Cancer Chemother Pharmacol (1993) 33: 1–9

Cancer Chemotherapy and Pharmacology

Review

Statistical approaches to pharmacodynamic modeling: motivations, methods, and misperceptions

Rosemarie Mick, Mark J. Ratain

Section of Hematology/Oncology, Department of Medicine, Cancer Research Center and Committee on Clinical Pharmacology, University of Chicago, Pritzker School of Medicine, Chicago, IL 60637, USA

Received 18 December 1992/Accepted 13 May 1993

Summary. We have attempted to outline the fundamental statistical aspects of pharmacodynamic modeling. Unexpected yet substantial variability in effect in a group of similarly treated patients is the key motivation for pharmacodynamic investigations. Pharmacokinetic and/or pharmacodynamic factors may influence this variability. Residual variability in effect that persists after accounting for drug exposure indicates that further statistical modeling with pharmacodynamic factors is warranted. Factors that significantly predict interpatient variability in effect may then be employed to individualize the drug dose.

In this paper we have emphasized the need to understand the properties of the effect measure and explanatory variables in terms of scale, distribution, and statistical relationship. The assumptions that underlie many types of statistical models have been discussed. The role of residual analysis has been stressed as a useful method to verify assumptions. We have described transformations and alternative regression methods that are employed when these assumptions are found to be in violation. Sequential selection procedures for the construction of multivariate models have been presented. The importance of assessing model performance has been underscored, most notably in terms of bias and precision.

In summary, pharmacodynamic analyses are now commonly performed and reported in the oncologic literature. The content and format of these analyses has been variable. The goals of such analyses are to identify and describe pharmacodynamic relationships and, in many cases, to propose a statistical model. However, the appropriateness and performance of the proposed model are often difficult to judge. Table 1 displays suggestions (in a checklist format) for structuring the presentation of pharmacodynamic analyses, which reflect the topics reviewed in this paper.

Correspondence to: Rosemarie Mick, University of Chicago Medical Center, 5841 S. Maryland Ave MC2115, Chicago, IL 60637, USA

This study was supported in part by grant N01-CM-07301 and Cancer Center Core grant CA-14599

Introduction

In this paper, we describe the statistical aspects of pharmacodynamic modeling as a practical guide for the novice. Variability in drug effect is defined and the motivations for its characterization are discussed. Pharmacokinetic and pharmacodynamic variability are contrasted as potential sources of variability in effect. This paper emphasizes several key points: (1) the selection of a measure of drug effect and explanatory variables, (2) the definition of valid statistical models, (3) the verification of model assumptions, (4) the application of multivariate regression techniques, and (5) the assessment of model performance. Important terminology is introduced in italics. Rigorous details on statistical theory and specialized methods (i.e., weighted or extended least squares) are not included. Lastly, a summary table is provided that offers suggestions for structuring the presentation of pharmacodynamic analyses.

Motivations: variability in effect

What it is

In contrast to *pharmacokinetics*, the study of what the body does to a drug (over time), *pharmacodynamics* is the study of what the drug does to the body, i.e., the study of drug effect, such as the degree of tumor shrinkage or severity of toxicity observed in a patient. Patients who receive an anticancer drug at some recommended phase II dose can exhibit substantial variation in a specific effect, known as *interpatient variability in effect,* which can be due to interpatient pharmacokinetic and/or pharmacodynamic variation [18, 33, 34].

Interpatient pharmacokinetic variability impacts on the absorption, distribution, metabolism, and/or elimination of a drug and, thus, determines drug exposure. Consider two patients treated with the same dose of carboplatin; one patient with a pretreatment creatinine clearance of 30 ml/min exhibits a markedly higher area under the time-concentration curve (AUC) than the patient with a normal

Table 1. Presentation checklist

The effect measure:	
	Demonstrate that the interpatient variability in effect is profound and scientifically and/or clinically relevant.
	Demonstrate that is it not related simply to varying drug dose.
	Describe the magnitude of interpatient variability (CV%).
The statistical model	
	Demonstrate that the model is appropriate for the relationship between effect and drug exposure and check the assumptions.
	Describe any transformations used and indicate the reason.
The proposed pharmacodynamic model:	
	Define all explanatory variables evaluated.
	Describe the multivariate selection method used.
	Indicate, if possible, which variables are believed to be pharmacokinetic or pharmacodynamic.
	Describe the criterion used for selecting the final variables.
	Indicate the *P* value, partial F, or chi-square statistic for each.
	Define the equation of the proposed model.
	Indicate the intercept and estimates of the regression parameters.
	Indicate and comment on the sign of the parameters in terms of the direction of worse outcome.
	Describe the overall assessment of the model (*F* statistic, R^2).
The residual error:	
	Describe measures of bias and precision (MPE, RMSE).
	Display a scatter plot of the observed effect by the predicted effect.
	Describe where and how the observed effect departs from the predicted effect.
Validation:	
	Demonstrate validation on an independent data set.
	Split the sample into training and test sets, if necessary.

creatinine clearance of 100 ml/min. The presence of interpatient pharmacokinetic variation in drug exposure is due to variation in renal function, which impacts on drug elimination. Creatinine clearance is then a recognized pharmacokinetic factor.

Variability in drug effect may be due to pharmacokinetic and/or pharmacodynamic factors, and although theoretically of differing biologic mechanisms, these factors may not be easily distinguished. For illustration, performance status might be correlated with drug clearance but should not be excluded from pharmacodynamic consideration. The biologic mechanisms influenced by pharmacodynamic factors are less clearly understood but are generally considered to be surrogates of overall patient health and tolerability. Again, consider two patients who exhibit similar drug exposure as measured by carboplatin AUC; one heavily pretreated patient experiences substantially worse platelet toxicity than the patient who was previously untreated. The presence of interpatient pharmacodynamic variation is due to the impact of prior treatment, a surrogate measure of drug tolerability. Prior treatment is then a recognized pharmacodynamic factor. In addition, prior treatment could alter pharmacokinetics if an end organ (i.e., the liver) were affected.

