The handle http://hdl.handle.net/1887/50818 holds various files of this Leiden University dissertation.

**Author:** Olofsen, E.
**Title:** Pharmacokinetic and pharmacodynamic analysis in anesthesia: a modeling odyssey
**Issue Date:** 2017-06-21
Chapter 1

Introduction

1.1 A Modeling Odyssey

When modeling data from experiments, many explicit and implicit assumptions have to be made about the deterministic and random processes that produced the data. One of the assumptions is that “unexplained” variation in the measurements is independent and identically distributed. “Independent” means that the value of a measurement is not influenced by previous (or subsequent) ones. One situation where this assumption is easily violated is with data from multiple subjects, because of differences between the characteristics of those subjects, which are the same within the data from each subject.

The collection of assumptions is referred to by the term “model”. To estimate both the parameters of the deterministic part of the model and the properties of the random processes, the program NONMEM was developed in the 1970s, and its development continues to this day. NONMEM is an acronym for “NONlinear Mixed-Effects Model”.

From the beginning, it was called into question if the complexity, and the associated computational cost, of NONMEM was worth the effort. Warnings that model parameter estimates and/or their standard errors could be biased if assumptions were violated, were balanced by examples that the parameter estimates and their standard errors were about the same with more naive estimation methods.

Part of the modeling process with NONMEM often involves obtaining successful estimation and covariance steps, specified as $ESTIMATION and $COVARIANCE records in a control file, which starts with the promising $PROBLEM record. Even when the syntax of the model specification records is correct, these commands usually lead to an unexpected (but correct!) response from the program - hence the leaflet interpretation of a famous dialogue in the science fiction literature. The estimation step typically takes a lot of computer time, with the possibility that it fails with the message that the objective function is infinite. The covariance step should output the standard errors of the estimated parameters. It also takes a lot of computer time, and it fails if it encounters a singular matrix, which means that, even with the model specified, there is still an infinite number of ways to describe the available data. The objective function is a number related to the likelihood of observing the data, conditional on the values of model parameters. Usually, and in NONMEM, it is calculated as minus two times the natural logarithm of the likelihood (discarding terms that are constant with respect to the model parameter values). So if the objective function is infinite, there is actually zero likelihood of observing the data. With a reasonable model, such a result does not
Chapter 1

seem to make sense. To obtain standard errors of the parameter estimates, the “Hessian” and/or “cross-product gradient matrix” needs to be inverted. If such a matrix is singular, it cannot be inverted. This might be caused by the model having too many parameters. But in this situation, there is no output of a possibly large standard error indicating which parameter causes overparameterization. One reason that NONMEM might fail to give successful and/or useful estimation and covariance steps, is that the model specified in the control file contains explicit or implicit assumptions about the data that are violated. For example, by writing

\[ \theta_i = \tilde{\theta} \cdot \exp(\eta_i) \]  

(1.1)

it will be assumed that model parameter \( \theta_i \) is lognormally distributed around \( \tilde{\theta} \) across the population. The lognormal distribution is unimodal, but for example a population with groups of low- and high responders could be bimodal, leading to estimation problems.

In the following chapters of this thesis, different conditions where assumptions are violated will be investigated. Most of the situations are explored using simulated data, but the characteristics of these data were based on studies in patients or volunteers. The next sections in this chapter introduce those situations by the people that first proposed solutions to handle these situations. In all cases, these solutions are given by single characteristic formulas.

1.2 The Limits of Agreement

One of the papers in Nature's Top 100 is called “Statistical methods for assessing agreement between two methods of clinical measurement”, written by Martin Bland and Douglas Altman in 1986. In this important paper, the authors explain why the correlation coefficient may not be a good measure of agreement. For example, the correlation coefficient may be high, while agreement is low, because the correlation coefficient is not sensitive to scaling factors such as a bias term. Bland and Altman proposed to calculate “limits of agreement” (LoA):

\[ \text{LoA} = \bar{d} \pm 2s, \]  

where \( \bar{d} \) and \( s \) denote the mean and standard deviation of the differences between the measurements, respectively. The mean is the bias of one measurement method with respect to the other. The above equation looks familiar: the LoA contain 95% of the differences, if these follow a normal distribution. However, the properties of the differences are not known and have to be estimated, so Bland and Altman also showed how to calculate \( \bar{d} \) and \( s \).

