Journal of Pharmacokinetics and Pharmacodynamics
https://doi.org/10.1007/s10928-022-09808-w

ORIGINAL PAPER

A pharmacokinetic and pharmacodynamic analysis of drug forgiveness

Noel P. McAllister[1] · Sean D. Lawley[1]

Received: 20 November 2021 / Accepted: 29 January 2022

Abstract
Nonadherence to medication is a major public health problem. To combat nonadherence, some clinicians have suggested using "forgiving" drugs, which maintain efficacy in spite of delayed or missed doses. What pharmacokinetic (PK) and pharmacodynamic (PD) factors make a drug forgiving? In this paper, we address this question by analyzing a linear PK/PD model for a patient with imperfect adherence. We assume that the drug effect is far from maximal and consider direct effect, effect compartment (biophase), and indirect response PD models. We prove that the average drug effect relative to the clinically desired effect is simply the fraction of prescribed doses actually taken by the patient. Hence, under these assumptions, drug forgiveness cannot be defined in terms of the average effect. We argue that forgiveness should instead be understood in terms of effect fluctuations. We prove that the rates of PK absorption, PK elimination, and PD elimination are exactly equivalent for determining effect fluctuations. We prove all the aforementioned results for any pattern of nonadherence, including late doses, missed doses, drug holidays, extra doses, etc. To obtain quantitative estimates of effect fluctuations, we consider a simple statistical pattern of nonadherence and analytically calculate the coefficient of variation of effect. We further show how effect fluctuations can be reduced by taking an extra "make up" dose following a missed dose if any one of the aforementioned PK/PD rates is sufficiently slow. We illustrate some of our results for a nonlinear indirect response model of metformin.

Keywords Medication adherence · Forgiveness · Missed doses · Stochastics

Introduction

Medication adherence is the process by which patients take their medications as prescribed [1]. Nonadherence to medication is a well-documented problem. It is estimated that medication nonadherence accounts for over 100,000 preventable deaths and over $100 billion in preventable healthcare costs every year in the United States [2]. In fact, the World Health Organization claimed that "increasing the effectiveness of adherence interventions may have a far greater impact on the health of the population than any improvement in specific medical treatments" [3, 4].

To combat the problem of nonadherence, some clinicians have suggested using so-called "forgiving" drugs [2]. Forgiveness is sometimes defined as the difference between the medication's post-dose duration of action and the prescribed time interval between doses [5]. However, it is difficult to precisely quantify the forgiveness of a specific drug in terms of a single number, as evidenced by the variety of mathematical definitions of drug forgiveness presented in the literature [6–13]. Nevertheless, the general, more qualitative notion of drug forgiveness is certainly an important characteristic of a drug. Indeed, it is well-established that some drugs require strict adherence to achieve therapeutic benefits (i.e. less forgiving drugs), whereas the benefits of some drugs are quite robust to lapses in adherence (i.e. more forgiving drugs) [14].

What makes a drug forgiving? In particular, what pharmacokinetic (PK) factors make a drug forgiving? A long drug half-life, which is related to a slow PK elimination rate, is generally considered to make a drug forgiving. Are there other PK factors which have an equally strong effect on forgiveness?

Furthermore, what pharmacodynamic (PD) factors make a drug forgiving? In an interesting computational study

✉ Sean D. Lawley
lawley@math.utah.edu

[1] Department of Mathematics, University of Utah, Salt Lake City, UT 84112, USA

Published online: 13 February 2022

[15], it was shown in one example that a drug with an effect compartment (biophase) PD model was more forgiving than a drug with a direct effect PD model. Does this example represent a more general principle about effect compartment PD models versus direct effect PD models? Is a drug with a long PK half-life and a direct effect PD model more or less forgiving than a drug with a short PK half-life and an effect compartment PD model? How do drugs with an indirect response PD model compare?

The purpose of this paper is to address these questions. We consider a patient who is instructed to take a drug repeatedly at a fixed dosing interval, but whose actual adherence deviates from this prescription. In the taxonomy of [1], this analysis thus concerns the implementation phase of adherence, which is the extent to which a patient's actual dosing follows the prescribed dosing regimen. We employ a PK/PD model, where the PK model consists of an absorption compartment and a plasma (main) compartment with first-order kinetics. For PD, we consider a direct effect model, an effect compartment (biophase) model, and an indirect response model [16]. Figure 1 illustrates these PK/PD models. We focus on the case that the drug effect is far from maximal, in which case the PD models are linear (see "Discussion" for more on this assumption). We use rigorous mathematical analysis to study these models, rather than computational simulations of specific numerical examples. This approach allows us to arrive at strong conclusions which apply to all such PK/PD models.

To describe our results, let $\langle E^{\text{perf}} \rangle$ denote the long-term average drug effect for a perfectly adherent patient. From a clinical perspective, $\langle E^{\text{perf}} \rangle$ is thus the desired drug effect. If $\langle E \rangle$ denotes the long-term average drug effect for an imperfectly adherent patient, then we prove that

$$\langle E \rangle = \mu \langle E^{\text{perf}} \rangle, \tag{1}$$

where μ denotes the long-term fraction of prescribed doses actually taken by the imperfectly adherent patient (μ is a multiplicative factor in (1)). This result implies that drug forgiveness cannot be defined in terms of the average drug effect. To see this, observe that (1) means that the ratio of the average drug effect to the desired drug effect, $\langle E \rangle / \langle E^{\text{perf}} \rangle$, is simply the fraction of doses taken, μ, which is independent of drug characteristics. For example, if the patient takes $\mu = 80\%$ of the prescribed doses, then the patient receives 80% of the clinically desired drug effect, regardless of the PK or PD drug parameters. We illustrate this point in Fig. 2, which plots time courses of the relative effect for a variety of adherence patterns and a variety of PK/PD parameters. Despite the different parameters and different adherence patterns, each curve has the same adherence rate μ (we take $\mu = 80\%$ in this figure), and thus the average relative effect $\langle E \rangle / \langle E^{\text{perf}} \rangle$ for each curve is μ.

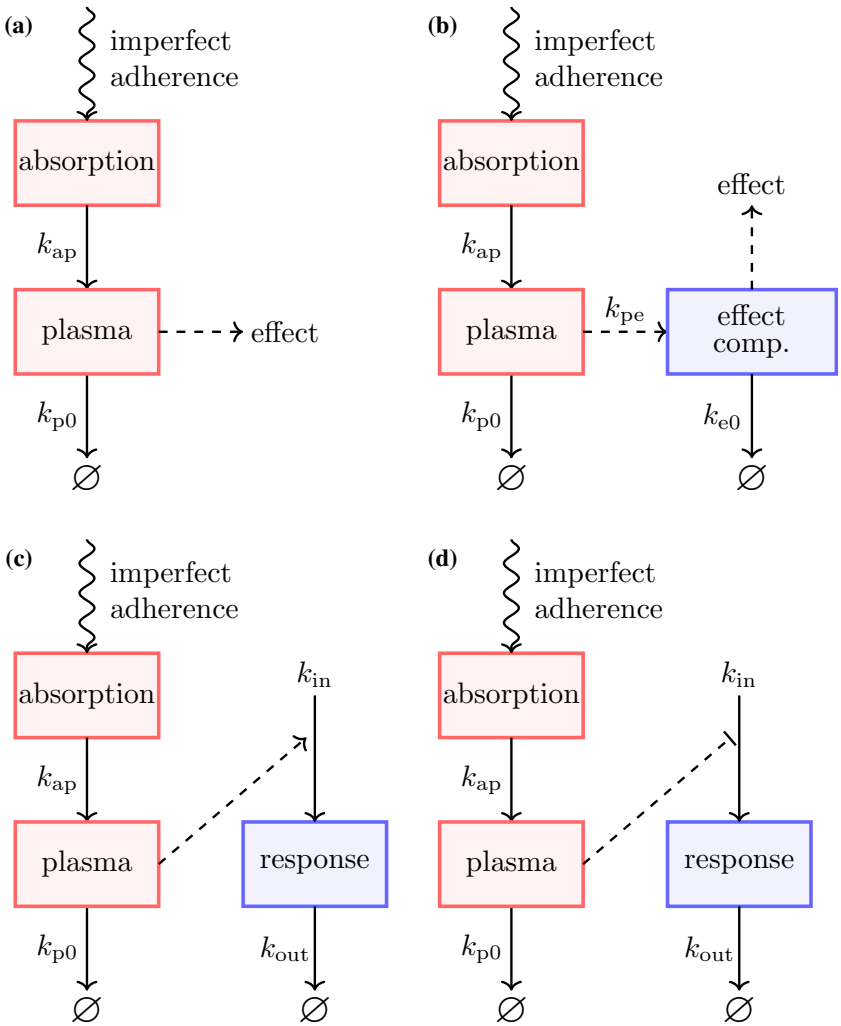
Rather than a property of the average drug effect, this analysis suggests that forgiveness is a property of drug effect fluctuations. Indeed, the usual objective during long-term pharmacotherapy is to maintain continuity of action of the prescribed drug [14]. A prominent feature of Fig. 2 is that the curves differ wildly in their fluctuations around their average. To quantify these fluctuations, we consider the coefficient of variation of the drug effect, denoted by $\text{CV}(E)$, which is the ratio of the standard deviation to the mean. We prove that $\text{CV}(E)$ is a symmetric function of the PK absorption rate, the PK elimination rate, and the PD elimination rate (meaning the value of $\text{CV}(E)$ is unchanged if we permute the values of these three rates). Hence, drug effect fluctuations, which are a proxy for drug forgiveness, depend equally on these three rates. Therefore, the PK elimination rate does not uniquely contribute to forgiveness, and these other rates should receive equal attention regarding forgiveness.

We prove all of the aforementioned results for any pattern of nonadherence, including any combination of late doses, missed doses, drug holidays, extra doses, etc. However, to obtain quantitative estimates of drug effect fluctuations, we must make assumptions on the patterns of patient nonadherence. Thus, for some of our analysis, we assume that the patient misses each dose with a fixed probability, independent of their prior behavior. We then obtain an explicit formula for $\text{CV}(E)$ as a function of this probability, the prescribed dosing interval, and the PK/PD rates described above. We note that this formula (and all the results above) are valid for the direct effect PD model, the effect compartment PD model, and the indirect response PD models. We illustrate the use of this formula by applying it to a PK/PD model of metformin [17]. We also obtain an explicit formula for $\text{CV}(E)$ in the case that the patient takes an extra "make up" dose when they take the drug following one or more missed doses. This analysis shows that taking make up doses reduces fluctuations in the drug effect if any of the PK/PD rates described above are sufficiently slow compared to the prescribed dosing interval.