Why study it?

The reasons for studying interpatient variability in effect are twofold: biologic and scientific relevance and clinical applicability. Identification of the biologic factors that promote variability in similarly treated patients can be a formidable challenge. Moreover, the results of pharmacodynamic analysis can have widespread clinical applicability. When pharmacokinetic and/or pharmacodynamic variation leads to an appreciable risk of lethal toxicity at what is assumed to be a safe dose, a modification in administration is warranted [30]. Pharmacodynamic models have been employed to individualize drug doses such that a safe and effective dose can be given to an individual patient [12]. Ideally, a pharmacodynamically guided dose will yield improved efficacy while simultaneously lessening the risk of unacceptable toxicity.

To determine whether the magnitude of interpatient variability in effect is worthy of investigation, the *coefficient of variation* (CV), the ratio of the sample standard deviation to the sample mean, is often employed. The CV is ideal for assessing the relative degree of variability across different effect measures. Consider a phase II study of a single agent that demonstrates an acceptable average nadir WBC of 2,800/μl but substantial interpatient variation (a standard deviation of 2,400/μl and an 86% CV%) such that pharmacodynamic investigation would be warranted.

When to study it

As discussed by Peck and colleagues [30], pharmacokinetic and pharmacodynamic (PK/PD) relationships should be studied continuously throughout the drug development process. In preclinical PK/PD analysis, defining dose-toxicity or concentration-toxicity relationships from rodent data may assist in establishing optimal dose escalation strategies for phase I trials. Likewise, analyses begun in phase I trials allow initial evaluation of dose-concentration-toxicity relationships over a wide range of delivered doses. Some particularly toxic (and never to be repeated) doses may provide valuable and unique clinical information.

Phase II studies provide a basis for most PK/PD investigations, as unexpected yet substantial variation in effect due to nontrivial factors may first be recognized. Dose-response relationships may also be revealed for the first time if no pharmacodynamic analysis was performed during phase I trials. Such analyses may lead to earlier identification of patient subgroups and justification for individualized drug dosing. PK/PD data bases should be established, augmented, and mined throughout the drug development process.

Amonafide is one example of a recently developed anticancer drug that may have benefited from PK/PD evaluation earlier in the development process. Two phase I studies of amonafide given on a 5-day schedule recommended divergent phase II doses: 250 vs 400 mg/m^2 daily [22, 38]. Shortly thereafter, Cancer and Leukemia Group B conducted a phase II study at an intermediate dose of 300 mg/m^2 daily, which revealed a striking (and unexpected) variation in the degree of leukopenia. A pharmacodynamic analysis later demonstrated a strong correlation between the plasma concentration of an active metabolite of amonafide (the rate of acetylation) and the extent of leukopenia, with additional prediction by three pharmacodynamic factors: pretreatment WBC, performance status, and race [36]. Continued study of amonafide at the University of Chicago established that individualized dosing based on acetylator phenotype was required [35, 37].

Confounding factors

An apparent correlation between an exposure and an effect may be due to a third factor, termed a *confounding factor*. A confounding factor can be an extraneous, sometimes trivial factor that may wholly or partially account for the apparent impact of an exposure. In the pharmacodynamic investigation of phase I data, an obvious confounding factor for toxicity is the varying drug dose. Observed interpatient variability in drug exposure (i.e., AUC) can be explained by two sources: variable dose levels (a trivial source) and true interpatient pharmacokinetic variability (as in the prior example, the effect of interpatient variation in creatinine clearance on drug elimination). A regression model may be used to assess the impact of drug exposure and other explanatory variables on drug toxicity while controlling for the varying dose levels at which patients are entered. Failure to control for dose may lead to the erroneous conclusion that drug exposure correlates with drug toxicity.

A phase I trial of the new anticancer drug pyrazine diazohydroxide (PZDH) was recently completed at the University of Chicago [46] and serves as an illustration for this point. Data on dose-limiting thrombocytopenia and drug exposure are displayed in Fig. 1 (presented with permission). First, a naive interpretation of Fig. 1A would lead to the erroneous conclusion that substantial interpatient variability in thrombocytopenia was related to pyrazine AUC. However, dose was an obvious confounding factor, as it correlated with platelet toxicity (Fig. 1B) and AUC (Fig. 1C). A multivariate model demonstrated that AUC was not correlated to platelet toxicity, after accounting for the effect of dose level.

Statistical methods: pharmacodynamic modeling

Measuring effect

Pharmacodynamic investigations describe drug effect, considered to be the result of the drug given. The effect, or *dependent variable*, can be a measure of toxicity or tumor response. The *functional form* of a pharmacodynamic model is dictated by properties of the effect: (1) the scale and distribution of the effect and (2) the relationship (nonlinear or linear) of drug exposure to effect.

An effect can be either categorical scaled or interval scaled. A categorical scaled effect can have two levels, termed dichotomous (i.e., toxicity/no toxicity or response/no response), or more than two levels, termed ordinal polychotomous [i.e., National Cancer Institute (NCI) toxicity grades 0, 1, 2, 3, and 4 or response levels complete, partial, minimal, or no response]. An ordinal categorical effect cannot be analyzed as an interval scaled effect, just as a toxicity grade of 3 cannot be interpreted as being 3 times more severe than a toxicity of grade 1.

An interval scaled effect is the result of some measurement, for example, the nadir blood count as a measure of bone marrow toxicity. Interval scaled posttreatment measures are widely used as dependent variables in pharmacodynamic modeling, and consideration should be given to pretreatment adjustment [20]. Evaluating a posttreatment measurement without taking into consideration the magnitude of the pretreatment value may yield an invalid result. Consider the toxicity data of a patient whose WBC fell from 6,000 to 3,000/μl after treatment, evaluated identically to that of a patient whose WBC fell only from 3,600 to 3,000/μl. Several measures are employed that reflect the relative change of a posttreatment or nadir measurement: absolute change (pretreatment minus nadir value), percent change [(pretreatment minus nadir value) divided by pretreatment value], and survival fraction (nadir value divided by pretreatment value). As described above, multivariate regression may be employed to assess the impact of explanatory variables (i.e., AUC or performance status) on the posttreatment or nadir value while controlling for the magnitude of the pretreatment value.

