Analyzing time-varying patterns of human exposure to xenobiotics and their biomedical impact

Bernard Vrijens

Aan mijn ouders,

Dankwoord

Nu de resultaten van een jarenlang onderzoek zijn neergeschreven in een doctoraal proefschrift is het moment aangebroken om een aantal mensen te bedanken.

Mijn dank gaat in de eerste plaats uit naar mijn promotor, Prof. Els Goetghebeur die met zeer veel belangstelling deze thesis heeft begeleid en mij zoveel (right to the point) suggesties deed. Haar betrokkenheid en enthousiasme heeft mij door de jaren heen gemotiveerd en gestimuleerd, en zonder Els zou dit proefschrift niet in deze vorm nooit geproduceerd zijn.

Naast mijn promotor wil ik ook Prof. John Urquhart bedanken voor de vele, buitengewone interessante discussies over drug exposure en intuitive hints. Daarnaast heeft zijn inzicht in de klinische relevantie en gevolgen van variabele drug exposure voor een verdieping van het proefschrift gezorgd.

Ik ben eveneens dank verschuldigd aan Prof. Jan Willems omdat hij mij de mogelijkheid geboden heeft om gedurende een jaar onderzoek te verrichten in zijn departement. Daarnaast wil ik hem bedanken voor de mogelijkheid een breder aspect te kunnen bestuderen van exposure.

Aangezien mijn onderzoek in het Europees Biomed II-project `THER-MOS' past, dank ik ook de leden van dit project.

Ik wil ook alle werkgevers bedanken voor de tijd en steun die ze de afgelopen jaren hebben aan mijn onderzoek hebben gegeven: Prof. Adelin Albert, Dr. Patrick Debouck and Dr. Jean-Michel Métry.

Eveneens dank aan alle collega's voor de fijne werksfeer en samenwerking.

Tenslotte wil ik mijn familie bedanken voor hun steun en geduld, met name tijdens de laatste maanden van afronding van dit proefschrift.

Bernard Vrijens
Gent, 30 April 2002

Contents

Chapter 1

Introduction

This thesis deals with the analysis of exposures to xenobiotic substances and their effects on human beings. The term xenobiotic covers all organic compounds that are foreign to the organism under study. This last century many new drugs with expected curative and palliative effects have lead to a major reduction in mortality and morbidity at all ages. As a result life expectancy has increased dramatically. Also, in many geographical areas, the earlier patterns of large epidemics and famine have been largely relieved, whilst in other areas the very same problems seem intractable. The improvement of well being has gone hand in hand with a drastic expansion of the industry. Not only the food production industry has grown but also a whole business has emerged aimed at improving the social condition of many people. The discovery of new technologies, new products, has exposed human beings to new and often not so well understood compounds (drug residue, fertilizer, pesticides, dioxins, PCB, ...). Coffee is a good example; over 1200 different compounds are identifiable in the natural product called coffee which was just another weed until the 18th century when the coffee craze hit Europe. As a consequence, we are daily exposed to hundreds of chemical compounds. Most of them are unlikely to have a major effect on health but this huge diversity in chemical compounds with their own way of action, combined with individuals who are highly variable (health status, diseases, age, race, ...) makes it difficult to assess the effectiveness and/or toxicity of the products.

Health scientists have routinely made decisions about the extent to which human exposure to chemical compounds is beneficial, hazardous or negligible. These decisions commonly focus on a target group, for

example, the general population of a country, patients suffering from a particular disease, a highly exposed sub-population, or specific individuals with some characteristics in common, such as children. Individuals within a target group are likely to face different levels of exposure due to differences in behavioral patterns (food consumption, working environment, adherence to prescribed therapy, ...), or physiological characteristics (absorption, distribution, drug metabolism, ...), or even over time. These differences contribute to the variability in health outcomes among individuals. At the beginning of the 20th century the expert decisions were largely a subjective qualitative assessment of evidence from individual case reports. Today judgments on drug efficacy and contaminant control are quantitatively guided with the help of the science of statistics. One synthesizes logical relationships, estimated magnitudes of effects and residual variation based on information contained in a given sample of chosen size.

When examining a dose-effect relationship, actual exposure to the studied chemical compound can be seen as the main driver of outcomes. As a consequence, uncontrolled variation in exposure will lead to variation in outcomes and will thus result in less powerful statistical tests. Therefore, researchers aim to study subgroups of the general population consisting of `standardized' subjects with predictably high exposure levels. The hope is then to highlight a `significant average effect' on a sizable sample of the studied population. Indeed economic pressures on sample size often force the researcher to reduce variability as much as possible. However in this \ideal" world, one looses representativeness and restricts observation to a narrow fraction of the outcome distribution and often loses important information. Further, for most of the xenobiotic substances, it would be an utopia to assume that in general ambulatory subjects are exposed to a constant dose over time. This thesis shows that assuming regular exposure and steady state in ambulatory subjects can result in biased conclusions concerning the pharmacokinetic/pharmacodynamic properties of the compounds. On the other hand, it suggests that using reliable measures of exposure, as observed in a more general population can result in a clinically relevant pharmacologic model that improves the information retrieved from a given number of subjects.

1.1 Chemical compound exposure

In modern society, humans are continuously exposed to numerous (several hundreds of thousands) chemical compounds. The exposure happens mainly via three different routes : continuous environmental exposure (daily exposure to air and/or food pollutants), intermittent exposure to specific products (war, occupational, smoking, accidental food contamination, ...) and, finally, prescribed drug exposure. While for the two first exposure routes health hazards are expected, the last route of exposure has usually been studied in detail within clinical trials and is only put into effect when a good balance between health benefit and adverse reactions is anticipated.

The impact on human health of environmental factors (daily exposure) can only be assessed through epidemiological studies. Since serious health outcomes show a low prevalence rate and are usually observed only after a long period of time, the number of people that must be monitored to detect an average increase is huge. Acknowledging the many possible interactions between the chemical compounds present in the environment, the study of their health hazard would require involving close to the entire world population in epidemiological studies. Not only is this prohibited by practical, ethical and economic reasons but within such large scale epidemiological studies it would be impossible to control for potential confounding factors.

When an incident of exceptionally high exposure occurs, it constitutes a unique opportunity to study in detail the biomedical effects of the xenobiotic substance in question. As an example, we will study exposures to dioxin-like substances during a short food contamination episode in Belgium. Our work however focuses mainly on exposures that happen through drug prescription in general.

Fortunately prescription drugs are a more common subset of exposure to xenobiotics and will be considered in more detail through several examples. In drug development, except for early phase studies where the patient is treated in hospital, patients tend not to take their medicine exactly as planned by the trialist. This lack of adherence results in a natural experiment in variation of drug exposure that can take a number of forms. The patient can be non persistent and drop out of the trial treatment, take none of the medicine whilst pretending to do so, forget to take the treatment from time to time, or take it at the wrong time.

1.1.1   Variation in exposure

The first problem we run into when discussing variation in exposure is one of terminology. Many terms exist and have been used interchangeably in the literature, e.g. compliance, duration, persistence, adherence, acute phase, drop out, ... As with any scientific undertaking, a sound taxonomy, with explicit definitions, is essential for scientific progress. To define the variation in exposure, we follow Urquhart's (Urquhart and Vrijens 2001) perspective and take a leaf from Aristotle. He viewed the drama as having three parts: a beginning, a middle, and an end. This stunningly obvious construct is nevertheless a good taxonomic starting point for describing variation in exposure, which also has these three main parts. The beginning is the time when a person is exposed to the studied chemical compound. The middle is the continuous exposure over time once that exposure has started. The end is an abrupt stop to the middle phase. The beginning and the end are abrupt changes in exposure at certain points in time. Usually, the beginning and end of an exposure can be predicted (prescription of a medication, decision to stop a medication, change in occupational exposure, relocation, food contamination crisis, ...).

In drug development, variation in exposure is controlled in designed clinical trials. Early phase studies are usually carried out within hospitals where variation in exposure is imposed and can be well controlled. Variation within-patients (drug escalation, cross over, ...) as well as between-patients (study groups with different doses, different regiments, ...) is of interest. This constitutes usually the basis of dose-finding through detailed traditional pharmacokinetic/pharmacodynamic studies. In later stage clinical development, when patients are treated for longer periods of time and multiple doses are prescribed, accurately recorded timing of repeated dosing is essential for accurate and precise estimation of the pharmacokinetic and pharmacodynamic properties of drugs.

In traditional ambulatory pharmacotherapy, once the drug has been prescribed, any variation in drug exposure is due to the patient (intentional variation or not). Therefore a specific taxonomy has been developed in this field. The beginning is the patient's acceptance, or not, of the recommended treatment. The middle is the patient's execution of the prescribed dosing regimen. The end occurs when treatment is discontinued, which in the trial setting is called \dropping out of treatment". The first and third phenomenon in the occurrence are a dichotomous outcome at a single point in time, whereas the second is in general a

continuous process over time, thus precluding a single, quantitatively useful parameter to cover all three. However, the literary notion of \ad- herence", may be useful as a blanket term for all three, provided that it be generally recognized that \poor adherence" may mean problems with any or all of the three. Thus, beneath \adherence", we have:

- **Acceptance, the dichotomous beginning,**

- **Execution, the continuous middle,**

- **Discontinuation, the dichotomous end.**

Time is not to be overlooked, for it is a factor of paramount importance. The event is discontinuation, but the duration of execution {persistence{ is a crucial measure.

1.1.2 Measuring exposure

While a study of the action of chemical compounds ideally relies on precise exposure assessments, in practice, it is difficult to collect reliable data over time. Measuring exposure is the first step to achieve this. Retrospective and/or prospective methods are used.

- **Retrospective data collection**
  The most commonly used retrospective method is self report. However the reliability and validity of self reports is highly variable (Waterhouse et al. 1993, George et al. 2000, Arnsten et al., 2001). When the questionnaire covers a long period of exposure time, we first face potential influences of elapsed time. This is especially true if detailed daily exposure is requested. Furthermore intentionally biased responses (face-saving, potential profit in exaggerating the toxic exposure, ...) is an important source of bias in this type of study. In practice, retrospective data collection via questionnaires, while still broadly used, is prone to producing inaccurate measures of long term exposure and can seriously confound the analysis and interpretation of exposure-response relationship. For some chemical compounds it is sometimes proposed to assess previous exposure via plasma concentrations. A measure of drug concentration in plasma is usually expensive and reflects exposure that occurred only during a preceding time-window, the width of which is ca. 3 times the plasma half-lives of the compound.

For compounds with long half-life (usually pollutants) this assessment can be used to summarize long term average exposure but cannot identify variation in exposure and cannot distinguish between short acute exposure and long background contamination. Knowing that 78 % of the drugs in therapeutic use have plasma half-lives that are less than 12 hours with 61 % of the plasma half-lives less than 6 hours (Hardman et al. 2001), this method restricts its estimate to the narrow intervals of time just prior to scheduled visits. Furthermore, we note that `white-coat adherence', the phenomenon whereby the patient improves adherence immediately before the consultation, can substantially exaggerate estimates of adherence outside this narrow interval of time just prior to a scheduled visit. As an alternative to plasma concentrations, measurements of biological effects are sometimes used. While it provides a quantitative assessment and is usually less expensive than plasma concentrations, incidental biological effects are not always present and are highly variable across human beings. Finally the so-called pill-counts, where the trialist counts the returned tablets, have been used extensively as a routine adherence measure in clinical trials. The technique is simple, cheap and non-invasive but provides only an average adherence summary over a given period of time. This method fails to distinguish between a regular taker and someone who balances periods of underdosing with periods of overdosing. Further pill-counts have often been shown to over-estimate patient adherence (Waterhouse et al. 1993, Pocock SJ and Abdalla M 1998). In summary retrospective data collection is only reliable for single day (or average longer period) exposure assessment and faces then the difficulty of interpretation. Having a single-point value does not readily improve the interpretation of exposure-response data nor does it translate easily into a plan of sensible intervention.

- **Prospective data collection**
Direct observation of exposure is the gold standard but is resource-intensive, requires continuous on-ward residence and thus may not have external validity for routine daily life. It is often used in animal studies or for restricted early drug development human studies. Prospective diaries are inexpensive but have been shown to be inaccurate and biased for similar reasons than for retrospective patient self report. Finally with the advent of micro electronic circuitry, electronic monitoring has become feasible. Electronic real

time monitoring allows unparalleled accuracy and convenience in assessing not only how much the exposure was, but also when it happened. It delivers long term temporal data and reflects the variation in exposure over time. It is usually easy to use but cost and reliability vary. Continuous monitoring of drug containers openings make the unobtrusive assessment of continuous medication intakes possible, yielding patterns of dosing history over time. However, each opening and closing of the monitor is recorded as a presumed dose, which obviously does not prove drug ingestion. Nevertheless, the view exists that electronic drug monitoring is a quite accurate surrogate of drug exposure because exceptional and sustained commitment is required to create a false record of good adherence (Urquhart and de Klerk 1998). This is especially true for long term drug monitoring. This kind of continuous monitoring results in a large amount of complex data and appropriate statistical analyses must be developed in order to deal optimally with them. Several classes of problems will be presented in this thesis; methodological answers will be proposed, applied and discussed.

1.2   The therapeutic cascade

The duration and intensity of an exposure is arbitrary and can have a completely different effect depending on the chemical compound. Therefore, understanding the relationship between exposure and blood concentration (pharmacokinetics, PK) and between blood or plasma concentration and biological effect (pharmacodynamics, PD) is important. It drives the rational development and the safe and effective use of a therapeutic agent as well as the regulation of tolerated intake of pollutants. In Figure 1.1, we schematically represent the biotransformation process between xenobiotic exposure and response.

To characterize these relationships and their variability adequately, studies in a representative population, using a relatively large number of subjects, are needed. However, for practical and ethical reasons, extensive pharmacokinetic and pharmacodynamic studies in large numbers of subjects representative for the population exposed are usually not possible. One must therefore derive pharmacokinetic and pharmacodynamic relationships based on sparse data (few data points per patient) collected under unbalanced designs. This rationale for population approaches in clinical evaluation was recognized already in the early papers

by Sheiner (Sheiner and Ludden 1992) in which this methodology was
first proposed.

In spite of this obvious need, the application of the population approach
was for many years quite limited. The main reasons being the complexity
of the statistical methodology, the lack of available software and the need
for relevant exposure assessment.

Xenobiotic exposure
(Drug or Toxin exposure)

?

PK : Absorption, distribution, metabolism
(Tissue drug concentration)

?

PD : Drug action on cellular process
(Measurable biological effect)

?

Clinical response
(Effect on patient health)
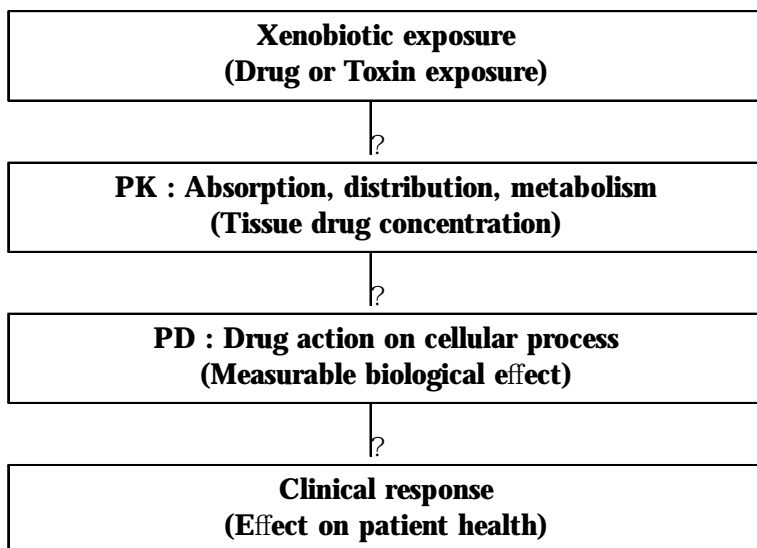
Figure 1.1: The therapeutic cascade

The chapters of this thesis illustrate several steps of the therapeutic
cascade. They were motivated by real problems, emerging from data
gathered to answer important scientific questions. Practical existing
problems related to variation in exposure and related effects were pro-
vided by Aardex, Abbott, Eli Lilly, M.S.D., Servier, and the Belgian
ministry of public health and will be described further in the thesis.

Whatever the design, experiments assessing exposure over time are usu-
ally expensive, demanding resources and generally resulting in large data
bases. These efforts could be wasted if the design or analysis was not
optimally performed. Traditionally, textbooks on exposure mainly deal
with the situation where average values or basic summary statistics are
considered. Summarizing the exposure history in a few measures may
lose too much information to detect relevant exposure patterns and es-

pecially changes over time. We found it necessary to go further and draw inference that exploits enough of the data in a way that captures the temporal evolution of exposure. This issue will be illustrated throughout this work. Then inclusion of richly detailed information on exposure can improve considerably the understanding of exposure-effect on human health. However if not done appropriately, it could confound the conclusions. As will be shown in this thesis, including the variability in exposure input may lead to much more precise and less biased pharmacologic and/or toxicological parameters.

1.3   Structure of the thesis

In Chapter 2, we investigate several summary measures that highlight different dimensions of adherence patterns and the drug context in which they may be meaningful. Further, we examine conditional and marginal models that enable comparisons of the full pattern of daily dosing indicators for subjects between groups. This highlights the need for detailed data in order to develop more powerful statistical methods to compare drug exposure between randomized treatments over time.

In chapter 3 we present methods for predicting patient compliance over time. To predict compliance outcomes, we use an auto-regressive model for binary data. We handle frailty by a normal distribution for random slopes and a random intercept. Model validation as well as two different types of clinically relevant model prediction are described and further discussed.

In chapter 4 we introduce the bases of population pharmacokinetics and discuss patient adherence to prescribed therapy as a leading source of variability in pharmacokinetic response. We quantify the impact of observed drug variation in population pharmacokinetic studies through simulations. Finally we conclude that when it is possible to observe and record irregular drug intake times, a substantial amount of precision in the pharmacokinetic parameters is retrieved from the same number of data points.

In chapter 5 we move on to estimate pharmacodynamic parameters, from just a few outcome measures on a sample of patients who are partial adherers. We investigate the information matrix for hierarchical nonlinear models and confirm that a substantial reduction in bias and gain in precision can be expected by observing irregular drug intake. Here we find support for the claim that nonadherence, as a rich natural

experiment of dosing variation, can be a blessing rather than a curse from the information/learning point of view. In general, the estimators of pharmacokinetic/pharmacodynamic parameters benefit greatly from information that enters through greater variation in the drug exposure process.

In chapter 6 we investigate the effect of patient's adherence to prescribed therapy on response assuming that adherence is non selective. In a first step we estimate individual pharmacokinetic parameters using a one compartment model and adjusting for observed dosing intakes. Then we relate pharmacological action to plasma concentration of drugs. We applied the methodology to the HIV field where it constitutes a first attempt to individualize therapy for the patient. Finally we derived a new adherence summary variable, inspired from the third moment of the interdose interval distribution, that allows to estimated a direct relation between dosing histories and response.

In chapter 7 we move on to estimate the true therapeutic biological effect of the treatment active component when potential selectivity is expected. The methodology focus on causal inference in randomized trials for comparing two active treatments. The method is applied to an antidepressant trial where a strong psychologic adherence effect can be expected on outcome.

In chapter 8, we introduce two dimensional Monte Carlo simulations as an appealing technique to combine data collected from different sources. Between-subjects variability and uncertainties due to non optimal data collection or biased sampling design are then propagated in two distinct dimensions allowing for proper interpretation. Simulation is hereby not a substitute for mathematics but an approach to calculation, resampling, similar to using bootstrap, for estimating the variance of a complicated point estimator. A simulation approach frees us from certain technical constraints to derive a probable causal relationship between exposure and outcome.

Finally, in the last chapter, we discuss the general ideas highlighted through the thesis and give topics for future research.

Chapter 2

Comparing adherence patterns between randomized treatments

When two equally efficacious drugs enter the market, the one with the better adherence is likely to be used by more patients. Special management of the drug delivery system may produce increased adherence. In this chapter we analyze a trial of a single drug dosing prescription with patients randomized to either daily self monitoring of the outcome (blood pressure) or not. The study used Medication Event Monitoring Systems (MEMS, Aardex Ltd.) to record each exact time and date when a patient opened the pill container. No established method is available for comparing these high-dimensional adherence patterns between groups. This paper investigates several summary measures that highlight different dimensions of the pattern and the drug context in which they may be meaningful. Further, we examine conditional and marginal models that enable comparisons of the full pattern of daily dosing indicators for subjects between the groups. We found no simple difference in average adherence levels, but we found an interesting interaction between treatment and time : similar adherence existed initially among patients in both randomized groups, with a stronger decline over time for patients who did not monitor their blood pressure. We discuss how a balance between simplicity of interpretation and efficiency of data use may be sought in this case.

2.1   Introduction

The likely crucial role of treatment adherence in the effectiveness of therapy and its potential to compete with other available drugs has become increasingly clear, both at the pharmacologic and at the clinical level (Cramer et al. 1989, Efron and Feldman 1991, Cramer and Spilker 1992, Sheiner and Rubin 1995). Many clinical trials, therefore, carefully measure adherence with assigned treatment, particularly when the trial is designed to evaluate the effect on adherence of different prescribed dosing strategies. Medication Event Monitoring Systems (MEMS, Aardex Ltd.), a relatively novel technique, record for each patient a series of exact times and dates at which the drug container was opened (\drug intakes") over the course of the study. The resulting data provide a detailed source of information on drug-taking behavior that must be summarized comprehensibly so it can be used to compare formally adherence between randomized groups. No standard method of comparison is available. In this paper we analyze MEMS data from a randomized clinical trial in which adherence was an outcome of interest. All patients in this trial received the same blood pressure reducing agent, enalapril, prescribed at one tablet of 20 mg a day. Patients were randomized to perform daily self-home-measurement of blood pressure or not. This paper looks at the daily pattern of drug intake to answer, the question \Does the prescription of daily self-home-measurement of blood pressure improve patient adherence ? ".

In the next section we describe the trial and MEMS data in more detail; section three examines a range of simple summary statistics; section four models the full pattern of daily dosing on a binary scale of yes/no, and compares the treatment groups. We discuss the advantages and disadvantages of conditional and marginal modeling in this context. We have found that the intermediate strategy of modeling weekly adherence recovers the essence of the detailed picture we had discovered earlier, with the added advantage of easier direct summaries for inspection by the physician.

2.2   Description of the trial

Six hundred twenty-eight patients with high blood pressure ($\geq$95 mmHg) were assigned to an initial daily treatment with one tablet of enalapril at breakfast and randomized to self-measurement or no self-measurement

of blood pressure at home (309 and 319 patients, respectively). The hemodynamic effects of the drug are seen within one hour of a single oral dose and the action of once-daily dosing lasts for about 24 hours. The two randomized groups were comparable with respect to sex, age, systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse, and weight.

Patients randomized to the Home Measure Group were asked to measure their blood pressure every morning before taking the medication. They received a form to transcribe SBP, DBP, and pulse each day. The trial coordinator supplied the apparatus (a Bosoprestige automatic) at the start of the trial, with directions for use. Measures were to be taken in the sitting position after five minutes of rest. Basic notions about SBP, DBP, pulse, variations in those measures, etc., were taught to the patients randomized to this group.

Patients have five scheduled visits with a physician. The first three are of interest for this chapter.

- Visit 1, at day 1 : Extensive baseline data, including SBP, DPB, and pulse, were collected from both groups. Randomization was performed. For the Home Measure Group the physician took two blood pressure measurements : once using his own apparatus and once using the patients apparatus.

- Visit 2, after 2 weeks : The physician recorded SBP, DBP, pulse, and reported symptoms.

- Visit 3, after 6 weeks : The same parameters were recorded as at Visit 2. Further, if the treatment was not effective, the physician was instructed to prescribe hydrochloro thiazid in addition to enalapril. A medication switch was specified at Visit 3 when the average of the last five home measures was greater than 90 mmHg for the patients in the Home Measure group and when the office measure of DBP was greater than 90 mmHg in the group randomized to No Home Measure.

This paper focuses on the period between Visit 1 and Visit 3, the time one can observe the effect of home measures for patients on enalapril, unconfounded by the different allocation rule for hydrochloro thiazid.

2.3   Measuring adherence

Irrespective of randomization, a subpopulation of 127 patients from both randomized groups (66 from the Home Measure Group and 61 from the No Home Measure Group) was chosen for evaluation of adherence using an electronic pillbox, the MEMS device. In practice, some of the physicians most likely to be able to collaborate with a more demanding treatment protocol were chosen to give MEMS to all their recruited patients. The child proof cover of the MEMS bottle contains a microprocessor that records the date and time of each opening. All monitors were returned for data retrieval (there was one missing data point because of a technical problem). Before entering the study, each patient gave informed consent, following the guidelines of the first International Symposium on adherence Monitoring, Heidelberg, 1988.

Adherence data are obtained as listings of consecutive dates and times of individual container openings. As a basic visual description of the data, we have graphed each patients pattern of drug intake from Visit 1 to Visit 5.

We observed vastly different adherence patterns among patients. Broadly, we saw two main types of adherers (also discussed in the next section): the punctual adherers, as exemplified by the patients in figure 2.1 (A and B), and the less regular adherers, as seen in figure 2.1 (C). Patient A, an extremely punctual adherer, takes the drug, every day at nearly the same time. Patient B, also quite a regular adherer, has a periodic trend in behavior: the drug intake happens later on weekend days. Patient C, on the other hand, has a lot of gaps between intake days, has one day of double dosing, and shows large variability in the timing of drug intake.

2.4   Summary statistics for adherence

Before modeling the detail of adherence patterns, we consider simple statistics that summarize one aspect of the patient's adherence history. We discuss the merits of six different measures. Figure 2.4 displays a series of two-by-two plots for the six summary measures.

2.4.1   Percentage of Prescribed Dosing Days with Correct Intake

The percentage of days with accurate dose intake is defined as follows:

$$C_{i1} = \frac{\text{Number of days with `correct' intake for subject i}}{\text{Number of prescribed dosing days for subject i}}$$

Where the Number of days with `correct' intake is the number of days during which the prescribed number of doses was taken. For the study at hand, the numerator is thus the number of days with one and only one drug intake. The higher this percentage, the better adherence. In practice this adherence summary statistics is often defined as \correct dosing".

This statistic captures some measure of closeness to \correct adherence". Depending on how the latter is defined, it may reflect the degree of regularity in lifestyle. However this summary measure gives no information concerning the timing of dose intake, it does not distinguish between days of overdosing and days of underdosing and thus, this focus on correct adherence may not capture deviations most relevant to the drug action.

2.4.2   Percentage of Prescribed Dose Taken

This measure, though often close to the previous one, focuses on the dose actually received :

$$C_{i2} = \frac{\text{Number of doses taken by subject i}}{\text{Number of prescribed doses for subject i}}$$

For patients with overdosing, multiple doses in one day are added up, thus this percentage can be larger than 100%. For once daily prescription, the number of prescribed doses is equal to the number of treatment days. A good adherer will have a percentage near 100%. Higher or lower percentage indicates \poor adherence". Several authors have used this measure to study the causal effect of adherence or to obtain an efficacy measure (Efron and Feldman 1991, Fischer-Lapp and Goetghebeur 1999).

This measure reflects the average dose received over a given period and hence also the total dose over that period. It accounts for period of time without drug intake and double dosing. However, it fails to distinguish between a regular taker and someone who balances periods of underdosing with periods of overdosing and it captures no information about the precise timing of drug intake.

A

B

C

Figure 2.1: Individual patient adherence. The horizontal axis displays the days since study entry. The vertical axis gives the time of drug intake on a 24-hour clock. The digits from 0 to 6 are used as a plotting symbol to characterize the day of the week (0 = Sunday, 1 = Monday,..., 6 = Saturday). The vertical line represents the third visit.

Plot of (dbp3-dbp1) against compliance
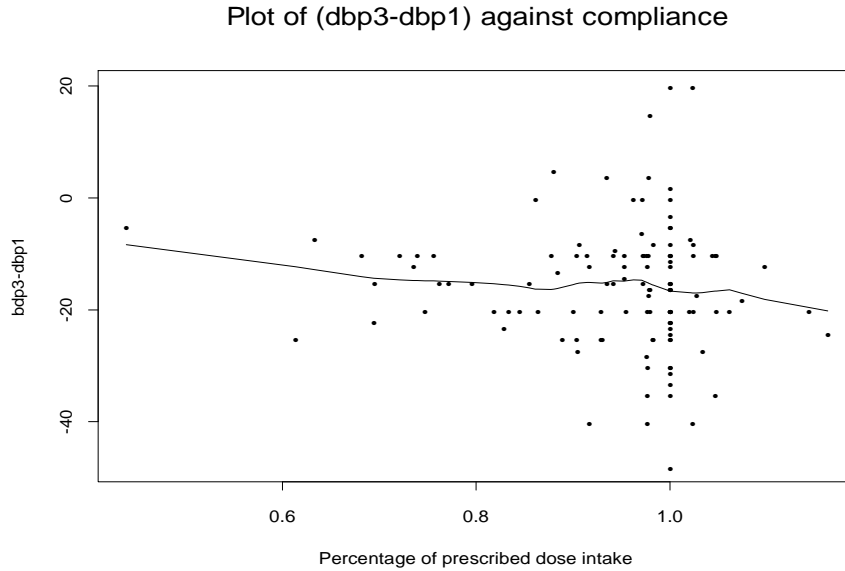
Figure 2.2: DBP reduction against adherence.

Figure 2.2 plots achieved DBP reduction by Visit 3 against the percentage of prescribed dose taken. A lowess curve (Cleveland 1979) drawn through the points spots at most a weak association over the range of observed adherence levels.

Figure 2.3 is a scatter plot of the percentage of dose taken (C_{i2}) against the percentage of prescribed dosing days with correct dosing (C_{i1}). Below the horizontal 100% line for the percentage of prescribed dose, we observe a strong positive correlation between the two summary measures (r = 0:92; n = 107). Above this line, we observe a negative correlation (r = −0:89; n = 20). The two points on the top of the figure are consistent with expectation and are not outliers.

2.4.3 Percentage of Drug Holidays

We define a \drug holiday" as a period of at least d (d $\geq$ 1) days without drug intake (that is, without MEMS opening). The statistic counts the number of such gaps among the prescribed days, ignoring the exact length of each individual gap :

$$C_{i3} = \frac{\textbf{Number of drug holidays for subject i}}{\textbf{Number of prescribed dosing days for subject i}}.$$
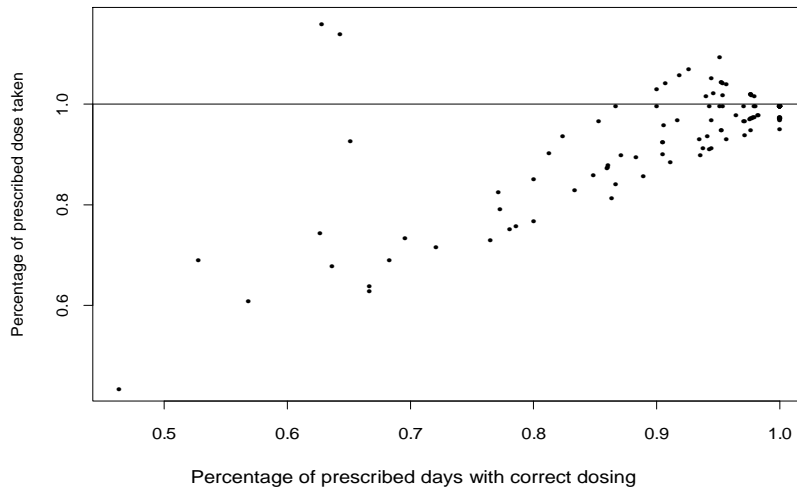
Figure 2.3: Percentage of prescribed dose taken against the percentage of prescribed days with correct dosing.

Where the number of prescribed dosing days for subject i is the number of days between the first intake and the day of the third visit for patient i. The lower this percentage, the better adherence.

This measure captures a feature important for a drug that could have a serious rebound effect (the sudden appearance of undesired effects after discontinuation of a drug). To decide on the length d in the numerator, one can rely on pharmacologic insights. For instance, for certain drugs, drug holidays of three or more consecutive days of no dosing may be hazardous to a patients life (Petri and Urquhart 1994). The main disadvantage is that the measure ignores double dosing and precise timing of drug intake.

2.4.4   Time Variability in Drug Intake

The three summary measures defined so far do not account for the precise timing of drug intake; they consider only the number of medication events each day. The following continuous adherence summary provides

a measure of the timing variability of intake :

$$C_{i4} = \frac{\sum_k \mid t_{ik} - median(t_i) \mid}{\text{Number of prescribed dosing days for subject i}}$$

The number of prescribed dosing days for subject i is defined above, $median(t_i)$, is the median times of drug intake on a 24-hour clock for patient i ($t_i = (t_{i1}, ..., t_{iK_i})$). k indexes the MEMS openings which happen at times t_{ik}, where $k = 1, ..., K_i$ (K_i = Total number of MEMS openings for patient i). For median dosing time near midnight we consider the day to end at 3 a.m. If during one day no dose was taken, we took $t_{ik} - median(t_i) = 12$ hours. For double doses, there are two t_{ik}'s contributing to the numerator, but the denominator is unaltered. In this way, both phenomena, underdosing and overdosing, though opposite in nature, add extra variability to the timing pattern. Thus, the lower this percentage, the better the adherence. Such measure will be interpreted in light of the duration of action of a drug. In the present study the drugs hemodynamic action starts at about one hour after intake and lasts for about 24 hours. Drug intake time schedule is thus of prime importance. This measure accounts for the timing of drug intake. Double dosing cannot compensate drug holidays as in C_{i2}. However absolute deviation is symmetric around zero and the measure is not proportional to the amount of drug taken. While intuitively appealing for assessing the precise timing of drug intake, this measure is difficult to generalize for multiple doses per day regimen.

2.4.5 Percentage of \Too Short" or \Too Long" Dosing Intervals

We introduce a separate measure for periods of \overdosing" (interval too short) and periods of \underdosing" (interval too long). Rather than averaging over the deviations in timing from the median, this measure looks at the number of deviations that exceed a crucial or meaningful threshold of dosing intervals that are either too short or too long. As one such measure, we define for each patient the percentage of dosing intervals shorter than δ^l hours and the percentage of dosing intervals longer than δ^u hours. Suppose that t_{ik} is the kth MEMS opening time for patient i. Then, we define a vector of dosing intervals for patient i: $d_{ik} = t_{i(k+1)} - t_{ik}, k = 1, ..., K_i - 1$ and the two measures can be formulated as follows:

$$C_{i5A} = \frac{\sum_k Ind(d_{ik} < \delta^l)}{K_i - 1}$$

$$C_{i5B} = \frac{\mathbf{P}_k \, \mathrm{Ind}(d_{ik} > \delta^u)}{K_i - 1}$$

Where Ind() is the event indicator.

The measures make a clear distinction between periods of overdosing and underdosing. Combining these measures with outcome data allows for identification of potential rebound effects. Figure 2.4 shows a strong correlation between C_{i5B} and the four first adherence summaries ($C_{i1}; ::::; C_{i4}$). Since C_{i5A} is less correlated with those measures, it may add substantial extra information to highlight differences between groups. The correlation between C_{i5A} and C_{i5B} is relatively low (r=0.37), indicating that patients having a high percentage of dosing intervals that are too long are not simply the same patients with a large percentage of dosing intervals that are too short. One of the main disadvantages is that a meaningful choice of cutoff points must be made. For this particular study we took the prescribed dosing interval of 24 hours plus minus 25%, resulting in $\delta^l = 18$ and $\delta^u = 30$. A discussion based on two summary measures can be difficult to interpret. To avoid the latter, some authors (Claxton et al. 2001) have combined the two summaries :

$$C_{i5} = \frac{\mathbf{P}_k \, \mathrm{Ind}(\delta^l \leq d_{ik} \leq \delta^u)}{K_i - 1}$$

The combined adherence summary is then labelled \dose timing".

For an alternative measure, rather than summing over just the number of intervals in excess of 30 hours, one may choose to add up the total time spent in excess of 30 hours before the next dose and divide this by the number of treatment days. These measures relate to one another like median and mean; here, the longest intervals count most. Because in this example, results from this approach were very similar to the observations on C_{i5A} and C_{15B}, we do not report them.

2.4.6 Median and Quantiles of Dosing Intervals

As a final summary measure capturing the variability in timing between doses we chose the Median and the 5th and 95th percentiles of the dosing intervals of each patient. Using the same notations as before and denoting ($d_i = (d_{i1}; ::::; d_{iK_i-1})$), we define:

$$C_{i6A} = \mathrm{median}(d_i)$$

and

$$C_{i6B} = Q_{05}(d_i)$$

$$C_{i6C} = Q_{95}(d_i)$$

The 5th and 95th percentiles reflect extremes of the dosing interval distribution. As before, we observe that patients with a large 95th percentile are different from the patients with a low 5th percentile. The median of the dosing intervals shows a low variation among subjects and thus it is not a good way to describe varying adherence levels. Some authors choose the mean of the dosing intervals, but because this statistic is in some sense equivalent to the percentage of drug taken, it adds nothing new.

2.5   Results

The first four summary measures ($C_{i1}; ...; C_{i4}$) show a distribution that is skewed to the left and bring to light a population of punctual adherers and some less regular ones. For example, figure 2.5 presents a histogram of the percentage of prescribed dose taken. Figure 2.4 shows bivariate associations between ($C_{i1}; ...; C_{i5}$). For clarity in plots, we took the square root of the time variability in drug intake, C_{i4}. Except for the percentage of dosing intervals that are too short (C_{i5A}), all those summary measures show a strong correlation in this study. In absolute value, the lowest reported correlation among them is 0.73 (n=127). Depending on the goal of the study and the nature of the drug, one can choose the more appropriate measure.

None of the six measures discussed shows a significant difference between randomized groups at the nominal 5% level. Although the measures provide a well-understood, simple description, patient adherence can vary in many different ways over time. Summarizing the history in just a few measures may lose power to detect relevant differences in adherence patterns. The next logical step is to draw inference using a simple test that exploits enough of the data in a way that captures the temporal evolution. The next sections describe relevant methodology.
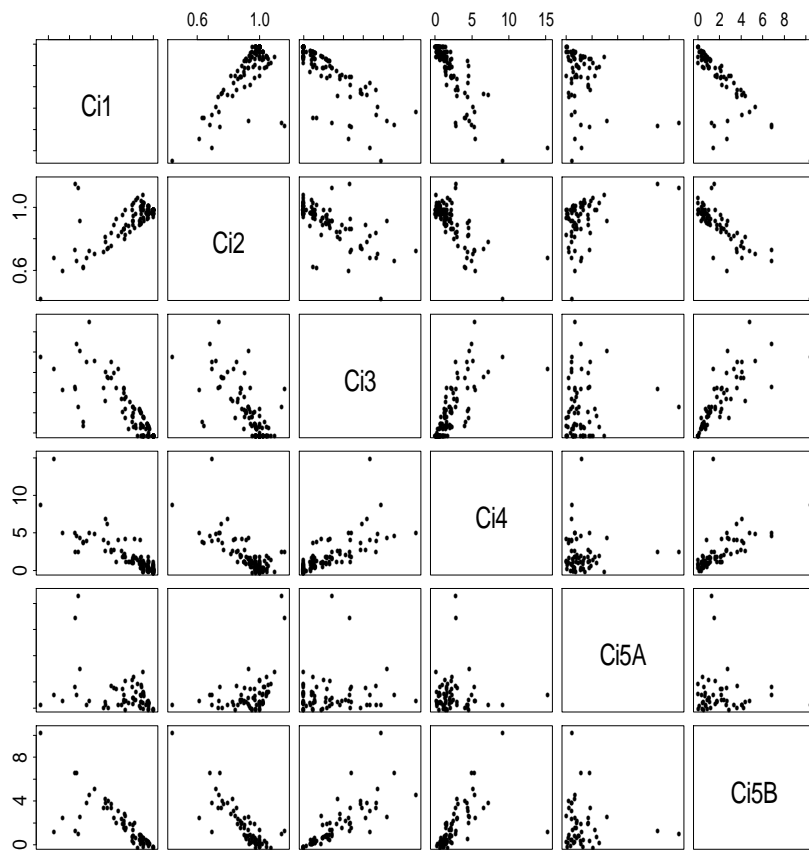
Figure 2.4: Two-by-two scatterplots of some summary measures of adherence for the patients in the No Home Measure group.
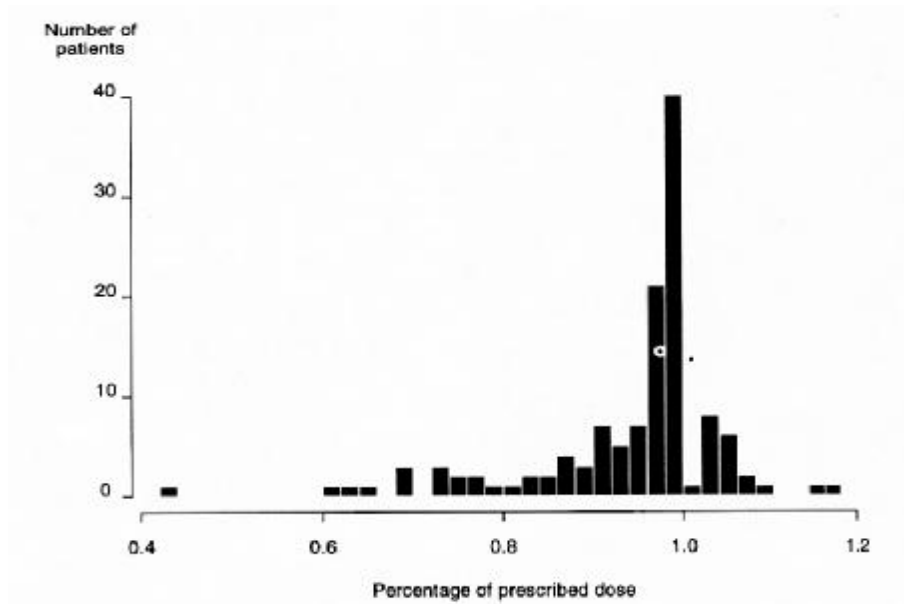
Figure 2.5: Histogram of the percentage of prescribed dose taken (C_{i2}).

2.6 Comparing daily binary adherence between randomized groups

2.6.1 Pattern of Binary Data

Our aim is to estimate the causal effect of randomized treatment on an informative measure of adherence. Because adherence to prescribed therapy measures the extend to which a person's behavior coincides with medical advice, it covers both a behavior and a measure. One influencing the other, delivers raw data that have a complex structure with unusual correlation patterns of dosing intervals and it is not obvious how to analyze them. Rather than study detailed patterns of adherence, we summarize the adherence pattern in a sequence of binary data indicating whether yes (1) or no (0), at least one dose has been taken on each consecutive day. This coding retains much of the temporal structure in the individual patterns, but not exact times of drug doses. We assume that different patients have independent sequences, but that the measures belonging to one patient may not be independent. Furthermore, we account for the possibility of a systematic trend in adherence over time and an effect of the day of the week (DOW) on adherence. Patients were

scheduled to return to the physician for the third visit after 6 weeks. For 77 (66%) of the 127 patients, the third visit took place within the first 6 weeks. In order to compare both groups on a reasonable number of patients over the full time period studied, we considered the individual time series of the 50 remaining patients up to 42 days (= 6 weeks) only. We have found that when only a few patients have been observed over a great number of time points, they can be extremely influential in a simple parametric model for the time dependence of adherence.

2.6.2   A Conditional Model

First, we implement a model conditional on past adherence for Z_{it} the yes/no drug indicator for patient i at day t, allowing for feedback from output to output, for a main treatment effect and first order interactions. We consider the following model:

$$
\begin{aligned}
\text{logit}(P\left(Z_{it} = 1 \mid Z_{it-1}; ....; Z_{i1}; H_{it}\right)) \;=\; & \alpha_t^0 T_i + \\
& \beta_1 Z_{it-1} + \beta_2 Z_{it-2} + ::: + \\
& \beta_{11}^0 DOW_i(t) + f(t; \nu) + \\
& \gamma^0 X_i + \\
& \delta^0 T_i * (X_i; Z_{it-1}; Z_{it-2}; :::);
\end{aligned}
$$

$$(2.1)$$

where H_{it} contains the information on baseline covariates and exogenous time dependent covariates. T_i is a dummy indicator for the randomized \treatment" assigned to individual i (i.e., enalapril with or without Home Measures). The ::: refer to possible further steps retained for the feedback from output to output (using time series terminology), DOW refers to the day of the week (6-variate dummy), $f(t; \nu)$ is a function parameterized by ν representing the time trend in adherence, X_i contains the baseline covariates (sex, age, DBP, pulse), and the last term allows for an interaction between the group and the other covariates in the model. We interpret the parameters as follows:

1. α_t can, in general, have a different parameter for each observed value of t. It then estimates non parametrically a time-dependent effect of treatment. Clearly, one pays for leaving the dependence on t unspecified by using a large number of degrees of freedom in

the data. Modeling a parametric dependence on t can easily be done and may sometimes be more desirable.

