# Understanding the Dose-Effect Relationship:
## Clinical Application of Pharmacokinetic-Pharmacodynamic Models

*Nicholas H.G. Holford and Lewis B. Sheiner*

Division of Clinical Pharmacology, Departments of Medicine and Pharmacy, Schools of Medicine and Pharmacy; and Department of Laboratory Medicine, School of Medicine, University of California, San Francisco, California

*Summary*

  *It is a major goal of clinical pharmacology to understand the dose-effect relationship in therapeutics. Much progress towards this goal has been made in the last 2 decades through the development of pharmacokinetics as a discipline. The study of pharmacokinetics seeks to explain the time course of drug concentration in the body. Recognition of the crucial concepts of clearance and volume of distribution has provided an important link to the physiological determinants of drug disposition. Mathematical models of absorption, distribution, metabolism and elimination have been extensively applied, and generally their predictions agree remarkably well with actual observations. However, the time course of drug concentration cannot in itself predict the time course or magnitude of drug effect. When drug concentrations at the effect site have reached equilibrium and the response is constant, the concentration-effect relationship is known as pharmacodynamics. Mathematical models of pharmacodynamics have been used widely by pharmacologists to describe drug effects on isolated tissues. The crucial concepts of pharmacodynamics are potency — reflecting the sensitivity of the organ or tissue to a drug, and efficacy — describing the maximum response. These concepts have been embodied in a simple mathematical expression, the $E_{max}$ model, which provides a practical tool for predicting drug response analogous to the compartmental model in pharmacokinetics for predicting drug concentration.*

  *The application of pharmacodynamics to the study of drug action* in vivo *requires the linking of pharmacokinetics and pharmacodynamics to predict firstly the dose-concentration, and then the concentration-effect relationship. This may be done directly by equating the concentration predicted by a pharmacokinetic model to the effect site concentration, but this simplistic approach is often not appropriate for various reasons, including delay in drug equilibrium with the receptor site, use of indirect measures of drug action, the presence of active metabolites, or homeostatic responses, thus often necessitating the use of more complex models.*

  *The relative pharmacodynamic bioavailability of different preparations of the same drug may be determined from the time course of a drug effect. Bioavailability determined in this way may differ markedly from bioavailability defined by measurements of drug concentration if active metabolites are formed or if effects are produced in the non-linear region of the concentration-effect relationship.*

*The influence of changes in the extent of plasma protein binding may be important in the interpretation of drug concentration measurements since it is generally held that only the unbound fraction is pharmacologically active. Clear examples of this phenomenon are few, but this reflects the general paucity of adequate observations rather than casting doubt on the usual assumption.*

*The design of rational dosing regimens for clinical therapeutics cannot be performed with a knowledge of pharmacokinetics alone. The time course of drug effect may be essentially independent of concentration when a dose produces near maximal effects throughout the dosing interval. If effects are between 20 and 80% of maximum, the response will decrease linearly even though concentrations are declining exponentially. Finally, at relatively small degrees of effect, the time course of drug effect and concentration will be in parallel. The usual 'rule of thumb' of dosing every half-life is a conservative strategy for limiting wide fluctuations in drug effect, but demands more from the patient in terms of dosing frequency than may be necessary to achieve consistent drug action. On the other hand, if therapeutic success is dependent more on cumulative response than moment to moment activity, the use of extended dosing intervals may markedly reduce the effectiveness of the same average dose. Considerations of these factors can be incorporated into a dosing scheme by combined application of the principles of pharmacokinetics and pharmacodynamics.*

Recent achievements in clinical pharmacology have rested heavily on the use of pharmacokinetics to describe drug concentrations in the body. Of greater clinical relevance is the relationship between these concentrations and clinical efficacy and toxicity. The study of the relationship between concentration and drug effect is the science of pharmacodynamics. Advances in the technology for measurement of drug concentrations in biological fluids have spurred the development of mathematical descriptions of the relationship between dose and concentration. These models are used extensively in pharmacokinetics and have been widely discussed and applied. Corresponding pharmacodynamic models have received much less attention from clinical pharmacologists, although these models are extensively used in classical pharmacology. This article reviews the development and application of pharmacodynamic models to the description of drug effects *in vivo* with particular attention to the synthesis of pharmacokinetics and pharmacodynamics. This integrated approach can be used to describe the overall relationship between dose and effect, and can lead to useful insights into rational dose regimen design.

The terms *pharmacokinetics* and *pharmacodynamics* are depicted in figures 1a and b.

The term *pharmacokinetics* is used to describe the process controlling drug concentration at any time after 1 or more doses. Drug concentrations may be predicted at the site of measurement, for example in plasma, or in additional hypothetical compartments which cannot be sampled directly. The usual pharmacokinetic models describe drug concentrations as a function of both dose and time. In contrast, pharmacodynamic models are in essence independent of time and describe the equilibrium (time-independent) relationship between concentration and effect. The parameters of pharmacodynamic models, e.g. max-
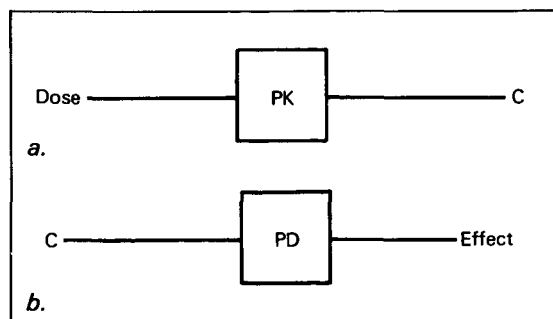


*Fig. 1.* Schematic depiction of the terms (a) pharmacokinetics (PK) and (b) pharmacodynamics (PD). C = drug concentration.
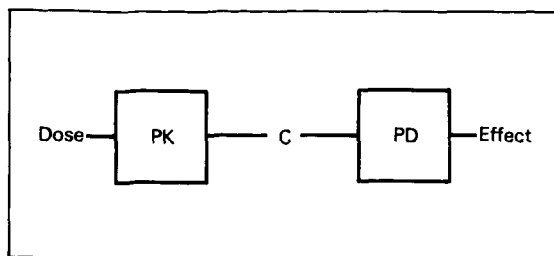
*Fig. 2.* Schematic representation of the dose-effect relationship using pharmacokinetic (PK) and pharmacodynamic (PD) models. C = drug concentration.

imum effect, may be time-dependent, but this is a special case which will be discussed separately. The rectangles in figure 1 are used to indicate the place of mathematical models in this scheme.

The pharmacokinetic and pharmacodynamic models share a common feature, concentration, and they can be combined to describe the overall dose-effect relationship (fig. 2).

Because we need to use concentrations in our pharmacodynamic model which reflect the action of drug at the effect site, and our pharmacokinetic model alone may not be able to predict such values directly, we will need to develop a further class of models to link them: pharmacokinetic-pharmacodynamic models, shown diagrammatically in figure 3. Plasma concentration is used only because it is commonly derived from a pharmacokinetic model, but other concentrations at other sites described by pharmacokinetic models could also be used.

Finally we must consider what is meant by effect, and how this is to be incorporated in the scheme. A drug may be referred to as an antihypertensive agent because its observable effect is to lower the blood pressure, but often the effect can be more precisely localised to a direct action, for example relaxation of vascular smooth muscle. In this case the pharmacodynamic model should strictly refer to the direct effect of the drug on the blood vessel and a physiological model is needed to relate this action to the observed effect of lowering the blood pressure. The dose effect scheme now becomes that shown in figure 4, where PE is a physiological model relating the direct

effect of the drug on some tissue or organ to a grossly observable effect.

A clear understanding of the overall dose-effect relationship requires knowledge of each of the 4 steps in figure 4. This review does not discuss the first of these, pharmacokinetics, because descriptions of these phenomena are widely available (e.g. Gibaldi and Perrier, 1975; Wagner, 1975). We will, however, discuss at least some aspects of the other 3 steps. Firstly, some generally useful pharmacodynamic models are described, and it is then shown how they may be incorporated into pharmacokinetic-pharmacodynamic and physiological models.

## 1. Pharmacodynamic Models

### 1.1 Types of Models

#### 1.1.1 Fixed Effect

When the observed effect is either present or absent, such as in prevention of seizures, or is defined by some criterion, such as greater than 70% suppres-
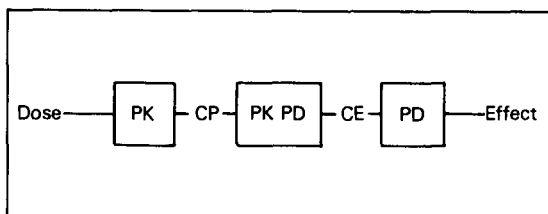


*Fig. 3.* Pharmacokinetic and pharmacodynamic model linked by an effect compartment. CP = plasma concentration, CE = effect site concentration.
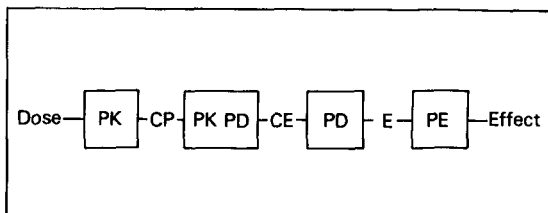


*Fig. 4.* Pharmacokinetic, pharmacodynamic and physiological models. E = grossly observable effect, PE = physiological effect at tissue level.

sion of ventricular extrasystoles, we can refer to the effect as 'fixed'. The degree of the effect is now immaterial, what is important is whether or not it occurs. This can be related to concentration in a statistical fashion expressing the probability of the effect given a particular concentration. For example, Beller et al. (1971) collected observations on an unselected series of patients receiving digoxin. The effect was defined as the presence or absence of digitalis toxicity. From their data we can calculate the probability of digitalis toxicity to be 50% at a concentration of 2ng/ml. The probability of the effect can be expressed over a range of concentrations by defining the nature of the underlying statistical distribution, for example cumulative normal distribution, with a particular standard deviation. When this is plotted, a sigmoid curve is obtained and, while this may resemble the sigmoid $E_{max}$ model described in section 1.1.2, the underlying mechanisms need not be similar.

