Research Article

A Hybrid Markov Chain–von Mises Density Model for the Drug-Dosing Interval and Drug Holiday Distributions

Kelly Fellows,¹ Vivian Rodriguez-Cruz,¹ Jenna Covelli,¹ Alyssa Droopad,¹
Sheril Alexander,¹ and Murali Ramanathan^{1,2}

Received 20 September 2014; accepted 23 December 2014; published online 22 January 2015

ABSTRACT. Lack of adherence is a frequent cause of hospitalizations, but its effects on dosing patterns have not been extensively investigated. The purpose of this work was to critically evaluate a novel pharmacometric model for deriving the relationships of adherence to dosing patterns and the dosing interval distribution. The hybrid, stochastic model combines a Markov chain process with the von Mises distribution. The model was challenged with electronic medication monitoring data from 207 hypertension patients and against 5-year persistence data. The model estimates distributions of dosing runs, drug holidays, and dosing intervals. Drug holidays, which can vary between individuals with the same adherence, were characterized by the patient cooperativity index parameter. The drug holiday and dosing run distributions deviate markedly from normality. The dosing interval distribution exhibits complex patterns of multimodality and can be long-tailed. Dosing patterns are an important but under recognized covariate for explaining within-individual variance in drug concentrations.

KEY WORDS: adherence; circular distribution; compliance; dosing patterns; runs.

INTRODUCTION

Pharmaceutical and medical advances have rendered many previously untreatable diseases, *e.g.*, human immunodeficiency virus (HIV) and cardiovascular disease, manageable as chronic conditions requiring long-term drug therapy. Lack of adherence to drug therapy is a serious and costly public health problem because it is a frequent but potentially preventable cause of hospitalizations and a risk factor for the emergence of drug-resistant pathogens.

In addition to dose omissions and overconsumption (1), several patterns of nonadherence have been reported; these include premature termination, the “drug holiday” (DH), defined functionally as ≥ 3 consecutive days without drug administration (2), and the “toothbrush effect,” wherein adherence is increased just prior to visits to the caregiver.

Adherence is a complex multifactorial phenomenon containing both random and nonrandom, patient-specific factors. Urquhart (3) cautioned that, basing drug decisions solely on adherence data could lead to erroneous conclusions. Because pharmacometric modeling of drug-dosing patterns is challenging, the majority of clinical reports have used

descriptive and exploratory statistical approaches. However, effective adherence models could be useful for pharmacokinetic–pharmacodynamic (PK–PD) modeling and for analyzing data from electronic medication event monitoring systems (MEMS). In PK–PD modeling, Monte Carlo approaches have been used to generate irregular dosing patterns that randomly drop doses in a multiple dosing regimen using, *e.g.*, a Bernoulli process. However, the adequacy of such stochastic modeling methods has neither been systematically evaluated nor challenged with a large real world data set.

Although the drug interval distribution is fundamental for a wide range of pharmacometric analyses and clinical trial simulations, there is a relative lack of rigorous mechanistic methodology that has been brought to bear on the problem. One of the main pharmacometric challenges is that the interval between dosing events contains a superposition of a discrete, periodic time-related variable, the dosing interval, with a continuous time variable, the variations in dose timing within each dosing interval.

It is important to distinguish between overall variability in adherence *versus* the variability in dosing patterns caused by differences in adherence behaviors. Patients with a given level of adherence can exhibit a diverse range of dosing patterns. The differences in dosing patterns result from variations in behavioral responses to missed doses. Our working hypothesis was that dosing patterns resulting from stochastic variability in adherence behaviors rather than adherence *per se* are the critical determinant of variability in drug-dosing intervals, which determine drug concentrations, effects, and outcomes in individuals and populations. This research therefore focuses on developing a theoretically

Electronic supplementary material The online version of this article (doi:10.1208/s12248-014-9713-5) contains supplementary material, which is available to authorized users.

¹Department of Pharmaceutical Sciences and Neurology, State University of New York, 355 Kapoor Hall, Buffalo, New York 14214-8033, USA.

²To whom correspondence should be addressed. (e-mail: Murali@Buffalo.Edu)

rigorous modeling framework and practical data analysis implementation for obtaining the drug-dosing interval distribution from adherence data.

METHODS

Modeling Terminology and Dosing Regimen Representations

We describe the modeling terminology in Supplementary Methods.

Experimental Data Sets

The data used for challenging the model were the observed dosing regimens for 207 patients on once daily hypertension treatments. Data on dose timing, number of dose units taken for each dose, and dosing intervals were extracted from the Pharmion Knowledge Center (PKC) database (<http://www.iadherence.org/www/pkc.adx>), a compilation of electronically monitored patient dosing histories.