Specific models exist for interval scaled and categorical scaled dependent variables. Linear and nonlinear models may be employed for interval scaled effects. Survival fraction (SF) is a widely used interval scaled measure that often displays a sigmoidal relationship to drug dose or drug exposure as expressed by the formula

$$SF = \exp[-kC_p],$$

where C_p is the plasma concentration and k is a rate constant that determines the slope of the decay curve. A linear relationship may be defined by a natural log transformation on SF to yield:

$$LN(SF) = -kC_p.$$

Substitution of pretreatment and nadir blood counts for SF yields the linear relationship:

$$LN(\text{nadir value}) = LN(\text{pretreatment value}) - kC_p.$$

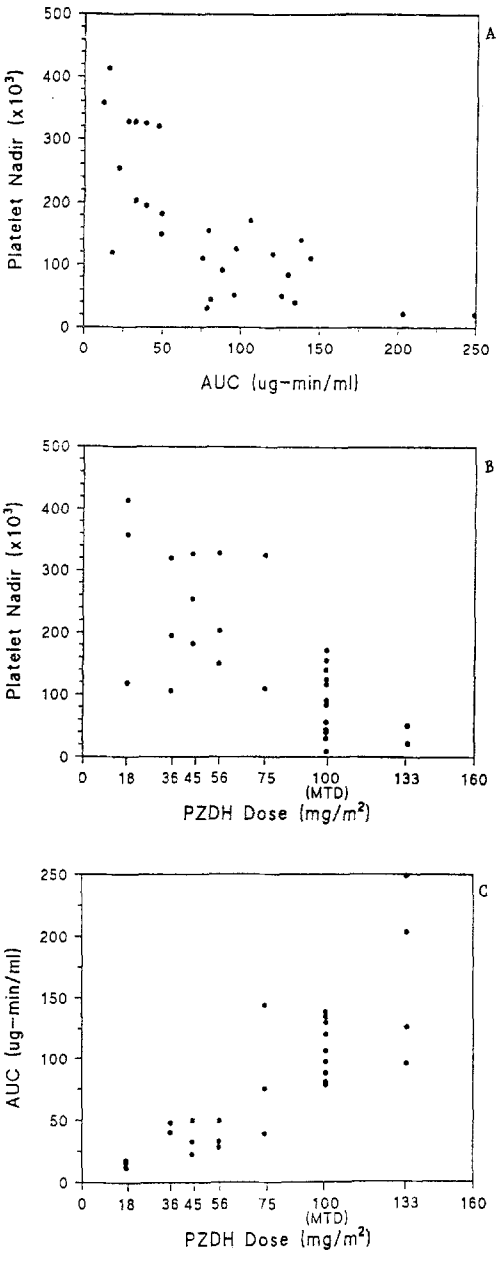

Fig. 1 A – C. Bivariate scatter plots from a phase I study of pyrazine diazohydroxide. *PZDH*, Pyrazine diazohydroxide; *AUC*, area under the time-concentration curve. **A** Platelet nadir by pyrazine AUC. **B** Platelet nadir by pyrazine dose. **C** Pyrazine AUC by pyrazine dose

The distribution of nadir blood count is known to be truncated at zero and skewed to the right. Shown in Fig. 2A is a histogram of pseudo-data of nadir WBC that fails the Kolmogorov-Smirnov normality test. In some circumstances a natural log transformation (Fig. 2B) on such data yields a *log normal distribution*, whereas in other circumstances a square root transformation (Fig. 2C) may yield a normal distribution.

Explanatory variables

As described above, pharmacokinetic and pharmacodynamic factors may be employed as *independent*, or explanatory, variables in the investigation of the variability in effect. Well-recognized pharmacokinetic factors that impact on known physiologic mechanisms are prior treatment, weight, genetic factors (i.e., genotype, race, sex), and measures of hepatic and renal function (i.e., bilirubin and creatinine clearance, respectively) [28]. The aging process is also likely to influence pharmacologic processes. Age is viewed as an accumulation of the physiologic effects of both the disease and aging processes [11]. As an illustration, recently published pharmacokinetic analyses have demonstrated that substantial interpatient variability in the clearance of 5-fluorouracil is predicted by gender and age [26] and dihydropyrimidine dehydrogenase (DPD) activity [9], a genetic factor.

Although less easily recognized, pharmacodynamic factors may also be predictive and serve as surrogate measures for patient health. Informative pharmacodynamic factors include pretreatment blood counts, transfusion requirement, performance status, prior treatment, and albumin. A population pharmacodynamic study of the relationships of carboplatin exposure (measured by AUC) to interpatient variability in the degree of thrombocytopenia and the likelihood of response in over 1,000 women with ovarian cancer serves as a prime example [19]. The likelihood of thrombocytopenia (graded 1, 2, 3, or 4) and the tumor response (complete or partial response) were analyzed by multivariate logistic regression, which demonstrated that three pharmacodynamic factors were predictive of outcome: prior treatment (yes/no), pre-treatment platelet count, and performance status.

Unlike the dependent variable, the scale of the independent variable is not restricted by regression methods, although categorical scaled variables that are dichotomous (yes/no) must be represented by *indicator variables* that are assigned a value of zero or 1. Categorical variables that are polychotomous, such as race (i.e., white, black, Hispanic, Asian), must be represented by a *set* of k-1 indicator variables for a variable with k levels.