### 1.1.2 $E_{max}$ Model

The simplest model which adequately describes drug effect over the whole range of concentrations is based on the hyperbolic relationship that we call the $E_{max}$ model:

$$E = \frac{E_{max} \cdot C}{EC_{50} + C} \qquad \text{[Eq. 1]}$$

where E is effect, C is concentration, $E_{max}$ is the maximum effect attributable to the drug and $EC_{50}$ is the concentration producing 50% of $E_{max}$. This model has been extensively used in other areas, for example in enzyme kinetics (Michaelis-Menten) and protein binding. Its use for pharmacodynamic phenomena may be justified on theoretical grounds (Ariens and Simonis, 1964a,b), but empirically it is useful because it has 2 important properties: (a) it predicts the maximum effect a drug can achieve, and (b) it predicts no effect when no drug is present. These properties may appear to be trivial, but they are not shared by one of the most widely used expressions — the log-linear model — discussed in section 1.1.3. The $E_{max}$ model
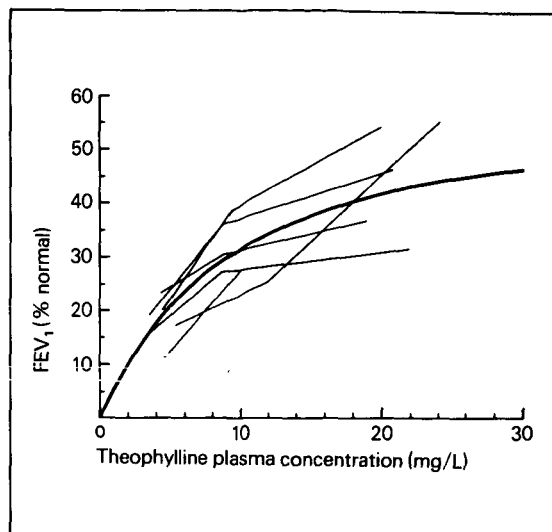


Fig. 5. Use of $E_{max}$ model to describe the effects of theophylline on change in normalised forced expiratory volume (FEV₁); $E_{max}$ = 63%, $EC_{50}$ = 10mg/L [see also text; after Mitenko and Ogilvie (1973) and Holford and Sheiner (1981b)].

incorporates the 'law of diminishing returns', reflecting the ever higher concentrations required to increase the effect by a given amount. This appears to describe a common biological reality. The use of the model is illustrated by the observations of Mitenko and Ogilvie (1973), who described the effect of theophylline on airway obstruction in a group of asthmatics. The effect was expressed as the change in forced expiratory volume (FEV₁) as a percentage of the predicted FEV₁ for a comparable person with normal lungs. When this effect is plotted as a function of theophylline concentration in serum it is clear that the 'law of diminishing returns' is in effect (fig. 5). The solid line in the figure represents the prediction of the $E_{max}$ model applied to these observations. The maximum effect, $E_{max}$, is 63% of predicted FEV₁ and reflects the typical non-reversible component (37%) of airway obstruction for this group of patients. The concentration producing 50% of $E_{max}$ ($EC_{50}$) is 10mg/L which expresses the clinical observation that adequate relief of bronchoconstriction is often obtained at this concentration. Doubling the

concentration to 20mg/L can be expected to achieve only a further 17% increase in effect, and indicates the ill-advisedness of increasing concentrations beyond this value; little gain can be expected and serious toxicity may ensue. The relationship between increase in concentrations and increase in effect predicted by the $E_{max}$ model is shown in table I.

When the drug effect is measured as inhibition of some biological phenomenon, equation 1 can be rewritten as:

$$E = NODRUG - \frac{E_{max} \cdot C}{IC_{50} + C} \qquad [Eq. 2]$$

where NODRUG is the effect when no drug is present and $IC_{50}$ is the concentration producing 50% of the maximum inhibition of effect ($E_{max}$). This inhibitory $E_{max}$ model can be illustrated by the work of Singh et al. (1980), who studied the effect of timolol, a $\beta$-adrenoceptor antagonist, on resting heart rate and its response to maximal exercise (fig. 6).



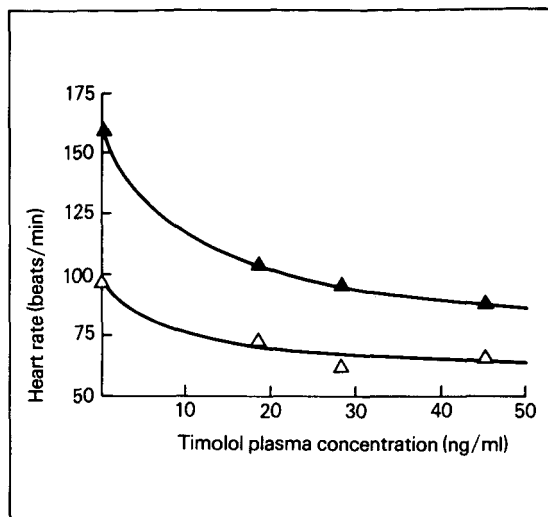Fig. 6. Use of inhibitory $E_{max}$ model to describe effects of timolol on exercise (▲) and resting (△) heart rate [see also text; after Singh et al. (1980) and Holford and Sheiner (1981b)].

When the heart rate is plotted as a function of timolol plasma concentration, it is clear that a maximum effect on these 2 measures of effect is being approached at the higher concentrations. Using equation 2 we estimated the values of $E_{max}$, NODRUG and $IC_{50}$ for both effects. The lowest heart rate achievable is given by the difference between NODRUG and $E_{max}$, and this is 56 beats per minute for resting heart rate and 68 beats per minute during exercise. The $IC_{50}$ for both effects was similar, i.e. 10ng/ml, and this can be used in conjunction with table I to predict 90% of maximum effect at a concentration of 90ng/ml. If a drug is capable of completely abolishing an effect, the value of $E_{max}$ is equal to NODRUG, which then becomes:

$$E = NODRUG \left(1 - \frac{C}{IC_{50} + C}\right) \qquad [Eq. 3]$$

The expression $C/(IC_{50} + C)$ is the 'fractional $E_{max}$' model. It describes the relationship between concentration and the fraction of maximal effect that can be attributed to the drug.

Table I. Relationship between increase in concentration (C) and increase in effect as predicted by the $E_{max}$ model

| Function E | $= \dfrac{E_{max} C}{EC_{50} + C}$ |
|---|---|

| C (multiples of $EC_{50}$) | Effect (% $E_{max}$) |
|---|---|
| 0.5 | 33 |
| 1 | 50 |
| 2 | 66.6 |
| 3 | 75 |
| 4 | 80 |
| 9 | 90 |
| 99 | 99 |

### 1.1.3 Linear Model

If drug concentrations are low in relation to $EC_{50}$ (eq. 1), the effect becomes proportional to concentration:

$$E = S \cdot C \qquad \text{[Eq. 4]}$$

where S is the slope of the line relating effect to concentration. This linear model can be derived from the $E_{max}$ model and shares with it the property of predicting no effect when drug is absent. It lacks, however, the ability to define the maximum effect. In practice it may not be possible to achieve concentrations which achieve effects approaching the maximum, and therefore the maximum effect cannot be known. The linear model may then be used as an empirical description of the drug effect over the observed concentration range.

### 1.1.4 Sigmoid $E_{max}$ Model

The concentration effect curve for some drugs is not represented simply by the hyperbolic form of the $E_{max}$ model. This deviation was noted by Hill (1910) in his studies of the saturation of haemoglobin by oxygen. He found empirically that his observations could be explained by the addition of an extra



*Fig. 7.* The sigmoid $E_{max}$ model, showing the change in shape from the simple hyperbola (N = 1) to other curves (N = 2 and N = 0.5) [after Holford and Sheiner (1981a)].

parameter which altered the simple hyperbolic form. This model we call the sigmoid $E_{max}$ model, and is defined as:

$$E = \frac{E_{max} \cdot C^N}{EC_{50}^N + C^N} \qquad \text{[Eq. 5]}$$

where N is a number influencing the slope of the curve (fig. 7).

When N = 1, we have the usual hyperbolic $E_{max}$ model, but when N is greater than 1 the curve becomes sigmoid with a steeper slope in its central portion. If it is less than 1 it is shallower in its central portion, but steeper at low concentrations and more shallow at higher ones. Hill discovered that when N was 4 it described the association between haemoglobin and oxygen on each haemoglobin molecule. This can be derived on theoretical grounds as the number of binding sites for oxygen, but, empirically, non-integer values of N may be found that have no physical interpretation in terms of receptor binding sites. We believe the sigmoid $E_{max}$ model is useful for describing concentration-effect relationships, but cautious interpretation of the meaning of N is necessary. For example, when this model was used by Meffin et al. (1977) to describe the steep response of patients to tocainide administered for suppression of ventricular extrasystoles, values of N up to 20 were found. These are unlikely to reflect receptor structure, but do emphasise the steepness of the curves (fig. 8).
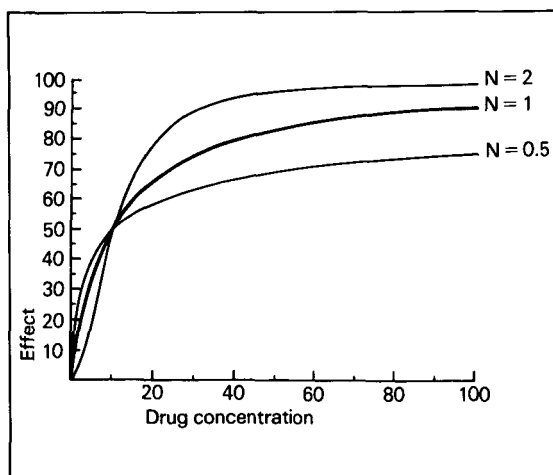
## 1.2 Baseline Effect

Many observable drug effects reflect changes from some baseline value, for example blood pressure. Incorporation of the baseline effect into a pharmacodynamic model is required to estimate the parameters describing drug effect. How this is done is determined by the properties of the baseline. Depending on the nature of the effect and experimental design, this baseline effect may be considered a fixed value with negligible measurement error, or may be
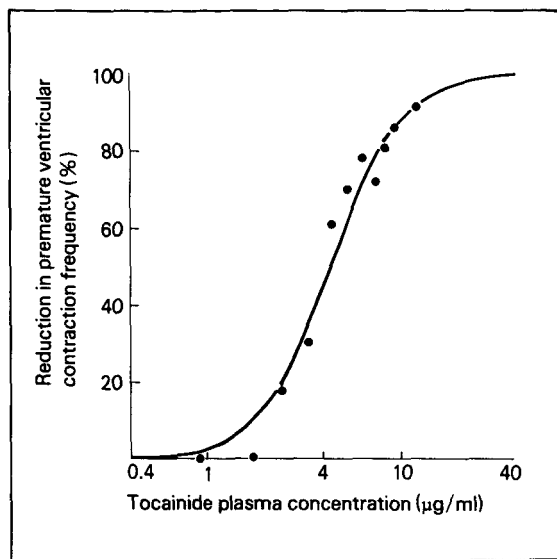
*Fig. 8.* Steep concentration response curve for tocainide requiring use of the sigmoid $E_{max}$ model [after Winkle et al. (1976), reproduced by permission of the American Heart Association, Inc.].

known only as well as any particular measurement made in the presence of the drug.

If the baseline effect is known to have the same measurement error as the other effect measurements, then we can add a parameter to estimate it ($E_0$) to one of the pharmacodynamic models described above, for example:

$$E = \frac{E_{max} \cdot C}{IC_{50} + C} + E_0 \quad (E_{max} \text{ model}) \quad [Eq. 6]$$

$$E = S \cdot C + E_0 \quad (\text{Linear model}) \quad [Eq. 7]$$

Notice that $E_0$, the estimated baseline effect, is not necessarily the same as the measured baseline value because of measurement error. The NODRUG parameter described above for the inhibitory $E_{max}$ model (eq. 2) is a baseline parameter exactly equivalent to $E_0$. The parameters of the model should be estimated using all the effect measurements, in-

cluding the baseline measurement. If $E_0$ is different from the measured baseline and cannot be explained by typical measurement error, then this raises serious questions about the appropriateness of the model. This is illustrated by the observations of Kramer et al. (1979) on the effect of digoxin on electromechanical systole ($QS_2$). Using a linear pharmacodynamic model with a baseline parameter and a pharmacokinetic-pharmacodynamic model to estimate the effect site concentration, we obtained the values shown in figure 9. The baseline parameter value was 10msec which is over 30% of the greatest observed effect and clearly different from zero. Kramer et al. (1979) concluded that an $E_{max}$ model could be more appropriate than a linear model, but the observations of Kelman and Whiting (1980) suggest the baseline parameter reflects the effect of the experiment independent of the drug, because this degree of effect was seen after placebo administration.