Very often, data are obtained from a group of subjects with multiple paired measurements in each subject. Bland and Altman warned that in this case the differences may not be considered as statistically independent, and estimated limits of agreement could be too narrow if this is not taken into account. However, the calculations become much more complicated, for which there is no readily available software.

Therefore, in Chapter 2, the development of a freely available implementation in JavaScript is described, which is able to run in a Web browser. We validate the implementation by giving a formal description of both the basic and more advanced Bland-Altman comparison methods, and by using simulated data so that it can be verified.
that the calculations are correct. We also study the effects caused by failing to take the presence of multiple paired measurements per subject properly into account. Because the results depend on the properties of the data, and analysis methods used, we list important items for a standard format of reporting comparison studies.

### 1.3 An Information Theoretic Criterion

Another entry in Nature’s Top 100 is called “A new look at the statistical model identification”, written by Hirotugu Akaike in 1974. He proposed “An Information theoretic Criterion” for model selection - AIC - which he invented while taking a seat in a commuter train according to his review of his Citation Classic. In his honor, the first letter of AIC is usually pronounced as “Akaike”. The criterion may be written as

\[
AIC = -2 \log L + 2p, \tag{1.3}
\]

where \( L \) is the likelihood of observing the data, and \( p \) the number of parameters of the model used to calculate \( L \). It is often stated that when AIC is used for model discrimination it leads to “overfitting”, i.e., it selects models with a higher dimension than the dimension of the model that generated the data. But for example with experimental pharmacometric data it may not be possible to identify the correct model, because of the complexity of the processes governing drug disposition and action. Instead of trying to find the correct model, a more useful objective might be to minimize the prediction error of drug concentrations or drug effects in subjects with unknown drug characteristics. In that case, the AIC might be the selection criterion of choice.

In Chapter 3, we perform Monte Carlo simulations using a model of pharmacokinetic data (a power function of time) with the property that fits with common multi-exponential models can never be perfect - thus resembling the situation with real data. AIC and AICc (the criterion with a correction for small sample sizes) values are calculated and averaged. The average predictive performances of the models, quantified using simulated validation sets, are compared to the means of the AICs. These simulations are also done at different levels of interindividual variability in the pharmacokinetic volume of distribution, to check that AIC remains a valid criterion under these circumstances.

### 1.4 The Kalman Filter

Another “Citation Classic” is on a paper by Rudolf Kalman. It contains the mathematical foundation for another paper from the same year, introducing the “Kalman filter”. The Kalman filter filters measurements to obtain optimal estimates of the state of the system and its uncertainty, by feeding back the difference between measurements and predictions multiplied by what is now called the “Kalman gain” \( K_k \):

\[
K_k = P_k C_k^T \cdot \left( C_k P_k C_k^T + R_k \right)^{-1} \tag{1.4}
\]

where the terms on the right side of the equation denote properties of the system and of the noise perturbing the system and the measurements.
In Chapter 4, the pharmacokinetic-pharmacodynamic (PK-PD) properties of buprenorphine transdermal patch in healthy volunteers on electroencephalographic (EEG) characteristics and pain tolerance are studied. Because the latter effect is often based on subjective measurements, electroencephalography offers a possibility to objectively quantify and track the changes in the the activity of the brain when an opioid is administered.

Usually, the pharmacokinetic and pharmacodynamic states are assumed to deterministically depend on drug administration only. However, there could be variability in the absorption rate from the patch and/or in the blood–effect-site equilibration rate. The estimates of buprenorphine’s properties could be biased if such variability is not taken into account. Therefore, a population PK-PD model with stochastic differential equations was implemented in the NONMEM to analyze the PK and PD data simultaneously.

