<http://www.diva-portal.org>

Preprint

This is the submitted version of a paper published in *Journal of Pharmacokinetics and Pharmacodynamics*.

Citation for the original published paper (version of record):

Ernest II, C., Nyberg, J., Karlsson, M., Hooker, A. (2014)
Optimal clinical trial design based on a dichotomous Markov-chain mixed-effect sleep model.
Journal of Pharmacokinetics and Pharmacodynamics, 41(6): 639-654
<http://dx.doi.org/10.1007/s10928-014-9391-z>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-208583>

Optimal clinical trial design based on a dichotomous Markov-chain mixed-effect sleep model

C. Steven Ernest II^{1,2}, Joakim Nyberg¹, Mats O. Karlsson¹, Andrew C. Hooker¹

(1) Department of Pharmaceutical Biosciences, Uppsala University, Sweden;

(2) Eli Lilly and Company, Indianapolis, IN, USA.

To whom correspondence should be addressed:

C. Steven Ernest II

Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285

Telephone: (317) 276-2350

Email: cse@lilly.com

Abstract.

Objectives: D-optimal designs for discrete-type responses have been derived using generalized linear mixed models, simulation based methods and analytical approximations for computing the Fisher information matrix (FIM) of non-linear mixed effect models with homogeneous probabilities over time. In this work, D-optimal designs using an analytical approximation of the FIM for a dichotomous, non-homogeneous, Markov-chain phase advanced sleep non-linear mixed effect model was investigated.

Methods: The non-linear mixed effect model consisted of transition probabilities of dichotomous sleep data estimated as logistic functions using piecewise linear functions. Theoretical linear and nonlinear dose effects were added to the transition probabilities to modify the probability of being in either sleep stage. D-optimal designs were computed by determining an analytical approximation the FIM for each Markov component (one where the previous state was awake and another where the previous state was asleep). Each Markov component FIM was weighted either equally or by the average probability of response being awake or asleep over the night and summed to derive the total FIM (FIM_{total}). The reference designs were placebo, 0.1, 1-, 6-, 10- and 20-mg dosing for a 2- to 6-way crossover study in 6 dosing groups. Optimized design variables were dose and number of subjects in each dose group. The designs were validated using stochastic simulation/re-estimation (SSE).

Results: Contrary to expectations, the predicted parameter uncertainty obtained via FIM_{total} was larger than the uncertainty in parameter estimates computed by SSE. Nevertheless, the D-optimal designs decreased the uncertainty of parameter estimates relative to the reference designs. Additionally, the improvement for the D-optimal designs were more pronounced using SSE than predicted via FIM_{total}.

Conclusion: Through the use of an approximate analytic solution and weighting schemes, the FIM_{total} for a non-homogeneous, dichotomous Markov-chain phase advanced sleep model was computed and provided more efficient trial designs and increased NLMEM parameter precision.

Keywords: D-optimality; discrete data; optimal designs; nonlinear mixed effects; Markov models.

Introduction

Mathematical model-based approaches can be used to describe pharmacological responses or to identify and quantify pharmacokinetic/pharmacodynamic (PKPD) relationships, thereby facilitating drug development in an increasingly complex, costly, and risky landscape where data and time are limited. Pharmacometrics¹ is a culmination of

pharmacology, PKPD, and statistics to qualitatively and quantitatively describe pharmacological responses², and provides the prospect to better understand and predict physiological events. Pharmacometric models can be utilized to make statistical inferences, individualize treatments, or find optimal trial designs³⁻⁵. These models are generally based on nonlinear mixed-effects modeling (NLMEM) approaches⁶⁻⁷ as first proposed by Sheiner et al.⁸,

and in general, PKPD data on which they are based can exist in the form of either continuous or discrete random variables. Discrete random variables, in turn, can take the form of, for example, ordered categorical, non-ordered categorical, dichotomous, or count variables.

One such series of discrete random variables is that observed during sleep. Sleep is not a homogeneous state of unconsciousness, but is instead characterized by different non-ordered categorical sleep stages that define the ‘sleep architecture’. Disturbances of the ‘normal’ sleep architecture have been estimated to affect approximately 10% of the world’s population, resulting in impaired memory and cognitive functions that result in compromised productivity and wellness^{9,10}. Sleep disturbances therefore constitute a substantial socio-economic burden and, for this reason, appropriate diagnoses and treatment options for these pathologies represent great challenges. This has led to vast investments to develop and test safer and more effective molecules to regulate the sleep-wake alternation^{11,12}. One type of clinical trial conducted for the development and evaluation of new sleep-inducing molecules is phase advanced sleep, wherein subjects go to sleep several hours before their usual bedtime to induce transient insomnia, thus disrupting their normal sleep architecture and subsequently testing a drug’s effectiveness on inducing sleep^{13,14}. The NLMEMs created to describe the sleep architecture for this type of polysomnography (PSG) data are partitioned into separate, smaller models (sub-models) based on the previous sleep state (Markov-chain) and estimate the transition probabilities to different sleep states¹⁵⁻¹⁹. To account for non-homogeneity over time, temporal dependence is implemented using piecewise linear functions. Thus, optimizing these phase advance sleep clinical trials can provide more accurate NLMEM parameter estimation and can make a major impact in reducing the associated costs for molecules developed to treat these pathologies.