The rest of the paper is organized as follows. We describe the PK/PD models in "Methods" and present the results of analyzing these models in "Results". We conclude in "Discussion" by discussing model limitations, relations to prior work, and future directions. We collect some technical aspects of the mathematical analysis in the "Appendix".

Methods

We now describe the PK/PD models used in this study.



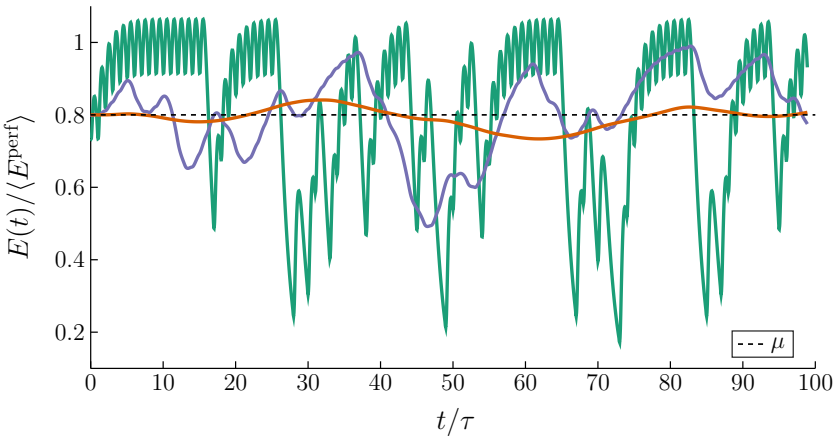


Fig. 1 Diagram of PK/PD models. The PK model is the same in **a**–**d**. The PD model is direct effect in **a**, effect compartment (biophase) in **b**, indirect response with stimulation of response in **c**, and indirect response with inhibition of response in **d**

Fig. 2 Drug effect $E(t)$ relative to the clinically desired drug effect $\langle E^{\text{perf}}\rangle$ as a function of time t relative to the dosing interval τ. The three solid curves correspond to different adherence patterns and different drugs with different PK/PD parameters. Nevertheless, the average drug effect relative to the clinically desired drug effect is simply the fraction of prescribed doses actually taken by the patient, which is $\mu = 80\%$ for each curve. However, the three curves differ markedly in their fluctuations

PK model

Consider the standard PK model consisting of an absorption compartment and a plasma (main) compartment as depicted in Fig. 1. Let $c_a(t)$ denote the concentration in the absorption compartment at time $t \geq 0$ and suppose c_a satisfies the following ordinary differential equation (ODE),

$$\frac{d}{dt}c_a = -k_{ap}c_a + \frac{I(t)}{V}, \tag{2}$$

where $I(t)$ describes the drug input and k_{ap} denotes the PK absorption rate. The PK absorption rate is commonly denoted "k_a" in the literature, but instead we adopt the convention that k_{ij} denotes the transfer rate from compartment i to j, where $i,j \in \{a, p, e, 0\}$ and "a" denotes absorption compartment, "p" denotes plasma compartment, "e" denotes effect compartment, and "0" denotes elimination. We let V denote the volume of distribution and without loss of generality let each compartment have volume V in order to simplify notation.

The concentration in the plasma compartment is denoted by $c_p(t)$ and satisfies

$$\frac{d}{dt}c_p = k_{ap}c_a - k_{p0}c_p, \tag{3}$$

where k_{p0} denotes the PK elimination rate. If the drug input is a single dose of size $D > 0$ taken at time zero, then the concentration in the plasma compartment is

$$c_p(t) = \frac{DF}{V}\frac{k_{ap}}{k_{ap} - k_{p0}}\left(e^{-k_{p0}t} - e^{-k_{ap}t}\right), \quad t \geq 0, \tag{4}$$

if $k_{ap} \neq k_{p0}$ and $F \in (0, 1]$ denotes the bioavailability.

PD: direct effect model

Perhaps the simplest PD model is the so-called direct effect model, in which the effect of the drug at time $t \geq 0$, $E^{de}(t)$, is the following function of the drug concentration in the plasma compartment [16],

$$E^{de}(t) := \frac{E_{max}c_p(t)}{EC_{50} + c_p(t)}. \tag{5}$$

We use the superscript "de" for "direct effect" to distinguish this model from other PD models described below. In (5), E_{max} is the maximum possible effect and EC_{50} is the drug concentration which produces one half of the maximum effect. The direct effect model is depicted in Fig. 1a.

PD: effect compartment model

Another common PD model is the "effect compartment" or "biophase" model [16] depicted in Fig. 1b. In this model, the drug can pass from the plasma compartment to an effect compartment, whose concentration c_e satisfies

$$\frac{d}{dt}c_e = k_{pe}c_p - k_{e0}c_e. \tag{6}$$

In (6), k_{pe} denotes the transfer rate from the plasma to effect compartments and k_{e0} denotes the PD elimination rate. In keeping with standard assumptions (for example, see [16]), the amount of drug moving in and out of the effect compartment is negligible compared to the amount in the plasma compartment and therefore does not influence the PK of the drug (i.e. c_e does not appear in (3)). The effect of the drug on the body is given by

$$E^{ec}(t) := \frac{E_{max}c_e(t)}{EC_{50} + c_e(t)}, \tag{7}$$

where the superscript "ec" distinguishes this effect compartment model from other PD models.

PD: indirect response model

Another common PD model is the "indirect response" model [16], in which the drug response $R(t)$ evolves according to an ODE of the form,

$$\frac{d}{dt}R = k_{in}\left(1 + \frac{S_{max}c_p}{SC_{50} + c_p}\right) - k_{out}R, \tag{8}$$

or

$$\frac{d}{dt}R = k_{in}\left(1 - \frac{I_{max}c_p}{IC_{50} + c_p}\right) - k_{out}R. \tag{9}$$

In (8)–(9), the response $R(t)$ models a biomarker whose production is either stimulated (in the case of (8)) or inhibited (in the case of (9)) by the concentration of the drug in the plasma compartment. The indirect response models in (8) and (9) are depicted respectively in Fig. 1c and d. We note that some indirect response models allow the drug to affect the dissipation of the biomarker [16], but we do not consider this type of model in this paper.

In order to unify our analysis of the indirect response model with the other PD models described above, we relabel the parameters in (8)–(9) as

$$\begin{aligned} &k_{out} = k_{e0}, \quad k_{in} = k_{e0}R_{base}, \quad IC_{50} = SC_{50} = EC_{50}, \\ &I_{max} = S_{max} = E_{max}k_{pe}/(R_{base}k_{e0}). \end{aligned} \tag{10}$$

In particular, the ODEs (8)–(9) can be written using the notation in (10) as

$$\frac{d}{dt}R = k_{e0}R_{base}\left(1 \pm \frac{E_{max}k_{pe}/(R_{base}k_{e0})c_p}{EC_{50} + c_p}\right) - k_{e0}R. \tag{11}$$

We emphasize that there is no loss of generality in using (11) rather than (8)–(9), as it amounts to merely relabeling parameters as in (10).

In the absence of the drug (i.e. $c_p = 0$), (9) implies that the baseline value of the biomarker R is R_{base}. It is therefore natural to define the effect of the drug to be the change from baseline,

$$E^{ir}(t) := \pm(R(t) - R_{base}), \tag{12}$$

where the superscript "ir" stands for indirect response. Plugging (12) into (9) implies that E^{ir} evolves according to

$$\frac{d}{dt}E^{ir} = \frac{k_{pe}E_{max}c_p}{EC_{50} + c_p} - k_{e0}E^{ir}. \tag{13}$$

Comparing the PD models

We now compare the PD models introduced above. We first note that the direct effect model is a limiting case of both the effect compartment model and the indirect response model. To see this, observe that the ODE for the effect compartment model in (6) is equivalent to

$$c_e(t) = \int_0^t e^{-k_{e0}(t-s)}k_{pe}c_p(s)\,ds, \tag{14}$$

assuming $c_e(0) = 0$ for simplicity. If we take $k_{pe} = k_{e0} \to \infty$ in (14) and use Laplace's method [18], then we obtain

$$\lim_{k_{pe}=k_{e0}\to\infty} c_e(t) = c_p(t).$$

Using the concentration–effect relations in (5) and (7), we thus have that

$$\lim_{k_{pe}=k_{e0}\to\infty} E^{ec}(t) = E^{de}(t).$$

Hence, the direct effect model is a special case of the more general effect compartment model.

Similarly, the ODE for the indirect response model in (13) is equivalent to

$$E^{ir}(t) = \int_0^t e^{-k_{e0}(t-s)}\frac{k_{pe}E_{max}c_p(s)}{EC_{50} + c_p(s)}\,ds, \tag{15}$$

assuming $E^{ir}(0) = 0$ for simplicity. Taking $k_{pe} = k_{e0} \to \infty$ in (15) and using Laplace's method yields

$$\lim_{k_{pe}=k_{e0}\to\infty} E^{ir}(t) = E^{de}(t).$$

Hence, the direct effect model is also a special case of the more general indirect response model.

We now show that the effect compartment model is equivalent to the indirect response model if the drug effects are much less than maximal. More precisely, if $c_e < EC_{50}$, then we can write the concentration–effect relation in (7) as a geometric series,

$$\begin{aligned} E^{ec}(t) &= E_{max}\sum_{j\geq 1}(-1)^{j+1}\left(\frac{c_e(t)}{EC_{50}}\right)^j \\ &= E_{max}\left(\frac{c_e(t)}{EC_{50}} - \left(\frac{c_e(t)}{EC_{50}}\right)^2 + \ldots\right). \end{aligned} \tag{16}$$

If

$$c_e \ll EC_{50}, \tag{17}$$

then $E^{ec}(t)$ is well-approximated by taking only the first term in (16) which yields the following linear concentration–effect relation,

$$E(t) := \frac{E_{max}}{EC_{50}}c_e(t) \approx E^{ec}(t). \tag{18}$$

By the same argument, if

$$c_p \ll EC_{50}, \tag{19}$$

then E^{ir} in (15) is well approximated by

$$E(t) = \frac{E_{max}}{EC_{50}}\int_0^t e^{-k_{e0}(t-s)}k_{pe}c_p(s)\,ds \approx E^{ir}(t). \tag{20}$$

However, notice that (14) implies that $E(t)$ in (20) is identical to $E(t)$ in (18).

To summarize, the direct effect model is a special case of both the effect compartment model and the indirect response model. Furthermore, if the drug effect is far from maximal (in which case the PD models are linear), then the indirect response model and the effect compartment model are equivalent. The upshot of this is that if the drug effect is far from maximal (meaning (17) or (19)), then we do not need to analyze the direct effect, effect compartment, and indirect response PD models separately. That is, our analysis of the effect $E(t)$ in (18) applies equally well to all of these PD models.

