Journal of Pharmacokinetics and Pharmacodynamics
https://doi.org/10.1007/s10928-023-09897-1

ORIGINAL PAPER

How drug onset rate and duration of action affect drug forgiveness

Elias D. Clark[1,2] · **Sean D. Lawley**[2]

Received: 9 August 2023 / Accepted: 7 December 2023

Abstract
Medication nonadherence is one of the largest problems in healthcare today, particularly for patients undergoing long-term pharmacotherapy. To combat nonadherence, it is often recommended to prescribe so-called "forgiving" drugs, which maintain their effect despite lapses in patient adherence. Nevertheless, drug forgiveness is difficult to quantify and compare between different drugs. In this paper, we construct and analyze a stochastic pharmacokinetic/pharmacodynamic (PK/PD) model to quantify and understand drug forgiveness. The model parameterizes a medication merely by an effective rate of onset of effect when the medication is taken (on-rate) and an effective rate of loss of effect when a dose is missed (off-rate). Patient dosing is modeled by a stochastic process that allows for correlations in missed doses. We analyze this "on/off" model and derive explicit formulas that show how treatment efficacy depends on drug parameters and patient adherence. As a case study, we compare the effects of nonadherence on the efficacy of various antihypertensive medications. Our analysis shows how different drugs can have identical efficacies under perfect adherence, but vastly different efficacies for adherence patterns typical of actual patients. We further demonstrate that complex PK/PD models can indeed be parameterized in terms of effective on-rates and off-rates. Finally, we have created an online app to allow pharmacometricians to explore the implications of our model and analysis.

Keywords Drug forgiveness · Medication nonadherence · Medication adherence · Stochastic modeling · Probability · Monte Carlo simulations

Introduction

Medication adherence is the process by which patients take their medications as prescribed [1]. One of the most significant problems in healthcare today is the issue of medication *non*adherence. Indeed, some studies have suggested that nonadherence accounts for over 100,000 avoidable deaths in the United States every year as well as over $100 billion in preventable excess healthcare costs [2, 3]. Medication adherence is often described in terms of the three phases of initiation, implementation, and discontinuation [1] (and the term "persistence" is sometimes used to describe the time from initiation to discontinuation [1, 4]). In this paper, we analyze the implementation phase, which is the extent to which a patient's actual dosing follows the prescribed dosing regimen [1].

Patient adherence in long-term pharmacotherapy is difficult to measure and control, with data often coming from limited clinical studies that may not accurately reflect the patterns of nonadherence in actual clinical practice [5]. Furthermore, quantifying the effects of nonadherence is challenging and often relies on numerical simulations of specific drugs in specific nonadherence scenarios [5, 6, 6–20] since pharmacokinetic/pharmacodynamic (PK/PD) models often consist of systems of nonlinear differential equations involving many parameters [21, 22]. Hence, such bespoke computer simulations hinder the discovery of general quantitative principles regarding nonadherence.

One common strategy to mitigate the effects of patient nonadherence is to prescribe what are called "forgiving" drugs [23]. While there is no universal definition for drug forgiveness, in general, a drug is considered forgiving if efficacy is maintained in spite of missed or delayed doses.

✉ Sean D. Lawley
lawley@math.utah.edu

Elias D. Clark
elias.clark@metrumrg.com

[1] Metrum Research Group, 2 Tunxis Road, Suite 112, Tariffville, CT 06081, USA

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

Published online: 10 January 2024

While this is an intuitive definition, it is rather qualitative and there is a need to determine what aspects of different medications lead to higher forgiveness. For example, under what conditions could it be more important to have a fast acting medication rather than one with a long half-life? Another question of great importance is how the nature of a patient's nonadherence may affect treatment efficacy. Some studies have shown that when patients miss doses, they are likely to miss either single doses or at least three sequential doses, so called 'drug holidays' [2, 24]. Hence, a simple model of patient dosing that assumes independence between doses (i.e. missing one dose does not change the probability the patient misses the next dose) may not accurately reflect adherence patterns in actual patients.

In this paper, we formulate and analyze a stochastic model of the effects of missed doses for patients undergoing long-term treatment. We use the model to estimate how drug characteristics and adherence statistics combine to yield the drug effect experienced by the patient. The model describes a medication in terms of an effective on-rate r_{on} (rate of onset of drug effect on a biomarker) and an effective off-rate r_{off} (rate of loss of drug effect on a biomarker). We assume the patient is instructed to take the medication at regular intervals of time τ, and we parameterize their adherence by the proportion of prescribed doses the patient takes, p, and the correlation between missed doses, q. That is, a patient takes proportion p of their doses and patients who miss a dose may be more likely (if $q > 0$), less likely (if $q < 0$), or equally likely (if $q = 0$) to miss the following dose.

We analyze this "on/off" model and derive explicit formulas for clinically relevant metrics such as the (i) average drug effect, (ii) standard deviation of drug effect, and (iii) long-term fraction of time above a given minimal threshold for effective treatment. Indeed, (iii) has been called "the most flexible and clinically most meaningful measure of noncompliance" [25]. We find that different drugs can have identical efficacies under perfect adherence, but vastly different efficacies for adherence patterns typical of actual patients. We illustrate this general result by comparing the effects of nonadherence on the efficacy of seven antihypertensive medications. We further demonstrate that complex PK/PD models consisting of systems of nonlinear differential equations and involving many parameters can indeed be accurately described by our simple on/off model. We have also created an online app to allow pharmacometricians to explore the implications of our model and analysis [26]. The app is available at https://seanlawley.shinyapps.io/OnOff/. The rest of the paper is organized as follows. In the "Methods" section, we formulate and analyze our mathematical model (details of the mathematical analysis are in the Appendix). In the "Results" section, we present the pharmacological implications of our model and analysis. In particular, we present the average drug effect (the "Average drug effect" section), drug effect variability (the "Drug effect variability" section), application to antihypertensive medications (the "Application to antihypertensive medications" section), and fraction of time the drug effect is in a desired therapeutic window (the "Time in therapeutic window" section). We then compare our simple on/off model to more complex PK/PD models in the "On/off model approximates complex PK/PD models" section, explore the effects of correlations in missed doses in the "Correlations in missed doses" section, and extend our analysis to a population model in the "Population PK/PD" section. We conclude by discussing some basic guidelines concerning the effects of nonadherence in the "Discussion" section.

Methods

In this section we formulate the mathematical model and analyze it to derive formulas for various metrics of drug efficacy.

Biomarker dynamics

Assume that a fixed dose size of a drug is prescribed at regular time intervals of length τ. The model describes the effect of this drug on a biomarker whose response level is scaled to be between 0 and 1, with 1 indicating full desired effect and 0 indicating no effect. The model tracks the average value of this biomarker response over each dosing interval, denoted by the sequence $\{X_n\}_n$. That is, the response levels $\{X_n\}_n$ model the proportion of the full drug effect the patient experiences on average over each dosing interval. For instance, a value of $X_n = 0.8$ means that the average drug effect the patient experiences between the n-th and $(n+1)$-th scheduled doses is 80% of the full drug effect under perfect adherence.

We assume that the average biomarker response over each dosing interval changes according to the relationship

$$X_{n+1} = \begin{cases} \alpha X_n & \text{dose } (n+1) \text{ not taken,} \\ 1 - \beta(1 - X_n) & \text{dose } (n+1) \text{ taken,} \end{cases} \tag{1}$$

where

$$\alpha = e^{-r_{\text{off}}\tau} \quad \text{and} \quad \beta = e^{-r_{\text{on}}\tau}.$$

In words, if the patient misses a dose, then over the next dosing interval the average biomarker response decreases exponentially at rate r_{off}. If the patient takes the dose, then the average biomarker response increases exponentially to its maximum value of 1 at rate r_{on}.

Patient adherence

We model the patient's dosing history by a sequence of identically distributed Bernoulli random variables $\{\xi_n\}_n$, where

$$\xi_n = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1-p. \end{cases} \tag{2}$$

That is, $\xi_n = 1$ means the patient took dose n, and $\xi_n = 0$ means that the patient missed dose n. The parameter $p \in (0,1)$ is probability the patient takes any given dose. Combining (2) with (1), the biomarker response levels follow the recursion relationship

$$X_{n+1} = \alpha X_n + (\beta - \alpha) X_n \xi_{n+1} + (1-\beta)\xi_{n+1}, \quad n \in \mathbb{Z}. \tag{3}$$

Note that the dosing time index n varies over the positive and negative integers, which models a patient who has been prescribed the medication for a long time (i.e. long-term pharmacotherapy).

We do not necessarily assume that $\{\xi_n\}_{n\in\mathbb{Z}}$ are independent. In particular, we assume that $\{\xi_n\}_{n\in\mathbb{Z}}$ is a two-state Markov chain with the following transition probabilities [27],

$$\mathbb{P}(\xi_{n+1} = 0 \mid \xi_n = 0) = p_0, \qquad \mathbb{P}(\xi_{n+1} = 1 \mid \xi_n = 0) = 1 - p_0 \tag{4}$$

$$\mathbb{P}(\xi_{n+1} = 0 \mid \xi_n = 1) = 1 - p_1, \quad \mathbb{P}(\xi_{n+1} = 1 \mid \xi_n = 1) = p_1. \tag{5}$$

In words, (4)–(5) means that if the patient takes a dose, then they have probability p_1 of taking the next dose, and if they miss a dose, then they have probability p_0 of missing the next dose.