On the other hand, when the baseline effect is well known *a priori* it can be subtracted from the measurements made in the presence of the drug, for example:

$$(E - E_0) = \frac{E_{max} \cdot C}{EC_{50} + C} \quad [Eq. 8]$$

$$(E - E_0) = S \cdot C \quad [Eq. 9]$$

In this case $E_0$ is the known baseline effect and is no longer estimated as a parameter of the model. The removal of this unneeded parameter can result in substantial improvements in the estimation of the other parameters if experimental error is relatively large.

The work of O'Reilly (1974) on the actions of warfarin stereoisomers illustrates this. O'Reilly studied the change in prothrombin time after separate doses of S(−) and R(+) warfarin in volunteers. He avoided the need to describe the complex relationship between warfarin concentration, time and anticoagulant effect (see section 3) by relating the area under the effect-time curve to the corresponding area under the concentration-time curve. He was then able to

show a linear relationship between these areas in this group of subjects (fig. 10). The effect measurement was the change in prothrombin time from a known baseline. However, by definition, the area under the effect-time curve must be zero when no drug is present. O'Reilly used the linear model with a baseline parameter (eq. 7), thus estimating the value of $E_0$, implicitly known to be zero. The slope of the line describing the effects of S( − ) warfarin is therefore calculated to be more shallow than it would be if equation 9 had been used. O'Reilly estimated the potency of S( − ) by this method to be 2.2 times that of R( + ), which is lower than the value of 3.8 estimated by Breckenridge et al. (1974), who used a fixed baseline effect model.

## 1.3 Transformation of Effect Measurements

It is implicit in the pharmacodynamic models described above that the observed effect is directly related to the action of the drug at its effect site. Sometimes the actual measurement used clearly violates
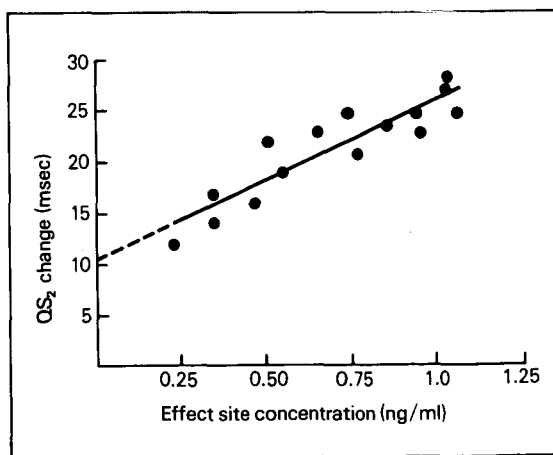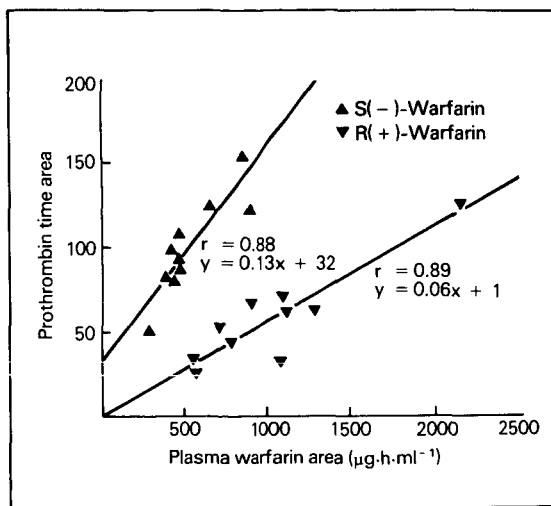
Fig. 10. Use of a linear model to determine relative potency of warfarin isomers by O'Reilly (1974) illustrates the pitfalls of estimating a baseline value unnecessarily [after Holford and Sheiner (1981b)].

this assumption. For example, heparin is known to interfere with blood coagulation, but its effects are most commonly measured by the time taken for blood to clot. As heparin approaches its maximum effect in inhibiting coagulation, this clotting time will tend towards infinity. Similarly, the prothrombin time as a measure of oral anticoagulant action will tend toward infinity as the concentration of vitamin K-dependent clotting factors approaches zero. This disproportionate relationship between the effect measurement and the actual drug effect may be corrected by a suitable transformation. Thus, Whitfield and Levy (1980) used the logarithm of the activated thromboplastin time to describe the effects of heparin, and O'Reilly (1974) used the logarithm of change in prothrombin time to describe the effects of warfarin. These transformations are quite empirical; others may be equally suitable. Whenever the actual index of drug effect is known to be an indirect measure of the drug's effect at its site of action some thought should be given to an appropriate transformation.

Desirable transformations often (a) increase the linearity of the relationship of the transformed varia-

Fig. 9. Linear relationship between change in electromechanical systole (QS₂) and digoxin concentration.
   The relatively large value of the baseline QS₂ (10msec) suggests that the data or the model may be inadequate [after Kramer et al. (1979) and Holford and Sheiner (1981b)].

ble to the explanatory one; (b) map a variable from a constrained space (e.g. 0-1) to ± infinity; or (c) make the measurement error of comparable magnitude throughout the range of the transformed variable.

## 1.4 The Log-Linear Pharmacodynamic Model — Use and Abuse

It has become traditional to illustrate the relationship between concentration and effect using the logarithm of concentration on the abscissa. The reasons for this transformation are 2-fold: firstly, it compresses the concentration scale, which is often convenient if a wide range of concentrations has been studied. Secondly, the relationship between log concentration and effect is a linear approximation to the $E_{max}$ model in the range of 20 to 80% maximum effect. This second property was valuable before the availability of non-linear regression methods because simple linear regression could be used to describe the slope of the line, and statistical tests of parallelism could be applied to compare the effects after addition of another drug such as a competitive antagonist. Unfortunately, the investigator was still required to estimate the value of $E_{max}$ visually and then calculate the value of $EC_{50}$ from the log-linear model prediction of 50% maximum effect. The log linear model can be written:

$$E = S \log (C) + I \qquad \text{[Eq. 10]}$$

where I is an arbitrary constant with no physical meaning. The log-linear model is unable to predict the obvious absence of effect when drug concentration is zero, and implicitly does not recognise the existence of a maximum effect. It is unable to accommodate the existence of a baseline effect, and this may lead to the abuse of a baseline measurement as if it were known without error. If the maximum effect is not known, it is impossible to be certain that the effect lies in the 20 to 80% $E_{max}$ range. If observations are indeed below 20% of $E_{max}$, data arising from the $E_{max}$ model will appear to be a curved line, concave upwards, when plotted according to equation 10. If the log-linear

model is applied without recognising that some of the measurements of effect approach $E_{max}$, the concentration dependence of the effect may be mistakenly discarded.

We see little value in the description of drug effects in terms of the log-linear model. There is no biological rationale for this transformation, and the model is unable to capture information about the parameters of interest. Graphic representation of concentration-effect data using a log scale for concentration distorts the data unnecessarily, and may obscure the existence of a maximum effect (Mitenko and Ogilvie, 1973), or make recognition of the need for the sigmoid $E_{max}$ model difficult.

## 2. Pharmacokinetic-Pharmacodynamic Models

### 2.1 Pharmacokinetic Model-independent

The most direct link between concentration and effect is made by using simultaneous measurements of effect and concentration at the effect site. This is illustrated for the effects of theophylline in figure 5. The principal disadvantages of this method are: (a) the requirement for simultaneous measurement of concentration and effect measurements when concentrations are rapidly changing; (b) the necessity of ignoring measurement error in the concentration value when relating it to effect; and (c) the necessity of sampling concentrations from the effect site. The use of an appropriate pharmacokinetic model, on the other hand, may permit better prediction of the actual concentration at the time of the effect.

### 2.2 Pharmacokinetic Compartment Model

The time course of drug concentrations in plasma, urine and other sampled sites is usually modelled by compartmental analysis. Other approaches are possible, for example perfusion models, but they are rarely practical with usual study designs. The concentration
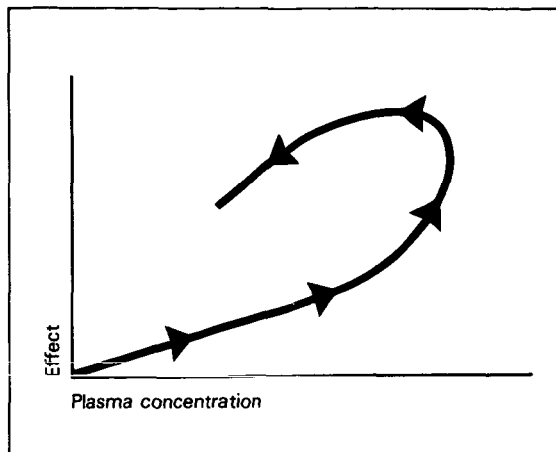
Fig. 11. Anticlockwise hysteresis loop indicating equilibration delay between plasma concentration and the effect site producing the effect [after Holford and Sheiner (1981a)].

of drug in one of the compartments of the model which is not directly able to be sampled may be predicted from the analysis of plasma concentration data (see Benet, 1972; Vaughan and Trainor, 1975; Veng Pedersen, 1980). The use of a pharmacokinetic model frees the investigator from simultaneous measurement of both concentration and effect. The timing of concentration samples and effect measurements can then be optimised independently according to the expected model (e.g. Box and Lucas, 1959; Di Stefano, 1981; St John and Draper, 1975). Plots of concentration against effect using concentrations predicted by a pharmacokinetic model for the central compartment are useful for the recognition of equilibration delays between plasma concentration and the effect site (fig. 11).

The anticlockwise hysteresis loop is characteristic of the equilibration delay. It may also occur if an active metabolite is formed and the metabolite-to-parent compound ratio increases with time.

If the plasma concentration-effect plot suggests equilibration delay, then a plot using concentration in another compartment of the model may eliminate the hysteresis (fig. 12) and suggest a suitable pharmacodynamic model. The use of the tissue concentration

predicted by a 2-compartment model for LSD-25 allowed Wagner et al. (1968) to use a linear pharmacodynamic model to describe the drug's effect on mental performance (fig. 13).

The principal disadvantages of the pharmacokinetic compartment approach is the requirement for the time course of effect-site concentration to parallel the distribution of drug to those tissues that determine the multi-exponential character of the plasma concentration-time course. This may be reasonable if the effect-site tissue or organ has a large drug capacity, but there is no *a priori* reason to assume this will be so. Clearly the distribution of drug to some small tissue such as the aqueous humour of the eye will not have a discernible influence on the plasma concentration-time course. Penetration of drug into the aqueous humour of the eye may significantly lag behind plasma concentration, and it would be impossible to predict concentrations in the eye from plasma measurements alone.