Drug effect for general adherence

If the concentration in the plasma compartment is given by (4) (corresponding to a single dose of the drug given at time zero), then we show in the "Appendix" that

$$E(t) = \frac{E_{max}}{EC_{50}}\frac{DF}{V}\left(b_{p0}e^{-k_{p0}t} + b_{ap}e^{-k_{ap}t} + b_{e0}e^{-k_{e0}t}\right), \tag{21}$$

where

$$b_{p0} = \frac{k_{pe}k_{ap}}{(k_{p0}-k_{ap})(k_{p0}-k_{e0})},$$
$$b_{ap} = \frac{k_{pe}k_{ap}}{(k_{ap}-k_{e0})(k_{ap}-k_{p0})}, \quad (22)$$
$$b_{e0} = \frac{k_{pe}k_{ap}}{(k_{e0}-k_{ap})(k_{e0}-k_{p0})}.$$

We assume k_{p0}, k_{ap}, and k_{e0} are distinct throughout this paper.

Suppose the patient takes a dose of an amount $Df_n \geq 0$ at time $t_n \geq 0$ for $n \geq 0$, where $\{f_n\}_{n\geq 0}$ is any nonnegative sequence of dose sizes and

$$0 = t_0 < t_1 < t_2 < \cdots \quad (23)$$

is any increasing sequence of times. In this case, the drug input $I(t)$ in (2) is

$$I(t) = DF \sum_{n\geq 0} \delta_{dirac}(t-t_n) f_n, \quad (24)$$

where δ_{dirac} denotes the Dirac delta function. Applying the superposition principle to (21) yields that effect at time $t \geq 0$ is

$$E(t) := \frac{E_{max}}{EC_{50}} \frac{DF}{V} \sum_{n:t_n \leq t} \Big(b_{p0} e^{-k_{p0}(t-t_n)} + b_{ap} e^{-k_{ap}(t-t_n)} + b_{e0} e^{-k_{e0}(t-t_n)} \Big) f_n, \quad (25)$$

where the sum is over all indices n such that $t_n \leq t$. We emphasize that (25) holds for any sequence of nonnegative dose sizes $\{f_n\}_{n\geq 0}$ taken at any increasing sequence of times $\{t_n\}_{n\geq 0}$.

If the patient is instructed to take a dose of size $D > 0$ every $\tau > 0$ units of time, then in the special case of perfect adherence we have for $n \geq 0$,

$$f_n = 1 \quad \text{and} \quad t_n = n\tau. \quad \text{(perfect adherence)} \quad (26)$$

In this case of perfect adherence, (25) can be written in the following form,

$$E^{perf}(N\tau + t) := \frac{E_{max}}{EC_{50}} \frac{DF}{V} \Big(b_{p0} e^{-k_{p0}t} \sum_{n=0}^{N} (e^{-k_{p0}\tau})^n + b_{ap} e^{-k_{ap}t} \sum_{n=0}^{N} (e^{-k_{ap}\tau})^n + b_{e0} e^{-k_{e0}t} \sum_{n=0}^{N} (e^{-k_{e0}\tau})^n \Big), \quad (27)$$

where $t \in [0, \tau)$ denotes the time elapsed since the $(N+1)$-st dose. If the patient continues their perfect adherence for a long time, then it follows from (27) that the effect at time $t \in [0, \tau)$ since the most recent dose is

$$\lim_{N\to\infty} E^{perf}(N\tau + t) = \frac{E_{max}}{EC_{50}} \frac{DF}{V} \Big(\frac{b_{p0} e^{-k_{p0}t}}{1 - e^{-k_{p0}\tau}} + \frac{b_{ap} e^{-k_{ap}t}}{1 - e^{-k_{ap}\tau}} + \frac{b_{e0} e^{-k_{e0}t}}{1 - e^{-k_{e0}\tau}} \Big).$$

To study the effects of imperfect adherence, we allow the dose sizes $\{f_n\}_{n\geq 0}$ and times $\{t_n\}_{n\geq 0}$ to deviate from (26).

Results

We now use the mathematical models introduced above to investigate how drug effects depend on patient adherence and PK/PD parameters. We study long-term average drug effects, and toward this end we denote the long-term average of any time course $\{x(t)\}_{t\geq 0}$ by

$$\langle x \rangle := \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\, dt.$$

We also study how drug effects fluctuate around their averages, and we denote the coefficient of variation of any time course $\{x(t)\}_{t\geq 0}$ by

$$CV(x) := \frac{1}{\langle x \rangle} \sqrt{\lim_{T\to\infty} \frac{1}{T} \int_0^T \big(x(t) - \langle x \rangle\big)^2\, dt}. \quad (28)$$

As in "Methods", we suppose that the patient is instructed to take a dose of size $D > 0$ every $\tau > 0$ units of time. We also assume that the drug effect is far from maximal so that the PD models are linear, and hence the effect $E(t)$ in (18) applies to all the PD models described in "Methods" (see "Comparing the PD models" for details).

Average relative effect is the average drug intake

Drugs are said to be "forgiving" if delayed or missed doses only mildly change the effect on the body compared to perfect adherence. For instance, drugs with long half-lives (i.e. slow PK elimination rates k_{p0}) are often said to be forgiving. The point of this subsection is to show that if one considers only how the long-term average drug effect compares to the clinically desired effect, then the drug characteristics are irrelevant and the only important quantity is the long-term fraction of prescribed doses actually taken by the patient.

If the patient has perfect adherence, then a straightforward calculation shows that the long-term average effect is

$$\langle E^{\text{perf}} \rangle = \frac{E_{\max}}{\text{EC}_{50}} \frac{DF}{V} \frac{1}{\tau} \frac{k_{\text{pe}}}{k_{\text{p0}} k_{\text{e0}}}. \tag{29}$$

Hence, for a given drug and patient with PK/PD parameters k_{pe}, k_{p0}, k_{e0}, V, $E_{\max}$, and EC_{50}, the dose size D and dosing interval τ are prescribed so that $\langle E^{\text{perf}} \rangle$ in (29) is the clinically desired average effect.

Now consider an imperfectly adherent patient. Imperfect adherence can take a variety of forms, such as delayed doses, missed doses, drug holidays, extra doses, etc. [2]. Regardless of the particular form of the imperfect adherence, let μ denote the fraction of prescribed doses actually taken by the patient. More precisely, assume that

$$\lim_{T \to \infty} \frac{\tau}{T} \sum_{n: t_n \leq T} f_n = \mu, \tag{30}$$

where the sum is over all times $t_n \leq T$. Hence, $\mu > 0$ is the average number of doses taken in each prescribed dosing interval.

Under the assumption in (30), we prove in the "Appendix" that the average effect $\langle E \rangle$ compared to the desired average effect $\langle E^{\text{perf}} \rangle$ is simply the fraction of prescribed doses actually taken,

$$\langle E \rangle = \mu \langle E^{\text{perf}} \rangle. \tag{31}$$

We emphasize that (31) means that the ratio of the actual average effect to the desired average effect is independent of the PK/PD parameters. We further emphasize that (31) follows from merely assuming the adherence rate assumption in (30). In particular, the imperfect adherence could be any combination of delayed doses, missed doses, drug holidays, extra doses, etc. We illustrate (31) in Fig. 2, as described in the Introduction. Of course, a striking feature of Fig. 2 is that the curves differ wildly in their fluctuations around the average (which we investigate below).

To summarize, if one only considers how the average drug effect compares to the clinically desired average effect, then the PK/PD parameters are irrelevant and the average drug intake is the only important quantity. Therefore, the notion of a "forgiving" drug requires considering the fluctuations in drug effects, rather than mere averages. We investigate how effect fluctuations depend on PK/PD parameters in the subsections below.

PK absorption, PK elimination, and PD elimination rates are equivalent for effect fluctuations

The PK/PD models in "Methods" involve the four rate parameters, k_{ap}, k_{p0}, k_{pe}, and k_{e0}. How do these parameters influence how $E(t)$ fluctuates around its average?

We measure fluctuations in drug effect via the coefficient of variation, $\text{CV}(E)$, defined in (28). Assuming merely that the patient has adherence μ as in (30), the relation in (31) implies that the coefficient of variation of the effect E can be written as

$$\begin{aligned} \text{CV}(E) &= \frac{1}{\langle E \rangle} \sqrt{\lim_{T \to \infty} \frac{1}{T} \int_0^T \big(E(t) - \langle E \rangle\big)^2 \, \mathrm{d}t} \\ &= \sqrt{\lim_{T \to \infty} \frac{1}{T} \int_0^T \left(\frac{E(t)}{\mu \langle E^{\text{perf}} \rangle} - 1\right)^2 \mathrm{d}t}. \end{aligned} \tag{32}$$

Using the general formula for $E(t)$ in (25) and the formula for $\langle E^{\text{perf}} \rangle$ in (29), we have that

$$\begin{aligned} \frac{E(t)}{\langle E^{\text{perf}} \rangle} = \sum_{n: t_n \leq t} \Big(& a_{\text{p0}} e^{-k_{\text{p0}}(t - t_n)} + a_{\text{ap}} e^{-k_{\text{ap}}(t - t_n)} \\ & + a_{\text{e0}} e^{-k_{\text{e0}}(t - t_n)} \Big) f_n, \end{aligned} \tag{33}$$

where

$$\begin{aligned} a_{\text{p0}} &= \frac{\tau k_{\text{ap}} k_{\text{p0}} k_{\text{e0}}}{(k_{\text{p0}} - k_{\text{ap}})(k_{\text{p0}} - k_{\text{e0}})}, \\ a_{\text{ap}} &= \frac{\tau k_{\text{ap}} k_{\text{p0}} k_{\text{e0}}}{(k_{\text{ap}} - k_{\text{e0}})(k_{\text{ap}} - k_{\text{p0}})}, \\ a_{\text{e0}} &= \frac{\tau k_{\text{ap}} k_{\text{p0}} k_{\text{e0}}}{(k_{\text{e0}} - k_{\text{ap}})(k_{\text{e0}} - k_{\text{p0}})}. \end{aligned} \tag{34}$$

By inspecting (33)–(34), we see that the ratio $E(t)/\langle E^{\text{perf}} \rangle$ is (i) independent of k_{pe} and (ii) a symmetric function of the rates k_{ap}, k_{p0}, and k_{e0}. Point (ii) means that the value of $E(t)/\langle E^{\text{perf}} \rangle$ is unchanged if we swap the values of k_{ap}, k_{p0}, and k_{e0}. More precisely, if $x_1 > 0$, $x_2 > 0$, and $x_3 > 0$ are any three distinct values, then

$$\left. \frac{E(t)}{\langle E^{\text{perf}} \rangle} \right|_{(k_{\text{ap}}, k_{\text{p0}}, k_{\text{e0}}) = (x_1, x_2, x_3)} = \left. \frac{E(t)}{\langle E^{\text{perf}} \rangle} \right|_{(k_{\text{ap}}, k_{\text{p0}}, k_{\text{e0}}) = (x_i, x_j, x_l)} \tag{35}$$

for any distinct indices $i, j, l \in \{1, 2, 3\}$. Therefore, (32) implies that the coefficient of variation $\text{CV}(E)$ is (i) independent of k_{pe} and (ii) a symmetric function of the rates k_{ap}, k_{p0}, and k_{e0}. In particular,

$$\left. \text{CV}(E) \right|_{(k_{\text{ap}}, k_{\text{p0}}, k_{\text{e0}}) = (x_1, x_2, x_3)} = \left. \text{CV}(E) \right|_{(k_{\text{ap}}, k_{\text{p0}}, k_{\text{e0}}) = (x_i, x_j, x_l)} \tag{36}$$

for any distinct indices $i, j, l \in \{1, 2, 3\}$.