Data from the calendar plot were obtained for each patient. A dosing event was defined as a “success” when ≥ 1 dose units were taken on a calendar day, and a calendar day in which no dose unit was taken was defined as a “failure” dosing event. The time interval between overdosing events and the persistence data were also derived from the calendar plots. The dosing interval data were extracted from graphical displays of the time in hours between two consecutive doses. Persistence data from two previous studies, Catalan and LeLorier (4) and Larsen *et al.* (5), were digitally extracted from the paper by Hughes and Walley (6).

Mathematica (Wolfram Research, Champaign, IL, USA) was used for all modeling and stochastic simulations. SPSS (IBM, Armonk, NY, USA) and the R circular statistics package were used for statistical analyses (7,8). A MacBook computer running Mac OS X operating system was used.

Modeling Individual Dosing Profiles

We modeled the short-range dependence of dosing using a two-state Markov chain with the following transition matrix A_i :

$$A_i = \frac{1}{F} \begin{bmatrix} p_{SS}(i) & p_{SF}(i) \\ p_{FS}(i) & p_{FF}(i) \end{bmatrix} = \frac{1}{F} \begin{bmatrix} p_{SS}(i) & 1-p_{SS}(i) \\ p_{FS}(i) & 1-p_{FS}(i) \end{bmatrix}$$

The probability of a success at the i th dosing event following a success at the $(i-1)$ th dosing event was denoted by $p_{SS}(i)$ and the probability of a success at the i th dosing event following a failure at the $(i-1)$ th dosing event was denoted by $p_{FS}(i)$. The matrix A_i is a square transition matrix whose rows sum to unity. Although our analysis algorithms are general, we assume that the transition matrix $A_i = A$ is time homogeneous, *i.e.*, p_{SS} and p_{FS} are constants.

$$A = \frac{1}{F} \begin{bmatrix} p_{SS} & p_{SF} \\ p_{FS} & p_{FF} \end{bmatrix} = \frac{1}{F} \begin{bmatrix} p_{SS} & 1-p_{SS} \\ p_{FS} & 1-p_{FS} \end{bmatrix}$$

Maximum-likelihood estimates (9) of p_{SS} and p_{FS} were obtained from N_{SS} , N_{FS} , N_{FF} , and N_{SF} , the frequencies of success followed by success, failure followed by success,

failure followed by failure and success followed by failure events, respectively, using:

$$\hat{p}_{SS} = \frac{N_{SS}}{N_{SS} + N_{SF}}$$

$$\hat{p}_{FS} = \frac{N_{FS}}{N_{FS} + N_{FF}}$$

The patient cooperativity index (PCI) was defined as:

$$PCI = \left(\frac{p_{FS}}{1-p_{FS}} \right) / \left(\frac{p_{SS}}{1-p_{SS}} \right)$$

The PCI represents a likelihood ratio of the likelihood of a success following a failure to the likelihood of a success following a success. Patients with low PCI values are less likely to take a dose following a lapse in dosing. Patients with a PCI value of unity have $p_{FS} = p_{SS}$, and their dosing patterns are independent of their prior dosing behavior, *i.e.*, their dosing patterns follow a Bernoulli distribution. Patients with high PCI values are more likely to take a dose following an omission.

We assessed adherence changes over time, *i.e.*, the time homogeneity of the Markov chain model, by estimating the adherence over 30-day blocks and plotting the data as a function of time as more formal procedures for challenging the time nonhomogeneity of Markov chains are difficult to construct (9).

Modeling the Dosing Runs Distribution

We extended the model using the theory of runs to compute the distribution of the dosing runs, which is critical to assessing the distributions of DH and dosing intervals. Exact formulae for run statistics require combinatorial probability techniques (10,11) and are surprisingly complicated even for the simplest independent, identically distributed Bernoulli random variables case. However, the Markov chain imbedding technique, developed by Fu and Koutras (12), has allowed run distributions to be derived using intuitive methodology (see Supplementary Data) for identically and non-identically distributed, Markov-dependent trials.

The probability of a run of r or more successes in n trials, $G_{n,r}$ was used for DH calculations. It was calculated using the computationally efficient triangular matrix decomposition method (13). The exact distribution of runs (*e.g.*, for Fig. 1q) was estimated using Monte Carlo methods with 100,000 simulations. Run calculations were implemented in Mathematica. Run distributions were compared using the Fisher exact test.