### 2.3 Effect Compartment Model

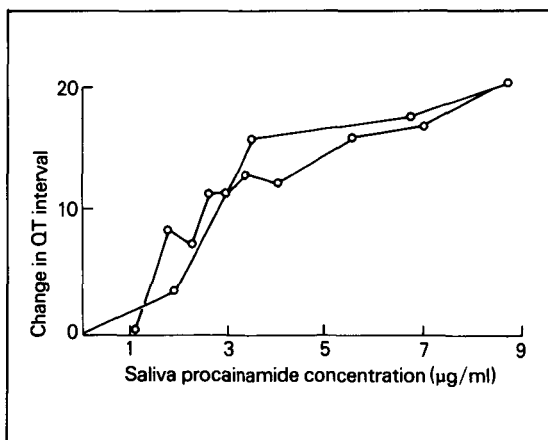If the drug enters and leaves the effect site by a first-order process, then the time course of drug ac-



Fig. 12. Use of saliva compartment concentrations to eliminate hysteresis of procainamide concentration on QT interval change [after Galeazzi et al. (1976)].
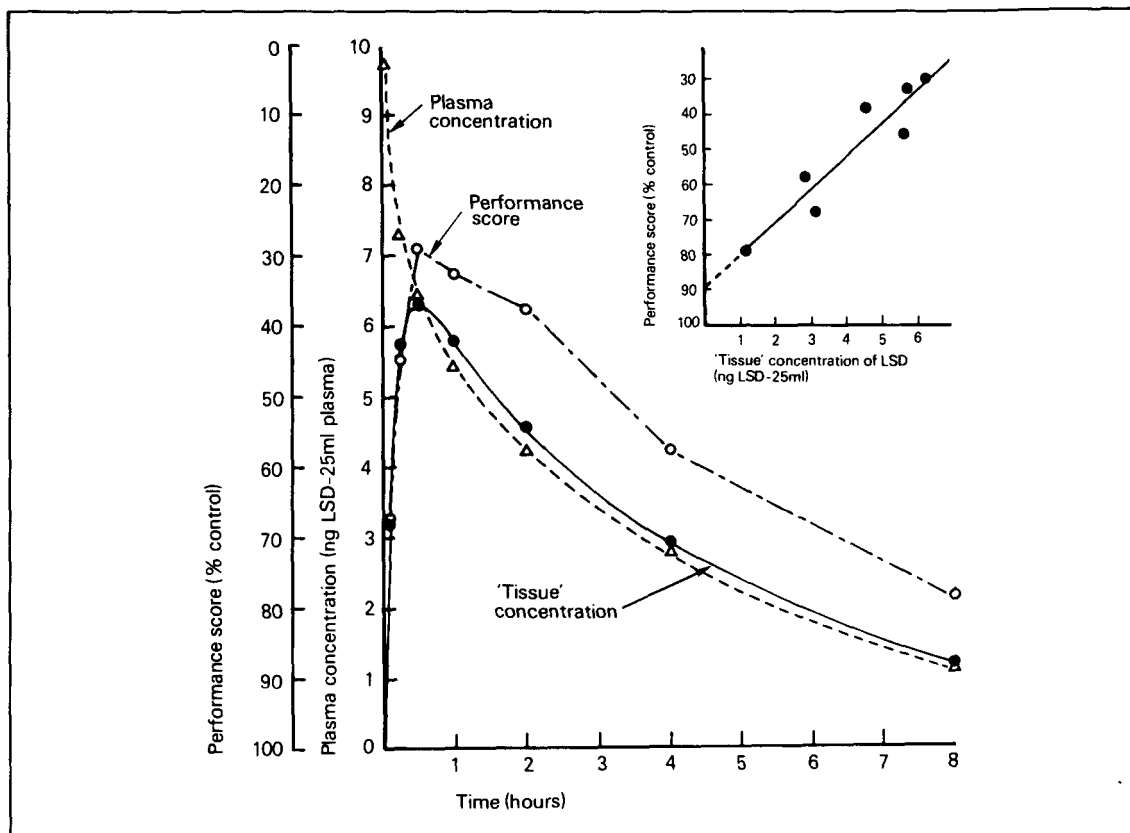
*Fig. 13.* Linear pharmacodynamic model describing the relationship between predicted tissue compartment concentration and effect of LSD on mental performance (inset) and explaining the time delay in onset of effect [after Wagner et al. (1968)].

cumulation at the effect site can be predicted if the plasma concentration is changed. Consider the hypothetical experiment illustrated in figure 14 in which plasma concentration is suddenly changed from zero to a new value C. The equilibrium effect corresponding to this concentration is $E_C$. The actual effect increases with time, and asymptotes to $E_C$. The rate of onset of effect will be controlled by the rate constant describing loss from the effect site, just as the time to steady-state concentration from a constant infusion is controlled by the drug elimination rate constant. The time to reach 50% of $E_C$ is the equilibration half-time and may be calculated from $\ln(2)/K_{e0}$ where $K_{e0}$ is the rate constant for drug loss from the effect site.

The first attempt to use this principle to estimate the equilibration half-time was by Forrester et al. (1974). They used the observations of Shapiro et al. (1970) on the time course of changes in electro-mechanical systole ($QS_2$) following bolus intravenous administration of a variety of cardiac glycosides (see fig. 15).

The equilibration half-time varied from 6 minutes for ouabain, to 23 minutes for digoxin and 58 minutes for digitoxin. Their calculations assume that plasma concentration remains constant throughout the development of drug effect. This was obviously not the case, and is demonstrated in figure 15 by the decreasing effect of ouabain after its peak.

*Table II.* Equations describing the effect-site concentration ($C_e$) for a variety of common pharmacokinetic models

| Drug input to plasma | Plasma pharmacokinetic model | Effect site concentration[1] |
|---|---|---|
| Bolus | 1-compartment | $\dfrac{D K_{e0}}{Vd} \left[ \dfrac{e^{-Kt}}{(K_{e0} - K)} + \dfrac{e^{-K_{e0}t}}{(K - K_{e0})} \right]$ |
| | 2-compartment | $\dfrac{D K_{e0}}{Vd} \left[ \dfrac{(K_{21} - \alpha) e^{-\alpha t}}{(\beta - \alpha)(K_{e0} - \alpha)} + \dfrac{(K_{21} - \beta) e^{-\beta t}}{(\alpha - \beta)(K_{e0} - \beta)} + \dfrac{(K_{21} - K_{e0}) e^{-K_{e0}t}}{(\alpha - K_{e0})(\beta - K_{e0})} \right]$ |
| | 3-compartment | $\dfrac{D K_{e0}}{Vd} \left[ \dfrac{(K_{21} - \alpha)(K_{31} - \alpha) e^{-\alpha t}}{(\beta - \alpha)(\gamma - \alpha)(K_{e0} - \alpha)} + \dfrac{(K_{21} - \beta)(K_{31} - \beta) e^{-\beta t}}{(\alpha - \beta)(\gamma - \beta)(K_{e0} - \beta)} + \dfrac{(K_{21} - \gamma)(K_{31} - \gamma) e^{-\gamma t}}{(\alpha - \gamma)(\beta - \gamma)(K_{e0} - \gamma)} + \dfrac{(K_{21} - K_{e0})(K_{31} - K_{e0}) e^{-K_{e0}t}}{(\alpha - K_{e0})(\beta - K_{e0})(\gamma - K_{e0})} \right]$ |
| First-order (absorption) | 1-compartment | $\dfrac{D KA K_{e0}}{Vd} \left[ \dfrac{e^{-Kt}}{(KA-K)(K_{e0} - K)} + \dfrac{e^{-KAt}}{(K-KA)(K_{e0} - KA)} + \dfrac{e^{-K_{e0}t}}{(K-K_{e0})(KA-K_{e0})} \right]$ |
| | 2-compartment | $\dfrac{D KA K_{e0}}{Vd} \left[ \dfrac{(K_{21} - \alpha) e^{-\alpha t}}{(\beta - \alpha)(KA - \alpha)(K_{e0} - \alpha)} + \dfrac{(K_{21} - \beta) e^{-\beta t}}{(\alpha - \beta)(KA - \beta)(K_{e0} - \beta)} + \dfrac{(K_{21} - KA) e^{-KAt}}{(\alpha - KA)(\beta - KA)(K_{e0} - KA)} + \dfrac{(K_{21} - K_{e0}) e^{-K_{e0}t}}{(\alpha - K_{e0})(\beta - K_{e0})(KA - K_{e0})} \right]$ |
| | 3-compartment | $\dfrac{D KA K_{e0}}{Vd} \left[ \dfrac{(K_{21} - \alpha)(K_{31} - \alpha) e^{-\alpha t}}{(\beta - \alpha)(\gamma - \alpha)(KA - \alpha)(K_{e0} - \alpha)} + \dfrac{(K_{21} - \beta)(K_{31} - \beta) e^{-\beta t}}{(\alpha - \beta)(\gamma - \beta)(KA - \beta)(K_{e0} - \beta)} + \dfrac{(K_{21} - \gamma)(K_{31} - \gamma) e^{-\gamma t}}{(\alpha - \gamma)(\beta - \gamma)(KA - \gamma)(K_{e0} - \gamma)} + \dfrac{(K_{21} - K_{e0})(K_{31} - K_{e0}) e^{-K_{e0}t}}{(\alpha - K_{e0})(\beta - K_{e0})(\gamma - K_{e0})(KA - K_{e0})} \right]$ |

| | 1-compartment | 2-compartment | 3-compartment |
|---|---|---|---|
| Zero-order | $$\frac{K^0 K_{e0}}{Vd} \left[ \frac{(e^{KT}-1)e^{-Kt}}{K(K_{e0}-K)} + \frac{(e^{K_{e0}t}-1)e^{-Kt}}{K_{e0}(K-K_{e0})} \right]$$ | $$\frac{K^0 K_{e0}}{Vd} \left[ \frac{(K_{21}-\alpha)(e^{\alpha T}-1)e^{-Kt}}{\alpha(\beta-\alpha)(K_{e0}-\alpha)} + \frac{(K_{21}-\beta)(e^{\beta t}-1)e^{-\beta t}}{\beta(\alpha-\beta)(K_{e0}-\beta)} + \frac{(K_{21}-K_{e0})(e^{K_{e0}T}-1)\, e^{-K_{e0}t}}{K_{e0}(\alpha-K_{e0})(\beta-K_{e0})} \right]$$ | $$\frac{K^0 K_{e0}}{Vd} \left[ \frac{(K_{21}-\alpha)(K_{31}-\alpha)(e^{\alpha T}-1)e^{-\alpha t}}{\alpha(\beta-\alpha)(\gamma-\alpha)(K_{e0}-\alpha)} + \frac{(K_{21}-\beta)(K_{31}-\beta)(e^{\beta T}-1)e^{-\beta t}}{\beta(\alpha-\beta)(\gamma-\beta)(K_{e0}-\beta)} \right.$$ $$\left. + \frac{(K_{21}-\gamma)(K_{31}-\gamma)(e^{\gamma T}-1)e^{-\gamma t}}{\gamma(\alpha-\gamma)(\beta-\gamma)(K_{e0}-\gamma)} + \frac{(K_{21}-K_{e0})(K_{31}-K_{e0})(e^{K_{e0}T}-1)e^{-K_{e0}t}}{K_{e0}(\alpha-K_{e0})(\beta-K_{e0})(\gamma-K_{e0})} \right]$$ |

1 *Abbreviations*: D = dose; Vd = central compartment volume of distribution; $K_{e0}$ = equilibration rate constant; $K^0$ = zero-order infusion rate; K, α, β, γ = exponential rate constants for 1-, 2- and 3-compartment models; KA = first-order absorption rate constant; $K_{21}$, $K_{31}$ = rate constant for transfer from compartment 2 or compartment 3 to compartment 1; t = time; T = t during infusion, then, after infusion has stopped, T = infusion duration.
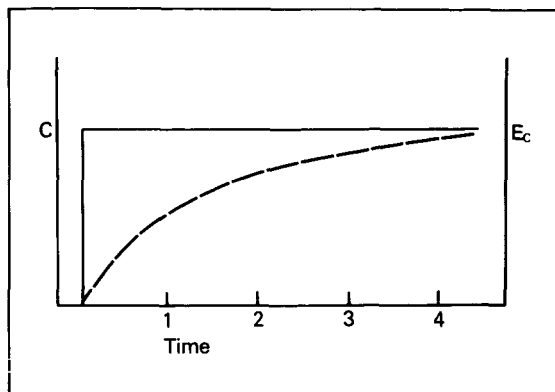
*Fig. 14.* The time course of onset of effect reaching an equilibrium value ($E_C$) corresponding to the concentration (C), which has been suddenly attained from a zero value.