Therefore, the PK absorption rate k_{ap}, the PK elimination rate k_{p0}, and the PD elimination rate k_{e0} are equally important for determining how the effect $E(t)$ fluctuates around its average. In particular, while the PK elimination rate k_{p0} is usually considered to be a determinative factor of

a drug's forgiveness (since k_{p0} controls the PK elimination half-life), this analysis shows that k_{ap} and k_{e0} are just as important for controlling effect fluctuations. We stress that this result holds for any pattern of nonadherence.

We illustrate (35)–(36) in Fig. 3 by plotting sample time courses of the relative plasma concentrations (left panel) and relative effects (right panel) for three different choices of k_{p0} and k_{e0}. We note that the values of k_{ap} and τ and the adherence is identical for these three curves (we take $k_{ap}\tau = 1$ and the patient takes or misses doses at the exact same times for the three curves). Looking at the plasma time courses in the left panel (i.e. only considering PK), the orange dotted curve seems to represent the most forgiving drug and the green solid curve seems to represent the least forgiving drug. Indeed, this matches with the PK elimination rates in that the orange dotted curve has a much slower value of k_{p0} compared to the value of k_{p0} for the green solid curve.

However, looking at the time courses of the relative effect in the right panel of Fig. 3, we see that the orange dotted curve is exactly the same as the green solid curve. This is due to (35) since the difference in the PK/PD parameters for the orange dotted curve and the green solid curve is merely that the values of k_{p0} and k_{e0} have been swapped. Furthermore, if one looks only at the plasma time courses in the left panel, the drug represented by the purple dashed curve seems to be much more forgiving than the drug represented by the green solid curve. However, the right panel of Fig. 3 shows that the purple dashed curve actually represents the least forgiving drug in terms of the effect fluctuations, and this is due to its very fast PD elimination rate k_{e0}.

Coefficient of variation estimate

To obtain quantitative estimates of the fluctuations in the effect, we must specify more details about the patient's adherence. In this section, we assume simply that the patient takes each scheduled dose with probability $p \in (0, 1]$ and misses each scheduled dose with probability $1 - p \in [0, 1)$, independent of their prior behavior. Mathematically, this means that the dosing times in (25) are $t_n = n\tau$ for $n \geq 0$, and the dose sizes are

$$f_n = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1 - p, \end{cases} \tag{37}$$

where f_n and f_m are independent if $n \neq m$. See "Discussion" for more on this independence assumption.

It is immediate that (30) is satisfied with

$$\mu = p, \tag{38}$$

and therefore (31) yields that the long-term average effect is

$$\langle E \rangle = p \langle E^{\text{perf}} \rangle.$$

In the "Appendix", we find the following exact formula for the coefficient of variation of the effect,

$$\text{CV}(E) = \sqrt{\frac{g_1}{p} + g_2 - 1}, \tag{39}$$

where

$$\begin{aligned} g_1 &:= h_1(k_{ap}\tau, k_{p0}\tau, k_{e0}\tau), \\ g_2 &:= h_2(k_{ap}\tau, k_{p0}\tau, k_{e0}\tau) + h_2(k_{p0}\tau, k_{ap}\tau, k_{e0}\tau) \\ &\quad + h_2(k_{e0}\tau, k_{p0}\tau, k_{ap}\tau), \end{aligned} \tag{40}$$

and h_1 and h_2 are the functions,

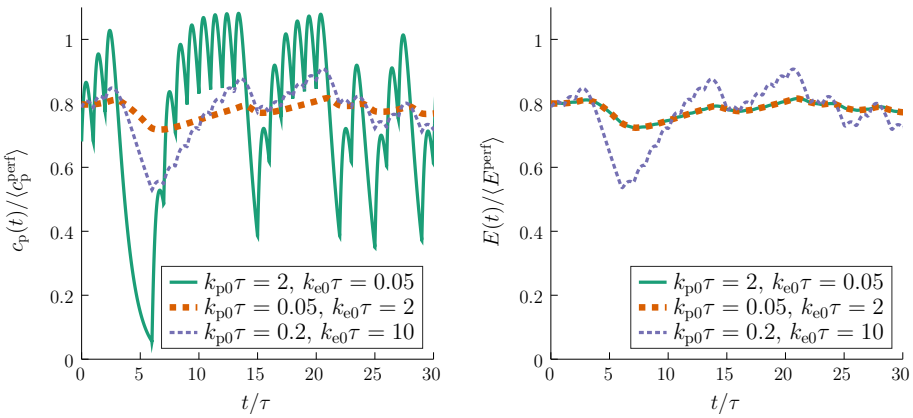


Fig. 3 Fluctuations in relative plasma concentration (left panel) and relative effect (right panel) for imperfect adherence. The plasma concentration time courses in the left panel are normalized by $\langle c_p^{\text{perf}} \rangle$, which denotes the average plasma concentration for perfect adherence. Though the solid green curve shows large fluctuations in plasma concentration in the left panel, the corresponding effect fluctuations in the right panel are quite small. See the text for details (Color figure online)

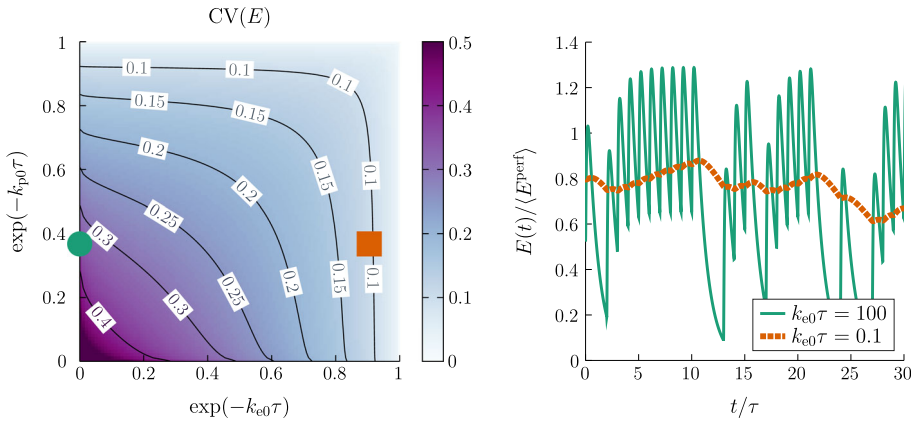


Fig. 4 Left: Contour plot of $\mathrm{CV}(E)$ in (39). Right: Sample time course of relative effect for the parameter values indicated by the green circle and orange square markers in the left panel. In both panels, we take $k_{\mathrm{ap}}\tau = 10$ and $p = 0.8$ (Color figure online)

$$h_1(x,y,z) := \frac{xyz(x+y+z)}{2(x+y)(x+z)(y+z)},$$
$$h_2(x,y,z) := \frac{xy^2z^2}{(x+y)(x+z)(x-y)(x-z)(e^x-1)}. \tag{41}$$

Even though we have assumed a very simple model of nonadherence in this section, the formulas in (39)–(41) are quite complicated. Nevertheless, it is straightforward to use these formulas to plot $\mathrm{CV}(E)$ as a function of the various parameters.

In the left panel of Fig. 4, we show a contour plot of $\mathrm{CV}(E)$ using the formula in (39). In this plot, we set $p = 0.8$ and fix τ and k_{ap} so that $k_{\mathrm{ap}}\tau = 10$, and we let k_{p0} and k_{e0} vary. Note that, due to the symmetry in (36), this is equivalent to fixing either k_{p0} or k_{e0} and letting the other rates vary. This plot shows that $\mathrm{CV}(E)$ increases if any of the rates k_{p0}, k_{e0}, or k_{ap} increases or if τ increases. Indeed, using the exact formula in (39), we have verified through extensive numerical tests that for any $p \in (0,1]$, $\tau > 0$, and any distinct values of the rates k_{p0}, k_{e0}, or k_{ap},

$$\frac{\partial}{\partial k}\mathrm{CV}(E) > 0, \quad \frac{\partial}{\partial \tau}\mathrm{CV}(E) > 0,$$

where k is any of the three rates k_{p0}, k_{e0}, or k_{ap}.

To illustrate what the $\mathrm{CV}(E)$ implies about actual time courses of the effect, in the right panel of Fig. 4 we plot sample time courses of the relative effect $E(t)/\langle E^{\mathrm{perf}}\rangle$ for two different values of the PD elimination rate k_{e0} (corresponding to the green circle and orange square markers in the left panel of Fig. 4). This plot shows that the effect of the drug with a fast PD elimination rate (green solid curve) fluctuates greatly, whereas the drug with the slow PD elimination rate maintains a fairly stable effect (orange dashed curve). This agrees with the predictions of the formula for $\mathrm{CV}(E)$ in (39), since $\mathrm{CV}(E) \approx 0.39$ for the green curve and $\mathrm{CV}(E) \approx 0.11$ for the orange curve. We emphasize that the only difference between these two time courses are the different values of k_{e0}; all of the other PK/PD parameters and the adherence patterns are identical. We note that the value of k_{e0} is sufficiently large that the PD is essentially the direct effect model (see "Comparing the PD models"). Hence, Fig. 4 illustrates how a drug which is well described by a direct effect model may yield very large fluctuations in the effect.

Application to metformin

We now illustrate our results for a patient with type 2 diabetes mellitus taking metformin with imperfect adherence. We take the PK/PD model developed by Hong et al. [17] as our starting point. Hong et al. [17] describes the PK of metformin with a model identical to our PK model. Further, the PD model of [17] is identical to our indirect response PD model. Specifically, Hong et al. [17] describes the PD of metformin via the inhibitory indirect response model in (9), where R is the plasma glucose, k_{in} is the zero-order rate constant for glucose production, k_{out} is the first-order rate constant for glucose utilization, c_{p} is the metformin concentration in the plasma compartment, IC_{50} is the concentration of metformin yielding half-maximal antihyperglycemic effect, and $I_{\max} = 1$.