Modeling the Drug-Dosing Interval Distribution

Drug-Dosing Interval Distribution. The drug-dosing interval distribution contains a discrete contribution from the dosing failure runs distribution derived from the Markov chain model and a continuous contribution due to variability in dose timing. The time interval Δt_i between doses was approximated using:

$$\Delta t_i = \tau(1 + f_i) + \delta_i$$

The term f_i represents the exact number of failures in the run of failures between the i th dosing event and the

Fig. 1. Observed and simulated dosing patterns. **a, e, i, m** Observed number of prescribed dose units consumed at each dosing event for four representative patients (denoted by #808, #813, #820, and #810) with different levels of overall adherence. The y-axis denotes the number of prescribed dose units consumed; the prescribed dose was 1 dose unit (0 denotes that the dose was not taken within the dosing interval). **b, f, j, n** Observed value of indicator variable representing whether or not the dose was taken. **c, g, k, o** Representative simulations based on the two-state Markov model. **d, h, l, p** Comparison of the running sum (or cumulative frequency) of the observed value of the indicator variable (blue symbols) to the simulated value of the indicator variable (green symbols) from the two-state Markov model. The x-axis represents a time index in **a–p**. **q** Comparison between the complete distributions of runs observed in patients #808 (orange filled circles), #813 (red filled circles), #810 (blue filled circles), and #820 (green filled circles) and to those predicted by the two-state Markov model. The solid line is the best-fit line and its Pearson correlation R^2 value is also shown. **r** Thirty-day moving average of adherence in patients #808 (orange filled circles), #810 (red filled circles), #813 (blue filled circles), and #820 (green filled circles)

immediately preceding successful dosing event, and τ represents the prescribed dosing interval. The δ_i is the dose timing deviation, *i.e.*, the time difference between the actual dosing time and the nearest reference dose.

The probability density function (PDF) of the dosing interval is the convolution of the PDF of f_i and δ_i because it is the sum of these two random variables. The PDF of f_i was obtained from the exact distribution of runs.

Dose Timing Deviation Distribution. We assume that the PDF of δ_i after transformation to angular coordinates is distributed according to the von Mises (VM) PDF function $VM(\mu, \kappa)$:

$$2\pi \frac{\delta_i}{\tau} \sim VM(\mu_{SS}, \kappa_{SS})$$

The R circular statistics package (7) was used to obtain maximum likelihood estimates for μ_{SS} and κ_{SS} , mean, and concentration parameters for the VM distribution of dosing events containing a success followed by a success.

von Mises Distribution. The VM distribution describes angular random variables and its PDF $p(\theta)$ at angular position θ radians is:

$$p(\theta) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)} \text{ for } \mu - \pi \leq \theta \leq \mu + \pi$$

The μ is the mean, and κ is a measure of how concentrated the data are around the mean; $1/\kappa$ is analogous to variance; I_0 is the Bessel function of order zero. We will represent the random variate θ from the VM distribution as $RV[\theta, VM(\mu, \kappa)]$.

Modeling Overdosing Events

Model distributions for the extent of overdosing (*i.e.*, the number of additional doses taken) and the interval between overdosing events were assessed visually using probability–probability plots. The approach of the model distribution to observed distribution was assessed using the Kolmogorov–Smirnov (KS) test. Maximum-likelihood estimation was used to calculate model parameters (*i.e.*, λ_{Poisson} and $\lambda_{\text{Exponential}}$, the rate constants for the Poisson and exponential distributions, respectively).

Modeling Long-Term Persistence

During chronic dosing, the number of individuals taking drug decreases over time. Persistence was defined as the period from the date of the first prescription to the date of discontinuation. Fractional persistence was defined as the proportion of patients remaining on treatment (5). We modeled the time dependence of persistence using the Weibull distribution function.

$$p(i) = 1 - e^{-\left(i/\lambda\right)^\kappa}$$

Weibull distribution parameters were obtained from survival functions using nonlinear regression.

RESULTS

Modeling Adherence with the Two-State Markov Dosing Model

The average adherence value for the 207 patients was 86.5%. Table I and Fig. 1 summarize dosing data and

modeling results from four representative patients: #808, #813, #820, and #810, who exhibited a wide range of overall adherence values: 94.5%, 73.2%, 54.5%, and 22.1%, respectively. Figure 1a, e, i, and m summarizes the observed number of dose units taken at each dosing event. The plots for the “whether or not dose was taken” dichotomous variable are shown in Fig. 1b, f, j, and n.

The two-state, time-homogeneous Markov model contains two parameters, p_{SS} and p_{FS} . Table I summarizes the maximum likelihood estimates of p_{SS} and p_{FS} and the PCI values. Stochastic simulations (Fig. 1c, g, k, and o) of dosing patterns from the two-state time-homogeneous Markov model with patient-derived estimates of p_{SS} and p_{FS} demonstrated qualitative concordance with the observed dosing patterns.

A 30-day simple centered moving average of the dosing sequence over time was conducted in four representative patients (Fig. 1r) to assess whether adherence changes within individuals over time. In Fig. 1r, adherence remains relatively constant over time, which lends support for using the time-homogeneous Markov chain model.