2. $\beta_{\mathbf{j}}$ relates to the post-randomization variable $\mathbf{Z_{i;t-j}}$ which may in its own right already capture some or most of the effect of treatment. This parameter expresses the dependence of a person's daily adherence on his own past adherence behavior.

3. $\beta_{\mathbf{11}}$ expresses a constant (across time, individuals and groups) effect of the day of the week, a covariate external to the observed process.

4. ν expresses the average evolution (across individuals, and groups) of adherence over time since entry into the study within the parametric form of $\mathbf{f}$.

5. γ expresses the usual effect of baseline covariates on adherence at day t. One should remember however, that this is after adjusting for previous adherence. In other words, a given $\mathbf{X}$ could set the baseline adherence standard but not enter into the picture after that, implying a parameter value that is zero.

6. Given the different meaning of adherence in both groups, we found it necessary to allow in principle for an interaction between the covariates influencing adherence and treatment.

For the initial time points where $\mathbf{Z_{i;t-j}}$ is not available, we propose a reduced right-hand side and allow initially for time-specific parameters γ and δ there.

2.6.3   Model Selection and Results

First, we observed a strong association between past adherence and the current adherence of the same patient. An analysis of variance table supports maintaining in the model indicators for adherence over the six previous days. Table 2.1 shows six days is a clear cutoff point for the deviance reduction. We next used a backward selection beginning from the most exhaustive model proposed above. At each step, we deleted the variable showing the smallest contribution to the model until all the variables remaining in the model were significant at the 0.05 level. Higher order effects and interactions between the remaining independent

Table 2.1: Deviance Table for Selecting Past adherence (Note: $Z_{i;t-j}$ is the yes/no drug indicator for patient i at time $t-j$)

Variable	Residual Deviance	DF	Residual Deviance Reduction	DF Reduction	p-value
NULL	2038	3394			
$Z_{i;t-1}$	1854	3393	184	1	< 0:0001
$Z_{i;t-2}$	1744	3392	110	1	< 0:0001
$Z_{i;t-3}$	1716	3391	28	1	< 0:0001
$Z_{i;t-4}$	1697	3390	18	1	< 0:0001
$Z_{i;t-5}$	1678	3389	19	1	< 0:0001
$Z_{i;t-6}$	1645	3388	34	1	< 0:0001
$Z_{i;t-7}$	1641	3387	4	1	0:0559
$Z_{i;t-8}$	1635	3386	6	1	0:0115
$Z_{i;t-9}$	1634	3385	1	1	0:2980

variables were subsequently analyzed and retained in the model if significant at the 0.05 level. This leads to modeling linearly the dependence of treatment effect on time t. Thus, in the notation of α_t, we can omit the subscript t. As shown in Figure 2.6 this is reasonable after day six, which is where the full conditional model starts. Table 2.2 presents the final model. The day of the week affects adherence. Specifically, the probability of taking a drug on a Sunday tends to be lower than the other days.

Two parameters (α and δ) in Equation 2.1 refer directly to the treatment effect. We found a significant interaction between time and treatment. Interactions between treatment and other covariates were not retained because they did not reach the 0.05 significance level. Thus, after conditioning on initial adherence behavior, we find an additional significant effect of treatment on adherence since $\alpha \ominus 0$. When patients start the treatment, the probability of taking the daily drug is similar for patients in both the Home Measure and the No Home Measure Groups with a similar adherence behavior over the past six days. Over time, however, this probability decreases in both groups. The decrease in the No Home Measure Group is significantly stronger (p < 0:05 for the interaction) than in the Home Measure Group.

Figure 2.6 presents a visual representation of the model estimation (af-
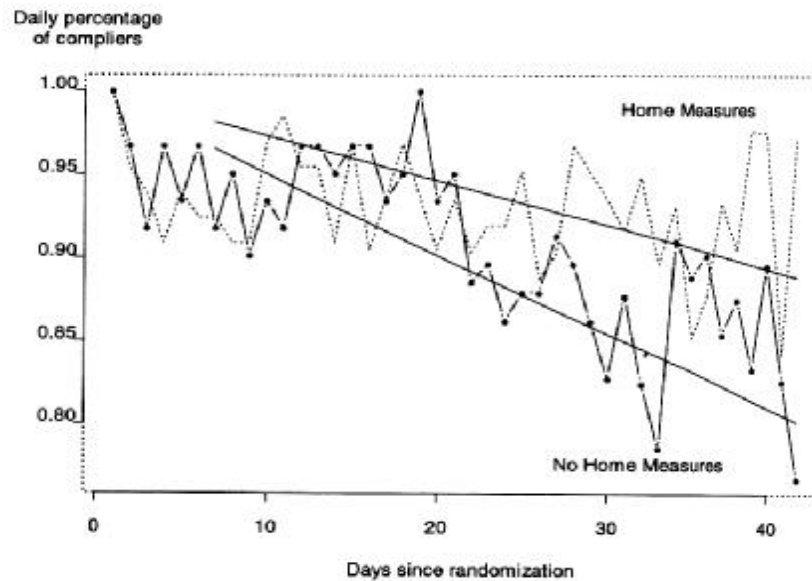
Figure 2.6: Daily percentage of adherers over time in both randomized groups. The dots represent daily percentages of adherers in each group. They are joint per group by broken lines. Full lines represent the logistic model fit of Table 2.2. In the observed range of days from randomization they look quite linear

Table 2.2: Parameter Estimates for Conditional Model (Note: $Z_{i;t-j}$ is the yes/no drug indicator for patient i at time $t-j$ and DOW(k) is an indicator for day k: 1=Monday, ..., 6=Saturday)

Variable	Estimated Coefficient	Standard Error	Coefficient/ Standard Error
INTERCEPT	-2.04	0.33	
$Z_{i;t-1}$	1.09	0.17	6.59
$Z_{i;t-2}$	1.36	0.16	8.31
$Z_{i;t-3}$	0.58	0.18	3.15
$Z_{i;t-4}$	0.64	0.18	3.51
$Z_{i;t-5}$	0.62	0.18	3.32
$Z_{i;t-6}$	1.01	0.18	5.77
DOW(1)	0.21	0.22	0.95
DOW(2)	0.37	0.23	1.62
DOW(3)	0.38	0.23	1.63
DOW(4)	0.48	0.24	2.01
DOW(5)	0.25	0.23	1.12
DOW(6)	0.12	0.22	0.53
Time*Home Measure Group	-0.011	0.007	-1.54
Time*NO Home Measure Group	-0.020	0.006	-3.11

ter adjusting for other covariates, including past adherence). Since we maintain indicators for past adherence over the six previous days, the reduced right-hand side model estimation is presented between day 6 and day 42.

In summary, the model estimates adherence to be less in the No Home Measure Group than in the Home Measure group.

2.6.4   Generalized Estimating Equation (GEE) Type Approach

Because model 2.1 conditions on past adherence, parameters may not capture the full effect of randomized treatment. For this reason, it can be more relevant to ask about the marginal effect of treatment and other external covariates on outcome at a given day : $E(Z_{it} \mid T_i; X_i; ...)$. Several methods have been proposed for this; see, for instance, (Bonney 1987, Prentice 1988, Zhao and Prentice 1990, Lipsitz et all. 1991, Carey et al. 1994, Diggle et al. 1994). We used the GEE method (Diggle et al. 1994) to fit the following model :

$$
\begin{aligned}
\text{logit}(P(Z_{it} = 1 \mid H_{it})) \;=\; & \alpha_t^0 T_i + \beta_{11}^0 DOW_i(t) + f(t; \nu) + \\
& \gamma^0 X_i + \delta^0 T_i * X_i.
\end{aligned}
\tag{2.2}
$$

The same covariates as those retained in the final model of the conditional regression yield marginal parameters which turn out very close to their conditional counterparts. A modeling procedure considering all the explanatory variables was impracticable in this setting. For long time series, the current implementation of the GEE approach may not converge (Diggle et al. 1994). Extensions of this are the topic of ongoing statistical research. Alternatively, we have examined the marginal effect of a summary vector of treatment adherence as, for instance, weekly total dose taken. That is, we considered:

$$
D_i = (D_{i1}; ...; D_{i6}) = \left(\sum_{t=1}^{7} Z_{it}; ...; \sum_{t=36}^{42} Z_{it} \right)
$$

a vector of the weekly total doses consumed. In comparison to the yes/no daily indicator, this last outcome measure has the advantage of taking double dosing events into account. On the other hand we now have a 6-dimensional rather than 42 dimensional summary of the observed adherence patterns. We use the Poisson distribution to model those

count data. Since the number of time points is small and each outcome can theoretically be any discrete number, a marginal model becomes more appealing.

In terms of the mean model, since observations are collected at regular weeks similarly for each patient, we can specify the marginal expectation as follows :

$$\log(E(D_{it} \mid H_{it})) = x_{it}^0 \beta:$$

Specifically in our situation, to estimate the overall treatment effect we use the following model:

$$\log(E(D_{it} \mid H_{it})) = \alpha_t^0 T_i + f(t) + \gamma^0 X_i + \delta^0 T_i * X_i: \qquad (2.3)$$

After summation, the day of the week effect no longer enters, but we still expect the marginal means of D_i to depend on T_i, X_i and the week $f(t)$.

Estimations of the coefficients in Equation 2.3 can be obtained by solving first a GEE equation of the following form :

$$\sum_{i=1}^{N} \left(\frac{@\mu_{D_i}}{@\beta} \right)^{!0} V_i^{-1} \left(D_i - \mu_{D_i} \right) = 0 \qquad (2.4)$$

with $\mu_{D_i} = E(D_i)$. V_i, the working covariance matrix for D_i, is a function of the marginal means and perhaps of additional q parameters α, i.e.:

$$(V_i)_{kl} = Cov(D_{ik}; D_{il}) = {}' (\mu_{D_{ik}}; \mu_{D_{il}}; \alpha) \qquad (2.5)$$

Where $k; l = 1; \dots; 6$, $k \ominus l$ and ' is a know function. Since inference on fixed effects is quite robust regarding the definition of the working covariance matrix, we chose the exchangeable covariance structure, which makes computations easier. It is defined as follows :

$$(V_i)_{kl} = \begin{cases} \alpha; & k; l = 1; \dots; 6 \text{ and } k \ominus l \\ 1; & \text{otherwise} \end{cases} \qquad (2.6)$$

The correlation parameters α may then be estimated by simultaneously solving Equation 2.4 and

Table 2.3: Parameter Estimates for the GEE Model

Variable	Estimated Coefficient	Standard Error	Coefficient/ Standard Error
INTERCEPT	1.93	0.016	
Treatment(T_i) (No Home)	0.048	0.026	1.87
Week*Home Measure Group	-0.0081	0.004	-1.98
Week*No Home Measure Group	-0.0271	0.008	-3.16

$$\sum_{i=1}^{N} \left(\frac{\partial \eta_i}{\partial \alpha} \right)^0 \mathbf{H}_i^{-1} (\mathbf{W}_i - \eta_i) = 0; \tag{2.7}$$

where $\mathbf{W}_i = (D_{i1}; D_{i2}; D_{i1}; D_{i3}; ::::; D_{i5}; D_{i6})$,
$\mathbf{H}_i = \text{diag}(Var(D_{i1}; D_{i2}); Var(D_{i1}; D_{i3}); ::::; Var(D_{i5}; D_{i6}))$,
and $\eta_i = E(\mathbf{W}_i)$.

We fitted this model using the S-PLUS function from Statlib by V. Carey (1994). Table 2.3 presents estimated coefficients of the model retained after backward selection from the most exhaustive model proposed above. Here again we modeled the dependence of treatment effect linearly over time t. At the start of the study, we observe that the weekly number of drugs taken tends to be a little higher for patients in the No Home Measure Group than for patients in the Home Measure Group. Over the weeks, however, this number decreases in both groups; the decrease is significantly stronger ($p < 0.05$) in the No Home Measure Group. The dose intake per week is now estimated to be reduced by exp (-0.008) = 99 % in the Home Measure Group and exp (-0.027) = 97 % in the No Home Measure Group. We find these results easier to interpret than the reduction in daily odds measured in the previous analysis. For example, for a patient in the No Home Measure Group, dose intake after 6 weeks is estimated to be reduced by $(0.97)^6 = 83$ % (94 % in the Home Measure Group). Thus, if we suppose that the patient starts the study taking 7 drugs a week, after 6 weeks the weekly estimated number of drug intake is 5.8 (6.6 in the Home Measure Group). The model confirms the visual picture we obtained in Figure 2.7. In line with daily intake, modeling weekly adherence confirms the decrease in adherence over time and has retained sufficient detail to detect the
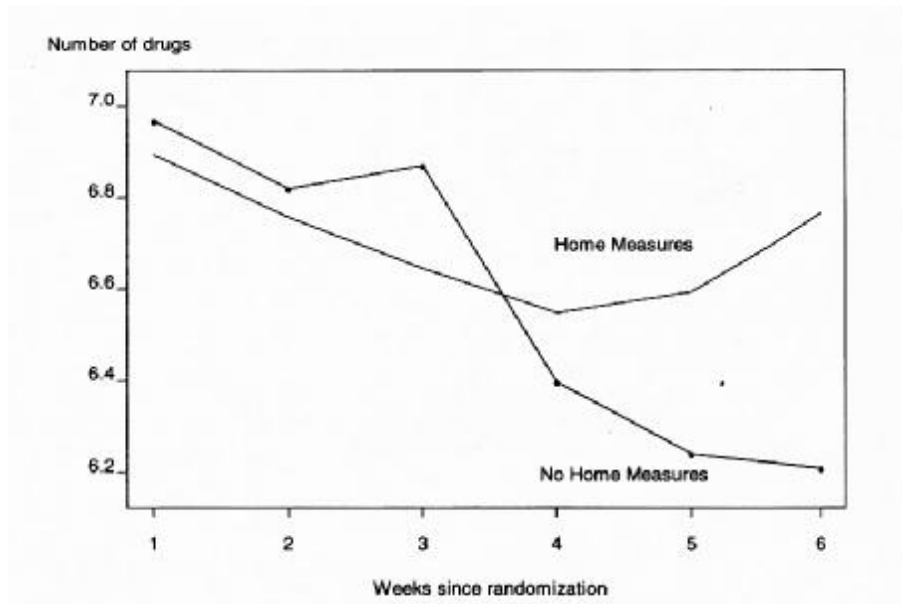
Figure 2.7: Average number of study pills taken each week in both randomized groups.

different rates of decline between both randomized groups. When incorporating the extra information on double dosing, we estimate a slight difference in the intercept: The No Home Measure Group appears to take more pills initially.

2.7   Comment

The conditional model presented above can be used to investigate the effect on adherence of other covariates which may be time varying or not. We illustrate it by studying the association between patient's age and adherence. In the No Home Measure group, the age of the patients ranges between 39 and 82 years with a mean age of 63 years. In Table 2.4, we present the results from the fitted conditional model 2.1 allowing on top of the Markov dependencies a marginal effect of age, weekend and their interaction.

The probability to take a medication on weekends is estimated to be lower than on a week day. The adherence to the treatment tends to decrease with age but we found a significant interaction between age and

Table 2.4: Parameter Estimates for Conditional Model ($Z_{i;t-j}$ is the yes/no drug indicator for patient i at time $t-j$, Weekend is a binary indicator: 0=week day, 1=weekend day and age is in years)

Variable	Estimated Coefficient	Standard Error	Coefficient/ Standard Error
INTERCEPT	-1.15	0.65	
$Z_{i;t-1}$	1.30	0.25	5.27
$Z_{i;t-2}$	1.27	0.24	5.20
$Z_{i;t-3}$	0.51	0.28	1.83
$Z_{i;t-4}$	1.32	0.25	5.20
$Z_{i;t-5}$	-0.28	0.32	-0.88
$Z_{i;t-6}$	1.29	0.26	5.05
Weekend	-2.30	0.85	-2.72
Age	-0.02	0.01	-1.81
Weekend * Age	0.03	0.01	2.29

weekends. In conclusion, the negative effect of weekends on adherence disappears for older people. This confirms the idea that younger people tend to go out on weekends and then forget their prescribed medication, while for older people the difference between week days and weekend days is probably less pronounced. Here again, we see the importance of drawing inference by exploiting enough of the data in a way that captures the temporal evolution.

2.8   Discussion

Although medication event monitoring systems have brought new accuracy and detail in measurement of adherence, most statistical methods that deal with MEMS data use only very simple summary measures like the percentage of total dose taken or mean and standard deviation of dosing intervals (Kass et al. 1986, Cramer et al. 1989, Psaty et al. 1990, Kruse and Weber 1990, Petri and Urquhart 1994). In this paper, we introduced a larger set of simple summary measures that capture well-defined relevant features of the adherence pattern. Furthermore, we have shown how a much more detailed summary can lead to a well-understood and informative analysis using recent methodology implemented in S-PLUS and derived packages. In the different one- and multi-dimensional

summary measures for adherence information, we have strived to balance the simplicity of a meaningful synthesis against enough information. Known or hypothesized characteristics of the drug under study will generally drive the choice, but some principles hold quite generally for classes of chemical drugs. The distinction between the behavioral focus, timing variability, total dose, and local and global underdosing or overdosing are elements to consider. For the trial we analyzed, none of the univariate summary measures revealed a difference between randomized groups; however, analyzing the binary dosing pattern over time showed a clear and meaningful difference, with implications for future practice. We believe that such binary summary combined with a marginal approach that allows for within-patient correlations is particularly suited for this type of analysis. One could argue that a multinomial response per day, with categories 0,1 and more than 1, would also reflect patterns of overdosing. Although in theory such a model could be fit using a similar approach, we found too few instances of double dosing in our study to make this meaningful. In between the day-to-day model and the univariate summaries lies the vector of weekly summaries. We found that a weekly summary measure captured the important distinguishing features and could more easily deal with more than one dose per day. Such data can be fit with fewer parameters, as the models do not need to include a day-of-the-week effect. The resulting parameters are easier interpreted.

Chapter 3

On predicting future patient compliance from its past

Clinical practice, clinical trials design and structural mean estimation for the causal effect of observed exposure in randomized trials, may benefit greatly from good predictors of future compliance based on a patients past measurements. While conditional and marginal models were developed in the previous chapter to describe such patient adherence general patterns over time, prediction is the harder problem. This paper proposes to extend the methods to allow for prediction based on different data structures. To predict compliance outcomes, we use an auto-regressive model for repeated binary data. We handle frailty through a normal distribution for the random slopes and random intercept. The first challenge is to find relevant measures of prediction error which can subsequently be minimized. We discuss several such measures. Model validation as well as two different types of clinically relevant model predictions are developed next. The approach is applied to analyze a clinical trial studying the compliance of patients suffering from major depressive disorders treated with selective serotonin reuptake inhibitors (SSRI).

3.1   Introduction

Actual exposure to treatment, a consequence of adherence, is typically not the primary outcome of a drug trial. The double status which such post-randomization pattern embodies is part of the reason. At the same time consequence of treatment assignment and driver of treatment effect, it suffers from possible selectivity: the fact that `natural' or treatment-free outcomes are spontaneously associated with observed dose-timing patterns. Several authors argued that adherence is reasonably seen as an attribute of the patient irrespective of treatment, i.e. a pre-randomization variable (Efron and Feldman 1991). Others have contested this (Lee et al. 1991), neither standpoint has convincingly been backed up with lots of data in a wide range of disease areas and for a wide range of scientific questions. If compliance were an attribute, good predictions of future patterns should be possible from a run-in period on placebo. Such predictions can in turn serve as an instrument to get at estimated compliance-response curves (White and Pocock 1996, Goetghebeur and Lapp 1997, White and Goetghebeur 1998) and to evaluate any shift in post-randomization adherence between two treatments. Clinical practice (Cramer and Spilker 1991), clinical trial design (Cramer and Spilker 1991, Lim 1992) and structural mean estimation for the causal effect of observed exposure in randomized trials (Goetghebeur and Lapp 1997), may benefit greatly from good patient specific predictors of future compliance. Models for prediction can be mere extrapolation or based on long term observations in different data sets with similar patients.

In this paper we work with detailed information provided by electronic monitors of drug-intake events. Data from a typical clinical trial involving MEMS devices (Medication Event Monitoring System, Aardex Ltd.) consist of one list per patient containing dates and times of consecutive openings of a pill container. These raw data have a complex structure with unusual correlation patterns of dosing intervals. There is no standard approach to their analyses. In most simple regimen studies, rather than studying exact time patterns of compliance, it is satisfactory to summarize compliance patterns as a sequence of binary data indicating whether yes (1) or no (0), at least one dose has been taken during each consecutive scheduled dosing period. Conditional and marginal models were developed to describe such patient adherence general patterns over time (Vrijens and Goetghebeur 1997, Girard et al. 1998, Smith and Diggle 1998). However none of those models were designed specifically for prediction. The current paper investigates predictability of patient

adherence with an anti depressive agent using an existing data base. In section 2, we describe a clinical example that we will use to illustrate the proposed methodology. In section 3, the statistical analysis approach is presented and implemented. Model validation as well as two different types of clinically relevant model predictions are developed. The results and methodology are discussed in section 4.

3.2 Clinical example

We analyze data from a controlled clinical trial designed to study adherence profiles of ambulatory patients suffering from major depressive disorders treated with selective serotonin reuptake inhibitors (SSRI). The double blind study included 76 patients aged 22 to 44 years. After a wash-out placebo lead-in period, patients were randomized to fluoxetine or paroxetine once a day. They were then followed for 6 weeks during acute phase. Two visits were scheduled during the acute phase. The first one, 2 weeks after starting treatment and the second one 4 weeks later. Finally patients entered a 4 month period called maintenance phase. During this phase, a monthly visit was scheduled. Individual adherence was assessed during the whole trial period using MEMS devices. All patients started intake following randomization. Because they are of different physical dimensions, we find it necessary to develop separate prediction models for the persistence and compliance. This paper focuses on modeling compliance as a daily pattern of indicators of whether yes or no the patient took any dose that day. We found too few instances of double dosing to warrant an ordinal variable per day. Several patients (29%) dropped out during the study leaving 54 patients to study their compliance, uncounfounded by the different persistence length. We study the level of intake during 93 days (3 first months of the maintenance phase) among those who are still taking the drug. In figure 3.1 we present the percentage of compliers over time during the maintenance phase and in figure 3.2 the individual cumulative number of doses taken over time. Since the drug was prescribed once a day the ideal patient dosing history is the diagonal line. Horizontal segments indicate the presence of drug holidays (i.e. extended time period without drug).
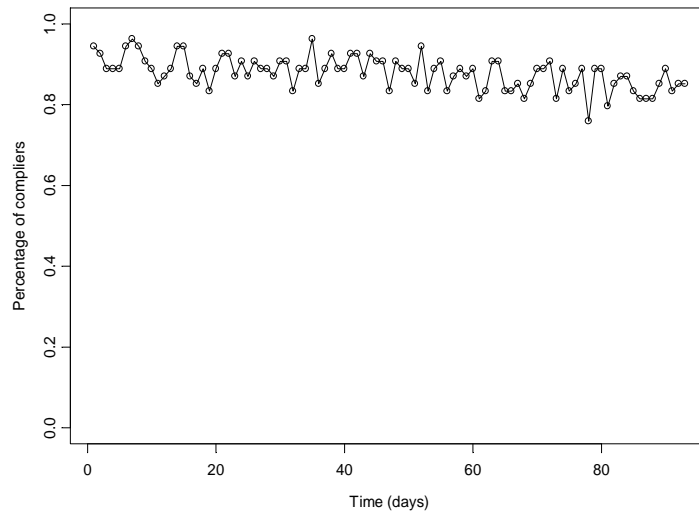
Figure 3.1: Observed proportion of compliers over time (days since randomization).
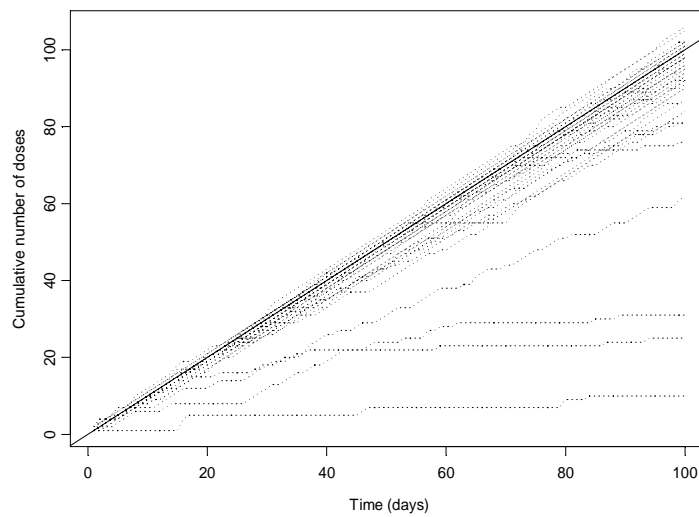
Figure 3.2: Accumulated number of doses per patient over time. Each dotted line represents a single patient and the diagonal full line represents a theoretical perfect complier

3.3 Auto-regressive model for binary compliance data

Methods acknowledging the dependence among successive compliance binary responses of an individual in a repeated measurements situations have already be developed (Vrijens and Goetghebeur 1997). However none of those models were designed specifically for prediction. In the current paper we move on to investigate predictability of dependent daily patterns of binary compliance data. Two types of such dependencies can be expected: those arising from heterogeneity among individuals, often called frailty, and those from serial correlation over time. Serial correlation can be accommodated by conditioning on previous observations, e.g. fitting Markov chains by standard logistic models.

Let y_{ik} represent the yes/no drug indicator for patient i on day k. A model accounting for frailty and serial correlation can be written as:

$$\eta(p_{ik}) = \eta_{ik} = \alpha_{0i} + y_{i(k-1)}\alpha_{1i} + y_{i(k-2)}\alpha_{2i} + ...; \qquad (3.1)$$

and

$$\alpha_i \sim MVN(\alpha; \gamma):$$

Where $p_{ik} = P(Y_{ik} = 1 \mid Y_{i(k-1)}; Y_{i(k-2)}; ....; Y_{i1}; \alpha_i)$, $\eta(:)$ is the logit link function and $\alpha; \gamma$ are a set of fixed parameters to be estimated.

Conditional on the random effect α_i, the probability of a particular compliance pattern, y_i, is assumed to be the product of the conditional probabilities of each of the n_i binary responses, namely

$$g(y_i \mid \alpha_i) = \prod_{k=1}^{n_i} \frac{e^{\eta_{ik}y_{ik}}}{(1 + e^{\eta_{ik}})}:$$

Thus the marginal probability of this pattern is given by

$$h(y_i \mid \alpha; \gamma) = \int_{-1}^{+1} ... \int_{-1}^{+1} g(y_i \mid \alpha_i)\phi(\alpha_i; \alpha; \gamma)d\alpha_i:$$

The log-likelihood for the pattern from the N patients can then be written as

$$
\begin{aligned}
\log \prod_{i=1}^{N} L(y_i \mid \alpha; \gamma) &= \sum_{i=1}^{N} l(y_i \mid \alpha; \gamma) \\
&= \sum_{i=1}^{N} \log[h(y_i \mid \alpha; \gamma)] \\
&= \sum_{i=1}^{N} \log[\int_{-1}^{+1} ... \int_{-1}^{+1} g(y_i \mid \alpha_i)\phi(\alpha_i; \alpha; \gamma)d\alpha_i]
\end{aligned}
$$

There is no closed form solution for the values α and γ at which the above integral reaches a maximum. If a single random intercept is used

(α_{0i}), the random effect can be eliminated by applying integrated likelihood techniques. We use Gaussian quadrature to numerically evaluate the above likelihood integral for a vector of random effects. Letting $\phi(:)$ take a normal parametric form, the integrated likelihood becomes:

$$\sum_{i=1}^{N} \log[\sum_{q=1}^{Q} g(y_i \mid \alpha_0; \gamma \times B_q) A(\gamma \times B_q)]$$

where B_q and $A(B_q)$, $q = 1; ::; Q$ are the Q fixed quadrature locations and weights, respectively, and the scale parameter γ represents now the standard deviation of the mixing distribution. When random regression coefficients for the Markov chain are investigated the parameters become more difficult to estimate and an MCMC approach is preferable. Win-BUGS (Gilks et al. 1996) version 1.3 was used to analyze the random parameters models. Convergence of the Gibbs sampler was evaluated by visual inspection of the trace plots of the population parameters in Win-BUGS. When the chain was thought to be stable, a formal diagnostic using the Raftery and Lewis (Gilks et al. 1996) methods was carried out using the Coda macro under S-PLUS. At least five thousand iterations were collected after convergence had been achieved for estimates of the parameters. We note that modeling serial correlation by conditioning can account for a considerable proportion of the heterogeneity. Therefore it is important to estimate which part is attributable to each of the two components.

3.3.1 Model validation

In order to estimate the level of the random effect(s), for a particular patient, we used the expected a posteriori value of the parameter. Using WinBUGS, the expected a posteriori value of the random effect are immediately available while after numerical integration equation 3.2 was used:

$$\hat{\alpha}_{0i} = E(\alpha_{0i} \mid Y_i) = \frac{1}{h(y_i)} \int_{-1}^{+1} \alpha_{0i} g(y_i \mid \alpha_i) \phi(\alpha_i; \hat{\alpha}\hat{\gamma}) d\alpha_i: \qquad (3.2)$$

These can be calculated using Gaussian quadrature as previously mentioned.

Having estimated the parameters based on all data, we set out to check whether the model actually fits the data (goodness-of-fit). For this purpose we check the fit of the different conditional models $f(y_{ik} \mid$

$y_{i(k-1)}; ::::; y_{i1})$ in turn by calculating $E(y_{ik} j y_{i(k-1)}; ::::; y_{i1})$ under the model and contrasting it with the observed values. Thus at each point in time we assume that previous compliance is known as observed and we are checking whether the data at each point in time fit the conditional distributions. This results in a unique deviance labelled \1-step Deviance" in the sequel. To assess the goodness of the joint fit, the analytical solution is hard to find but we proceed in a practically convenient way as follows. Based on the estimated random effect, one can estimate a first day outcome as a random drawn from the expected Bernouilli variable. Then plugging in this value as input into the auto regressive model , we estimate the second day outcome based on the distribution conditional on this information, and so forth for longer sequences of compliance. Since stochastic imputation is used at each step of the conditional model, the estimated sequence $\hat{y}_{ik}$ and thus the different conditional models $f(y_{ik} j \hat{y}_{i(k-1)}; ::::; \hat{y}_{i1})$ are not unique. We repeated the simulation 100 times to evaluate the modeled distribution. For each simulation, the discrepancies between the observed and estimated compliance patterns is assessed through the deviance as described previously and the subject average deviance is reported (Joint Deviance) in column 3 of the following tables. We note that this technique is closely related to `Baysian predictive distribution'. It represents our current estimation of the compliance patterns (y_{ik}) taking into account both the uncertainty about the value of the parameters and the residual variability about (y_{ik}) when the parameters are known.

To acknowledge dependence in future patterns, we inspect the Euclidean distance between observed and predicted numbers of drug holidays (period of time without drug intake) of specified length. When studying patient's compliance, it can be clinically more relevant to focus on the dependencies among the missed doses than to be confused by the whole correlation structure. Because the percentage of drug taken (taking compliance) is an often used compliance summary, we found it relevant to estimate also the mean of the Euclidean distance between observed and predicted sequences. The results for the model validation are presented in table 3.1 for different length of the Markov order.

3.3.2 Model prediction

In the previous section, parameter estimation was based on all available data collected over 93 days (3 months). In this section we wish to evaluate the ability of the model to predict new data. This approach

Table 3.1: Model validation

Markov Order	1-step Dev.	Joint Dev.	Taking Compliance	Drug holiday of length				
				1	2	3	4	5
Fixed effect models								
1	3010	4276	0.20	18.7	16.3	14.5	13.1	12.0
2	2788	4538	0.21	19.3	16.8	14.9	13.4	12.6
Random intercept models								
0	2383	2383	0.028	2.6	3.4	3.9	4.2	4.4
1	2337	2434	0.031	2.9	3.1	3.3	3.5	3.7
2	2323	2460	0.035	3.3	3.5	3.5	3.5	3.6
Random parameters models								
1	2224	2520	0.036	3.3	2.6	2.6	2.6	2.8
2	2227	2507	0.038	3.5	3.0	2.9	2.9	3.0
3	2229	2507	0.039	3.6	3.1	3.2	3.0	3.1

is conceptually different from the goodness-of-fit question addressed before. While the latter is a measure of the fit to the same data on which the model parameters were based, here we intent to assess the fit to future observations drawn from the same population under similar circumstances.

To study properties of predictors of new observations, two types of predictions were investigated as illustrated in figure 3.3. First we based our model estimation on 31 days of observed compliance and predicted the next 3 months. This is for example the situation when in a clinical trial setting, a group of patients are observed during a run-in period and the future of the trial must be predicted. The results of this prediction approach are presented in table 3.2 for different lengths of the Markov order of the fitted model. In a next step we estimate the influence of the length of the run-in period used for parameter estimation. In table 3.3 we fitted the first order Markov prediction model with random parameters estimated on 31, 21 and 11 days of compliance data. The model was then used to predict the next 3 months of compliance.

Another way to look at prediction is to suppose that a group of patients are already observed on the therapy and we want to predict the compliance of a new similar patient based on patient specific initial compliance data. In this situation the prediction model is derived from all data on

Type 1 prediction
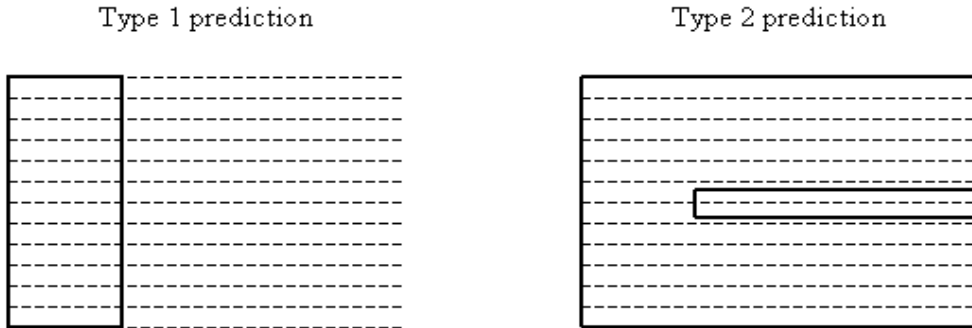
Type 2 prediction

Figure 3.3: Schematic illustration of the two prediction types investigated. Patient compliance patterns over time are schematically represented by the horizontal dotted lines. The bold boxes indicate the data that are used to estimate the parameters of the prediction model for the two types of prediction described.

Table 3.2: Model prediction

Markov Order	1-step Dev.	Joint Dev.	Taking Compliance	Drug holiday of length				
				1	2	3	4	5
Random intercept models								
0	2656	2656	0.077	7.1	8.1	8.5	8.1	7.7
1	2549	2698	0.077	7.1	7.4	7.6	7.3	7.1
2	2503	2698	0.078	7.3	7.2	7.3	7.1	6.9
Random parameters models								
1	2553	2687	0.077	7.2	6.8	6.9	6.8	6.5
2	2514	2703	0.080	7.4	7.0	7.1	6.8	6.6
3	2553	2723	0.079	7.3	6.7	6.8	6.7	6.5

Table 3.3: model prediction type 1 - First order Markov

Nb days used	1-step Dev.	Joint Dev.	Taking Compliance	Drug holiday of length				
				1	2	3	4	5
Random intercept model								
31	2549	2698	0.077	7.1	7.4	7.6	7.3	7.1
21	2636	2844	0.085	7.9	6.9	7.1	7.1	7.0
11	2714	3024	0.121	11.2	10.6	10.7	10.4	10.2
Random parameters model								
31	2553	2687	0.077	7.2	6.8	6.9	6.8	6.5
21	2696	2875	0.092	8.6	7.0	6.8	6.6	6.4
11	2727	3173	0.138	12.8	11.3	10.9	10.4	10.0

all but one of the patients. This rule is then applied to predict the omitted patient. Evaluation of this cross-validation approach is typically computationally intensive. For each patient left out, model parameters have to be re-estimated. For cross-validation, much faster results can be obtained by considering a quadratic approximation of $l_{(i)}(\theta)$ with $(\theta = (\alpha; \gamma))$, the log likelihood obtained after deleting the i-th case (Cook and Weisberg 1995),

$$l_{(i)}(\theta) \simeq l_{(i)}(\hat{\theta}) + (\theta - \hat{\theta})^T l_{(i)}^0(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^T l_{(i)}^{00}(\hat{\theta})(\theta - \hat{\theta})$$

where the elements of $l_{(i)}^0(\hat{\theta})$ and $l_{(i)}^{00}(\hat{\theta})$ are defined as follows

$$[l_{(i)}^0(\hat{\theta})]_m = \frac{\partial l_{(i)}(\theta)}{\partial \theta_m}\Big|_{\theta = \hat{\theta}} \text{ and } [l_{(i)}^{00}(\hat{\theta})]_{mn} = \frac{\partial^2 l_{(i)}(\theta)}{\partial \theta_m \partial \theta_n}\Big|_{\theta = \hat{\theta}}:$$

If $-l_{(i)}^{00}(\hat{\theta})$ is strictly positive definite, the quadratic approximation is uniquely maximized at

$$\theta_{(i)}^1 = \hat{\theta} - (l_{(i)}^{00}(\hat{\theta}))^{-1} l_{(i)}^0(\hat{\theta})$$

the one-step approximation to $\hat{\theta}_{(i)}$. Denoting by $\sum_{(i)}$ the summation over all but the ith patient, we have in general

$$\frac{\partial l_{(i)}(\theta)}{\partial \theta_m} = \sum_{(i)} \frac{1}{h(y_i \mid \theta)} \frac{\partial h(y_i \mid \theta)}{\partial \theta_m}$$

Table 3.4: model prediction type 2 - First order Markov

Nb days used	1-step Dev.	Joint Dev.	Taking Compliance	Drug holiday of length				
				1	2	3	4	5
Random intercept model								
31	2530	2675	0.08	7.4	7.8	8.0	7.6	7.3
21	2614	2775	0.08	7.8	7.3	7.4	7.3	7.2
11	2678	2853	0.10	9.2	8.6	8.8	8.7	8.6
Random parameters model								
31	2596	2713	0.08	7.2	7.0	7.1	6.8	6.5
21	2710	2880	0.09	8.3	6.9	6.7	6.4	6.2
11	2776	2957	0.11	9.9	8.5	8.3	8.1	8.0

$$\frac{\partial^2 l_{(i)}(\theta)}{\partial \theta_m \partial \theta_n} = \sum_{(i)} [\frac{1}{h(y_i \mid \theta)} \frac{\partial^2 h(y_i \mid \theta)}{\partial \theta_m \partial \theta_n} - \frac{1}{h^2(y_i \mid \theta)} \frac{\partial h(y_i \mid \theta)}{\partial \theta_m} \frac{\partial h(y_i \mid \theta)}{\partial \theta_n}]$$

Denoting $\eta_{iqk} = X_{ik}\beta + \gamma B_q$, with B_q from the Gauss-Hermite approximation, the derivatives of the individual marginal likelihood can be approximated for the logit link as

$$\begin{cases} \frac{\partial h(y_i \mid \theta)}{\partial \theta_m} = \sum_{q=1}^{Q} \prod_{k=1}^{n_i} \frac{y_{ik} + (y_{ik}-1)e^{\eta_{iqk}}}{1+e^{\eta_{iqk}}} \frac{\partial \eta_{iqk}}{\partial \theta_m} g(y_i \mid \theta; B_q) A(B_q) \\ \frac{\partial^2 h(y_i \mid \theta)}{\partial \theta_m \partial \theta_n} = \sum_{q=1}^{Q} \prod_{k=1}^{n_i} [\frac{y_{ik} + (y_{ik}-1)e^{\eta_{iqk}}}{1+e^{\eta_{iqk}}} \frac{\partial \eta_{iqk}}{\partial \theta_m} \frac{\partial g(y_i \mid \theta; B_q)}{\partial \theta_n} - \\ \frac{e^{\eta_{iqk}}}{(1+e^{\eta_{iqk}})^2} \frac{\partial \eta_{iqk}}{\partial \theta_n} \frac{\partial \eta_{iqk}}{\partial \theta_m} g(y_i \mid \theta; B_q)] A(B_q) \end{cases}$$

If $\hat{\theta}_{(i)}$ is not too different from $\hat{\theta}$ and $l_{(i)}(\theta)$ is locally quadratic, the one-step estimator should be close to the fully iterated value (Cook and Weisberg 1995). Applied to our data, its computation time was just 1 percent of the original time and corresponding results are presented in table 3.4.

As with observed compliance data, we find it useful to plot the cumulative predicted number of doses taken over time against time (days). In figures 3.4, each plot represents for a given patient the observed compliance (full line) and simulated predictions (dotted lines). The model used for prediction is the type 1, first order Markov random parameters model estimated on 31 days. We conclude that the model predicts compliance well.
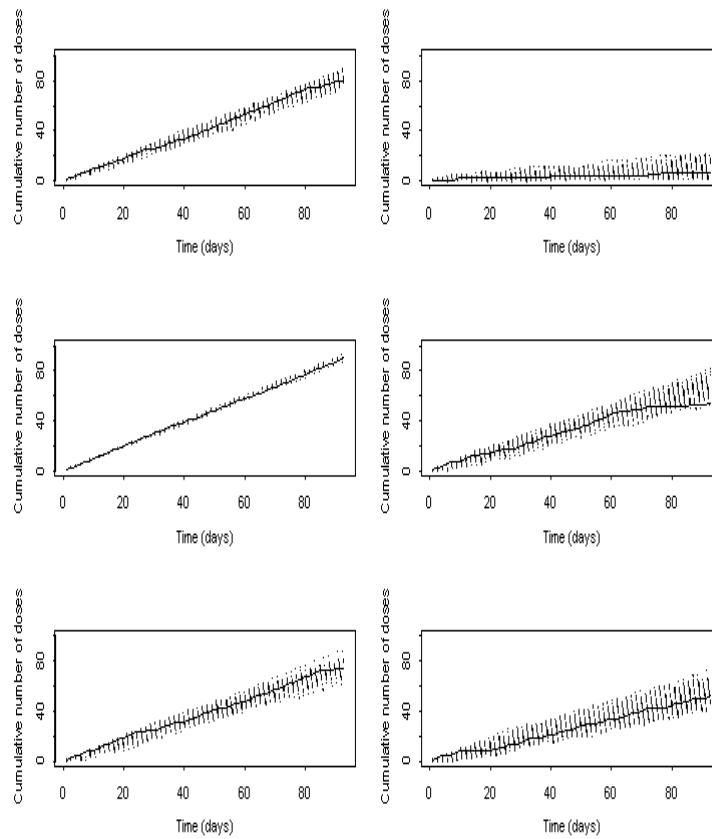
Figure 3.4: Prediction assessment for 6 individual patients

3.4   Discussion

Having distinguished 3 qualitatively different dimensions of adherence: acceptance, execution and discontinuation, we have set out to study and predict the patterns of drug exposure over time during the execution phase (compliance). While being a summary of the available data, depending on the pharmacodynamics of the drug, the daily binary representation of the drug intake is usually seen as a relevant measure. In this study we found it relevant to base our daily compliance assessment on the daily drug indicator. For drugs with shorter half-lives, the timing of the intake can be used in the rule to dichotomize the daily compliance.

Model validation and two different types of prediction were investigated. Four different loss functions help assess the model fit and the quality of predictions. After model validation, we conclude that three months of binary patient compliance data can be summarized in just a few parameters per patient without loosing much information. A random slope and intercept improve the model, especially when drug holidays are to be predicted. With this model, the average goodness-of-fit error in `taking compliance' is less than 4 percent and as indicated by the number of drug holidays predicted, the relevant dependencies among the observations are reasonably well captured. The proposed model-based prediction of compliance, uses parameters derived from measurements of a patient's prior compliance. In the setting of this trial, both prediction techniques deliver similar results. This finding will be different when considering a longer follow-up during which a time trend in compliance is seen. When 31 days were used for prediction, the average prediction error in taking compliance is 8 percent and the relevant dependencies among the observations remain reasonably predicted as illustrated in figure 3.4 for 6 different patients. We note that, to check whether the predictions actually fit the data, a variety of different statistics (loss functions) can be investigated. We proposed and discussed 4 different ones that are relevant in the present settings. The choice will in general depend upon the drug under study and the goal of the medical researcher.

We feel that the issues raised by considering adherence to therapy as an attribute of the patient are somewhat more complex than usually argued. There is probably no clear dichotomous answer. Having split the adherence in a persistence and a compliance components will help to understand this complex process. Results presented here indicate that the compliance of some patients is intrinsically better than of others. This difference corresponds to the importance of the random effects

(frailty) in the models studied. These results support the idea that once a patient accepts to take a drug, he enters into a daily routine that is probably closely linked to his lifestyle. Some patients are intrinsically better compliers than others. The end of this continuous process occurs when the patient takes a new decision. The discontinuation can eventually be influenced by the drug itself (adverse reactions, sickness improvement, relapses, ...) and perhaps is not an attribute of the patient. The prediction of the time to discontinuation based on baseline data has to be investigated drugs/diseases specifically and is the topic of further research.