Table 1 Parameter values for metformin PK/PD model developed in [17]

Parameter	Numerical value
k_{ap}	$2.15\,\mathrm{h}^{-1}$
k_{p0}	$0.1219\,\mathrm{h}^{-1}$
k_{in}	$195.2\,\mathrm{mg/(dL \cdot h)}$
k_{out}	$0.8\,\mathrm{h}^{-1}$
τ	12 h
$I_{\max}$	1
IC_{50}	0.423 mg/dL
D	500 mg
V/F	6480 dL

If we relabel the PD parameters as in (10) and define the effect of metformin to be the change in glucose from the baseline glucose $R_{\text{base}} = k_{\text{in}}/k_{\text{out}}$ as in (12),

$$E^{\text{ir}}(t) := R_{\text{base}} - R(t),$$

then $E^{\text{ir}}(t)$ satisfies the nonlinear ODE in (13).

Using the numerical values of Hong et al. [17] (see Table 1), the long-term average plasma compartment concentration for a $D = 500\,\text{mg}$ dose administered twice per day ($\tau = 12\,\text{h}$) is

$$\langle c_{\text{p}}^{\text{perf}} \rangle = \frac{DF}{V}\frac{1}{k_{\text{p0}}\tau} \approx 0.05\,\text{mg/dL}. \tag{42}$$

Furthermore, Hong et al. [17] estimated that $\text{IC}_{50} = \text{EC}_{50} = 0.423\,\text{mg/dL}$. Thus, the average plasma concentration of metformin for perfect adherence is almost an order of magnitude less than the concentration which yields one half of the maximal effect,

$$\frac{\langle c_{\text{p}}^{\text{perf}} \rangle}{\text{EC}_{50}} \approx \frac{0.05}{0.423} \approx 0.12 \ll 1. \tag{43}$$

Hence, this model agrees with the assumption of "Methods" that the effect for perfect adherence is much less than the maximal effect.

In particular, (43) suggests using the linear concentration–effect approximation in (20),

$$E(t) \approx E^{\text{ir}}(t), \tag{44}$$

to estimate the drug effect. In the left panel of Fig. 5, we plot $\text{CV}(E)$ using the formula in (39) as a function the adherence p. The green solid curve uses the population average parameter values of [17] (see Table 1), including the PK absorption rate of $k_{\text{ap}} = 2.15\,\text{h}^{-1}$ corresponding to the immediate release metformin used in [17] (between subject variability in a population of patients is addressed in "Discussion"). The orange and purple curves use the same parameter values except for slower PK absorption rates of $k_{\text{ap}} = 0.215\,\text{h}^{-1}$ (orange curve) and $k_{\text{ap}} = 0.0215\,\text{h}^{-1}$ (purple curve) corresponding to hypothetical extended release metformin formulations. This plot shows how $\text{CV}(E)$ decreases for slower absorption rates.

The square markers in the left panel of Fig. 5 are values of the coefficient of variation of the effect E^{ir} for the full nonlinear indirect response model described above. These values of $\text{CV}(E^{\text{ir}})$ are obtained from numerical simulations of E^{ir} over many dosing intervals. In support of the approximation in (44), these values of $\text{CV}(E^{\text{ir}})$ are each within 10% of the corresponding value of $\text{CV}(E)$ computed from the formula in (39).

Therefore, this analysis predicts that extended release formulations dosed at the same frequency as immediate release formulations have the potential of maintaining a more stable drug effect in spite of imperfect adherence. This is illustrated in the right panel of Fig. 5, where we plot sample time courses of the drug effect E^{ir} for the parameters used in the left panel. We stress that these time courses of E^{ir} are the effect for the full nonlinear indirect response model described above. This plot shows how the drug effect (in this case, a decrease in glucose) persists following a missed dose if the drug absorption rate is slow.

'Make up' doses reduce variation for slow PK or PD

Is it ever appropriate for a patient to take an extra "make up" dose to compensate for a missed dose? To address this question, we modify the adherence model presented in "Coefficient of variation estimate". In "Coefficient of variation estimate", we assumed that the patient either takes or misses each scheduled dose with respective probabilities p and $1-p$. Importantly, the patient never takes more than a single dose at a time in the model of "Coefficient of variation estimate".

In this section, we suppose that the patient takes a double dose whenever they take their medication if they happened to have missed their prior dose. More precisely, we assume that the dosing times in (25) are $t_n = n\tau$ for all $n \geq 0$. To describe the dose sizes, let $\{\xi_n\}_{n\geq 0}$ be a sequence of independent Bernoulli random variables with

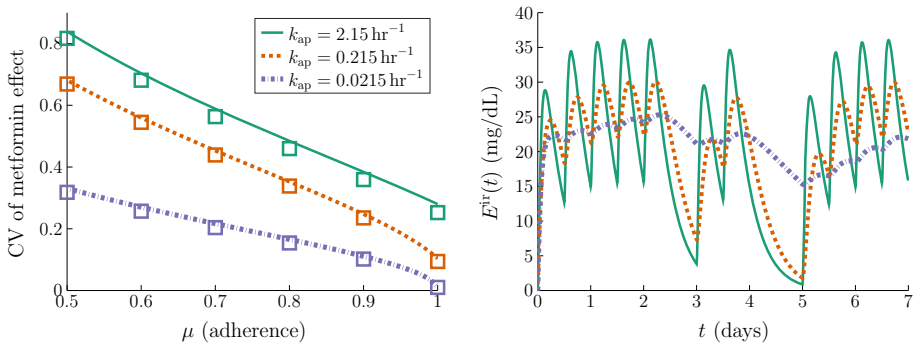


Fig. 5 Left: Coefficient of variation of effect of metformin for linear E model (curves) and nonlinear E^{ir} model (square markers). Right: Sample time courses of decrease in blood glucose for nonlinear E^{ir} metformin model. See the text for details

$$\xi_n = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1-p. \end{cases} \tag{45}$$

The dose sizes in (25) are then $f_0 = \xi_0$ and for $n \geq 1$,

$$f_n = \begin{cases} 0 & \text{if } \xi_n = 0, \\ 1 & \text{if } \xi_n = \xi_{n-1} = 1, \\ 2 & \text{if } \xi_n = 1,\ \xi_{n-1} = 0. \end{cases} \tag{46}$$

In words, (45) means that the patient "remembers" or "forgets" to take their medication at the nth dosing time with respective probabilities p and $1-p$. Further, (46) means that the patient does not take a dose when they forget, they take a single dose if they remember and they took their last scheduled dose, and they take a double dose if they remember and they happened to have missed their last scheduled dose. This adherence model first appeared in [19] as an input to a simpler PK model. We refer to this model as the "double dose protocol" and the model in "Coefficient of variation estimate" as the "single dose protocol."

In the "Appendix", we prove that the double dose protocol has the long-term average drug intake in (30) with

$$\mu = \mu^{\text{double}} := p + p(1-p).$$

As expected, the double dose protocol yields a higher average drug intake than the intake rate of $\mu^{\text{single}} := p$ in (38) for the single dose protocol. Hence, from the perspective of increasing the average drug effect, the double dose protocol is always superior to the single dose protocol.

How do the drug effect fluctuations for the double dose protocol compare to the single dose protocol? Letting $E^{\text{single}}(t)$ and $E^{\text{double}}(t)$ denote the respective drug effects for the single and double dose protocols, $\text{CV}(E^{\text{single}})$ is given in (39). We prove in the "Appendix" that the coefficient of variation for the double dose protocol is

$$\text{CV}(E^{\text{double}}) = \sqrt{\frac{g_3}{p(2-p)^2} - 1}, \tag{47}$$

where

$$g_3 := h_3(k_{\text{ap}}\tau, k_{\text{p0}}\tau, k_{\text{e0}}\tau) + h_3(k_{\text{p0}}\tau, k_{\text{ap}}\tau, k_{\text{e0}}\tau) + h_3(k_{\text{e0}}\tau, k_{\text{p0}}\tau, k_{\text{ap}}\tau),$$

and h_3 is the function

$$\begin{aligned} &h_3(x,y,z) \\ &:= (xyz)^2 \\ &\times \left[\frac{e^{-x}((-2p^2+7p-4)e^x + 2p(p^2-3p+2) + (4-3p)e^{2x})}{2(e^x-1)x(y-x)^2(z-x)^2} \right. \\ &+ \frac{2(p^2+3pe^y-5p-4e^y+4)}{(e^y-1)(y-x)(z-x)(y-z)^2(y+z)} \\ &- \frac{2(p^3-4p^2+4p)}{(e^z-1)(y-x)(z-x)(y-z)^2(y+z)} \\ &+ \frac{2(p-2)(p-1)pe^{-z}}{(y-x)(z-x)(y-z)^2(y+z)} \\ &\left. - \frac{2(p-2)(p-1)pe^{-y}}{(e^y-1)(y-x)(z-x)(y-z)^2(y+z)} \right]. \end{aligned}$$

In Fig. 6, we use the formulas in (39) and (47) to plot the following normalized difference between the effect coefficients of variation for the single and double dose protocols,

$$\delta := \frac{\text{CV}(E^{\text{single}}) - \text{CV}(E^{\text{double}})}{\text{CV}(E^{\text{single}})} \in \mathbb{R}.$$

Note that $\delta < 0$ means that the single dose protocol yields smaller fluctuations in drug effect and $\delta > 0$ means that the double dose protocol yields smaller fluctuations in drug effect. Figure 6 shows that the double dose protocol yields smaller fluctuations in drug effect compared to the single dose protocol except in the case that all three rates k_{ap}, k_{p0}, and k_{e0} are much faster than $1/\tau$.

To illustrate what these results imply about actual time courses, in Fig. 7 we plot a path of the relative drug effect for the single dose protocol (dashed orange curve) and the double dose protocol (solid purple curve). Evidently, the drug effect is much less perturbed by missed doses for the double dose protocol compared to the single dose protocol. Indeed, the time courses in Fig. 7 illustrate that the double dose protocol has a larger average drug effect ($\mu^{\text{double}} = p + p(1-p) = 0.96 > 0.8 = p = \mu^{\text{single}}$). Further, these time courses show that the double dose protocol has smaller effect fluctuations. In fact, for the values of $p = 0.8$, $e^{-k_{\text{ap}}\tau} = 10^{-6}$, $e^{-k_{\text{p0}}\tau} = 0.5$, and $e^{-k_{\text{e0}}\tau} = 0.8$ used in this plot, (39) and (47) yield

$$\text{CV}(E^{\text{double}}) \approx 0.08 < \text{CV}(E^{\text{single}}) \approx 0.15, \tag{48}$$

and thus $\delta \approx 0.45$. We emphasize that the times of missed doses and all the PK/PD parameters are identical for the two curves in Fig. 7.