We conducted several additional assessments of the dosing pattern distribution predicted by the Markov model. Figure 1d, h, l, and p compares the observed cumulative dose at each time index (blue lines) to simulated cumulative dose (green lines). The concordance indicates that the model is capable of describing the time dependence of cumulative dose satisfactorily.

Because the primary goal of this work was to model the distribution of DH and drug-dosing intervals, we compared key statistics of DH and the distributions of both success and failure runs in the analyses (Fig. 1q and Table I). The model-predicted values of the number of DH and the average DH length are concordant with the observed values for each patient (Table I). The observed and predicted numbers of success and failure runs (Fig. 1q) are clustered around the line of identity, which indicates that the Markov model adequately describes dosing patterns. The Fisher exact test p values (Table I) are not significant indicating that the observed and predicted run distributions are not different.

Modeling Drug Holidays

We used the model to obtain the distribution of DH. Figure 2a and b shows the dependence of the number and average length of DH for a 90-day-long, once-daily dosing regimen on the percentage of missed doses for PCI values of 0.2, 1, and 5. The number of DH increases initially with increased percentage of missed doses, reaches a peak value, and then decreases rapidly. The initial increases are to be expected given the increase in nonadherence. However, at higher values of percent nonadherence longer DH become more frequent. The average DH length increases monotonically (Fig. 2b), and the probability of observing no DH decreases rapidly (Fig. 2c) with increased percentage of missed doses. The average DH length shows inflection points as the formation of longer DH and merging of adjacent DH become more likely. Eventually, as the percentage of missed doses approaches 100%, the average DH length approaches the length of the dosing regimen and the average number of DH approach unity.

More interestingly, the number, average length of DH, and the probability of observing no DH (Fig. 2a–c) at any

Table 1. Observed Adherence and Overdosing Characteristics, Markov Model Parameters, and Observed and Predicted Values from the Markov Model of the Number and Length of Drug Holidays in Four Representative Patients

Pt ^a	Length (days)	Observed		Markov model			Number of DH		Length of DH		<i>p</i> value ^c
		Adherence (%)	Overdosing (%)	<i>p</i> _{SS} (%)	<i>p</i> _{FS} (%)	PCI	Obs	Pred	Obs	Pred ^a	
808	176	94.5	8.50	93.9	100	ND ^b	0.00	0.00	1.00	0.999±0.01	0.83
813	179	73.2	12.3	70.8	79.2	1.57	3.00	1.63	3.00	3.03±0.65	0.84
820	90	54.5	11.1	56.3	53.7	0.900	4.00	4.59	4.25	3.85±0.66	0.34
810	104	22.1	2.90	36.4	18.5	0.400	10.0	10.1	7.40	7.41±1.65	0.37

^a Mean±SD^b The odds ratio is not defined, *i.e.*, it is infinity, because *p*_{FS} is unity^c *p* values are from the Fisher exact test comparing the observed distribution of all success and failure runs in each subject to the predicted distribution of all success and failure runs in that subject

fixed level of adherence (or nonadherence) are strongly dependent on the PCI. The effects of PCI are particularly prominent at intermediate levels of 10–50% missed doses that are frequently seen in the population and in clinical practice. For example, when 33% of prescribed doses (*i.e.*, one in three doses) are missed, the number and average length of DH in subjects with a low PCI of 0.2 are 4.11 and 4.32±SD 0.98, respectively. In subjects with PCI of 5, the number of DH was nearly 10-fold lower at 0.43, and the average DH length was 45% lower at 2.38±SD 0.64.

One of the strengths of our method is that it is capable of providing the complete distribution of DH for homogeneous (and nonhomogeneous) Markov chains. Figure 2d and e shows the probability of observing 1, 2, or 3 DH (data for ≥4 DH is not shown) for 90- and 30-day dosing regimens. Again, the strong effect of PCI identified in Fig. 2a–c was found—the peak probability density values for PCI=5 occurred at a higher value of percentage of missed doses compared to PCI=0.2. DH occurred with greater frequency and at lower levels of nonadherence in patients with low PCI values compared to those with high PCI. Because of the relative values of *p*_{SS} and *p*_{FS}, the dosing event failures in high PCI subjects are more likely to be interrupted by successes, whereas the dosing event failures in low PCI subjects are more likely to be followed by failures that cause consecutive missed doses and DH.

A 30-day dosing regimen exhibits a more complex probability density pattern than a 90-day dosing regimen (Fig. 2e) because the shorter total length of the regimen constrains the frequency and length of DH.

Modeling Dose-Timing Deviations

We evaluated whether the distribution of the dose-timing deviations was described using a VM distribution. Since the VM PDF has not been previously used to model drug adherence, we provide a concise background on its salient characteristics (Fig. 3a–c and Supplementary Results).