Optimal design using NLMEM was first elaborated by Mentré et al. by deriving a first-order analytic expression of the Fisher Information Matrix (FIM) for continuous data²⁰ and optimization of the design variables provided more robust and efficient NLMEM parameter estimates. Recently, optimization of design variables for homogenous discrete-type pharmacodynamic responses has been explored²¹⁻²³. These types of models are commonly

constructed to describe the individual likelihood of observing a discrete response and are often more non-linear with respect to the random effects. Hence, a higher order linearization (i.e. greater than first order) is required to accommodate the nonlinear random effects, such as the Laplace approximation as implemented in NONMEM²⁴.

The purpose of this work was to investigate D-optimal experimental design methodologies, maximizing the determinant of the FIM, for a phase advanced sleep NLMEM where the individual likelihood of observing a discrete response was non-homogeneous over time. However, to minimize the number of sub-models explored in this work, sleep data were dichotomized (awake and asleep) and as such, the total FIM was derived from two separate sub-models (one where the previous state was awake and another where the previous state was asleep). The expression for the FIM was based on an approximation of the likelihood using a second order Taylor expansion as described previously by Breslow and Clayton²⁵, Longford²⁶, Wand²⁷, Ogungbenro and Aarons²¹ and Nyberg et al.^{22,23}, which, in general, does not acknowledge correlation between the fixed and random effect parameters^{21-23,25-27}. Thus, through the use of an approximate analytic solution of parts of the FIM, the FIM was computed without extensive Monte Carlo simulations as previously employed by Nyberg et al.^{22,23}. The experimental designs considered within this work included 2- to 6-way crossover designs with 6 different dosing groups, and the design variables optimized were the dose and the number of subjects per dosing group.

Methods and Materials

More expansive nonlinear mixed-effects models have previously been developed to estimate the transition probabilities to different sleep states¹⁵⁻¹⁹; however, for this work, phase advanced sleep data were dichotomized (awake and asleep) to form the framework to investigate D-optimality for non-homogeneous (over time) discrete responses.

The analyses consisted of: (1) characterize the sleep stage transition probabilities between awake and asleep in phase advanced subjects over a 13-hr PSG measurement period after placebo dosing using a NLMEM approach; (2) incorporate theoretical linear and non-linear dose effects to some of the transition probabilities to alter these the transition probabilities obtained from fitting the placebo data; (3) establish

reference designs to evaluate the predicted precision of parameter estimates using the developed optimal design methodology and compare the predicted precision to that derived from simulation/estimation; and (4) optimize either the dose or the number of subjects per dosing group using the developed optimal design methodology and compare the predicted precision to that derived from simulation/estimation. Specific details pertaining to these analyses are provided below.

Dichotomous Markov mixed-effect sleep model

The model used in the present work was derived from the placebo arm of two phase advanced sleep clinical trials with an investigational drug, of which 27 subjects received a single dose of placebo. Data collected from these two studies were conducted at two different sites utilizing a randomized, placebo-controlled, 4-period crossover, incomplete block design with a washout duration of at least 7 days between periods to minimize carry-over effects related to drug concentration and sleep pattern in previous periods. Subjects were phase advanced on the first day of each period of the study for the purpose of assessing drug effects during what would normally be an 'awake' state. For each dosing day, study medication was administered at approximately 05:30 pm, on a single occasion, 30 minutes prior to the artificially early bedtime. Subjects remained awake until lights were turned off at approximately 06:00 pm (30 minutes after dosing), and underwent 13 hours of PSG measurements while in bed. The first 3 periods of the study were devoted to PSG measurements, whereas the fourth period was used to obtain plasma concentrations of the investigational drug. The placebo data consisted of 42050 sleep stage observations (~1560 epochs per subject). For the current work, the epochs other than awake ($m=0$) were merged into a single sleep stage termed 'asleep' ($m=1$).

The probability of an event one step into the future conditioned exclusively on the previous state ('First-order Markov chain process') is equal to the conditional probability of the future event given just the present one. Therefore, partitioning the sleep data into two, different, smaller models (sub-models), one where the previous state was awake ('From awake') and another where the previous state was asleep ('From asleep'), allows the independent

estimation of parameters related to the transitioning to a different sleep state given the previous sleep state. Estimation of the probability of transitioning to another state, or transition probability (TP), from each particular sleep stage to another was obtained through the implementation of a non-homogeneous Markov-chain model, using a nonlinear mixed-effect approach, similar to that previously reported^{16,17}.

If $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{in})$ is the vector of observations for the i^{th} subject, then the probability that Y_{it} is equal to the stage m ($m=0$ or 1) at epoch = t , given that the preceding observation was k ($k \neq m$), has the following general structure:

$$\text{logit}(P(Y_t = m | Y_{t-1} = k)_i) = \text{logit}(p_{imk}) = g_{m|k} + \eta_i, \quad (1)$$

where

$$\text{logit}(p_{imk}) = \ln \left(\frac{p_{imk}}{1 - p_{imk}} \right), \quad (2)$$

and,

$$g_{m|k} = \ln \left(\frac{TP_{m|k}}{1 - TP_{m|k}} \right), \quad (3)$$

where $TP_{m|k}$ is the estimated transition probability, θ , to m given previous observation of k . Therefore,

$$P(Y_t = m | Y_{t-1} = k)_i = \frac{e^{g_{m|k} + \eta_i}}{1 + e^{g_{m|k} + \eta_i}}, \quad (4)$$

and the probability Y_{it} is equal to the stage m ($m=0$ or 1) at epoch = t , given that the preceding observation was m is

$$P(Y_t = m | Y_{t-1} = m)_i = 1 - \frac{e^{g_{m|k} + \eta_i}}{1 + e^{g_{m|k} + \eta_i}}. \quad (5)$$

The logit transform is used to ensure the probability to be between zero and one, and $g_{m|k}$ defines the two sub-models. Inter-individual variability (IIV, η_i) of the population was included by assuming each individual logit to be normally distributed with a mean of zero and variance, ω^2 . No correlation was assumed between the two sub-models.