Statistical models

In the simplest terms, the regression model is a framework with which one can study the relationship of an independent variable to a dependent variable. In a *functional* relationship the value of the dependent variable is exactly determined by the function and value of the independent variable. As illustrated by the functional relationship of pounds to kilograms (1 lb = 0.454 kg), a 100-lb patient weighs exactly 45.4 kg. By contrast, in a *statistical* relationship the value of the dependent variable is not exactly determined by the function and value of independent variable, but rather the dependent variable has a tendency to vary in some systematic fashion with the independent variable. An example is the commonly held observation that on average, worse hematologic toxicity is more likely to occur

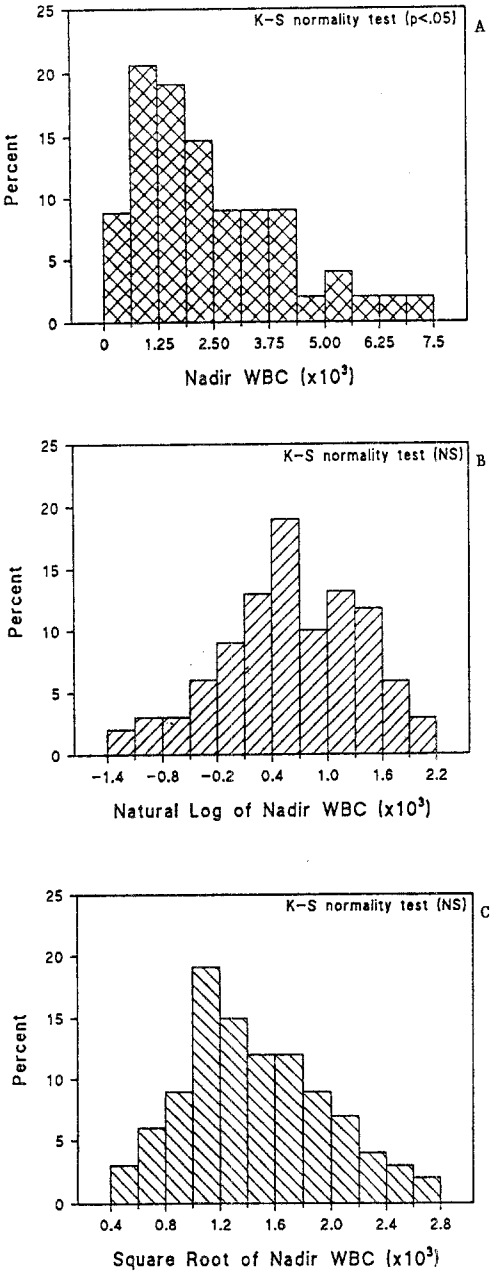

Fig. 2 A – C. Histograms of pseudo-data of nadir white blood cell counts and transformed nadir white blood cell counts. *WBC*, white blood cell count; *K-S*, Kolmogorov-Smirnov test; *NS*, not significant. **A** Distribution of nadir WBC, truncated at zero and skewed to the right (fails the normality test). **B** Distribution of natural log-transformed nadir WBC (passes the normality test). **C** Distribution of square root-transformed nadir WBC (passes the normality test)

in previously treated patients than in untreated patients. As such, a statistical relationship is defined by both a regression component (the functional relationship) and an error component (the deviation from expected). Understanding the magnitude and distribution of the error component is a fundamental requirement of any regression application.

We have thus far discriminated between interval and categorical scaled effects. Next we define specific statistical models for each type of effect. For interval scaled effects, linear [10, 15, 21, 27, 47] and nonlinear [8, 10, 27] pharmacodynamic models are appropriate. For dichotomous and ordinal polychotomous categorical scaled effects, logistic [27, 47] and proportional odds [1, 23] pharmacodynamic models, respectively, are appropriate.

Consider a phase II study of 50 patients treated at some fixed dose of an anticancer drug, with leukopenia as measured by natural log-transformed nadir WBC (LNWBCN) being the primary toxicity of interest. If substantial variation in leukopenia is observed, it may be pharmacokinetic and/or pharmacodynamic in orgin. Since nadir WBC is an interval scaled measure, a linear or nonlinear model may be applied. Suppose three independent variables are to be evaluated: dose, age, and natural log-transformed pretreatment WBC (LNWBCP). Dose and age, assessed as pharmacokinetic factors, would be expected to impact on drug exposure, and LNWBCP, assessed as a pharmacodynamic factor (or confounding factor), may reflect overall patient health.

A linear model may be employed if the relationship with the independent variables is *linear in the parameters*, that is, if no parameter appears as an exponent or is multiplied or divided by another parameter, and if the relationship is *linear in the independent variable(s)*, that is, if the variables appear only in the first power. Let B_0 be the intercept and B_1, B_2 and B_3 be the estimates (betas) of the regression parameters for each of the three independent variables. The regression component of the model is represented by a linear combination of the betas and the independent variables. The error component is represented by the model error (ε), the deviation of the observed nadir WBC from the nadir predicted from the true regression. The linear statistical model for this example would be:

$$\text{LNWBCN} = \underbrace{B_0 + B_1 \times \text{dose} + B_2 \times \text{age} + B_3 \times \text{LNWBCP}}_{\text{regression}} + \underbrace{\varepsilon}_{\text{error}}. \quad (1)$$

If the relationship of effect to the independent variables is nonlinear in the parameters or nonlinear in the independent variables, a nonlinear model is appropriate. An example of a nonlinear relationship is that of survival fraction to plasma concentration (C_p):

$$\text{SF} = \frac{(E_{MAX}) \times (C_p)^H}{(C_{50})^H + (C_p)^H} + \varepsilon. \quad (2)$$

In the above Hill E_{max} model, the parameters estimated are E_{MAX}, the maximal effect, C_{50}, the concentration associated with a half-maximal effect; and H, the Hill constant [14].