Chapter 4

The impact of patient adherence in pharmacokinetic studies

A successful drug is not just a molecule and a formulation but a unit dose and a dosing regimen. To establish these, it is necessary to know at what rates the drug is absorbed and eliminated from the body and what the therapeutic window is for plasma concentration. Pharmacokinetics is the mathematics of the time course of absorption, distribution and elimination of drugs in the body. To establish this, exhaustive studies, with directly observed drug intake and intensive blood sampling, are traditionally undertaken early in drug development. However, studies in early phases of drug development are usually small, very well controlled and not representative of the general population. Therefore it is necessary to investigate the pharmacokinetic properties of the drug in all phases of clinical development. Large scale, intensive pharmacokinetic sampling in ambulatory patients is resource demanding, very costly and difficult to implement. It is in this context that population pharmacokinetics becomes practically appealing but still raises methodological problems. In population pharmacokinetic studies, one observes just a few concentration measures spread out in time, on a sizable sample of the target population. One thus moves beyond more traditional designs where a nonlinear model is fit to (a series of) data from a single patient, measured frequently under controlled conditions over a small period of time. Advantages of the population method include the potential to estimate between-patient variation in pharmacokinetic parameters as well

as the ability to measure the more realistic impact of drugs when taken over a longer period of time. A disadvantage stems from the fact that, naturally, one needs not only good measures of concentration but also reliable information on the drug dosing history at those measuring times. Often the latter has been lacking. The population PK literature warns against the naive assumption that patients are in steady state, as an overestimation of adherence may lead to seriously biased estimates (Girard et al. 1996, Vrijens and Goetghebeur 1999). Studies of simulated adherence behavior have helped quantify the problem with naive adherence estimators and pointed towards a solution. In this chapter we look at actually observed adherence patterns recorded via electronic monitoring in a clinical trial (Bovet et al. 1997). We simulate a documented pharmacokinetic model from the hypertensive literature on top of these and come to some interesting findings. In this clinical trial the problem of nonadherence is much more dramatic than simulated adherence patterns suggested so far. The systematic errors made by adherence naive estimators can be corrected when using timing explicit hierarchical nonlinear models and accurate information on a number of previous dose timings. When it is possible to observe irregular drug intake times in a well controlled study (with monitoring of the drug intakes), a substantial amount of precision is retrieved from the same number of data points. In general, the estimators of pharmacokinetic parameters benefit greatly from information that enters through greater variation in the drug exposure process. Here we find support for the claim that nonadherence as a rich natural experiment of dosing variation can be a blessing rather than a curse from the information/learning point of view.

4.1   Introduction

Pharmacokinetic studies aim to describe the kinetics of absorption, distribution and elimination of a drug in the body after drug administration. It is one area where even the simplest statistical model is nonlinear whilst repeated (correlated) measures over time are the rule, not the exception: reasons enough to aim for efficiency in observation and data analysis.

When plasma concentrations are observed following drug intake in just one or a few chosen subjects, the role of statistics is reduced to dealing with small errors of observation around a deterministic concentration-time curve for a number of separate individuals (Gibaldi and Perrier 1982). No attempt is made to formally aggregate these few curves and estimate the population variation in concentration profiles after a first drug intake. In a next phase, several consecutively scheduled drug intakes are studied where a small reservoir of drug builds up slowly in the body (compartments) until a regular oscillating pattern of plasma concentrations indicates that `steady state' is reached. The body has arrived at a limiting state where an optimal drug effect is hopefully being produced. A good kinetic model for the absorption and elimination of the drug can help understand drug action and guide a range of medical decisions (Vozeh et al. 1996). Since the work of Sheiner and Beal in the 70's (Sheiner and Ludden 1992), the field of population pharmacokinetics (and pharmacodynamics), based on hierarchical nonlinear statistical analysis, has found increasing application in drug development. A useful discussion of the statistical issues can be found in Davidian and Giltinan (1995). Pharmacokinetic study is feasible on a larger scale by embedding it in phase III trials, and measure plasma concentration each time a hospital visit is scheduled during follow-up.

In this scenario, fewer concentration measures are taken on a larger number of patients. Furthermore, one has typically lost direct observation of the last or previous dose intake, and certainly of the longer history of intake. Statistical analyses have tended to ignore the longer intake history and only consider information about the last dose intake whilst assuming that patients have been in steady state for a while before the blood is taken. This assumption works well in populations of near perfect adherers with prescribed regimens. However, electronic monitoring systems (MEMS, Aardex Ltd.) have repeatedly revealed that intakes of patients are not as regular as one might hope (Cramer et al. 1990, Urquhart 1994). Consistent with an irregular exposure process is the fact

that established pharmacokinetic/pharmacodynamic (PK/PD) models, when applied to population observations, show more convergence problems than expected: as if the model does no longer fit the data. Also, the between- and within-patient variation is large, suggesting some source of major variation that has not been accounted for. Several authors have acknowledged this problem and pointed out the dangerous impact nonadherence may have on standard population PK/PD analysis and its naive interpretation.

In this chapter, we review the literature on the topic highlighting problems and the potential of newer approaches. Section 2 examines the literature to find population PK analyses ignoring nonadherence, and some initial approaches towards incorporating it. Section 3 introduces one PK model in more detail, that will serve as an illustration throughout the chapter. We found no documented use of actually observed drug intake histories along with PK measures. Therefore section 4 simulates plasma concentration profiles, based on documented PK models and actually observed adherence patterns in the context of a clinical trial. Several statistical approaches towards data analyses are implemented on these data sets in section 5. New approaches allowing for time dependent dosing are proposed and compared to traditional methodology. Finally, the results are presented in section 6 and discussed in section 7.

4.2 Patient adherence to the prescription is a leading source of variability in PK response

Variability in dosing due to patient nonadherence is often seen as a major source of variation in drug response in both clinical trials and medical practice (Cramer and Spilker 1991, Efron and Feldman 1991, Harter and Peck 1991, Goetghebeur and Pockok 1993, Urquhart 1998). The pharmacokinetic literature too presents erratic adherence with prescribed therapy as a major concern (Antal et al. 1989, Cramer 1990, Rubio et al. 1992, Jerling et al. 1994, Urquhart 1994, Girard et al. 1996, Vanhove et al. 1996, Vozeh et al. 1996, Wang 1996, Kastrissios and Blaschke 1997, Urquhart 1998, Wang 1998) and an often overlooked aspect in drug development. Pharmacokinetic studies intend to model the `output', i.e. plasma concentrations at various times, as a function of the `input': amount and timing of each preceding drug dose. The analysis of retrospectively collected data on dosing behavior (input) suffers from the reliability of patient recall being suspect, and has been known to

yield biased estimates of the pharmacokinetic parameters and markedly upward biased estimates of both inter- and intra- individual variability (Antal et al. 1989, Girard et al. 1996). Antal and co-workers (1989) assessed the impact of different data collection schemes on the quality of input data for model fitting. They conclude that good estimates of population pharmacokinetic parameters can be obtained with data arising from clinical trials provided accurate measures of drug administration histories and time(s) of sampling are recorded. Since then, electronic monitoring systems enable much better collection of this information. Such devices record the exact date and time of each opening of a drug container. The assumption that each such opening corresponds to one dose ingestion is often reasonable. To use these data, however, presents new methodological and practical challenges linked to the large imbalance between extensive input data available from the MEMS device and sparse output data characteristic of population pharmacokinetic studies. Girard and co-workers (1996) investigated six different strategies for summarizing dosing information. Some yield nearly the same parameter estimates as using complete records but reduce drastically the computing time. It becomes thus feasible to integrate nonadherence information using standard methods and available software.

In practice, retrospective data collection via questionnaires is still broadly used. This method, however, is prone to producing inaccurate measures of adherence (Urquhart 1994) and can seriously confound the analysis and interpretation of the adherence-exposure-response relationship (Kastrissios and Blaschke 1997). Perhaps even more unfortunate it has lead in practice to deleting large numbers (up to 50%) of patients from analyses due to missing information on the dosage schedule (Grasela et al. 1993, Jerling et al. 1994).

A procedure that is currently in use to overcome this problem is to select patients with good adherence during a run-in period and to further stimulate adherence by pill counts, digital drug dispenser, patient diaries, financial support, adherence counselling, ... Others (Vander Stichele 1991, Urquhart 1998) protest that using observational data, under a wide range of adherence patterns, is an underestimated opportunity to study the offset of drug action as dosing lapses occur on the patients own initiative. Especially when it appears unethical to impose random drug holidays, the introduction of MEMS devices within experimental studies would enable one to switch from artificial dose-response designs, typically done by varying the dose and holding constant the interval between doses, towards more realistic adherence-response designs where

patients hold the dose constant and vary the interval (Urquhart 1998).

Our results will show that estimating pharmacokinetic parameters based on actually observed adherence patterns can yield much more precise parameter estimates than estimation within a population of perfect adherers.

4.3   A population pharmacokinetic model

In this section we describe the population PK model that will be laid on top of dose-timing patterns in the next section. We need a structural model that relates concentrations over time to an individual's pharmacokinetic parameters and a statistical model that describes the joint distribution of pharmacokinetic parameters between individuals as well as the variation of observed measures around an individuals curve. We start from a clinical trial describing the PK/PD modeling of anti hypertensive response to amlodipine (Donnelly et al. 1993).
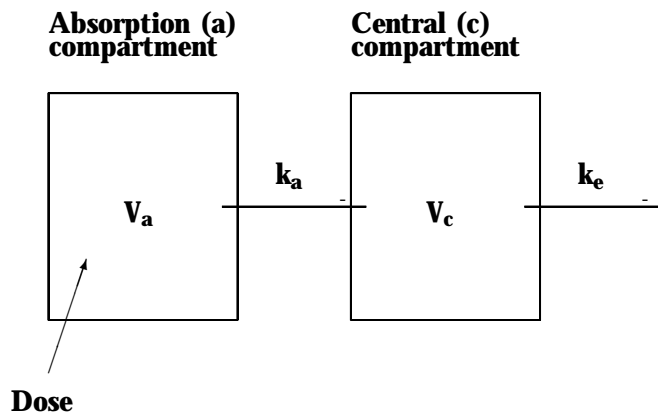

Figure 4.1: One compartment open model with as parameters, the volume of the absorption compartment (V_a), volume of the central compartment (V), absorption rate (k_a) and elimination rate (k_e).

They describe the pharmacokinetics of the drug assuming that the drug enters the body by an apparent first order absorption process, is elim-

inated by a first order process and distributes in the body according to a one-compartment model. The latter depicts the body as a single kinetically homogeneous unit. A first order elimination process assumes that the rate of elimination of drug from the body at any time is proportional to the amount of drug in the body at that time. This simple model is illustrated in figure 4.1 and is broadly used for pharmacokinetic analysis of drugs that distribute relatively rapidly throughout the body. The model parameters are thus, ratio of volume to bioavailability (V=F), elimination rate constant (k_e) and absorption rate constant (k_a). In this paper, we assume that 100% of the oral dose reaches the systemic circulation (F = 1) and denote by V the volume of the central compartment. For patient i on oral therapy prescribed at one 5 mg dose per day, let $C_i(t)$ and $C_{a;i}(t)$ be the typical value of concentration at time t in the central and absorption compartment, respectively. Let further $t_{0;i}$ indicate start of study (no dose yet) and $t_{d;i} = t_{1;i}; t_{2;i}; \dots$ the dosing time of dose number d = 1; 2; $\dots$ with $t_{d;i} \leq t < t_{d+1;i}$. Then in recursive form, the model relating plasma concentration at time t to drug history and concentrations at the previous intake times is given by (Davidian and Giltinan 1995):

$$C_i(t) = C_{a;i}(t_{d;i}) \frac{(k_a)_i}{(k_a)_i - (k_e)_i} \left\{ e^{-(k_e)_i(t-t_{d;i})} - e^{-(k_a)_i(t-t_{d;i})} \right\}$$
$$+ C_i(t_{d;i}) e^{-(k_e)_i(t-t_d)}$$

$$C_{a;i}(t_{d;i}) = C_{a;i}(t_{d-1;i}) e^{-ka_i(t_{d;i}-t_{d-1;i})} + \frac{Dose}{(V)_i}$$

(4.1)

with starting conditions:

$$C_i(t_{0;i}) = 0$$

$$C_{a;i}(t_{0;i}) = 0;$$

Note that the volume of the absorption compartment, V_a, has canceled out of the recursive concentration equation and is not specified.

For the simulations presented in the next section, we assume log-normal statistical models for the between-patient variability in k_e and V as follows:

$$(k_e)_i = k_e \times e^{\eta_{1;i}}; (V)_i = V \times e^{\eta_{2;i}}; \qquad (4.2)$$

where V and k_e are the population parameters and the between-patient random effects, $\eta_{1;i}$ and $\eta_{2;i}$ have a multivariate normal distribution,

with means 0 and variance-covariance matrix:

$$\begin{pmatrix} 0.09 & -0.02 \\ -0.02 & 0.09 \end{pmatrix}.$$

Because few samples are typically collected during the absorption phase, it is common practice to estimate k_a as a fixed parameter across the population. We fixed the population parameters at the values proposed by Donnelly et al. (1993):

$$V = 1550 (\text{liters}); k_e = 0.02 (\text{hour}^{-1}); k_a = 0.29 (\text{hour}^{-1}).$$

Finally, the model allows for log-normal within-patient variation as follows:

$$c_i(t) = C_i(t) e^{\epsilon_i(t)} \tag{4.3}$$

with independent $\epsilon_i(t) \sim N(0, 0.01)$.

4.4 Observed nonadherence and plausible effects

For a set of 60 patients (Bovet et al. 1997) electronically monitored (MEMS) adherence patterns are observed over a median follow-up time of 271 days (minimum = 58 days, maximum = 460 days). Figure 4.2 summarizes the group's adherence by showing for each patient the accumulated number of doses against time. With a prescription of one dose per day, the perfect adherer is represented by a straight line with equation: \accumulated number of doses = time". Horizontal segments indicate the presence of drug holidays (i.e. extended time periods without drug intake), and segments making an angle with the horizontal axis larger than 45° indicate multiple dosing. This plot depicts over time the percentage of prescribed dose taken, a popular measure for analysis. Since each patient is represented by a broken line and those lines overlap sometimes, it is difficult to determine the course of a given patient. Ideally one would use a different color for each patient. Nevertheless, from the plot we learn that underdosing is much more common than overdosing.

To show the impact of perfect adherence with assigned regimen (Scenario 1), normal timing errors (Scenario 2) and true observed adherence (Scenario 3) on population PK estimation we simulate plasma concentration curves on top of each of the adherence scenarios (dosing histories D_i). Specifically for each patient we consider:
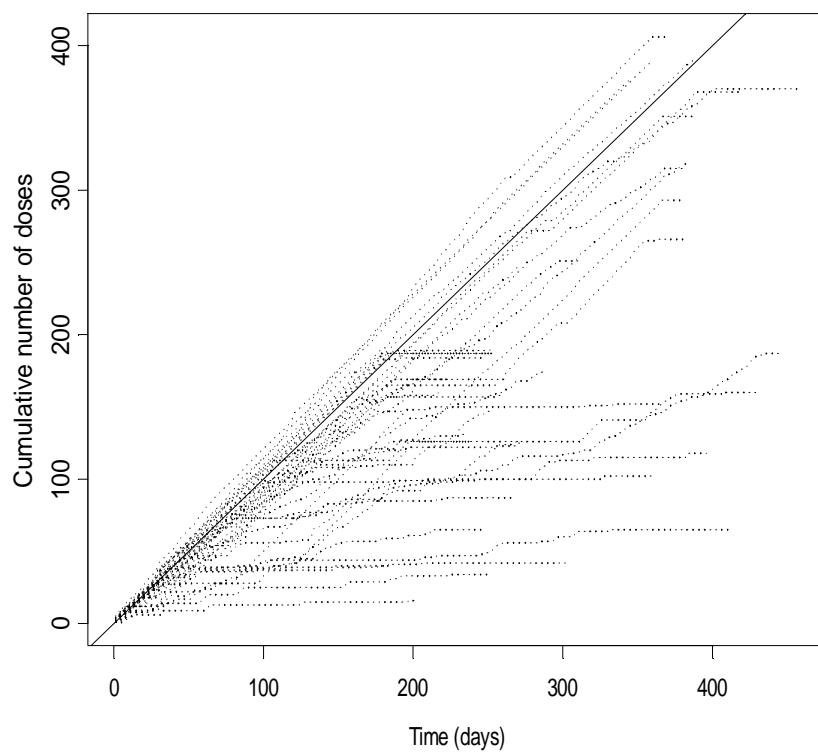
Figure 4.2: Adherence summary plot: accumulated number of doses over time

- **Scenario 1 (prescribed simulated intake):** the drug is taken exactly every 24 hours as prescribed for the same length as the observed period.

- **Scenario 2 (simulated intake):** the intervals between drug intakes are generated from normally distributed values with mean 24 and standard deviation 12. The CV is then 0.5 as proposed by Nony and co workers (1998) in their simulations. However when a negative value was generated for the dosing interval, we set it equal to zero, representing a double intake at that time.

- **Scenario 3 (observed intake):** actually observed dosing intervals (Bovet et al. 1997) are used as recorded by the MEMS devices.

Then, we generate the concentration curves from the one compartment PK model with parameter set $(k_a; k_{e;i}; V_{c;i})$. Specifically, we consider for each of the 60 individuals:

1. One draw $(k_{e;i}; V_{c;i})$ from the joint distribution (4.2) between individuals.

2. The scenario-specific dosing history D_i, combining it with the parameter set and random observation errors as in equations (4.3) to generate a concentration curve over time.

3. A set of selected observations at discrete times, 3 to 11 random time points (hospital visits) which are drawn under certain pragmatic constraints. The times must come after one month of prescribed therapy, cannot fall within the same week and are ignored if associated with a plasma concentration lower than 1 ng/ml. For each scenario, one plasma concentration measure is considered per visit.

In summary, the data vector per individual i consists of n_i time points $T_i = (t_1; ....; t_{n_i})$ at which concentrations in the central compartment, $C_i = (c_1; ....; c_{n_i})$, are recorded along with the dosing information, D_i. For the three scenarios we generated 100 data sets of 60 individuals, each time regenerating steps 1 and 3 in an independent random fashion. Figure 4.3 shows, for three typical patients, how the concentrations simulated on top of observed MEMS adherence patterns (Scenario 3) deviate dramatically from what would be seen if patients took the drug as prescribed (Scenario 1). They are also much more irregular than the
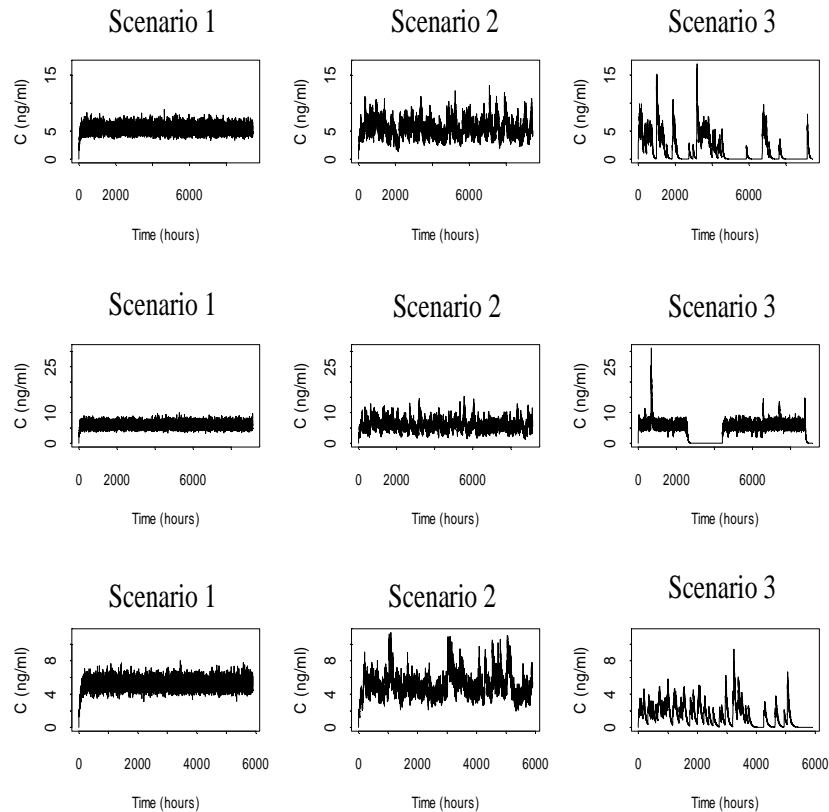
Figure 4.3: Concentration profiles simulation for three typical patients given the three proposed scenarios

concentration curves induced by the simulated timing errors (Scenario 2).

Having generated concentrations at `true' observation times, we now consider three different strategies for using the dosing information (input):

- **Strategy T: We observe the exact intake time for just the previous dose before blood sampling. This happens for instance when a dose is given in hospital a few hours before blood sampling and no recording of previous dosing histories is used.**

- **Strategy E: This strategy is similar to the previous one but we**

do not assume that the last dose was given in hospital. Then the patient is asked about the timing of last intake and the reported time since last dose is likely subject to errors. We consider that recent dose intake is easier to remember than earlier ones, hence we simulate larger errors for longer time elapsed since last dose, as follows:

$$\text{time since last dose} = (\text{exact time since last dose})e^{\psi}$$

where $\psi \sim N(0; 0.09)$.

- **Strategy +: Assuming electronic monitoring was used, we consider as model input the seven previous dosing event records.**

The corresponding `data sets' are labelled T, E, and + in figures and tables.

Three adherence scenarios and 3 strategies of time recall account for 9 different combinations.

4.5   Estimation method

Given a well specified hierarchical nonlinear model, one can use the linearization approach described in Davidian and Giltinan (1995) to estimate its parameters. Let for instance y_i be the outcome variable C_i and let ϵ_i be the $(n_i \times 1)$ vector of random intra-individual errors for individual i. The hierarchical nonlinear models described in the previous sections may be written, for $i = 1; ::::; N$, as

$$\left(\begin{array}{ll} \mathbf{y_i} = \mathbf{f_i}(\theta_i; \mathbf{T_i}) + \epsilon_i & \epsilon_i \, \mathbf{j} \, \eta_i \sim \mathbf{N}(\mathbf{0}; \mathbf{R_i}) \\ \theta_i = \mathbf{M} + \eta_i & \eta_i \sim \mathbf{N}(\mathbf{0}; \mathbf{\Gamma}) \end{array} \right. \tag{4.4}$$

where Γ is a $(k \times k)$ covariance matrix and $\mathbf{R_i}$ depends on i only through its dimension. In the PK model we assumed the multiplicative error model $\mathbf{y_i^*} = \mathbf{f_i^*}(\theta_i; \mathbf{T_i})e^{\epsilon_i}$, so we based its estimation on the log transformed model $\ln(\mathbf{y_i^*}) = \ln(\mathbf{f_i^*}(\theta_i; \mathbf{T_i})) + \epsilon_i$. Since the individual parameters were assumed to be log normally distributed, let θ_i represent the logarithm of the individual parameters, and M and Γ be their mean and variance-covariance respectively. Model (4.4) can then be rewritten

$$\mathbf{y_i} = \mathbf{f_i}(\mathbf{M} + \eta_i; \mathbf{T_i}) + \mathbf{R_i^{1=2}}(\mathbf{M} + \eta_i)\xi_i \tag{4.5}$$

where $R_i^{1=2}(M + \eta_i)$ is the Cholesky decomposition of R_i and explicitly indicates its dependence on M and η_i. A Taylor series expansion of (4.5) in η_i about η_i^0 close to η_i, retaining the two first terms in the expansion for the mean response vector and the leading term of the expansion of the residuals, yields

$$
\begin{aligned}
y_i &\approx f_i(M + \eta_i^0; T_i) + B_i(M + \eta_i^0; T_i)(\eta_i - \eta_i^0) + R_i^{1=2}(M + \eta_i^0)\xi_i \\
&= f_i(M + \eta_i^0; T_i) - B_i(M + \eta_i^0; T_i)\eta_i^0 + B_i(M + \eta_i^0; T_i)\eta_i + \epsilon_i^0
\end{aligned}
$$
(4.6)

where $B_i(M + \eta_i^0; T_i) = \frac{\delta f_i(\theta_i; T_i)}{\delta \theta_i}\mathbf{j}_{\eta_i = \eta_i^0}$ is the ($n_i \times k$) matrix of derivatives of $f_i(\theta_i)$ with respect to θ_i evaluated at $\theta_i = M + \eta_i^0$. As in the linear hierarchical case, the random effects η_i and within-individual errors ϵ_i^0 now enter the model in the same linear, additive fashion. Therefore, estimation follows from an iterative generalized least squares type algorithm in two steps (Lindstrom and Bates 1990):

- In a first 'pseudo data' step, given the current estimate of the variance components (intra- R_i and inter- individual Γ), joint estimation of the random and fixed effects (η_i and M) can be realized by jointly minimizing the sum of the negative log of the conditional density of the random effects for given M and twice the negative log likelihood for M for given η_i. We note that this is accomplished by specifying an augmented nonlinear least squares problem analogous to the linear case. The estimates of the fixed effects at this step will be used in the next step to specify the non systematic part of the model.

- Estimation of the variance components and fixed effects is then accomplished by a generalization of the maximum likelihood technique for linear models to the nonlinear case using equation (4.6).

Iterating this process to convergence will yield the final estimates.

Because it is current practice, we start by estimating parameters under the assumption that patients are in steady state at the time of their hospital visit and the last dose intake is the only one that is recorded. Steady state is theoretically acceptable since the first visit happens at least one month after starting the prescription. At steady state, the recursive equation set (4.1) can be simplified and the structural model

can be rewritten as:

$$C_i^{SS}(t) = \frac{Dose}{(V)_i} \frac{(k_a)_i}{(k_a)_i - (k_e)_i} \left\{ \frac{1}{1 - e^{-(k_e)_i\tau}} e^{-(k_e)_i t} - \frac{1}{1 - e^{-(k_a)_i\tau}} e^{-(k_a)_i t} \right\}$$

$$(4.7)$$

where $C_i^{SS}(t)$ is the typical value of concentration for patient i at time t after previous dose. Observations $c_i(t)$ show an error around outcome following equation (4.3) but we note that taking the logarithm of both sides of this expression results in a model that is linear in $\epsilon_i(t)$; this was used in further analyses. Following the original protocol, the dosing interval τ and the dose are assumed fixed at 24 hours and 5000 μg respectively, also k_a is treated as fixed across the population. We will call this approach `Naive Steady State Estimation' (NSSE) and use it to analyze the data issued from the three scenarios with (strategy E) and without (strategy T) error on the `observed' time interval since last dose.

With the observation of complete dosing histories (strategy +), the estimation approach uses data with seven previous dose intake records for each of the 3 adherence scenarios and assumes steady state beforehand. The recursive model given by equation (4.1) was then considered. We use the same assumption to model the within-patient variability as in the NSSE approach. We call this approach `Timing Explicit Estimation' (TEE).

The model used for estimating the concentration in the central compartment cannot distinguish between the parameters sets: (V; ka; ke) and (V ke=ka; ke; ka). Previous studies on amlodipine suggest that ka $\gg$ ke and this is imposed here to solve the ambiguity.

In practice, to estimate population PK studies, starting values are taken from early PK studies where the parameters are estimated on a few subjects who are intensively sampled. To speed up computations within the simulation context, starting values were chosen to be the exact population parameters perturbed by a normal random variable with mean 0 and variance 0.25. In case of non convergence, a non invertible matrix, or unrealistic covariance estimates, the fitting procedure was repeated ten times by changing the starting values.

Table 4.1: Convergence rates

	Scenario		
Strategy	**1**	**2**	**3**
T	100	96	99
E	100	97	97
+	100	100	100

4.6   Results

Figure 4.4 displays for each of the 9 different data types box plots summarizing the 100 parameter estimates. The horizontal lines indicate true median values used for data generation.

Under NSSE, we encountered some convergence problems under strategies 2 and 3 but as shown in table 4.1, convergence was achieved in at least 96% of the cases.

Comparing E and T results shows, as expected, that errors in time recall yield less accurate estimates. The impact on V and k_e is particularly striking under the regular scenario 1.

In terms of bias, NSSE works reasonably well with 1T, but serious biases occur everywhere else. As already pointed out by Wang and co-workers (1996), it is clear that the increase in variability due to nonadherence cannot be explained by a perfect adherence multiple dose model. In most population pharmacokinetic studies, parameter estimates that relate patient characteristics to pharmacokinetic parameters, here clearance and volume, have the greatest relevance. We find that those parameters are clearly biased when NSSE is used for data sets incorporating real observed adherence (i.e. scenario 3 T and E), it also significantly overestimates the within-patient variability (denoted sd in the figures). TEE on data sets with seven previous dosing records and assuming steady state beforehand yields results closest to the truth for all parameters when applied to scenario 3. This finding suggests that from an informational point of view the ideal situation is not intake history perfectly coinciding with the prescription but a correctly measured variation in drug exposure timing which is independent of the individual PK parameters, and is clearly seen when comparing scenario 1T and 3+ through the density plots from the simulated point estimates in figure 4.5. Densities have narrower spread and are centered closer to the `true' value
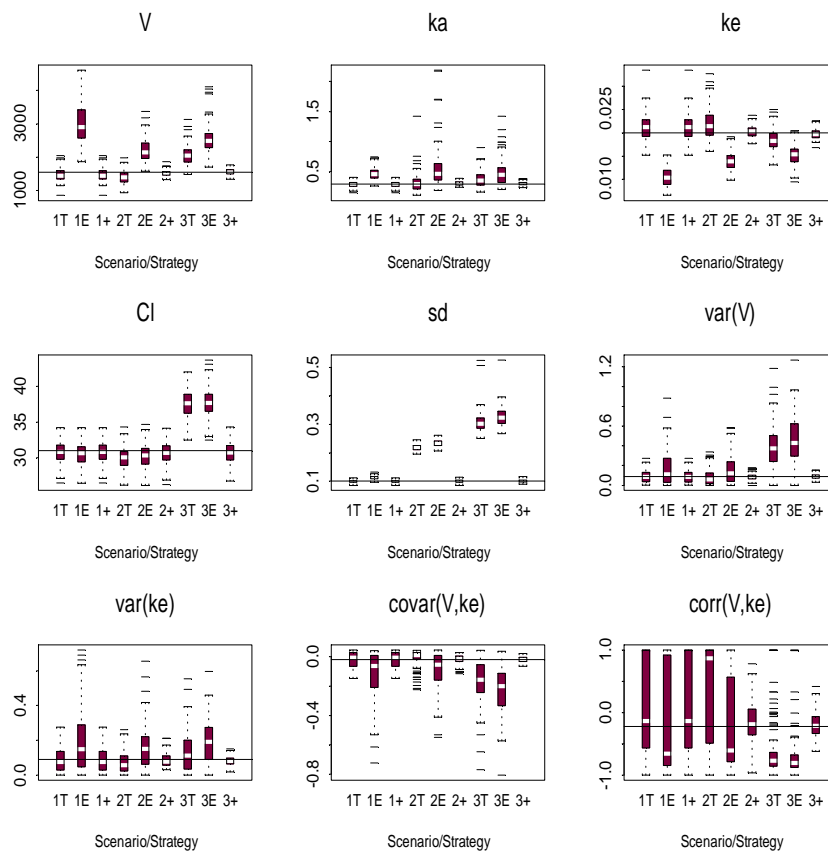
Figure 4.4: Box plot summaries the PK parameter estimates from 100 pseudo data sets; nine scenario/estimation strategies are labelled on the horizontal axis; the horizontal line in each plot denotes median values used for data generation

after applying TEE on real adherence data. In the same vein, Antal and co-workers (1989) noted that observing concentrations at irregular times since last dose also helped enrich the observed information. The opposite happens for example when the last dose is taken in hospital and blood samples are drawn at a fixed time interval thereafter.

Using computer simulation methods, different authors (Girard et al. 1996, Wang et al. 1996, Nony et al. 1998) assessed the variation in pharmacological effects when adherence with the prescribed treatment is a homogeneous process over time (similar to our strategy 2). Comparing our results with theirs, we find higher pharmacokinetic variability with real observed nonadherence data. In a recent paper, Lunn and Aarons (1997) claim that factors influencing inter occasion variability are either unmeasurable or unknown. Based on what we see, we suggest that variation in adherence over time induces high within-patient variability in exposure and could be an important source of variation of an individual's naively estimated PK parameters between dosing occasions. This situation is for instance suggested under scenario 3 for the second patient represented in figure 4.3. It would be interesting to reconsider the problem of Lunn and Aarons incorporating MEMS information.

We finally note that typically in studies relying on patient recall of timing, on top of the biases due to incorrect exposure assumptions, a high percentage (up to 50%) of patients tend to be discarded in final analyses because the requested information is not available. Because such missing data need not to arise `at random' an even more serious bias could result from such pharmacokinetic/pharmacodynamic investigations.

4.7 Discussion

Population PK studies have the great advantage of enabling the study of drug concentrations in a representative sample of patients who take the drug over a possibly long period of time. Statistical methodology that enables accurate determination of population characteristics based on data sets with few observations per subject are thus important for PK/PD research. Unfortunately, exposure histories are often poorly recorded and the extent of nonadherence tends to get underestimated. Adherence data shown here, come from patients who are assigned to hypertensive medication for a period close to one year. In shorter running phase III trials, one may hope for a lesser degree of nonadherence but this is not guaranteed. Accurate observation of drug intake history
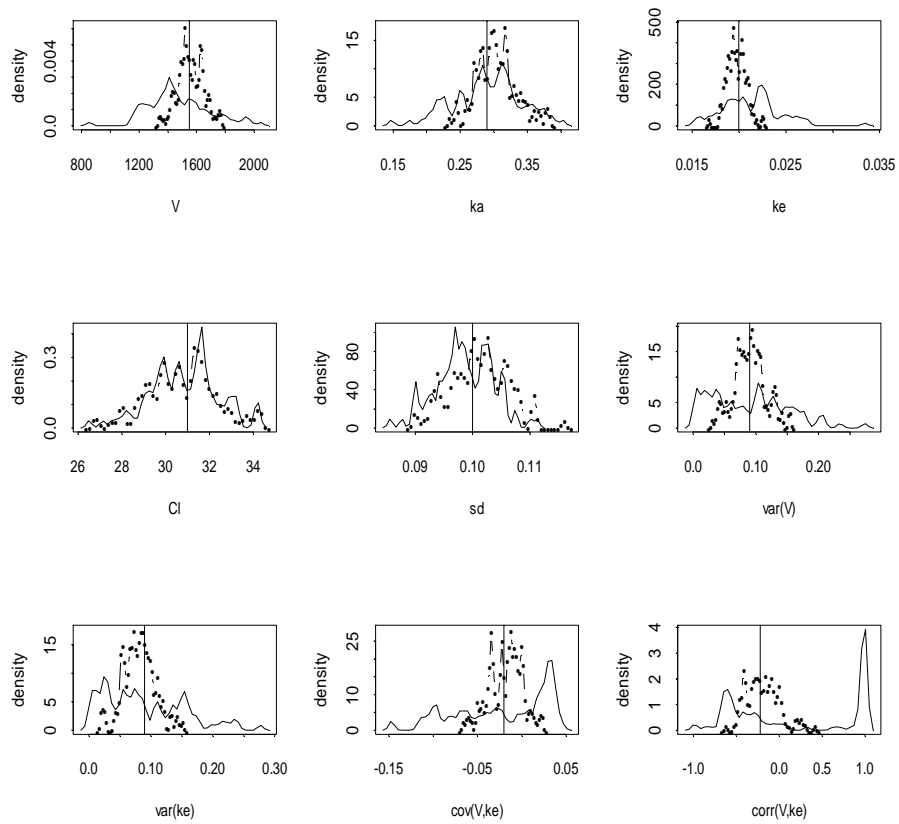
Figure 4.5: Non-parametric estimation of the probability density based on 100 PK Parameter estimates under scenarios 1T (full lines) and 3+ (mixed lines); the vertical line in the plots denotes median values used for data generation

remains key. Electronic monitoring systems may help in this process and perhaps even ultimately rescue population PK/PD studies. However electronic monitoring of dosing histories has not often been used in practice. Results from the presented simulations have re-emphasized that traditional modeling assuming a priori steady state incurs the risk of yielding biased and imprecise estimates. They also showed that better estimates of population PK parameters can be obtained from irregular intake times than from perfect adherence to a prescribed schedule. This consideration suggests that the trial designer might chose to assign drug intake at irregular times in well controlled PK trials.

The implemented TEE assumed steady state prior to the seven dosing records included in the analysis. We investigated an additional estimation strategy (labelled `-' in figures 4.6 and 4.7) assuming no dose was taken beforehand. Not surprisingly, the latter resulted in less accurate estimates for scenario 1 and 2 (see figure 4.6), but very similar estimates for scenario 3 (see figure 4.7). The level of robustness is quite encouraging but of course depends on the length of the number of previous dosing histories used compared to the half-life of the drug studied (Girard et al. 1996). Patient adherence in practice lies somewhere in between the `non' and the `perfect' taker and perhaps better estimates could be obtained by assuming a steady state prior to the seven dosing events using as `steady' dose, the assigned dose multiplied by an average adherence measure (Vrijens and Goetghebeur 1997).

This chapter showed data generated under the popular assumption that absorption rate k_a is fixed. Such assumption however is made more out of mathematical convenience than biological plausibility. For completeness, we ran additional simulations generating similar data but with random k_a. Related results will be presented in the next chapter. When the estimation model nevertheless assumed fixed k_a, we still recovered the average population k_a as one would hope. However, since the estimation model did not account for an added source of variation, this emerged in the form of overestimated variation in other components.
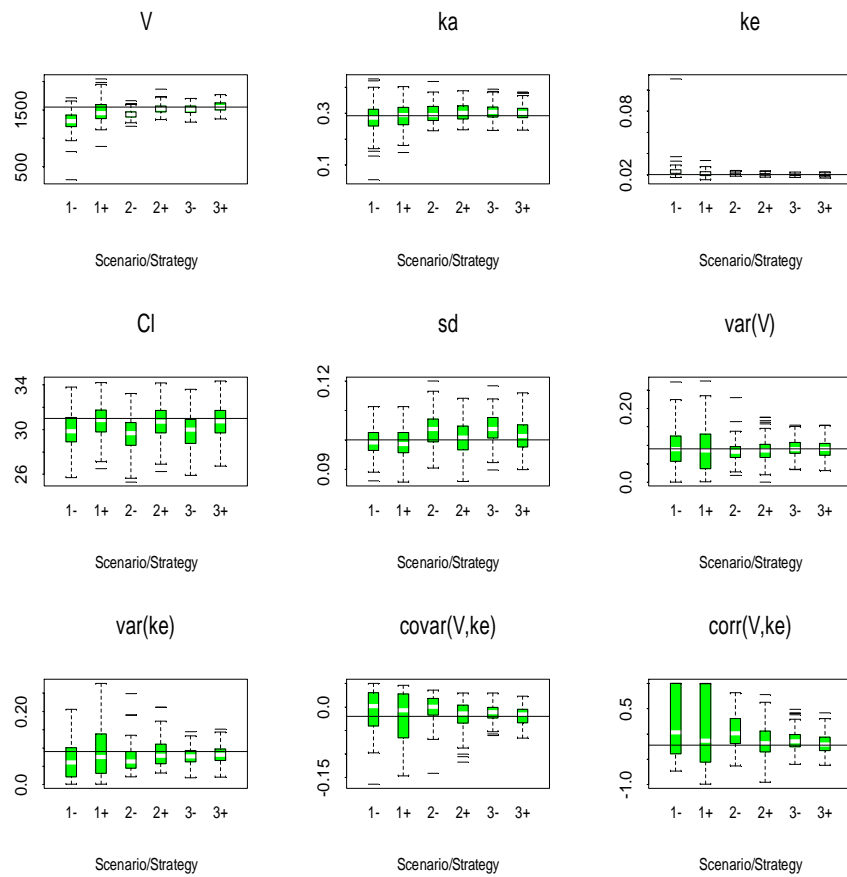
Figure 4.6: Box plot summaries the PK parameter estimates from 100 pseudo data sets; 3 scenario combined with strategy - and + are labelled on the horizontal axis; the horizontal line in each plot denotes median values used for data generation
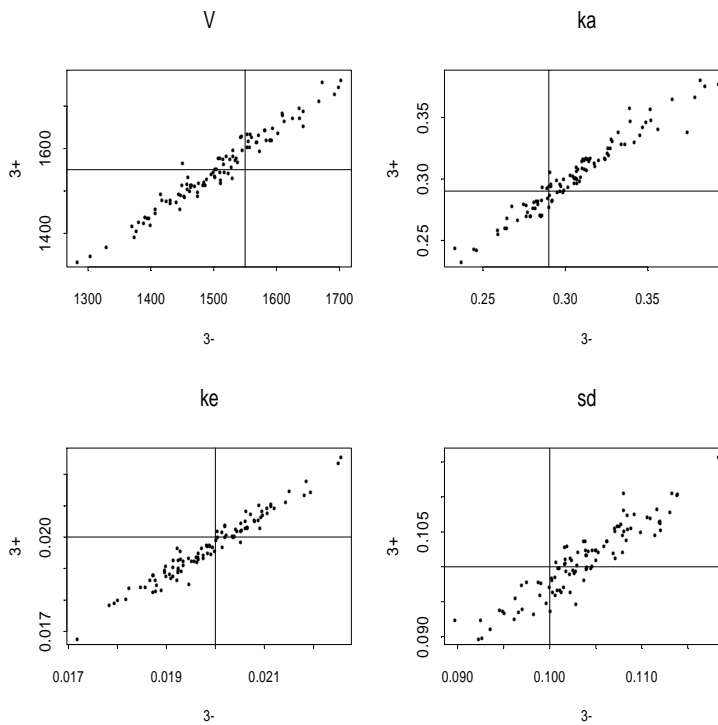
Figure 4.7: Comparison of strategies - and + for four PK parameter estimates under scenario 3; the lines in each plot denote median values used for data generation

Chapter 5

Irregular drug intake reduces bias and improves precision in PK/PD population studies

The next natural step is to study the pharmacodynamic (PD) properties of the drug on top of the PK results derived in the previous chapter. In this context, a pharmacodynamic model describes the relationship between concentration and effect in an individual. A common modeling approach is to relate a single measure of drug exposure, like the area under the concentration-time curve, to a single measure of drug effect, e.g. time to maximum effect. In contrast, and in line with previous chapter, we will focus on PD modeling approaches that relate the entire time course of drug concentration to the time course of drug effect. Population PK/PD studies the variability between-individuals in drug concentration and pharmacological effect when standard dosage regimens are assigned. They constitute a scientific basis for the determination of the optimal dosage of a new drug. Population PK/PD relies on relatively sparse data, but has the great advantage of studying a representative sample of patients who take the drug over a possibly long period of time. We confirm the problem of biased PK/PD estimators in the presence of partial adherence with assigned treatment as it occurs in practice. We propose a global solution by using timing explicit hierarchical nonlinear models and accurate information on a number of previous dose timings. In practice, electronic monitoring methods enable reliable estimation of ambulatory patient's drug dosing histories. Especially for nonlinear PD estimation we found that not only bias is reduced but higher precision is

retrieved from the same number of data points when non-selective irregular drug intake times occur in well controlled studies. We apply methods proposed by Mentré et al. (1997) to investigate the information matrix for hierarchical nonlinear models. This confirms that a substantial gain in precision can be expected due to irregular drug intakes. Intuitively this is explained by the fact that regular takers experience a relatively small range of concentrations which makes it hard to estimate any deviation from linearity in the effect model. We conclude that estimators of PK/PD parameters can benefit greatly from information that enters through greater variation in the drug exposure process.

5.1 Introduction

Careful characterization of the dose-response relationship of a drug is an essential feature of drug development. Early clinical pharmacology studies certainly provide a clue to the correct dose but they have rarely been followed up with studies designed to explore dose-response relationship in the clinical setting. The present research was part of an European project that aimed to design a new clinical trial for the estimation of the pharmacodynamic parameters for a novel negative chronotropic agent for the prevention of myocardial ischaemia. Since we found no documented use of actually observed drug intake histories along with PK and PD measures, we investigated the properties of such model strategies based in part on simulated data. The hope is that ultimately, a better understanding of patient-specific PK/PD and realistic adherence patterns will lead to a better choice of dosing, and a more nearly optimal, cost-efficient treatment of patients.

In this chapter we move on to estimate pharmacodynamic parameters, from just a few outcome measures on a sample of patients, who are partially adherent with a fixed dosing assignment. The estimation problem is harder as the PD model derives its model input from estimated PK parameters. We start by deriving steady state equations for a PD model which builds an Emax model on concentrations in an effect compartment added to the PK model (Holford and Sheiner 1981).

Next we compare estimators which assume steady-state with adherence-explicit estimation under two scenarios: perfect regular adherence and drug intake behavior as observed in the clinical trial. Besides a study of bias, we evaluate precision building on expressions for the information matrix which Mentré et al. (1997) derived for the nonlinear mixed effects model. To assess the potential impact of irregular drug intake, we consider ranges of effect concentrations reached under regular and irregular drug intake and show how they influence an individuals contribution to the information matrix.