The probability that the patient takes a given dose is then (see the Appendix)

$$p = \frac{1-p_0}{2-p_0-p_1} \in (0,1). \tag{6}$$

Further, the correlation coefficient between any two successive doses is then

$$q = \mathrm{Corr}\left(\xi_n, \xi_{n+1}\right) = \frac{\mathbb{E}[\xi_n\xi_{n+1}] - \mathbb{E}[\xi_n]\mathbb{E}[\xi_{n+1}]}{\mathrm{Var}\left(\xi_0\right)} = \frac{p_1 - p}{1-p} \in [-1,1].$$

Note that, in general, correlation coefficients can range from $q = -1$ to $q = 1$, but fixing the value of $p \in (0,1)$ places restrictions on the range of q. Specifically,

$$\text{if } p < 1/2, \text{ then } q > -p/(1-p),$$
$$\text{if } p \geq 1/2, \text{ then } q > (p-1)/p.$$

Mean and variance of biomarker response

The distribution of X_n is independent of $n \in \mathbb{Z}$ and thus we omit the subscript when describing its distribution or statistics (i.e. $X_n = X$). In the Appendix, we show that the mean biomarker response is given by the following explicit formula,

$$\mathbb{E}[X] = \frac{p(q\alpha - 1)(1-\beta)}{\alpha(1-p+pq) + \beta(q+p-pq) - q\alpha\beta - 1}. \tag{7}$$

We further show that the second moment of the biomarker response is given by

$$\begin{aligned} \mathbb{E}[X^2] = (1-\beta)^2 p\left(\alpha^2 q - 1\right)\big[&\alpha + \alpha p(q-1) \\ &+ \beta p(q-1) + \alpha\beta q - \beta q - 1\big] \\ &\times \big[\alpha + \alpha p(q-1) + p(\beta - \beta q) \\ &- \alpha\beta q + \beta q - 1\big]^{-1} \\ &\times \big[\alpha^2\left(p(q-1) - \beta^2 q + 1\right) \\ &- \beta^2 p(q-1) + \beta^2 q - 1\big]^{-1}. \end{aligned} \tag{8}$$

With these two formulas, we immediately obtain formulas for the variance of the biomarker response via the relation $\mathrm{Variance}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X]^2)$, as well as the standard deviation, $\mathrm{SD}(X) = \sqrt{\mathrm{Variance}(X)}$.

While we have computed the first and second moments of the biomarker response X, it is to our knowledge not possible to obtain explicit formulas for the full probability distribution of X. Indeed, our model generalizes so-called infinite Bernoulli convolutions [28–31], which are well-known to have highly irregular probability distributions [32]. The study of infinite Bernoulli convolutions dates back to at least the 1930s [33–35], and in more recent years has been applied to pharmacokinetic models [36–39].

Nevertheless, we find that the distribution of X can be well-approximated by a simple, smooth probability distribution in many parameter regimes of pharmacological interest. Specifically, we approximate the distribution of X by the distribution of a Beta random variable B whose two parameters, denoted a and b, are chosen so that the first and second moments of B and X agree. The explicit formulas for this Beta distribution fit are collected in the Appendix in the "Beta distribution formulas" section.

Results

We now explore the pharmacological implications of our mathematical model and analysis. Recall that $X \in [0,1]$ denotes the biomarker response relative to perfect adherence (so that $X = 1$ corresponds to a perfectly adherent patient and $X = 0$ corresponds to a patient who never takes any medication). The medication is parameterized by its on-rate r_{on} (effective rate of onset of effect when the medication is taken) and its off-rate r_{off} (effective rate of loss of effect when a dose is missed). The prescribed dosing interval is τ and patient adherence is described by the fraction $p \in (0,1)$

of doses taken and the correlation $q \in (-1, 1)$ between successive doses.

Average drug effect

We first consider the case where $q = 0$ (i.e. no correlation between missed doses). The average drug effect (relative to perfect adherence) in (7) then reduces to

$$\mathbb{E}[X] = \frac{p(1 - e^{-r_{\text{on}}\tau})}{1 - e^{-r_{\text{off}}\tau} + p(e^{-r_{\text{off}}\tau} - e^{-r_{\text{on}}\tau})}. \tag{9}$$

It is straightforward to check that this formula for $\mathbb{E}[X]$ increases if we increase p, increase r_{on}, or decrease r_{off}, which we express mathematically in terms of partial derivatives,

$$\frac{\partial \mathbb{E}[X]}{\partial p} > 0, \quad \frac{\partial \mathbb{E}[X]}{\partial r_{\text{on}}} > 0, \quad \frac{\partial \mathbb{E}[X]}{\partial r_{\text{off}}} < 0. \tag{10}$$

Since improving patient adherence can be difficult, equations (9)–(10) delineate how drug efficacy can be improved by adjusting drug formulations to increase r_{on} or decrease r_{off}. Essentially, a high on-rate means that when the patient takes the drug after missing a dose(s), they will quickly recover the drug effect. A low off-rate means that when the patient misses a dose(s), they will not lose much drug effect before they start taking the medication again.

Figure 1 plots $\mathbb{E}[X]$ as a function of $r_{\text{on}}\tau$ for $p = 80\%$. This plot shows that $\mathbb{E}[X]$ can differ substantially from p. Indeed, though $p = 80\%$ is often considered to be an adequate level of adherence (the so-called 80% rule) [40, 41], Fig. 1 shows that depending on drug formulation (i.e. depending on r_{on} and r_{off}), a patient taking $p = 80\%$ of their doses can experience a drug effect as low as $\mathbb{E}[X] = 28\%$ or as high as $\mathbb{E}[X] = 98\%$ of the drug effect they would receive with perfect adherence (as $r_{\text{on}}\tau$ and $r_{\text{off}}\tau$ range between 0.1 and 10). This analysis thus exposes a very serious limitation to the 80% rule and implies that thresholds for adequate adherence levels must be drug-specific.

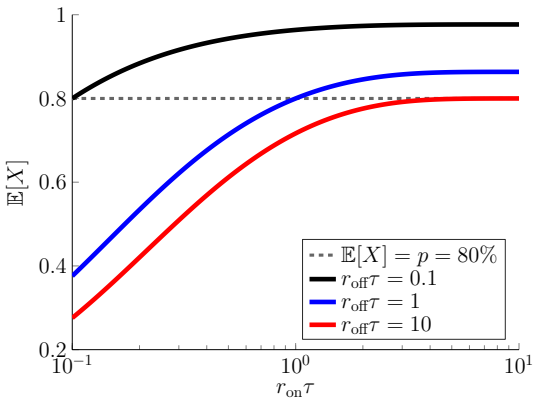

Fig. 1 Average drug effect relative to perfect adherence as a function of the on-rate. We set $p = 0.8$ and $q = 0$

In general, (9) implies that

$$\begin{aligned} \mathbb{E}[X] > p \quad &\text{if } r_{\text{on}} > r_{\text{off}}, \\ \mathbb{E}[X] < p \quad &\text{if } r_{\text{on}} < r_{\text{off}}, \end{aligned} \tag{11}$$

and $\mathbb{E}[X] = p$ if $r_{\text{on}} = r_{\text{off}}$. In words, (11) says that a drug is forgiving of missed doses if its on-rate is faster than its off-rate.

In Fig. 2a, we show a contour plot of $\mathbb{E}[X]$ as a function of $r_{\text{on}}\tau$ and $r_{\text{off}}\tau$ for $p = 80\%$. In addition to illustrating that fast on-rates and slow off-rates maximize $\mathbb{E}[X]$, this plot also shows that $\mathbb{E}[X]$ is more sensitive to r_{on} than r_{off} in much of parameter space. For example, notice that if $r_{\text{off}}\tau = 1$ and $r_{\text{on}}\tau$ varies from $r_{\text{on}}\tau = 0.1$ to $r_{\text{on}}\tau = 10$, then $\mathbb{E}[X]$ varies from less than $\mathbb{E}[X] = 40\%$ to more than $\mathbb{E}[X] = 85\%$. In contrast, if $r_{\text{on}}\tau = 1$ and $r_{\text{off}}\tau$ varies from $r_{\text{off}}\tau = 0.1$ to $r_{\text{off}}\tau = 10$, then $\mathbb{E}[X]$ only varies between about $\mathbb{E}[X] = 95\%$ to about $\mathbb{E}[X] = 70\%$. Overall, Figs. 1 and 2a demonstrate that a moderate level of nonadherence can be countered by prescribing medications with high on-rates and low off-rates.

Drug effect variability

In addition to the average drug effect, the variability in drug effect is also clinically important. In the case $q = 0$ (i.e. no correlation between missed doses), using (7)–(8) yields that the standard deviation of the drug effect (relative to perfect adherence) is given explicitly by

$$\text{SD}(X) = \sqrt{\frac{(1-\alpha)^2(1-\beta)^2(1-p)p}{(\alpha - p\alpha + \beta p - 1)^2(1 - \alpha^2(1-p) - \beta^2 p)}}, \tag{12}$$

where $\alpha = e^{-r_{\text{off}}\tau}$ and $\beta = e^{-r_{\text{on}}\tau}$. In Fig. 2b, we show a contour plot of the standard deviation as a function of $r_{\text{on}}\tau$ and $r_{\text{off}}\tau$ for $p = 80\%$. In this plot, the smallest standard deviation (i.e. lowest variability in drug effect) occurs for fast on-rates and slow off-rates.