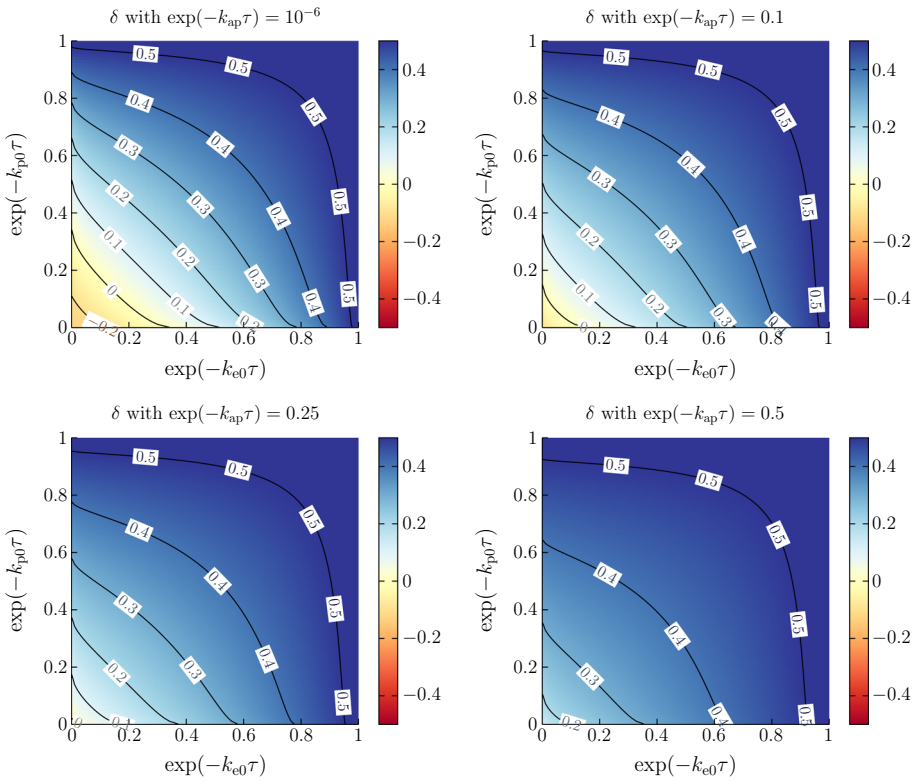


Fig. 6 Comparison of effect fluctuations for the single dose and double dose protocols. In each panel, we plot the normalized difference $\delta := (\mathrm{CV}(E^{\mathrm{single}}) - \mathrm{CV}(E^{\mathrm{double}}))/\mathrm{CV}(E^{\mathrm{single}})$ of effect coefficient of variations. We take $p = 0.8$ in each panel. The regions in which $\delta > 0$ (blue) are the regions where the double dose protocol yields smaller fluctuations than the single dose protocol (Color figure online)

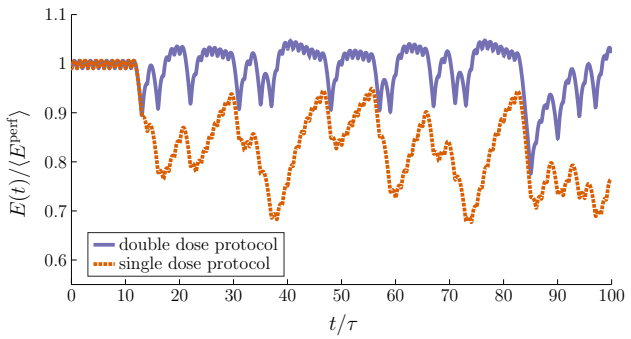


Fig. 7 Relative effect time courses for the single and double dose protocols. The times of missed doses and all the PK/PD parameters are identical for the two curves. See the text for details

Discussion

In this paper, we used mathematical analysis to study how PK and PD contribute to drug forgiveness. Assuming that the drug effect is far from maximal (in which case the PD models are linear), we found that drug forgiveness cannot be defined in terms of the average drug effect, since the average drug effect relative to the clinically desired effect is simply the fraction of prescribed doses actually taken by the patient, regardless of the drug characteristics and adherence patterns. We therefore argued that forgiveness should instead be understood in terms of fluctuations in effect. We found that the PK absorption rate, PK elimination rate, and PD elimination rate are equally important for determining fluctuations and thus forgiveness.

We further considered a simple model of patient nonadherence which allowed us to calculate the drug effect fluctuations as a function of the PK and PD rates. Using this model of patient nonadherence, we also investigated different ways of handling missed doses, referred to as the single dose protocol (i.e. skip any missed doses) and the double dose protocol (i.e. take an extra "make up" dose to compensate for a missed dose). As one would expect, we found that the double dose protocol increases the average drug effect compared to the single dose protocol. In addition, we found that the double dose protocol decreases effect fluctuations compared to the single dose protocol if any one of the aforementioned PK/PD rates is slow compared to the prescribed dosing rate.

Taking a double dose might be avoided out of concern that it could cause the drug concentration to rise dangerously high. This issue was recently addressed in [20]. Assuming a slow PK absorption or PK elimination rate, Theorem 5 in [20] implies that the double dose protocol can at most cause the plasma concentration to rise only slightly above the concentration in a perfectly adherent patient. In the "Appendix", we briefly review this result using the notation of the present paper.

In our model and analysis, we assumed fixed values of the model parameters (PK/PD rates, volume of distribution, bioavailability, etc.). It is well-known that there is often significant population variability in these parameters. This means that, for example, any two individual patients in a population likely have different values of k_{ap}, k_{p0}, and k_{e0}. Our analysis can thus be interpreted as concerning any single patient in a population. Our result that the average drug effect relative to the clinically desired effect is the fraction of prescribed doses taken by the patient thus holds for any patient in a population. The reason for this is that the desired effect is defined as the effect that this particular patient would have if they had perfect adherence, which depends on parameters specific to that patient. Put another way, this result is expressed mathematically in (1), and the patient-specific PK/PD parameters in the lefthand side of (1) cancel out. In addition, our result that k_{ap}, k_{p0}, and k_{e0} are equally important for determining forgiveness also holds for any patient in the population by the same reasoning. Finally, it is straightforward to use the formulas (39) and (47) for $\mathrm{CV}(E)$ to investigate how population variability affects the coefficient of variation of effect for a particular drug. In particular, one obtains a distribution of $\mathrm{CV}(E)$ by merely evaluating these formulas for many realizations of k_{ap}, k_{p0}, and k_{e0} sampled from given probability distributions which describe the population variability in these PK/PD rates.

The present work follows a line of previous works which have used PK, PD, and coupled PK/PD models to study medication nonadherence. Much of this previous work employed numerical computations [6, 11, 15, 21, 21–32], which allowed for detailed study of specific PK/PD models, some of which are too complicated for mathematical analysis. Previous works which have utilized detailed mathematical analysis include [19, 20, 33–36]. These prior analytical works considered isolated PK models rather than the coupled PK/PD models considered in the present work.

As a theoretical investigation, our model and analysis necessarily makes many simplifying assumptions. For example, while the PK/PD models studied in this paper are quite standard, certainly many drugs have been described by more complicated models (for example, models with more compartments such as transit compartment models [37, 38]). In addition, we assumed a continuous drug effect, whereas the more clinically relevant effect in some settings is binary. In a more general model, we would not be assured that the PK absorption, PK elimination, and PD elimination rates would be exactly equivalent in their contributions to effect fluctuations, as we found for the models in this paper. Nevertheless, this work demonstrates the general result that the PK elimination rate is not a uniquely determinative factor for drug forgiveness. This result is not surprising, but it emphasizes the point that other PK rates and PD rates must be considered when estimating forgiveness.

While the results in "Average relative effect is the average drug intake"–"PK absorption, PK elimination, and PD elimination rates are equivalentfor effect fluctuations" were proven for any pattern of nonadherence, the simple model of patient adherence in "Coefficient of variation estimate"–""Make up" doses reduce variation for slow PK or PD" assumes that a patient misses a dose independently of their prior behavior. However, some data shows correlations in the doses taken by some patients [39], meaning that, for example, a patient may be more likely to miss a dose if they missed their prior dose. We ignored any such correlations and assumed independence in order to simplify the analysis and results. Indeed, even in this simplified scenario, the formulas obtained in "Coefficient of variation estimate" and ""Make up" doses reduce variation for slow PK or PD" are quite complicated. Furthermore, the existence of such correlations for all or even most patients is debatable. Indeed, in a detailed statistical analysis of the data in [40], Sun et al. [41] found that only one third of the patients studied showed sufficient evidence to reject the hypothesis of independent dosing.

One key assumption of this work is that the drug effect is far from maximal, which makes the PD models linear. This assumption follows the computational study of Boissel and Nony [15], who noted that this is the exact scenario in which adherence is most critical. In particular, due to the plateauing nature of typical dose–response curves (such as the concentration–effect relations in the PD models in "Methods"), missing doses only mildly reduces the drug effect if the prescribed dose is near the effect plateau. In contrast, in the scenario studied in the present paper in which the prescribed dose is much less than the dose yielding maximal effect, missed doses cause a significant decrease in drug effect. Nevertheless, investigating the effects of nonadherence in the case that the drug effect is not necessarily far from maximal remains an important area for future research.

Appendix

In this appendix, we present the mathematical proofs and derivations of some results in the main text.