As the TPs were non-homogeneous over the night, temporal dependence was implemented using piecewise linear functions. TPs and associated IIVs were estimated at the beginning of the night, at the last observation (1560 epoch) and at additional, estimated, break points. Inclusion of break points was tested for significance using the Akaike information criterion and log-likelihood ratio test (p -value < 0.05). All break points were assumed to be common to the entire population. Using piecewise linear functions, the TPs between the

estimated logits were provided by linear interpolation.

As these models describe the individual likelihood of observing a discrete response, a higher order linearization (than the first order) is typically required due to the non-linear random effects^{16,17}. Therefore, parameter estimation was performed using NONMEM 7.1.2[®] using the LAPLACE option²⁴.

After the placebo NLMEM was derived, theoretical linear (slope-dose) and non-linear ($E_{\max}$ and ED_{50}) dose effects were added to some of the TPs. These parameter values were selected to allow dosages of 1-, 6- and 10-mg to alter these TPs obtained from fitting the placebo data.

Optimal Design

The FIM for the i^{th} individual with the vector of design parameters X_i is defined as the negative expectation of the second-order partial derivatives of the population log-likelihood function ($\ln L$) of a model

$$FIM(\Theta, X_i) = -E \left[\frac{\partial^2 \ln L(X_i, \Theta)}{\partial \Theta_i \partial \Theta^T} \right], \quad (6)$$

where Θ is the vector of population parameters, $\Theta = [\theta_1, \dots, \theta_n, \omega^2_{1, \dots, n}]$, and X_i is the vector of design parameters. The FIM above is analytically intractable and approximations have been used. The expression for the FIM requiring higher order linearization, as prerequisite for the current work, has been described based on an approximation of the likelihood using a second order Taylor expansion for GLMMs by Breslow and Clayton²⁵, Longford²⁶, Wand²⁷. Recently, Ogungbenro and Aarons²¹ and Nyberg et al.^{22,23} have utilized this FIM approximation for discrete response GLMMs and NLMEMs. The proposed FIM²¹ yields a block diagonal matrix which does not acknowledge direct correlation between the ω^2 and θ

$$FIM(\Theta, X_i) = \begin{bmatrix} J_i^T V_i^{-1} J_i & 0 \\ 0 & \frac{1}{2} F \end{bmatrix}, \quad (7)$$

where

$$J_i = \frac{\partial P_i(Y_i = 0 | Y_{i-1} = k)}{\partial \theta}, \quad (8)$$

and

$$F = \text{tr} \left(V_i^{-1} \frac{dV_i}{d\omega_r} V_i^{-1} \frac{dV_i}{d\omega_s} \right), \quad (9)$$

for $r, s = 1 \dots n$ and the approximate variance-covariance matrix is given by

$$V_i \approx Z_i \Omega Z_i^T + W_i, \quad (10)$$

where

$$\Omega = \text{diag}[\omega_1^2 \dots \omega_n^2], \quad (11)$$

$$Z_i = \frac{\partial P_i(Y_i = 0, Y_{i-1} = k)}{\eta} \Big|_{\eta=0}, \quad (12)$$

and for the dichotomous response, W_i is a $t_i \times t_i$ diagonal element matrix representing the variance of the binomial distribution, i.e.

$$W_i = \text{diag}(P_i(Y_i = 0, Y_{i-1} = k) \cdot (1 - P_i) \cdot (Y_i = 0, Y_{i-1} = k)). \quad (13)$$

The population FIM can be considered the sum of all individual FIM

$$FIM(\Theta, X) = \sum_{i=1}^N g_i \cdot FIM(\Theta, X_i), \quad (14)$$

where g_i is the number of individuals in each design group and N is the total number of groups.

Theoretically, the FIM could be based on the covariance matrices associated with maximum-likelihood estimates for all the model parameters in the NLMEM. However, the NLMEM for the dichotomous phase advanced sleep data was partitioned into two sub-models and correlation of parameters between the sub-models are block diagonal. Therefore, for the current work, two separate FIMs ($FIM_{k=0}$ and $FIM_{k=1}$) were computed

$$FIM(\Theta, X)_0 = \sum_{i=1}^N g_i \cdot FIM_i(\Theta, X_i) \Big|_{k=0}, \quad (15)$$

and

$$FIM(\Theta, X)_1 = \sum_{i=1}^N g_i \cdot FIM_i(\Theta, X_i) \Big|_{k=1}, \quad (16)$$

and correlation of parameters between the sub-models are block diagonal.