Furthermore, Fig. 2b shows that $\text{SD}(X)$ is more sensitive to r_{off} than r_{on} in much of parameter space (in contrast to $\mathbb{E}[X]$ which tends to be more sensitive to r_{on} than r_{off}). For example, notice that if $r_{\text{on}}\tau = 1$ and $r_{\text{off}}\tau$ varies from $r_{\text{off}}\tau = 0.1$ to $r_{\text{off}}\tau = 10$, then $\text{SD}(X)$ varies from less than $\text{SD}(X) = 5\%$ to $\mathbb{E}[X] = 38\%$. In contrast, if $r_{\text{off}}\tau = 1$ and $r_{\text{on}}\tau$ varies from $r_{\text{on}}\tau = 0.1$ to $r_{\text{on}}\tau = 10$, then $\text{SD}(X)$ only varies between $\text{SD}(X) = 21\%$ to $\text{SD}(X) = 28\%$

Application to antihypertensive medications

We now use our model to study the effects of nonadherence on medications for hypertension. In 2011, Lowy et al.

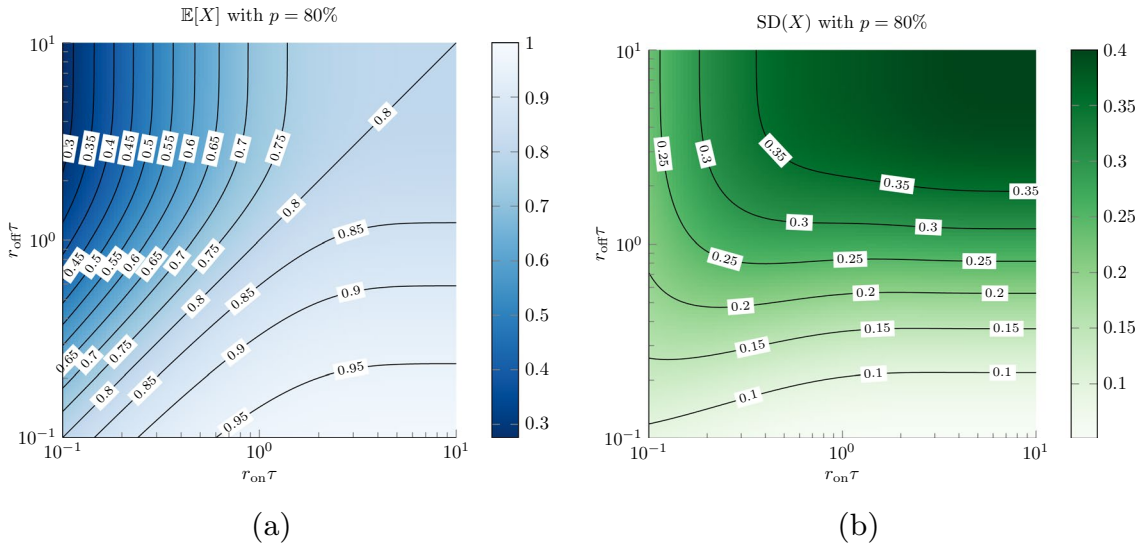

Fig. 2 Average drug effect (**a**) and standard deviation in drug effect (**b**) relative to perfect adherence for $p = 0.8$ and $q = 0$

Table 1 Antihypertensive medications dosed once per day and corresponding off-rates [42–47]

Medication	Effect remaining after missed dose	r_{off} (1/day)
Amlodipine 5-10 mg [47]	95%	0.05
Irbesartan 300 mg [44]	73%	0.32
Ramipril 10 mg [44]	64%	0.45
Losartan 100 mg [45]	59%	0.53
Atenolol 100 mg [46]	21%	1.6
Enalapril 20 mg [43]	10%	2.3
Quinapril 20 mg [42]	0%	> 5.3

[5] reviewed several clinical studies [42–47] to analyze the effects of nonadherence on seven different antihypertensive medications. Table 1 lists these antihypertensive medications, their estimated percentage of full effect remaining after a single missed dose (from [42–47]), and the corresponding off-rates for our model. Specifically, these off-rates were chosen so that $e^{-r_{\mathrm{off}}\tau}$ equals the percentage of full effect remaining after a missed dose, where $\tau = 1$ day is the dosing interval.

To estimate the on-rates for these antihypertensive medications, Lowy et al. [5] assumed that the medications yield their full effect of 15 mmHg decrease in systolic blood pressure after a dose is taken on 3 consecutive days. To approximate this assumption, we set the on-rate in our model to be $r_{\mathrm{on}} = 0.69/\mathrm{day}$ so that $e^{-r_{\mathrm{on}}\tau} = 0.5$ and the medications yield 87.5% of their full effect after a dose is taken on 3 consecutive days.

Figure 3 shows contour plots of the average drug effect relative to perfect adherence (left panel) and the standard deviation of the drug effect relative to perfect adherence (right panel). These plots are functions of the adherence percentage p and the off-rate r_{off}. The horizontal dashed lines correspond to the estimated off-rates of the medications in Table 1. We set $q = 0$ in this figure, though one could readily perform similar analysis using nonzero correlations using (7)–(8).

We emphasize three important features of Fig. 3a. First, this plot shows how the different off-rates of these seven medications (horizontal dashed lines) yield vastly different drug effects if the patient has imperfect adherence (i.e. if $p < 100\%$). For example, our model predicts that a patient taking $p = 80\%$ of their doses would receive 98% of the full drug effect if they are taking Amlodipine, whereas they would receive less than 72% of the full drug effect if they are taking Atenlol, Enalapril, or Quinapril. If we assume full drug effect is a 15 mmHg decrease in systolic blood pressure, then this translates to a 14.7 mmHg decrease for Amlodipine and less than a 10.8 mmHg decrease for Atenlol, Enalapril, or Quinapril. This plot also shows that a patient on Quinapril would have to take more than $p = 98\%$ of their doses to match the average drug effect of a patient on Amlodipine taking only $p = 70\%$ of their doses. In addition to higher average drug effects, Fig. 3b shows how the medications with slow off-rates also yield markedly less variable drug effects for patients with imperfect adherence.

Second, we emphasize that the drastic differences in drug effect between these different antihypertensive medications is only evident when the patient adherence dips below $p = 100\%$. Hence, since adherence is generally thought to be higher in clinical trials than in clinical practice [5, 48], it may be the case that two drugs show similar efficacy in

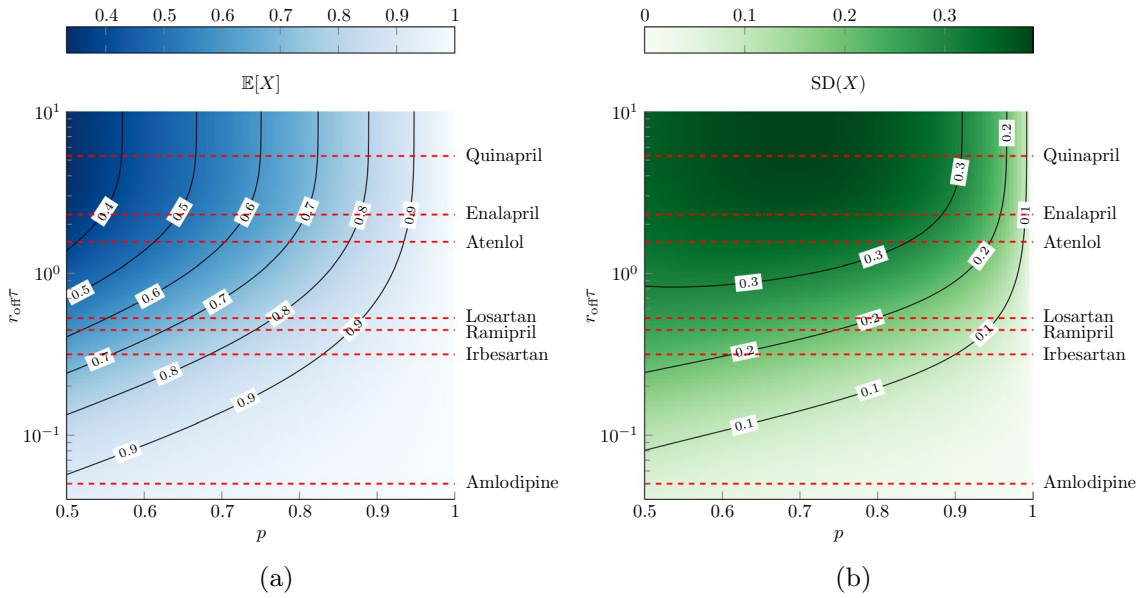

Fig. 3 Average drug effect (**a**) and standard deviation in drug effect (**b**) relative to perfect adherence for $r_{\text{off}}\tau = 0.69$ and $q = 0$

clinical trials but yield quite different effects in clinical practice when the adherence percentage drops.