In Section 2, we describe the family of PK/PD models studied throughout the chapter. In Section 3 we consider documented PK/PD models (Holford and Sheiner 1981, Gibaldi and Perrier 1982, Reid and Meredith 1990, Donnelly et al. 1993) and study derived plasma concentration and effect profiles for a range of plausible parameters through simulation. Several analysis approaches are implemented and compared on these datasets in Section 4. Section 5 presents a method for the evaluation

of the information matrix and uses it to assess the gain in information when estimating PD parameters from non regular drug intake. Detailed results are presented in Section 6 and further discussed in Section 7.

5.2   Population PK/PD models

Generalizing the concepts of the previous chapter, the complete population PK/PD model consists generally of two parts. First, a structural model expresses a systematic expected effect outcome over time in function of a dosing pattern and an individual's pharmacokinetic, pharmacodynamic parameters. Second, a statistical model describes the joint distribution of pharmacokinetic and pharmacodynamic parameters between individuals as well as the variation of observed effects around the individuals expected effect evolution over time. As an example, we reconsider the model fitted in a clinical trial studying the PK/PD of anti hypertensive response to amlodipine (Donnelly et al. 1993).

As previously described the open, one-compartment model with first order absorption and first order elimination has as parameters the volume of the central compartment V_c (liters), elimination rate k_e (hour $^{-1}$) and absorption rate k_a (hour $^{-1}$). To allow for a lag-time in the concentration-effect relationship (PD model) a standard PK model is augmented by an \effect" compartment (Holford and Sheiner 1981) as illustrated in figure 5.1. Parameter (k_{e0}) of the effect model describes a constant removal rate of drug from the effect compartment and characterizes the phase lag between the change in plasma drug concentration and its effect.

The observed effect over time, $E(t)$, is systolic blood pressure (SBP) reduction, and can be modeled as follows :

$$E(t) = \phi(C_e(t); \theta) = \phi[(\psi * C_c)(t); \theta] \tag{5.1}$$

where ϕ is a known function depending on unknown parameters θ and relates the PK response at the effect site, C_e, to the PD effect. ψ accounts for the non equilibrium between the measured PK response, i.e. $C_c(t)$, and the actual PK response at the effect site, $C_e(t)$. In practice, the only outcomes that are directly measured at one or more points in time t are the concentration in the central compartment $C_c(t)$ and the effect $E(t)$. The symbol $*$ indicates the convolution integral,
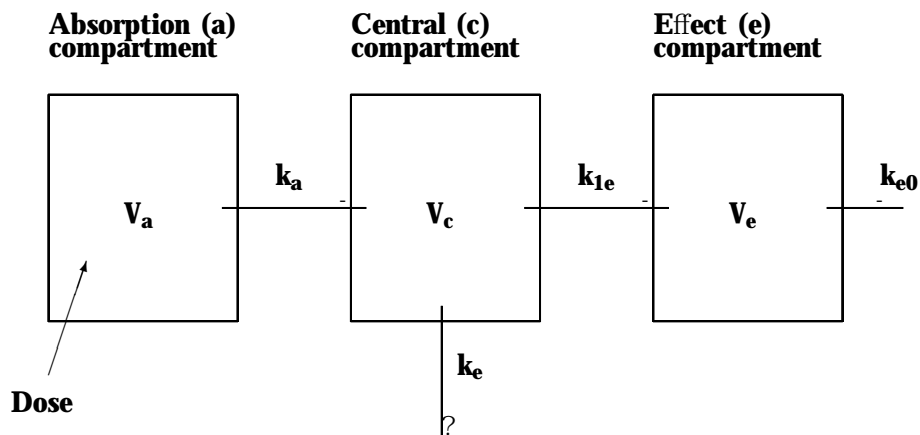
Figure 5.1: One compartment open model augmented by an effect compartment with as parameters, the apparent volume of the absorption compartment (V_a), of the central compartment (V_c), and of the effect compartment (V_e), first-order absorption rate constant (k_a), first-order elimination rate constant (k_e), and equilibration rate constant (k_{e0})

i.e.,

$$C_e(t) = (\psi * C_c)(t) = \int_0^t \psi(t - u)C_c(u)du: \qquad (5.2)$$

In general, ψ can have various biological interpretations. For instance, it can represent the transfer of a drug from one physiological site to another or the transformation of the drug into a metabolite responsible for the measured effect. For the PD model ϕ almost all practical applications make use of one of three forms which are, with increasing complexity:

$$
\begin{array}{lll}
\text{-the linear model} & E(t) = & \theta_0 - \theta_1 \times C_e(t); \\[4pt]
\text{-the Emax model} & E(t) = & \theta_0 - \frac{\theta_1 \times C_e(t)}{\theta_2 + C_e(t)}; \\[4pt]
\text{-the sigmoid Emax model} & E(t) = & \theta_0 - \frac{\theta_1 \times C_e(t)^{\theta_3}}{\theta_2^{\theta_3} + C_e(t)^{\theta_3}}:
\end{array}
\qquad (5.3)
$$

The PD model combined with the PK model of the previous chapter, yield a model for the observable concentration in the central compartment and the effect.

5.3 PK/PD model and plausible effects

For patient i on oral therapy prescribed at one 5 mg dose per day, let $C_{c;i}(t)$, $C_{a;i}(t)$ and $C_{e;i}(t)$ be the typical value of concentration at time t, measured in hours since first dosing, in the central, absorption and effect compartment, respectively. Then in recursive form, the model relating plasma concentration at time t to drug history and concentrations at the previous intake time is given by:

$$
\begin{aligned}
C_{e;i}(t) &= C_{e;i}(t_{d;i})e^{-k_{e0;i}(t-t_{d;i})} \\
&+ C_{c;i}(t_{d;i})\frac{k_{e0;i}}{k_{e;i}-k_{e0;i}}(e^{-k_{e0;i}(t-t_{d;i})}-e^{-k_{e;i}(t-t_{d;i})}) \\
&+ C_{a;i}(t_{d;i})k_{e0;i}k_{a;i}\times \\
&\quad [\frac{e^{-k_{e0;i}(t-t_{d;i})}}{(k_{a;i}-k_{e0;i})(k_{e;i}-k_{e0;i})} + \frac{e^{-k_{e;i}(t-t_{d;i})}}{(k_{a;i}-k_{e;i})(k_{e0;i}-k_{e;i})} \\
&\quad + \frac{e^{-k_{a;i}(t-t_{d;i})}}{(k_{e;i}-k_{a;i})(k_{e0;i}-k_{a;i})}] \\[6pt]
C_{c;i}(t) &= C_{a;i}(t_{d;i})\frac{k_{a;i}}{k_{a;i}-k_{e;i}}\{e^{-k_{e;i}(t-t_{d;i})}-e^{-k_{a;i}(t-t_{d;i})}\} \\
&+ C_{c;i}(t_{d;i})e^{-k_{e;i}(t-t_{d;i})} \\[6pt]
C_{a;i}(t_{d;i}) &= C_{a;i}(t_{d-1;i})e^{-k_{a;i}(t_{d;i}-t_{d-1;i})} + \frac{Dose}{V_{c;i}}
\end{aligned}
\tag{5.4}
$$

with starting conditions:

$$
\begin{aligned}
C_{e;i}(t_{0;i}) &= 0 \\
C_{c;i}(t_{0;i}) &= 0 \\
C_{a;i}(t_{0;i}) &= 0;
\end{aligned}
$$

where $t_{0;i}$ indicates start of study (no dose yet), $t_{d;i} = t_{1;i}; t_{2;i}; ...$ is the dosing time of dose number $d = 1; 2; ...$ with $t_{d;i} \leq t < t_{d+1;i}$. Note that the volume of the absorption compartment, V_a, cancels out of the recursive concentration equation and is not specified. The unknown parameters are thus $k_{a;i}; V_{c;i}; k_{e;i}$ and $k_{e0;i}$.

According to Donnelly et al.(1993), the effect, in this case Systolic Blood Pressure (SBP) change from baseline, is then related to the drug concentration in the effect compartment by means of a nonlinear Emax model.

$$
E_i(t) = \phi(C_{e;i}(t); \theta_i = (E_{max;i}; C_{50;i})) = \frac{-E_{max;i} \times C_{e;i}(t)}{C_{50;i} + C_{e;i}(t)}: \tag{5.5}
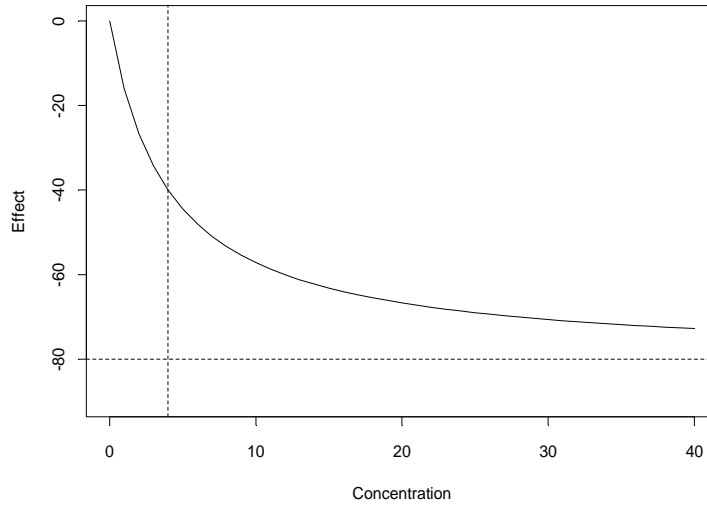$$

Figure 5.2: Evolution of the effect, E(t), in function of the concentrations, C(t), according to an Emax model with parameters $E_{max} = 80$ and $C_{50} = 4$

Assuming $\theta_1 = E_{max;i} \geq 0$, the model describes a decrease in SBP with increasing drug concentration in the effect compartment. The parameter $E_{max;i}$ can be interpreted as the maximum effect for an individual and $\theta_2 = C_{50;i}$ as the concentration in the effect compartment at which half of the maximum effect is reached. We assume no baseline effect ($\theta_0 = 0$). Figure 5.2 presents equation (5.5) graphically.

The combined PK/PD models lead to a structural model for the observable concentration in the central compartment and the effect. The model allows for within-patient variation as follows:

$$
\begin{aligned}
c_i(t) &= C_{c;i}(t)e^{\epsilon_{1;i}(t)} \\
e_i(t) &= E_i(t) + \epsilon_{2;i}(t)
\end{aligned}
\tag{5.6}
$$

with independent $\epsilon_{1;i}(t) \sim N(0; 0.1)$ and $\epsilon_{2;i}(t) \sim N(0; 3)$.

Next, the joint distribution of PK/PD parameters follows a log normal

distribution. Specifically,

$$
\begin{aligned}
k_{a;i} &= k_a \times e^{\eta_{1;i}} \\
k_{e;i} &= k_e \times e^{\eta_{2;i}} \\
V_{c;i} &= V_c \times e^{\eta_{3;i}} \\
k_{e0;i} &= k_{e0} \times e^{\eta_{4;i}} \\
E_{max;i} &= E_{max} \times e^{\eta_{5;i}} \\
C_{50;i} &= C_{50} \times e^{\eta_{6;i}};
\end{aligned}
\tag{5.7}
$$

where k_a, k_e, V_c, k_{e0}, E_{max} and C_{50} are the population parameters. The between-patients random effects, $\eta_i = (\eta_{1;i}; ::::; \eta_{6;i})$ have a multivariate normal distribution, with mean 0 and variance-covariance matrix:

$$
\begin{pmatrix}
0.09 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.09 & -0.02 & 0 & 0 & 0 \\
0 & -0.02 & 0.09 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.09 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.09 & 0.045 \\
0 & 0 & 0 & 0 & 0.045 & 0.09
\end{pmatrix}:
$$

A negative covariance (-0.02) was found between V_c and k_e and a positive one (0.045) between E_{max} and C_{50}. Other parameters were considered independent. The population parameters are also fixed at the values proposed by Donnelly et al. (1993):

$$
\begin{aligned}
& V_c = 1550(\text{liters}) && k_e = 0.02(\text{hour}^{-1}) && k_a = 0.29(\text{hour}^{-1}) \\
& k_{e0} = 0.52(\text{hour}^{-1}) && E_{max} = 80(\text{mmHg}) && C_{50} = 4(\text{ng=ml}):
\end{aligned}
\tag{5.8}
$$

To show the impact of perfect adherence with an assigned regimen of 1 pill per day (Scenario regular) as opposed to true observed adherence (Scenario irregular) on population PK/PD estimation we simulate plasma concentration curves and effect profiles on top of each of the adherence scenarios (dosing histories). Specifically for 60 patients we consider:

- <u>Scenario r</u>: the drug is taken exactly every 24 hours as prescribed for the same length as the observed period. This scenario is similar to scenario 1 in the previous chapter but here we used a random k_a.

- <u>Scenario i</u>: actually observed dosing intervals (Bovet et al. 1997) are used as recorded by the MEMS devices and described in the previous chapter. This scenario is similar to scenario 3 in the previous chapter, with random k_a.

We generate the concentration and effect curves from the one compartment PK model augmented by an effect compartment (5.4) and the Emax PD model (5.5) with parameter set $(k_{a;i}; k_{e;i}; V_{c;i}; k_{e0;i}; C_{50;i}; E_{max;i})$. Specifically, we consider for each of the 60 individuals:

1. One draw $(k_{a;i}; k_{e;i}; V_{c;i}; k_{e0;i}; C_{50;i}; E_{max;i})$ from the joint distribution between individuals.

2. The scenario-specific dosing history D_i, combining it with the parameter set and random observation errors as in equations (5.6) to generate a concentration and an effect curve over time.

3. A set of selected observations at discrete times, 3 to 11 random time points (hospital visits) which are drawn under certain pragmatic constraints. The times must come after one month of prescribed therapy, cannot fall within the same week and are ignored if associated with a plasma concentration lower than 1 ng/ml. For each scenario, one plasma concentration and one effect measure are considered per visit.

In summary, the data vector per individual i consists of n_i time points $T_i = (t_1; ....; t_{n_i})$ at which concentrations in the central compartment, $C_i = (c_1; ....; c_{n_i})$, and effects, $E_i = (e_1; ....; e_{n_i})$ are recorded along with the dosing information (D_i).

For both scenarios we generated 100 data sets of 60 individuals, each time regenerating steps 1 and 3 in an independent random fashion.

5.4   Estimation methods

To estimate parameters of a well specified hierarchical nonlinear model, one can use the linearization approach (Davidian and Giltinan 1995) to maximize the full likelihood describing PK and PD outcomes. However this approach frequently encounters convergence problems due to the complexity of the combined PK/PD model. The data are therefore typically analyzed sequentially. The two stage strategy analyzes PK data first. Then individual PK parameters are estimated and fixed before fitting PD data. We will further compare both approaches.

Steady model

When the last dose intake is the only one that is recorded and the assumption is made that patients are in steady state at the time of their hospital visit, the recursive equations describing plasma concentrations over time can be simplified. Concentrations in the central and effect compartments at steady state can be expressed in closed form as (proof is given in appendix):

$$C_i^{SS}(t) = \frac{\text{Dose}}{V_{c;i}} \frac{k_{a;i}}{k_{a;i} - k_{e;i}} f\frac{1}{1 - e^{-k_{e;i}\tau}}e^{-k_{e;i}t} - \frac{1}{1 - e^{-k_{a;i}\tau}}e^{-k_{a;i}t}g; \quad (5.9)$$

$$
\begin{aligned}
C_e^{SS}(t) = \;& \frac{D}{V_c}\frac{1}{1-e^{-k_a\tau}}f(k_a;k_e;k_{e0};t) \\
& + \frac{D}{V_c}f(k_a;k_e;k_{e0};\tau)\frac{e^{-k_{e0}t}}{(1-e^{-k_a\tau})(1-e^{-k_{e0}\tau})} \\
& + \frac{D}{V_c}\frac{k_a k_{e0}}{(k_a-k_e)(k_e-k_{e0})}\frac{(e^{-k_{e0}t}(1-e^{-k_e\tau})-e^{-k_e t}(1-e^{-k_{e0}\tau}))(e^{-k_e\tau}-e^{-k_a\tau})}{(1-e^{-k_e\tau})(1-e^{-k_{e0}\tau})(1-e^{-k_a\tau})}
\end{aligned}
$$
$$\tag{5.10}$$

where

$$f(k_a;k_e;k_{e0};t) = k_a k_{e0} \; [\frac{e^{-k_{e0}t}}{(k_a-k_{e0})(k_e-k_{e0})} + \frac{e^{-k_e t}}{(k_a-k_e)(k_{e0}-k_e)} + \frac{e^{-k_a t}}{(k_e-k_a)(k_{e0}-k_a)}]:$$

Combined with equation (5.5) this defines the structural PD model for the observed effect at steady state. Estimation based on the Traditional 'naive steady state' assumption is applied to the data issued from the regular dosing scenario (r) and irregular dosing scenario (i) and labelled respectively 'rT' and 'iT' in the results.

Recursive model

When the data include complete dosing histories one is led to the complete recursive model (equation 5.4) to model the concentrations in the central and effect compartments. One saves computation time with negligible loss in accuracy (Girard et al. 1996) by explicitly including only the seven previous dose intake records (D_i) and assuming steady state before that. This 'timing explicit' estimation on data following irregular dosing scenario (i) is labelled 'i+' in the sequel.

Further details of the estimation procedure

We argued before that few samples are typically collected during the absorption phase, which leads to the common practice of treating k_a as constant across the population. Similarly for k_{e0}, which accounts for the non-equilibrium observed between the central and the effect compartment. Hence the data below are generated from random k_a and k_{e0}'s but are analyzed as if they were fixed.

The model for concentrations in the central compartment cannot distinguish between parameter sets, $(V_c; k_a; k_e)$ and $(V_c k_e = k_a; k_e; k_a)$. Previous studies on amlodipine suggest that $k_a \gg k_e$ and this is imposed here to solve the ambiguity.

In our simulations, starting values were exact population parameters perturbed by a normal error with variance 0.25. Unrealistically large values, i.e. $E_{max} > 500$(mmHg) and/or $C_{50} > 50$(ng=ml) were considered a sign of non-convergence or convergence to a local maximum only. In case of non convergence, convergence to a local maximum, a non-invertible matrix, or unrealistic estimates, the fitting was repeated five times by changing the starting values. We used the NLME macro (version 3.2, available from http://nlme.stat.wisc.edu/) for S-PLUS (version 4.5) to implement this. The code of the NLME macro was adapted in order to implement the recursive models.

5.5 Assessing the gain in PD-information from irregular intake

According to (5.5) and (5.6), the relationship between the predicted effect and the concentration in the effect compartment is assumed to be:

$$
\begin{aligned}
E_i(T_i) &= \phi(C_{e;i}(T_i); \theta_i) + \epsilon_{2;i} \\
&= \frac{-e^{\ln(E_{max;i})} \times C_{e;i}(T_i)}{e^{\ln(C_{50;i})} + C_{e;i}(T_i)} + \epsilon_{2;i}
\end{aligned}
\tag{5.11}
$$

where ϵ_i is a n_i-vector of normally distributed errors with zero mean and variance-covariance matrix R_i which depends on i only through its dimension.

In our simulations, we let $\theta_i = (\ln(E_{max;i}); \ln(C_{50;i}))$ be normally distributed with mean vector and variance-covariance matrix respectively

given by:

$$\mathbf{M} = (\ln(E_{max} = 80); \ln(C_{50} = 4))$$

$$\Gamma = \begin{pmatrix} 0.09 & 0.045 \\ 0.045 & 0.09 \end{pmatrix} : \tag{5.12}$$

The block of the Fisher information matrix I corresponding to the 5 effect parameters (M and Γ) is denoted Ie and based on the data obtained from N individuals, it can be written as (Mentré et al. 1997):

$$\mathbf{Ie} = \sum_{i=1}^{N} \mathbf{Ie_i} = \sum_{i=1}^{N} \begin{pmatrix} \mathbf{Ie_i^M} & 0 \\ 0 & \mathbf{Ie_i^\Gamma} \end{pmatrix} \tag{5.13}$$

where $\mathbf{Ie_i^M}$ is the 2×2 information matrix for the mean vector M and $\mathbf{Ie_i^\Gamma}$ is the 3×3 information matrix for the (1,1), (1,2), and (2,2) elements of the covariance matrix Γ. Let $\mathbf{B_i}$ be the $n_i \times 2$ Jacobian matrix

$$\mathbf{B_i}(\theta_i; \mathbf{T_i}) = \frac{\delta}{\delta \theta_i} \phi(\mathbf{C_{e;i}}(\mathbf{T_i}); \theta_i)$$

$$= [\frac{-e^{\ln(E_{max;i})} C_{e;i}(T_i)}{e^{\ln(C_{50;i})} + C_{e;i}(T_i)}; \frac{e^{\ln(E_{max;i})} e^{\ln(C_{50;i})} C_{e;i}(T_i)}{(e^{\ln(C_{50;i})} + C_{e;i}(T_i))^2}] \tag{5.14}$$

It can be shown (Mentré et al. 1997) that in first order approximation:

$$\mathbf{Ie_i^M} = \Gamma^{-1}[\Gamma - f\mathbf{B_i}(M; T_i)^0 R_i^{-1} \mathbf{B_i}(M; T_i) + \Gamma^{-1} g^{-1}]\Gamma^{-1} : \tag{5.15}$$

A simple algorithm derives $\mathbf{Ie}^\Gamma$ from $\mathbf{Ie}^M$ as follows. Let for two indices k and l ($l \leq k$), $k \bullet l$ be defined as $k \bullet l = l + k(k-1)=2$. The $k \bullet l; m \bullet n$ element for the matrix $\mathbf{Ie_i^\Gamma}$, for $k = 1; ....; p; l \leq k; m = 1; ....; p; n \leq m$ is then given by:

$$[\mathbf{Ie_i^\Gamma}]_{k\bullet l; m\bullet n} = 2^{-\delta_{kl}} 2^{-\delta_{lm}} f[\mathbf{Ie_i^M}]_{l;m}[\mathbf{Ie_i^M}]_{k;n} + [\mathbf{Ie_i^M}]_{l;n}[\mathbf{Ie_i^M}]_{k;m} g \tag{5.16}$$

where $\delta_{kl} = 1$ if $k = l$ and 0 otherwise.

To examine the PD information gain due to irregular intake, we focus on a `typical patient' whose parameters equal the population medians (5.8). When such patient achieves a perfectly regular intake, the steady state $C_e(t)$ ranges between 5.84 and 7.36 (ng/ml). A different dose-timing typically increase the observed concentration range. This is seen in figure 4.2 as well as figure 5.3 where we have plotted for each patient the expected SBP reduction as a function of the concentration in the effect compartment for perfect takers and for observed dosing histories. As often reported (Urquhart 1994, Vrijens and Goetghebeur 1997), we face much more under- than over- dosing.
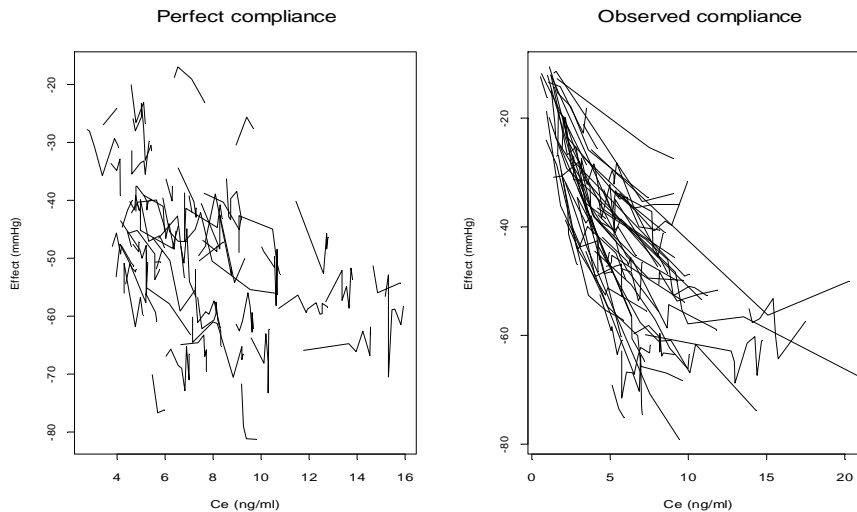
Figure 5.3: Profile plots of the SBP reduction in function of the concentration in the effect compartment. The first picture is as realized by perfect adherers, the second one is based on observed adherence

This phenomenon stretches the range of observed concentrations towards smaller concentrations, whilst only few high concentrations are observed. Hence to quantify the contribution of a single `non adherent' patient to the information matrix, we consider different sets of observed concentrations as described in figure 5.4.

For the evaluation of a typical patient who under-doses, we consider the concentration range of a perfect adherer (labelled SEQ1) ranging from 5:84 to 7:36 ng=ml and stretch it sequentially (SEQ2-SEQ5) towards 0, reducing the lowest concentration by 1 ng=ml each time. We consider then 10 concentrations at equally spaced intervals within those defined ranges. For a typical patient who overdoses, we consider again the reference interval of a perfect adherer (SEQ6=SEQ2) but now simply add one observation at a systematically 1 ng=ml higher concentration resulting in SEQ7-SEQ19 respectively. In (SEQ 21-24) both such under- and over-dosing are considered by combining SEQ2-SEQ5 and SEQ7-SEQ10 respectively. For reference purposes, SEQ20=SEQ6=SEQ1=interval of a perfect adherer.

Table 5.1: Convergence rates

	rT	iT	i+
PK	98%	98%	100%
PD conditional on PK	43%	71%	99%
Simultaneous PK, PD	63%	91%	94%

5.6   Results

5.6.1   Simulated datasets

Figures 5.5 (PK) and 5.6 (PD) display for each of the 6 different combinations of dosing histories and estimation methods box plots summarizing the estimates from 100 pseudo data sets. In each plot we present three scenario/estimation strategies with two implementations: sequential and simultaneous PK/PD estimation. The horizontal lines indicate true median values used for data generation.

As far as the PK part of the model goes, results for sequential estimation process are very similar to the one reported in the previous chapter. Small changes can be expected from the simultaneous modeling as same PK-information is then borrowed from effect-data thru the complete data model. In summary, PK results with random k_a indeed confirm findings of the previous chapter. The increase in variation due to nonadherence cannot be explained by a perfect adherence multiple dose model. Timing explicit estimation (strategy +) applied to irregular drug intake (scenario i) yields results closest to this truth for all PK parameters.

In what follows, we focus on new results for PD.

Whereas traditional naive steady state estimation applied to perfect adherers (rT) worked reasonably well for PK estimation, it gives the worst PD estimates. In addition, we encountered much more convergence problems under rT and the box plots represent only the relatively small proportion (about 50 %) of estimators that converged. Table 5.1 shows the convergence rates.

Surprisingly and in contrast with the PK situation, even iT yields a higher percentage of convergence and more precise PD estimates than rT. This indicates that the `exact' model for perfect adherers contains

less information than a steady state approximation for irregular drug intakes. Timing explicit estimation (+) on data sets with seven previous dosing records, assuming steady state beforehand, yields results closest to the truth for all parameters when analyzing irregular intake. Comparing rT with i+, we observe to what extent estimators have a narrower spread and are centered closer to the `true' value after applying timing explicit estimation on real adherence data. Simultaneous PK/PD estimation yields slightly more precise PD estimates than does the two stage method. When naive steady state estimation is used to analyze observed adherence (iT) , the within-patient variability in effect measurements is substantially overestimated, just as for PK.

5.6.2 Gain in information

Figure 5.7 shows in the five first plots for each sequence (as defined in figure 5.4) the standard deviation of the diagonal elements of the information matrix (I_e). They are computed as the square root of the diagonal elements of the inverse of the information matrix. The three last plots represent the values of the determinant of the information matrix for the mean, for the variance-covariance and the combined matrix respectively.

Sequences 21 to 24 are the most plausible in practice. One sees clearly that the standard deviation decreases and a substantial gain in information can be expected from irregular dosings. The three last plots show that this gain is growing faster when delaying intake time (underdosing: SEQ 2-5) than when shortening the dosing interval (overdosing: SEQ 7-20).

To show that those qualitative results continue to hold when C_{50} is larger than the observed concentration range, we repeated the exercise for an hypothetical value of $C_{50} = 10$ ng/ml. Figure 5.8 still shows a clear decrease in the standard deviation of all parameters. We thus conclude that substantial information can be gained from irregular dosings. This intuitively follows from the fact that regular drug intake restricts the observational concentration range as illustrated in figure 5.3.

5.7   Discussion

The complex structure of nonlinear hierarchical models used for population PK/PD makes it hard to examine the influence certain design

features may have on the information matrix. We have adapted the theory developed by Mentré et al. (1997) to examine the impact of non regular drug intake patterns on the PK/PD information. Focusing on adherence patterns which occur in practice, we studied the influence of those patterns under the assumption that they evolve independently of: a) the PK/PD parameters of the patient, and b) the random highs or lows of the concentration and effect measures at the times of measurement. The latter are not satisfied when partial adherence is selective (Fischer and Goetghebeur 1999). In those cases, also the dose-timing-specific analysis presented here will lead to biased results. It is generally expected that adherence is less selective when patients suffering from an asymptomatic disease and are treated with an asymptomatic drug. Sometimes selectivity can be eliminated by adjusting for baseline covariates like sex, age or gene information.

Whereas population PK analysis is more and more reported in the literature, combined population PK/PD analyses are less often seen because of problems with hysteresis loops, lack of convergence, unrealistic parameter estimates, ... Recent recommendations in the field (Aarons et al. 1996) teach the researcher to stimulate adherence by digital drug dispensers, patient diaries, pill counts, financial support, adherence counselling, a run-in period to preselect the good adherers, ... This practice aims to include `standardized' patients which are close to the `ideal' perfect adherer. However, we have shown that regular drug intake restricts observations to a narrow window of the concentration-effect relationship. This is much more important for nonlinear PD models than for compartmental PK models. In the former case, this often results in limited information which may prevent fitting of the appropriate model. This is particularly true when `hard nonlinearities' occur, which may be expected at the beginning or at the end of the dosing period, e.g. rebound effects when dosings stop (Kleinbloesem et al. 1987, Houston and Hodge 1988, Psaty et al. 1990, Gilligan et al. 1991, Rosenbaum et al. 1998). Our results suggest that acknowledging actual dose timing as observed in current practice may help solve the problem. Indeed, estimators of PK and particularly PD parameters can gain serious precision from information that enters through greater variation in the drug exposure process. In practice patients do naturally vary the dosing interval and electronic monitoring of a wide range of over- and under-dosing patterns is an underestimated opportunity to observe and model realistic concentration-effect relationships. Alternatively one might use the results of the present research to design independent irregular in-

take schedules under controlled conditions where perfect adherence can be imposed (Duffull et al., 2001). Knowing that 78 % of the drugs in therapeutic use have plasma half-lives that are less than 12 hours, our simulations based on a half-life of 36 hours are conservative in their assessment of the impact of irregular intakes.

5.8 Appendix

In this appendix we derive steady state equations for the one compartment open model augmented by an effect compartment. If $y_i(t)$ represents the quantity of drug in compartment $i = (a; c; e)$ at time t after dosing, then according to figure (5.1) the differential equations describing the rate of change over time of drug in the three compartments is given by

$$\begin{cases} \frac{dy_a(t)}{dt} = -k_a y_a(t) \\ \frac{dy_c(t)}{dt} = k_a y_a(t) - k_e y_c(t) \\ \frac{dy_e(t)}{dt} = k_{1e} y_c(t) - k_{e0} y_e(t) \end{cases} \tag{5.17}$$

where $k_a; k_e$ and k_{e0} are apparent first order rate constants of transfer between compartments and assuming $k_{1e} <<< k_e$. To describe the time course of the amount of drug in the body after drug intake, equations (5.17) must be integrated. The method of Laplace transforms will be employed. The transform of (5.17) is

$$\begin{cases} s\bar{y}_a(t) - y_a(t = 0) = -k_a \bar{y}_a(t) \\ s\bar{y}_c(t) - y_c(t = 0) = k_a \bar{y}_a(t) - k_e \bar{y}_c(t) \\ s\bar{y}_e(t) - y_e(t = 0) = k_{1e} \bar{y}_c(t) - k_{e0} \bar{y}_e(t) \end{cases} \tag{5.18}$$

where s is the Laplace transformation, and $(t = 0)$ indicates the intake time. For a single dose (D), where $y_a(t = 0) = \text{Dose}$ and $y_c(t = 0) = y_e(t = 0) = 0$, rearrangement of (5.18) yields

$$\bar{y}_e(t) = \frac{Dk_{1e}k_a}{(s + k_{e0})(s + k_a)(s + k_e)} \tag{5.19}$$

which when solved using a table of Laplace transforms gives

$$y_e(t) = Dk_{1e}k_a \left[\frac{e^{-k_{e0}t}}{(k_a - k_{e0})(k_e - k_{e0})} + \frac{e^{-k_e t}}{(k_a - k_e)(k_{e0} - k_e)} + \frac{e^{-k_a t}}{(k_e - k_a)(k_{e0} - k_a)} \right] \tag{5.20}$$

The concentration of drug in the effect compartment (C_e) is then derived by dividing y_e by the effect compartment volume (V_e). At equilibrium,

the rates of drug transfer between the central and effect compartments are equal, i.e.

$$k_{1e}y_c(t) = k_{e0}y_e(t) \tag{5.21}$$

or

$$k_{1e}V_cC_c(t) = k_{e0}V_eC_e(t) \tag{5.22}$$

or

$$\frac{k_{1e}}{V_e} = \frac{k_{e0}C_e(t)}{V_cC_c(t)} = \frac{k_{e0}}{V_c}K_p(t) \tag{5.23}$$

where $K_p(t)$ is the partition coefficient and was set equal to 1 in this paper.

Then rearranging (5.20) will yield:

$$C_e(t) = \frac{D}{V_c}K_p(t)k_{e0}k_a \left[\frac{e^{-k_{e0}t}}{(k_a-k_{e0})(k_e-k_{e0})} + \frac{e^{-k_et}}{(k_a-k_e)(k_{e0}-k_e)} + \frac{e^{-k_at}}{(k_e-k_a)(k_{e0}-k_a)}\right] \tag{5.24}$$

For multiple doses (D), rearrangement of (5.18) yields

$$\overline{y}_e(t) = \frac{y_e(t=0)}{(s+k_{e0})} + \frac{y_c(t=0)k_{1e}}{(s+k_{e0})(s+k_e)} + \frac{y_a(t=0)k_{1e}k_a}{(s+k_{e0})(s+k_a)(s+k_e)} \tag{5.25}$$

which when solved using a table of Laplace transforms gives

$$\begin{aligned}
y_e(t) = \quad & y_e(t=0)e^{-k_{e0}t} \\
+ \quad & y_c(t=0)\frac{k_{1e}}{k_e-k_{e0}} \ (e^{-k_{e0}t} - e^{-k_et}) \\
+ \quad & y_a(t=0)k_{1e}k_a \ \left[\frac{e^{-k_{e0}t}}{(k_a-k_{e0})(k_e-k_{e0})} + \frac{e^{-k_et}}{(k_a-k_e)(k_{e0}-k_e)} \right. \\
& \left. + \frac{e^{-k_at}}{(k_e-k_a)(k_{e0}-k_a)}\right]
\end{aligned} \tag{5.26}$$

Dividing both sides by the apparent volume of the effect compartment (V_e) and assuming (5.23) will yield the following concentration-time relationship in the effect compartment

$$\begin{aligned}
C_e(t) = \quad & C_e(t=0)e^{-k_{e0}t} \\
+ \quad & C_c(t=0)K_p(t)\frac{k_{e0}}{k_e-k_{e0}} \ (e^{-k_{e0}t} - e^{-k_et}) \\
+ \quad & C_a(t=0)K_p(t)k_{e0}k_a \ \left[\frac{e^{-k_{e0}t}}{(k_a-k_{e0})(k_e-k_{e0})} + \frac{e^{-k_et}}{(k_a-k_e)(k_{e0}-k_e)} \right. \\
& \left. + \frac{e^{-k_at}}{(k_e-k_a)(k_{e0}-k_a)}\right]
\end{aligned} \tag{5.27}$$

We note that for mathematical convenience, the concentration in the absorption compartment is defined as the amount of drug in that compartment divided by the volume of the central compartment, however this quantity is difficult to interpret biologically.

For drugs administered in a constant dose at regular intervals (τ), Gibaldi and Perrier (Gibaldi and Perrier 1982) have shown that, at the time of the Nth dose, the concentrations in the absorption and central compartment are respectively given by

$$\left(\begin{aligned} C_a(N) &= \frac{D}{V_c} \frac{1 - e^{-Nk_a\tau}}{1 - e^{-k_a\tau}} \\ C_c(N) &= \frac{D}{V_c} \frac{k_a}{k_a - k_e} \left[\frac{1 - e^{-Nk_e\tau}}{1 - e^{-k_e\tau}} - \frac{1 - e^{-Nk_a\tau}}{1 - e^{-k_a\tau}} \right] \end{aligned} \right. \tag{5.28}$$

It can be shown by simple algebra that the concentration in the effect compartment at the time of the Nth dose is given by

$$\begin{aligned} C_e(N) =\ & \frac{D}{V_c} k_a k_{e0} \left[\frac{e^{-k_{e0}\tau}}{(k_a - k_{e0})(k_e - k_{e0})} + \frac{e^{-k_e\tau}}{(k_a - k_e)(k_{e0} - k_e)} + \frac{e^{-k_a\tau}}{(k_e - k_a)(k_{e0} - k_a)} \right] \\ & \times g(k_a; k_{e0}; N - 1) \\ +\ & \frac{D}{V_c} \frac{k_a k_{e0}}{(k_a - k_e)(k_e - k_{e0})} (e^{-k_{e0}\tau} - e^{-k_e\tau}) \\ & \times f e^{-k_e\tau} g(k_e; k_{e0}; N - 2) - e^{-k_a\tau} g(k_a; k_{e0}; N - 2) g \end{aligned} \tag{5.29}$$

where g is defined as follows

$$g(k_1; k_2; N) = \overset{N}{\underset{i=0}{P}} \overset{N-i}{\underset{j=0}{P}} e^{-ik_1\tau} e^{-jk_2\tau}$$

For prolonged periods of dosing, a steady state is reached when the amount of drug in the body reaches a level where the input rate equals the output rate. The equation describing the time course of the drug in the effect compartment at the plateau or steady state, can be obtained by setting N to 1. Then g can be rewritten as

$$\begin{aligned} g(k_1; k_2; N) =\ & \overset{N}{\underset{i=0}{P}} \overset{N}{\underset{j=0}{P}} e^{-ik_1\tau} e^{-jk_2\tau} - \overset{\aleph}{\underset{i=1}{}} \overset{\aleph}{\underset{j=N-i+1}{}} e^{-ik_1\tau} e^{-jk_2\tau} \\ & \underbrace{\phantom{\hspace{4cm}}}_{\{z\phantom{\hspace{3cm}}\}} \\ & \quad !\ 0\ \text{if } N\ !\ 1 \\ \sim\ & (1 + e^{-k_1\tau} + \cdots + e^{-Nk_1\tau})(1 + e^{-k_2\tau} + \cdots + e^{-Nk_2\tau}) \\ \sim\ & \frac{1 - e^{-Nk_1\tau}}{1 - e^{-k_1\tau}} \frac{1 - e^{-Nk_2\tau}}{1 - e^{-k_2\tau}} \\ \sim\ & \frac{1}{1 - e^{-k_1\tau}} \frac{1}{1 - e^{-k_2\tau}} \end{aligned}$$

Assuming N is large, substituting $C_a(N)$, $C_c(N)$ and $C_e(N)$ for initial concentrations at ($t = 0$) in (5.27) will yield

$$\begin{aligned} C_e^{SS}(t) =\ & \frac{D}{V_c} \frac{1}{1 - e^{-k_a\tau}} f(k_a; k_e; k_{e0}; t) \\ & + \frac{D}{V_c} f(k_a; k_e; k_{e0}; \tau) \frac{e^{-k_{e0}t}}{(1 - e^{-k_a\tau})(1 - e^{-k_{e0}\tau})} \\ & + \frac{D}{V_c} \frac{k_a k_{e0}}{(k_a - k_e)(k_e - k_{e0})} \frac{(e^{-k_{e0}t}(1 - e^{-k_e\tau}) - e^{-k_e t}(1 - e^{-k_{e0}\tau}))(e^{-k_e\tau} - e^{-k_a\tau})}{(1 - e^{-k_e\tau})(1 - e^{-k_{e0}\tau})(1 - e^{-k_a\tau})} \end{aligned} \tag{5.30}$$

where

$$f(k_a; k_e; k_{e0}; t) = k_a k_{e0} \left[\frac{e^{-k_{e0}t}}{(k_a - k_{e0})(k_e - k_{e0})} + \frac{e^{-k_e t}}{(k_a - k_e)(k_{e0} - k_e)} + \frac{e^{-k_a t}}{(k_e - k_a)(k_{e0} - k_a)} \right] \tag{5.31}$$
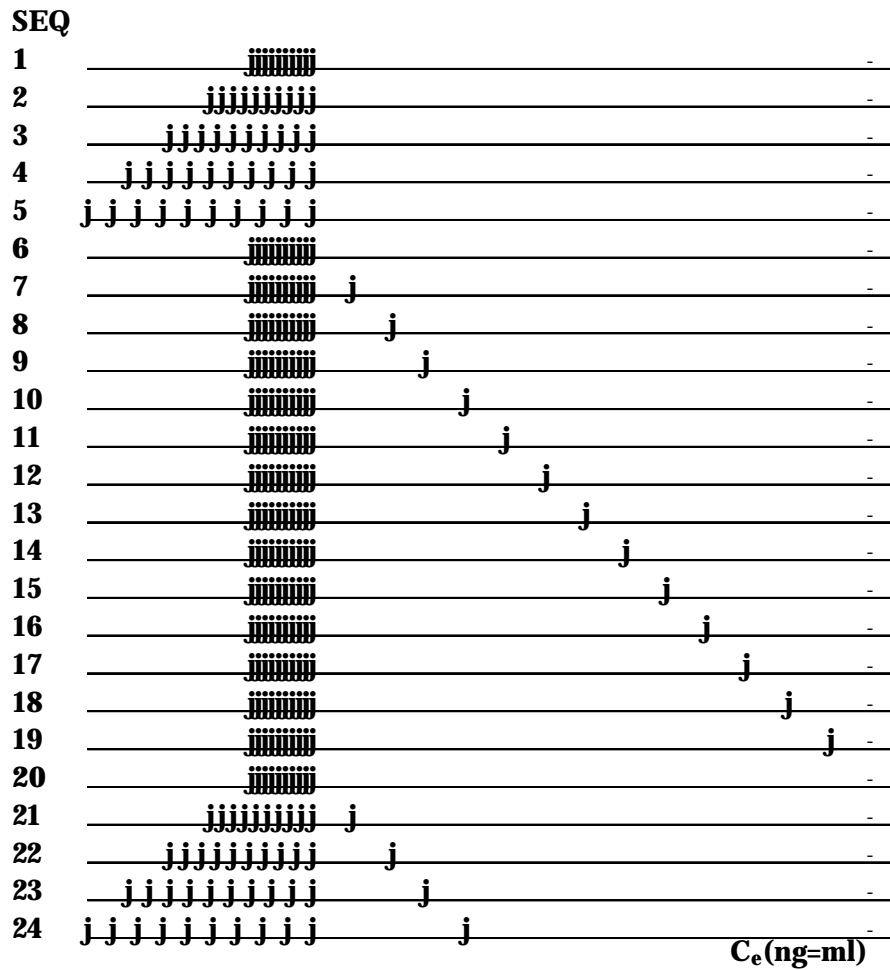
SEQ

