Bioequivalence, bootstrapping and case-deletion diagnostics in a biologic: a model-based analysis of the effect of formulation differences in a monoclonal antibody

Justin J Wilkins, Aurélie Gautier, Phil J Lowe

Modeling & Simulation, Novartis Pharma AG, Basel, Switzerland

Introduction and Objectives

The primary objective of this work was to ascertain, through an integrated PK/PD model-based approach, whether the pharmacokinetics of a monoclonal antibody in development and the pharmacodynamic responses of free and total IgE to this drug were similar for three different formulations – a reference formulation (A), and two alternatives (B and C).

Methods

Two studies employing an open-label, randomized, two-parallel-group, single subcutaneous injection design were used in the analysis (see Table 1).

Table 1. Structure of the dataset.

Study	Formulation	*n*	Obs	Demographic data (mean ± SD & range)		
				Age [y]	Weight [kg]	Baseline IgE [ng.mL^{-1}]
1	A	74	3402	33 ± 12 (18–64)	71 ± 12 (48–91)	190 ± 126 (55–616)
1	B	79	3649	36 ± 13 (18–63)	71 ± 12 (48–91)	186 ± 124 (47–620)
2	B	10	470	41 ± 12 (24–62)	64 ± 9 (56-82)	323 ± 316 (66–928)
2	C	29	1363	34 ±10 (19–58)	69 ± 13 (48–90)	175 ± 93 (44–459)

A previously-published instantaneous equilibrium drug-ligand binding and turnover population model [1, Figure 1] was adapted in NONMEM VI to allow estimation of the effects of formulation and study on key model parameters relative to formulation A, in a proportional manner such that an effect of zero would deliver an estimated parameter value of unity – allowing intuitive estimation of the relative bioequivalence of formulations B and C for each parameter. An example of how this was done for any model parameter P, with typical covariates, is given by Equation1:

$$P = \theta_P \cdot \left(\frac{WT_i}{70}\right)^{\theta_{P,WT}} \cdot \left(\frac{IgE_{0,i}}{143}\right)^{\theta_{P,IgE}} \cdot \left(\theta_{P,STDY}\right)^{STDY} \cdot \left[\theta_{P,FORMB} \middle| \theta_{P,FORMC}\right] \cdot \exp(\eta_i) \qquad (1)$$

Here, θ_P is the population mean parameter value, WT_i is body weight in individual i, $\theta_{P,WT}$ is the effect of body weight relative to 70 kg, $IgE_{0,i}$ is baseline IgE in individual i, $\theta_{P,IgE0}$ is the effect of baseline IgE relative to 143 ng·mL^{-1}, $\theta_{P,STDY}$ is the influence on P exerted by study 2, $STDY$ is study (0 or 1), $\theta_{P,FORMB}$ and $\theta_{P,FORMC}$ are the influence on P exerted by formulations B or C respectively, and η_i is interindividual variability of P in individual i, normally distributed with mean 0 and variance ω^2.

Using this approach, if the population mean value of θ_P is unchanged between formulations A, B and C, $\theta_{P,FORMB}$ and $\theta_{P,FORMC}$ should equal unity. Bioequivalence may thus be assessed by ascertaining whether the confidence intervals of these three 'bioequivalence' parameters overlap the range 0.8–1.25.

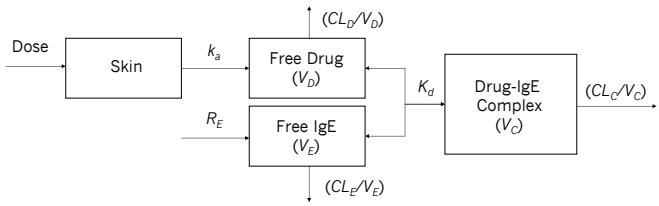

Figure 1. The drug-IgE turnover model. CL_n, V_n = clearance and volume of distribution respectively, for free drug (D), free IgE (E) and drug-IgE complex (C); K_d = dissociation constant; R_E = rate of IgE production.

Resampling-based diagnostics, as implemented in PsN 2.2.5 RC1 [2], were employed at key decision points in addition to standard comparisons of NONMEM objective function value (OFV) and diagnostic plots. The sensitivity of the model to unusual individuals or segments of the data was tested using case-deletion diagnostics. The asymptotic standard errors produced by NONMEM assume a normal distribution, not ideal for an exercise such as this, and so bootstrapping was used to provide more robust estimates of these, as well as to gauge model robustness. The model dataset was resampled 2500 times with replacement, and each was then fitted by the final model. Nonparametric confidence intervals were determined by inspecting the percentiles.

Table 2. Results for bioequivalence and study effect. Bootstrap-generated 90% CIs that extend beyond the acceptable range are highlighted.

Drug or IgE parameter	Formulation B		Formulation C		Study	
	Estimate	90% CI	Estimate	90% CI	Estimate	90% CI
CL_X/F	0.973	0.887-1.07	1.00	0.778-1.33	1.31	1.03-1.62
CL_E/F	1.01	0.918-1.12	0.907	0.745-1.11	0.949	0.827-1.1
CL_C/F	1.10	0.978-1.22	0.913	0.743-1.10	0.995	0.844-1.20
V_X/F & V_E/F	1.03	0.971-1.08	0.849	0.761-0.948	1.24	1.14-1.35
V_C/F	1.05	0.974-1.13	0.811	0.680-0.961	1.26	1.09-1.48
R_E/F	1.01	0.915-1.12	0.933	0.765-1.13	0.914	0.801-1.05
k_a	1.11	0.952-1.27	1.17	0.773-1.74	0.732	0.519-1.03
K_d	0.989	0.924-1.06	1.11	0.945-1.31	0.833	0.725-0.97

*The 97.5th percentiles were also <1, or the 2.5th percentile >1, therefore can be regarded as statistically significant at the 95% level.