Figure 3d–g shows quantile–quantile plots that enable comparisons of the observed distribution of dose timing deviations transformed to the angular coordinate system for the four representative patients to the VM distribution function. Visually, the linearity of the quantile–quantile plots of all four subjects indicates a good fit of the VM model to the observed data. We tested the goodness of fit of the VM distribution using the Watson test. The *p* values for patient #808, #810, and #813

were all *p*>0.1, whereas for patient #820, 0.01≤*p*≤0.05 was obtained. We conclude that the VM distribution is a suitable approach for modeling dose-timing deviations.

Modeling the Drug-Dosing Interval Distribution

We combined key results from the two-state Markov chain runs distribution with the VM distribution of dose-timing deviations to formally derive the distribution of drug-dosing intervals (Fig. 4). The calculated distribution of dosing intervals describes salient qualitative and quantitative features of the observed distribution of dosing intervals including its step-like pattern. The PDF of the dosing interval deviates considerably from normality and exhibits multimodality. The number of peaks in the PDF generally decreases in patients with higher adherence, *e.g.*, the PDF for patient #808 (94.5% adherence) has only two peaks, whereas the PDF for patient #810 (22.1% adherence) has numerous peaks. The model underestimated the observed dosing interval distribution modestly in part of the domain for patients #810 and #820. The exact reasons for the underestimation are not known. However, contributions from sampling variability and deviations from randomness caused by patient-specific behavioral patterns could be involved.

Modeling Overdosing Events

The dosing records of 95.6% of patients had at least one overdosing event, *i.e.*, an event in which the patient was taking more doses than prescribed (>1 each day).

The Poisson distribution provided a satisfactory model for the distribution of the extent of overdosing (*i.e.*, the number of additional doses taken)—the KS test *p* values were 1 for all four patients in Fig. 1.

To obtain a more complete assessment, we investigated the distribution of the interval between overdosing events. An exponential distribution can describe the time interval between events whose counts are Poisson distributed (14). An exponential distribution model fitted the interval between overdosing events satisfactorily as assessed with probability–probability plots (Supplementary Figure 1) and the KS test (*p* values were not significant for any of six patients analyzed).

In the next step, we assessed whether our Poisson model for the extent of overdosing could be combined with the

Fig. 2. Distribution of drug holidays for a markov process. **a** Average number of drug holidays (operationally defined as ≥ 3 dosing events missed) for 90-day-long, once-daily dosing regimen as a function of the percentage of dosing events missed (*i.e.*, percentage nonadherence) on the *x*-axis. Three values of cooperativity index 5 (green open symbols and curve), 1 (blue open symbols and curve), and 0.2 (red open symbols and curve) are shown. **b** Dependence of average length of a drug holiday (note logarithmic *y*-axis) on the percentage of dosing events missed for the cooperativity index values of 0.2, 1, and 5, respectively. **c** Probability of not observing a drug holiday as a function of the percentage of dosing events missed for subjects with cooperativity index values of 5, 1, and 0.2. **d** Probability density of observing one drug holiday (*top graph*), two drug holidays (*middle graph*), or three drug holidays (*bottom*) as a function of the percentage of dosing on the *x*-axis for a 90-day-long dosing regimen. **e** Same results as **d** but for a 30-day-long dosing regimen

exponential model for the overdosing interval to obtain a more parsimonious model. We found that the rate constant of the exponential distribution describing the interval between overdosing events was estimated as the product of adherence and the rate constant of the Poisson distribution describing the extent of overdosing. Figure 5 highlights the excellent correlation between the two parameters. The three outliers in Fig. 5 were investigated more closely and found to result from patients who had less overdosing occurrences; two outliers had three overdosing intervals, and one had four overdosing intervals. Thus, the extent of overdosing and the interval between overdosing events can be described with adherence and a single additional parameter: the rate constant for the Poisson distribution (λ_{Poisson}) model for the distribution of the extent of overdosing.

Modeling Persistence

In chronic dosing regimens, individuals may stop taking medication altogether, which leads to a decline in persistence. A trend of declining persistence was observed in the PKC

adherence data set of 207 hypertension patients. The Kaplan-Meier curves (Fig. 6a) summarize the fractional persistence, *i.e.*, the fraction of subjects remaining on treatment.

We compared several parametric time-to-event models for describing persistence including the exponential and gamma distributions. We found that the Weibull distribution effectively describes the persistence data in each of the subgroups in the PKC adherence dataset. The subgroups shown by the green and blue lines (Fig. 6a) both had a 6-month observation period and overall adherence values of 68.5% and 92.8%, respectively. The subgroup with the red line was followed up for 9 months and had 92.5% adherence. The Weibull model provides a satisfactory description of the persistence in the combined dataset over a 6-month period (data not shown).