Third, Fig. 3a is very similar to Fig. 3 in [5] which was obtained therein by many computer simulations of a different model. A key difference between the computational model in [5] and our model is that the model in [5] assumed that following a missed dose, the drug effect decreases linearly, whereas we assumed that the drug effect decreases exponentially. The close agreement between Fig. 3a in the present work and Fig. 3 in [5] shows that the precise nature of the decline in drug effect following a missed dose (i.e. linear or exponential) is not critical when quantifying the effects of nonadherence. A benefit of our model is that we were able to obtain an exact analytical formula for the average drug effect (see (7)). Furthermore, our analysis reveals the variability in drug effect that stems from nonadherence (i.e. through the exact formula for the standard deviation plotted in Fig. 3b), and such variability was not reported for the model in [5].

Time in therapeutic window

One metric of clinical interest is the expected proportion of time the medication effect will be above some minimal threshold for effective treatment. This metric has been called "the most flexible and clinically most meaningful measure of noncompliance" [25]. To express this metric in our model, suppose that the relative biomarker response level X needs to be above some threshold θ for the treatment to be effective. The percentage of time that $X > \theta$ is then given by the probability $\mathbb{P}(X > \theta)$. While the exact distribution function of X does not have a closed form expression, it can be approximated by the Beta distribution derived in the "Beta distribution formulas" section. Specifically, in terms of the parameters a and b of the Beta distribution given in (27) (which are explicit functions of p, $r_{\text{on}}\tau$, $r_{\text{off}}\tau$, and q via (7)–(8)), the percentage of time that the drug effect is above some threshold θ can be approximated by

$$\begin{aligned}\mathbb{P}(X > \theta) &\approx \int_\theta^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\,\mathrm{d}x \\ &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} - B_\theta(a,b),\end{aligned} \tag{13}$$

where $\Gamma(z)$ denotes the Gamma function and $B_z(x,y)$ denotes the incomplete Beta function (both of which have implementations in most computational software programs). This approximation is very accurate for slow on/off rates (i.e. if $r_{\text{on}}\tau$ and/or $r_{\text{off}}\tau$ are small). The accuracy decreases for fast on/off rates (i.e. if $r_{\text{on}}\tau$ and/or $r_{\text{off}}\tau$ are large) since the distribution of X becomes highly irregular in such a parameter regime.

In Fig. 4a, we plot the approximation in (13) as solid curves and the distribution of X computed numerically as dashed curves (details of the numerical calculation are in the Appendix) for three of the antihypertensive medications considered in the "Application to antihypertensive medications" section above (we set $p = 80\%$ and $q = 0$). This plot illustrates that the approximation (13) is very accurate for Amlodipine, as the red solid and dashed curves are nearly indistinguishable. This reflects the slow off-rate and on-rate of Amlodipine ($r_{\text{off}}\tau = 0.05$

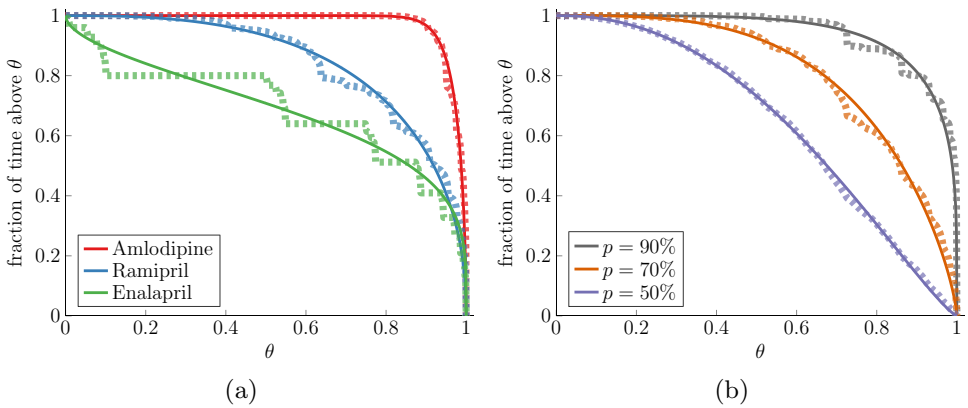

Fig. 4 Fraction of time the relative drug effect is above a threshold θ. The solid curves are the approximation in (13) and the dashed curves are the numerically computed values of $\mathbb{P}(X > \theta)$. **a** Different drugs for adherence $p = 80\%$. **b** Irbesartan with different values of the adherence p

and $r_{\text{on}}\tau = 0.69$). The approximation is less accurate for Enalapril since Enalapril has a faster off-rate ($r_{\text{off}}\tau = 2.3$), but even in this case the approximation still captures the general behavior for Enalapril.

In Fig. 4b, we plot the approximation in (13) (solid curves) and the distribution of X computed numerically (dashed curves) for Irbesartan for different values of the adherence percentage p (we set $q = 0$). This plot shows that the approximation (13) is quite accurate, which reflects the relatively slow off-rate and on-rate of Irbesartan ($r_{\text{off}}\tau = 0.32$ and $r_{\text{on}}\tau = 0.69$).

On/off model approximates complex PK/PD models

We were able to readily apply our model to seven antihypertensive medications in the "Application to antihypertensive medications" and "Time in therapeutic window" sections above since the effective on-rates and off-rates for these drugs have been previously estimated [5, 42–47]. However, effective on-rates and/or off-rates are not necessarily known for many drugs. Indeed, drugs are typically modeled with much more complex PK/PD models consisting of systems of differential equations involving many parameters. In this section, we show that some complex PK/PD models can be closely approximated by our simple on/off model.

Suppose the PD effect of some drug can be modeled as a function $R(t)$ with baseline value R_0. Define $E(t) = \pm(R_0 - R(t))$, where the sign is chosen so that $E(t)$ increases as drug effect increases and $E(t) = 0$ for no drug effect. For instance, in an inhibition of production model $R(t)$ decreases as drug effect increases, so $E(t) = R_0 - R(t)$ ensures $E(t) = 0$ at no drug effect and $E(t)$ increases as drug effect increases. For a patient with perfect adherence ($p = 100\%$), let $E_{\text{avg}}^{\text{perf}}$ denote the average drug effect at steady state over a single dosing interval τ. Consider the following metric,

$$E_n = \left(\frac{1}{\tau E_{\text{avg}}^{\text{perf}}}\right) \int_{(n-1)\tau}^{n\tau} E(t)\,\mathrm{d}t, \tag{14}$$

which is the average drug effect over a single dosing interval relative to perfect adherence (note that $E_n \in [0, 1]$ due to the scaling by $E_{\text{avg}}^{\text{perf}}$). Hence, E_n is the equivalent in a full PK/PD model to the variable X_n in our on/off model. Ideally, if the PK and PD of some medication can be described in terms of an effective on-rate and off-rate, then the effects of nonadherence can be more easily analyzed with our on/off model, rather than relying on computer simulations of the full PK/PD model.

We now demonstrate in two case studies that complex PK/PD models can indeed be closely approximated with our simple on/off model. Each of these two PK/PD models consists of systems of nonlinear differential equations, and these two models involve respectively 9 and 6 nontrivial parameters.

Case Study 1: Warfarin

Warfarin is an anticoagulant commonly prescribed to treat blood clots and prevent strokes [49]. Warfarin is often used as a case study in PK/PD modeling, with industry-standard software such as Monolix [50] having multiple tutorials involving Warfarin data [51, 52]. Furthermore, determining the effects of nonadherence on the efficacy of Warfarin treatment is a matter of great importance [53–56], as missing doses leads to a greater risk of strokes.

Based on output from Monolix, the PK of Warfarin can be satisfactorily modeled by a one compartment model with first-order absorption (at rate k_{a}) and linear elimination (at rate k_{e}). The drug concentration is given by the value $C(t)/V$, where V is the apparent volume of distribution and the drug amount $C(t)$ is computed from the following differential equations,

$$\begin{aligned}\frac{dA}{dt} &= -k_a A(t) + I(t)\\ \frac{dC}{dt} &= k_a A(t) - k_e C(t).\end{aligned} \tag{15}$$

The function $I(t)$ represents the drug input, which is

$$I(t) = D\sum_{n\in\mathbb{Z}} \delta(t - n\tau - T_{lag})\xi_n,$$

where T_{lag} is a lag time, τ is the dosing interval, $\{\xi_n\}_{n\in\mathbb{Z}}$ is a sequence of Bernoulli random variables describing dosing history, and $\delta(\cdot)$ is the Dirac delta distribution. The PD of Warfarin can be modeled by an indirect turnover model with inhibition of production, described by the nonlinear differential equation

$$\frac{dR}{dt} = k_{in}R_0\left(1 - \frac{I_{max}(C(t)/V)^\gamma}{IC_{50}^\gamma + (C(t)/V)^\gamma}\right) - k_{out}R, \tag{16}$$

where k_{in} and k_{out} are the in- and out-rates, R_0 is the baseline value, and γ is the Hill coefficient. For simplicity, we assume $k_{in} = k_{out}$. The response value R is the percent reduction in prothrombin complex activity; that is, $R_0 = 100$ and a value of $R(t) = 80$ means the prothrombin complex activity is 80% of full activity. The parameter estimates from Monolix are in Table 2.