The time scale is in units of the equilibration half-time [after Holford and Sheiner (1981a)].

A simple solution to the estimation of equilibration half-time when plasma concentrations are changing was proposed by Hull et al. (1978) and elaborated upon by Sheiner et al. (1979). This is illustrated in figure 16, where the plasma compartment is considered to be connected to an effect compartment site, with drug entry and loss from the effect compartment controlled by first-order processes.

For the hypothetical experiment described above (fig. 15) we proposed that the concentration in compartment 1 was suddenly changed from zero to a new constant value. If we now consider a more complex model to describe the changes in concentration in compartment 1, for example the predictions of some pharmacokinetic model describing plasma concentrations, then we can readily derive an expression for the concentration of drug in the effect compartment [see Holford and Sheiner (1981b) for a more detailed description].

Equations describing the effect-site concentration, $C_e$, for a variety of common pharmacokinetic models are listed in table II. These expressions use the parameters of the pharmacokinetic model for plasma concentrations and, additionally, an equilibration rate constant $K_{e0}$. This rate constant cannot be estimated directly from plasma concentrations of drug, but

given a fully specified pharmacokinetic model and a specific form for the pharmacodynamic model it may be estimated from the time course of drug effect. For instance, if we use the $E_{max}$ model for equation 1 we can write:

$$E_t = \frac{E_{max} \cdot C_e(t)}{EC_{50} + C_e(t)} \qquad \text{[Eq. 11]}$$

where $E_t$ are the effects and $C_e(t)$ the effect-site concentrations predicted by one of the equations in table II at time t. This combination of a pharmacokinetic model for the effect-site concentration and a pharmacodynamic model for effect can now be used to estimate the equilibration rate constant ($K_{e0}$); and also $E_{max}$ and $EC_{50}$, using a suitable non-linear regression method. This approach has been used for a variety of drugs, and some equilibration half-times are listed in table III.

## 3. Physiological Effect Models

When the effect of a drug on the body is not directly measurable, for example with enzyme inhibi-
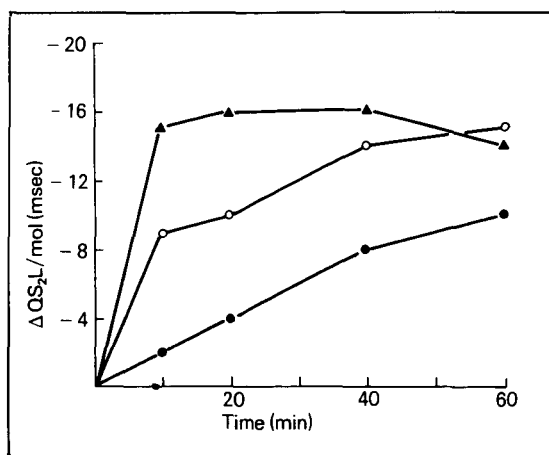


*Fig. 15.* The time course of onset of changes in electromechanical systole ($QS_2$) after bolus administration of ouabain (▲), digoxin (○) and digitoxin (●) [after Forrester et al. (1974), reproduced by permission of the American Heart Association Inc.].
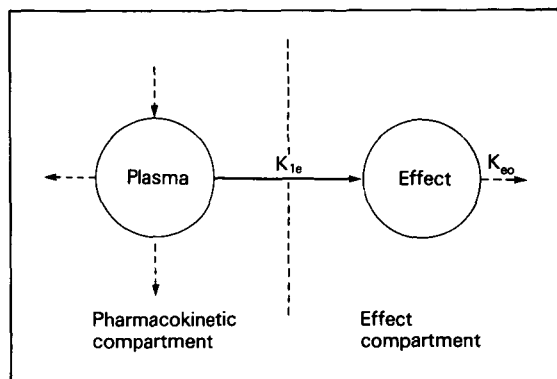
*Fig. 16.* Schematic representation of the link between the central compartment of a pharmacokinetic model and the effect compartment of a pharmacodynamic model.
The rate of loss of drug from the effect compartment is $K_{eo}$ and controls the equilibration half-life [after Holford and Sheiner (1981b)].

tion, the physiological consequences of this effect are often used as a measure of drug effect. Incorporation of what is known about the physiological system into a physiological effect model can enhance not only the understanding of the pharmacodynamics of the drug, but also of the physiological process. This concept may be illustrated by the effects of oral anticoagulants such as warfarin.

Warfarin is known to inhibit the conversion of a vitamin K metabolite, vitamin K epoxide, back to the parent compound, thereby limiting the availability of vitamin K to the carboxylation reaction forming prothrombin complex from precursors (Whitlon et al., 1978). A simple way to measure the activity of warfarin is to determine the prothrombin time, which can be converted to an index of prothrombin complex concentration, P. If warfarin is assumed not to influence the degradation of promthrombin complex, then the degradation rate can be estimated, along with the parameters of a pharmacodynamic model for warfarin's effect on prothrombin synthesis, by observing the time course of change in P.

The most commonly applied pharmacodynamic model for relating warfarin concentration to inhibition of prothrombin synthesis is the log-linear model

(Nagashima et al., 1969; Shepherd et al., 1979; Yacobi et al., 1980), but other models may be used. For example, Sheiner (1969) used a linear model with constraints on the fractional effect to lie between zero and 1, and Powers et al. (1980) have used the $E_{max}$ model. A potential advantage of the last 2 models is their ability to use equation 2 over the whole range of observations of prothrombin complex activity. The log-linear model can only be used with a pharmacokinetic model for warfarin starting at a non-zero warfarin concentration.

However, the log-linear model has been a powerful tool for describing the effects of the oral anticoagulants. Shepherd et al. (1979) studied the pharmacodynamics of warfarin in young and elderly patients. They concluded that the rate of prothrombin complex elimination was reduced by 50% in older people, but the concentration of warfarin required to reduce synthesis by 50%, and thus achieve the same steady-state prothrombin complex concentration, was the same in both populations. However, Routledge et al. (1979) claimed that older people are more sensitive to warfarin, based on the warfarin concentration required to prolong the prothrombin time by 20% (a fixed-effect model). It is not clear why this discrepancy exists.

Yacobi et al. (1980) examined the interaction between phenylbutazone and warfarin in rats. They clearly identified 2 separate mechanisms for the potentiation of anticoagulant effect. Phenylbutazone decreases warfarin elimination and displaces it from plasma proteins. Using total warfarin concentration to describe the effect on prothrombin complex synthesis, phenylbutazone appeared to potentiate warfarin by a factor of 30. However, when unbound warfarin concentrations were used the potentiation was only 3-fold, suggesting phenylbutazone has some other action. Lewis et al. (1974) and O'Reilly et al. (1980) have shown in man that phenylbutazone selectively decreases the elimination of the more potent S(−) isomer of warfarin, which could explain the findings in rats.

This application of a combined pharmacodynamic and physiological effect model is especially powerful

*Table III.* Equilibration half-times of some drugs, calculated using the approach discussed in the text (see section 2.3)

| Drug | Equilibration half-time (minutes) | Effect[1] | Reference |
|---|---|---|---|
| Disopyramide | 2 | QT prolongation | Whiting et al. (1980) |
| D-Tubocurarine | 4 | Muscle paralysis | Sheiner et al. (1979) |
| Quinidine | 8 | QT prolongation | Holford et al. (1981) |
| Cimetidine | 30 | Inhibition of gastric acid secretion | Holford et al. (in preparation) |
| Digoxin | 200 | LVET shortening | Kelman and Whiting (1980) |

1 *Abbreviations:* QT = electrocardiographic QT interval; LVET = left ventricular ejection time.

because it allows the estimation of the rate of prothrombin complex elimination without direct measurement of the prothrombin complex itself.

## 4. Dose-dependent Pharmacodynamics — Tolerance and Sensitisation

The pharmacodynamic models described so far may be applied to predict drug effects from a combination of doses by simply adding the concentrations to be expected from each dose. The application of this principle of additivity rests upon the parameters of the model, for example $E_{max}$ or $EC_{50}$, remaining fixed. It is well recognised that this is not the case for many drug actions and a variety of terms have been coined, usually to describe the decreased effect with repeated doses or prolonged exposure to the drug; for example tolerance, tachyphylaxis, down-regulation.

With improved understanding of receptor biology (Snyder, 1979), 2 receptor-based mechanisms for tolerance have been defined. The first, a decrease in receptor number, has been implicated in the reduced effectiveness of β-adrenoceptor agonists, such as isoprenaline (isoproterenol), and the term 'down-regulation' is now preferred for this specific mechanism. This process is dependent on drug concentration and duration of exposure and may take several hours to days to develop. A more rapid decrease in receptor activity has been shown in isolated cell systems and may involve the translocation of receptors within the membrane without a decrease in number. The combination of these slow and rapid mechanisms may be part of a cellular regulation system capable of adaptation to intense stimulation. The converse process of increased sensitivity may occur when receptor agonists are withdrawn, for example following pharmacological or surgical denervation, or when the receptor is blocked by an antagonist, for example propranolol. Prolonged propranolol use may cause an increase in receptor number which is expressed when the drug is withdrawn and enhanced sensitivity to β-agonists ensues (Boudoulas et al., 1977).

The decreased effect of repeated administration of a drug such as tyramine, which causes release of noradrenaline (norepinephrine) from nerve terminals, can best be explained in terms of a physiological model. The limiting process is not that the intrinsic ability of the nerve terminal to release stored noradrenaline has decreased, but that the rate of synthesis of noradrenaline is too low to maintain intraneuronal stores. The pharmacodynamic parameters for tyramine are probably unchanged, but its interaction with a component of a physiological model can be used to explain the apparent decrease in its effectiveness.

Many examples of tolerance that have no explanation are known in clinical practice. Psychoactive agents, for example opiates and barbiturates, are the best known examples. Because of the development of tolerance to the effects of drugs on higher mental functions across a wide variety of chemical agents, it seems reasonable to suppose that the underlying mechanism may involve a homeostatic process like that described above. A good example of rapid development of tolerance is shown by the work of van Dyke et al. (1978). They studied the euphoria ('high') following cocaine administration by the intranasal route in 'experienced' volunteers. When the degree of euphoria is plotted against the cocaine plasma concentration, the time sequence of points follows a clockwise hysteresis loop, i.e. less effect is seen at the same concentration at later points. This form of hysteresis loop is pathognomic of tolerance (fig. 17). Mayersohn et al. (1978) analysed the concentration effect data from the same study after oral dosing and used it to demonstrate Levy's principle (1964) of linear disappearance of effect with time (see below). However, the concentration-effect plot of the data is linear in the range of observations studied, and it seems likely that an exponential decay of effect would describe the data as well. It is not feasible to distinguish between the models with such a limited range of concentrations (5-fold). It is interesting to note following oral administration that the sensitivity to cocaine in the later part of the experiment was greater than that following intranasal doses, and there was no evidence of development of tolerance.