Figure 5.4: Observed concentrations in the effect compartment; SEQ 1, 6, 20 refer to the perfect drug taker; SEQ 2-5 and 7-19 represent respectively underadherence and over-adherence; SEQ 21-24 represent a combination of under- and over-adherence.
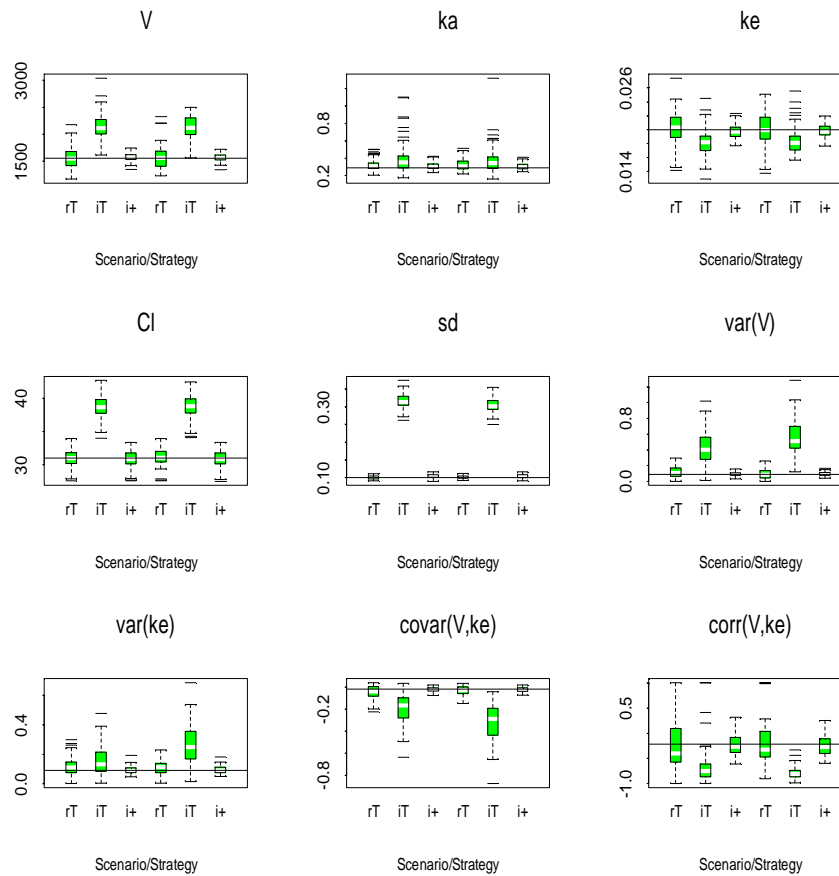
Figure 5.5: Box plot summaries the PK parameter estimates from 100 pseudo data sets; three scenario/estimation strategies are labelled on the horizontal axis (rT: regular intake assuming steady state, iT: irregular intake assuming steady state, i+: irregular intake assuming the recursive model). The two stage estimators are shown in the 3 first box plots while results from the simultaneous estimation process are in the 3 last box plots of each figure; the horizontal line in each plot denotes median values used for data generation
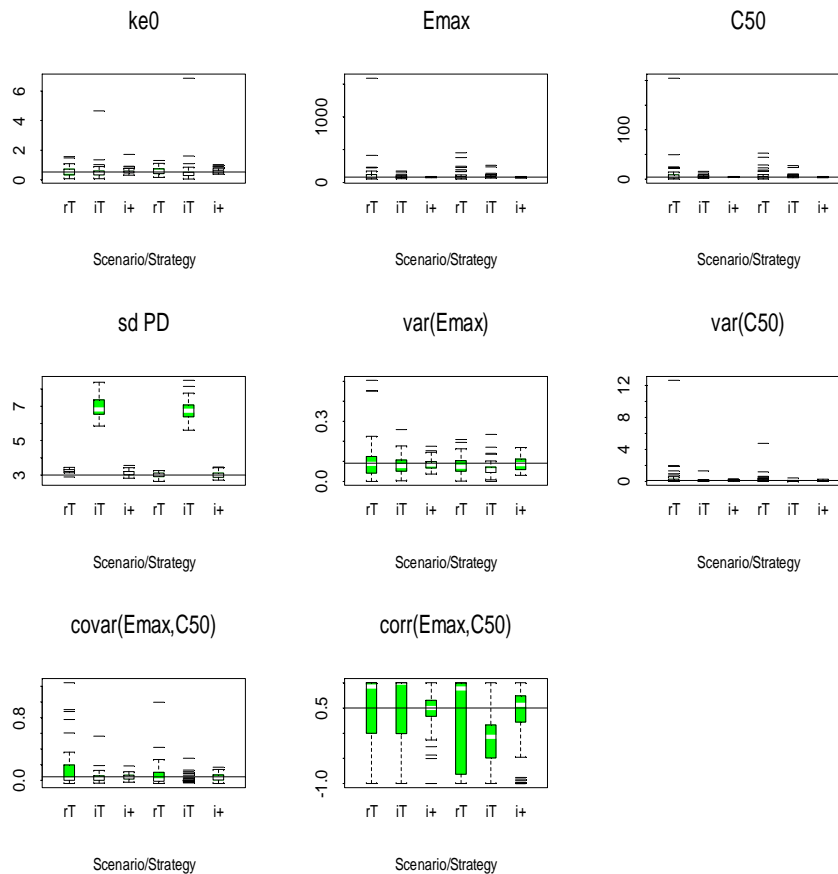
Figure 5.6: Box plot summaries the PD parameter estimates from 100 pseudo data sets; three scenario/estimation strategies are labelled on the horizontal axis (rT: regular intake assuming steady state, iT: irregular intake assuming steady state, i+: irregular intake assuming the recursive model). Two stage estimators are shown in the 3 first box plots while results from the simultaneous estimation process are in the 3 last box plots of each figure; the horizontal line in each plot denotes median values used for data generation
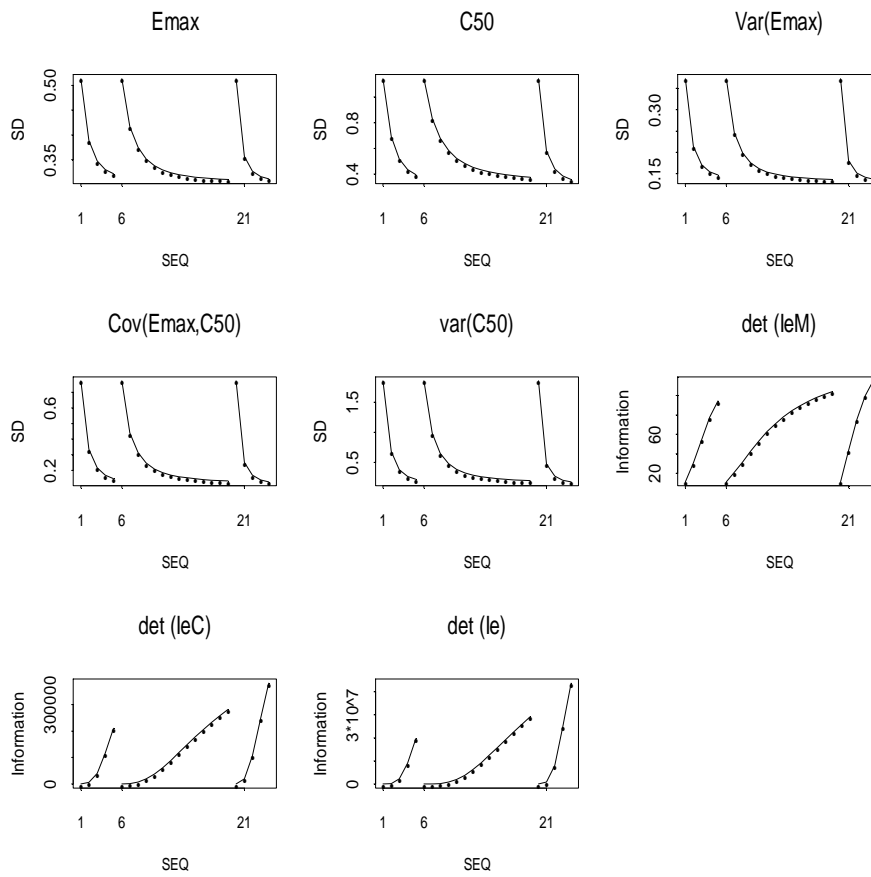
Figure 5.7: Diagonal elements and determinant of the information matrix
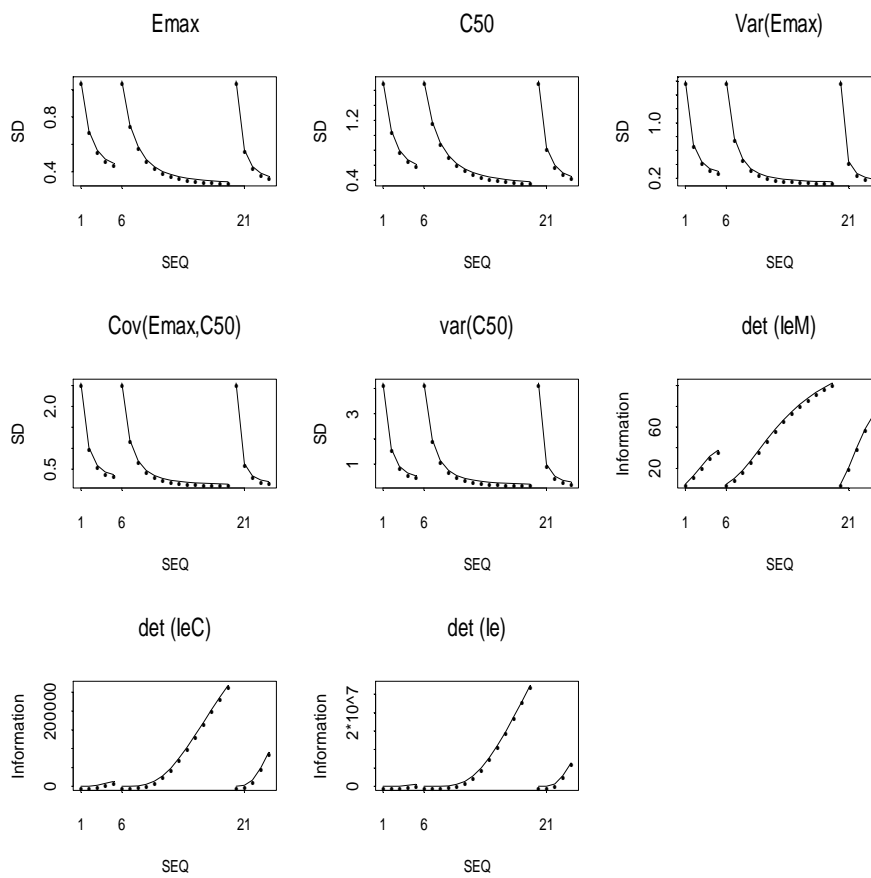
Figure 5.8: Diagonal elements and determinant of the information matrix for a hypothetical situation where C50=10 ng/ml.

Chapter 6

Modeling the association between adherence and viral load in HIV-infected patients

The primary objective of this paper is to investigate the effect of adherence to prescribed therapy on virologic response measured repeatedly over time in HIV patients. To this end observations on viral load assessed in cp/ml are categorized into 4 clinically meaningful states, $[0-50[$, $[50-400[$, $[400-2000[$, $[2000$ and up. A time-dependent continuation ratio model is used to analyze longitudinal individual ordinal responses. The main challenge lies in modeling dependencies over time and using the information in the data efficiently to establish a dynamic relation between observed adherence and viral load. Among the several measures of adherence investigated, two account specifically for long periods of time without intake. One is derived from the third moment of the inter-dose interval while the second reflects internal drug exposure using pharmacokinetic features. The approach is applied to a clinical trial involving 35 patients who were followed over 12 months. The data show a significant relation between patient adherence and virologic response.

6.1 Introduction

The last years, we experienced a revolution in HIV treatment. New tests that can accurately measure levels of HIV in blood became commercially available. Resulting data showed plasma viral load to be the strongest predictor of the risk of progression to AIDS and death (Mellors et al. 1996, Hughes et al. 1997, Marchner et al. 1998, Delta Coordinating Committee 1999). New drugs, including potent protease inhibitors (PI's) also became available. The new treatments, however, are not a cure for HIV disease. While many patients are benefiting from treatment, others are not or have experienced only a temporary benefit. There are several reasons why treatment fails. Poor patient adherence to therapy is the most cited (Montaner et al. 1998, Paterson et al. 2000), and indeed a probable cause for many of the treatment failures. The study of why and how HIV antiretroviral treatment fails is important for the development of new treatment strategies for those patients for whom therapy is currently failing.

The statistical methods presented in this paper were motivated by clinical studies reporting measurements of human immunodeficiency virus (HIV) RNA in plasma from HIV-infected subjects. Typically, technical measurements in the assay lead to measurements that are either left or right censored. For example, the assay can have a quantifiable range from 50 to 75,0000 copies of RNA per milliliter of plasma (cp/ml). Outside this range, quantification is unreliable or the virus may not be detectable. Review of recent reports from HIV clinical trials shows that a common analytic approach is to replace left censored values by imputed values, for example replacing values < 50 cp/ml by 49 cp/ml. Modeling then proceeds as if these are the actual values. The statistical literature warns against this kind of imputation, and methods acknowledging censored observations have been recommended (Hughes 2000). Furthermore, measurement error and/or large within-patient variations, especially at higher viral load levels, have to be taken into account when interpreting the results from traditional methods. While a continuous scale is often appealing, in the field of HIV there are clinically meaningful cut points in viral load reduction that are used in practice. The goal is to eliminate the virus but in practice 50 cp/ml is the detection limit. A viral load < 50 is considered a successful response with the current assessment technology (ultra-sensitive assay). However an outcome of 49 versus 1 makes a big difference for the time the virus needs to replicate to a substantial infection level and hence for the protective ability

of the drug and its forgiveness for treatment lapses. Until early 1999, the lower quantifiable limit was < 400 cp/ml and, therefore, this level remains meaningful to clinicians. Finally a viral load between 400 and 2,000 cp/ml is traditionally considered as a partial response to the therapy. Therefore, an average decrease of 500 cp/ml will have a different medical interpretation if it is a reduction from $20,000$ cp/ml than if it is from $1,000$ cp/ml. This paper intends to use a clinically meaningfully categorized measure of viral load and provides a model for analyzing the repeated ordinal responses. The model allows one to estimate the effect of explanatory variables and especially patient adherence not only on success rate (viral load < 50), but also on the rate at which patients progress through the infection.

The proposed model helps to assess how drug exposure adherence patterns, thru viral load levels, explain treatment failure and thus provides a rationale and a focus for an adherence intervention strategy.

6.2   Study design and data

A phase II, randomized, open-label, multi-center, multi-country study was conducted to assess the safety, tolerability, pharmacokinetics and anti-viral activity of high dose lopinavir (one dose a day regimen{QD) and standard dose lopinavir (two doses a day regimen{BID) with additional ritonavir in thirty eight protease inhibitor and non-nucleoside reverse transcriptase inhibitor-experienced HIV-infected subjects. For three patients, adherence data were not available, resulting in 35 evaluable subjects. Eighteen were randomized to QD (5 females) and 17 to BID (4 females). The baseline visit included a discussion of the importance of adherence to the study drug administration schedule and details of the lopinavir regimen to which the subject has been randomized. Subjects were then instructed to begin their dispensed regimen the following day. Adherence to the subject's assigned regimen was discussed during each following study visit. Subjects returned to the clinic for viral load assessment at week 4, 8, 12, 16, 20, 24, 32, 40 and 48. Blood samples were collected for measurement of lopinavir and ritonavir levels at day 21 for intensive PK measurements and then at weeks 8, 16, 24, 48 for trough PK, where the blood sampling is assumed to be taken before the next scheduled dose ingestion. At each of those visits the timing of the last dose intake (or two last doses for BID regimen) was recorded on a diary by the patient and reported to the study monitor.

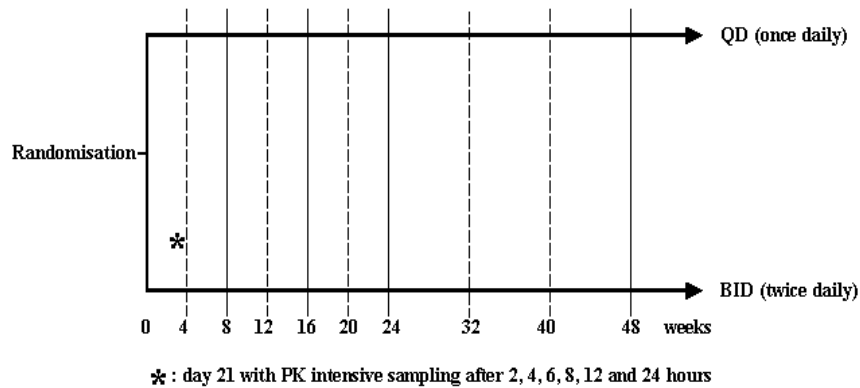* : day 21 with PK intensive sampling after 2, 4, 6, 8, 12 and 24 hours

Figure 6.1: Schematic illustration of the study design. Dotted vertical lines represent visit dates with viral load assessment. Solid lines represent visit dates with viral load assessment and PK sampling. The star indicates the intensive PK sampling on day 21.

For the \intensive PK" at day 21, a first blood sample was collected at arrival to the center (0-hour), immediately prior to the morning dosing. Then a standardized breakfast was served with the morning dose and finally blood samples were successively collected after 2, 4, 6, 8, 12 and 24 (only for QD) hours. Study design is illustrated in figure 6.1

6.3 Adherence to the prescribed therapy

Protease inhibitor adherence was assessed using an electronic monitor. Medication Event Monitoring System (MEMS, Aardex Ltd.) record for each patient a series of exact times and dates at which the drug container is opened over the course of the study. As a basic visual description of the data, we have graphed each patient's pattern of drug intake combined with the PK and viral load information as illustrated in figure 6.2 for a single subject. Visits with viral load observation are represented by vertical lines, while the blood sampling times are represented by stars. Actual viral load values (cp/ml) and lopinavir trough concentrations (ng/ml) are shown on the top and lower bottom lines of the plot respectively. The combined data provide a detailed source of information on
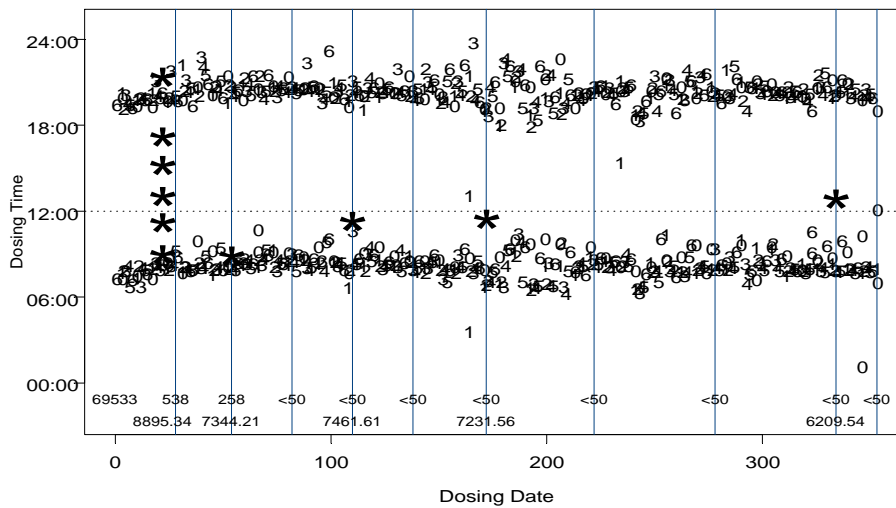
Figure 6.2: Individual patient adherence. The horizontal axis displays the dosing days relative to study entry. The vertical axis gives the time of drug intake on a 24-hour clock. The digits from 0 to 6 are used as plotting symbols to characterize the day of the week (0=Sunday, 1=Monday, ..., 6=Saturday). Vertical lines represent the viral assessment visit. The stars represent the PK sampling day and time. The bottom line of numbers the blood concentration (ng/mL) of ABT-378/r at trough and the line above it shows viral load (cp/ml) respectively.

drug-taking behavior, pharmacokinetics and pharmacodynamics. Figure 6.2 shows the viral load declining systematically over time until it becomes undetectable on day 82 after which it stays stable.

Several adherence summary statistics have already been proposed (Vrijens and Goetghebeur 1997) and are widely used in the literature. In the sequel, we investigate the percentage of prescribed doses taken (taking compliance), the percentage of prescribed dosing days with the correct number of doses taken (correct dosing) and the percentage of prescribed dosing days with dosing intervals within the prescribed inter dose interval $[\delta_0$ - 3 hours;δ_0 + 3 hours] (timing compliance), where δ_0 is the prescribed dosing interval. For example, $\delta_0 = 12$ hours for a BID medication. In this study we found it useful to introduce a new adherence summary statistic derived from the observed dosing intervals. The measure intends to penalize more severely for larger dosing intervals or in other words longer drug holidays. If we denote by δ_{ik} the kth observed dosing interval for the ith patient, then δ_{ik}^* is the corresponding standardized dosing interval defined as $\delta_{ik}^* = \frac{\delta_{ik}-\delta_0}{\delta_0}$. Thus for a perfect adherer, $\delta_{ik}^* = 0$ for all k. While the mean standardized dosing interval contains information pertaining to dosing interval discrepancies, higher moments of the within-patient distribution of δ_{ik}^*, like variance and particularly skewness, are relevant. The variance can be interpreted as a measure of the irregularity in drug intake over time while skewness to the right points to longer drug holidays. We considered for a given patient i,

$$CD_i = (\frac{1}{n_i} \sum_i (\delta_{ik}^*)^3)^{1=3};\tag{6.1}$$

the cubic distance between the observed and prescribed rather than average standardized dosing intervals. Compared to the Euclidean distance it lets dosing intervals which are too short be compensated by dosing intervals which are too long.

6.4   Internal exposure

Since we focus on long term effects of adherence on viral load (monthly), we choose to ignore short term kinetic detail and use a single compartment model rather relying on more complex but uncertain modeling. We model the macroscopic kinetic behavior of the drug by a one-compartment model with first-order absorption and first-order elimination. We consider then the volume of the central compartment (V), elim-

ination rate constant (k_e) and absorption rate constant (k_a) as model parameters. For parameter estimation we used all available lopinavir concentration data, hence from the intensive blood sampling day as well as from the trough concentration sampling. We use the conditional first-order method developed by J. Pinheiro and D. Bates as implemented in NLME v3 (S-PLUS 2000) to fit the nonlinear compartmental model. By including all available data over time, estimation is more representative of long term drug activity under real circumstances. We assumed log normal models for the between-patient variability in k_e and V while k_a was treated as fixed across the population due to limited observations during the absorption phase. When modeling steady state concentrations based on just the timing of the last dose (QD) or the two last doses (BID) as model input, the numerical procedure to maximize the likelihood does not converge. This non-convergence is likely due to a conflict between the steady state assumption and the true data generating mechanism. The problem is solved by using the detailed records of the dosing history prior to the blood sampling. For repeated dosing, the patient's estimated concentration at time t, labelled the internal exposure IE_t in the sequel, can be recursively estimated by solving the set of differential equations associated with a one compartmental model (Gibaldi and Perrier 1982). Estimates of individual pharmacokinetic parameters (ka_i, ke_i and V_i) were derived using the expected a posteori value of the random effects and are then combined with observed dosing histories to simulate IE_t profiles over time. Such profile is illustrated for a single patient in figure 6.3. To summarize internal exposure over an observed time period, we used the median value, the total time and percentage of time that the IE_t falls below the EC50{the average concentration at which half of the maximum effect is reached.

6.5 Modeling the viral load

To interpret a patient's viral load evolution, we found it useful to stage the patient clinically. Based on popular threshold values in the HIV field, we stage it as follows: $[0-50[; [50-400[; [400-2000[$ and $[2000-1[$. The ordinal response y_{it} thus indicates one of four HIV states for patient i at time t. A natural model for these events compares the probability of belonging to each given category with the probability of belonging to any of the lower categories. Thus, as graphed in figure 6.4 the three last categories are compared to the first one, the last two to
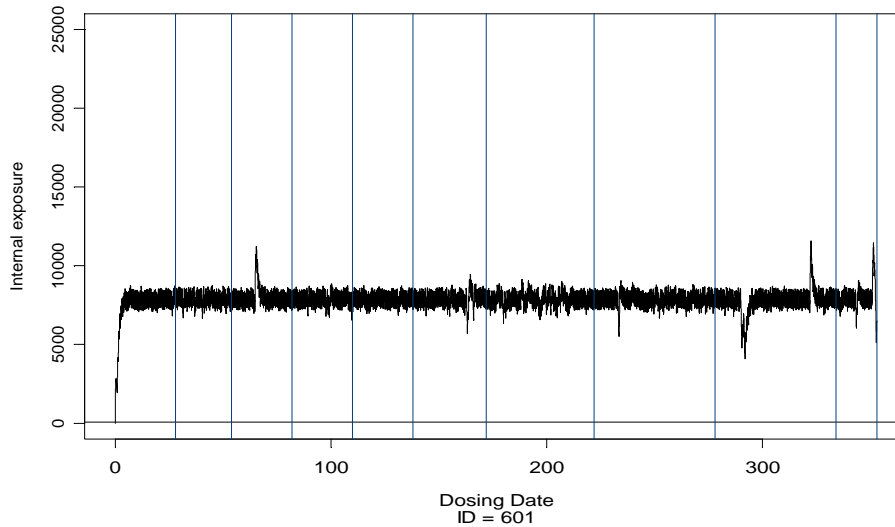
Figure 6.3: Individual patient internal exposure. The vertical lines represent the visit dates when viral loads were assessed

the second, and the third to the last. We thus consider the probability of improvement conditional on the present position. For an ordinal response with 4 categories, there are now 3 comparisons ($c = 1; 2; 3$) to make. Alternatively, one can start from the bottom of the scale, and describe the probability of deteriorating one step further on the scale, given the present position (graphically presented on the right hand side of figure 6.4).

The model describes the probability of moving further on the scale, given the present position. The continuation ratio model has the advantage of being a simple decomposition of a multinomial distribution, with parameters that can be estimated by applying a standard logistic regression to the 3 reconstructed subtables making the comparisons described above (Agresti 1990). Technically, to fit the separate logistic models, we denote by $y_{c=it}$ the transformed binary response corresponding to the cth comparison involving the tth event for the ith patient. Specific binary coding of $y_{c=it}$ is given in tables 6.1 and 6.2 for the two modeling strategies.

With this reparametrization, the original multinomial distribution becomes a product of three binomial distributions and the model can be written:
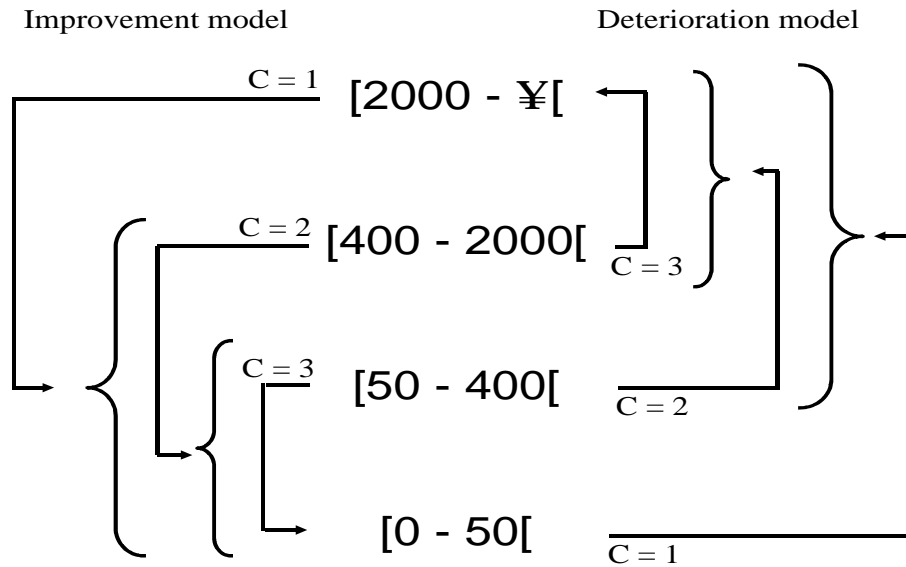
Figure 6.4: Schematic representation of the probabilities in a continuous ratio model. Two models, one for improvement and one for deterioration are presented separately. For both models the three comparisons are labelled c = 1; 2; 3.

Table 6.1: $y_{c=it}$ coding for the improvement model. The coding is undefined in the empty cells

c	$[0-50[$	$[50-400[$	$[400-2000[$	$[2000-1[$
			y_{it}	
1	1	1	1	0
2	1	1	0	
3	1	0		

Table 6.2: $y_{c=it}$ **coding for the deterioration model. The coding is undefined in the empty cells**

c	$[0 - 50[$	$[50 - 400[$	y_{it} $[400 - 2000[$	$[2000 - 1[$
1	0	1	1	1
2		0	1	1
3			0	1

$$\eta(p_{c=it}) = C_{it}^0\alpha + X_{it}^0\beta + Y_{it-1}\gamma + \delta Z_i \qquad (6.2)$$

where $p_{c=it} = P(Y_{c=it} = 1 \mid C_{it}; X_{it}; Y_{it-1}; Z_i)$, $\eta(:)$ is the logit link function, α, β, γ, δ are a set of fixed parameters to be estimated, C_{it} represents the dummy indicators for the comparison subtables, X_{it} potential explanatory variables, and the two last terms account for dependencies among individuals. The model allows one to estimate the effect of explanatory variables, not only on the rate of success (viral load < 50), but also on the rate at which patients progress through the viral load categories. By studying interactions between explanatory variables and the comparison subtables we can test whether their effect is identical across the categories.

For each patient, we have a sequence of ordinal responses over time. When estimating model parameters we must consider the dependence among successive responses of an individual. Two types of dependence are expected: those arising from heterogeneity among individuals, often called frailty, and those from serial correlation over time within an individual. A possible approach to handle frailty is by incorporating individual-specific normal random effects into an event-specific binary logistic model (Dos Santos and Berridge 2000). To this end we assume that on the scale of the logit link function η, each subject's probability of response is shifted according to the subject-specific value Z_i, drawn from a Standard Normal random variable with unit variance. δ is then the standard deviation of the patient random effect. Serial correlation can be modeled by fitting Markov chains to the transformed binary data. In such model, the present response at any time point is made conditional on that in the previous time period. Conditioning can happen on the lagged value of the binary response ($Y_{c=it-1}$), but interpretation is easier if the lagged ordinal response (Y_{it-1}) variable is used directly

Table 6.3: Number of observations in each state

State	Nb observations
1 : $[0 - 50[$	195
2 : $[50 - 400[$	81
3 : $[400 - 2000[$	30
4 : $[2000 - 1[$	46

as an explanatory factor variable (Lindsey et al. 1997). The additional effect of covariates and especially patient adherence can then be tested through the $X_{it}^0\beta$ term in the model.

6.6   Results

Table 6.3 gives the number of available observations that were collected among the 35 patients who completed the 12 months observation period.

To detect dependence among repeated responses we first fit reference models assuming independence. The first column of table 6.4 shows results of fitting such models which completely ignore the fact that several responses have been observed on each patient. To handle frailty, a patient-specific random effect has been incorporated into the model whose results are shown in column 2. We used Gaussian quadrature to numerically maximize the likelihood. On the other hand serial correlation is allowed by letting the response on the previous visit be an explanatory variable for the next. The corresponding deviances are given in the third column of the table. The last column gives results when frailty and serial correlation are jointly incorporated.

Table 6.4 reveals that serial correlation is the main source of stochastic dependence among successive responses for the same individual. Heterogeneity among individuals accounts only for a very small proportion of the dependence. Our serial correlation model does the job of taking dependence into account and it is not necessary to take the additional heterogeneity among individual into account. Furthermore since we had a baseline viral load observation, we do not loose the first observation by introducing such lagged variables as explanatory variables. In this particular situation, the model can be fitted by standard logistic regression after a simple restructuring of the data. This maneuver will become

Table 6.4: Deviance table

| Model deviance	Independent	Frailty	Serial correlation	Frailty and serial correlation
		Improvement model		
Null model	581.5	554.1	404.5	400.2
Correct dosing	579.3	553.7	402.8	398.9
Taking compliance	577.2	552.1	396.6	392.7
Timing compliance	581.5	553.9	403.8	399.7
Cubic distance	581.5	573.9	388.9	384.7
Median	572.1	548.9	395.7	395.8
% time(IE < EC50)	575.7	548.8	392.1	387.4
time(IE < EC50)	576.0	550.1	389.2	386.0
		Deterioration model		
Null model	581.5	559.1	418.1	413.5
Correct dosing	578.4	558.6	411.4	408.5
Taking compliance	574.8	555.8	401.5	398.5
Timing compliance	581.1	559.0	414.9	411.3
Cubic distance	571.9	554.2	389.1	388.3
Median	571.1	559.3	402.0	403.4
% time(IE < EC50)	573.3	553.5	392.9	391.1
time(IE < EC50)	574.0	554.7	390.6	389.7
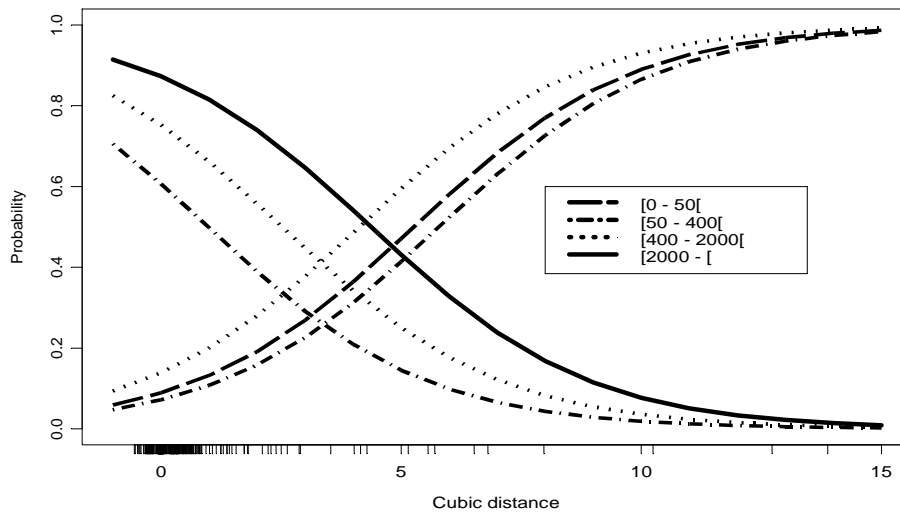
Figure 6.5: model based probabilities of improvement (decreasing curves) and deterioration (increasing curves)

especially convenient when more data will become available and larger database will be under investigation.

Two adherence explanatory variables perform significantly better than the other ones, the cubic distance and the amount of time during which the internal exposure of a patient falls below the EC50. Both measures are strongly influenced by drug holidays{one or more days without drug intake. The amount of time during which the internal exposure falls below the EC50 has the advantage of a nice pharmacologic interpretation but suffers from the practical necessity of collecting blood concentration measures. The cubic distance establishes a direct relation between the dosing histories and the viral load observation, but suffers from a more difficult interpretation. The model fitted conditional probabilities of improvement and deterioration are graphically presented in figure 6.5.

The decreasing curves represent the probabilities of improvement in function of the cubic distance, while the increasing curves are the probabilities of deterioration. The model fitted is conditional on past assessment of the viral load staged in four categories. This condition is expressed by the type of the lines that indicate the virologic state in which the patient was at the previous visit. When the cubic distance

Table 6.5: Probabilities estimated directly from the continuous ratio models

Improvement model	Deterioration model
$P(< 2000)$	$P(\geq 50)$
$P(< 400 \mid < 2000)$	$P(\geq 400 \mid \geq 50)$
$P(< 50 \mid < 400)$	$P(\geq 2000 \mid \geq 400)$

increases, the probability of improvement drops substantially quickly while the probability of deterioration increases. The relation between the time that the internal exposure remains below the EC50 delivers a very similar figure and is not reported here.

The model parameters are fitted through a decomposition of the multinomial distribution, therefore the estimated probabilities are not only conditional on past viral load history and adherence but also on the subtables used for the construction of the model. Conditional on past viral load history and adherence, the probabilities directly estimated by the continuous ratio models and plotted in figure 6.5 are given in table 6.5. They can be misleading for direct medical interpretation. Denoting $a_1 = 2000$, $a_2 = 400$, and $a_3 = 50$ allows us to derive the marginal multinomial probabilities using the recursive equations 6.3:

$$\left(\begin{array}{ll} P(< a_i) = P(< a_i \mid < a_{i-1})P(< a_{i-1}); & \text{for } i = 2, 3 \\ P(\geq a_i) = P(\geq a_i \mid \geq a_{i+1})P(\geq a_{i+1}); & \text{for } i = 1, 2 \end{array} \right. \tag{6.3}$$

Those can then be used to estimate patient specific marginal chances of improvement or deterioration conditional on his past viral load history and adherence patterns.

6.7 Discussion

As we emphasized in the introduction, we wanted to investigate the effect of patient adherence to therapy on an individual-specific viral load response. We found it appropriate to categorize the viral load in a clinically meaningful way. The categorization proposed is not unique and other cut points could be investigated. The essential feature of this structure is that `improvement' or `deterioration' is judged only when there has been a substantial change in viral load, thus lifting the

discussion out of the range of error in the test for viral load, and out of the range of changes too small to have any evident clinical meaning.

The model proposed can deal with repeated ordinal data and especially with uncommon measurement times and unequal numbers of repeated measures for each patient. The two main forms of stochastic dependence among viral load on a individual (frailty and serial correlation) were studied. Acknowledging for serial correlation takes this dependence into account.

A one compartmental pharmacokinetic model adjusted for observed dosing histories, fitted the sparse data on lopinavir concentration well. Further, viral load was significantly related to plasma concentration of the drug. However, the most significant PK-derived predictor, the time that the IE falls below the EC50, appeared not very sensitive to the between-patient variability in pharmacokinetics. It is largely driven by longer non-dosing periods (drug holidays). This explains why for this particular drug, the patient-specific pharmacokinetics could usefully be replaced by a common summary of patient specific adherence patterns, the cubical distance.

While the internal exposure approach allows for a direct interpretation of the adherence variable, the cubical distance has its own advantages for comparing the impact of adherence on viral load unconfounded by the kinetic properties of the drugs studied. This will especially be useful to assess the forgiveness of the drugs for lapses in dosing. Rewriting equation 6.1 as $(\frac{1}{n_i} \sum_i \text{sign}(\delta^*_{ik}) j \delta^*_{ik} j^F)^{1=F}$, one can define different orders of central moment and see $F > 0$ as a measure of the degree of forgiveness. Higher values for F will identify more forgiveness drugs. As illustrated in figure 6.6, the described approach allows one to assess not only the drug specific degree of forgiveness for improvement but also for deterioration. We conclude that lopinavir is more forgiving for deterioration than in the improvement phase. When more data will become available on different drugs, we shall be able to compare the degree of forgiveness between the compounds.

While the main objective of the study was to compare the standard BID regimen with the high dose QD regimen, we used the data to derive a relation between dosing histories and viral loads. However after adjusting for the observed dosing histories, no relevant difference in improvement and deterioration probabilities were found between the two regimens.

Several extensions of the model could be investigated when appropriate data will become available. For example, a trend over time in response
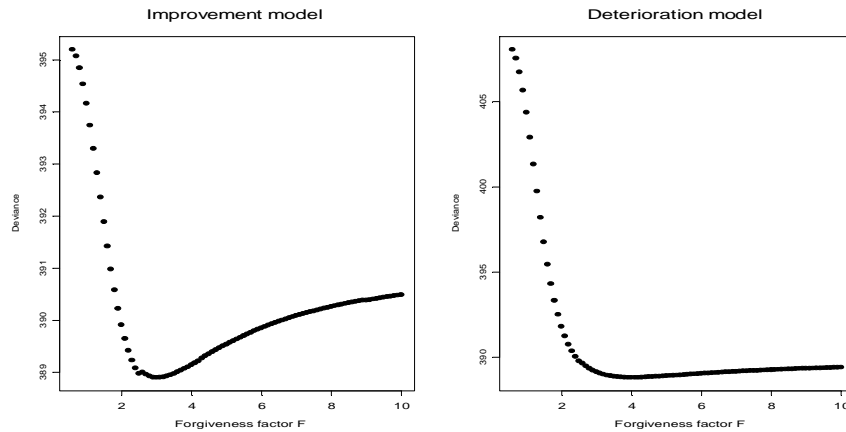
Figure 6.6: Graphical presentation of the deviance in function of the `degree of forgiveness' F

or interaction between subtable and the adherence variables might be expected but were not significant at the 5 percent level.

The study at hand was part of an adherence programme and thus patient adherence may not be representative of the general HIV population. Regular takers experience a relatively small range of concentration which makes it hard to estimate deviation from normality in the viral load model. Estimators of the proposed model parameters could benefit greatly from information that enters through greater variation in the drug exposure process (Vrijens and Goetghebeur 2001). This could be collected on less good adherers as observed in usual practice.

The model proposed in this paper is not only useful to assess how viral load predictions can help to explain the reasons for treatment failure but it can also guide the practitioner in how best to focus an adherence intervention strategy. Anti-HIV medicine has to be taken for life, so this approach to medication management can bring long-term value.

Finally, this research suggest that monitoring a patient for his adherence could allow for early detection for patients with extremely hazardous adherence patterns. Consequently, through monitoring of dosing histories, and perhaps the study of patient pharmacokinetics, the physician could be able to individualize therapy for the patient, with perhaps lower doses and less severe side effects of treatment. Thereby encouraging long persistence with treatment.

Chapter 7

A structural model to compare two active treatments

A randomized clinical trial is designed to answer a relevant clinical question. The more simple a trial's design and primary data analysis, the more likely the answer is robust, enabling results to achieve widespread understanding and acceptance. The simplicity is lost when some subjects fail to adhere to the prescribed therapy. Claxton et al. (2001) show that in long term therapy, patient adherence to prescribed drug therapy is far from optimal. Ignoring this fact leads to problems. Long term clinical trials can then give rise to an underestimation of the true biological treatment effect which in today's circumstances of pressing public health need for effective medicines is unacceptable. Hence, non-adherence must be addressed and has the drawback to open large and well designed trials to criticism. On the other hand it gives the opportunity to study dose-response relationships, which would not be visible if adherence had been perfect. The latter is a very appealing feature but unfortunately adherence is an uncontrolled covariate and can be \select"-tive". For example as illustrated in Efron and Feldman (1991), patients with better adherence in the placebo group can have better outcomes on average. Different adherence patterns can \select" different outcomes in the following ways:

- **Feedback selectivity is defined as the effect of disease progression and/or treatment side effects on the patients adherence. In other**

words, changes in patient adherence during the course of treatment can depend on the disease evolution and/or appearance of side effects during that period. For example, patients can stop the drug because they feel better or worse. This selectivity will be less important for asymptomatic diseases and for drugs free of side effect.

- **Prognostic selectivity** takes place when a baseline prognostic factor confounds the relation between dose and response. This happens when the prognostic factor is associated with adherence. Presence of non trial medicines in secondary prevention studies is a typical example. Then the measure of adherence with placebo is a surrogate measure of adherence with non trial medicines. The latter being a potential prognostic factor for outcome, it could confound the adherence-effect relationship. We note that prognostic factor can be unknown and difficult to formalize. For example some psychologic characteristic of an individual could induce some psychologic effect on outcome and be correlated with adherence at the same time.

- **Interacting selectivity**
  Potential interactions between the two above effects. Also time-dependent confounders may occur.

Until now, in this thesis, we assumed that adherence was not selective and derived dose-response associations from empirical data. Results from such investigations can only be interpreted as causal when the above mentioned selection mechanisms are not present. In chapters 4 and 5, when simulating PK/PD data on top of the adherence patterns, we generated nonselectivity and this was thus not a concern at the analysis stage. In chapter 6, where we derived an association between exposure and viral loads. To give a causal interpretation to this association, we have to assume nonselectivity. Considerations listed below supports the idea that in this particular field, the observed dose-response relationship, would not be substantially confounded by any selective mechanism of adherence. 1) It is an asymptomatic disease and thus no feedback from outcome was anticipated. 2) No psychologic effect was expected on viral load measures. 3) No association between PK parameters and adherence was found. 4) side effects could influence adherence. For example, the appearance of side effects could be correlated with health status of the patient. However, PI treatment achieve

blood concentrations far above the EC50, and thus their biological effect is very strong. It is expected to outweight any effect the forementioned selectivity might have, making the latter negligible for our analysis.

Moving to the field of depression, we observe a strong persistence problem and the assumption of nonselectivity is unlikely to hold. Depression is no longer an asymptomatic disease and patients can stop the treatment because they might perceive that treatment is not working or the opposite, they feel better. Further, in this field a strong placebo effect is expected and a psychological component of adherence probably applies, better adherers are expected to be better patients to begin with. Those are several reasons to seek proper interpretation of the dose-effect relationships. We must acknowledge that nonadherence is a post randomization variable, imparting to a controlled clinical trial some of the character of an observational study, where the subjects are no longer randomly assigned to the observed exposure levels.

7.1   Introduction

Most clinical trials aim to show superiority of an active molecule compared to placebo. However, the World Medical Assembly recently decreed it is no longer ethical to use placebos in any trial situation where an effective treatment exists. One option is to make ever more subtle distinctions in medical conditions, so as to define progressively narrower areas for which a proven therapy does not exist. The problem with that approach, aside from its transparency, is that those who invest in clinical trials strive for broader markets. A leading alternative to placebo control is the active control trials, in which the reference level is an agent of proven efficacy. The chief difficulty with that is that results of a badly executed trial can create the illusion of effectiveness. This prospect leads to the concept of a trial's assay sensitivity (Urquhart 2001) where compliance with prescribed therapy is recognized as the leading factor that can compromise the credibility of an active control trial. Imagine a placebo controlled trial in which no participant, for one reason or another, took any of either the active or the placebo. The result would be failure to reject the null hypothesis, and the conclusion that the new medicine doesn't work. In a active control trial in which the same thing happened, the result would be equivalence of test agent and proven agent, and the conclusion that the new medicine works.

While this example is extreme, it illustrates the point that one should adjust analyzes of randomized clinical trials for the fact that not all patients comply fully with the prescribed treatment. Today's circumstances of pressing public health need for effective medicines make critical the avoidance of type II errors. Especially in active control trials, beside recommended trial sensitivity, it is therefore important to give insight into the actual treatment efficacy and effect variation across different compliance levels.