If the effect is categorical and dichotomous, such as toxicity/no toxicity or response/no response, then a logistic regression model is employed. If P denotes the probability that dose-limiting toxicity (DLT) would be observed for a patient given the three predictive factors defined above, then the *odds* that DLT would be observed would be equal to the ratio $P/1\text{-}P$ and the log(odds), or *logit*, would be equal to natural $\log(P/1\text{-}P)$. The probability of DLT is defined by a sigmoidal logistic function:

$$\text{Prob(DLT)} = \frac{\exp[B_0 + B_1 \times \text{dose} + B_2 \times \text{age} + B_3 \times \text{WBCP}]}{1 + \exp[B_0 + B_1 \times \text{dose} + B_2 \times \text{age} + B_3 \times \text{WBCP}]}. \quad (3)$$

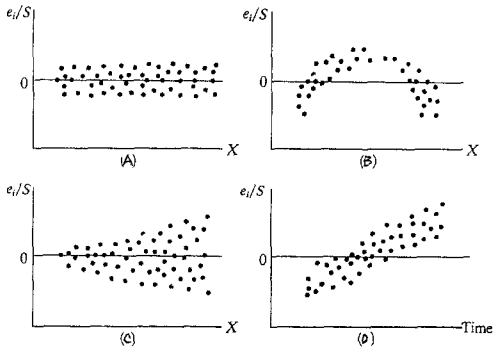

Fig. 3 A – D. Bivariate scatter plots of standardized residuals by independent variable X (from [21], p. 239). **A** Residuals display a mean equal to zero, constant variance, and no correlation to X. **B** Residuals display a trend (nonlinear) with X. **C** Residuals display a trend of increasing variance with increasing X. **D** Residuals display a trend (linear) with time, indicating autocorrelation

The logit is linear in the parameters and linear in the independent variables, corresponding to a *linear logistic model* (note the similarity to model 1):

$$\text{logit} = \log(P/1\text{-}P) = B_0 + B_1 \times \text{dose} + B_2 \times \text{age} + B_3 \times \text{WBCP}. \quad (4)$$

If the effect is ordinal polychotomous, then a proportional odds model (or cumulative logistic model) is employed. For a toxicity with five grades (0, 1, 2, 3, and 4), the proportional odds model defines the log of the ratio of the odds for a split of the levels (i.e., the odds for patients graded ≤ 2 to the odds for patients graded >2) by a linear combination of the independent variables.

Assumptions and transformations

When a statistical model is selected, one cannot know in advance whether the particular model is appropriate for the pharmacokinetic/pharmacodynamic data. Theory may steer us to specific models, but the appropriateness of the model for the actual data must be tested. To perform this evaluation, one must verify that the *assumptions* of the statistical model have been met. Useful transformations and alternative regression methods should be considered when violations are apparent.

For all models considered, critical distributional assumptions about the *model error* (ε) must be met. As described above, the model error reflects the deviation between the observed effect and the effect expected from the true linear model if the exact relationship is known. Similarly, *residual error* (e), the deviation between the observed effect and the effect predicted from the fitted model, reflects our inability to predict the effect exactly based on a relatively small number of predictors and an allowance for error in the measurement of effect.

The *ordinary least squares* (OLS) method is one method employed to derive estimates (betas) of the regression parameters [10, 27, 47]. First, the relationships of effect to the interval scaled variables are assumed to be linear, verifiable by visual inspection of bivariate scatter plots. If this assumption is in violation, transformations (i.e., natural log) may be considered or nonlinear models can be investigated. Second, OLS linear models require that the model errors have a mean equal to zero and a constant variance and are independent between individuals. The additional assumption that the errors are normally distributed is required for *maximum likelihood* (ML) estimation, which provides the distributional properties necessary for statistical inference (i.e., hypothesis testing).

These assumptions are verified by evaluating the residual errors (or residuals). Normal probability plots and scatter plots of the residuals by the predicted effect and by each explanatory variable allow confirmation of the assumptions. Scatter plots that reveal distinct patterns indicate that the assumptions are in violation. Figure 3A presents a plot of the residuals by an explanatory variable X, in which the horizontal band of data points shows no apparent trend, indicating that the assumptions have been met. Figure 3B clearly shows that the relationship between effect and X is not linear, as the residuals follow a curvilinear pattern. Figure 3C demonstrates that the error variance is not constant, as it increases as X increases, whereas Fig. 3D indicates that the residuals are not independent, as there is a positive correlation with time.

Transformations are the most commonly used method to correct for model violations. Popular transformations are log, square root, reciprocal, and arcsin. ACE transformations [6] maximize the correlation between effect and an explanatory variable but may not address the violations discussed. A transformation on a variable with a narrow range of values is not likely to be effective. Transformations on both dependent and independent variables may linearize their relationship, whereas a transformation on the dependent variable alone may normalize the error distribution and/or stabilize the error variance. Other methods to deal with nonconstant error variance are *weighted least-squares* and *extended least-squares regression* [41]. If the residuals appear to be correlated, specific autocorrelation models should be considered. Once the data have been transformed, resulting residuals should again be examined, as controlling one violation by transformation may reveal or enhance another type of violation.

As with linear models, assumptions for nonlinear models are based on the model errors and require verification. If the model errors are independent, with the mean being equal to zero, and have a common variance, then OLS estimates are obtained by iterative methods. If the model errors are also normally distributed, then ML estimates are possible. Again, scatter plots are useful for residual analyses and the weighted least-squares method is commonly employed when the errors do not display a common variance [29].

Logistic and proportional odds models are two examples of models that fall within the class of *generalized linear models* [24]. Linear regression cannot be applied directly to dichotomous or polychotomous categorical effects, since these are discrete measures. As noted earlier, defining the outcome as the probability of an event (i.e., dose-limiting toxicity or response) yields a sigmoidal relationship, which may be modeled with nonlinear regression. The logit transformation yields a linear logistic model for which the assumptions of OLS and ML also hold. OLS and ML estimates can be obtained by iterative methods.