Besides a change in receptor sensitivity or number, the clockwise hysteresis loop may also be explained by accumulation of an antagonistic metabolite. Collste et al. (1980) have shown that the effects of metoprolol, a β-adrenoceptor antagonist, on heart rate are less at the same concentration when patients receive regular intermittent doses of 100mg compared with 50mg. Because metoprolol is extensively metabolised, this may be explained by the accumulation of a metabolite which itself is chronotropic. Furthermore there may be physiological adjustment of the cardiovascular system, for example a reduction

in vagal tone, which may be dependent on the dose-related fall in blood pressure.

When the mechanism of reduced drug effectiveness following prolonged or repeated exposure is unknown, the term 'tolerance' is used to describe the phenomenon. If the mechanism involves a drug metabolite, we can classify it more precisely as a drug combination (see section 5). The converse process of increased drug effect with continued exposure may be called sensitisation. The work of Bean et al. (1979) provides an example of sensitisation which occurred following chronic low level infusion of angiotensin in dogs. The blood pressure response to an acute increase in infusion rate was greater the longer the chronic infusion had been maintained.

## 5. Drug Combinations

The combination of 2 drugs may produce alterations in drug responses attributable to pharmacokinetic or pharmacodynamic changes. This discussion is restricted to the pharmacodynamic mechanisms of drug interaction. Complex models of the effects of drug combinations have been proposed and applied *in*
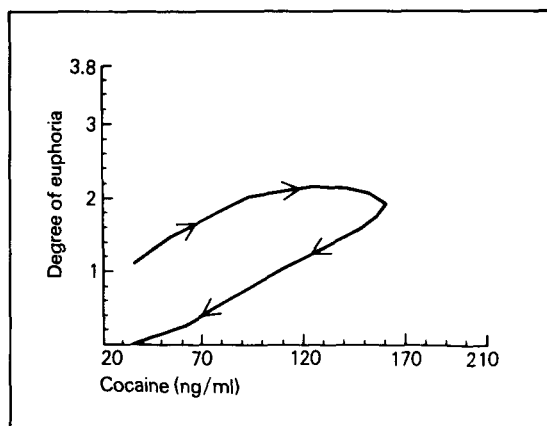
*Fig. 17.* Clockwise hysteresis loop typical of tolerance is seen after intranasal administration of cocaine when related to degree of euphoria experienced in volunteers [after van Dyke et al. (1978)].

*vitro* (e.g. Ariens and Simonis, 1964a,b). We have reviewed elsewhere those models of this type that may find practical application *in vivo* (Holford and Sheiner, 1981b). Most pharmacodynamic drug interactions of clinical importance can be explained by simple predictions of the pharmacodynamic models described above.

## 5.1 Drug-Drug Interactions

Drug combinations are frequently used in the treatment of malignant disease (e.g. cyclophosphamide and prednisone), hypertension (e.g. propranolol and chlorothiazide), airways obstruction (e.g. theophylline and ephedrine), and epilepsy [e.g. phenytoin and phenobarbitone (phenobarbital)]. The advantage of this approach is the ability to achieve a given response by additive therapeutic effects using drugs without additive toxicity. It is usually believed that greater efficacy is possible with the combination than with either drug alone. For example, if an $E_{max}$ model is used to describe the effects of 2 drugs, A and B, independently, then their combination could be described by:

$$E_{AB} = \frac{E_{max_A} A}{EC_{50_A} + A} + \frac{E_{max_B} B}{EC_{50_B} + B} \quad \text{[Eq. 12]}$$

If the observed effect of the combination $E_{AB}$ is greater than predicted by the additive model above, then the drugs are said to act synergistically. The application of such a model to the study of first the individual and then the combined effects of drugs would allow precise statements to be made about the additive or synergistic nature of interactions.

If the drugs are thought to share a common intermediate step, for example the accumulation of cyclic AMP following treatment with a β-adrenoceptor agonist or a phosphodiesterase inhibitor, a more complex model may be required to describe the additive effects because the maximum response of either drug may be limited by the efficacy of cyclic AMP (see for example Gero, 1978).

Steiness et al. (1980) have used a linear model for digoxin to demonstrate a pharmacodynamic interaction with quinidine. The effects of digoxin on the heart were measured by changes in left ventricular ejection time (LVET) and when combined with quinidine the effects of digoxin at the same plasma concentration were abolished.

The pharmacokinetic interaction between digoxin and quinidine involving a reduction in both clearance and volume of distribution (Hager et al., 1979) of digoxin may be viewed as an example of a pharmacodynamic process, i.e. a drug effect of quinidine. Powell et al. (1980) have provided evidence that the changes in digoxin pharmacokinetics are quinidine dose- and, presumably, concentration-dependent. These effects may be incorporated into a pharmacokinetic model for digoxin which includes a pharmacodynamic model for the effect of quinidine. This is elaborated below to provide an example of the types of models considered here.

For simplicity we may describe the plasma concentration of digoxin during elimination by a 1-compartment model. If we ignore (for clarity) the effect of quinidine on digoxin clearance we can write:

$$\frac{dD}{dt} = \frac{-Cl}{Vd} \cdot D \quad \text{[Eq. 13]}$$

where D is the plasma concentration of digoxin, Cl is digoxin clearance, and Vd is the volume of distribution of digoxin without quinidine. We can modify this expression to include the effects of quinidine on Vd:

$$\frac{dD}{dt} = \frac{-Cl}{Vd \cdot f(Q)} \cdot D \quad \text{[Eq. 14]}$$

where f(Q) is a pharmacodynamic model for the effect of quinidine concentration Q, as a function of time on the volume of distribution. For example, the equation would now become:

$$f(Q) = \left(1 - \frac{FV \cdot Q}{IC_{50} + Q}\right) \quad \text{[Eq. 15]}$$

where $IC_{50}$ is the quinidine concentration reducing by 50% that fraction of the volume of distribution of

digoxin that quinidine can modify (FV). Using this differential equation definition of digoxin plasma concentration and a pharmacokinetic model for quinidine, it should be possible to estimate the value of $IC_{50}$ by observing the concentration-time course of digoxin after administration of quinidine. In practice, the model would need to be expanded to describe the multicompartmental disposition of digoxin and the effects of quinidine on digoxin clearance.

### 5.2 Drug-Metabolite Interactions

The metabolism of a parent compound and accumulation of a metabolite may be responsible for apparent changes in response based on measurement of the parent-drug concentration alone. If the ratio of parent to metabolite concentration is relatively constant, the contribution of the metabolite to the overall effect may be impossible to determine from administration of the parent drug alone. Advantage may be taken of heterogeneity in the ability of a population to metabolise the parent compound in order to study the influence of differing parent-metabolite ratios.

The antiarrhythmic efficacy of N-acetyl-procainamide (NAPA), a major metabolite of procainamide, has been established in man (e.g. Lertora et al., 1979). Schroeder et al. (1979) examined the contribution of N-acetyl-procainamide to the suppression of ventricular extrasystoles in patients following a myocardial infarction who were receiving procainamide. The ratio of N-acetyl-procainamide to procainamide varies several-fold depending on the acetylation phenotype and renal function. Schroeder and co-workers examined the antiarrhythmic efficacy of procainamide with relatively high or low concentrations of procainamide in different patients, and came to the conclusion that N-acetyl-procainamide appeared to antagonise the effects of procainamide. This conclusion has not been confirmed, however, and is not supported by other studies of procainamide and N-acetyl-procainamide combinations (Elson et al., 1975).

Parent to metabolite concentration ratios may also be altered by administration of the drug by intravenous and oral routes. If there is extensive first-pass metabolism following oral dosing the concentration of metabolite may be much higher at the same parent-drug concentration when compared with intravenous administration. We have used this phenomenon to examine the effects of quinidine metabolites on cardiac repolarisation. Following intravenous administration the change in QT interval of the electrocardiogram (ECG) was only 20msec for each mg/L change in estimated quinidine concentration in the effect compartment compared with 34msec per mg/L change in quinidine concentration after an oral dose (Holford et al., 1981) [fig. 18]. We attribute this difference to the presence of one or more metabolites of quinidine formed during absorption. We have applied the same method to study the contribution of disopyramide metabolites and conclude that their contribution is negligible after a single oral dose (Whiting et al., 1980). Eichelbaum et al. (1980) observed the effect of verapamil on the PR interval of the ECG after intravenous and oral doses. They showed that PR interval prolongation was greater at the same verapamil concentration after intravenous doses and suggested that there was stereoselective metabolism of verapamil following oral doses, with the appearance of a larger fraction of a relatively inactive metabolite in the plasma.

## 6. Pharmacodynamic Bioavailability

The extent of drug bioavailability is usually defined in terms of the area under the concentration-time curve in comparison with the area following administration of a standard preparation, for example following intravenous administration. Pharmacodynamic bioavailability may also be defined in terms of the area under the pharmacological effect-time curve in comparison with a standard. This is particularly appropriate when adequate assays for the plasma concentration of the drug are not available. In the same way that pharmacokinetic bioavailability studies are limited to drugs with concentration-inde-

pendent or linear pharmacokinetic models, pharmacodynamic bioavailability studies should be limited to linear pharmacodynamic models. Thus, suppose 2 drugs are administered with identical rates of absorption but their extent of absorption differs 10-fold. The time course of concentration will be identical for the 2 drugs, but will differ 10-fold in magnitude. If the concentrations of the less available drug cause a maximum drug effect for a large fraction of the time, the drugs will have similar areas under the effect-time curve and appear to have the same pharmacodynamic bioavailability. Repeating the studies with a much lower dose where concentrations are now in the linear part of the concentration-effect curve will reveal that one drug is indeed less bioavailable.

The pharmacodynamic equivalence of different erythrityl tetranitrate preparations on blood vessels has been assessed using the approach of calculating the area under the effect-curve (Hannemann et al., 1981). Extrapolation of the conclusions to lower doses may not be applicable, but there is some evidence that effects were not at their maximum in this study. We have shown that the effect of quinidine on the QT interval is linear in the concentration range studied (up to 4mg/L), and we used the area under the effect-curve to show that oral doses are bioequivalent to intravenous doses even though the pharmacokinetic bioavailability is only 70% by the oral route (Guentert et al., 1979; Holford et al., 1981).