Until now, most authors propose causal inference tools for comparing an active treatment arm with a control arm receiving nothing (Sommer and Zeger 1991) or placebo (Efron and Feldman 1991, Goetghebeur and Lapp 1997). A comparison of active treatment with placebo has the advantage that we actually observe the treatment-free response on a random subset of individuals in the trial. That helps to identify the potential treatment-free response on patients of different compliance levels on active arm and estimate the treatment effect for given compliance level.

Although the general ideas of Robins (1994) do not require existence

of an inactive arm, we have seen few implementations of these ideas in comparing two active treatments. In this chapter, we consider a clinical trial comparing two different active molecules in the treatment of depression. In section 2, we start by introducing the clinical example. In section 3, we define an appropriate structural model for this setting, then discuss identifiability issues and consider the optimal estimator. Results are presented in section 4 and further discussed in section 5.

7.2   The anti-depressant trial

We illustrate the methodology on data issued from the clinical trial described in chapter 3 of this thesis. The trial was designed to study adherence profiles of ambulatory patients suffering from major depressive disorders treated with selective serotonin reuptake inhibitors (SSRI). After a one week run-in period, 76 patients were randomized to fluoxetine or paroxetine once a day. Adherence to prescribed therapy and Hamilton depression scores were the primary outcomes variables. While an intent to treat analysis comparing responses between both randomized groups constitutes the primary trial analysis, the relation between actual exposure and Hamilton scores is of prime interest. For the current analysis we focus on data collected during the 6 weeks acute period that started just after randomization. During this period one can study the dose-effect relationship of both treatments uncounfounded by missing final Hamilton scores in the maintenance period due to patient loss of follow-up.

7.3   A Structural model for 2 active treatments

We consider an active control trial, where patients i are randomized to one of 2 active treatments, denoted A or B in the sequel. Randomization indicators are defined as : R_i^A, having value 1 when patient i is randomized to treatment A and 0 otherwise. Then $R_i^B = 1 - R_i^A$, with similar interpretation. Formally, we can consider for each subject i both responses Y_i^A and Y_i^B following the possible assignment to treatment A or B respectively. Clearly for each patient i only one of both is observed, the other being the potential response if the patient would have been randomized to the other treatment group. In addition, we denote by Y_i^0 the potential treatment-free response that is not supposed to be

observed in an active control trial. C_i^A and C_i^B are the (potential) univariate adherence summary measures following assignment to A or B. Finally, we denote by X_i the vector of baseline covariates

A simple structural model assumes a linear effect of dose (adherence) on the expected outcome on each arm. Subtracting this effect from the outcome should result on average in the expected treatment-free response, conditional on compliance and baseline characteristics. Hence we assume :

$$E(Y_i^A - \beta^A C_i^A | X_i; C_i^A) = E(Y_i^0 | X_i; C_i^A)$$
$$\text{and } E(Y_i^B - \beta^B C_i^B | X_i; C_i^B) = E(Y_i^0 | X_i; C_i^B): \qquad (7.1)$$

Note that this model does not parameterize the distribution of C_i^A and C_i^B nor $E(Y_i^0 | X_i; C_i^A)$ and $E(Y_i^0 | X_i; C_i^B)$.

Estimating equations can be derived from the implied equality:

$$E(Y_i^A - \beta^A C_i^A | X_i) = E(Y_i^B - \beta^B C_i^B | X_i):$$

Clearly, the two distinct parameters β^A and β^B cannot be identified, when $E(C_i^A | X_i) = E(C_i^B | X_i)$; but one can still identify their difference: $\beta^A - \beta^B$ in that case. In this chapter, we develop the case where $E(C_i^A | X_i) \neq E(C_i^B | X_i)$ and thus β^A and β^B are separately estimable.

We denote β the vector containing the causal parameters $(\beta^A; \beta^B)$ and β^0 is its assumed true value. Unbiased estimating equations are then obtained from (7.1) as

$$\sum_{i=1}^{X} g(R_i^A; R_i^B; X_i)[H_i(\beta) - q(X_i)] = 0;$$

where

$$H_i(\beta) = Y_i - R_i^A \beta^A C_i^A - R_i^B \beta^B C_i^B \qquad (7.2)$$

and $E[g(R_i^A; R_i^B; X_i) | X_i] = 0$. The function q may be any function of X_i. Semi-parametric efficiency is obtained with optimal choices of g and q. $g(R_i^A; R_i^B; X_i)$ is a vector of dimension 2 and the corresponding matrix is assumed to be of full rank.

Following Robins (1994), the optimal function q_{opt} is

$$q_{opt}(X_i) \equiv E[H_i(\beta_0) | X_i]:$$

Since $E[H_i(\beta_0)|X_i] = E[Y_i^0|X_i]$, the expected treatment-free response, this is not directly estimable from the data from a trial comparing two active treatments. A different choice of $q(X_i)$ does not create a systematic biases, only the precision of resulting estimates suffers from a \bad" choice of q. One could approximate q_{opt} by the \best guess" for expected treatment-free response (e.g. baseline value of the outcome variable, predictions based on time trends observed in the run-in period). Alternatively, one could start with an initial choice of $q = q_0$ and obtain parameter estimates $\hat{\beta}^{(0)}(q_0)$ first, to subsequently estimate q_{opt} as $E[H_i(\hat{\beta}^{(0)}(q_0))|X_i]$.

Given q, the optimal function g_{opt} can be found as follows. Let

$$\varepsilon_i = \varepsilon_i(\beta_0) = H_i(\beta_0) - q(X_i):$$

Now define

$$w_{opt;i} = w_{opt}(R_i^A; R_i^B; X_i) = \{E[\varepsilon_i^2|R_i^A; R_i^B; X_i]\}^{-1};$$

$$\mu_{opt;i} = \mu_{opt}(R_i^A; R_i^B; X_i) = E[\partial H_i(\beta_0)/\partial\beta_0|R_i^A; R_i^B; X_i];$$

$$g_{opt;i} \equiv w_{opt;i}\{\mu_{opt;i} - E[w_{opt;i}|X_i]^{-1}E[w_{opt;i}\mu_{opt;i}|X_i]\}$$

(7.3)

With q having its optimal form q_{opt}, $Var\{n^{1=2}[\hat{\psi}(q_{opt}; g_{opt})]\}$ attains the semi parametric efficiency bound.

The estimates for all above expectations can be estimated by regression methods.

Fischer-Lapp and Goetghebeur (1999) showed that in placebo-controlled trials q_{opt} can be estimated from the placebo group and predictions from the obtained model used also for the treatment group subjects. In this case it was realistic to assume that ε_i had constant variance, given X_i and the randomization indicator R_i. However, this is no longer true when comparing two active treatments, because of a much poorer prediction of q_{opt}.

Suppose G is a matrix with ith row $(g_i^A; g_i^B)$ and C is a matrix with ith row $(R_i^A C_i^A; R_i^B C_i^B)$.

Now the estimated parameter vector will have a matrix representation

$$\hat{\beta} = (G^0 C)^{-1} G^0 (Y - q):$$

(7.4)

Identifiability of the parameters depends on the existence of the inverse of $G'C$. The components of G involve conditional expectations of C_i^A and C_i^B, given X_i. For identifiability it is sufficient that:

$$|Cor[E(C_i^A|X_i); E(C_i^B|X_i)]| < 1:$$

The variance estimator is :

$$(G'C)^{-1}\frac{1}{n}\sum_i f\hat{g}_{opt;i}(H_i - \hat{\phi}_{opt;i})^2 \hat{g}_{opt;i}^0 g(G'C)^{0-1}$$

at $\beta = \hat{\beta}$ and where n is the total number of subjects.

If $Cor[E(C_i^A|X_i); E(C_i^B|X_i)] \approx 1$, $G'C$ will be nearly singular and the parameter estimates will be highly correlated and have a high variance.

7.4   Results

7.4.1   Descriptive statistics

The structural mean model described previously allows only for a univariate summary of adherence. We use the percentage of prescribed dose, actually taken by the patient. However we found it useful to present a more informative graph depicting the evolution of adherence over time. Figure 7.1 presents the cumulative number of doses taken against the days elapsed since randomization. Separate plots are drawn for both randomized groups. Those plots allows us to distinguish between the compliance and persistence components of patient adherence. The horizontal line segments reveal that two patients are non persistent in the fluoxetine group, while 5 are non persistent in the paroxetine group. For those who persist, we observe a larger variability in compliance under fluoxetine than under paroxetine as revealed by more substantial deviation from the diagonal line. During the 6-week acute period, the general patient adherence was relatively high.

Figure 7.2 depicts the evolution of the Hamilton scores over time. Patients were included at visit 1 if their Hamilton score was greater or equal to 21. The period between visit 1 and visit 2 is the placebo run-in period. At visit 2, patients were randomized and started their active treatment. From then on, we observe the strongest decrease in Hamilton scores.
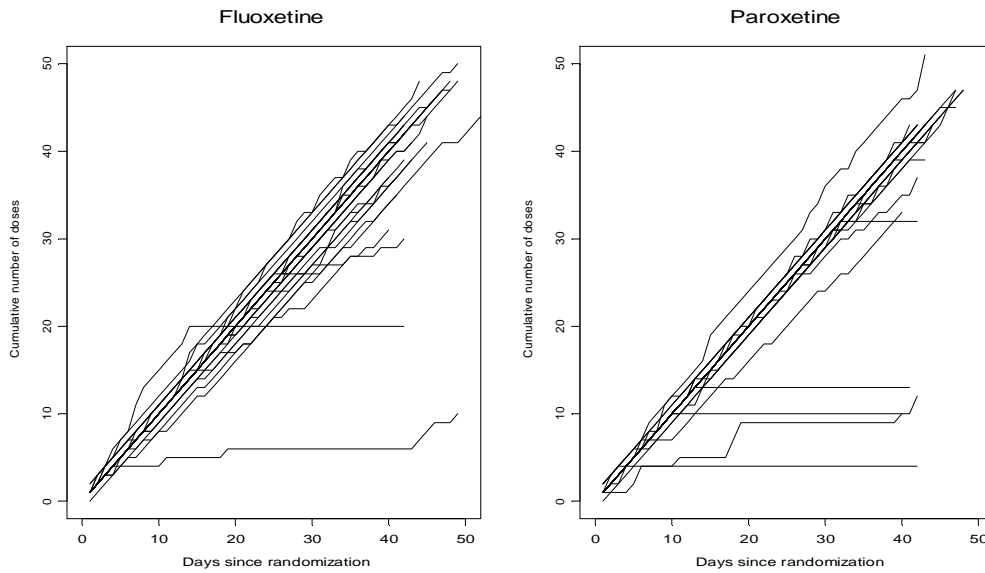
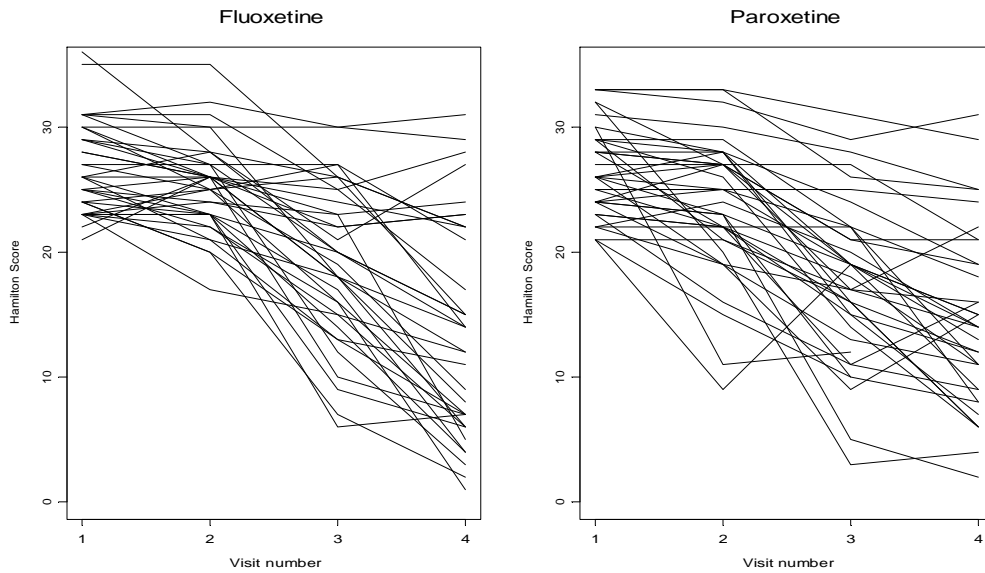**Figure 7.1: Adherence summary plots for both randomized groups.**

Figure 7.2: Evolution of Hamilton scores during the run-in (visit 1 and visit 2) and acute (visit 3 and visit 4) phases in both randomized groups.

Table 7.1: Intent to treat analysis. Comparison of Hamilton Scores and Hamilton score differences from baseline between both randomized groups.

	Hamilton score : visit 4					
	n	**mean**	**sd**	**min**	**max**	**p-value**
Fluoxetine	37	13.7	8.6	1	31	
Paroxetine	39	14.5	7.1	2	31	0.66
	Hamilton score :visit4-visit2					
	n	**mean**	**sd**	**min**	**max**	**p-value**
Fluoxetine	37	-11.6	6.8	-24	1	
Paroxetine	39	-9.8	6.3	-22	4	0.26

7.4.2   Intent to treat analysis

In table 7.1 we compare the Hamilton scores and decrease in Hamilton scores (Hamilton scores at visit 4 - Hamilton scores at visit 2) between both randomized groups. Observed mean (differences) between both randomized groups are not significantly different (t-test) neither are they clinically relevant. We thus conclude that there is no significant difference in Hamilton scores between both treatments.

However, this does not imply that the true biological effect of both treatments is equal. A difference in adherence between both groups could bias this interpretation.

To compare adherence patterns between both randomized groups, we use the methods described in chapter 2. We summarize thus the adherence pattern as a sequence of binary data indicating whether yes (1) or no (0), at least one dose has been taken on each consecutive day. A conditional logistic regression accounting for dependencies within individuals is then implemented by fitting Markov chains. The model allows to estimate the effect of explanatory variables on adherence. As in chapter 2, on top of the Markov dependencies, time (in days) since randomization and weekend were significant contributors to the variability in adherence. In addition two answers collected at baseline through an anti-depressant adherence questionnaire were predictive for adherence.

- **Question 4: My depression is caused by external factors: (1) completely agree; (2) agree; (3) disagree; (4) completely disagree.**

- **Question 27: When I am more depressed, I can take more of the**

Table 7.2: Intent to treat analysis. Comparison of adherence patterns between both randomized groups. Group is coded as follows: 0=fluoxetine, 1=paroxetine.

Variable	Estimated Coefficient	Standard Error	Coefficient / Standard Error
Intercept	-1.68	0.37	-4.53
$Z_{i;t-1}$	1.52	0.21	7.20
$Z_{i;t-2}$	1.30	0.22	5.78
$Z_{i;t-3}$	1.20	0.24	5.06
$Z_{i;t-4}$	0.75	0.26	2.91
$Z_{i;t-5}$	0.94	0.25	3.71
Time	-0.02	0.01	-2.43
Weekend	-0.92	0.23	-4.06
Question 4	-0.12	0.25	-0.49
Question 27	0.44	0.17	2.55
(Treatment) group	-0.03	0.25	-0.13
Question 4:group	-0.73	0.35	-2.10
Weekend:group	0.84	0.34	2.45

prescribed dose: (1) completely agree; (2) agree; (3) disagree; (4) completely disagree.

We found a dichotomous answer to both questions most predictive. Therefore we coded the answers as follows : 1 if completely disagree and 0 otherwise. In table 7.2 we present results from the fitted conditional model. Covariates were included in the model according to a forward selection procedure. We observe a linear decrease in adherence over time. In the fluoxetine group, patients tend to omit doses during the weekends. This effect cancels out in the paroxetine group. This explains probably the larger variability in adherence observed in the fluoxetine group than in the paroxetine one. Those who do not believe they can modify they prescribed dosing regimen and take more drugs when they are more depressed tend to be better adherers. Finally, patients who deny that their depression is caused by external factors tend to be worse adherers. This effect is especially true for patients randomized to the paroxetine group.

We note that the differences in adherence observed between the two randomized groups is suggestive of potential selectivity and will further

allow us to estimate separately β_A and β_B in the causal model.

7.4.3 Structural mean model

As the main outcome, we consider here the change in Hamilton score over the 6-week acute period (Hamilton score at visit 4 - Hamilton score at visit 2).

In the structural model, the set of baseline covariates X_i includes not only predictors for adherence but also predictors for treatment free outcomes. The last ones being difficult to identify in an active control trial. WHO well-being score is a good predictor candidate and we found it useful to include it in the set of baseline covariates. WHO well-being score is obtained by summing the answers (0 - minimal agreement, 5 - maximal agreement) to the two following questions collected at visit 2:

- I feel calm and peaceful

- I am full of energy

Using this set of baseline covariates, we estimate parameters of the structural mean model:

$$E[Y^A - Y^0 | X; C^A] = \delta^A C^A$$

in the first group and of the model:

$$E[Y^B - Y^0 | X; C^B] = \delta^B C^B$$

in the second group, based on estimating equations that are derived from:

$$E[Y^A - \delta^A C^A | X] = E[Y^0 | X] = E[Y^B - \delta^B C^B | X]:$$

The estimates of the parameters of interest, δ_A and δ_B are presented in table 7.3.

Correlation between the two estimated coefficients is estimated to be 0.799.

We see that both parameters are identifiable with relatively small variance. Thus one could conclude a significant biological effect of treatment : a 100% adherer would experience on average a Hamilton Score reduction of 8.9 points on fluoxetine, or 6.9 on paroxetine on top of the

Table 7.3: Causal analysis. Parameter estimates.

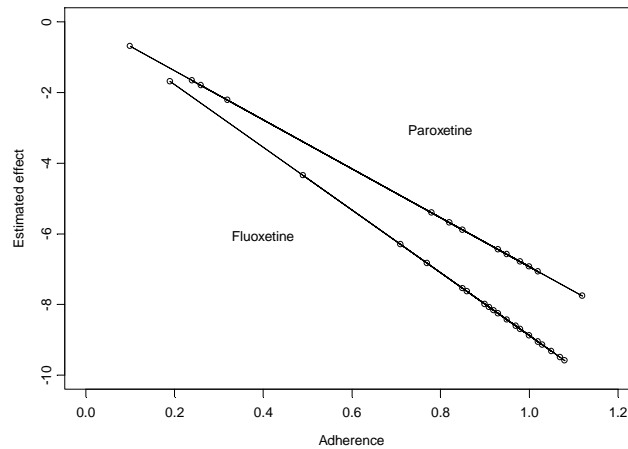| Treatment Group	Estimated Coefficient	Standard Error
fluoxetine	-8.88	2.89
paroxetine	-6.93	2.98

Figure 7.3: Plot of estimated treatment effect against observed adherence levels (Percent of prescribed dose taken)

placebo effect . The absolute difference of the effect on full adherers is estimated to be 1.95 (se=1.86), hence no significant difference can be found between treatment effects at equal adherence levels. The results are graphically presented in figure 7.3.

Since we are able to estimate two separate parameters here, we can also predict potential treatment-free response on each arms, using $Y_i^A - \hat{\beta}^A C_i^A$ and $Y_i^B - \hat{\beta}^B C_i^B$ respectively, and assuming the proposed simple model is correct. Figure 7.4 presents for each treatment groups separately, observed Hamilton scores at visit 2 and visit 4 as well as back transformed treatment free Hamilton scores at visit 4. Here one can clearly depicts the estimated additional effect of the active molecule on top of the placebo effect.
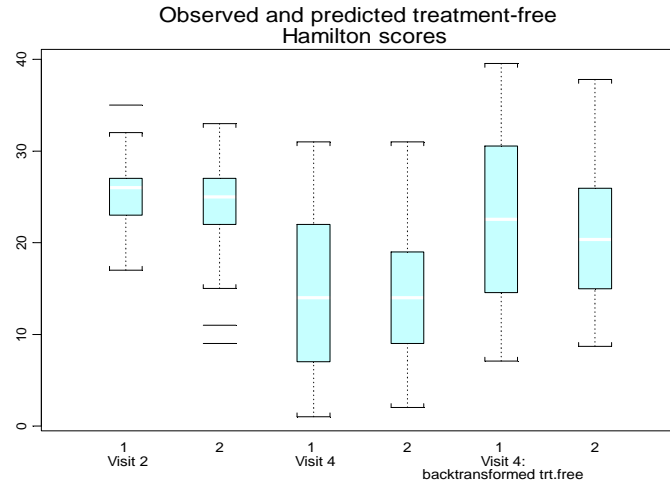
Figure 7.4: Box plots of observed baseline Hamilton score, outcome score and back transformed treatment-free score

7.5   Discussion

In this chapter we have worked thru a structural model for the effect of treatment actually received, thus making a first link between the more traditional approaches developed in the first chapters and a new school of causal inference emerging in the literature. While we have added something to the existing literature, we make no attempt to be complete here. Some observations and practical difficulties encountered are still worth mentioning.

It is well known that the precision of structural estimators can gain from good baseline predictors of adherence. Our analysis involved simple baseline predictors which yielded quite precise estimators nevertheless. Our intention, however, had been to use results from chapter 3 to predict patient adherence from the placebo run-in period. This turned out impossible because the run-in period covered only 7 days and observed adherence during this period was nearly perfect.

In this trial plasma concentrations of the drug (PK data) were not collected. However in a field like depression, where a strong placebo effect is recognized, one should not expect to find a better relation between plasma concentrations and effect than between dose taken and effect.

In the present work the initial choice for q_{opt} is 0. However, in practice

this could be estimated from a placebo group previously observed in similar conditions. Then the predictions from the obtained model could be used for the treatment group subjects. The necessary data were not available to us right now. We will investigate this at a later stage.

Chapter 8

Probabilistic intake assessment of dioxin-like substances

This chapter highlights the importance of collecting reliable measurements of exposure to estimate potential effects of xenobiotic substances in crisis situations. Practically, emergency decisions must be taken and the data are often collected according to suboptimal designs or not at all. Bias in the sampling scheme and changes in the procedures over time can invalidate direct conclusions. Probabilistic assessment of exposure becomes then a challenge and simulation an appealing technique to combine data collected from all sources.

With today's computing technology it is indeed possible to perform realistic simulations of potential exposures from multiple sources and even from multiple compounds. Simulation is hereby not a substitute for mathematics but an approach to calculation, resampling similar to using bootstrap for estimating the variance of a complicated point estimator. Theoretically, if your sample is small enough, you can actually calculate exactly what the bootstrap distribution is. You could work out what every possible sample could be and what the probability distribution would be. In practice, for a reasonably large sample, you just can not do this. Instead, one runs a simulation of the problem, which provides an empirical distribution which would be very close to what the theoretical distribution would be if only one had the time and the ability to perform the computation. That is the way in which the simulation actually overcomes a technical difficulty. Within the context of the therapeutic

cascade, many of the realistic analysis problems are just too difficult to solve analytically in the observational settings (mathematics) or by simpler experimental designs (ethics). A simulation approach frees us from certain technical constraints to analyze more realistic examples. Numerical simulation, third pillar of research, next to theory and observation becomes thus necessary when experimentation is impossible.

Probabilistic modeling techniques allow in principle for realistic estimates of exposure and risk by computing the full distribution of potential exposures rather than a single `worst case' average exposure. However, these techniques require additional considerations regarding the appropriate input data and models.

In this chapter we propose a 2-dimensional Monte Carlo simulation as a general approach towards population risk evaluation in the presence of uncertainties in exposure levels. In traditional probabilistic analyses, the between-subject variation and associated uncertainties are not separated but are uni-dimensionally propagated through the exposure model. The output is then presented as an hybrid distribution possibly leading to erroneous inferences. Therefore we found it necessary to extend the traditional Monte Carlo simulation and distinguish between inter subject variability and parameter uncertainty. The two dimensional approach thus allows adequate representation the between-subjects variation in intake and identification of potential subgroups at higher risk.

As an illustration, we evaluate the intake dioxin like contaminants via the food chain during the 1999 Belgian PCB-Dioxin incident.

8.1   Introduction

In Belgium, in February 1999, PCB oils and dioxins contaminated animal feed and, from there on, part of the human food chain. This dioxin incident became public at the end of May 1999 and gave rise to great concern about possibly serious health effects among the population. As a result, stringent food marketing and export limitations were imposed by the European Commission, which resulted in a recession for the Belgian food sector. How severely was the Belgian population contaminated and what risks are associated? In this chapter, we develop methodology to answer this question. Specifically we assess the risk regarding polychlorinated dibenzo-p-dioxins (PCDD), dibenzofurans (PCDF) and dioxin-like biphenyls (PCB), the most hazardous part of the contamination. Ideally one would have measured PCDD, PCDF and dioxin-like PCB concentrations in biological fluids in the general population. However such projects were not funded because of associated costs and because it appeared nearly impossible to identify sub-groups of subjects more exposed than others. The next best thing is to perform a risk assessment through relevant exposure data representative for the whole incident period. As it happens, detailed data on the distribution of intake of dioxin-like substances via food were not available for the period February - May. As a result we must rely on data sources with some uncertain link to the exposure of interest. Hence in our assessment we must confront two intrinsically different sources of variation: the inter subject variability and uncertainty in the input parameters resulting from a non-optimal sampling scheme. In a first step, we thus perform retrospectively a probabilistic exposure assessment based on 1) available food consumption data, sufficiently representative for the Belgian population during the incident period, and 2) the best estimates available regarding background and incident-related PCDD, PCDF, and dioxin-like PCB food contamination. Finally, we use the information on intake to estimate the variation in body burden of dioxin-like substances in the Belgian population which, for substances that are slowly eliminated from the body, form the basis for the health risk assessment. After lengthy discussions with the authorities this exercise fully started in the spring of 2000.

8.2 Daily intake of dioxin like contaminants

The probabilistic assessment of exposure aims to estimate the distribution of contaminant intake in the population under study. The between-subject distribution in exposure is important because it will allow us to identify high risk percentiles of the population. On the other hand, there is uncertainty in the subject-specific input parameters both concerning the consumed food items and their item-specific contamination level. A two dimensional Monte Carlo simulation will be developed to separate both sources of variation.

8.2.1 Model structure

The overall structure of the simulation process is illustrated in the flowchart given in figure 8.1.

For each subject i, the average daily intake of dioxin-like contaminants via food, DI_i, is computed according to the equation $DI_i = Y_i/(bw_i)$; where Y_i is the subjects average daily dose over the incident period (in pg TEQ) and bw_i his body weight (in kg). To estimate Y_i, we proceed as follows. We distinguish 9 different fat origins for fat in human consumption items. Let T indicate the length of the incident period in days. Then we denote by $X_{\nu;i;t}$ the amount (g) of fat from origin ν, consumed by subject i, at day t (t = 1; ::::; T), $\delta_{\nu;i;t}$ is a binary indicator determining whether the particular food item, $X_{\nu;i;t}$, is issued from an incident related production unit ($\delta = 1$) or not ($\delta = 0$), and $C_{\nu;i;t}$ is the concentration of the dioxin-like contaminant expressed in pg TEQ/g fat. The superscript I or B indicates whether the contamination stems from an incident related production unit or background contamination. Then, the subject-specific average daily dose is computed by combining individual fat consumption data (derived from food consumption data) and concentrations of the dioxin-like contaminants in the specific food items according to equation (8.1):

$$Y_i = \frac{1}{T} \sum_{\nu} \sum_{t} [(X_{\nu;i;t}C^I_{\nu;i;t})\delta_{\nu;i;t} + (X_{\nu;i;t}C^B_{\nu;i;t})(1 - \delta_{\nu;i;t})] \qquad (8.1)$$

While the model remains deterministic, $X_{\nu;i;t}$, $C_{\nu;i;t}$ and $\delta_{\nu;i;t}$ vary randomly between people according to a model derived from several data sources. The model parameters carry themselves some uncertainties.
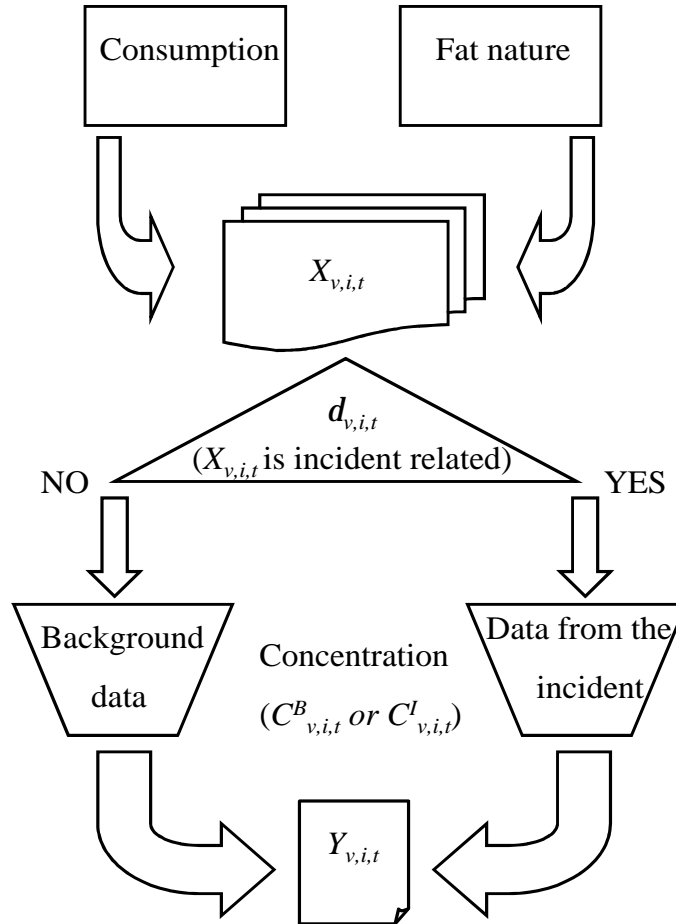
Figure 8.1: Flowchart illustrating the major inputs into the model. For fat origin ν and subject i at time t, we have: 1) food consumption $X_{\nu;i;t}$, 2) $\delta_{\nu;i;t}$ indicating if product $X_{\nu;i;t}$ is issued from an incident related production unit (1) or not (0), and 3) concentrations of dioxin-like contaminants in food items in background condition ($C^B_{\nu;i;t}$) or related to the incident ($C^I_{\nu;i;t}$). The model output $Y_{\nu;i;t}$, is the subject's average daily dose. For further explanation see text.

Monte Carlo simulations allow one to estimate variation in the outcome Y_i by repeating the execution of the model for different generation of sample values for input parameters. We can identify two types of such variation. The first source of variation is characterized by uncertainties in the parameters and will be denoted τ. The second source of variation is due to variation in characteristics between individuals and variation in product contamination. It will be denoted ξ in the sequel. The model input will be described in detail in the next section. Specific contribution to τ will be clearly stipulated.

8.2.2   Model input

Food consumption data

The only available Belgian study with food intake data at the individual level based on more than one day{a requirement for the estimation of within-subject variation in intake{is a survey in adolescents aged 14-18 years, carried out in the spring of 1997. A representative sample of 341 adolescents (both boys and girls) from the region of Ghent in the Dutch speaking part of Belgium completed a 7-day food diary in a standardized way (De Henauw et al. 2001). Instructions for the completion of the structured diary and regular checks for quality and completeness of the diaries were carried out by experienced dietitians. The check list contained altogether 745 different food items. From these, 527 food items are characterized as contributing in a measurable way to the fat intake of this population. For these 527 food items, the total fat content and the relative content of fat from different origins (vegetable, milk, egg, pork, beef, chicken, sheep, horse or fish) is determined on the basis of different food composition databases (NEVO 1996, NUBEL 1999, Souci et al. 2000), recipe books and in a limited number of cases also on the basis of information from the food industry. The available data on food consumption are translated in a three-dimensional database presenting, for each combination of fat origin ν, subject i and time in days t, the amount of fat consumed $X_{\nu;i;t}$ (g). It is further assumed that for each subject food items from the same fat origin, consumed on a single day, are issued from the same production farm. The food consumption input contributes to the variability ξ of the simulations proposed. The subjects with their corresponding food consumption and weight are empirically sampled. Denoting $X_{\nu;i;:} = \frac{1}{T} \sum_t X_{\nu;i;t}$, table 8.1 summarizes the `observed' between-subjects average daily fat consumption ($\overline{X}_{\nu;i;:}$),

Table 8.1: Average daily fat intake ($\overline{X}_{\nu;\mathbf{i};:}$ in g/day) from different origins, over 7 days, across 341 students.

Origin of fat	Minimum	Median	Mean	Maximum	% without intake
Chicken	0.00	1.13	1.55	11.08	23 %
Pork	0.00	10.17	11.41	39.36	1 %
Beef	0.00	4.31	4.80	27.11	1 %
Egg	0.00	2.38	2.66	8.95	1 %
Milk	0.01	19.46	20.56	69.99	0 %
Fish	0.00	0.18	0.71	11.07	30 %
Vegetables	6.19	42.67	44.68	111.70	0 %
Sheep	0.00	0.00	0.68	12.50	84 %
Horse	0.00	0.00	0.23	4.61	75 %

and their origin, as estimated from 7 days consumption data for the students, aged 14-18. While we acknowledge that some approximations have been made to determine the `observed' fat intake, we consider these values as given for the sequel. In the discussion section we will return to the impact of possible uncertainties at this level.

A first analysis of the data reveals the importance of fish in cases of high intakes of dioxin-like contaminants. For this reason we additionally study a dietary pattern{low in total fat and high in fruit, vegetables and fish{simulated on the basis of the dietary guidelines for coronary patients as issued by the Joint Task Force of European and other societies on Coronary Prevention (Wood et al. 1998). Table 8.2 presents the consumption data ($X_{\nu;\mathbf{i};:}$ in g/day) for a typical cardiac patient following a 14 day cholesterol poor, fish rich diet.

The probability that a food item comes from an incident related production unit

Incident related production units are farms that received contaminated animal feed during the incident. $\delta_{\nu;\mathbf{i};\mathbf{t}}$ is then derived according the following steps:

- According to data received from the Dioxin Cell of the Ministry of Agriculture and data from official statistics (Nationaal Instituut voor de Statistiek 2000), the proportion (p_{ν}^{0}) of incident related production units, was estimated as given in table 8.3.

Table 8.2: Daily fat intake ($X_{\nu;i;}$ in g/day) from different origins, over 14 days, in a typical prescribed cholesterol poor, fish rich diet.

Origin ofat	Daily minimum	Mean	Daily maximum
Chicken	0.00	1.36	9.68
Pork	0.00	1.15	4.38
Beef	0.00	0.93	2.95
Egg	0.00	1.74	6.11
Milk	5.46	8.60	13.92
Fish	0.00	4.15	19.62
Vegetables	50.89	56.95	107.40
Sheep	0.00	1.80	25.25
Horse	0.00	0.00	0.00

- p_ν^0 are estimations on a farm basis, whereas we really need to know what proportion p_ν of each product comes from a contaminated source. To acknowledge uncertainty at the product level, the probabilities $p_{\nu;\tau}$ were randomly drawn from a uniform distribution with probability density function:

$$f(p_{\nu;\tau}) = \frac{1}{p_{\nu;l}^0 - p_{\nu;u}^0} \text{ where } p_{\nu;l}^0 \leq p_{\nu;\tau} \leq p_{\nu;u}^0$$

and the uncertainty ranges $p_{\nu;l}^0$; $p_{\nu;u}^0$ are given in table 8.3.

- Furthermore, it is expected that during the dioxin-incident, items issued from incident related production units were not uniformly distributed across the population. In other words, it remained possible that some individuals had nearly continuously access to food from contaminated production units, while others had less. Therefore, for each contaminated fat nature ν, the inter-subject variability in the percentage of items consumed from a incident related production unit is modeled according to a uni-modal, right skewed, beta distribution. The individual probabilities $p_{\nu;\tau;i}$ are then randomly drawn with probability density function given in equation 8.2.

$$f(p_{\nu;\tau;i}) = \frac{p_{\nu;\tau;i}^{b_{1;\tau}-1}(1 - p_{\nu;\tau;i})^{b_{2;\tau}-1}}{B(b_{1;\tau}; b_{2;\tau})} \qquad (8.2)$$

Table 8.3: Number of incident related production units and total number of Belgian production units, estimated percentage and arbitrary chosen minimum and maximum uncertainty bounds.

Matrix	Numbers contaminated/total	p_ν^0 (%)	Range of uncertainty $[p_{\nu;l}^0; p_{\nu;u}^0]$ (%)
Chicken	580/2703	21	10-32
Pork	1934/7487	26	15-37
Beef	184/20775	1	0-5
Egg	175/4786	4	1-15

for $0 \leq p_{\nu;\tau;i} \leq 1$, $b_{1;\tau}; b_{2;\tau} > 0$, and $B(b_{1;\tau}; b_{2;\tau}) = \frac{\Gamma(b_{1;\tau})\Gamma(b_{2;\tau})}{\Gamma(b_{1;\tau}+b_{2;\tau})}$: $\Gamma(x)$ represents the Gamma function, such that, $\Gamma(x) = (x-1)!$. In order to let our uncertainty cover a large range of potential distributions, the parameter $b_{1;\tau}$ is randomly drawn between 1 and 10 as illustrated in figure 8.2, and $b_{2;\tau}$ is derived according to $b_{2;\tau} = b_{1;\tau} \times (1-p_{\nu;\tau})=p_{\nu;\tau}$ in order to achieve $p_{\nu;\tau}$ as the mean probability that an item is issued from an incident related production unit. Due to the structure of the food market, with separate production and distribution of different food items, we did not introduce a dependence between food items issued from different fat natures ν.

- For each subject i, $\delta_{\nu;i;t}$ is then drawn from a Bernouilli distribution with probability $p_{\nu;\tau;i}$.

Dioxin-like contaminants

As opposed to drug development where the focus lies in the exposition to a well defined chemical compound, accidental exposure is often concerned with several xenobiotic substances. Therefore, in this context, it is necessary to briefly describe the contaminants present during the Belgian incident. Dioxins is the general collective term for chlorinated aromatic compounds, consisting of a group called polychlorinated dibenzo-p-dioxins (PCDDs) and polychloro dibenzofurans (PCDFs). The group of PCDDs consists of 75 congeners (similar but not isomeric molecular structure), whereas the groups of PCDFs consists of 135 congeners. PCDDs and PCDFs have been identified worldwide in very diverse environmental media (Safe, 1994). Moreover, they accumulate and biomagnify in the food chain due to fat solubility and pronounced resis-
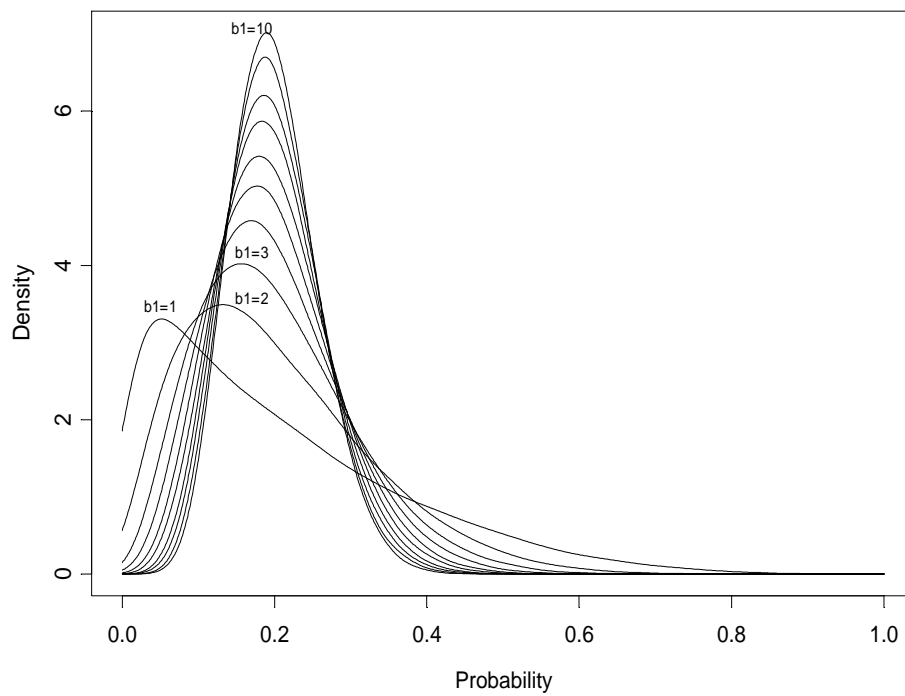
Figure 8.2: Beta distribution with parameters $b_1 = 1; ....; 10$ and $b_2 = b_1 \times (1 - p_{\nu;\tau})=p_{\nu;\tau}$.

tance towards metabolic degradation. In contrast to the dioxins group which does not have any known use and is not synthetized intentionally, polychlorinated biphenyls (PCBs) had well specified commercial uses such as dielectric, heat-exchange and hydraulic fluids in capacitors and transformers. Their production has ceased decades ago, but they are still among the main global contaminants. Chemically, the term PCBs denotes a family of 209 congeners, which exhibit a wide range of toxicological properties. Some of them, known as dioxin-like, are potentially the most toxic given their structural similarity to PCDDs and PCDFs. 2,3,7,8- Tetrachloro dibenzo-p-dioxin (TCDD) is the prototype for this class of aromatic hydrocarbons which have similar patterns of toxicity. Because of the need to estimate toxicity, the international community came up with the concept of toxic equivalency factors (TEFs) to address and facilitate risk assessment. Toxic potencies of the different congeners are evaluated relative to TCDD and are indicated by their corresponding TEF values. Dioxin-like contaminants, e.g. PCDDs, PCDFs and dioxin-like PCBs are then expressed in toxic equivalents ($TEQ = weight \times TEF$) concentrations in pg TEQ/g, or intakes as pg TEQ/kg bw. The consensus TEFs were used (Van Den Berg et al. 1998). In the sequel, the term \dioxins" will be used for the sum of 2,3,7,8-chloro-substituted PCDDs and PCDFs, whereas the term \dioxin-like substances" or \dioxin-like contaminants" will include the dioxins and the dioxin-like PCBs.

Before the incident, dioxin-like PCBs were never assessed in Belgium. Since then, new methodological procedures were investigated and out of 12 dioxin-like PCBs, four congeners (77, 81, 126, and 169) are now systematically assessed together with PCDDs and PCDFs, in a number of food items, within an actual Belgian contamination control program (Focant et al. 2000). Whereas the PCB concentrations in incident related samples were significantly higher than in non-incident related or background samples, there is no significant difference in the PCB congener profiles (Dioxin Body Burden Working Group 2001), indicating that both contaminations originate from a similar source. We, therefore, rely on the more recent results in order to estimate the ratio r(r > 1) of the concentration of total dioxin-like substances ($C_{\nu;i;t}$) to the concentrations of PCDDs and PCDFs ($c_{\nu;i;t}$) in fat items of interest for our intake assessment:

$$C_{\nu;i;t} = r \times c_{\nu;i;t}:$$

Ratios based on dioxin and dioxin-like PCB concentrations < 0:1 pg TEQ/g fat are excluded. In the simulations, empirical sampling within

Table 8.4: Concentrations of PCDD and PCDF, expressed in pg TEQ/g fat, ventilated over the different fat natures, as determined in Belgian and Dutch food items in background conditions.

Fat nature	1st Qu.	Median	Mean	3rd Qu.	n
Milk	0.79	1.35	1.39	1.79	38
Fish	6.38	12.93	20.23	28.22	68
Horse	4.22	6.48	7.23	10.48	15
Sheep	0.07	0.27	0.67	0.38	5
Vegetables	0.08	0.09	0.09	0.10	6
Chicken	0.13	0.27	0.30	0.38	70
Egg	0.71	1.62	2.29	2.55	16
Pork	0.01	0.05	0.21	0.15	55
Beef	0.90	1.42	1.69	1.87	49

the distribution of ratios is realized separately for each fat origin. For fish, all samples available had been analyzed for PCDDs, PCDFs and dioxin-like PCBs. $C_{\nu;i;t}$ could thus be directly sampled.

- **When the food item is assumed to come from a non-incident related production unit, background levels of dioxins (PCDD/F) are applied. The degree of contamination, $c^B_{\nu;i;t}$, is randomly sampled from the empirical distribution of Belgian and Dutch background levels (summarized in table 8.4) and contributes to the variability ξ of the simulation proposed. Since background levels are quite similar in samples taken in 1999 and in 2000, background data from both years are pooled in the analysis.**

- **As mentioned before, it is hard to get reliable information on incident related dioxin concentrations. The degree of contamination, $c^I_{\nu;i;t}$ in incident related samples represents the concentration of PCDD, PCDF expressed in TEQ. However, because of high associated costs, dioxin analysis were not carried out at random but were only performed if the PCB analysis exceeded the maximum tolerated value. Once the incident became public, PCB assessments were realized randomly in all products that enter the food chain. We are thus confronted with two major uncertainties : 1) the representativeness of the PCB samples for the period before the incident was made public and 2) the prediction of TEQ values from PCB measurements. They are described explicitly here be-**

low. To simplify our notations, for each fat nature ν, dioxin and PCB concentrations measured in sample k will be denoted $\mathbf{TEQ}_{\nu;\mathbf{k}}$ and $\mathbf{PCB}_{\nu;\mathbf{k}}$ respectively.