Here, these two matrices can be simply added to construct the total FIM, FIM_{total} . This simple addition implies equal weighting for the two sub-models; ‘Equally Weighted FIM’. This method was employed previously when computing the FIM_{total} for a categorical model with a Markov element²³. However, the probability of observing a specific $TP_{m|k,t}$ for the phase advanced dichotomous sleep data is not equally weighted over the night. Therefore, a weighting for FIM_0 and FIM_1 that corresponds to the probability of the response being awake, $p(m=0)$, or asleep, $(1-p(m=0))$, over the night was also included; ‘Average Probability Weighted FIM’. In order to include a weighting but reduce the computational burden, FIM_0 and FIM_1 were weighed according to the average (over time) probability of the response being $p(m=0)$ or $(1-p(m=0))$ over the night based on the typical individual ($\eta=0$) for each design group i

$$FIM_{total,i} = p_i(m=0) \cdot FIM_{0,i} + (1-p_i)(m=0) \cdot FIM_{1,i}, \quad (17)$$

which makes $p_i(m=0)$ time independent. The probabilities directly inherited from the model were the conditional probabilities, $P_i(Y_{t_i}=0 \text{ or } 1 | Y_{t-1}=0 \text{ or } 1)$, and hence, Bayes theorem was used to calculate the overall probability of being awake, $p_i(m=0)$, or being asleep, $p_i(m=1)$.

According to the Cramer-Rao inequality (Equation 2)^{28,29}, the FIM_{total} provides an asymptotic lower bound for the precision of any unbiased maximum likelihood estimator:

$$COV(\Theta, X) \geq FIM_{total}(\Theta, X)^{-1}. \quad (18)$$

Note that, because the FIM_{total} is an asymptote, the covariance of a model's fitted parameters in an actual experiment will generally be greater than (less precise), and asymptotically approach, the inverse of the FIM based on the Cramer-Rao inequality³⁰.

By adjusting design parameters, X , one can change the value of the FIM_{total} and thus, with some metric, optimize a design based on the FIM_{total} . The most common optimization approach is to minimize the joint confidence volume of the parameters, which is equivalent to maximizing the determinant of the FIM_{total} . This criterion is called D-optimality. D-optimal design seeks to minimize the variance and covariance of parameter estimates.

The expected improvement of one design, X_1 , versus another, X_2 , based on the FIM_{total} was computed via an D-efficiency calculation, $D_{eff,Pred}$,

$$D_{eff,Pred} = \left(\frac{|FIM_{total}(\Theta, X_1)|}{|FIM_{total}(\Theta, X_2)|} \right)^{1/p}, \quad (19)$$

where p is the total number of estimated parameters in the NLMEM.

Design Evaluation and Optimization

The optimal experimental design software PopED 2.13 was used for computing the predicted precision of parameter estimates via the FIM_{total} and for optimizing the trial design based on D-optimality³¹⁻³³. Initial evaluation of the predicted precision of NLMEM parameter estimates and FIM_{total} approximation using both 'Average Probability Weighted FIM' and 'Equally Weighted FIM' was conducted with 27 subjects receiving a single dose of placebo. This initial evaluation also explored the impact of the

number of epochs/subject uniformly distributed over the night ranging from 100 to 1560.

The experimental designs considered within the context of this work included a 2- to 6-way crossover design (arm) with 6 different dosing groups (group). The reference designs consisted of placebo, 0.1-, 1-, 6-, 10-, and 20-mg and 7 subjects per group. Evaluation of the FIM_{total} of these reference designs using both 'Average Probability Weighted FIM' and 'Equally Weighted FIM' was performed to provide a foundation to assess the optimized designs. Optimization using both 'Average Probability Weighted FIM' and 'Equally Weighted FIM' focused on either the dose (the doses were optimized within [0.01] units) or the number of subjects per group. Constraints incorporated in the optimization included: 1) at least 1 subject per group; 2) maximum total number of subjects was 42 and 3) minimum and maximum doses were 0.05-mg and 50-mg, respectively. Optimization was performed varying the number of epochs/subject/arm over the night in 100 epoch increments from 100 to 1200 (uniformly distributed).

Simulations and Statistical Comparison

A joint simulation NLMEM was constructed from both sub-models ('From awake' and 'From asleep') and used to generate 100 datasets to simulate the sleep stages over the night. A visual predictive check (VPC) was performed to test the models' ability to properly describe the frequencies of transitions and sleep stages throughout the night. The median value and corresponding 95% confidence interval derived from the simulated datasets were plotted against the observed median across the night to evaluate the models' predictive performance. The first 8 hours were divided into 10 intervals of equal duration (48 minutes), whereas the final 5 hours were divided into 5 intervals of equal duration (60 minutes).

Simulations and re-estimations (SSE) of the simulated datasets were performed in NONMEM to compare the empirical parameter estimate precision to those via predictions from the FIM_{total} . In all SSEs, the LAPLACE option was used for estimation.

Summary statistics computed to compare the predicted parameter precision via the FIM to those of the SSE were (Equations 6-7)

$$RSE = \frac{SD(\theta_s)}{\theta_T}, \quad (20)$$

$$MRE = \frac{1}{N * \theta_T} \sum_{s=1}^N (\theta_s - \theta_T), \quad (21)$$

where the RSE is the relative standard error, MRE is the mean relative error, N is the number of simulated and re-estimated datasets, s is the s^{th} simulated data set ($s = 1, 2, \dots, N$), θ_s is a parameter estimate from the s^{th} data set, θ_T is the true parameter value used in the simulation, and SD is the standard deviation of the estimates. The RSE measures the precision of the parameters and the MRE represents the size of the bias. For an unbiased estimator, the RSE is asymptotically linked to the square root of the diagonal of the inverse FIM and the MRE should be 0.