## 7. Influence of Protein Binding

It is commonly held that drug effects are mediated by unbound drug, but there are few examples where this has clearly been shown to be the case. Reidenberg et al. (1971) observed that patients with renal failure receiving phenytoin for epilepsy were adequately controlled at total drug concentrations which would be ineffective in patients with normal renal function. Phenytoin binding was found to be markedly reduced in renal failure, and unbound drug concentrations were found to be the same as in patients without renal failure. This application of a fixed-effect pharmaco-
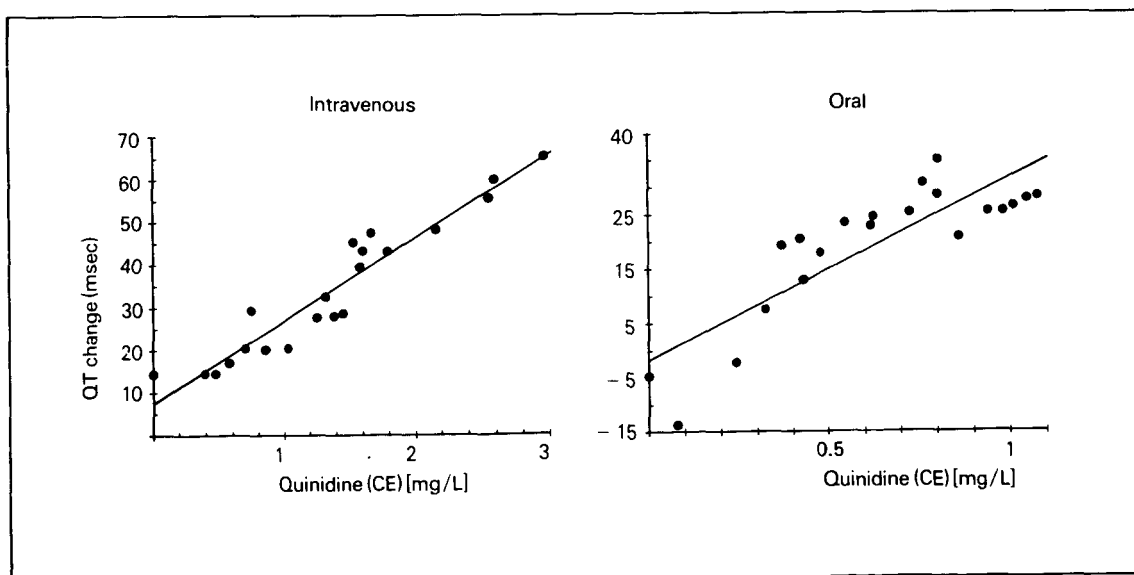


Fig. 18. Effect of quinidine on QT interval after intravenous and oral dosing.

The change in QT interval is greater after oral dosing per mg/L change in quinidine concentration estimated in the effect compartment (CE) [after Holford et al. (1981)].

dynamic model shows the importance of the unbound drug concentration in determining the anticonvulsant effect of phenytoin.


## 8. Dosing Regimen Design

The combination of pharmacokinetic and pharmacodynamic models permits rational designing of dosing regimens. The choice of dose size and dosing interval depends not only on the parameters of drug disposition and the relationship between concentration and effect, but also on the nature of the desired therapeutic response. If the desired response is related directly to the drug effect at any time, for example muscle paralysis produced by d-tubocurarine, then we can use an instantaneous pharmacokinetic-pharmacodynamic model to maintain drug concentrations within the range producing the therapeutic response. On the other hand, if the therapeutic response is not related directly to the effect at any instant, but to the overall or integrated response over some period of time, then we need to use an integrated pharmacokinetic-pharmacodynamic model. For example, the treatment of peptic ulceration may depend more upon the reduction in average daily acid output than the instantaneous inhibition of secretion with a drug like cimetidine.

### 8.1 Dosing for Instantaneous Response

The time course of drug effect after administration of a dose can be considered in 3 phases:

*Phase 1 — $EC_{80} < Cp$:* When the plasma concentration (Cp) is greater than the concentration producing 80% of drug effect ($EC_{80}$), the relationship between concentration and effect is shallow. A 9% change in effect, from 90 to 99%, requires a 90-fold change in drug concentration (table I). In the range of concentrations greater than $EC_{80}$, changes in drug concentration appear to have little influence on drug effect. For a drug like propranolol, whose elimination half-life is only 2 to 3 hours, a single daily dose may

nonetheless maintain blood levels greater than the $EC_{50}$ for most of the day [using an $EC_{50}$ of 10ng/ml (Esler et al., 1977)]. This may explain the apparent efficacy of once-daily dosing for the treatment of hypertension.

*Phase 2 — $EC_{20} < Cp < EC_{80}$:* Levy (1964) pointed out that response will decline linearly with time with concentrations in the range of 20 to 80% of maximum response, in contrast to the exponential decay of plasma concentration. He illustrated this with the time course of recovery of muscle strength after d-tubocurarine administration observed by Bellville et al. (1964).

*Phase 3 — $Cp < EC_{20}$:* When the effect is less than 20% of maximum, for practical purposes it will be directly proportional to concentration and will decline exponentially with time in parallel with the plasma level.

The instantaneous effect after a given series of doses can be predicted at any time through the use of a pharmacokinetic model to predict the concentration at that time by summing the contribution of the individual doses. If a non-linear pharmacodynamic model is used it is important to sum the concentrations due to multiple doses before predicting the effect, rather than summing the individual effects predicted by the contribution of each dose to the concentration.

### 8.2 Dosing for Overall Response

Wagner (1968) pointed out the consequences of a non-linear concentration-effect relationship when the same dose is administered in divided doses at different intervals. He used the observations of Murphy et al. (1961) to propose a non-linear concentration-effect relationship for the effects of chlorothiazide. When the same dose per day, 2g, was given at intervals of from 3 hours up to 24 hours, the sodium excretion decreased with lengthening dosing interval. This can be understood in terms of phases 1 and 3 of the effect-time course described above (see section 8.1). During phase 3 the instantaneous sodium excretion is directly

proportional to concentration, but in phase 1 the excretion is proportionately much less. Although the average daily concentration is the same, long dosing intervals cause concentrations to be in phase 3 for more of the time than when shorter intervals are used. Therefore the overall effect of sodium excretion is less with longer intervals. The overall effect, $\bar{E}$ can be calculated using

$$E = \frac{E_{max}}{\tau} \int_0^\tau \frac{C(t,\tau) \cdot dt}{EC_{50} + C(t,\tau)} \qquad [Eq.\ 16]$$

where $C(t,\tau)$ is a function yielding the concentration at time t in the dosing interval $\tau$. Using this expression we have calculated the overall effect of cimetidine on daily gastric acid secretion for total daily doses up to 2000mg administered as a constant infusion ($\tau = 0$) or at intervals of 6, 12 and 24 hours (fig. 19). Typical pharmacokinetic parameters for cimetidine in patients with duodenal ulcer were used (Somogyi et al., 1980) and an $EC_{50}$ of 0.5mg/L used for inhibition of gastric acid secretion (Burland et al., 1975; Kaojarern et al., 1981). Figure 19 illustrates that the same overall inhibition of acid output can be achieved with a lower dose if the dosing interval is shortened.
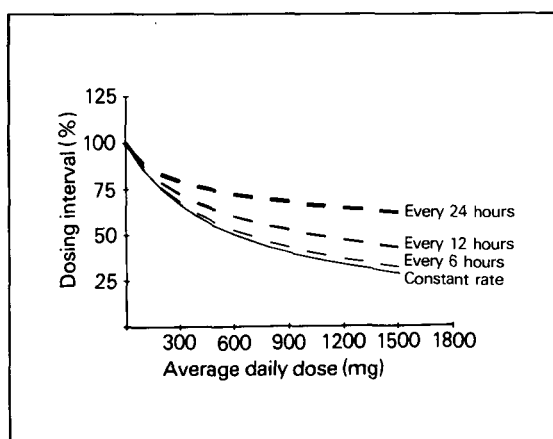
The simulation assumes rapid drug absorption, but enhancement of effect could be produced for each dosing interval if a sustained-release preparation were used.

As a further example of this concept, differences in therapeutic outcome have been reported for alternate-day corticosteroid treatment compared with once-daily dosing. The efficacy of prednisone in the treatment of cranial arteritis was impaired to a clinically important degree when the drug was given on alternate days (Hunder et al., 1975). We may deduce that the pharmacodynamics of prednisone are non-linear for the treatment of this condition (Holford and Sheiner, 1981b).

As a final example, the enhanced nephrotoxicity in dogs of gentamicin and tobramycin when administered as a bolus every 24 hours when compared with a constant infusion (Reiner et al., 1978) suggests a non-linear concentration-toxicity relationship for these drugs. A mechanism for this non-linearity is suggested by the work of Whelton et al. (1978) who have shown saturable uptake of aminoglycosides into renal cortical tissue in the range of concentrations achieved by Reiner et al. (1978).

### Acknowledgement

### References

Ariens, E.J. and Simonis, A.M.: A molecular basis for drug action. Journal of Pharmacy and Pharmacology 27: 137-257 (1964a).

Ariens, E.J. and Simonis, A.M.: A molecular basis for drug action. The interaction of one or more drugs with different receptors. Journal of Pharmacy and Pharmacology 16: 289-312 (1964b).

Bean, B.L.; Bronn, J.J.; Casals-Stenzel, J.; Fraser, R.; Lerer, A.F.; Millar, J.A. and Norton, J.J.: The relation of arterial pressure and plasma angiotensin II concentration. Circulation Research 44: 452-458 (1979).



Fig. 19. The influence of dosing interval on the overall inhibition of acid secretion by cimetidine using an $IC_{50}$ of 0.5mg/L.

Beller, G.A.; Smith, T.W.; Abelman, W.H.; Haber, E. and Hood, W.B., Jr: Digitalis intoxication. A prospective clinical study with serum level correlations. New England Journal of Medicine 284: 989-997 (1971).

Bellville, J.R.; Cohen, E.N. and Hamilton, J.: The interaction of morphine and d-tubocurarine on respiration and grip strength in man. Clinical Pharmacology and Therapeutics 5: 35 (1964).

Benet, L.Z.: General treatment of linear mammillary models with elimination from any compartment as used in pharmacokinetics. Journal of Pharmaceutical Sciences 61: 536-541 (1972).

Boudoulas, H.; Lewis, R.P.; Kates, R.E. and Dalamangas, G.: Hypersensitivity to adrenergic stimulation after propranolol withdrawal in normal subjects. Annals of Internal Medicine 87: 433-436 (1977).

Box, G.E.P. and Lucas, H.L.: Design of experiments in non-linear situations. Biometrika 46: 77-90 (1959).

Breckenridge, A.; Orme, M.; Wesseling, H.; Lewis, J.R. and Gibbons, R.: Pharmacokinetics and pharmacodynamics of the enantiomers of warfarin in man. Clinical Pharmacology and Therapeutics 15: 424-430 (1974).

Burland, W.L.; Duncan, W.A.M.; Hesselbo, T.; Mills, J.G. and Sharpe, P.C.: Pharmacological evaluation of cimetidine, a new histamine $H_2$-receptor antagonist in healthy man. British Journal of Clinical Pharmacology 2: 481-486 (1975).

Collste, P.; Haglund, K. and von Bahr, C.: Plasma levels and effects of metoprolol after single and multiple oral doses. Clinical Pharmacology and Therapeutics 27: 441-449 (1980).

Di Stefano, J.J.: Optimized blood sampling protocols and sequential design of kinetic experiments. American Journal of Physiology 240: R259-R265 (1981).

Eichelbaum, M.; Birkel, P.; Grube, E.; Gutgemann, U. and Somogyi, A.: Effects of verapamil on P-R intervals in relation to verapamil plasma levels following single i.v. and oral administration and during chronic treatment. Klinische Wochenschrift 58: 919-925 (1980).

Elson, J.; Strong, J.M.; Lee, W.K. and Atkinson, A.J.: Antiarrhythmic potency of N-acetylprocainamide. Clinical Pharmacology and Therapeutics 17: 134-140 (1975).

Esler, M.; Zweifer, A.; Randall, O. and DeQuattro, V.: Pathophysiologic and pharmacokinetic determinants of the antihypertensive response to propanolol. Clinical Pharmacology and Therapeutics 22: 299-308 (1977).

Forrester, W.; Lewis, R.P.; Weissler, A.M. and Wilke, T.A.: The onset and magnitude of the contractile response to commonly used digitalis glycosides in normal subjects. Circulation 49: 517-521 (1974).

Galeazzi, R.L.; Benet, L.Z. and Sheiner, L.B.: Relationship between the pharmacokinetics and pharmacodynamics of procainamide. Clinical Pharmacology and Therapeutics 20: 278-289 (1976).