Multivariate regression methods

In multivariate regression, the key challenge may be the systematic evaluation of the explanatory variables. Since a number of pharmacokinetic and/or pharmacodynamic factors may play a role, problems arise when numerous variables are considered. First, variables that are correlated (termed *multicollinear*) hinder the accurate prediction of effect. Second, parsimonious models are easier to understand and apply. Overfit models of marginally informative explanatory variables are of little use. Model reduction, not model augmentation, should be considered whenever possible.

For linear, logistic, and proportional odds regression models, many statistical computer packages offer several approaches to multivariate regression based on a sequential selection of explanatory variables. Each variable is assessed for its contribution to the reduction of unexplained variation in the effect. Two techniques for multivariate regression are *all possible subsets* regression and *automatic search* methods [10, 17, 27, 47].

All possible subsets regression, just as it sounds, involves fitting all possible models from a list of independent variables. For 10 independent variables, there are 2^{10} = 1,024 possible models. Although preferable to search methods, this approach may exert an appreciable demand on computing resources. All subsets regression begins with the identification of the "best" one-variable model from all possible one-variable models, followed by identification of the best two-variable model, and so forth. The best model is determined by a criterion that measures reduction in unexplained residual error, such as R^2, the *coefficient of multiple determination*. One should be aware that some criteria, such as R^2, monotonically increase as more variables are added to the model; thus, there is a tradeoff of larger R^2 for model complexity.

Somewhat less exhaustive yet more widely used are the automatic search methods, which include forward selection, backward elimination, and stepwise regression. By these methods, a sequential selection of the "best" explanatory variables is made based on specific criteria. In forward selection, each explanatory variable is assessed individually for correlation with the effect. The variable with the strongest correlation based on a criterion (*partial F statistic* or *partial* r^2, the *coefficient of simple determination*) is entered into the model. A partial r^2 refers to the correlation of the explanatory variable to the dependent variable adjusted for the contribution of the other variables present in the model. After a variable has been entered, those that remain are again assessed for correlation with effect, controlling for the variable previously entered into the model. The process continues until none of the variables that remain out of the model meet the level of significance for inclusion. Backward elimination starts with a full model of all potential variables entered, after which the variable with the weakest correlation is removed. The process continues until none of the variables that remain in the model meet the level of significance for exclusion. In stepwise regression an important flexibility is added; variables that have been entered previously are also allowed to be deleted from the model in later steps.

These multivariate regression methods are for the most part computationally trivial with the current computing resources but must be conducted with careful thought. One should be aware of the number of explanatory variables being assessed in relationship to the number of patients [5]. The stepwise regression method does not construct accurate models if the sample size is small relative to the number of potential explanatory variables. Also, the consequence of multicollinearity among the predictors will bias and lessen the precision of the model. The modeling process should not be mechanical. Aside from utilizing these automatic selection procedures, one should build models that make biologic "sense." Knowledge of plausible pharmacokinetic and/or pharmacodynamic relationships is an important tool.

The methods for parameter estimation of nonlinear models do not feature the theoretical (or computational) simplicity of linear and generalized linear models. The statistical theory of OLS and ML estimation for nonlinear models is complicated by iterative numerical procedures required to obtain a solution to estimation of the regression parameters [8]. Moreover, multiple solutions may exist. Nonlinear modeling computer packages commonly offer several search procedures for parameter estimation, including Gauss-Newton interation, method of steepest descent, and Marquardt's compromise (a middle ground between the first two methods). These procedures require either starting values for the parameters to be estimated or boundaries for a grid that may be searched for the best starting values. Initially, a nonlinear function relating drug exposure to effect must be defined, and then the evaluation of pharmacodynamic factors may be performed. We investigated the pharmacodynamics of the widely used anticancer drug etoposide using a Hill E_{max} model to relate nadir WBC to pretreatment WBC (WBCP) and steady-state etoposide concentration, Css [25], as expressed by the formula:

$$\mathrm{WBCN} = \mathrm{WBCP} \times \left[1 - \frac{C_{ss}^{H}}{C_{ss}^{H} + (K_{PD})^{H}}\right] + \varepsilon. \quad (5)$$

K_{PD} was defined as an estimable pharmacodynamic constant viewed as unexplained interpatient pharmacodynamic variability. We then extended the model to examine potential pharmacodynamic factors by replacing the K_{PD} term with potential explanatory variables (i.e., serum albumin or performance status).

Measures of model performance and validation

All statistical models are wrong, yet some statistical models are "less wrong" than others. One should maintain some degree of skepticism about any model. Formal procedures can be employed to test the fit of a model, such as the F test for linear models and a chi-square test for logistic models. For selection among several nonlinear models, the *Akaike information criterion* (AIC) is a quantitative measure that can be employed, in which a lower value of AIC indicates a superior model [2].

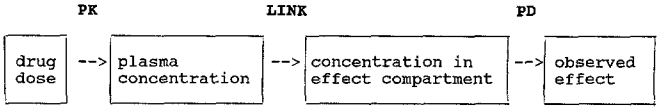

Fig. 4. Structural framework of PK/PD modeling; the PK model relates the dose to the concentration in the central compartment; the Link model relates the concentration in the central compartment to the hypothetical effect compartment; the PD model relates the concentration in the effect compartment to the observed effect

The R^2 measure is often employed to assess model performance, since it reflects the total variation in effect associated with (or explained by) the explanatory variables in the model. However, it does not reflect the predictive merit of the model and can be misleading. Although R^2 may be sizeable, the variation in the residual errors may be too large to yield accurate predictions. Thus, the model would not be appropriate for clinical application.