In Fig. 5, we plot the average drug effect and drug affect standard deviation (both relative to perfect adherence as in (14)) as a function of the adherence percentage p. The square markers are computed from the full PK/PD model and the solid curves are our analytical formulas in (9) and (12) for our on/off model with

$$r_{on} = 0.016\,\text{/hour}, \quad r_{off} = 0.009\,\text{/hour}, \tag{17}$$

with dosing interval $\tau = 24$ hours so that

$$r_{on}\tau = 0.38, \quad r_{off}\tau = 0.22.$$

For the full PK/PD model, each square marker in Fig. 5 was computed from 1000 independent stochastic simulations of 7 mg dosed every $\tau = 24$ hours, where each simulation was for 250 days. The values of r_{on} and r_{off} in (17) were chosen to fit the simulation data from the full PK/PD model.

Table 2 Warfarin and Metformin model parameters

	Warfarin [50]	Metformin [58]
T_{lag}	0.93 h	0
k_a	1.6 /hour	2.15 /hour
k_e	0.0165 /hour	0.1219 /hour
V	7.89 L	6480 L
I_{max}	0.9	1
IC_{50}	1.15 mg/L	4.23 mg/L
k_{out}	0.06 /hour	0.8 /hour
R_0	100	2410 mg/L
γ	1.74	1

The close agreement between the square markers and the solid curves in Fig. 5 demonstrate that the on/off model can accurately approximate a full PK/PD model of Warfarin. Since (i) the full PK/PD model is a system of nonlinear differential equations with 9 parameters and must be studied computationally and (ii) the on/off model has only 2 parameters and allows analytical investigation, this is a major simplification.

This analysis predicts that Warfarin is a relatively forgiving drug in the sense that its on-rate is about three times greater than its off-rate (see (17)). Indeed, Fig. 5a predicts that a patient taking only $p = 75\%$ of their doses would still reap nearly 90% of the full drug effect for perfect adherence.

Case Study 2: Metformin

Metformin is a medication used to treat Type 2 diabetes [57] whose PK and PD (with regards to reduction of fasting plasma glucose) can be described by the same differential equation systems as Warfarin given in (15) and (16) [58, 59]. The parameter estimates from [58] are in Table 2. As with Warfarin, medication nonadherence to Metformin is an area

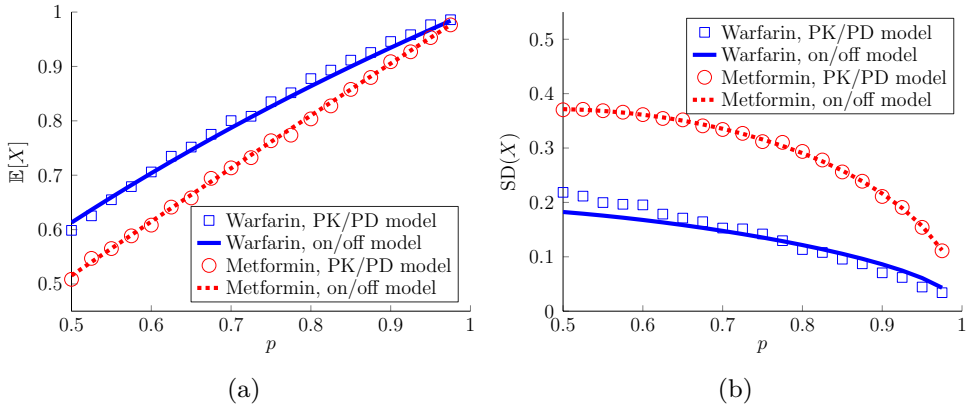

Fig. 5 Comparison of full PK/PD models of Warfarin (blue) and Metformin (red) to the simple on/off model

of great interest [60–62], with one meta-analysis concluding that patients taking Metformin have among the lowest adherence rates within Type 2 diabetes patients [63], likely due to the side effects.

Similar to Warfarin, in Fig. 5 we plot the average drug effect and drug affect standard deviation as a function of p. The circle markers are computed from the full PK/PD model and the dashed curves are our analytical formulas in (9) and (12) for our on/off model with

$$r_{\text{on}} = 0.110\,\text{/hour}, \quad r_{\text{off}} = 0.098\,\text{/hour}, \tag{18}$$

with dosing interval $\tau = 12$ hours so that

$$r_{\text{on}}\tau = 1.32, \quad r_{\text{off}}\tau = 1.18.$$

For the full PK/PD model, each circle marker in Fig. 5 was computed from 1000 independent stochastic simulations of 500 mg dosed every $\tau = 12$ hours, where each simulation was for 250 days. The values of r_{on} and r_{off} in (18) were chosen to fit the simulation data from the full PK/PD model. The same values of r_{on} and r_{off} in (18) were also found to fit simulations of the full PK/PD model with 250 mg dosed every $\tau = 6$ hours.

As in the case of Warfarin, the close agreement between the circle markers and the dashed curves in Fig. 5 demonstrate that the on/off model can accurately approximate a full PK/PD model of Metformin, which is again a major simplification. This analysis shows that Metformin is not especially forgiving or unforgiving of nonadherence in the sense that $r_{\text{on}} \approx r_{\text{off}}$ in (18), and thus the average drug effect $\mathbb{E}[X]$ is approximately the adherence percentage p.

Correlations in missed doses

The analysis above assumed $q = 0$, which means that there are no correlations in missed doses (i.e. after a missing a dose, the patient is no more or less likely to take the next scheduled dose). However, our formulas in (7)–(8) allow us to readily investigate how such correlations affect both the average drug effect and the drug effect variability. In particular, analyzing these formulas show that the general qualitative results above still hold in the case of correlations between missed doses. Indeed, (10) still holds if $q \neq 0$, which means that the average drug effect for an imperfectly adherent patient increases when either the patient's adherence level p increases, the drug on-rate r_{on} increases, or the drug off-rate r_{off} decreases.

Interestingly, including correlations can either increase or decrease the average drug effect compared to the the average drug effect formula in (9) for $q = 0$ depending on the values of r_{on} and r_{off} and whether the correlations are positive ($q > 0$) or negative ($q < 0$). Specifically, the average drug effect with correlations will be higher than the average drug effect without correlations if and only if

$$qr_{\text{off}} > qr_{\text{on}}.$$

That is, a model that does not consider correlations between missed doses will understate the true average drug effect if either (i) the correlation between missed doses is positive (a patient who misses a dose is more likely to miss the next dose as well) and $r_{\text{off}} > r_{\text{on}}$, or (ii) the correlation between missed doses is negative and $r_{\text{on}} > r_{\text{off}}$. In the special case $r_{\text{on}} = r_{\text{off}}$, we have $\mathbb{E}[X] = p$ regardless of the value of q.

Furthermore, we find that for most parameter values, positive correlations tend to increase the drug effect variability, whereas negative correlations tend to decrease drug effect variability. To understand this intuitively, notice that increasing the correlation q means that a patient is more likely to both (i) miss consecutive doses which allows the drug effect to drop lower than if these missed doses were interspersed between taken doses and (ii) take consecutive doses which allows the drug effect to rise higher than if these taken doses were punctuated by missed doses.

It is worth pointing out that the importance of correlations in patient adherence data is uncertain. For example, Vrijens et al. [48] noted a high prevalence of so-called drug "holidays" (three or more consecutive days of missed doses), which implies positive correlations. On the other hand, Sun et al. [64] carried out a careful statistical analysis of the adherence data in [65] and concluded that only one third of the patients studied showed sufficient evidence to reject the hypothesis of no correlations in missed doses.

Population PK/PD

In the analysis above, we considered fixed parameters so that the results can be considered as describing treatment efficacy for a single, 'average' patient. Due to the analytical formulas we obtained in (7)–(8) (9), and (12), it is straightforward to extend our on/off model to a patient population whose parameters vary per some given probability distribution. This allows one to quickly investigate how nonadherence affects treatment efficacy across a group of patients. We now briefly demonstrate how to make this extension and investigate how between subject variability modifies our results.

Consider a population of N patients, indexed by $i = 1, 2, \ldots, N$. Suppose that in this population, the patients' on- and off-rates are log-normally distributed (though this analysis can be performed in a similar fashion for any other choice of distribution). The on- and off-rates for the i-th patient thus are given by

$$r_{\text{on}}^{(i)} = \frac{\overline{r_{\text{on}}}}{\exp(\sigma_{\text{on}}^2/2)} \exp\left(\sigma_{\text{on}} Z_{\text{on}}^{(i)}\right),$$
$$r_{\text{off}}^{(i)} = \frac{\overline{r_{\text{off}}}}{\exp(\sigma_{\text{off}}^2/2)} \exp\left(\sigma_{\text{off}} Z_{\text{off}}^{(i)}\right), \tag{19}$$

where $\overline{r_{\text{on}}}$ and $\overline{r_{\text{off}}}$ are the population typical values for the on- and off-rates, σ_{on} and σ_{off} are positive parameters describing population variability, and $Z_{\text{on}}^{(i)}$ and $Z_{\text{off}}^{(i)}$ are standard normal random variables. For this population, determining the effects of nonadherence on treatment efficacy becomes merely a matter of computing the derived metrics using the parameters $r_{\text{on}}^{(i)}$ and $r_{\text{off}}^{(i)}$.