1.5 The Entropy of Permutations

In the 1940s, Claude Shannon proposed the following measure of uncertainty:

\[
\text{Entropy} = - \sum p_i \log p_i, \tag{1.5}
\]

where \( p_i \) is the probability of event \( i \) to be observed. Shannon selected this function of the probabilities because of its desirable mathematical properties. For example, entropy is maximal if all events are equally likely to occur, and entropy is zero if none of the events are likely to occur, except one.

A frequently used electroencephalographic index of central anesthetic drug effect is the “Bispectral Index“ (BIS). The algorithm that is used to calculate the BIS is not in the public domain, and is subject to repetitive updates with unknown changes. On the other hand, open source indices are often highly sensitive to artifacts due to muscle activity (e.g., eye blinking).

The EEG waveform can be described as a sequence of ordinal patterns. The permutation entropy (PE) describes the relative occurrence of each of these patterns. The normalized PE is high (almost maximal) when the signal has predominantly high frequencies and low (approximately 40% of maximal) when the signal consists of only low frequencies. The permutation entropy was shown to be insensitive to eye blink artifacts. With high amplitude eye blinks, the low amplitude high frequency components indicative of awakeness still dominate the permutation entropy, because permutations are not dependent on amplitude per se, but only on amplitude rankings.

The cost of sampling EEG indices is quite independent on the amount of samples, so these can be acquired at high sampling rate. However, PK-PD model fits then show correlations between the samples. For example, if a measurement is above the line representing the model fit, there is a high probability that the next measurement is also above the model output. This represents another situation where the data are not independent, and model parameter estimates may be biased unless a Kalman filter is used. Therefore, in Chapter 5 models using stochastic differential equations for the state of the brain are studied using simulated and experimental data.
1.6 Location, Location, Location, of the Sampling Site

About 40 years ago Lewis Sheiner\textsuperscript{[73]} (but see also Hull\textsuperscript{[41]} and Segre\textsuperscript{[74]}) pointed out that it is important to take into account a delay between drug concentration in the arterial blood and the effect-site:

\[
\frac{dC_e(t)}{dt} = k_{e0} \cdot (C_b(t) - C_e(t)),
\]  \hspace{1cm} (1.6)

where \(k_{e0}\) is the blood–effect-site equilibration rate. This equation is able to describe, for example, the phenomenon of increasing drug effect with decreasing blood concentration - namely for as long as \(C_b(t)\) is higher than \(C_e(t)\). In pharmacokinetic-pharmacodynamic modeling studies, venous plasma samples are sometimes used to derive pharmacodynamic model parameters. In principle, the same equation can be used, but because there is an arterio-venous delay, \(k_{e0}\) will be estimated with a bias towards a smaller value.

In Chapter 6, the extent of arteriovenous concentration differences of morphine-6-glucuronide is quantified based on arterial and venous blood samples in volunteers. An extended pharmacokinetic model is described with standard compartments for the arterial data that are linked to additional compartments for the venous data. The extent of bias in pharmacodynamic model parameter estimates is explored via simulation studies with NONMEM. Furthermore, simulations are presented where a pharmacokinetic model based on arterial data is connected to a pharmacodynamic model based on venous data, to assess the influence of this mismatch on the predicted effect.

1.7 NONMEM and Beyond the Infinite

As may be expected, violations of modeling assumptions may have just mild but also significant effects on the data analysis results. As may also be expected, violations may also have significant modeling effects that are perhaps not expected in advance. Therefore, in Chapter 7, the findings of the five studies that are presented in this thesis will be revisited in detail. The existence of increasingly more powerful computers and increasingly more sophisticated software like NONMEM is invaluable for studies like those described, because it provides the means to look one or more dimensions deeper than the present state-of-the-art.