel group studies for treatment of Major Depressive Disorders (MDD), with HAMD-17 total as primary efficacy measurement and with study duration of 8 weeks (study 810, 448 and 449).

Study	Arm 1 Treatment	Arm 2		Arm 3		Nb of Centers	Nb of Patients
		Treatment	Dose	Treatment	Dose		
448	Placebo	Parox_IR	20-50mg -flex	Parox_CR	25-62.5mg -flex	19	299
449	Placebo	Parox_IR	20-50mg -flex	Parox_CR	25-62.5mg -flex	20	333
810	Placebo	Parox_CR	12.5mg-fix	Parox_CR	25mg-fix	38	489

Treatment responses in the 3 treatment arms were defined by the time-varying scores of the HAMD-17 clinical scale. The time-course of this measure usually shows a nonlinear decrement behavior from a high initial score (~25) to a lower value (~9) associated with clinical remission within a time-window of 6-8 weeks. The HAMD-17 time course in the 3 treatment arms were independently analyzed using a mixed Weibull/linear equation $f(t) = Ae^{-(t/td)^b} + h_{rec}t$ where A is the baseline HAMD-17 score, td is the time corresponding to 63.2% of the maximal change from baseline, b is the shape or sigmoidicity factor, and h_{rec} is the remission rate [1]. The model parameters were estimated using the non-linear mixed effects modeling approach as implemented in the NONMEM VI software using the FOCE INTERACTION estimation method.

Clinical trial simulation was used to assess how different distributions of placebo response in the different centers may affect the overall active drug performances (detectable magnitude of clinical efficacy). The post-hoc individual trajectories of the HAMD-17 scores were used to estimate the mean center response for the 3 treatment arms. Treatment response in each center was simulated by selecting the set of individuals, re-sampled from the 3 treatment arms overall populations, with HAMD-17 scores simultaneously falling within the center-specific 95% confidence intervals of HAMD-17 at baseline and at week 8. For the study 810, five trial simulation scenarios were evaluated to explore the impact of different level of placebo response in the recruitment centers. For each scenario the magnitude of treatment effect was derived by averaging the results from 20 simulated trials. The 5 scenarios were: 1) reference scenario corresponding to the observed distribution, 2) centers with placebo mean HAMD-17 at week 8 < the first quartile (q1) of the overall center distribution were replaced with subjects belonging to centers with HAMD-17 at week 8 > q1, 3) centers with the placebo mean HAMD-17 at week 8 > than the 3rd quartile (q3) of the overall center distribution were replaced with subjects belonging to centers with HAMD-17 at week 8 < q3, 4) centers with the placebo mean HAMD-17 at week 8 > the median of the overall center distribution were replaced with subjects belonging to centers with HAMD-17 at week 8 < median, and 5) centers with the placebo mean HAMD-17 at week 8 < the median of the overall center distribution were replaced with subjects belonging to centers with HAMD-17 at week 8 > median.

A probabilistic approach was used to characterize the performance of each recruitment center and to provide criteria for ranking center performances. The placebo response was considered as a background noise that interferes with the probability of detecting the signal of an effective active treatment. The magnitude of clinical efficacy (Δ) per center was calculated as the mean difference between the HAMD-17 total score of active and placebo arm at week 8 minus the same difference at baseline. A change in HAMD-17 total score at week 8 of at least 3 units was considered of clinical relevance and indicative of antidepressant effect. A logistic model was used to describe the probability of observing a clinical efficacy ($\Delta > 3$) considering the mean placebo HAMD-17 score at study-end per center as potential predictor variables. The probability associated with each center was used as a ranking criterion.

Results: The placebo response (HAMD-17 at week 8) in each recruitment center was correlated with the active treatment response showing that 'placebo effect' affects the signal of clinical response (HAMD-17 score) measured either in the active or in the placebo arm.

The detectable magnitude of treatment effect (HAMD-17 difference between active and placebo at study-end) was directly correlated with the distribution of placebo HAMD-17 score at week 8 in the different recruitment centers. Simulations of the expected outcomes of study 810 were performed according to different scenarios accounting for different proportions of recruitment centers with high and low placebo response. The simulations showed that the detectable magnitude of treatment effect varies from 4.1 to 1.23 points when the median placebo HAMD-17 score at week 8 varies from 15.5 to 9.8 points. This finding suggests that the treatment effect cannot be considered as an absolute number but it should be interpreted in relation with the level of placebo response. Despite similarity in the study design and dosage regimen, the studies 448 and 449 showed different detectable magnitude of treatment effect. This was mainly due to the different distribution of the placebo response in the recruitment centers. Trial simulation showed that the discrepancy in treatment effect disappeared with the normalization of placebo response across trials.

The logistic analysis provided criteria for ranking each center as a function of the placebo response according to the probability to detect a clinically relevant signal of efficacy. The probability threshold of 20% to detect clinically significant response was associated with a placebo HAMD-17 at week 8 of ~ 9 (remission) and was used as criteria to 'normalize' placebo response distribution by filtering-out un-informative centers.

Conclusions: The appropriate assessment of the magnitude of antidepressant treatment effect, the definition of dose/exposure response and the comparison of the drug response across studies would require the 'normalization' of the quantifiable treatment effect by the level of placebo response. The results of the analyses presented indicate that the risk of false negative results for an active antidepressant treatment is associated with the proportion of recruitment centers with high placebo response. Finally, the probability to detect a clinically relevant signal of efficacy in each recruitment center could represent a valuable criterion for ranking center performances and for 'normalizing' placebo response across trials.

References:

[1] Br J Clin Pharmacol. 2007 May;63(5):595-613.