We derive PCB information from the database from the Ministry of Agriculture based on an intensive food sampling program for non-dioxin-like PCBs expressed in weight units as the sum of 7 marker PCBs (28, 52, 101, 118, 138, 153, and 180). From a total of 42,336 marker PCB analyses carried out during the 1999 control program, 15,261 originate from incident related production units. From those, a total of 5,264 originate from basic food items that directly enter the human food chain and are clearly identified with regard to fat origin ν and date of sampling $\mathbf{d}_{\nu;\mathbf{k}}$. However, not all 5,264 data are representative for the food contamination during the incident. Figure 8.3 shows that the number of samples from the initial period, before the incident became public end of May, is quite limited. Moreover, in chicken and eggs, the food types that were most contaminated, concentrations decrease in function of time.

A major uncertainty therefore concerns the concentrations in food items in the initial period which is most important for our incident linked exposure assessment. To take this uncertainty into account, the PCB concentration of a sample, $\mathbf{PCB}_{\nu;\mathbf{k}}$, is used if the corresponding date of sampling $\mathbf{d}_{\nu;\mathbf{k}} < \Lambda_{\nu;\tau}$. The arbitrary cut off dates Λ_{ν}, determining which samples are introduced into the simulation, are summarized in table 8.5 and graphically presented in figure 8.3for each fat nature ν. In order to test the sensitivity of the end result towards this uncertainty, a second, less severe simulation (B) is run.

{ For eggs and chicken, the end date of sample concentrations used in simulation A is allowed to vary between May 15 and July 15. Taking May 15 as the end date, subjects that consumed incident related products are supposed to have consumed mainly highly contaminated eggs and chicken. When data until July 15 are included, it is assumed that also less contaminated eggs and chicken had been consumed from incident related production units. For the less severe simulation B, the end date of sampling for eggs and chicken is allowed to vary between June 15 and July 15.

{ For pork, showing a much longer turnover production period,

Table 8.5: Uncertainty interval about the end date the PCB samples are introduced into the simulation. A and B indicates the two type of simulations described in the text.

$\Lambda_{\nu;\tau}$	Uncertainty interval	Number of samples included	
		Min	Max
$\Lambda^{\mathbf{A}}_{\mathbf{egg};\tau}$	[May 15 - July 15]	4	302
$\Lambda^{\mathbf{A}}_{\mathbf{chicken};\tau}$	[May 15 - July 15]	14	802
$\Lambda^{\mathbf{B}}_{\mathbf{egg};\tau}$	[June 15 - July 15]	252	302
$\Lambda^{\mathbf{B}}_{\mathbf{chicken};\tau}$	[June 15 - July 15]	393	802
$\Lambda_{\mathbf{pork};\tau}$	[June 15 - July 15]	19	949
$\Lambda_{\mathbf{beef};\tau}$	[July 15 - December 15]	163	481

the sampling end date for both simulations vary between June 15 and July 15.

{ For beef, where contaminations remained much more stable over time, for both simulations, it vary between July 15 and December 15.

In this way a maximum of 2,534 contamination data, out of 5,264, are used in our simulations. Among these, 221 are simultaneously analyzed for PCDD and PCDF. For a number of PCB concentrations below time dependent, maximum tolerated concentrations{ identified at decreasing concentrations during the course of the incident, from 1000 ng/g fat down to 200 ng/g fat{the actual concentrations are either not reported or they are reported as being below a given limit: which means that a left censored observation is reported. Since the dioxin analyses were not carried out randomly but mainly in samples positive for marker PCBs, dioxin results do not constitute a random sample from the dioxin concentrations in the overall food items and they cannot be used for random sampling from the empirical distribution. Therefore, random sampling is realized in the non-dioxin-like PCB concentrations and, according to equation 8.3 the dioxin concentrations are predicted from the selected marker PCB concentrations:

$$
(\quad
\begin{aligned}
&\ln(\mathbf{TEQ}_{\nu;\mathbf{k}}) = \alpha_{0;\nu} + \epsilon_{\nu;\mathbf{k}}; \text{ for censored PCB samples} \\
&\ln(\mathbf{TEQ}_{\nu;\mathbf{k}}) = \alpha_{1;\nu} + \beta_{1;\nu}\ln(\mathbf{PCB}_{\nu;\mathbf{k}}) + \epsilon_{\nu;\mathbf{k}}; \text{ else.}
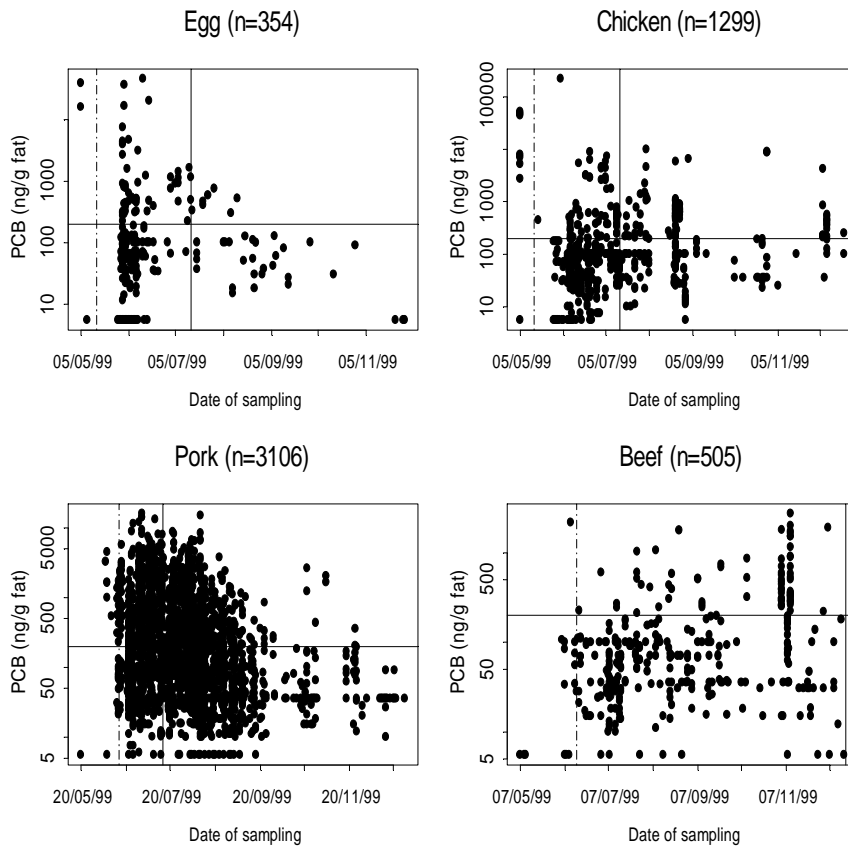\end{aligned}
\qquad (8.3)
$$

Figure 8.3: Scatter plots of the observed PCB concentrations in incident related samples over time for the 4 contaminated fat items. The horizontal line indicates the level of 200 ng PCB/g fat, the maximal tolerated level that was eventually chosen for marker PCBs. The vertical lines indicate, for each matrix, the minimal and maximal end of the period during which analytical results were introduced into the simulation (for explanation see text).

Where $\epsilon_{\nu;k} \sim N(0; \sigma_{\nu})$. $\alpha_{0;\nu}$ is thus the mean value for samples of fat origin ν that are censored for the marker PCB value; $\alpha_{1;\nu}$ and $\beta_{1;\nu}$ are the intercept and slope when regressing the logarithm of the dioxin concentrations in function of the logarithm of the sum of the seven marker PCBs. Typically, technical measurement in the assay lead to TEQ values that are left and/or right censored. Therefore, we use parametric regression, allowing for censored observations in order to estimate the regression parameters. Uncertainty about these parameters, due to finite sample size estimation, is reflected in their covariance structure Ω. To acknowledge this, we use a multivariate normal distribution to represent the population mean parameters and error variance of the log transformed TEQ values. Specifically, $(\alpha_{0;\nu}; \beta_{1;\nu}; \beta_{2;\nu}; \sigma)_{\tau} \sim MVN((\alpha_{0;\nu}; \beta_{1;\nu}; \beta_{2;\nu}; \sigma); \Omega)$.

Duration of the incident

The exposure assessment presented in this paper is designed to produce an estimate of the average daily dose during the Belgian dioxin-incident. It is officially accepted that the incident started on February 1 and ended end of May, 1999. Therefore, we are evaluating a risk of increased intake during 16 weeks (T = 112 days). Since the period at risk is longer than the record data we have on food intake, we repeat the latest as many times as necessary to cover the risk period. Doing this, we ignore potential within-subject variation over time.

8.2.3   Model estimation

The Monte Carlo simulation of the TEQ distribution runs the model several times, each time using different values for each of the input parameters. We used two dimensional simulations in order to distinguish two different sources of variation. First, uncertainty is a property of the analyst and is characterized by uncertainty in the incident parameters, and by uncertainty due to sample size. Since there are multiple input parameters with uncertainty, one value from each is sampled simultaneously in each repetition in the simulation. Each Monte Carlo simulation represents a random uncertainty from an m dimensional uncertainty space, where m is the number of uncertainties that are inputs in the model. In order to represent uncertainty over the course of the simulation, 200 draws were taken from the joint uncertain model. Then, for each random

uncertain vector, the between-subject variation in output was assessed. This is a property of the system studied and is characterized by variations in characteristics between individuals (between-subject variation in food consumption $X_{\nu;i;t}$ and weight), and variations in product contamination ($C_{\nu;i;t}$, $\delta_{\nu;i;t}$). The result is a two dimensional table that gives, for each uncertainty vector, a set of values for the model output variable DI_i, which can be analyzed as if they were an experimentally or empirically observed set of data. This enables us to evaluate the cumulative distribution function in the directions of both between-subject variation and uncertainty.

The estimates of the cumulative distribution function are numerically more stable than estimates of probability density and are thus more meaningful. Furthermore, the cumulative distribution function allows for quantitative insight regarding the percentiles of the distribution. Although the generation of sample values for model input parameters is probabilistic, the execution of the model for a given set of samples, in repetition, remains deterministic. The advantage of Monte Carlo methods, however, is that these deterministic simulations are repeated in a manner that yields important insights into the sensitivity of the model to variations in the input parameters, as well as into the likelihood of obtaining any particular outcome. It allows the modeler to use any type of probability distribution for which values can be generated on a computer, rather than restricted to forms which are analytically tractable. All the simulations presented were run using the S-PLUS 2000 software (Professional release 1, Math Soft Inc).

8.3   Body burden

Having determined the daily intake in background condition as well as during the incident, the next step is to evaluate the potential effect of the incident on human health. The dioxin-like contaminants of interest for this study are lipophylic substances, showing a slow metabolic degradation in the organism. They accumulate over time and their kinetic behavior can be described by a single compartmental model with almost instantaneous distribution. It is assumed that the doses may be considered as rapidly absorbed daily pulses. The dose is the dietary intake, corrected for gastrointestinal absorption. For repeated dosing, the body burden (bb, in ng TEQ/kg bw) can be recursively estimated according to equation(8.4):

$$bb_{(day=t)} = f \times DI + bb_{(day=t-1)} e^{-k_e}: \qquad (8.4)$$

Where DI is the daily intake (in pg or ng TEQ/kg bw/d), f is the fraction of the dose absorbed and k_e is the constant elimination rate (day^{-1}). Following a regular daily intake of a compound, the body burden will increase over a period of time until an equilibrium, or steady state, is established. At equilibrium, the amount of compound eliminated each day exactly equals the intake. The average steady state body burden ($\overline{bb}^{SS}$) derived from equation (8.4) is given by:

$$\overline{bb}^{SS} = \frac{f \times DI \times t_{1=2}}{\ln(2)} \qquad (8.5)$$

Where the elimination half life ($t_{1=2}$) can be expressed in terms of the elimination rate k_e, $t_{1=2} = \frac{\ln 2}{k_e}$, and is considered constant during the relevant lifetime. The values for f and $t_{1=2}$ for dioxin-like substances are quite variable between congeners and for many unknown. Since they are assumed to behave similarly in the organism as TCDD, it is generally accepted to apply the TCDD values to the whole of the congeners that play a role in the total TEQ. We used f = 0:5, and $t^{1=2}$ = 7:5 years (WHO-ECEH-IPCS 2000). Starting at zero body burden at birth ($bb_{(day=0)} = 0$), we first estimated life time body burdens under background conditions. As input we used the background daily intake estimated in simulation A (adolescents) up to 50 years of age. At the age of 50, we introduced the background daily intake of dioxin-like substances estimated in simulation C (coronary patient) to calculate the body burden up to the age of 60. A separate estimation was realized to assess the increase in body burden due to the dioxin-incident. To this end, two four month periods{the assumed duration of the dioxin-incident, February till May{of increased intake we are superimposed to the background situation, one at the age of 16 (intake estimation from simulation A), and one at the age of 60 (intake estimation from simulation C). In separate estimation, both the most likely daily intake at median uncertainty and a worst case daily intake at 95% uncertainty were introduced in the calculation.
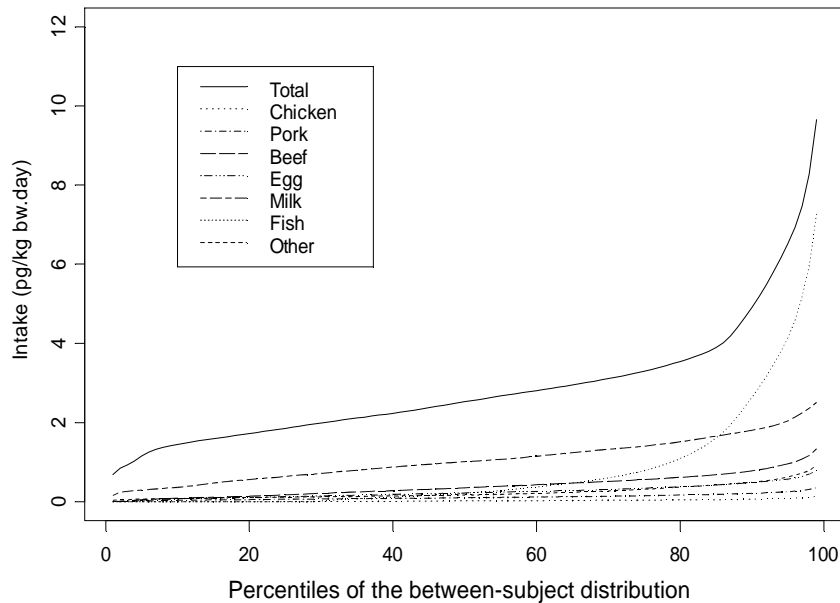
Figure 8.4: **At the median uncertainty level, the inverse cumulative distribution function is plotted for background fat intake in adolescents, first total fat intake and then for each fat nature separately. For clarification see insert.**

8.4 Results

8.4.1 Monte Carlo simulations of daily intake of dioxin-like substances via food in background conditions

Background daily intake in the adolescent subgroup

The combination of adolescent food consumption and background food contamination data (simulation A) led to the daily background intake of dioxin-like compounds illustrated in figures 8.4 and 8.5. In these conditions no uncertainty is introduced into the parameters, the only uncertainty being due to sample size. Since we performed 1500 replicates in the simulation of between-subject intake variation, the uncertainty with regard to background intake becomes negligible and the curve at median uncertainty can be accepted as the best probabilistic estimation.

Figure 8.4 represents median uncertainties of the cumulative distribution functions of intake representing the between-subject variation for total fat intake and for each fat nature separately. The figure shows that an estimated 3% of the adolescent population remains below a daily intake of 1 pg TEQ/kg bw/day (the lower WHO tolerated daily intake) for the sum of PCDDs, PCDFs and dioxin-like PCBs; 85% of the population is estimated to remain below 4 pg TEQ/kg bw/day (the higher WHO tolerated daily intake). The analysis of each fat nature separately indicates that a significant contribution to the total intake in background conditions, for the entire population, happens through milk and milk derived products. In the subjects with higher intakes, above the 80th percentile, fish and fish oils become the major source of intake of dioxin-like substances. The importance of the other contamination sources (at the 95th percentile) are, in decreasing order: beef, remaining sources, egg, pork and chicken. Remaining sources are horse, sheep and vegetables taken together. It should be stressed that the variation for each fat nature is estimated independently, therefore the total intake at a certain percentile is not the sum of the different fat contributions at that percentile.

A different way to look at the relative magnitudes in variation in intake and associated uncertainty is presented in figure 8.5, showing the 5th, median, and 95th variation percentile of the subject specific cumulative uncertainty distribution function in intake, as well as the individual curves for 50 randomly selected adolescents. The between-subject variation{indicated by the degree of spread across the 5th, median, and 95th percentile{is much greater than the uncertainty about any of the individuals as indicated by the range covered by any single line. Moreover, all the subject specific cumulative uncertainty distribution functions in intake are fairly horizontal, indicating very small individual uncertainty in magnitude. The large between-subject variation can thus be interpreted as subject specific risk for different degrees of TEQ dioxin intake. This means that some people are more prone to high TEQ dioxin intakes than others. This is linked to the individual amount and nature of fat consumption.

Background daily intake in the coronary patient

A similar combination of a typical coronary diet and background food contamination data (simulation C) led to the daily background intake of dioxin-like compounds{total fat and each fat nature separately{
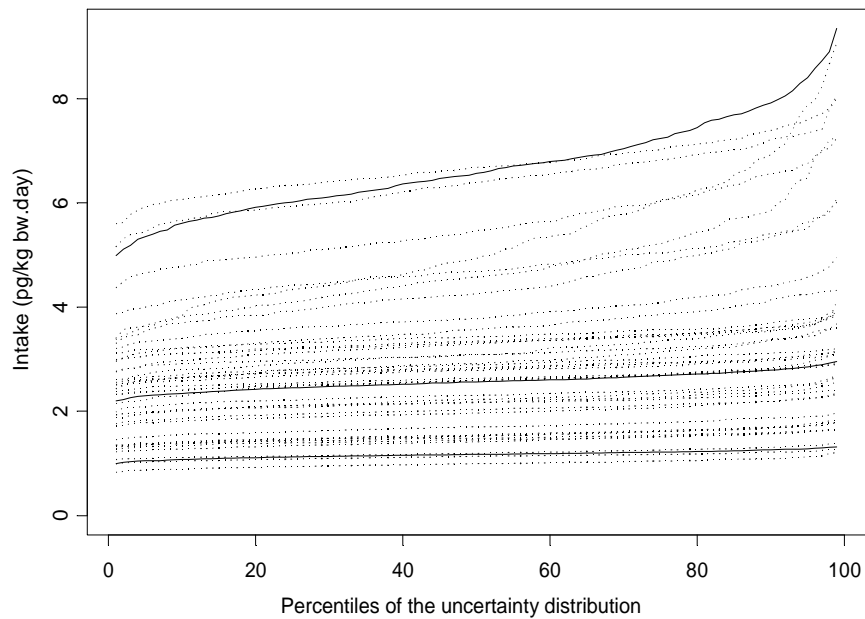
Figure 8.5: The full lines correspond to the 5th, median and 95th percentile of the between-individuals distribution of total background fat intake. At each of those uncertainty values, the inverse of the cumulative uncertainty distribution in background intake is shown in the full line, together with the corresponding functions for 50 randomly selected adolescents in dotted lines.
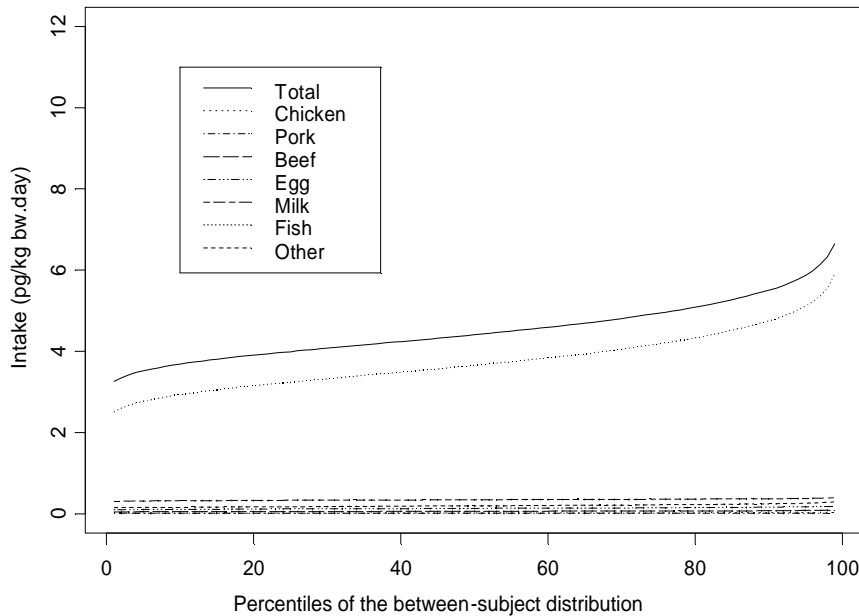
Figure 8.6: At the median uncertainty level, the inverse cumulative distribution function is plotted for background fat intake with the coronary diet, first total fat intake and then for each fat nature separately. For clarification see insert.

described in figure **8.6**. For similar reasons, as before the curve at median uncertainty can be accepted as the best probabilistic estimation. The figure shows that a coronary patient, carefully following the prescribed diet, remains above the daily intake of 1 pg TEQ/kg bw/day for the sum of PCDDs, PCDFs and dioxin-like PCBs while only 25% of patients would remain below 4 pg TEQ/kg bw/day. The analysis of each fat nature separately reveals the dominant role of fish in the daily intake of dioxin-like contaminants with this diet. Increasing the fish intake in a coronary patient diet results in a marked increase in dioxin intake in comparison with the adolescent subgroup. Still milk and milk derived products remain a significant contributor to the total intake. At the 95th between-subjects percentile, the other contamination sources are, in decreasing order: remaining sources, egg, beef, chicken, pork, but all negligible compared to the intake via the fish.
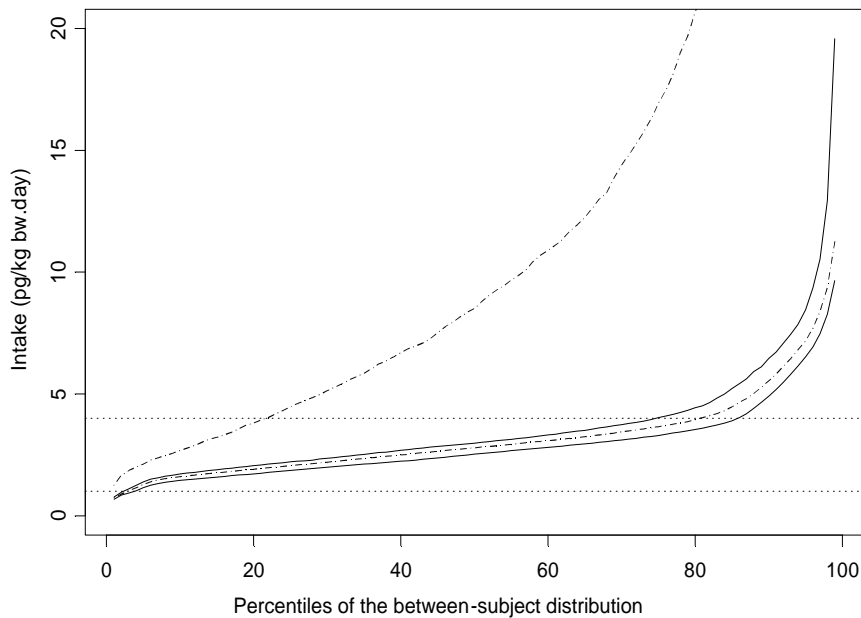
Figure 8.7: At 5% (lower dashed line), median (upper full line), and 95% (upper dashed line) uncertainty level, the inverse cumulative distribution function is plotted for incident total fat intake in adolescents. For comparison, corresponding median background total intake curve (lower full line) is added to the plot. Horizontal dotted lines represents the WHO recommended TDIs of 1 and 4 pg/kg bw/day.

8.4.2 Monte Carlo simulations of daily intake of dioxin-like substances via food during the Belgian dioxin-incident

Incident daily intake in the adolescent subgroup

Introducing incident contamination data for chicken, egg, pork and beef, and background contamination data for milk, fish, vegetables, sheep and horse in simulation A led to the incident daily intake of dioxin-like contaminants illustrated in figures 8.7 and 8.8. The results presented acknowledge the range of possible uncertainty combinations. In figures 8.7, the 5% and 95% uncertainty curves represent the 5% lower and 95% upper bound probabilities of exposure to lower or higher degrees of food contamination.

Figure 8.7 shows that between-subject variation in incident intake within this group of adolescents is greater than the uncertainty in the estimation. Nevertheless, uncertainty becomes larger when moving towards the highest percentiles of intake. During the dioxin-incident, the intake of dioxin-like substances by the 50th percentile of adolescents is estimated to be increased by 0.25 (5% uncertainty), over 0.45 (median uncertainty) towards 5.96 (95% uncertainty) pg TEQ/kg bw/day. This increase is accepted to be significant since the curve representing the background intake lays fully below the 5% uncertainty curve during the incident.

Figure 8.8 represents median uncertainties of the cumulative distribution functions in intake representing the between-subject variation for total fat intake and for each fat nature separately. Milk and milk derived products remain important as source for all adolescents. For most of the adolescents the moderate increased intake follows the contamination of chicken, egg, pork and beef. Fish remains an important input for the higher percentiles. Only from the 90th percentile on, chicken becomes more important as source and overtakes fish at the highest percentiles.

Incident daily intake in the coronary patient

Simulation C, now combining the incident and background contamination data with the coronary diet, led to an incident daily intake of dioxin-like contaminants illustrated in figures 8.10 and 8.11. In figures 8.10, the 5% and 95% uncertainty curves represent the 5% lower and 95% upper bound probabilities of exposure to lower or higher degrees of food contamination. During the dioxin-incident, the intake of dioxin-like substances via this diet might have increased, at the 50th percentile, by 0.07 (5% uncertainty), over 0.21 (median uncertainty) towards 4.57 (95% uncertainty) pg/kg bw/day. The lesser degree of increase, in comparison with the adolescents, results from the greater importance of fish over meat in this diet, the fish contamination staying unchanged during the dioxin-incident.

Figure 8.11 represents median uncertainties of the cumulative distribution functions in intake for total fat intake and for each fat nature separately. Fish remains important as source. The overall increased intake follows the increased contamination in chicken, egg, pork and beef. Only from the 95th percentile on, chicken becomes more important as source and overtakes fish at the highest percentiles.
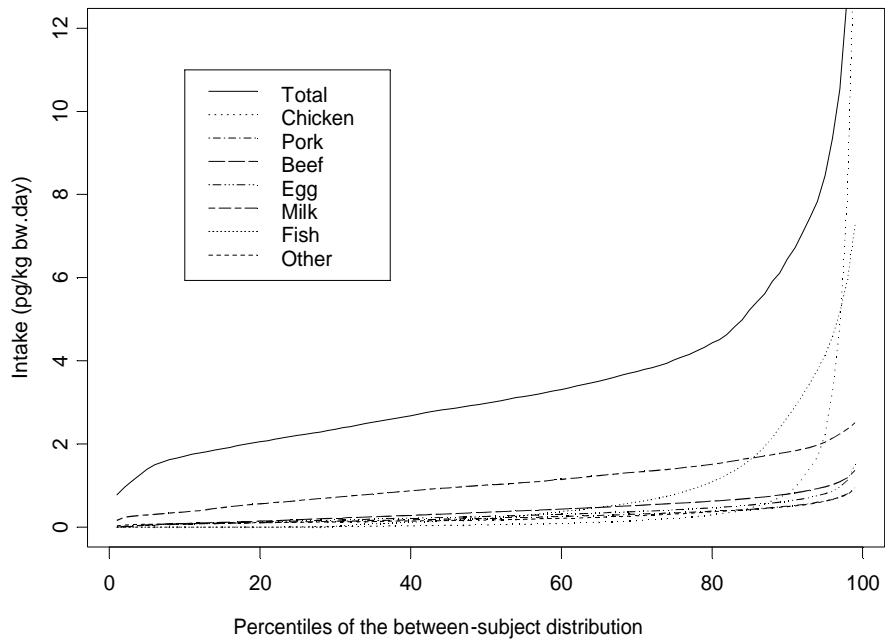
Figure 8.8: At the median uncertainty level, the inverse cumulative distribution function is plotted for incident fat intake in adolescents, first total fat intake and then for each fat nature separately. For clarification see insert.
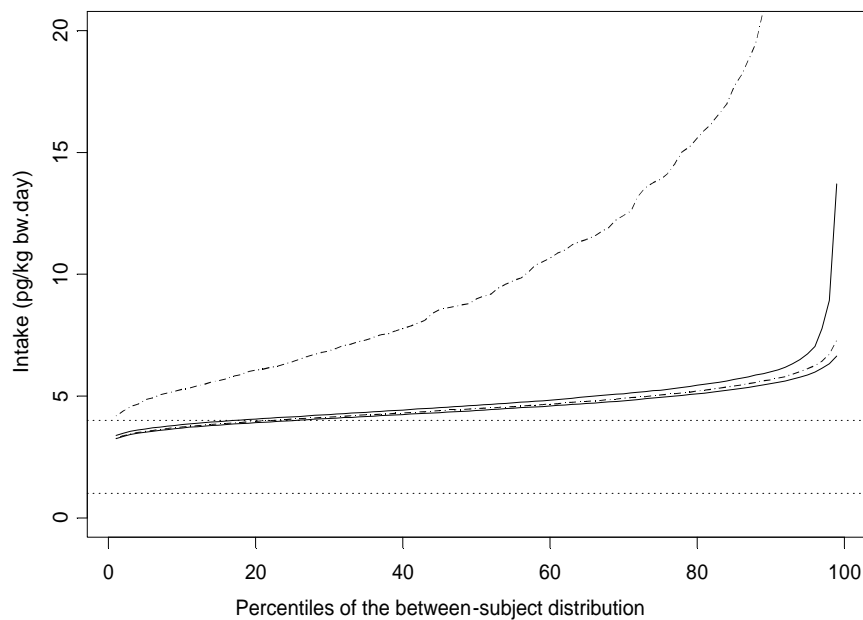
Figure 8.9: At 5% (lower dashed line), median (upper full line), and 95% (upper dashed line) uncertainty level, the inverse cumulative distribution function is plotted for incident total fat intake with the coronary diet. For comparison, corresponding median background total intake curve (lower full line) is added to the plot. Horizontal dotted lines represents the WHO recommended TDIs of 1 and 4 pg/kg bw/day.

Figure 8.10: 5% (lower dashed line), median (upper full line), and 95% (upper dashed line) uncertainties of the cumulative distribution functions in incident total fat intake, together with the median background intake curve (lower full line) in a coronary patient. Horizontal dotted lines represents the WHO recommended TDIs of 1 and 4 pg/kg bw/day.
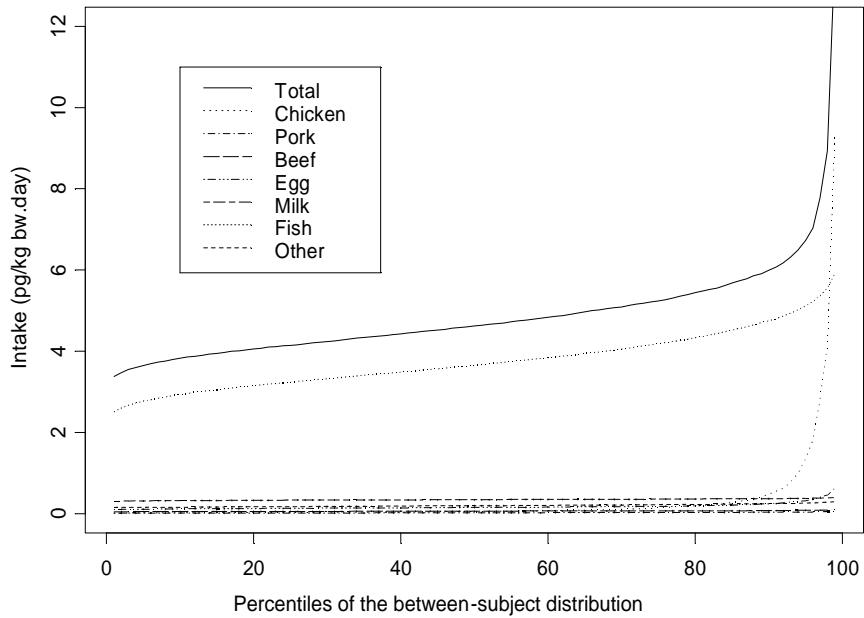
Figure 8.11: At the median uncertainty level, the inverse cumulative distribution function is plotted for incident fat intake with the coronary diet, first total fat intake and then for each fat nature separately. For clarification see insert.

8.4.3 Identification of key contributors to the between-subject variation and associated uncertainty in the intake estimation

In order to improve the understanding and interpretation of the simulation, we intend to identify which inputs \drive" the outcome results. To this end, key contributors to between-subject variation and associated uncertainty were identified separately.

Between-subject variation : regression analysis

In order to identify the most influential inputs with regard to between-subject variation in intake of dioxin-like substances, we considered the group of 341 adolescents in a background situation. For each of them we simulated the dioxin intake across 400 uncertainty replications. Assuming a log normal distribution, the median uncertainty in intake was modeled via a regression model. The following potential covariates were tested: age, gender, and the average daily intakes of the different fat origins: fish, pork, beef, milk, chicken, vegetables, eggs, and other. Parameters from a multivariate regression were estimated by forward stepwise regression allowing for first order interactions. In line with figure 8.4, fish, milk and beef were the main contributors to the between-subject variation in dioxin intake. Gender effect, while identified as a significant contributor to dioxin intake in the univariate analysis, does not enter in the multivariate analysis, probably because differences between sexes are due to the fact that for a same body weight, on average, males consume more fat than females.

Uncertainty: sensitivity analysis

The simulation introduces four main sources of uncertainty with regard to the consumption of the possibly contaminated fat natures, chicken, egg, pork and beef:

1. the probability of consuming food from an incident related production unit,

2. the inter-subject variation in the percentage of items issued from an incident related production unit,

3. the sampling period which is representative for the main contamination period,

Table 8.6: Comparison of median intakes for the different percentiles of the adolescent subgroup population between background conditions, incident basal conditions, and incident conditions skewed towards the high initial contamination concentrations. For further explanation see text.

Percentile	Background median uncertainty	Incident 95% uncertainty simulation A	simulation B
1	0.68	1.24	0.79
3	0.93	1.84	1.13
5	1.15	2.11	1.41
25	1.85	4.46	2.22
50	2.53	8.49	3.02
75	3.30	16.91	4.02
95	6.52	47.23	8.43
97	7.47	58.04	10.34
99	9.65	94.62	19.63

4. the regression coefficients for predicting dioxins from PCB.

In order to identify the most influential uncertainties, we ran several simulations in which all but one uncertainty were set to their central values. We identified the sampling period (uncertainty 3) as by far the major source of uncertainty. Since no reliable information can guide us in this choice, we included in the uncertainties of simulation A the most severe situation, where the period considered as representative for the incident for chicken and egg data was only until May 15. In other words, we allowed that chicken and egg products issued from incident related production units were all highly contaminated. In order to quantify the impact of this uncertainty, we compared this with less stringent conditions and ran a simulation B where the period considered as representative for the incident for chicken and egg data was at least from May 1 till June 15. The 95% uncertainty results of simulations A and B, the worst cases for each, are given in table 8.6, illustrating the important impact of the choice of the data period.

In conclusion, the majority of the between-subject variation in the output distribution is attributable to 1) variation in the amount of fat intake via fish, milk and beef and 2) the uncertainty in the time period over which the data are considered representative for the incident.

8.4.4   Body burden

Figure 8.12 gives a graphical representation of the results obtained. The first body burden estimation accepts a daily intake at the 50th percentile{simulation A up to the age of 50, simulation C up to the age of 60{in background conditions; the second estimation accepts a daily intake at the 95th percentile; the third estimation accepts a daily intake at the 97th percentile.

A second run follows for each of those percentiles, but now super imposing at the age of 16 the dioxin-incident intakes of simulation A, at the median and 95% uncertainty levels, and at the age of 60 the dioxin-incident intakes of simulation C, also at the median and 95% uncertainty levels.

Changing at the age of 50 from an adolescent daily intake towards that of the coronary patient produces an increase in body burden at the 50th percentile, but a decrease at the 95th and 97th percentiles. This distinction might be artificial since the coronary patients daily intakes at the 95th and 97th percentiles are most probably underestimations, based on an assumed rigorous adherence to the theoretical diet. It is further obvious from the figure that the temporary dioxin-incident related increases in body burden at the median level of uncertainty are quite low for each of the percentiles in intake. At the 95% and 97% uncertainty level a more marked increase in body burden becomes apparent, which is maintained up to the age of 50.

8.5   Comments and discussion

Material and methods

In traditional probabilistic analyses, where the between-subject variation and associated uncertainties are not separated and are uni- dimensionally propagated through the exposure model, the output would have presented an hybrid distribution possibly leading to erroneous inferences. This helped to adequately quantify the between-subject variation in intake that identified significant sub populations, e.g. people eating fish, that merit more focused attention.

The full probabilistic intake assessment has been limited to adolescents, the only subgroup of the Belgian population for which recent, sufficiently detailed food consumption data are available. However, previous sur-
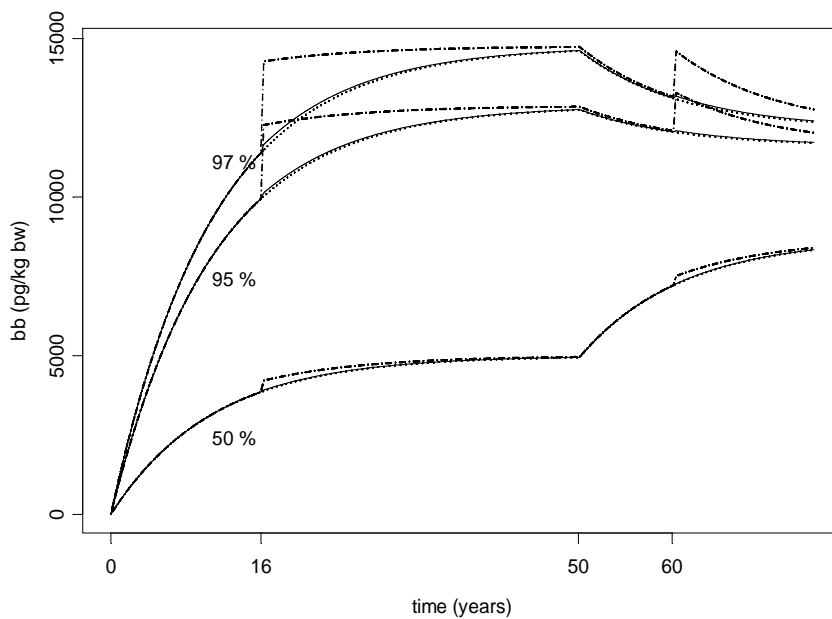
Figure 8.12: **The first body burden estimation accepts a daily intake at the 50th percentile{simulation A up to the age of 50, simulation C up to the age of 60{in background conditions (dotted lines); the second estimation accepts a daily intake at the 95th percentile; the third estimation accepts a daily intake at the 97th percentile. A second run follows for each of those percentiles, but now super imposing at the age of 16 the dioxin-incident intakes of simulation A, at the median (full lines) and 95% uncertainty (dashed lines) levels, and at the age of 60 the dioxin-incident intakes of simulation C, also at the median (full lines) and 95% uncertainty (dashed lines) levels. Please remark that the increase at median level of uncertainty can hardly be seen at this ordinate scale**

veys would indicate that total daily fat consumption in our region is quite similar in adults as in adolescents (Kornitzer and Bura 1989), as it is also in the Netherlands (Voedingscentrum 1998). A preliminary analysis of food consumption habits in 1340 senior Belgian persons (65-92 years) further suggests that fat intake becomes substantially less after 65 years of age, whereas the distribution of fat origin is quite similar to that in adolescents (Dioxin Body Burden Working Group 2001). Therefore, the extrapolation to the more general population{as has been done more especially in the body burden calculation up to the age of 50{ seems acceptable and might even be an overestimation. At the end, the estimated median fat intakes (see table 8.1) appear to be of the same order as those estimated for the Dutch general population (Liem and Theelen 1997). Further we did an attempt to assessed the intake based on a typical coronary patient.

Our dioxin intake assessment includes PCDDs, PCDFs and dioxin-like PCBs, leading towards an overall intake expressed in pg TEQ/kg bw/day. This intake might be somewhat underestimated since for several food items background dioxin-like PCB contaminations were only measured for the non-ortho PCBs, but included PCB 126 and 169, most important congeners with regard to toxicity (TEF values 0.1 and 0.01). In summing up PCDDs, PCDFs and dioxin-like PCBs, we stay in line with the Scientific Committee on Food (SCF) (2000) that stated that the tolerated daily intake (TDI){and, therefore, also the daily intake (DI){for 2,3,7,8-TCDD could be extended to include all 2,3,7,8-substituted PCDDs and PCDFs, and dioxin-like PCBs, each multiplied by their respective TEF. An important observation is that, during the dioxin incident, a major part of food items contained levels of dioxin-like substances that did not exceed background contamination.

Several major uncertainties related to the incident period were taken into account in the probabilistic assessment. The major uncertainty was the nearly absence of food samples from between February and May 1999, before the intensive food monitoring programs started. This was solved by allowing the worst case approach in simulation A. As far as the contribution of dioxin-like PCBs towards the total dioxin-like contamination is concerned, we had to rely mainly on background data from 2000. It is legitimate to do this because PCB profile analysis had shown that the nature of the contamination source was similar in background as in incident-related conditions.

Background daily intake of dioxin-like contaminants

Our intake assessment in adolescents at the 50th percentile, 2.53 pg TEQ/kg bw/d, is quite comparable to the mean daily intake of PCDDs, PCDFs, and dioxin-like PCBs in Dutch and US male and female adolescents, respectively 2.5 and 2.1 pg TEQ/kg bw/day (Patandin et al. 1999) and 3.5 and 2.7 pg/kg bw/day (Schecter et al. 2001). The European Scientific Committee on Food (2000) mentions an average intake of dioxin-like contaminants over different European countries of 1.2 to 3.0 pg TEQ/kg bw/day. Further comparisons, more especially with regard to higher percentiles of intake, are not possible because the data from the literature usually report averages and do not address the between-subject variation issue within a particular group.

Our dioxin intake simulation clearly identifies fish, milk and other dairy products as major sources of the Belgian background intake of dioxin-like contaminants. The question therefore arises whether the findings in this paper warrant any specific changes in dietary recommendations in the population at large or in particular subgroups{especially in pregnant women as adverse health effects in the foetus can occur at the lowest dioxin doses (WHO-ECEH-IPCS 2000). Further simulations are currently carried out to address this problem.

Incident related daily intake of dioxin-like contaminants

The median uncertainty curve of the simulated incident intake data represents the most likely estimate for the intake during the dioxin incident period, the 95th upper bound uncertainty curve a reasonable worst case. The uncertainty is markedly higher than in background conditions, due to the uncertainties inherently linked to food contamination data before the extent of the contamination became fully understood. The fraction of the adolescent population with an intake of less or equal to 1 pg TEQ/kg bw/day decreased from 3% towards 2% during the incident. The fraction of the population with an intake of less or equal to 4 pg TEQ/kg bw/day decreased from 85% towards 76% during the incident.

For most of the subjects, the moderate increase in intake during the incident follows the increase in contamination of chicken, egg, pork and beef. Fish remains an important input for the higher percentiles. Only from the 90th percentile on, highly contaminated chicken becomes more important as source to overtake fish at the highest percentiles.

Besides our simulation, only two other groups have tried to assess the possible increase in intake of dioxin-like substances during the Belgian dioxin incident (Bernard et al. 1999; Van Larebeke et al. 2001). Their estimations are not based on detailed food consumption data but on some, arbitrarily selected, food habits, and on the assumption, possible but not substantiated, that either 1 g TEQ of PCDD/F (Bernard et al. 1999) or 300 mg TEQ of PCDD/F and 900 mg TEQ of dioxin-like PCB (Van Larebeke et al. 2001) were introduced into the Belgian food chain at the start of the incident. By extrapolating dioxin concentrations and dioxin-like PCBs from a very limited number of samples, heavily contaminated by PCBs, to the total list of available food items in which marker PCBs have been measured (Van Larebeke et al. 2001), bias is introduced. Assuming overall high degrees of contamination in some particular food items, high exposure rates are obtained which might be true in some particular real cases, but extrapolation from there to the more general population remains rather questionable.