To further challenge the Weibull model, we assessed whether it effectively described previously reported long-term persistence data. The persistence data from Catalan and LeLorier (4) was for 983 Canadian patients prescribed statin therapy over an 5-year study period; the data analyzed by Larsen *et al.* (5) was for 3623 patients from a Danish

Fig. 3. Assessments of the von Mises distribution for modeling angular dose-timing deviations. **a** Cartesian coordinate representation of von Mises probability density functions for values of mean $\mu=0$, $\pi/2$, and $-\pi/2$, and a fixed value of $\kappa=2$. The x -axis represents the value of the angular random variable. **b** Cylindrical coordinates representing the von Mises probability density functions shown in **a** wrapped around a unit circle. **c** von Mises probability density functions with mean fixed to $\mu=0$ but with differing $\kappa=0$, 0.5, 1, and 2

population-based cohort on lipid lowering drugs over a 5-year study period. For each of these data sets, the Weibull model provided a parsimonious description of the persistence trend over the longer observation period (Fig. 6b).

DISCUSSION

We investigated a novel approach for analyzing, modeling, and simulating dosing patterns and dosing intervals in patients. The run distribution from a two-state Markov process was convolved with the VM distribution to obtain the drug-dosing interval distribution. This parsimonious stochastic modeling strategy was found to describe key

characteristics of dosing patterns including the probability of DH, the distribution of dose timing deviations and dosing intervals, the occurrence of overdosing events, and persistence. The model was challenged with a large data set containing MEMS dosing data for 207 patients on once-daily treatment for hypertension.

The starting point for our modeling strategy is the two-state Markov process component, which contains only two estimated parameters (p_{FS} and p_{SS}). The PCI is calculated from these two parameters. The PCI was a very effective parameter for describing the substantial variations in dosing patterns at a given level of overall adherence. For two subjects with the *same* adherence, a subject with lower PCI

Fig. 4. Modeling dosing intervals. The *red circles* in **a–d** show the observed distribution of dosing intervals, whereas the *solid line* shows the predicted distribution from the Markov-von Mises model for patients #808, #813, #820, and #810, respectively. The *insets* show the predicted probability density functions corresponding to the distribution functions

exhibits more DH than a subject with a higher PCI. The subject with a lower PCI is less likely to follow a failed dose with a successful dose causing the failed doses to be distributed in a manner that creates more DH. The PCI has

the form of an odds ratio (OR). The sampling distribution of OR has been extensively characterized in the statistical literature. The OR is insensitive to the type of sampling, *i.e.*, the OR estimate from a convenience sample, *e.g.*, case–

Fig. 5. Modeling the extent of overdosing events and the intervals between overdosing events. The rate constant ($\lambda_{\text{Exponential}}$) for the exponential distribution (y-axis) is plotted against rate constant (λ_{Poisson}) for the Poisson distribution multiplied by the adherence (x-axis). Patients with ≥ 3 overdosing events, which corresponds to ≥ 2 overdosing intervals are plotted. A total 166 patients are represented on the graph (blue circles), and 3 patients were outliers (red circles). The dashed line is the regression line, whose equation and regression coefficient (R^2) value are also shown

control study design, is the same as that from a population sampling design. For inference, the logarithm of the OR, which is symmetric about OR of 1 (logarithm OR equals 0) is approximately normally distributed, is used.

Our use of the theory of runs is an innovative contribution that has potential clinical pharmacology applications. Many multiple dosing regimens for chronic diseases require a run of successful dosing events so that steady-state concentrations in the therapeutic range are reached and maintained. Poor efficacy and drug resistance to cell cycle specific antitumor agents, antibiotics, antivirals, and life-cycle-specific antiparasitic agents can occur as a result of nonadherence. Dosing frequency, PK, and PD factors can make one drug less “forgiving” of nonadherence than another (15). Drugs with short half-lives, steep concentration–effect relationships, and less frequent dosing regimens are more sensitive to nonadherence (15). The Markov chain used is a binary two-state Markov chain. Overdosing is addressed in a separate modeling component. An alternative approach would be to superimpose overdosing into a multistate Markov model. This approach is more complex because additional parameters are required to characterize each overdosing state in the Markov chain.

