Pharmacokinetic and pharmacodynamic analysis in anesthesia : a modeling odyssey

Olofsen, E.

Chapter 7

Summary and Conclusions

7.1 Expected and Unexpected Findings

AS ALLUDED TO in **Chapter 1**, violations of modeling assumptions may have significant effects that are hard to foresee. In the present chapter both expected and unexpected effects are summarized. The findings suggest that, if the various sources of bias and variability are not properly taken into account, the following warnings may be warranted:

- A measurement method may mistakenly seem to be interchangeable with its golden standard (Chapter 2).
- The prediction error (the weighted difference between measurement and model output), when used as a validation criterion, may be more variable than necessary, leading to a higher probability of selecting a less than optimal model (Chapter 3).
- Intra-individual process noise may be mistaken as variability between measurements and/or individuals (Chapters 4 and 5).
- Blood–effect-site equilibration half-life and potency of a drug may be underestimated (Chapter 6).

7.2 The 95% Confidence Intervals of the Limits of Agreement

Bland-Altman methods to assess agreement between two measurement methods were studied in **Chapter 2**. The warning in the literature was confirmed that when multiple measurements have been obtained in several individuals, an analysis that does not take this into account may be expected to give limits of agreement that are too narrow. However, using simulations it was shown that this has even more of an effect on the confidence intervals around these limits. When reporting how closely measurements between two devices are related, the accuracy of the limits of agreement is just as important as the limits themselves. Clearly, suitable software that permits easy calculation of these confidence intervals can be helpful in assessing the value of medical devices. To that end, an open-source web application was developed so that a Bland-Altman analysis can be performed without the need to install any software apart from the ubiquitous

web browser. In previous studies wrong conclusions on agreement between two methods may have been reached, particularly when the number of subjects was small. To avoid inconclusiveness, it is proposed that studies that use Bland-Altman methods of comparison should follow a standard format. By providing sufficient data on the assumptions underlying an analysis of agreement next to the results, especially the 95% confidence intervals of the limits of agreement and inter- and intra-individual variation, ambiguity can be reduced and confidence in the results increased.

7.3 Akaike's Information Theoretic Criterion

Akaike's Information Theoretic Criterion (AIC) is a number representing a model's goodness of fit, relative to competing models. The simulations performed in **Chapter 3** demonstrated that, at least in a relatively simple mixed-effects modeling context with a set of prespecified models, minimum mean AIC coincided with best predictive performance.

It was found that in the presence of interindividual variability, prediction error by itself becomes a less suitable validation criterion, because it does not take into account whether estimated interindividual variability matches the variability in the validation data. The context of AIC is the one where the random effects have been integrated out, with the parameters at their (estimated) population values, which is to be done when all data are acquired. This holds also for the validation data, so this context is different from the case where prediction errors are calculated with the random effects set to zero. In other words, interindividual variability is predicted as well; the distributions of the model parameters are estimated to allow optimal prediction of a new set of data, even when the individualized model parameter values remain unknown until enough data are gathered.

7.4 Kalman-Filtered Concentrations and Measures of Analgesia

The opioid buprenorphine significantly increased the resting state EEG ratio (a surrogate EEG measure of analgesia) and skin pain tolerance compared with placebo, as was demonstrated in **Chapter 4**. A stochastic model was applied to the data, which adequately characterized the concentration-time and effect-time courses for both the skin heat stimulation and the resting state EEG ratio outcomes, with variations in the drug's absorption rate during a 144-hour treatment period. As measured by the potency parameter, the EEG effect was about 10 times more sensitive to buprenorphine than the skin pain test. The findings suggest that the resting state EEG ratio is an objective alternative for assessing opioid effect.

The stochastic PK-PD analysis was successful, in the sense that three kinds of random sources could be identified: variability between individuals, variability within individuals, and variability in measurements. This allowed the computation of a time-dependent variability in drug absorption from patch to blood. However, the effects of ignoring this variability remained unknown.

7.5 Kalman-Filtered Surrogate EEG Measures of Anesthesia

An example where the standard two-stage (combining results from separate fits for each individual) and nonlinear mixed-effects modeling (NONMEM) approach yielded nearly identical parameter estimates was encountered in **Chapter 5**. Furthermore, it was found that the interindividual variability identified by a mixed-effects but otherwise standard PK-PD analysis was for a large part actually intra-individual variability, namely process noise.

Analysis of permutation entropy data calculated from raw EEG measurements with a first Kalman filter design displayed a large value of the steepness parameter ($\gamma \approx 17$) of the sigmoid-$E_{\max}$ model. As a consequence, the model output is very sensitive in the region of loss of consciousness, but very insensitive in deeper levels of anesthesia. In contrast, analysis of the data with an alternative Kalman filter design showed a relatively low value ($\gamma \approx 2.5$), where the model output responds smoothly to changes in anesthetic concentration.

Simulations showed that the model parameters could be reliably estimated. Estimated parameter values were similar if the Kalman filter was present or absent in the model, except for the interindividual variability estimates. Without the Kalman filter, these were overestimated with a factor of 10–30. Furthermore, the steepness parameter was not overestimated. Interestingly, the largest interindividual variability (coefficient of variation about 25%) was found to be present on the standard deviation of the process noise.

7.6 Sampling Site Bias

Arteriovenous morphine-6-glucuronide (M6G) concentration differences were analyzed in **Chapter 6**. Arterial plasma concentrations were higher just after infusion, whereas at later times venous concentrations exceeded arterial concentrations. An extended pharmacokinetic model adequately described the data; it consisted of three arterial compartments, one central venous compartment, and one peripheral venous compartment.

The simulation studies revealed large biases in model parameters derived from venous concentration data. The biases were dependent on the value of $t_{½,k_{e0}}$, the blood–effect-site equilibration delay. Assuming that the true value of M6G's $t_{½,k_{e0}}$ may be in the range of 120 to 240 minutes (depending on the endpoint measured), we would have underestimated $t_{½,k_{e0}}$ by 30%, whereas the potency parameter would have been overestimated by about 40%, when using venous plasma samples.

A delay between arterial and venous concentrations would not be unexpected. If $t_{½,k_{e0}}$ would be estimated based on venous data, we would expect a smaller value than if it would be based on arterial data. There are two other PD parameters, C_{50} and γ; and interestingly their values may also be biased when estimated from venous data. Most of the duration of the experiment, venous concentrations were higher than arterial concentrations, so when the effect occurs at higher concentrations, this would lead to an upward biased C_{50}.

While this was not investigated, the biases are likely dependent on the administration schedule. That means that the biases are very hard to know in advance, if the admin-