The observed improvement of one design versus another based on these SSEs, $D_{\text{eff,Obs}}$, were computed using the efficiency calculation in equation 11 where the determinant of the inverse empirical covariance matrix, $|COV|^{-1}$, was used in place of $|FIM_{\text{total}}(\Theta, X)|$.

Results

NLMEM for Phase Advanced Subjects

For the two sub-models, three break points (BPB, BPC and BPD) were included in the 'From awake' sub-model and one break point (BPB) was included in the 'From asleep' sub-model. The estimated break points, TPs logit values, and IIV on the TP logits are listed in Table 1. Figure 1 illustrates the VPC for the frequency of transitions and sleep stages. The VPC plots show a good agreement between the observed and the simulated statistics as the observed profiles generally fall within the confidence intervals obtained from the simulations. Overall, the VPC provide evidence that the final NLMEM effectively describes the progression of sleep transitions along the night for phase advanced subjects after placebo dosing.

The linear and non-linear dose effect parameter estimates added to the TPs are listed in Table 1. Figure 2 demonstrates the effect of these dose effect parameters for dosages of 0.1-, 1-, 6-, 10-, and 20-mg on the TPs and sleep stages.

Fig. 1. Visual predictive check on frequency of transitions (first two rows) and stage frequencies (last row).

Table 1. Model parameter values for NLMEM.

From model	Parameters	Parameter estimates for transition to awake						
awake	Logits	TP1 ^a	TP2 (BPB)	TP3 (BPC)	TP4 (BPD)	TP5 ^c		
		2.36	2.09	1.96	1.06	2.92		
	BP	BPB ^c	BPC	BPD ^d				
		0.462	901	0.536				
	Dose Effect	D ₅₀ ^e (TP1)	E _{max} ^e (TP1)	D ₅₀ ^e (TP2)	E _{max} ^e (TP2)	D ₅₀ ^e (TP3)	E _{max} ^e (TP3)	Slope ^f (TP5)
		1.60	-0.311	5.13	-1.83	0.0943	-0.491	-0.186
	σ^2_{g}	TP1	TP2	TP3	TP4	TP5		
		0.963	2.09	1.73	1.44	1.89		
asleep	Logits	TP1 ^a	TP2 (BPB)	TP3 ^b				
		-2.58	-3.71	-3.12				
	BP	BPB ^c						
		0.405						
	Dose Effect	Slope ^f (TP2)	D ₅₀ ^e (TP3)	E _{max} ^e (TP3)				
		-0.124	0.466	-0.0948				
	σ^2_{g}	TP1	TP2	TP3				
		2.05	0.474	0.495				

Abbreviations: BP – break points, IIV – inter-individual variability; TP – transition probability.

^a refers to the beginning of the night

^b refers to the last observation

^c $BPB = (BPC - BPA) \cdot \theta_n + BPA$

^d $BPD = (BPE - BPC) \cdot \theta_n + BPC$

^e $TP_n = \theta_n \cdot (1 + E_{max} \cdot Dose / (D_{50} + Dose))$; where TP_n is the transition probability at a particular break point, θ_n is the typical logit estimate, and the unit for D_{50} is mg.

^f $TP_n = \theta_n + Slope \cdot Dose$; where TP_n is the transition probability at a particular break point and θ_n is the typical logit estimate.

^g variance estimate for IIV.

Design evaluation

Evaluation of only placebo dosing in 27 subjects demonstrated the expected increase in the determinant of FIM_{total} with each successive increase in the number of epochs/subject for both the ‘Average Probability Weighted FIM’ and ‘Equally Weighted FIM’ (data not shown). The predicted precision of parameter estimates generally improved with the successive increase of the number of epochs (Figure 3). However,

there was relatively little difference between the %RSE for the parameters with epochs/subject ≥ 1200 . The %RSE were not that dissimilar when using the ‘Average Probability Weighted FIM’ and ‘Equally Weighted FIM’ for the ‘From asleep’ sub-model parameters. The parameter %RSE were notably decreased when using the ‘Equally Weighted FIM’ as opposed to the ‘Average Probability Weighted FIM’ for the ‘From awake’ sub-model.

Fig 2. Typical transition probabilities for each sub-model and each sleep stage. Computation was based on the typical parameter estimates obtained for each sub-model over the night not including the variance estimates for IIV.

Fig 3. Predicted parameter %RSE via the FIM_{total} (400, 800, 1200, and 1560 epochs/subject) and SSE (1560 epochs/subject) based on placebo dosing in 27 subjects.

The %RSE resulting from SSE were generally less than the predicted %RSE via the FIM_{total} based on 1560 epochs/subject (Figure 3). The precision of the TPs, break point and IIV parameters from SSE were reasonably good (21 to 38 %RSE) and relatively unbiased (-2 to 10 %MRE).