Figure 6 plots the probability density functions of the mean drug effect $\mathbb{E}[X]$ and the drug effect standard deviation $\text{SD}(X)$ for a medication with population typical values of $\overline{r_{\text{on}}}\tau = 0.69$ and $\overline{r_{\text{off}}}\tau = 0.53$. The three curves correspond to coefficients of variation for $r_{\text{on}}^{(i)}$ and $r_{\text{off}}^{(i)}$ equal to 10%, 20%, and 30%. We take $Z_{\text{on}}^{(i)}$ and $Z_{\text{off}}^{(i)}$ to be independent in this figure, though extending to dependent values is straightforward. Since for each patient in the population the mean and standard deviation of the drug effect with nonadherence can be nearly instantaneously computed from our derived formulas, determining the proportion of the population that is expected to have drug efficacy below or above a certain threshold when nonadherence is present can be done easily and without the need for computationally expensive numerical simulations.

Discussion

It is difficult to improve patient adherence. One way to compensate for nonadherence is to specifically formulate drugs to mitigate its effects. Our model identifies and quantifies two ways to improve average drug efficacy for patients with imperfect adherence (outside of directly improving patient adherence). First, increasing the on-rate ensures that drug effect loss from missed doses can be quickly recovered. Second, decreasing the off-rate ensures that when the patient misses doses, the drug effect persists.

Naturally, our mathematical model made a number of simplifying assumptions. For example, we assumed that the patient takes their medication at only the scheduled dosing times. In reality, patients do not always take their medications at exactly the same time intervals (e.g. if a patient is supposed to take medication daily, they may not necessarily take the medication at exactly the same time every day). As another limitation, while we allowed for correlations in missed doses via the correlation coefficient q, patient behavior is surely a very complex process that cannot be encapsulated merely in a few statistical parameters. Additionally, we assumed that the drug efficacy does not change over time; that is, that r_{on} and r_{off} are constant, regardless of the current biomarker response level, drug concentration, or length of treatment. Tachyphylaxis, which occurs when a drug rapidly loses efficacy after repeated administration, is not considered in this model and could be an issue for patients with relatively high adherence. Furthermore, the biomarker response level X as defined in our model is bounded between 0 (no effect) and 1 (desired full effect), and the possibility of the biomarker response rising to unsafe levels is not considered in this model since extra dosing and overdosing are not considered. Indeed, we assumed that the dose size is constant over time. However, patients are sometimes recommended to take a double dose after missing a dose [14, 38, 66, 67]. Higher (or lower) dosing levels may affect a drug's on-rate and length of time before the biomarker response level begins to decrease, which in turn may affect the guidelines presented.

Despite these limitations, our simple on/off model provides an analytical framework for understanding, quantifying, and mitigating the deleterious effects of medication nonadherence. Our analytical formulas for the (i) average drug effect, (ii) standard deviation of drug effect, and (iii) fraction of time the drug effect is above a threshold allow one to study how treatment efficacy depends on drug parameters (effective on-rate r_{on} and off-rate r_{off}) and adherence

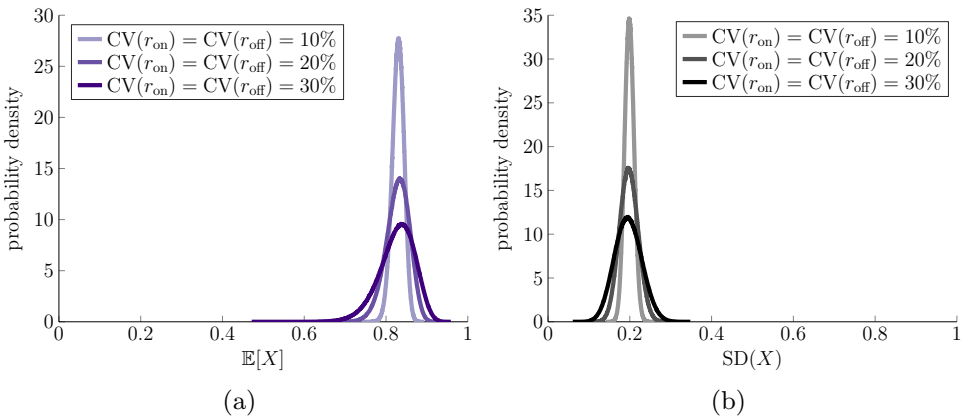

Fig. 6 Distributions of means and standard deviation of drug effect across a population of patients for a medication with $\overline{r_{\text{on}}}\tau = 0.69$ and $\overline{r_{\text{off}}}\tau = 0.53$ and $p = 80\%$

statistics (fraction of doses taken p and correlation q). Furthermore, we have shown that complex PK/PD models can indeed be well-described by our simple on/off model. Finally, we have created an online app to allow pharmacometricians to explore the implications of our model and analysis [26].

Appendix

In this Appendix, we collect details of the mathematical analysis.

Moments of drug effect

We now compute the first and second moments in (7)–(8). First, for the transition probabilities in (4)–(5), the transition matrix is

$$P = \begin{pmatrix} p_0 & 1-p_0 \\ 1-p_1 & p_1 \end{pmatrix}, \tag{20}$$

and the stationary distribution is

$$\pi = \begin{pmatrix} \dfrac{1-p_1}{2-p_0-p_1} & \dfrac{1-p_0}{2-p_0-p_1} \end{pmatrix}.$$

Hence, the proportion of doses taken is

$$p = \pi(1) = \frac{1-p_0}{2-p_0-p_1}.$$

Note that the first moment of X is given by the sum

$$\mathbb{E}[X_0] = \mathbb{E}[X_0 \mathbb{1}_{\xi_{-1}=0}] + \mathbb{E}[X_0 \mathbb{1}_{\xi_{-1}=1}],$$

where $\mathbb{1}_E \in \{0,1\}$ denotes the indicator function, which takes on a value of 1 if the event E occurs and 0 otherwise. At steady state (i.e. the patient has been taking the medication for sufficiently long at adherence rate p), we have

$$X \overset{\mathrm{d}}{=} \alpha X + (\beta-\alpha) X\xi + (1-\beta)\xi, \tag{21}$$

where $X \overset{\mathrm{d}}{=} Y$ denotes equality in distribution. Therefore,

$$X_0 \overset{\mathrm{d}}{=} X_1 = \alpha X_0 + (\beta-\alpha) X_0 \xi_1 + (1-\beta)\xi_1,$$

which means that

$$X_0 \mathbb{1}_{\xi_0=j} \overset{\mathrm{d}}{=} X_1 \mathbb{1}_{\xi_1=j} = \big(\alpha X_0 + (\beta-\alpha) X_0 \xi_1 + (1-\beta)\xi_1\big)\mathbb{1}_{\xi_1=j}.$$

Taking the expected value, we have

$$\begin{aligned} \mathbb{E}[X_0 \mathbb{1}_{\xi_0=j}] &= \mathbb{E}[X_1 \mathbb{1}_{\xi_1=j}] \\ &= \begin{cases} \mathbb{E}[\alpha X_0 \mathbb{1}_{\xi_1=0}], & j=0, \\ \mathbb{E}[\big(\beta X_0 + (1-\beta)\big)\mathbb{1}_{\xi_1=1}], & j=1, \end{cases} \\ &= \begin{cases} \alpha \mathbb{E}[X_0 \mathbb{1}_{\xi_1=0}], & j=0, \\ (1-\beta)\pi(1) + \beta \mathbb{E}[X_0 \mathbb{1}_{\xi_1=1}], & j=1. \end{cases} \end{aligned} \tag{22}$$

where $\pi(1) = p$. Note that

$$\begin{aligned} \mathbb{E}[X_0 \mathbb{1}_{\xi_1=j}] &= \sum_{i=0}^{1} \mathbb{E}[X_0 \mathbb{1}_{\xi_1=j} \mathbb{1}_{\xi_0=i}] \\ &= \sum_{i=0}^{1} \mathbb{E}[\mathbb{E}[X_0 \mathbb{1}_{\xi_1=j} \mathbb{1}_{\xi_0=i} \mid \{\xi_n\}_{n\le -1}]], \end{aligned} \tag{23}$$

where the second equality comes from the tower property of conditional expectation. The term within the sum can be simplified. For a function $f(X_0, j)$,

$$\begin{aligned} &\mathbb{E}[\mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_1=j}\mathbb{1}_{\xi_0=i} \mid \{\xi_n\}_{n\le -1}]] \\ &\quad = \mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_0=i}\mathbb{E}[\mathbb{1}_{\xi_1=j} \mid \{\xi_n\}_{n\le -1}]] \\ &\quad = \mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_0=i} P_{ij}] \\ &\quad = P_{ij}\mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_0=i}], \end{aligned} \tag{24}$$

where P_{ij} is the ij-th entry of the transition matrix (20). This means that

$$\begin{aligned} \mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_1=j}] &= \sum_{i=0}^{1} P_{ij}\mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_0=i}] \\ &= P_{0j}\mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_0=0}] \\ &\quad + P_{1j}\mathbb{E}[f(X_0,j)\mathbb{1}_{\xi_0=1}]. \end{aligned}$$