Body burden estimation and risk assessment

We only took into account the intake of dioxin-like contaminants via food. This introduces an underestimation of the body burden, that is, however, small since it is generally accepted that at least 90% of the intake of dioxin-like substances occurs via food. We further applied a single compartmental model with kinetic data from 2,3,7,8-TCDD- in stead of using more complex, physiologically based, pharmacological models that include absorption, distribution, binding and elimination kinetics (Van der Molen et al. 1996, Kreuzer et al. 1997). However, pharmacokinetic data for the different congeners of interest are incomplete and different parameters of the kinetic models are still fraught with uncertainties. We further started from a zero body burden, not taking into account the dioxin body burden at birth or the possible high intake with breast-feeding. A preliminary analysis, starting from literature body burden data at age 1, indicated that in that approach the equilibrium body burden was identical to the body burden we calculated. Finally, we applied the simulated intake of simulation A (adolescents) from age 0 to age 50. As said before, the limited recent food records for Belgian adults do not allow a similar intake assessment, but would, nevertheless suggest that intakes are similar or lower than in adolescents. An intake estimation of dioxin-like substances in US adolescents, adults and senior persons indeed demonstrated a tendency of lower intakes with

advancing age (Schecter et al. 2001). The body burdens estimated at the age of 50 years, therefore, remain a good estimate for most of the Belgian population.

Our background estimations are quite similar to background body burdens estimated by others. Equilibrium body burdens of 2.38 - 5.95 ng TEQ/kg bw can be calculated from the average European daily intakes for PCDD, PCDF and dioxin-like PCBs of 1.2 - 3 pg TEQ/kg bw/day (Scientific Committee on Food 2000). Extrapolation from dioxin-like activity in blood lipids from three different Flemish regions (Schoeters 2000, Staessen et al 2001){assuming 20% body fat-leads to average body burden values of 5.0 to 9.2 ng TEQ/kg bw for 18 years old, and 6.2 to 8.6 ng TEQ/kg bw for women at 50 to 65 years of age. These literature data, however, do not address the higher percentiles of their populations. This similarity provides an indirect validation of our probabilistic approach, supporting also our extrapolation of body burdens calculated on the basis of a daily intake in adolescents towards the total population. By switching from background intake towards the intake during the dioxin incident period, we implicitly assume that the individuals at the 50th, 95th and 97th percentile are the same in background conditions as during the dioxin incident period. This is not so, in fact the simulations in different conditions and at different levels of uncertainty are run independently. This means that, by assuming that the most exposed subjects in background conditions are also the most exposed subjects during the dioxin incident, we obtain an overestimation of the body burden and create a worst case situation.

The background and dioxin incident-related body burdens, estimated from median uncertainty intake simulations, lie around the body burdens that, in the experimental animal, are accompanied by biochemical and functional effects that may or may not lead to adverse effects{ e.g., CYP1A1/2 induction, EGF-receptor down-regulation and oxidative stress (3-10 ng TEQ/kg bw){and by increased viral sensitivity (10 ng TEQ/kg bw) (Scientific Committee on Food 2000, WHO-ECEH-IPCS 2000). Even in the worst case estimation, they remain below the body burdens that were accompanied in the experimental animal with observed increased incidence of adverse health effects, e.g., developmental neurotoxicity (25-37 ng TEQ/kg bw), reproductive toxicity (30-40 ng TEQ/kg bw), and tumor promoting effects (294 ng TEQ/kg bw). For those effects, threshold doses can be accepted below which the risk is thought to be negligible, without quantifying this extra low risk (Scientific Committee on Food 2000, WHO-ECEH-IPCS 2000). The incident-

related median body burden we simulated at median uncertainty, 3.9 ng/kg bw/d, is 5 to 22 times lower than median body burdens attained in Seveso residents{as extrapolated from the 50th percentile TCDD serum concentrations, lipid adjusted, among exposed in zone B (94 ppt) and A (447 ppt) and assuming 20% body lipids (Needham et al. 1997). The body burden estimated at the 99th percentile, under worst case conditions, e.g., 23.7 ng TEQ/kg bw, is 470 times lower than the body burden extrapolated from the highest Seveso serum concentrations (56000 ppt) (Needham et al. 1997). The range of Seveso exposure levels, substantially higher than the Belgian dioxin incident levels, are now accompanied by an increase in cancer mortality (Bertazzi et al. 2001). A trend of decreased male/female sex ratio became apparent in Seveso from a body burden of 16 ng TEQ/kg bw on (serum concentration of 80 ppt) (Mocarelli et al. 2000), which is within the range of body burdens estimated for the Belgian population at the higher levels of intake, both background and incident-related.

Others do not accept a threshold dose for tumor promoting agents like dioxins and calculate a cancer risk on the basis of a low dose extrapolation. Following the risk estimates from Becher and co-workers (Becher et al. 1998), a figure of up to 8,000 extra cancer deaths has been proposed as a result of the increased dioxin intake during Belgian dioxin incident (Van Larebeke et al. 2001). However, a prediction in a particular real life situation, e.g., a short time food contamination incident, based on upper bound risk estimates deduced from another quite different condition, high industrial exposures, remains questionable.

Non-dioxin-like PCBs

Our risk assessment is fully based on the quantitative exposure assessment towards dioxin-like contaminants. Non-dioxin-like PCBs were not included. The reason for that is 1) that the 1999 concentrations present in Belgian food items in background conditions hardly exist and 2) that the quantitative hazard assessment of non-dioxin-like PCBs has not been studied in depth as it has been done for the dioxin-like substances. On the basis, however, of a very preliminary analysis we would accept that the main hazard of this kind of food contamination resides in the dioxin-like substances.

8.6   Practical conclusions

We performed a probabilistic intake assessment of dioxin-like substances via food in adolescents, in background conditions and for the 1999 Belgian dioxin incident, leading towards the estimation of the variation in intake among the general population. During the 1999 dioxin incident, increases in daily intake were limited for most of the population, but potentially higher for those that regularly consumed highly contaminated chicken. As a result, body burdens of dioxin-like substances increased slightly above background levels, except at the highest percentiles of intake under worst case scenario conditions. Estimated background and incident-related body burdens are at levels where, in the experimental animal, biochemical and functional changes are observed but below levels that are accompanied by an increased incidence of adverse health effects. They are below levels observed in subjects exposed during the Seveso incident. It is unlikely that the 1999 Belgian dioxin incident will have an observable impact on public health, but it is not excluded that the incidence of adverse health effects might increase in particular subgroups of the population with some unusual food habits.

Discussion and further research

The introduction listed important methodological challenges which dose-response analysts face. A number of them were addressed in the 7 chapters of this thesis. The object under scrutiny is complex and controversial. Pathology, disease severity, the active molecule, its dose, its way of action, ... are so many factors that influence the patient's response to drug that we do not believe that a general analytic approach can fit all. We investigated separately different classes of problems motivated by real clinical studies.

In chapters 2 and 3, we started by analyzing time-varying exposure following nonadherence to prescribed drug therapy. We highlighted the need for detailed data on exposure and proposed statistical models to describe its evolution over time.

The phenomenon of nonadherence is often characterized as noise in clinical trials but, chapters 4 and 5 have shown that acknowledging natural variation in the exposure process is possible and can furthermore solve a general problem posed by nonlinear systems. Indeed, diversity in temporal patterns of input is needed to allow adequate characterization of input-output relations.

We applied this methodology in chapter 6 to a clinical trial assessing viral load decrease in HIV patients. We assume that adherence is not selective and derive a dose-response association through PK/PD models.

In chapter 7, we move on to seek a causal interpretation of dose-effect relationships in a situation where the assumption of nonselectivity is unlikely to hold. To this end, we extend existing methodology on structural models for trial with placebo control to the situation of active control trials.

In contrast to world of drug development, in an environmental setting, one face often the problem of lack of relevant exposure data. Chapter 8 was devoted to this type of exposures. We showed how to derive causal conclusions in this context acknowledging for potential uncertainties.

Here, we discuss the general ideas highlighted in previous chapters and give topics for future research.

Variation in exposure and clinical studies

To preserve the type one error of a randomized trial guidelines recommend and drug regulators require performance of a traditional intent to treat analysis. Hence patients are analyzed as randomized rather than on the basis of the amount of medication actually received. However, results from this type of analysis can be severely affected by the extent to which patients adhere with treatment. Intent to treat analysis risks underestimation of the true biological effect of treatment in the presence of suboptimal exposure. One solution lies in assessing patient adherence to prescribed therapy and judging if enough adherence has been achieved to support relevant conclusions of the trial. An other solution is to make efficient use of the observed variation in exposure to establish a dose-effect relationship. From the scientific point of view, there is a clear interest in accounting for partial adherence when analyzing and interpreting clinical trials. However, the bias protection offered by proper randomization must be respected as far as possible. Indeed, the potential bias introduced by considering partial compliance should be smaller than the bias resulting from the intent to treat analysis. In the next section, we consider the evaluation of adherence levels per se, after that we evaluate its link with outcome.

Assessing adherence to prescribed therapy

To draw inference on patient exposure, one must first define the xenobiotic substance in question and measure it. Exposure assessment may be highly diverse but this thesis supports the idea that reliability of exposure measurements over time is of prime importance. Several approaches towards assessing exposure were discussed in the introduction. We use more detail than pill counts can provide and have found the added information makes quite a difference. Therefore, we focus mainly on statistical methods to deal with data issued from electronic monitoring (MEMS, Aardex Ltd.) of patient's adherence. MEMS devices record precise timing of pill box openings but do not prove ingestion. While this method may underestimate exposure, it could be solved by new package design. With the advance of electronics more research should take place on electronic devices that would provide a count of the number of tablets removed. The size of the devices could be reduced and could include the recording of more specific behavioral details. For example,

one could record if the drug is taken during meals, in combination with other drugs, ...

We proposed several methods for the analysis of adherence as a time-varying pattern. However, we may be able to go further in the details. For example an important problem that was only partially tackled in this thesis is that the time window of maximum possible treatment efficacy must not stretch over the whole treatment period. We could go further by deriving adherence assessments that give more weight to observations located in a more relevant time period. Estimation of such weights is the topic of further research.

Schematically, we discussed the concept of adherence as a combined persistence phase and compliance process. It would be useful to further model those two different concepts jointly. Time until discontinuation (persistence) could be modeled as a time to event outcome. Then a relevant compliance measure that summarizes patient compliance from the start of the study until discontinuation could be derived and jointly modeled. A further extension of the two dimensional structure is to consider repeated discontinuations for a given patient, representing a patient who changes his decision over time: starts, stop, starts again, ... The main challenge here would be to define a rule to establish when a patient takes a new decision and breaks his daily routine.

In clinical trials, an analysis of adherence patterns should routinely be carried out. First it will indicate if reasonable adherence has been achieved or if it is worthwhile to investigate an adherence-effect relationship based on available data.

However, even with exhaustive analysis of time varying patterns of adherence, the difficulty remains to decide how much adherence is enough. It is obviously a matter that is disease-specific, drug-specific, formulation-specific and patient-specific.

In the current context of intent to treat analysis, researchers who want to validate their findings often aim to show that acceptable adherence was achieved in their clinical trial. Therefore, they could opt for an overestimation of adherence through pill counts and report an average adherence level over the whole study period ignoring individual timing patterns of exposure and pharmacologic properties of the drug.

Exposure as an explanatory variable in clinical trials

In a clinical trial, if adherence to prescribed drug therapy is variable and is accurately recorded together with outcome, we can investigate how treatment effect depends on the amount and timing of treatment actually received. Unbiased answers are hard to derive due to the possibly selective nature of patient adherence to prescribed therapy. Patients with different adherence patterns can have different outcomes for many reasons besides variability in the amount and timing of drug administration. By studying the association between observed adherence and outcome from empirical data, we still ignore the true effect of `a forced dose' to the patient. Indeed, study of the dose-effect associations in empirical data ignores potential selectivity (as defined in chapter 7). This type of analysis can only be interpreted as causal if the selection mechanisms are not present. A detailed analysis of the study adherence would probably deliver important indications for potential selectivity. Whatever the selective nature, statistical techniques are essential to disclose information hidden in the outcome data. They allow complex modeling of the pharmacological process. If a strong association is found between adherence and effect, further analyses to identify the causal effect would be worthwhile.

In the presence of selectivity, structural models allow one to estimate the average causal effect of the drug for differing degrees of exposure. Specifically it is possible to estimate the average difference in effect between two randomized groups for somebody who would be a perfect adherer with one of the assignments. As the data contain rather weak information on causality in the presence of unspecified selectivity, one is typically limited to rather simple structural models. A linear relation between dose and effect is often assumed. Current state of causal inference methodology has thus other limitations.

In this context, it is getting increasingly difficult to choose between methodological and biological complexity. However, the validity of the methodological approach can in no case be better than the quality of the data themselves. Having conceptually split adherence into two dimensions, persistence and compliance, allows us to refine the selectivity discussion. Results presented in this thesis indicate that the compliance of some patients is intrinsically better that of others. These results support the idea that once a patient accepts to take a drug, he enters into a daily routine closely linked to his lifestyle. Future studies enabling simultaneous monitoring of several drugs, including a placebo, will help

clarify those ideas. The end of the continuous process occurs when the patient takes a new decision. The discontinuation can eventually be influenced by the drug itself (adverse reactions, sickness improvement, relapses, ...) and therefore is probably not an attribute of the patient. Therefore, to avoid selectivity it is absolutely essential to maintain persistence in the trial as long as possible,which is especially true when nonpersistence with drug therapy results in trial drop out and no further evaluation of the patient is available. Such type of missing value are clearly not missing at random and could further bias the interpretation of the results. Some preliminary work has appeared in the literature on how to treat the simultaneous occurrence of non adherence and missing outcomes.

Variation in exposure and PK/PD studies

Population pharmacokinetic/pharmacodynamic studies represent one specific area where ignoring potential selectivity is the rule, not the exception. In such clinical trials plasma concentration is an uncontrolled covariate and can thus be selective, if for example, fast metabolizer happen to be the poor adherer to start with: reason enough to identify important source of variation in the process and especially actual drug exposure. This could then help to identify subgroups responsible for the selectivity by investigating the relations between adherence to prescribed therapy and PK parameters. This type of studies tends to take place in early/mid phase of drug development and thus one still wants to understand the process of action rather than only proving a population average effect. Selectivity problem in PK/PD studies could be solved by using concentration-controlled trials (CCT) as proposed by Sanathanan et al. (1991). In this type of trial, subjects are randomized to predetermined levels of average plasma drug concentrations. In addition to safety concerns which strongly suggest the use of CCTs for drugs with narrow therapeutic windows, sample size and power considerations favor the choice of CCTs in many situations. CCTs are designed to minimize the inter individual PK variability within comparison groups and consequently decrease the variability in clinical response within these groups. One can then incorporate a titration design within the exit rule, to provide sufficient information on the concentration-response relationship. For example increasing the concentration level of the placebo group and decreasing the concentration level of the treated group. Prac-

tical constraints have, however, hampered the implementation of these designs.

Exposure in environmental studies

Since non adherence is a post randomization variable, it imparts to a controlled clinical trial some of the characteristics of an observational study, where the subjects themselves decide about their exposure to the drug. However observational studies face additional difficulties. Compared to clinical trials usually they face more often exposure to multiple xenobiotic substances. Since it is typically impossible to monitor them all, it is current practice to assess some of them, assuming they constitute a surrogate for the total exposure. Because substances with potential toxic effect on health show a long half life, one tends to assess exposure through a single measurement of blood concentration in time. This approach can then not distinguish between short term contamination and long term background exposure and can not identify potential sources of exposure. In addition in environmental studies, when an incident is the focus, there is often a huge imbalance between the exposure assessment before the incident is identified compared to after. We show here that the study of toxic exposure could benefit greatly from a detailed record of exposure over time.

Simulation

Simulation was considered in chapter 8 to accommodate uncertainty in exposure measurements in the context of an environmental study. In drug development the process of action following exposure is much better studied and a lot of data is available via well controlled clinical trials. Nevertheless, we believe the methodology of two dimensional simulations can provide a new tool for instance to communicate with clinicians about the outcome to expect from the trial facing explicitly the uncertainties inherent to the assumptions. This approach can then be used to bridge the gap of transfer of knowledge between earlier phases of drug development and the clinical development of the product acknowledging potential uncertainty on the bridging links. A well established PK/PD model can be used to simulate clinical trials. Clinical trial simulation requires a multi disciplinary effort to put the information together. Traditional sample size and power analysis are part of what one can expect

from a clinical trial simulation. However, the true value of this method goes much deeper, allowing to evaluate a broader variety of \what if" situations and estimate many performance measures quantitatively, rather than address simple yes/no questions. \What if the patient does not take the drug as prescribed in an intent to treat trial?" and thus the impact of nonadherence on trial design (e.g. sample size) can be investigated via simulation. Such investigation could be done by using adherence patterns previously recorded in similar settings. Association between dosing patterns and responses, responsible for potential selectivity, could then be propagated in the uncertainty dimension, leading to more appropriate interpretation and more realistic sample size estimation. Upon identification of substantial uncertainty with possible serious consequences, new study designs should be contemplated. This technology is pointing far beyond what we now consider power analysis, perhaps towards a second or a third generation of power analysis. As these tools grow in reliability, acceptance, and sophistication we will come to expect much more of \power analysis".

Conclusion

In conclusion, we found that the help of careful statistical analysis can yield a much better understanding of xenobiotic levels in practice and their causal effect on outcome. The work of this thesis added new tools for several classes of problems which we trust are useful in practice. We are however well aware that this is not the end of the road. The ultimate challenge lies in combining, if possible, the sometimes complex biologically relevant models with causal analysis of practical clinical effects. While we made a start with this, several open problems await further research.

Bibliography

[1] Aarons L, Balant LP, Mentré F, Morselli PL, Rowland M, Steimer J-L and Vozeh S (1996) Practical experience and issues in designing and performing population pharmacokinetic/pharmacodynamic studies. Eur J Clin Pharmacol, 49, 251-254.

[2] Agresti A (1990) Categorical data analysis. John Wiley and sons.

[3] Antal EJ, Grasela TH, Smith RB (1989) An evaluation of population pharmacokinetics in therapeutic trials. Part III. Prospective data collection versus retrospective data assembly. Clin Pharmacol Ther, 46 (5), 552-559.

[4] Arnsten J, Demas P, Farzadegan H, Grant R, Gourevitch M, Chang C, Buono D, Eckholt H, Howard A, Schoenbaum E (2001) Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. Clinical Infectious Diseases, 33, 1417-23.

[5] Becher H, Steindorf K, Flesch-Janys D (1998) Quantitative cancer risk assessment for dioxins using an occupational cohort. Environ Health Perspect, 106 Suppl 2, 663-670.

[6] Bernard A, Hermans C, Broeckaert F, De Poorter G, De Cock A, Houins G (1999) Food contamination by PCBs and dioxins. An isolated episode in Belgium is unlikely to have affected public health. Nature, 401, 231-232.

[7] Bertazzi PA, Consonni D, Bachetti S, Rubagotti M, Baccarelli A, Zocchetti C, Pesatori AC (2001) Health Effects of Dioxin Exposure: A 20-Year Mortality Study. Am J Epidemiol, 153, 1031-1044.

[8] Bonney GE (1987) Logistic regression for dependent binary observations. Biometrics, 43, 951-973.

[9] Bovet P, Madeleine G, Perret F, Gervasoni JP, Paccaud F and Waeber B (1997) One-year compliance to medication in newly diagnosed hypertensive patients in a developing country. Can J Cardiol, 13 (Suppl B), 206B.

[10] Carey V, Zeger S, Diggle P (1994) Modeling multivariate binary data with alternating logistic regressions. Biometrika, 80, 517-526.

[11] Claxton AJ, Cramer J and Pierce C (2001) A systematic review of the associations between dose regimens and medication compliance. Clinical Therapeutics, 23 (8), 1296-1310.

[12] Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. J Am Stat Ass,74, 829-836.

[13] Cook RD, Weisberg S. Residuals and influence in regression. Chapman and Hall: New York, 1995.

[14] Cramer JA, Mattson RH, Prevey ML, Scheyer RD, Ouelette VL (1989) How often is medication taken as prescibed ? A novel assessment technique. JAMA, 261, 3273-3277.

[15] Cramer JA, Scheyer RD, Mattson RH (1990) Compliance declines between clinic visits. Arch Intern Med, 150, 1509-1510.

[16] Cramer JA, Spilker B (1991) Patient compliance in medical practice and clinical trials. New-York: Raven Press.

[17] Cramer JA, Spilker B (1992) Non-compliance in clinical trials. New-York: Raven Press.

[18] Davidian M and Giltinan DM (1995) Nonlinear models for repeated measurement data. London: Chapman and Hall.

[19] De Henauw S, Maithys C, De Backer G (2001) Difference in overall food and nutrient intake-profile between breakfast users and breakfast skippers in a representative sample of 14-18 years old Belgian adolescents. Public Health Nutrition, In press.

[20] Delta Coordinating Committee and Delta Virology Committee (1999) HIV-1 RNA response to antiretroviral treatment in 1280 participants in the Delta trial: An extended virology study. AIDS, 13, 57-65.

[21] Diggle PJ, Liang KY and Zeger QL (1994) Analysis of longitudinal data. Oxford Science Publication

[22] Dioxin Body Burden Working Group (2001) The Belgian PCB-dioxin incident 1999. Report presented before the Belgian Health Council, 7300/1, Brussels, Belgium.

[23] Donnelly R, Meredith PA, Miller SHK, Howie CA and Elliott HL (1993) Pharmacodynamic modeling of the anti hypertensive response to amlopidine. Clinical Pharmacology and Therapeutics, 54 (3), 303-310.

[24] Dos Santos DM, Berridge DM (2000) A continuation ratio random effect model for repeated ordinal responses. Statistics in Medicine, 19, 3377-3388.

[25] Duffull SB, Mentré F, Aarons L (2001) Optimal design of a population pharmacodynamic experiment for ivabradine. Pharmaceutical Research, 18, 83-89.

[26] Efron B and Feldman D (1991) Compliance as an explanatory variable in clinical trials. Journal of the American Statistical Association, 86 (413), 9-17.

[27] Fischer-Lapp K and Goetghebeur E (1999) Practical properties of some structural mean analyses of the effect of compliance in randomized trials. Controlled Clinical Trials, 20 (6), 531-546.

[28] Focant J-F, Eppe G, Houziaux J-S, Xhrouet C, Andre J-E, Dipede D, De Pauw E (2000) Contribution and importance of non-ortho (coplanar) PCBs for the I-TEQ evaluation in "dioxin analysis" of biological matrices. Organohalogen Compounds, 48, 312-325.

[29] George CF, Peveler RC, Heiliger S, Thompson C (2000) Compliance with tricyclic antidepressants: The value of four different methods of assessment. British Journal of Clinical Pharmacology, 50, 166-171.

[30] Gibaldi M and Perrier D (1982) Pharmacokinetics. New-York: Marcel Dekker Inc.

[31] Gilligan DM, Chan WL, Stewart R and Oakley CM (1991) Adrenergic hypertensivity after beta-blocker withdrawal in hypertropic cardiomyopathy. American Journal of cardiology, 68, 766-772.

[32] Gilks WR, Richardson S, Spiegelhalter DJ. Markov chain monte Carlo in practice. Chapman and Hall: London, 1996.

[33] Girard P, Sheiner LB, Kastrissios H and Blaschke TF (1996) Do we need full compliance data for population pharmacokinetic analysis ?. Journal of Pharmacokinetics and Biopharmaceutics, 24 (3), 265-282.

[34] Girard P, Blaschke TF, Kastrissios H, Sheiner LB. A Markov mixed effect regression model for drug compliance. Statistics in Medicine 1998; 17 (20): 2313-2334.

[35] Goetghebeur E, Pockok S (1993) Statistical issues in allowing for noncompliance and withdrawal. Drug Inf J, 27, 837-845.

[36] Goetghebeur E, Lapp K (1997) The effect of treatment compliance in a placebo-controlled trial: regression with unpaired data. Applied Statistics 1997, 3, 351-364.

[37] Grasela TH, Fiedler-Kelly JB, Salvadori C, Marey C, Jochemsen R, Loo H (1993) Predictive performance of population pharmacokinetic parameters of tianeptine as applied to plasma concentrations from a post-marketing study. Eur J Clin Pharmacol, 45, 123-128.

[38] Hardman JG, Limbird LE, Molinoff PB, Ruddon RW, Gilman AG (2001) Goodman and Gilman's The Pharmacological Basis of Therapeutics, 9th ed. New York: Mc Graw Hill, Table A-II-1, pp. 1712-1792.

[39] Harter JG, Peck CC (1991) Chronobiology: Suggestions for integrating it into drug development. Annals of the New York Academy of Sciences, 618, 563-571.

[40] Holford NHG and Sheiner LB (1981) Understanding the dose-effect relationship: Clinical application of pharmacokinetic-pharmacodynamic models. Clinical Pharmacokinetics, 6, 429-453.

[41] Houston MC and Hodge R (1988) Beta-adrenergic blocker with withdrawals syndromes in hypertension and other cardiovascular diseases. American Heart Journal, 116, 515-523.

[42] Hughes MD, Johnson VA, Hirsch MS, et al. (1997) Monitoring plasma HIV-1 RNA levels in addition to CD4+ lymphocyte count improves assessment of antiretroviral therapeutic response. Ann Intern Med, 126, 929-938.

[43] Jerling M, Merle Y, Mentré F, Mallet A (1994) Population pharma-cokinetics of nortriptyline during monotherapy and during concomitant treatment with drugs that inhibit CYP2D6- an evaluation with the nonparametric maximum likelihood method. Br J clin Pharmac, 38, 453-462.

[44] Kass MA, Meltzer DW, Gordon MG, Cooper D, Goldberg J (1986) Compliance with topical pilcarpine treatment. Am J Ophth, 101, 515-523.

[45] Kastrissios H, Blaschke TF (1997) Medication compliance as a feature in drug development. Annu. rev. Pharmacol. Toxicol., 37, 451-475.

[46] Kleinbloesem CH, van Brummelen P, Danhof M, Faber H, Urquhart J and Breimer DD (1987) Rate of increase in the plasma concentration of nifedipine as a major determinant of its hemodynamic effects in humans. Clinical Pharmacol Ther, 41, 26-30.

[47] Kornitzer M, Bura L (1989) Clinical and anthropometric data, blood chemistry and nutritional patterns in the Belgian population according to age and sex. For the B.I.R.N.H. Study group. Acta Cardiologica, 44, 101-144.

[48] Kreuzer PE, Csanady GA, Baur C, Kessler W, Papke O, Greim H, Filser J G (1997) 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and congeners in infants. A toxicokinetic model of human lifetime body burden by TCDD with special emphasis on its uptake by nutrition. Arch Toxicol, 71, 383-400.

[49] Kruse W, Weber E (1990) Dynamics of drug regimen compliance: its assessment by microprocessor-based monitoring. Eur J Clin Pharmacol, 38, 561-565.

[50] Lee J, Ellenberg J, Hirtz D, Nelson K. Analysis of Clinical Trials by Treatment Actually Received: Is it Really an Option ? Statistics in Medicine 1991; 10: 1595-1605.

[51] Levy G (1993) A pharmacokinetic perspective on medicament non-compliance. Clinical Pharmacology and Therapeutics, 54 (3), 242-244.

[52] Lim LL-Y. Estimating compliance to study medication from serum drug levels: Application to an AIDS clinical trial of Zidovudine. Biometrics 1992; 48: 619-630.

[53] Liem AKD, Theelen RMC (1997) Dioxins: chemical analyssis, exposure and risk assessment. PhD. Thesis, University Utrecht, The Netherlands.

[54] Lindsey JK, Jones B, Ebbutt AF (1997) Simple models for repeated ordinal responses with an application to a seasonal rhinitis clinical trial. Statistics in Medicine, 16, 2873-2882.

[55] Lindstrom MJ and Bates DM (1990) Nonlinear mixed effects models for repeated measures data. Biometrics, 46, 673-687.

[56] Lipsitz SR, Laird N, Harrington DP (1991) Generalized estimating equations for correlated binary data using the odds ratio as a measure of association. Biometrika, 78, 153-160.

[57] Lunn DJ, Aarons LJ (1997) Markov chain monte carlo techniques for studying interoccasion and intersubject variability: application to pharmacokinetic data. Appl. Statist., 46 (1), 73-91.

[58] Marschner IC, Collier AC, Coombs RW et al. (1998) Use of changes in plasma levels of Human Immunodeficiency Virus Type 1 RNA to assess the clinical benefit of antiretroviral therapy. Journal of Infectious Diseases, 177, 40-47.

[59] Mellors JW, Rinaldo CR, Gupta P et al. (1996) Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science, 272, 1167-1170.

[60] Mentré F, Mallet A and Baccar D (1997) Optimal design in random-effects regression model. Biometrika, 84 (2), 429-442.

[61] Mocarelli P, Gerthoux PM, Ferrari E, Patterson DGJ, Kieszak SM, Brambilla P, Vincoli N, Signorini S, Tramacere P, Carreri V, Sampson EJ, Tumer WE, Needham LL (2000) Paternal concentrations of dioxin and sex ratio of offspring. Lancet, 355, 1858-1863.

[62] Montaner JS, Hogg R, Raboud J, Harrigan R, O'Shaughnessy M (1998) Antiretroviral treatment in 1998. The Lancet, 352, 1919-1922.

[63] Nationaal Instituut voor de Statistiek (2000) Landbouwstatistieken, landbouw- en tuinbouwtelling op 15 mei 1999. Ministerie van Economische Zaken, Brussels, Belgium.

[64] Needham LL, Gerthoux PM, Patterson DGJ, Brambilla P, Turner WE, Beretta C, Pirkle JL, Colombo L, Sampson EJ, Tramacere PL, Signorini S, Meazza L, Carreri V, Jackson RJ, Mocarelli P (1997) Serum dioxin levels in Seveso, Italy, population in 1976. Teratog Carcinog Mutagen, 17, 225-240.

[65] NEVO (1996) NEVO Tabel. Nederlands voedingsstoffenbestand, Velotekst, Den Haag, The Netherlands.

[66] Nony P, Cucherat M, Boissel J-P (1998) Revisiting the effect compartment through timing errors in drug administration. Tips, 19, 49-54.

[67] NUBEL (1999) Belgische voedingsmiddelentabel, Brussels, Belgium.

[68] Patandin S, Dagnelie PC, Mulder PGH, Op de Coul E, van der Veen JE, Weisglas-Kuperus N, Sauer PJJ (1999) Dietary exposure to polychlorinated biphenyls and dioxins from infancy until adulthood: a comparison between breast-feeding, toddler, and longterm exposure. Environ Health Perspect, 107, 45-51.

[69] Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squler C, Wagener MM, Singh N (2000) Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Journal of Internal Medicine, 133, 21-30.

[70] Petri H, Urquhart J (1994) Patient compliance with beta-blocker medication in general practice. Pharmacoepidemiol and Drug Safety, 3, 251-256.

[71] Pocock SJ, Abdalla M (1998) The hope and the hazards of using compliance data in randomized controlled trials. Statistics in Medicine, 17, 303-317.

[72] Prentice R (1988) Correlated binary regression with covariates specific to each binary observation. Biometrics, 44, 1033-1048.

[73] Psaty BM, Koepsell TD, Wagner EH, LoGerfo JP and Inui TS (1990) The relative risk of incident coronary heart disease associated

with recently stopping the use of beta blockers. JAMA, 263, 1653-1657.

[74] Reid JL and Meredith PA (1990) Concentration-Effect analysis of anti hypertensive drug responses. Hypertension, 16, 12-18.

[75] Reuven Y. Rubinstein (1981) Simulation and the Monte Carlo Method. Wiley.

[76] Robins JM (1994) Correcting for noncompliance in randomized trials using structural nested mean models. Communs Statist. Theory Meth., 23, 2379-2412.

[77] Rosenbaum JF, Fava M, Hoog SL, Ascroft RC and Krebs WB (1998) Selective serotonin reuptake inhibitor discontinuation syndrome: a randomized clinical trial. Biol. Psychiatry, 44, 77-87.

[78] Rosenberg M and Waugh MS (1999). Causes and consequences of oral contraceptive noncompliance. Am. J. Obstet. Gynecol., 180, 276-279.

[79] Rubio A, Cox C, Weintraub M (1992) Prediction of diltiazem plasma concentration curves from limited measurements using compliance data. Clinical Pharmacokinetics, 22 (3), 238-246.

[80] Sanathanan LP, Peck C, Temple R, Lieberman R and Pledger G (1991) Randomization, PK-controlled dosing, and titration: an integrated approach for designing clinical trials. Drug Information Journal, 25, 425-431.

[81] Schecter A, Cramer P, Boggess K, Stanley J, Papke O, Olson J, Silver A, Schmitz M (2001) Intake of dioxins and related compounds from food in the U.S. population. Journal Toxicol Environ Health A, 63, 1-18.

[82] Schoeters G (2000) Samenvatting van het toxicologie luik 'Milieu en gezondheid'. VITO, Mol, Belgium.

[83] Scientific Committee on Food (2000) Opinion of the SCF on the risk assessment of dioxins and dioxin-like PCBs in food. SCF/CS/CNTM/DIOXIN/8 Final, p.1-141, Brussels, Belgium.

[84] Sheiner LB, Ludden TM (1992) Population pharmacokinetics/pharmacodynamics. Ann Rev Pharmacol. Toxicol., 32, 185-209.

[85] Sheiner LB, Rubin DB (1995) Intention to treat analysis and the goals of clinical trials. Clin Pharmacol Ther, 57, 6-15.

[86] Smith DM, Diggle PJ (1998) Compliance in an anti-hypertension trial: A latent process model for binary longitudinal data. Statistics in Medicine 1998, 17 (3), 357-370.

[87] Sommer A, Zeger SL (1991) On estimating efficacy from clinical trials. Stat in Med, 10, 45-52.

[88] Souci SW, Fachmann W, Kraut H (2000) Souci Fachmann Kraut Food composition and nutrition tables. CRC Press, London, UK.

[89] Staessen JA, Nawrot T, Hond E D, Thijs L, Fagard R, Hoppenbrouwers K, Koppen G, Neten V, Schoeters G, Vanderschueren D, Van Hecke E, Verschaeve L, Vlietinck R, Roels HA (2001) Renal function, cytogenetic measurements, and sexual development in adolescents in relation to environmental pollutants: a feasibility study of biomarkers. Lancet, 357, 1660-1669.

[90] Stephenson J (1999) Noncompliance may cause half of anti hypertensive drug 'failures'. J. Am. Med. Ass., 282, 313-314.

[91] Urquhart, J (1994) Role of patient compliance in clinical pharmacokinetics: a review of recent research. Clinical Pharmacokinetics, 27 (3), 202-215.

[92] Urquhart, J (1997) The electronic medication event monitor. Clinical Pharmacokinetics, 32 (5), 345-356.

[93] Urquhart J (1998) Variable patient compliance as a source of variability in drug response. Excerpta Medica Congress Series, 1178, 188-198.

[94] Urquhart J, De Klerk E (1998) Contending paradigms for the interpretation of data on patient compliance with therapeutic drug regimens. Statistics in medicine, 17, 251-267.

[95] Urquhart J (2001) Demonstrating effectiveness in a post placebo era. Clin Pharmacol Ther, 70, 115-120.

[96] Urquhart J, Vrijens B (2001) Taxonomy of patient compliance-related events in drug trials Abstract of paper presented at the Emory University Conference on Causal Inference, Snowbird Utah, August 10, 2001.

[97] van den Berg M, Birnbaum L, Bosveld BTC, Brunstrom B, Cook P, Feeley M, Giesy JP, Hanberg A, Hasegawa R, Kennedy SW, Kubiak T, Larsen JC, van Leewen FXR, Liem AKD, Nolt C, Peterson RE, Poellinger L, Safe S, Schrenk D, Tillitt D, Tysklind M, Younes M, Waern F, Zacharewski T (1998) Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ Health Perspect, 106, 775-792.

[98] Van Der Molen GW, Kooijman SALM, Slob W (1996) A generic toxicokinetic model for persistent lipophilic compounds in humans: an application to TCDD. Fundam Appl Toxicol, 31, 83-94.

[99] Vander Stichele R (1991) Measurement of patient compliance and the interpretation of randomized clinical trials. European Journal of Clinical Pharmacology, 41, 27-35.

[100] Vanhove GF, Schapiro JM, Winters MA, Merigan TC, Blaschke TF (1996) Patient compliance and drug failure in protease inhibitor monotherapy. Journal of the American Medical Association, 276 (24), 1955-1956.

[101] Van Larebeke N, Hens L, Schepens P, Covaci A, Baeyens J, Everaert K, Bernheim JL, Vlietinck R, De Poorter G (2001) The Belgian PCB and dioxin incident of january-june 1999: exposure data and potential impact on health. Environ Health Perspect, 109, 265-273.

[102] Voedingscentrum (1998) Zo eet Nederland. Resultaten van de Voedselconsumptiepeiling 1997-1998. ISBN 90 5177 036 7, Voedingscentrum Den Haag, The Netherlands.

[103] Vozeh S, Syeimer J-L, Rowland M, Morselli P, Mentre F, Balant LP, Aarons L (1996) The use of population pharmacokinetics in drug development. Clinical Pharmacokinetic, 30 (2)2, 81-93.

[104] Vrijens B and Goetghebeur E (1997) Comparing compliance patterns between randomized treatments. Controlled Clinical Trials, 18, 187-203.

[105] Vrijens B and Goetghebeur E (1999) The impact of compliance in PK studies. Statistical Methods in Medical Research, 8, 247-262.

[106] Vrijens B and Goetghebeur E (1999) On predicting compliance from its past. University of Ghent - Center for Statistics. Technical report (1999/1)

[107] Vrijens B and Goetghebeur E (2000) Irregular drug intake reduces bias and improves precision in PK/PD population studies. University of Ghent - Center for Statistics. Technical report (2000/1)

[108] Vrijens B, De Henauw S, Dewettinck K, Talloen W, Goeyens L, De Backer G, Willems J (2002) Probabilistic intake assessment and body burden estimation of dioxin-like substances in background conditions and during a short food contamination episode. Food Additives and Contaminants, accepted.

[109] Wang W, Husan F, Chow S-C (1996) The impact of patient compliance on drug concentration profile in multiple doses. Statistics in Medicine, 15, 659-669.

[110] Wang W, Ouyang SP (1998) The formulation of the principle of superposition in the presence of non-compliance and its applications in multiple dose pharmacokinetics. J Pharmacokinet. Biopharm., 26(4), 457-469

[111] Waterhouse D, Calzone C, Brenner D (1993) Adherence to oral tamoxifen: a comparison of patient self-report, pill counts, and microelectronic monitoring. Proceedings Amer Soc Clin Oncology, 12, 167, abstract 464.

[112] White IR, Pocock SJ (1996) Statistical reporting of clinical trials with individual changes from allocated treatment. Statistics in Medicine, 15, 249-262.

[113] White IR, Goetghebeur EJT (1998) Clinical trials comparing two treatment policies: Which aspects of the treatment policies make a difference ? Statistics in Medicine, 17, 319-339.

[114] WHO-ECEM-IPCS (2000) Consultation on assessment of the health risks of dioxins; re-evaluation of the tolerable daily intake (TDI): Executive summary. Food Addit Contam, 17, 223-240.

[115] Wood D, De Backer G, Faergeman O, Graham I, Mancia G, Pyorala K (1998) Prevention of coronary heart disease in clinical practice: recommendations of the Second Joint Task Force of European and other Societies on coronary prevention. Atherosclerosis, 140, 199-270.

[116] Zhao LP, Prentice R (1990) Correlated binary regression using a quadratic exponential model. Biometrika, 77, 642-648.

Samenvatting

Deze thesis handelt over het analyseren van tijdsafhankelijke bloot-stellingspatronen en hun impact op uitkomsten relevant voor de menseli-jke gezondheid. Terwijl lange tijd de precieze blootstelling van mensen aan heel wat stoffen verborgen bleef zijn in recente jaren gesofisticeerde technieken op de markt gekomen die ons met een weelde aan gede-tailleerde data overladen. De statistische methoden om uit die data relevante informatie te distilleren moesten echter nog grotendeels on-twikkeld worden en vormen het voorwerp van deze thesis. In de kontekst van klinische studies analyzeren we enerzijds blootstellingspatronen op zich die we vervolgens voorspellen, anderzijds brengen we die bloot-stellingen in verband met gezondheidsuitkomsten. We vormen besluiten over de associatie-patronen en geven aan hoe men de stap naar causale besluitvorming kan zetten. In tweede instantie zijn we niet blind voor het feit dat in minder gecontroleerde omstandigheden, bijvoorbeeld bij een milieu-crisis, data over blootstelling vaak veel schaarser en minder goed gemeten zijn. In zo'n situatie is het noodzakelijk om de onzek-erheid over het blootstellingsniveau expliciet onder ogen te zien en op inzichtelijke wijze te betrekken in een statistische evaluatie van de bloot-stelling en haar risico's. Ook voor zo'n situatie stellen we een algemene analysemethode voor. We passen ze toe op beschikbare gegevensbron-nen voor de Belgische dioxine-crisis in 1999. Een bondig overzicht van de verschillende hoofdstukken volgt hieronder.

In hoofdstuk 2 hebben we voor verschillende variabelen samengevat welk aspect van variabele therapietrouw ze belichten en onderzocht in welke context zij van belang kunnen zijn. Daarnaast onderzoeken we condi-tionele en marginale modellen om zo herhaalde metingen van dageli-jkse binaire blootstelling tussen groepen te kunnen vergelijken. Het hoofdstuk benadrukt de noodzaak om te beschikken over gedetailleerde gegevens om zo krachtiger statistische technieken te kunnen ontwikkelen

voor de vergelijking van gerandomizeerde groepen over de tijd.

In hoofdstuk 3 stellen we methoden voor om tijdsafhankelijke therapi-
etrouw te voorspellen. Hiervoor werd een auto-regressief model voor bi-
naire data ontwikkeld. We behandelen 'frailty' met behulp van een nor-
male verdeling voor random 'slopes' met een random 'intercept'. Mod-
elvalidatie en twee verschillende types klinisch relevante voorspellingen
worden beschreven en bediscussieerd.

In hoofdstuk 4 introduceren we de basis van populatie-pharmacokinetiek
en bespreken de invloed van variabele therapietrouw als grootste bron
van variatie in pharmaco-kinetische respons. We kwantificeren de im-
pact van geobserveerde variabiliteit in populatie-farmacokinetische stud-
ies met behulp van simulatie technieken. We ontwikkelen een methode
om een gedetailleerd blootstellingspatroon en onregelmatige uitkomsten
simultaan te analyseren. We concluderen dat wanneer het mogelijk is
om onregelmatige inname van een geneesmiddel te observeren, een sub-
stantile toename in de precisie van de schatting van pharmaco-kinetische
parameters kan worden bereikt.

In hoofdstuk 5 worden pharmaco-dynamische parameters geschat op ba-
sis van slechts enkele uitkomstmaten in een groep patiënten die variabele
therapietrouw vertonen. We onderzoeken de informatie matrix voor
hiërarchische non-lineaire modellen en stellen vast dat er tegelijk een
significante reductie in vertekening en toename in precisie bereikt kan
worden door het observeren van onregelmatige inname. Dit hoofdstuk
bevestigt de claim dat variabele therapietrouw, als natuurlijk experi-
ment in variatie van doseren, een verbetering kan betekenen vanuit het
informatie standpunt. In het algemeen kan men stellen dat de schattin-
gen van pharmaco-kinetische en pharmaco-dynamische parameters fors
kunnen verbeteren indien de informatie over onregelmatige inname goed
gebruikt wordt.

In hoofdstuk 6 onderzoeken we het effect van therapietrouw op respons
onder de aanname dat therapietrouw niet selectief is. Als eerste stap
schatten we de individuele pharmaco-kinetische parameters met behulp
van een n-compartiment model, aangepast voor geobserveerde patronen
van inname. Daarna relateren we de farmacologische werking aan de
plasmaconcentraties. We passen deze methodologie toe in een HIV on-
derzoek als een eerste stap in de richting van een geïndividualiseerde
therapie voor de patiënt. Tenslotte ontwikkelen we een nieuwe vari-
abele om therapietrouw samen te vatten, geïnspireerd op het derde mo-
ment van de interdosis-interval verdeling, die toelaat om de relatie tussen

doseringsgeschiedenis en respons direct te schatten.

In hoofdstuk 7 bekijken we het werkelijke biologische effect van het geneesmiddel indien potentiële selectiviteit verwacht wordt. De methodologie spitst zich toe op causale inferentie in gerandomiseerd onderzoek bij vergelijking van twee behandelingen. De methode wordt toegepast op een studie met antidepressiva waar een sterk psychologisch effect verwacht kan worden op respons.

In hoofdstuk 8 introduceren we twee dimensionale Monte Carlo simulaties als techniek om data uit verschillende bronnen te combineren. Tussen-subject variabiliteit en onzekerheid door niet optimale datacollectie of een vertekend steekproefplan worden geëvalueerd onder een extreem en minder extreem scenario. Simulatie is daarbij geen substituut voor wiskunde maar levert een benadering voor het rekenwerk, vergelijkbaar met de manier waarop de bootstrap methode de variantie van een gecompliceerde puntschatting helpt berekenen. De simulatietechniek geeft ons meer mogelijkheden om complexe causale relaties tussen blootstelling en uitkomst te evalueren.

Tenslotte bespreken we in het laatste hoofdstuk enkele algemene ideeën die in de hele thesis terugkomen en verwijzen naar onderwerpen voor verder onderzoek.