Evaluation of the reference designs demonstrated that increasing the number cross-over arms markedly decreased the predicted %RSE for D_{50} parameters in both sub-models, E_{max} (TP1) in the ‘From awake’ sub-model and E_{max} (TP3) ‘From asleep’ sub-model (Figure 4). For all other parameters a relatively small decrease in the predicted %RSE with increase in the number of cross-over arms was observed (Figure 4). The determinants of FIM_{total} based on the ‘Equally Weighted FIM’ were greater than

the ‘Average Probability Weighted FIM’ (data not presented), but the increase in the predicted efficiency with an increase in the number cross-over arms was relatively consistent with both the ‘Equally Weighted FIM’ and the ‘Average Probability Weighted FIM’ (Table 2). The relative magnitude of change of the predicted efficiency decreased with an increase in the number of cross-over arms: 1) going from a 2-way cross-over to a 3-way cross-over resulted in a 47-49 point difference, 2) going from a 3-way cross-over to a 4-way cross-over resulted in a 39-40 point difference, 3) going from a 4-way cross-over to a 5-way cross-over resulted in a 31-32 point difference, and 4) going from a 5-way cross-over to a 6-way cross-over resulted in a 23-25 point difference.

Fig 4. Predicted parameter %RSE via the FIM_{total} based on ‘Average Probability Weighted FIM’ and ‘Equally Weighted FIM’ for the reference designs based on 1200 epochs/subject/arm.

The parameter %RSE based on SSE were generally less than or similar to the predicted %RSE via the FIM_{total} (Figure 5; 4-way crossover design presented). However, several dose related effect parameter %RSE were greater than predicted via the FIM_{total} . The TPs, break point and IIV parameters from SSE were relatively unbiased (-3 to 12 %MRE). In addition, all dose related effect parameters except for D_{50} (TP1 and TP3) in the ‘From awake’ sub-model and D_{50} (TP3) in the ‘From asleep’ sub-model were relatively unbiased (-0.1 to 23 %MRE) across all crossover designs. Once the number of cross-overs was greater than 2, D_{50} (TP3) in the ‘From awake’ sub-model was relatively unbiased (-2.8 to 12 %MRE). As the number of crossovers

increased, the bias for D_{50} (TP1) in the ‘From awake’ sub-model and D_{50} (TP3) in the ‘From asleep’ sub-model decreased, but still displayed relatively high bias; D_{50} (TP1) in the ‘From awake’ sub-model decreased from 208%MRE (2-way crossover) to 64 %MRE (6-way crossover) and D_{50} (TP3) in the ‘From asleep sub-model decreased from 283 %MRE (2-way crossover) to 41 %MRE (6-way crossover). The observed efficiency was more pronounced than predicted via FIM_{total} with an increase in the number of cross-overs, but the trends in efficiency were the same; the relative magnitude of change decreased with an increase in the number of cross-over similar to the predicted efficiency via FIM_{total} (Table 2).

Table 2. Efficiency of the reference crossover designs.

Crossover	Group	# Subjects	Dosage (mg)	‘Average Probability Weighted FIM’ $D_{eff,Pred}^a$ (%)	‘Equally Weighted FIM’ $D_{eff,Pred}^a$ (%)	$D_{eff,Obs}^b$ (%)
2	1	7	0.00, 20.00	100	100	100
	2	7	0.10,0.00			
	3	7	1.00,0.10			
	4	7	6.00,1.00			
	5	7	10.00,6.00			
	6	7	20.00,10.00			
3	1	7	0.00, 20.00, 10.00	149	147	194
	2	7	0.10,0.00, 20.00			
	3	7	1.00,0.10, 0.00			
	4	7	6.00,1.00, 0.10			
	5	7	10.00,6.00, 1.00			
	6	7	20.00,10.00, 6.00			
4	1	7	0.00, 20.00, 10.00, 6.00	189	186	237
	2	7	0.10,0.00, 20.00, 10.00			
	3	7	1.00,0.10, 0.00, 20.00			
	4	7	6.00,1.00, 0.10, 0.00			
	5	7	10.00,6.00, 1.00, 0.10			
	6	7	20.00,10.00, 6.00, 1.00			
5	1	7	0.00, 20.00, 10.00, 6.00, 1.00	220	218	277
	2	7	0.10,0.00, 20.00, 10.00, 6.00			
	3	7	1.00,0.10, 0.00, 20.00, 10.00			
	4	7	6.00,1.00, 0.10, 0.00, 20.00			
	5	7	10.00,6.00, 1.00, 0.10, 0.00			
	6	7	20.00,10.00, 6.00, 1.00, 0.10			
6	1	7	0.00, 0.10, 1.00, 6.00, 10.00, 20.00	245	241	307
	2	7	0.00, 0.10, 1.00, 6.00, 10.00, 20.00			
	3	7	0.00, 0.10, 1.00, 6.00, 10.00, 20.00			
	4	7	0.00, 0.10, 1.00, 6.00, 10.00, 20.00			
	5	7	0.00, 0.10, 1.00, 6.00, 10.00, 20.00			
	6	7	0.00, 0.10, 1.00, 6.00, 10.00, 20.00			

^a Estimates were determined using Equation 19 relative to the 2-way crossover design based on 1200 samples/subject/arm using the corresponding weighting scheme of the FIM_{total} .

^b Estimates were determined using Equation 19 where the $|COV^{-1}|$ from SSE replaced $|FIM(\theta, \bar{X})|$ relative to the 2-way crossover design.