Combining this result with (22) we have

$$\begin{aligned} \mathbb{E}[X_0 \mathbb{1}_{\xi_0=0}] &= \mathbb{E}[X_1 \mathbb{1}_{\xi_1=0}] = \alpha \mathbb{E}[X_0 \mathbb{1}_{\xi_1=0}] \\ &= \alpha\big(P_{00}\mathbb{E}[X_0 \mathbb{1}_{\xi_0=0}] + P_{10}\mathbb{E}[X_0 \mathbb{1}_{\xi_0=1}]\big), \\ \mathbb{E}[X_0 \mathbb{1}_{\xi_0=1}] &= \mathbb{E}[X_1 \mathbb{1}_{\xi_1=1}] = (1-\beta)\pi(1) + \beta \mathbb{E}[X_0 \mathbb{1}_{\xi_1=1}] \\ &= (1-\beta)\pi(1) + \beta\big(P_{01}\mathbb{E}[X_0 \mathbb{1}_{\xi_0=0}] \\ &\quad + P_{11}\mathbb{E}[X_0 \mathbb{1}_{\xi_0=1}]\big). \end{aligned}$$

These lead to the system of equations

$$\begin{aligned} \mathbb{E}[X_0 \mathbb{1}_{\xi_0=0}] &= \frac{\alpha(1-p_1)}{1-\alpha p_0}\mathbb{E}[X_0 \mathbb{1}_{\xi_0=1}], \\ \mathbb{E}[X_0 \mathbb{1}_{\xi_0=1}] &= \frac{(1-\beta)\pi(1)}{1-\beta p_1} + \frac{\beta(1-p_0)}{1-\beta p_1}\mathbb{E}[X_0 \mathbb{1}_{\xi_0=0}]. \end{aligned}$$

The solution to this system is

$$\mathbb{E}[X_0\mathbb{1}_{\xi_0=0}] = \frac{\pi(1)\alpha(1-\beta)(1-p_1)}{(1-\beta p_1)(1-\alpha p_0)-\alpha\beta(1-p_0)(1-p_1)},$$
$$\mathbb{E}[X_0\mathbb{1}_{\xi_0=1}] = \frac{\pi(1)(1-\beta)(1-\alpha p_0)}{(1-\beta p_1)(1-\alpha p_0)-\alpha\beta(1-p_0)(1-p_1)}.$$

Thus,

$$\begin{aligned}\mathbb{E}[X_0] &= \mathbb{E}[X_0\mathbb{1}_{\xi_0=0}] + \mathbb{E}[X_0\mathbb{1}_{\xi_0=1}]\\ &= \frac{\pi(1)(1-\beta)\big(\alpha(1-p_1)+(1-\alpha p_0)\big)}{(1-\beta p_1)(1-\alpha p_0)-\alpha\beta(1-p_0)(1-p_1)}.\end{aligned}$$

The second moment can be computed in a similar fashion. Note that

$$X_0^2 \overset{\mathrm{d}}{=} X_1^2 = \big(\alpha X_0 + (\beta-\alpha)X_0\xi_1 + (1-\beta)\xi_1\big)^2,$$

and thus

$$\begin{aligned}\mathbb{E}[X_0^2\mathbb{1}_{\xi_0=j}] &= \mathbb{E}[X_1^2\mathbb{1}_{\xi_1=j}]\\ &= \begin{cases}\mathbb{E}[\alpha^2X_0^2\mathbb{1}_{\xi_1=0}], & j=0,\\ \mathbb{E}[\big(\beta X_0+(1-\beta)\big)^2\mathbb{1}_{\xi_1=1}], & j=1,\end{cases}\end{aligned} \tag{25}$$

The results in (23) and (24) can be applied to $\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=j}]$, which gives

$$\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=j}] = P_{0j}\mathbb{E}[X_0^2\mathbb{1}_{\xi_0=0}] + P_{1j}\mathbb{E}[X_0^2\mathbb{1}_{\xi_0=1}]$$

Thus,

$$\begin{aligned}\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=0}] &= \mathbb{E}[X_1^2\mathbb{1}_{\xi_2=0}] = \alpha^2\mathbb{E}[X_0^2\mathbb{1}_{\xi_2=0}]\\ &= \alpha^2\big(P_{00}\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=0}] + P_{10}\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=1}]\big)\\ &= \alpha^2\big(p_0\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=0}] + (1-p_1)\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=1}]\big),\\ \mathbb{E}[X_0^2\mathbb{1}_{\xi_1=1}] &= \mathbb{E}[X_1^2\mathbb{1}_{\xi_2=1}]\\ &= (1-\beta)^2\pi(1) + 2\beta(1-\beta)\mathbb{E}[X_0\mathbb{1}_{\xi_1=1}]\\ &\quad + \beta^2\mathbb{E}[X_0^2\mathbb{1}_{\xi_2=1}]\\ &= (1-\beta)^2\pi(1) + 2\beta(1-\beta)\mathbb{E}[X_0\mathbb{1}_{\xi_1=1}]\\ &\quad + \beta^2\big((1-p_0)\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=0}] + p_1\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=1}]\big).\end{aligned}$$

From this system of equations we have

$$\begin{aligned}&\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=0}]\\ &\quad= \left(\frac{\alpha^2(1-p_1)}{(1-\alpha^2p_0)(1-\beta^2p_1)-\alpha^2\beta^2(1-p_0)(1-p_1)}\right)\\ &\qquad\times\big((1-\beta)^2\pi(1)+2\beta(1-\beta)\mathbb{E}[X_0\mathbb{1}_{\xi_1=1}]\big),\\ &\mathbb{E}[X_0^2\mathbb{1}_{\xi_1=0}]\\ &\quad= \left(\frac{(1-\alpha^2p_0)}{(1-\alpha^2p_0)(1-\beta^2p_1)-\alpha^2\beta^2(1-p_0)(1-p_1)}\right)\\ &\qquad\times\big((1-\beta)^2\pi(1)+2\beta(1-\beta)\mathbb{E}[X_0\mathbb{1}_{\xi_1=1}]\big),\end{aligned}$$

which means that the second moment is

$$\begin{aligned}\mathbb{E}[X_0^2] &= \left(\frac{\alpha^2(1-p_1)+(1-\alpha^2p_0)}{(1-\alpha^2p_0)(1-\beta^2p_1)-\alpha^2\beta^2(1-p_0)(1-p_1)}\right)\\ &\quad\times\big((1-\beta)^2\pi(1)+2\beta(1-\beta)\mathbb{E}[X_0\mathbb{1}_{\xi_1=1}]\big).\end{aligned}$$

Numerical computation of $\mathbb{P}(X > \theta)$

Assuming independence of $\{\xi_n\}_{n\in\mathbb{Z}}$, the distribution function of the biomarker response level can be numerically computed using the recursion relationship (3). By (21), we have that

$$\begin{aligned}\mathbb{P}(X_n \le x) &= \mathbb{P}(\alpha X_n + (\beta-\alpha)X_n\xi_{n+1} + (1-\beta)\xi_{n+1} \le x)\\ &= (1-p)\mathbb{P}(X_n \le x/\alpha) + p\mathbb{P}(X_n \le (x-1)/\beta + 1).\end{aligned}$$

Denoting the distribution function of X by $F(x) = \mathbb{P}(X \le x)$, we thus have

$$F(x) = (1-p)F\left(\frac{x}{\alpha}\right) + pF\left(\frac{x-1}{\beta}+1\right). \tag{26}$$

By starting with an initial guess for the distribution function of X and iteratively applying (26), the true distribution function of X can be numerically determined. Specifically, we obtain a sequence of approximating distribution functions $\{F_m(x)\}_{m\ge0}$ where $F_0(x) = x$ and $F_{m+1}(x)$ is defined by $F_m(x)$ via

$$F_{m+1}(x) = (1-p)F_m\left(\frac{x}{\alpha}\right) + pF_m\left(\frac{x-1}{\beta}+1\right).$$

The dashed curves in Fig. 4 are obtained via this recursion once successive iterations $F_m(x)$ and $F_{m+1}(x)$ differ by less than 10^{-10} for all $x \in [0,1]$.

Beta distribution formulas

In the "Mean and variance of biomarker response" section, we described how we can approximate the full probability distribution of X by a Beta random variable B chosen so that the first and second moments of X and B agree. We now give the explicit formulas for this Beta distribution fit.

The probability density function of a Beta random variable B is given by

$$f_B(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}, \quad \text{if } x \in (0,1),$$

and $f_B(x) = 0$ for $x \notin (0,1)$, where $\Gamma(z) = \int_0^\infty u^{z-1}e^{-u}\,\mathrm{d}u$ denotes the Gamma function. The first and second moments of B are

$$\mathbb{E}[B] = \frac{a}{a+b}, \quad \mathbb{E}[B^2] = \frac{a(a+1)}{(a+b)(a+b+1)}.$$

Therefore, choosing a and b so that $\mathbb{E}[B] = \mathbb{E}[X]$ and $\mathbb{E}[B^2] = \mathbb{E}[X^2]$ implies

$$a = \frac{\mathbb{E}[X](\mathbb{E}[X] - \mathbb{E}[X^2])}{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}, \quad b = \frac{(1 - \mathbb{E}[X])(\mathbb{E}[X] - \mathbb{E}[X^2])}{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}. \tag{27}$$

Acknowledgements This work was supported by the National Science Foundation (Grant Nos. DMS-1944574 and DMS-1814832).

Author contributions EDC and SDL jointly devised and carried out the research and wrote the manuscript.

Declarations

Conflict of interest The authors declare no competing interests.