Superior measures of model performance as described by Sheiner and Beal [40] are bias and precision. Bias reflects whether the model error is random or varies systematically (consistently over or under predicts the effect). Precision reflects the magnitude of the model error. The model is *unbiased* if the mean of the model errors is zero and *precise* if the variance of the model errors is small. For N patients, the mean prediction error (MPE), mean squared error (MSE), and root mean square error (RMSE) are calculated from the residuals (e) by the respective equations:

$$\text{MPE} = 1/N \Sigma e,$$
$$\text{MSE} = 1/N \Sigma e^2, \text{ and}$$
$$\text{RMSE} = [1/N \Sigma e^2]^{1/2}.$$

The *mean prediction error* (MPE), the mean of the residuals, is a measure of bias. A confidence interval that is constructed in the standard manner by $\text{MPE} \pm t_{N-1}$ SEM (standard error of the mean) and includes zero is evidence of an unbiased model. The *mean squared error* (MSE) is the mean of the squared residuals. The *root mean square error* (RMSE), the square root of the MSE, is a measure of precision. The RMSE is an estimate of the standard deviation of the model errors; the smaller the magnitude of the RMSE, the more precise the model.

All statistical models reflect some degree of prediction bias. Naturally, a model is most predictive for the data upon which it is based. To a lesser degree, it may be predictive for a new data set. The most preferable method used to test the predictive performance of the final model is to *validate* the model on an independent set of data. If no independent data set is available and the original data set is sufficiently large, it may be divided into a training (model building) and a test (validation) data set.

The primary intent of a pharmacodynamic analysis is to identify and quantitate pharmacodynamic relationships. If the ultimate goal is to propose a statistical model for future individualized dosing schemes, then the error of the model must be fully described such that the safety and utility of the model can be judged. Lack of bias, adequately high precision, and successful validation of the proposed model should be demonstrated.

Advanced techniques

Population pharmacokinetics refers to the estimation of pharmacokinetic parameters for a drug based on a group of patients [32, 39]. These population parameters reflect "on average" the pharmacokinetics of the drug but may differ vastly from any individual patient's parameters. There are three types of population pharmacokinetic parameters. Pharmacokinetic parameters that reflect pharmacokinetic relationships are known as *fixed effect* parameters (i.e., variability in drug exposure due to interpatient variability in creatinine clearance). Parameters that take into account the variability of these pharmacokinetic relationships either among patients (interpatient variability) or within patients (intrapatient or residual variability) are known as *random effect* parameters.

Approaches to population pharmacokinetic modeling may or may not account for random effects. The naive pooled data method estimates the pharmacokinetic factors (fixed effects) simply by pooling all the patients and fitting a compartmental model as though the data had come from one individual. The complex roles of intra- and interpatient variability are ignored. Somewhat more rigorous approaches are the general two-stage or iterative two-stage methods [44]. In the first stage, pharmacokinetic parameters are estimated for each patient individually, and in the second stage, these estimates are combined to yield population estimates of the fixed and random effects. Although two-stage methods produce adequate estimates for fixed effects, they tend to yield biased estimates of the random effects. A third method involves estimation of both fixed and random effects in a nonlinear mixed effect model [42]. Although mixed effect models provide more appropriate estimation of random effects, they require explicit models of the error structure. Such complex data analyses have been widely performed using the computer program NONMEM of Beal and Sheiner [3].

The simultaneous modeling of the pharmacokinetics and pharmacodynamics of a drug is based on a parametric structural approach assuming a time course of drug effect and exposure [7]. Three theoretical parametric models are required to define this structure: a pharmacokinetic model, a link model, and a pharmacodynamic model (Fig. 4). The pharmacokinetic (PK) model relates the dose to the concentration in the central (or plasma) compartment. A link model relates the concentration in the plasma compartment to that in the hypothetical effect compartment. The pharmacodynamic (PD) model relates the concentration in the effect compartment to the observed effect. Parametric PK models (i.e., compartmental models) and PD models (i.e., sigmoid E_{max} models) and link models have most often been utilized [4, 16, 31, 43], but *nonparametric* PK and PD models with a parametric link model have recently been proposed [13, 45].

Acknowledgements. The authors are very grateful to Nicholas J. Vogelzang for permission to present data from the pyrazine diazohydroxide phase I trial, to Richard L. Schilsky for helpful suggestions on pharmacologic content, and to the referees who made constructive suggestions on an earlier draft of this paper.