Optimized Designs

The D-optimal designs for the dose and the number of subjects per group for the different crossover designs was examined with the number of epochs/subject/arm increasing from 100 to 1200. Regardless of the number of epochs/subject/arm for each individual cross-over

design and weighting scheme, the optimized dose (within [0.01] units) and number of subjects per group (Tables 3 and 4) did not change. However, with each successive increase in the number of epochs/subject/arm (100 epoch increments) or number of cross-over arms, the time to optimize each design increased by ~1-fold. These optimized designs had relatively little

Fig 5. Parameter %RSE for the reference 4-way crossover design based on 'Average Probability Weighted FIM' and 'Equally Weighted FIM' (1200 epochs/subject/arm) and SSE.

effect on the predicted %RSE for the TPs, break point and IIV parameters (2 to -5 %RSE difference) compared to the reference designs. However, the optimized designs, despite a few instances for the D_{50} parameters in the 'From awake' sub-model and E_{max} (TP3) in the 'From asleep' sub-model, generally improved the predicted precision of parameter estimates for the dose related parameters (Figure 6). The few instances where the precision of the parameters was not improved occurred when the number of cross-overs was less than 4. Once the number of cross-overs was 4 or greater, the predicted precision for these parameters improved (Figure 6). As the number of cross-overs increase, the relative decrease in the predicted %RSE generally diminished. Similar to the reference designs, the determinants of FIM_{total} based on the 'Equally Weighted FIM' were greater than the 'Average Probability Weighted FIM' (data not presented). The increase in the predicted efficiency relative to the reference designs when dose was optimized was relatively consistent with both the 'Equally Weighted FIM' and the 'Average Probability Weighted FIM' (Table 3). However, optimization of the number of subjects per group led to greater predicted efficiency based on the 'Average Probability Weighted FIM' than the 'Equally Weighted FIM' (Table 4). As the number of cross-over arms increased, the efficiency decreased for both the 'Equally Weighted FIM' and the 'Average Probability Weighted FIM' (Tables 3 and 4).

The parameter %RSE based on SSE displayed a generally similar trend to the predicted %RSE via the FIM_{total} as was observed for the reference designs; most parameter %RSE were generally less than or similar to the predicted %RSE except for several dose related

effect parameters (data not presented). The TPs, break point and IIV parameters bias and %RSE from SSE based upon the D-optimal designs for both weighting schemes were relatively unchanged compared to the corresponding reference design; difference of -3.1 to 2.1 %MRE and -3.7 to 1.8 %RSE. All dose related effect parameters except for D_{50} (TP1 and TP3) in the 'From awake' sub-model and D_{50} (TP3) in the 'From asleep' sub-model were relatively unbiased (-0.3 to 19 %MRE) across all crossover designs. Once the number of cross-overs was greater than 2, D_{50} (TP3) in the 'From awake' sub-model was relatively unbiased (-2.6 to 11 %MRE). Generally, the %MRE and %RSE decreased for the dose related effect parameters in both sub-models with the D-optimal designs for both weighting schemes when contrasted to the reference designs. However, there was a greater decrease in %RSE when using the D-optimal designs based on the 'Equally Weighted FIM' (Figure 7; 4-way crossover design presented). Nevertheless, for cross-overs less than 5, D_{50} (TP1) in the 'From awake' sub-model and D_{50} (TP3) in the 'From asleep' sub-model still displayed relatively high bias (> 50%) and %RSE (> 240%) using the D-optimal designs. The dose effect parameter D_{50} (TP1) in the 'From awake' sub-model decreased from 208%MRE (2-way crossover) to 64 %MRE (6-way crossover) and the dose effect parameter D_{50} (TP3) in the 'From asleep sub-model decreased from 283 %MRE (2-way crossover) to 41 %MRE (6-way crossover). Once the number of cross-overs was 5 or 6, the bias and %RSE decreased for D_{50} (TP1) in the 'From awake' sub-model and D_{50} (TP3) in the 'From asleep' sub-model relative to a lower number of cross-overs, but was still high; bias (> 29%) and %RSE (> 103%).

Table 3. D-optimal design (dose) and efficiency for ‘Average Probability Weighted FIM’ (APWF) and ‘Equally Weighted FIM’ (EWF) crossover designs.