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. Kim J, Combs K, Downs J, Tillman F III (2018) Medication adherence: the elephant in the room. US Pharmacist 43:30–34
4. De Geest S, Zullig LL, Dunbar-Jacob J, Helmy R, Hughes DA, Wilson IB, Vrijens B (2018) ESPACOMP medication adherence reporting guideline (EMERGE). Ann Intern Med 169(1):30–35
5. 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
6. Stauffer ME, Hutson P, Kaufman AS, Morrison A (2017) The adherence rate threshold is drug specific. Drugs R &D 17(4):645–653
7. Assawasuwannakit P, Braund R, Duffull SB (2015) Quantification of the forgiveness of drugs to imperfect adherence. CPT 4(3):204–211
8. 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
9. Boissel J-P, Nony P (2002) Using pharmacokinetic-pharmacodynamic relationships to predict the effect of poor compliance. Clin Pharmacokinet 41(1):1–6
10. 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
11. 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
12. 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
13. 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
14. 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
15. Gidal BE, Majid O, Ferry J, Hussein Z, Yang H, Zhu J, Fain R, Laurenza A (2014) The practical impact of altered dosing on perampanel plasma concentrations: pharmacokinetic modeling from clinical studies. Epilepsy Behav 35:6–12
16. Brittain ST, Wheless JW (2015) Pharmacokinetic simulations of topiramate plasma concentrations following dosing irregularities with extended-release vs. immediate-release formulations. Epilepsy Behav 52:31–36
17. 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
18. Hard ML, Wehr AY, Sadler BM, Mills RJ, von 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
19. 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
20. 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
21. Gibaldi M, Perrier D (1982) Pharmacokinetics, 2nd edn. Marcelly Dekker, New York
22. Felmlee MA, Morris ME, Mager DE. Mechanism-based pharmacodynamic modeling. In: Computational toxicology, pp 583–600. Springer, Berlin (2012)
23. 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
24. Urquhart J (2002) The odds of the three nons when an aptly prescribed medicine isn't working: non-compliance, non-absorption, non-response. Br J Clin Pharmacol 54(2):212–220
25. Levy G (1993) A pharmacokinetic perspective on medicament noncompliance. Clin Pharmacol Ther 54(3):242–244
26. Web app at `https://seanlawley.shinyapps.io/OnOff/`
27. Norris JR (1998) Markov chains. In: Statistical & probabilistic mathematics. Cambridge University Press, Cambridge
28. Yuval P, Wilhelm S, Boris S (2000) Sixty years of Bernoulli convolutions. In: Bandt C, Graf S, Zähle M (eds) Fractal geometry and stochastics II. Birkhäuser, Basel, pp 39–65
29. Solomyak B (1995) On the random series $\sum \pm\lambda^n$ (an erdos problem). Ann Math 142(3):611–625
30. Escribano C, Sastre MA, Torrano E (2001) Moments of infinite convolutions of symmetric Bernoulli distributions. J Comput Appl Math 153(1):191–199, 2003. In: Proceedings of the 6th international symposium on orthogonal polynomials, special functions and their applications, Rome, Italy, 18–22
31. Tian-You H, Lau K-S (2008) Spectral property of the Bernoulli convolutions. Adv Math 219(2):554–567
32. Kershner R, Wintner A (1935) On symmetric Bernoulli convolutions. Am J Math 57(3):541–548
33. Jessen B, Wintner A (1935) Distribution functions and the Riemann zeta function. Trans Am Math Soc 38(1):48–88
34. Wintner A (1935) On symmetric Bernoulli convolutions. Bull Am Math Soc 41:137–138
35. Erdös P (1939) On a family of symmetric Bernoulli convolutions. Am J Math 61:974

36. 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
37. 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
38. Counterman ED, Lawley SD (2021) What should patients do if they miss a dose of medication? A theoretical approach. J Pharmacokinet Pharmacodyn 48(6):873–892
39. Counterman ED, Lawley SD (2022) Designing drug regimens that mitigate nonadherence. Bull Math Biol 84(1):1–36
40. Burnier M (2019) Is there a threshold for medication adherence? Lessons learnt from electronic monitoring of drug adherence. Front Pharmacol 9:1540
41. Brian HR (1976) Compliance with therapeutic regimens, chapter A critical review of "determinants'' of patient compliance with therapeutic regimens. Johns Hopkins University Press, Baltimore, pp 26–39
42. Radauceanu A, Virion J-M, Boivin J-M, Zannad F (2002) Time-effect profile of antihypertensive agents assessed with trough/peak ratio, smoothness index and dose omission: an ambulatory blood pressure monitoring study with trandolapril vs. quinapril1. Fundam Clin Pharmacol 16(6):545–554
43. Vaur L, Bobrie G, Dutrey-Dupagne C, Dubroca I, Vaisse B, Bouvier d'Yvoire M, Elkik FE, Chatellier G, Menard J (1998) Short-term effects of withdrawing angiotensin converting enzyme inhibitor therapy on home self-measured blood pressure in hypertensive patients. Am J Hypertens 11(2):165–173
44. Palatini P, Jung W, Shlyakhto E, Botha J, Bush C, Keefe DL (2010) Maintenance of blood-pressure-lowering effect following a missed dose of aliskiren, irbesartan or ramipril: results of a randomized, double-blind study. J Hum Hypertens 24(2):93–103
45. Yves Lacourciere and for the Candesartan/Losartan study investigators Asmar, Roland. A comparison of the efficacy and duration of action of candesartan cilexetil and losartan as assessed by clinic and ambulatory blood pressure after a missed dose, in truly hypertensive patients*: a placebo-controlled, forced titration study. Am J Hypertens 12(12):1181–1187 (1999)
46. Johnson BF, Whelton A (1994) A study design for comparing the effects of missing daily doses of antihypertensive drugs. Am J Ther 1(4):260–267
47. Smilde JG (1997) A comparison of amlodipine and felodipine extended release in the treatment of hypertension at steady state and after two missed doses. Curr Ther Res 58(3):141–153
48. 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
49. AHFS DI Essentials. Warfarin. https://www.drugs.com/monograph/warfarin.html. Accessed 8 Feb 2023
50. Lixoft SAS, a Simulations Plus company. Monolix 2021r2. https://lixoft.com/products/monolix/
51. Lixoft SAS, a Simulations Plus company. Case studes. https://monolix.lixoft.com/case-studies/. Accessed 8 Feb 2023
52. O'Reilly RA, Aggeler PM (1968) Studies on coumarin anticoagulant drugs. Circulation 38(1):169–177
53. Waterman AD, Milligan PE, Bayer L, Banet GA, Gatchel SK, Gage BF (2004) Effect of warfarin nonadherence on control of the International Normalized Ratio. Am J Health Syst Pharm 61(12):1258–1264
54. Kimmel SE, Chen Z, Price M, Parker CS, Metlay JP, Christie JD, Brensinger CM, Newcomb CW, Samaha FF, Gross R (2007) The influence of patient adherence on anticoagulation control with warfarin: results from the international normalized ratio adherence and genetics (IN-RANGE) study. Arch Intern Med 167(3):229–235
55. Platt AB, Localio AR, Brensinger CM, Cruess DG, Christie JD, Gross R, Parker CS, Price M, Metlay JP, Cohen A, Newcomb CW, Strom BL, Laskin MS, Kimmel SE (2008) Risk factors for nonadherence to warfarin: results from the in-range study. Pharmacoepidemiol Drug Saf 17(9):853–860
56. Yao X, Abraham NS, Alexander GC, Crown W, Montori VM, Sangaralingham LR, Gersh BJ, Shah ND, Noseworthy PA (2016) Effect of adherence to oral anticoagulants on risk of stroke and major bleeding among patients with atrial fibrillation. J Am Heart Assoc 5(2):e003074
57. AHFS DI Essentials. Metformin. Data retrieved on 8 February 2023 from https://www.drugs.com/monograph/metformin.html
58. 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
59. McAllister N, Lawley S (2022) A pharmacokinetic and pharmacodynamic analysis of drug forgiveness. J Pharmacokinet Pharmacodyn 49:1–17
60. Tseng Y-J, Steinberg G, Fox KP, Armstrong J, Mandl KD (2017) Antihyperglycemic medications: a claims-based estimate of first-line therapy use prior to initialization of second-line medications. Diabetes Care 40(11):1500–1505
61. Parente M, Pignata I, Gnavi R, Spadea T, Di Martino M, Baratta F, Brusa P (2020) Assessment of non-adherence to oral metformin and atorvastatin therapies: a cross-sectional survey in piedmont (Italy). Patient Prefer Adherence 14:261–266
62. Christofides EA (2019) Practical insights into improving adherence to metformin therapy in patients with type 2 diabetes. Clin Diabetes 37(3):234–241
63. McGovern A, Tippu Z, Hinton W, Munro N, Whyte M, de Lusignan S (2018) Comparison of medication adherence and persistence in type 2 diabetes: a systematic review and meta-analysis. Diabetes Obes Metab 20(4):1040–1043
64. 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
65. 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
66. Jonklaas J, Bianco AC, Bauer AJ, Burman KD, Cappola AR, Celi FS, Cooper DS, Kim BW, Peeters RP, Rosenthal MS et al (2014) Guidelines for the treatment of hypothyroidism: prepared by the American thyroid association task force on thyroid hormone replacement. Thyroid 24(12):1670–1751
67. Clark ED, Lawley SD (2022) Should patients skip late doses of medication? A pharmacokinetic perspective. J Pharmacokinet Pharmacodyn 49(4):429–444