References

1. Agresti A (1989) A survey of models for repeated ordered categorical response data. Stat Med 8: 1209
2. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19: 716
3. Beal SL, Sheiner LB (1989) NONMEM users guides. NONMEM Project Group, University of California, San Francisco
4. Boyd RA, Chin SK, Don-Pedro O, Verotta D, Sheiner LB, Williams RL, Giacomini KM (1989) The pharmacokinetics and pharmacodynamics of diltiazem and its metabolites in health adults after a single bolus dose. Clin Pharmacol Ther 46: 408
5. Breiman L, Freedman D (1983) How many variables should be entered in a regression equation? J Am Stat Assoc 78: 131
6. Breiman L, Friedman JH (1985) Estimating optimal transformations for multiple regression and correlation. J Am Stat Assoc 80: 580
7. Colburn WA (1981) Simultaneous pharmacokinetic and pharmacodynamic modeling. J Pharmacokinet Biopharm 9: 367
8. Davies M, Whitting IJ (1972) Numerical methods for non-linear optimization. Academic Press, New York
9. Diasio RB, Beavers TL, Carpenter JT (1988) Familial deficiency of dihydropyrimidine dehydrogenase. Biochemical basis for familial pyrimidinemia and severe 5-fluorouracil-induced toxicity. J Clin Invest 8: 47
10. Draper H, Smith H (1981) Applied regression analysis. Wiley & Sons, New York
11. Egorin MJ (1993) Cancer pharmacology in the elderly. Semin Oncol 20: 1
12. Evans WE, Rodman JH, Relling MV, Crom WR, Rivera GK, Crist WM, Pui CH (1991) Individualized doses of chemotherapy as a strategy to improve response for acute lymphocytic leukemia. Semin Hematol 28: 15
13. Fuseau E, Sheiner LB (1984) Simultaneous modeling of pharmacokinetics and pharmacodynamics with a nonparametric pharmacodynamic model. Clin Pharmacol Ther 35: 733
14. Gilbaldi M, Perrier D (1982) Pharmacokinetics. Marcel Dekker, New York
15. Green P (1978) Analyzing multivariate data. Dryden, Hinsdale, Illinois
16. Gupta SK, Ritchie JC, Ellinwood EH, Wiedemann K, Holsboer F (1992) Modeling the pharmacokinetics and pharmacodynamics of dexamethasone in depressed patients. Eur J Clin Pharmacol 43: 51
17. Hocking RR (1976) Analysis and selection of variables in linear regression. Biometrics 32: 1
18. Holford NH, Sheiner LB (1981) Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. Clin Pharm 6: 429
19. Jodrell DI, Egorin MJ, Canetta RM, Langenberg P, Goldbloom EP, Burroughs JN, Goodlow JL, Tan S, Wiltshaw E (1992) Relationships between carboplatin exposure and tumor response and toxicity in patients with ovarian cancer. J Clin Oncol 10: 520
20. Kaiser L (1989) Adjusting for baseline: change or percentage change? Stat Med 8: 1183
21. Kleinbaum DG, Kupper LL (1978) Applied regression analysis and other multivariate methods. Duxbury, Boston
22. Leiby JM, Malspeis L, Staubus AE, Kraut EH, Grever MR (1988) Amonafide (NSC 308847): a clinical phase I study of two schedules of administration (abstract). Proc Am Assoc Cancer Res 29: 278
23. McCullagh P (1980) Regression models for ordinal data. J R Stat Soc [B] 42: 109
24. McCullagh P, Nelder J (1983) Generalized linear models. Chapman & Hall, London
25. Mick R, Ratain MJ (1991) Modeling interpatient pharmacodynamic variability of etoposide. J Natl Cancer Inst 83: 1560
26. Milano G, Etienne MC, Cassuto-Viguier E, Thyss A, Santini J, Frenay M, Renee N, Schneider M, Demard F (1992) Influence of sex and age on fluorouracil clearance. J Clin Oncol 10: 1171
27. Neter J, Wasserman W, Kutner MH (1985) Applied linear statistical models. Irwin, Homewood, Illinois
28. Newell DR (1989) Pharmacokinetic determinants of the activity and toxicity of antitumor agents. Cancer Surv 8: 557
29. Peck CC, Sheiner LB, Nichols A (1984) The problem of choosing weights in nonlinear regression analysis of pharmacokinetic data. Drug Metab Rev 15: 113
30. Peck CC, Barr WH, Benet LZ, Collins J, Desjardins RE, Furst DE, Harter JG, Levy G, Ludden T, Rodman JH, Sanathanan L, Schentag JJ, Shah VP, Sheiner LB, Skelly JP, Stanski DR, Temple RJ, Viswanathan CT, Weissinger J, Yacobi A (1992) Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. Clin Pharmacol Ther 51: 465
31. Porchet HC, Piletta P, Dayer P (1992) Pharmacokinetic-pharmacodynamic modeling of the effects of clonidine on pain threshold, blood pressure and salivary flow. Eur J Clin Pharmacol 42: 655
32. Racine-Poon A, Smith AF (1990) Population models. In: Berry DA (ed) Statistical methodology in the pharmaceutical sciences. Marcel Dekker, New York, pp 139 – 162
33. Ratain MJ (1992) Therapeutic relevance of pharmacokinetics and pharmacodynamics. Semin Oncol 19: 8
34. Ratain MJ, Schilsky RL, Conley BA, Egorin MJ (1990) Pharmacodynamics in cancer therapy. J Clin Oncol 8: 1739
35. Ratain MJ, Mick R, Berezin F, Janisch L, Schilsky RL, Williams SF, Smiddy J (1991) Paradoxical relationship between acetylator phenotype and amonafide toxicity. Clin Pharmacol Ther 50: 573
36. Ratain MJ, Berezin F, Allen SL, Costanza ME, Rosner G, Van Echo DA, Schilsky RL (1992) Population pharmacodynamic (PD) study of amonafide (bida): CALGB 8862 (abstract). Proc Am Soc Clin Oncol 10: 290
37. Ratain MJ, Mick R, Berezin F, Janisch L, Schilsky R, Vogelzang NJ, Lane LB (1993) Phase I study of amonafide dosing based on acetylator phenotype. Cancer Res 53: 2304
38. Saez R, Craig JB, Kuhn JG, Weiss GR, Koeller J, Phillips J, Havlin K, Harman G, Hardy J, Melink TJ (1989) Phase I clinical investigation of amonafide. J Clin Oncol 7: 1351
39. Sheiner LB (1984) The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. Drug Metab Rev 15: 153
40. Sheiner LB, Beal SL (1981) Some suggestions for measuring predictive performance. J Pharmacokinet Biopharm 4: 503
41. Sheiner LB, Beal SL (1985) Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. J Pharmacokinet Biopharm 14: 185
42. Sheiner LB, Grasela TH (1991) An introduction to mixed effect modeling: concepts, definitions and justifications. J Pharmacokinet Biopharm 19: 11S
43. Sheiner LB, Stanski DR, Vozeh S, Miller RD, Ham J (1979) Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to D-tubocurarine. Clin Pharmacol Ther 25: 358
44. Steimer JL, Mallet A, Golmard J, Boisvieux JF (1984) Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed effect model. Drug Metab Rev 15: 265
45. Unadkat JD, Bartha F, Sheiner LB (1986) Simultaneous modeling of pharmacokinetics and pharmacodynamics with nonparametric kinetic and dynamic models. Clin Pharmacol Ther 40: 86
46. Vogelzang NJ, Janisch L, Berezin F, Mick R, Schilsky R, Walter P, Ratain MJ (1992) Pyrazine diazohydroxide (PZDH): a phase I trial (abstract). Proc Am Assoc Cancer Res 33: 3182
47. Weisberg S (1985) Applied linear regression. Wiley & Sons, New York