Gero, A.: Intimate study of drug action III: Mechanisms of molecular drug action; in DiPalma (Ed) Drill's Pharmacology in Medicine, pp.67-98 (McGraw-Hill, New York 1971).

Gibaldi, M. and Perrier, D.: Pharmacokinetics (Marcel Dekker, New York 1975).

Guentert, T.W.; Holford, N.H.G.; Coates, P.C.; Upton, R.A. and Riegelman, S.: Quinidine pharmacokinetics in man: Choice of a disposition model and absolute bioavailability studies. Journal of Pharmacokinetics and Biopharmaceutics 7: 315-330 (1979).

Hager, W.D.; Fenster, P.; Mayersohn, M.; Perrier, D.; Graves, P.; Marcus, F.I. and Goldman, S.: Digoxin-quinidine interaction. New England Journal of Medicine 302: 1238-1241 (1979).

Hannemann, R.E.; Randall, J.E.; Stoltman, W.P.; Bronsen, E.C.; Williams, E.J.; Long, R.A.; Hull, J.H. and Starbuck, R.R.: Digital plethysmography for assessing erythrityl tetranitrate bioavailability. Clinical Pharmacology and Therapeutics 29: 35-39 (1981).

Hill, A.V.: The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. Journal of Physiology 40: iv-vii (1910).

Holford, N.H.G.; Coates, P.E.; Guentert, T.W.; Riegelman, S. and Sheiner, L.B.: The effect of quinidine and its metabolites on the electrocardiogram and systolic time intervals: Concentration-effect relationships. British Journal of Clinical Pharmacology 11: 187-195 (1981).

Holford, N.H.G. and Sheiner, L.B.: Pharmacokinetic and pharmacodynamic modelling in vivo. Critical Revues in Bioengineering. In press (1981a).

Holford, N.H.G. and Sheiner, L.B.: Kinetics of pharmacologic response. Pharmacology and Therapeutics. In press (1981b).

Hull, C.J.; van Beem, H.B.H.; McLeod, K.; Sibbals, A. and Watson, M.J.: A pharmacodynamic model for pancuronium. British Journal of Anaesthesia 50: 1113-1123 (1978).

Hunder, G.G.; Sheps, S.G.; Allen, G.L. and Joyce, J.W.: Daily and alternate-day corticosteroid regimens in treatment of giant cell arteritis. Annals of Internal Medicine 82: 613-618 (1975).

Kaojarern, S.; Feldman, M.; Richardson, C.T. and Brater, D.C.: Tiotidine and cimetidine — kinetics and dynamics. Clinical Pharmacology and Therapeutics 29: 198-202 (1981).

Kelman, A.W. and Whiting, B.: Modelling of drug response in individual subjects. Journal of Pharmacokinetics and Biopharmaceutics 8: 115-130 (1980).

Kramer, W.G.; Kolibash, A.J.; Lewis, R.P.; Bathala, M.S.; Visconti, J.A. and Reuning, R.H.: Pharmacokinetics of digoxin: Relationship between intensity and predicted compartmental drug levels in man. Journal of Pharmacokinetics and Biopharmaceutics 7: 47-61 (1979).

Lertora, J.J.L.; Atkinson, A.J.; Kushner, W.; Nevin, M.J.; Lee, W.K.; Jones, C. and Schmid, F.R.: Long term antiarrhythmic therapy with N-acetylprocainamide. Clinical Pharmacology and Therapeutics 25: 273-282 (1979).

Levy, G.: Relationship between rate of elimination of tubocurarine and rate of decline of its pharmacological activity. British Journal of Anaesthesia 36: 694-695 (1964).

Lewis, R.J.; Trager, W.F.; Chan, K.K.; Breckenridge, A.; Orme, M.; Roland, M. and Schary, W.: Warfarin-stereochemical aspects of its metabolism and the interaction with phenylbutazone. Journal of Clinical Investigation 53: 1607-1617 (1974).

Mayersohn, M. and Perrier, D.: Kinetics of pharmacologic response to cocaine. Research Communications in Chemical Pathology and Pharmacology 22: 465-474 (1978).

Meffin, P.J.; Winkle, R.A.; Blaschke, T.F.; Fitzgerald, J. and Harrison, D.C.: Response optimization of drug dosage: Antiarrhythmic studies with tocainide. Clinical Pharmacology and Therapeutics 22: 42-57 (1977).

Mitenko, P.A. and Ogilvie, R.I.: Rational intravenous doses of theophylline. New England Journal of Medicine 289: 600-603 (1973).

Murphy, J.; Casey, W. and Lasagna, L.: The effect of dosage regimes on the diuretic efficacy of chlorothiazide in human subjects. Journal of Pharmacology and Experimental Therapeutics 134: 286-290 (1961).

Nagashima, R.; O'Reilly, R.A. and Levy, G.: Kinetics of pharmacologic effects in man: The anticoagulant action of warfarin. Clinical Pharmacology and Therapeutics 10: 22-35 (1969).

O'Reilly, R.A.: Studies on the optical enantiomorphs of warfarin in man. Clinical Pharmacology and Therapeutics 16: 348-354 (1974).

O'Reilly, R.A.; Trager, W.F.; Motley, C.H. and Howald, W.: Stereoselective interaction of phenylbutazone with $[12_C/13_C]$ warfarin pseudoracemates in man. Journal of Clinical Investigation 65: 746-753 (1980).

Powell, R.; Fenster, P.; Wandell, M.; Hager, D.; Graves, P.; Conrad, K. and Goldman, S.: Quinidine-digoxin interaction: Multiple-dose pharmacokinetics. Clinical Pharmacology and Therapeutics 27: 279 (1980).

Powers, W.F.; Abbrecht, P.J. and Covell, D.G.: Systems and microcomputer approach to anticoagulant therapy. IEEE Transactions on Bio-Medical Engineering. BME-27: 520-523 (1980).

Reidenberg, M.; Odar-Cederlof, I.; Bahr, von C.; Borga, O. and Sjoqvist, F.: Protein binding of diphenylhydantoin and dismethylimipramine in plasma from patients with poor renal function. New England Journal of Medicine 285: 264-267 (1971).

Reiner, N.E.; Bioxham, D.D. and Thompson, W.L.: Nephrotoxicity of gentamicin and tobramycin given once daily or continuously in dogs. Journal of Antimicrobial Chemotherapy 4: 85-101 (1978).

Routledge, P.A.; Chapman, P.H.; Davies, D.M. and Rawlins, M.D.: Pharmacokinetics and pharmacodynamics of warfarin at steady state. British Journal of Clinical Pharmacology 8: 243-247 (1979).

St John, R.C. and Draper, N.R.: D-optimality for regression designs — a review. Technometrics 17: 15-23 (1975).

Schroeder, P.; Klitgaard, N.A. and Simensen, E.: Significance of the acetylation phenotype and the therapeutic effect of procainamide. European Journal of Clinical Pharmacology 15: 63-68 (1979).

Shapiro, W.; Narhara, K. and Taubert, K.: Relationship of sma digoxin and digoxin to cardiac response following intravenous digitalization in man. Circulation 42: 1065-1072 (1970).

Sheiner, L.B.: Computer-aided long-term anticoagulation therapy. Computers and Biomedical Research 2: 307-318 (1969).

Sheiner, L.B.; Stanski, D.R.; Vozeh, S.; Miller, R.D. and Ham, J.: Simultaneous modelling of pharmacokinetics and pharmacodynamics: Application to d-tubocurarine. Clinical Pharmacology and Therapeutics 25: 358-371 (1979).

Shepherd, A.M.M.; Wilson, M. and Stevenson, I.H.: Warfarin sensitivity in the elderly; in Crooks and Stevenson (Eds) Drugs and the Elderly, pp.199-209 (University Park Press, Baltimore 1979).

Singh, B.N.; Williams, F.M.; Whitlock, R.M.; Collett, J. and Chew, C.: Plasma timolol levels and systolic time intervals. Clinical Pharmacology and Therapeutics 27: 159-166 (1980).

Snyder, S.H.: Receptors, neurotransmitters and drug responses. New England Journal of Medicine 300: 465-472 (1979).

Somogyi, A.; Rohner, H.G. and Gugler, R.: Pharmacokinetics and bioavailability of cimetidine in gastric and duodenal ulcer patients. Clinical Pharmacokinetics 5: 84-94 (1980).

Steiness, E.; Waldorff, S.; Hansen, P.B.; Kjaergård, H.; Buch, J. and Egeblad, H.: Reduction of digoxin-induced inotropism during quinidine administration. Clinical Pharmacology and Therapeutics 27: 791-795 (1980).

Van Dyke, C.; Jatlow, P.; Ungerer, J.; Barash, P.G. and Byck, R.: Oral cocaine; plasma concentration and central effects. Science 200: 211-213 (1978).

Vaughan, D.P. and Trainor, A.: Derivation of general equations for linear mammillary models when the drug is administered by different routes. Journal of Pharmacokinetics and Biopharmaceutics 3: 203-218 (1975).

Veng Pedersen, P.: Model-independent method of analyzing input in linear pharmacokinetic systems having polyexponential impulse response. I: Theoretical analysis. Journal of Pharmaceutical Sciences 69: 298-305 (1980).

Wagner, J.G.; Agahajanian, G.K. and Bing, O.H.: Correlation of performance test scores with 'tissue concentration' of lysergic acid diethylamide in human subjects. Clinical Pharmacology and Therapeutics 9: 635-638 (1968).

Wagner, J.G.: Kinetics of pharmacologic response. I: Proposed relationships between response and drug concentration in the intact animal and man. Journal of Theoretical Biology 20: 171-201 (1968).

Wagner, J.G.: Fundamentals of Clinical Pharmacokinetics (Drug Intelligence, Hamilton 1975).

Whelton, A.; Carter, G.G.; Craig, T.J.; Bryant, H.H.; Herbst, D.V. and Walker, W.G.: Comparison of the intrarenal disposition of tobramycin and gentamicin: Therapeutic and tox-

icologic answers. Journal of Antimicrobial Chemotherapy 4: 13-22 (1978).

Whitfield, L.R. and Levy, G.: Relationship between concentration and anticoagulant effect of heparin in plasma of normal subjects: Magnitude and predictability of interindividual differences. Clinical Pharmacology and Therapeutics 28: 509-516 (1980).

Whiting, B.; Holford, N.H.G. and Sheiner, L.B.: Quantitative analysis of the disopyramide concentration-response relationship. British Journal of Clinical Pharmacology 9: 67-75 (1980).

Whitlon, D.S.; Sadowski, J.A. and Suttie, J.W.: Mechanism of coumarin action: Significance of vitamin K epoxide reductase inhibition. Biochemistry 17: 1371-1377 (1978).

Winkle, R.A.; Meffin, P.J.; Fitzgerald, J.W. and Harrison, D.C.: Clinical efficacy and pharmacokinetics of a new orally effective antiarrhythmic tocainide. Circulation 54: 884-889 (1976).

Yacobi, A.; Chii-Ming, L. and Levy, G.: Comparative pharmacokinetics of coumarin anticoagulants. LXV: Pharmacokinetic and pharmacodynamic studies of acute interaction between warfarin and phenylbutazone in rats. Journal of Pharmaceutical Sciences 69: 14-20 (1980).

Author's address: *Nicholas H.G. Holford*, Division of Clinical Pharmacology, 926 Science, University of California Medical Center, *San Francisco, CA 94143* (USA).