Results

The core model parameter estimates were well-estimated and consistent with those obtained previously. For formulation B, all ratio parameters were close to unity, with bootstrap-derived 90% confidence intervals within the range 0.8–1.25. With formulation C, the confidence intervals were outside the acceptance range, such that bioequivalence with formulation A could not be shown. Results for the bioequivalence parameters are given in Table 2.

Case-deletion diagnostics identified and characterized the effects of several influential outliers (Figure 2), as well as a study effect on several binding-related parameters, most notably K_d and its associated nonlinearity factor α (Figure 2 inset). 61.0% (1524) of the requested 2500 bootstrap iterations completed successfully ('minimization successful'). The majority of the remaining runs failed due to rounding errors (98.2% of failed runs). Only successful runs were included in the calculation of percentile-based confidence intervals (see Figure 3).

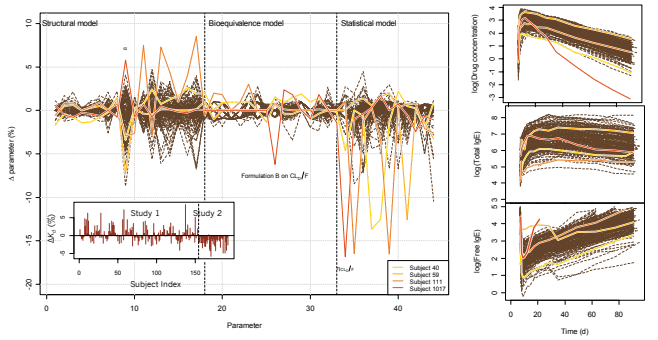

Figure 2. Influence of outliers on the model fit. The differences are represented as ratios of the model-estimated parameters omitting each of the four outliers to those from the model incorporating all subjects. Covariances have been omitted for brevity. In the inset, the case-deletion diagnostic procedure identified a study effect on K_d.

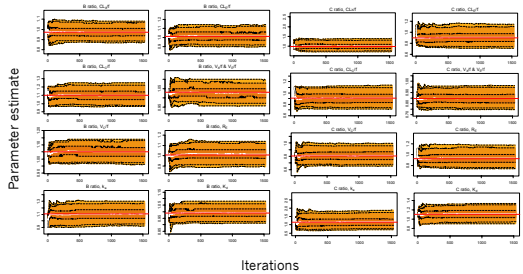

Figure 3. Results of the bootstrapping procedure. The shaded areas represent (from center) the 50%, 90% and 95% confidence intervals, the red line is the model-estimated value of the parameter, and the white line is the median bootstrap estimate.

Discussion and Conclusion

The model-based approach was effective in showing bioequivalence between the formulations A and B, but the low number of patients treated with formulation B in study 2 were not sufficient to allow successful bridging, clearly shown by the wide confidence intervals on the ratio parameters for formulation C (Table 2). The large differences between studies may relate to season: Study 1 completed in summer, whereas Study 2 completed in winter. IgE receptors are known to vary in their expression levels during the year [3,4]. The significantly lower K_d in Study 2 suggests higher affinity binding, which may be a result of there being fewer endogenous receptors available to compete with drug for binding to IgE. Assay differences are another possible source of differences, although laboratory analyses were carried out under GLP conditions in both cases.

Case-deletion diagnostics identified significant, previously-undetected study effects in volume- and binding-related parameter estimates, which produced a significant change in the overall results of the modeling exercise when properly accounted for. Several highly-influential individuals were identified; their inclusion produced significant changes to the model fit. Ultimately, they were retained for descriptive purposes, but had the model been developed for simulation, this decision might have been different. Bootstrapping produced robust 95% confidence intervals, of critical importance when assessing bioequivalence. Outermost confidence intervals stabilized sufficiently for confident interpretation after 600-700 successful runs. Based on the 61% success rate, the final model is likely to be somewhat overparameterized – there is probably insufficient information in the data to support all 59 model parameters.

A model-based approach to showing bioavailability through parameter similarity was shown to be effective given sufficient appropriate data. Bootstrapping and case-deletion diagnostics were pivotal in determining confidence intervals sufficiently robust to judge bioavailability criteria, and in highlighting previously-unidentified study differences. It is clear that these techniques are of value in an industry setting despite the large amounts of time and processing power required for their use.

References

[1] Hayashi N, Tsukamoto Y, Sallas WM, Lowe PJ. A mechanism-based binding model for the population pharmacokinetics and pharmacodynamics of omalizumab. *Brit J Clin Pharmacol* 2006;**63**: 548-561.
[2] Lindbom L, Pihlgren P, Jonsson EN. PsN-Toolkit – A collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Methods Programs Biomed* 2005; **79**: 241-257.
[3] Spiegelberg HL, Simon RA. Increase of lymphocytes with Fc receptors for IgE in patients with allergic rhinitis during the grass pollen season. *J Clin Invest* 1981;**68**: 845-852.
[4] Jung CM, Prinz JC, Rieber EP, Ring J. A reduction in allergen-induced FcERII/CD23 expression on peripheral B cells correlates with successful hyposensitization in grass pollinosis. *J Allergy Clin Immunol* 1995; **95**: 77-87