Proof of formula for $E(t)$ in (21)

To verify (21), note first that (21) and the definitions in (22) imply that $E(0) = 0$, as desired. Next, the definition of $E(t)$ in (18) and the ODE for c_e in (6) imply that $E(t)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}E = \frac{E_{\max}}{\mathrm{EC}_{50}} k_{\mathrm{pe}} c_{\mathrm{p}} - k_{\mathrm{e0}} E. \tag{49}$$

Differentiating (21) with respect to t and using the value of c_{p} in (4) shows that (21) indeed satisfies (49), which verifies (21). □

Proof that $\langle E \rangle = \mu \langle E^{\mathrm{perf}} \rangle$

To prove $\langle E \rangle = \mu \langle E^{\mathrm{perf}} \rangle$, we let $k > 0$ and consider the large T behavior of

$$F(T,k) := \frac{1}{T}\int_0^T \sum_{n:t_n \le t} e^{-k(t-t_n)} f_n \, \mathrm{d}t.$$

Interchanging the sum and integral yields

$$\begin{aligned} F(T,k) &= \frac{1}{T}\sum_{n:t_n \le T} \int_{t_n}^T e^{-k(t-t_n)} f_n \, \mathrm{d}t \\ &= \frac{1}{k}\frac{1}{T}\sum_{n:t_n \le T} f_n - \frac{1}{k}\frac{1}{T}\sum_{n:t_n \le T} e^{-k(T-t_n)} f_n. \end{aligned} \tag{50}$$

By (30), the first term in (50) has the large T limit,

$$\lim_{T\to\infty} \frac{1}{k}\frac{1}{T}\sum_{n:t_n \le T} f_n = \mu \frac{1}{k\tau}. \tag{51}$$

We claim that the second term in (50) vanishes as $T \to \infty$. To see this, let $\varepsilon \in (0,1)$ and observe that

$$\begin{aligned} &\frac{1}{T}\sum_{n:t_n \le T} e^{-k(T-t_n)} f_n \\ &\quad = \frac{1}{T}\sum_{n:t_n \le (1-\varepsilon)T} e^{-k(T-t_n)} f_n + \frac{1}{T}\sum_{n:(1-\varepsilon)T < t_n \le T} e^{-k(T-t_n)} f_n \\ &\quad \le e^{-\varepsilon k T}\frac{1}{T}\sum_{n:t_n \le (1-\varepsilon)T} f_n + \frac{1}{T}\sum_{n:(1-\varepsilon)T < t_n \le T} f_n. \end{aligned}$$

Since (30) implies that

$$\lim_{T\to\infty} e^{-\varepsilon k T}\frac{1}{T}\sum_{n:t_n \le (1-\varepsilon)T} f_n = 0,$$

it follows that

$$\lim_{T\to\infty} \frac{1}{T}\sum_{n:t_n \le T} e^{-k(T-t_n)} f_n \le \lim_{T\to\infty} \frac{1}{T}\sum_{n:(1-\varepsilon)T < t_n \le T} f_n. \tag{52}$$

To estimate the upper bound in (52), let $\varepsilon_2 > 0$ and observe that (30) implies that if T is sufficiently large, then

$$\frac{1}{T}\sum_{n:t_n \le T(1-\varepsilon)} f_n + \frac{1}{T}\sum_{n:T(1-\varepsilon) < t_n \le T} f_n \le \mu/\tau + \varepsilon_2, \tag{53}$$

and

$$\mu/\tau - \varepsilon_2 \le \frac{1}{T(1-\varepsilon)}\sum_{n:t_n \le T(1-\varepsilon)} f_n. \tag{54}$$

Combining (53)–(54) yields

$$\begin{aligned} \lim_{T\to\infty}\frac{1}{T}\sum_{n:T(1-\varepsilon) < t_n \le T} f_n &\le \mu/\tau + \varepsilon_2 - (1-\varepsilon)(\mu/\tau - \varepsilon_2) \\ &= 2\varepsilon_2 + \varepsilon\mu/\tau - \varepsilon\varepsilon_2. \end{aligned}$$

Since $\varepsilon \in (0,1)$ and $\varepsilon_2 > 0$ are arbitrary, we conclude from (52) that

$$\lim_{T\to\infty}\frac{1}{T}\sum_{n:t_n \le T} e^{-k(T-t_n)} f_n = 0,$$

and therefore (50) and (51) imply

$$\lim_{T\to\infty} F(T,k) = \mu\frac{1}{k\tau}. \tag{55}$$

Finally, using the expression for $E(t)$ in (25), the values of f_n and t_n in (26) for perfect adherence, and the result in (55), we conclude that

$$\langle E \rangle = \mu \frac{E_{\max}}{\mathrm{EC}_{50}}\frac{DF}{V}\left(\frac{b_{\mathrm{p0}}}{k_{\mathrm{p0}}\tau} + \frac{b_{\mathrm{ap}}}{k_{\mathrm{ap}}\tau} + \frac{b_{\mathrm{e0}}}{k_{\mathrm{e0}}\tau}\right) = \mu\langle E^{\mathrm{perf}}\rangle.$$

□

Derivation of CV(E^{single})

Using the model of nonadherence introduced in "Coefficient of variation estimate", applying (33) gives

$$\begin{aligned} \psi(t) := \frac{E(t)}{\mu\langle E^{\mathrm{perf}}\rangle} = \frac{1}{\mu}\sum_{n=0}^{\lfloor t/\tau \rfloor} \Big(& a_{\mathrm{p0}} e^{-k_{\mathrm{p0}}(t-n\tau)} \\ & + a_{\mathrm{ap}} e^{-k_{\mathrm{ap}}(t-n\tau)} + a_{\mathrm{e0}} e^{-k_{\mathrm{e0}}(t-n\tau)} \Big) f_n, \end{aligned}$$

where the coefficients a_{p0}, a_{ap}, and a_{e0} are defined in (34), $\lfloor t/\tau \rfloor$ denotes the largest integer less than or equal to t/τ, and $\{f_n\}_{n\in\mathbb{Z}}$ are as in (37). Note that we now take the sequence $\{f_n\}_{n\in\mathbb{Z}}$ to be bi-infinite (i.e. $n \in \mathbb{Z}$), which is convenient for the analysis below.

For $t \in [0,\tau]$, define

$$\Psi(t) := \frac{1}{\mu}\left(a_{\mathrm{p0}} e^{-k_{\mathrm{p0}}t} A_{\mathrm{p0}} + a_{\mathrm{ap}} e^{-k_{\mathrm{ap}}t} A_{\mathrm{ap}} + a_{\mathrm{e0}} e^{-k_{\mathrm{e0}}t} A_{\mathrm{e0}}\right), \tag{56}$$

where

$$A_{ij} := \sum_{n=0}^{\infty} (e^{-k_{ij}\tau})^n f_{-n}, \quad \text{for } i,j \in \{\mathrm{a},\mathrm{p},0\}. \tag{57}$$

Note that A_{ij} converges almost surely by the Weierstrass M-test since $|f_n|$ is bounded by a deterministic constant for all $n \in \mathbb{Z}$ (namely, $|f_n| \le 1$ in this case). Random variables

of the form in (57) are often referred to as random pullback attractors [42–46].

It then follows from (32) and Theorem 1 in [20] that

$$\begin{aligned} \mathrm{CV}(E^{\mathrm{single}}) &= \sqrt{\lim_{T\to\infty}\frac{1}{T}\int_0^T \Big(\psi(t)-1\Big)^2\,\mathrm{d}t} \\ &= \sqrt{\frac{1}{\tau}\int_0^\tau \mathbb{E}\Big[\big(\Psi(t)-1\big)^2\Big]\,\mathrm{d}t}, \end{aligned} \tag{58}$$

where $\mathbb{E}$ denotes mathematical expectation. Using the definition of $\Psi(t)$ in (56), equation (58) implies that $\mathrm{CV}(E^{\mathrm{single}})$ requires calculating expectations of the form

$$\mathbb{E}[A_{ij}] \quad \text{and} \quad \mathbb{E}[A_{ij}A_{lm}], \quad \text{for } i,j,l,m \in \{\mathrm{a},\mathrm{p},0\}. \tag{59}$$

These expectations are obtained immediately from Corollary 3 in [20]. In particular, we have that

$$\begin{aligned} \mathbb{E}[A_{ij}] &= \frac{p}{1-e^{-k_{ij}\tau}}, \\ \mathbb{E}[A_{ij}A_{lm}] &= \frac{p}{1-e^{-k_{ij}\tau}e^{-k_{lm}\tau}} \\ &\quad + p^2\Big(\frac{1}{(1-e^{-k_{ij}\tau})(1-e^{-k_{lm}\tau})} \\ &\quad - \frac{1}{1-e^{-k_{ij}\tau}e^{-k_{lm}\tau}}\Big). \end{aligned}$$

Plugging these formulas into (58) and performing the integration yields for the formula for $\mathrm{CV}(E^{\mathrm{single}})$ in (39).

Derivation of μ^{double}

For the double dose protocol in "“Make up” doses reduce variation for slow PK or PD", it is immediate that $\{f_n\}_{n\in\mathbb{Z}}$ is an irreducible, time-homogeneous Markov chain with the following transition probabilities,

$$\begin{aligned} \mathbb{P}(f_{n+1}=0\,|\,f_n=i) &= 1-p \quad \text{if } i\in\{0,1,2\}, \\ \mathbb{P}(f_{n+1}=1\,|\,f_n=i) &= \begin{cases} 0 & \text{if } i=0, \\ p & \text{if } i=1, \\ p & \text{if } i=2, \end{cases} \\ \mathbb{P}(f_{n+1}=2\,|\,f_n=i) &= \begin{cases} p & \text{if } i=0, \\ 0 & \text{if } i=1, \\ 0 & \text{if } i=2. \end{cases} \end{aligned}$$

Using standard Markov chain results (for example, see section 1.7 in [47]), a quick calculation shows that the stationary distribution of $\{f_n\}_{n\in\mathbb{Z}}$ is

$$\mathbb{P}(f_n=i) = \begin{cases} 1-p & \text{if } i=0, \\ p^2 & \text{if } i=1, \\ p(1-p) & \text{if } i=2. \end{cases}$$

Using (30), Birkhoff's ergodic theorem (for example, see section 7.2 in [48]) thus implies that the long-term fraction of doses taken for the double dose protocol is

$$\mu^{\mathrm{double}} = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} f_n = \sum_{i=0}^{2} i\,\mathbb{P}(f_n=i) = p+p(1-p).$$

Derivation of CV(E^{double})

The derivation of the formula for $\mathrm{CV}(E^{\mathrm{double}})$ in (47) is identical to the derivation presented in "Derivation of $\mathrm{CV}(E^{\mathrm{single}})$" above except the formulas for the expectations in (59) are given by the formulas for the double dose protocol in Corollary 3 in [20].

Maximum plasma concentration for double dose protocol

Theorem 5 in [20] yields the following upper bound for the maximum possible plasma compartment concentration obtained by following the double dose protocol,

$$c_{\mathrm{p}}^{\mathrm{double}}(t) \le \max_{s\ge 0} c_{\mathrm{p}}^{\mathrm{perf}}(s) + \varphi\langle c_{\mathrm{p}}^{\mathrm{perf}}\rangle, \tag{60}$$

where φ is the following dimensionless factor,

$$\varphi := \tau(k_{\mathrm{ap}})^{\frac{k_{\mathrm{p0}}}{k_{\mathrm{p0}}-k_{\mathrm{ap}}}}(k_{\mathrm{p0}})^{\frac{k_{\mathrm{ap}}}{k_{\mathrm{ap}}-k_{\mathrm{p0}}}}, \tag{61}$$

and $\langle c_{\mathrm{p}}^{\mathrm{perf}}\rangle = \frac{DF}{V}\frac{1}{\tau}\frac{1}{k_{\mathrm{p0}}}$ is the long-term average plasma concentration for perfect adherence. The upper bound in (60) is valid for any time $t\ge 0$ and any sequence of remembering or forgetting $\{\xi_n\}_{n\ge 0}$ for the adherence model in "“Make up” doses reduce variation for slow PK or PD" above. The first term in the righthand side of (60) is the maximum plasma concentration obtained by a patient with

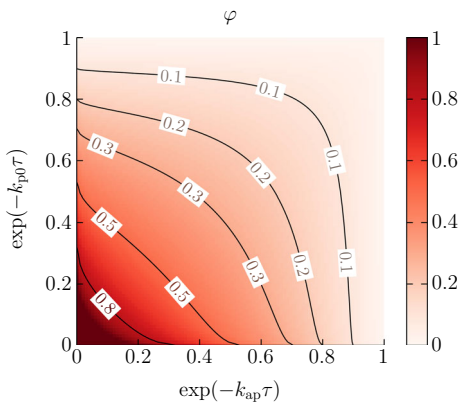


Fig. 8 Contour plot of the factor φ in (61) which bounds the maximum increase in plasma concentration caused by following the double dose protocol (see (60))

perfect adherence. The second term in the righthand side of (60) bounds the maximum possible increase in plasma concentration caused by the double dose protocol.