A rigorous model for drug-dosing intervals can be directly leveraged to understand the effects of adherence on PK. In linear models, the concentration profiles can be obtained using the principle of superposition; the mean and variance of the

Fig. 6. Modeling persistence. **a** Observed persistence distribution as the fraction of the patients remaining on drug treatment for the 207 patients on once-daily hypertensive drug treatment. The graph contains data for three separate subgroups (green, blue, and red lines) in the PKC adherence dataset, and a Weibull function (dashed lines) was fit to each group separately. **b** Fit of the Weibull distribution to two literature studies of patient persistence with long-term drug therapies. The upper line (green) corresponds to the observed data from Larsen *et al.* (5) and the lower line (red) to that from Catalan and Leloir (4). In **a** and **b**, the dashed line represents the fit of the observed data to a Weibull distribution

concentration profile are proportional to the mean and variance of the dosing interval distribution (16). Simulation software is needed for drugs with nonlinear PK and for PD models, which are frequently nonlinear (17). Li and Nekka examined the effect of adherence on several integral pharmacokinetic indices including area under the curve (AUC). Kiwuwa-Muyingo *et al.* (18) used coarse-grained patient-reported adherence with missing data information, to cluster clinical outcome data in HIV patients on anti-retroviral therapy.

The DH distribution and a substantive portion of the drug-dosing interval distribution are explained by run patterns in the Markov process. Girard *et al.* (19,20) utilized a three-state Markov model but used it in conjunction with a mixed effect regression model for describing HIV patients' compliance to three-times-daily zidovudine treatment; dose-timing deviations were assumed to be multivariate normally distributed with covariates. Modi *et al.* (21) used a two-state Markov model that leveraged the partition function approach for characterizing helix–coil transitions in biophysics. A distinctive contribution of this research is its rigorous

mechanistic approach to understanding within-subject variability in dosing patterns.

We developed a novel, angular coordinate representation of drug-dosing events that provides an intuitive analysis framework. It directly reflects the periodicity inherent in chronic drug-dosing regimens and enables the use of the VM distribution and the formalism of circular statistics. Although we assumed that the prescribed dosing intervals are evenly spaced for convenience, the approach can be easily modified to accommodate three- and four-times-daily dosing regimens that have uneven time intervals to accommodate sleep and regimens in which the prescribed dose is intermittent.

We modeled persistence with a Weibull function. Because the PKC adherence data were limited to approximately 6–9 months, we critically evaluated whether the approach was capable of modeling loss of persistence over the longer term. Analyses of the Catalan and LeLorier and Larsen *et al.* (4,5) data sets demonstrated that the model was generalizable to long-term patient persistence data.

Our analysis methods could be used to more informatively analyze MEMS data, *e.g.*, to derive the PCI. MEMS have allowed for more accurate and non-intrusive estimation of patient adherence. Incorrect adherence data can result in misinterpretation of clinical trial results and lead to less accurate dosing information (22,23). MEMS are more convenient than, and superior to self-reports and diaries, clinic, or pharmacy-based refill counting and therapeutic drug monitoring via blood or saliva tests, which are prone to bias, *e.g.*, due to forgetfulness, incorrect entry and the “toothbrush effect.” Despite their many advantages, MEMS nonetheless have some limitations. There may be bias favoring greater adherence if patients are aware of the monitoring device. It is assumed that each opening represents one dose and that no dose is taken when no opening is recorded (24). However, some patients may remove multiple doses for subsequent use and curiosity openings may be recorded as dosing events (25). MEMS are limited in their capacity to detect simultaneous overdosing that occurs concomitantly with another dosing event. We expect that our modeling strategy will generalize to improvements of MEMS technology that overcome these limitations.

We envisage our modeling and simulation methods to complement and enhance clinical trial simulation and population pharmacometric modeling. For example, the approaches could be used to derive parameter estimates from MEMS devices. The parameter estimates could be utilized for simulating dosing patterns. The dosing patterns could be used as input for PK/PD models to better predict population drug concentrations, drug effects, and efficacy and thus enhance clinical trial simulations.

CONCLUSIONS

Our approach, which combines Markov chains, the VM distribution and runs theory, was effective at describing the complex, skewed, and multimodal distributional features of DH and dosing intervals resulting from imperfect adherence and patient responses to missed doses. Our findings highlight the critical importance of dosing patterns to drug efficacy in real-world settings. Our model could prove useful in population PK-PD and in clinical trial simulations.

ACKNOWLEDGMENTS

Support from the National Multiple Sclerosis Society (RG4836-A-5) to the Ramanathan laboratory is gratefully acknowledged.

Financial Conflicts See disclosure statement.

Confidentiality Use of the information in this manuscript for commercial, non-commercial, research or purposes other than peer review not permitted prior to publication without expressed written permission of the author.

Author Contributions Kelly Fellows – Conducted experiments, data analysis, manuscript preparation.

Sheril Alexander, Alyssa Droopad, Jenna Covelli, Vivian Rodriguez-Cruz – Data acquisition.

Murali Ramanathan – Study concept and design, data analysis, manuscript preparation.

Disclosure Dr. Murali Ramanathan received research funding from the National Multiple Sclerosis Society and the Department of Defense. He received compensation for serving as an Editor from the American Association of Pharmaceutical Scientists. These are unrelated to the research presented in this report.

Kelly Fellows, Sheril Alexander, Alyssa Droopad, Jenna Covelli and Vivian Rodriguez-Cruz have no conflicts to disclose.

REFERENCES

1. Paes AH, Bakker A, Soe-Agnie CJ. Impact of dosage frequency on patient compliance. *Diabetes Care*. 1997;20(10):1512–7.
2. Urquhart J, De Klerk E. Contending paradigms for the interpretation of data on patient compliance with therapeutic drug regimens. *Stat Med*. 1998;17(3):251–67. discussion 387–259.
3. Urquhart J. Ascertaining how much compliance is enough with outpatient antibiotic regimens. *Postgrad Med J*. 1992;68 Suppl 3:S49–58. discussion S59.
4. Catalan VS, LeLorier J. Predictors of long-term persistence on statins in a subsidized clinical population. *Value Health J Int Soc Pharmacoecono Outcomes Res*. 2000;3(6):417–26.
5. Larsen J, Andersen M, Kragstrup J, Gram LF. High persistence of statin use in a Danish population: compliance study 1993–1998. *Br J Clin Pharmacol*. 2002;53(4):375–8.
6. Hughes DA, Walley T. Predicting “real world” effectiveness by integrating adherence with pharmacodynamic modeling. *Clin Pharmacol Ther*. 2003;74(1):1–8.
7. Circular Statistics (version 0.4–7) (computer program). Venice, Italy, San Luis Obispo, CA, USA: Department of Environmental Sciences, Informatics and Statistics, Foscari University, Venice, Italy; 2013.
8. R: A language and environment for statistical computing (computer program). Vienna, Austria: R Foundation for Statistical Computing; 2014.
9. Rajarshi MB. Markov chains and their extensions. *Statistical inference for discrete time stochastic processes*. New York, Springer; 2012:19–38.

10. Feller W. An introduction to probability theory and its applications. 3rd ed. New York: Wiley; 1968.
11. Uspensky JV. Introduction to mathematical probability. 1st ed. New York: McGraw-Hill Book Company; 1937.
12. Fu JC, Koutras MV. Distribution theory of runs—a Markov-chain approach. *J Am Stat Assoc*. 1994;89(41):1050.
13. Koutras MV, Alexandrou VA. Runs, scans and urn model distributions: a unified Markov chain approach. *Ann Inst Stat Math*. 1995;47(19):743.
14. Johnson NL. Poisson distribution. In: Balding DJ, Fisher NI, Johnstone IM, Kadane JB, Molenberghs G, Ryan LM, Scott DW, Smith AFM, Teugels JL, editors. *Univariate discrete distributions*. 3rd ed. United States: Wiley; 2005. p. 156–207.
15. Levy G. A pharmacokinetic perspective on medicament non-compliance. *Clin Pharmacol Ther*. 1993;54(3):242–4.
16. Gallivan S, Pagel C, Utley M, Franklin BD, Taxis K, Barber N. A technical note concerning non-adherence to drug therapy: exact expressions for the mean and variance of drug concentration. *Health Care Manag Sci*. 2008;11(3):296–301.
17. Boissel JP, Nony P. Using pharmacokinetic–pharmacodynamic relationships to predict the effect of poor compliance. *Clin Pharmacokinet*. 2002;41(1):1–6.
18. Kiwuwa-Muyingo S, Oja H, Walker SA, Ilmonen P, Levin J, Todd J. Clustering based on adherence data. *Epidemiol Perspect Innov EP+I*. 2011;8:3.
19. Girard P, Sheiner LB, Kastrissios H, Blaschke TF. Do we need full compliance data for population pharmacokinetic analysis? *J Pharmacokinet Biopharm*. 1996;24(3):265–82.
20. Girard P, Blaschke TF, Kastrissios H, Sheiner LB. A Markov mixed effect regression model for drug compliance. *Stat Med*. 1998;17(20):2313–33.
21. Wong D, Modi R, Ramanathan M. Assessment of Markov-dependent stochastic models for drug administration compliance. *Clin Pharmacokinet*. 2003;42(2):193–204.
22. Urquhart J. Role of patient compliance in clinical pharmacokinetics. A review of recent research. *Clin Pharmacokinet*. 1994;27(3):202–15.
23. Urquhart J. Patient non-compliance with drug regimens: measurement, clinical correlates, economic impact. *Eur Heart J*. 1996;17(Suppl A(14)):8–15.
24. Olivieri NF, Matsui D, Hermann C, Koren G. Compliance assessed by the Medication Event Monitoring System. *Arch Dis Child*. 1991;66(12):1399–402.
25. Arnet I, Haefeli WE. Overconsumption detected by electronic drug monitoring requires subtle interpretation. *Clin Pharmacol Ther*. 2000;67(1):44–7.