In Fig. 8, we plot the factor φ in (61). This plot shows that $\varphi \ll 1$ if $k_{ap}\tau \ll 1$ and/or $k_{p0}\tau \ll 1$. This means that if the PK absorption rate is slow compared to the dosing interval ($k_{ap}\tau \ll 1$) and/or the PK elimination rate is slow compared to the dosing interval ($k_{p0}\tau \ll 1$), then the double dose protocol can at most cause the plasma concentration to rise only slightly above the concentration for perfect adherence.

Funding SDL was supported by the National Science Foundation (Grant Nos. DMS-1944574 and DMS-1814832).

References

1. Vrijens B, De Geest S, Hughes DA, Przemyslaw K, Demonceau J, Ruppar T, Dobbels F, Fargher E, Morrison V, Lewek P et al (2012) A new taxonomy for describing and defining adherence to medications. Br J Clin Pharmacol 73(5):691–705
2. Osterberg L, Blaschke T (2005) Adherence to medication. N Engl J Med 353(5):487–497
3. Eduardo S et al (2003) Adherence to long-term therapies: evidence for action. World Health Organization, Geneva
4. Brian HR, Pauline MH, Amit G, Patty M (2002) Interventions for helping patients to follow prescriptions for medications. Cochrane Database Syst Rev. https://doi.org/10.1002/14651858.CD000011
5. Urquhart J (1998) Pharmacodynamics of variable patient compliance: implications for pharmaceutical value. Adv Drug Deliv Rev 33(3):207–219
6. Assawasuwannakit P, Braund R, Duffull SB (2015) Quantification of the forgiveness of drugs to imperfect adherence. CPT: Pharmacomet Syst Pharmacol 4(3):204–211
7. Assawasuwannakit P, Braund R, Duffull SB (2016) A framework for quantifying the influence of adherence and dose individualization. Clin Pharmacol Ther 99(4):452–459
8. Pellock JM, Brittain ST (2016) Use of computer simulations to test the concept of dose forgiveness in the era of extended-release (XR) drugs. Epilepsy Behav 55:21–23
9. Morrison A, Stauffer ME, Kaufman AS (2017) Relationship between adherence rate threshold and drug 'forgiveness'. Clin Pharmacokinet 56(12):1435–1440
10. Véronique D (2011) Drug forgiveness and interpatient pharmacokinetic variability in tuberculosis. J Infect Dis 204:1827–1829
11. Lowy A, Munk VC, Ong SH, Burnier M, Vrijens B, Tousset EP, Urquhart J (2011) Effects on blood pressure and cardiovascular risk of variations in patients? adherence to prescribed antihypertensive drugs: role of duration of drug action. Int J Clin Pract 65(1):41–53
12. Nony P, Boissel J-P (2002) Use of sensitivity functions to characterise and compare the forgiveness of drugs. Clin Pharmacokinet 41(5):371–380
13. Gohore DG, Fenneteau F, Barrière O, Li J, Nekka F et al (2010) Rational drug delineation: a global sensitivity approach based on therapeutic tolerability to deviations in execution. Pharmacol Pharm 1(02):42
14. Osterberg LG, Urquhart J, Blaschke TF (2010) Understanding forgiveness: minding and mining the gaps between pharmacokinetics and therapeutics. Clin Pharmacol Ther 88(4):457–459
15. Boissel J-P, Nony P (2002) Using pharmacokinetic-pharmacodynamic relationships to predict the effect of poor compliance. Clin Pharmacokinet 41(1):1–6
16. Melanie AF, Marilyn EM, Donald EM (2012) Mechanism-based pharmacodynamic modeling. Computational toxicology. Springer, New York, pp 583–600
17. Hong Y, Rohatagi S, Habtemariam B, Walker JR, Schwartz SL, Mager DE (2008) Population exposure-response modeling of metformin in patients with type 2 diabetes mellitus. J Clin Pharmacol 48(6):696–707
18. Bender CM, Orszag SA (2013) Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory. Springer, New York
19. Counterman ED, Lawley SD (2021) What should patients do if they miss a dose of medication? A theoretical approach. J Pharmacokinet Pharmacodynam
20. Counterman ED, Lawley SD (2022) Designing drug regimens that mitigate nonadherence. Bull Math Biol 84(1):1–36
21. Stauffer ME, Hutson P, Kaufman AS, Morrison A (2017) The adherence rate threshold is drug specific. Drugs in R&D 17(4):645–653
22. Garnett WR, McLean AM, Zhang Y, Clausen S, Tulloch SJ (2003) Simulation of the effect of patient nonadherence on plasma concentrations of carbamazepine from twice-daily extended-release capsules. Curr Med Res Opin 19(6):519–525
23. Reed RC, Dutta S (2004) Predicted serum valproic acid concentrations in patients missing and replacing a dose of extended-release divalproex sodium. Am J Health Syst Pharm 61(21):2284–2289
24. Dutta S, Reed RC (2006) Effect of delayed and/or missed enteric-coated divalproex doses on valproic acid concentrations: simulation and dose replacement recommendations for the clinician 1. J Clin Pharm Ther 31(4):321–329
25. Ding J, Zhang Y, Jiao Z, Wang Y (2012) The effect of poor compliance on the pharmacokinetics of carbamazepine and its epoxide metabolite using monte carlo simulation. Acta Pharmacol Sin 33(11):1431–1440
26. Chen C, Wright J, Gidal B, Messenheimer J (2013) Assessing impact of real-world dosing irregularities with lamotrigine extended-release and immediate-release formulations by pharmacokinetic simulation. Ther Drug Monit 35(2):188–193
27. Gidal Barry E, Oneeb M, Jim F, Ziad H, Haichen Y, Jin Z, Randi F, Antonio L (2014) The practical impact of altered dosing on perampanel plasma concentrations: pharmacokinetic modeling from clinical studies. Epilepsy Behav 35:6–12
28. Brittain ST, Wheless JW (2015) Pharmacokinetic simulations of topiramate plasma concentrations following dosing irregularities with extended-release vs immediate-release formulations. Epilepsy Behav 52:31-36D
29. Sunkaraneni S, Blum D, Ludwig E, Chudasama V, Fiedler-Kelly J, Marvanova M, Bainbridge J, Phillips L (2018) Population pharmacokinetic evaluation and missed-dose simulations for eslicarbazepine acetate monotherapy in patients with partial-onset seizures. Clin Pharmacol Drug Dev 7(3):287–297
30. Hard ML, Wehr AY, Sadler BM, Mills RJ, Moltke L (2018) Population pharmacokinetic analysis and model-based simulations of aripiprazole for a 1-day initiation regimen for the long-acting antipsychotic aripiprazole lauroxil. Eur J Drug Metab Pharmacokinet 43(4):461–469
31. Elkomy MH (2020) Changing the drug delivery system: Does it add to non-compliance ramifications control? A simulation study on the pharmacokinetics and pharmacodynamics of atypical antipsychotic drug. Pharmaceutics 12(4):297

32. Jia-qin G, Guo Y, Jiao Z, Ding J, Li G-F (2020) How to handle delayed or missed doses: a population pharmacokinetic perspective. Eur J Drug Metab Pharmacokinet 45(2):163–172
33. Li J, Nekka F (2007) A pharmacokinetic formalism explicitly integrating the patient drug compliance. J Pharmacokinet Pharmacodyn 34(1):115–139
34. Lévy-Véhel P-E, Lévy-Véhel J (2013) Variability and singularity arising from poor compliance in a pharmacokinetic model I: the multi-IV case. J Pharmacokinet Pharmacodyn 40(1):15–39
35. Fermín LJ, Lévy-Véhel J (2017) Variability and singularity arising from poor compliance in a pharmacokinetic model II: the multi-oral case. J Math Biol 74(4):809–841
36. Jie M (2017) *Stochastic modeling of random drug taking processes and the use of singular perturbation methods in pharmacokinetics*. Ph.D. Thesis, The University of Utah
37. Hof F, Bridge LJ (2021) Exact solutions and equi-dosing regimen regions for multi-dose pharmacokinetics models with transit compartments. J Pharmacokinet Pharmacodyn 48(1):99–131
38. Savic RM, Jonker DM, Kerbusch T, Karlsson MO (2007) Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies. J Pharmacokinet Pharmacodyn 34(5):711–726
39. Vrijens B, Vincze G, Kristanto P, Urquhart J, Burnier M (2008) Adherence to prescribed antihypertensive drug treatments: longitudinal study of electronically compiled dosing histories. BMJ 336(7653):1114–1117
40. Hammer SM, Vaida F, Bennett KK, Holohan MK, Sheiner L, Eron JJ, Wheat LJ, Mitsuyasu RT, Gulick RM, Valentine FT et al (2002) Dual vs single protease inhibitor therapy following antiretroviral treatment failure: a randomized trial. JAMA 288(2):169–180
41. Sun J, Nagaraj HN, Reynolds NR (2007) Discrete stochastic models for compliance analysis based on an AIDS clinical trial group (ACTG) study. Biom J 49(5):731–741
42. Crauel H (2001) Random point attractors versus random set attractors. J Lond Math Soc (2) 63(2):413–427
43. Mattingly JC (1999) Ergodicity of 2D Navier–Stokes equations with random forcing and large viscosity. Commun Math Phys 206(2):273–288
44. Schmalfuß B (1996) A random fixed point theorem based on Lyapunov exponents. Random Comput Dynam 4(4):257–268
45. Lawley SD, Mattingly JC, Reed MC (2015) Stochastic switching in infinite dimensions with applications to random parabolic PDE. SIAM J Math Anal 47(4):3035–3063
46. Lawley SD, Keener JP (2019) Electrodiffusive flux through a stochastically gated ion channel. SIAM J Appl Math 79(2):551–571
47. Norris JR (1998) Markov chains. Statistical & probabilistic mathematics. Cambridge University Press, Cambridge
48. Durrett R (2010) Probability: theory and examples, 4th edn. Cambridge University Press, Cambridge