Crossover	Group	# Subjects	Dosage (mg) (‘APWF’)	Dosage (mg) (‘EWF’)	‘APWF’ D _{eff,Pred} ^a (%)	‘EWF’ D _{eff,Pred} ^a (%)
2	1	7	0.00, 0.96	0.00, 34.07	134	141
	2	7	0.07,19.91	0.09,0.00		
	3	7	0.96,0.96	0.56,50.00		
	4	7	0.96,50.00	4.21,0.09		
	5	7	19.91,0.00	34.07,0.56		
	6	7	50.00,0.07	50.00,4.21		
3	1	7	0.00, 0.34, 0.34	0.00, 4.72, 50.00	124	124
	2	7	0.05,3.43, 0.00	0.09,0.00, 4.72		
	3	7	0.34,0.34, 41.47	0.96,0.09, 0.00		
	4	7	0.34,41.47, 0.05	4.72,50.00, 40.33		
	5	7	3.43,0.34, 0.00	40.33,0.96, 0.09		
	6	7	41.47,0.05, 3.43	50.00,40.33, 0.96		
4	1	7	0.00, 50.00, 1.87, 0.20	0.00, 50.00, 4.43, 34.37	119	121
	2	7	0.05,0.00, 50.00, 1.87	0.08, 0.00, 50.00, 4.43		
	3	7	0.20,24.27, 0.05, 0.00	0.74, 0.08, 0.00, 50.00		
	4	7	1.87,0.20, 24.27, 0.05	34.37, 0.74, 0.08, 0.00		
	5	7	24.27,0.05, 0.00, 50.00	4.43, 34.37, 0.74, 0.08		
	6	7	50.00,1.87, 0.20, 24.27	50.00, 4.43, 34.37, 0.74		
5	1	7	0.00, 39.17, 3.30, 0.49, 0.05	0.00, 0.08, 4.56, 50.00, 0.69	118	119
	2	7	0.05,0.00, 39.17, 3.30, 0.49	34.36, 0.00, 0.08, 4.56, 50.00		
	3	7	0.05,0.05, 0.00, 39.17, 3.30	0.69, 34.36, 0.00, 0.08, 4.56		
	4	7	0.49,0.05, 0.05, 0.00, 39.17	50.00, 0.69, 34.36, 0.00, 0.08		
	5	7	3.30,0.49, 0.05, 0.05, 0.00	4.56, 50.00, 0.69, 34.36, 0.00		
	6	7	39.17,3.30, 0.49, 0.05, 0.05	0.08, 4.56, 50.00, 0.69, 34.36		
6	1	7	0.00, 0.05, 0.21, 1.93, 24.55, 50.00	0.00, 0.08, 0.67, 4.50, 34.69, 50.00	116	118
	2	7	0.00, 0.05, 0.21, 1.93, 24.55, 50.00	0.00, 0.08, 0.67, 4.50, 34.69, 50.00		
	3	7	0.00, 0.05, 0.21, 1.93, 24.55, 50.00	0.00, 0.08, 0.67, 4.50, 34.69, 50.00		
	4	7	0.00, 0.05, 0.21, 1.93, 24.55, 50.00	0.00, 0.08, 0.67, 4.50, 34.69, 50.00		
	5	7	0.00, 0.05, 0.21, 1.93, 24.55, 50.00	0.00, 0.08, 0.67, 4.50, 34.69, 50.00		
	6	7	0.00, 0.05, 0.21, 1.93, 24.55, 50.00	0.00, 0.08, 0.67, 4.50, 34.69, 50.00		

^a Estimates were determined using Equation 19 relative to the corresponding reference cross-over design based on 1200 samples/subject/arm using the corresponding weighting scheme of the FIM_{total}.

Table 4. D-optimal design (number of subject/group) and efficiency for ‘Average Probability Weighted FIM’ (APWF) and ‘Equally Weighted FIM’ (EWF) crossover designs.

Crossover	Group	# Subjects (‘APWF’)	# Subjects (‘EWF’)	Dosage (mg)	‘APWF’ D _{eff,Pred} ^a (%)	‘EWF’ D _{eff,Pred} ^a (%)
2	1	23	25	0.00, 20.00	122	114
	2	3	1	0.10,0.00		
	3	13	11	1.00,0.10		
	4	1	1	6.00,1.00		
	5	1	1	10.00,6.00		
	6	1	3	20.00,10.00		
3	1	1	14	0.00, 20.00, 10.00	119	109
	2	25	19	0.10,0.00, 20.00		
	3	13	4	1.00,0.10, 0.00		
	4	1	3	6.00,1.00, 0.10		
	5	1	1	10.00,6.00, 1.00		
	6	1	1	20.00,10.00, 6.00		
4	1	1	1	0.00, 20.00, 10.00, 6.00	114	107
	2	6	13	0.10,0.00, 20.00, 10.00		
	3	25	25	1.00,0.10, 0.00, 20.00		
	4	8	1	6.00,1.00, 0.10, 0.00		
	5	1	1	10.00,6.00, 1.00, 0.10		
	6	1	1	20.00,10.00, 6.00, 1.00		
5	1	1	1	0.00, 20.00, 10.00, 6.00, 1.00	108	104
	2	1	1	0.10,0.00, 20.00, 10.00, 6.00		
	3	13	21	1.00,0.10, 0.00, 20.00, 10.00		
	4	25	17	6.00,1.00, 0.10, 0.00, 20.00		
	5	1	1	10.00,6.00, 1.00, 0.10, 0.00		
	6	1	1	20.00,10.00, 6.00, 1.00, 0.10		

^a Estimates were determined using Equation 19 relative to a the corresponding reference cross-over design based on 1200 samples/subject/arm using the corresponding weighting scheme of the FIM_{total}.

Fig 6: Predicted difference in parameter %RSE for optimized designs relative to the reference designs via the FIMtotal based on 'Average Probability Weighted FIM' and 'Equally Weighted FIM' based on 1200 epochs/subject/arm.

Overall, the observed efficiency compared to the corresponding reference design was not that dissimilar than predicted via FIM_{total} (Table 3-5), and D-optimal designs based on the 'Equally Weighted FIM' were more efficient than based on the 'Average Probability Weighted FIM' (Table 5).

Discussion

The application of experimental design theories for NLMEM to improve efficiencies and parameter precision has become increasingly widespread since Mentré et al. derived a first-order analytic expression of the Fisher Information Matrix (FIM) for continuous data²⁰. Optimization of design variables for homogenous discrete type responses has recently been explored²¹⁻²³, and the work presented here provides a complement by examining non-homogenous discrete type responses. The FIM_{total} approximation provided an efficient method for design evaluation and trial optimization for non-homogeneous discrete type response NLMEMs as exemplified by a dichotomous, non-homogeneous, Markov-chain phase advanced sleep NLMEM.

By employing the approximate analytic solution and different weighting schemes for the FIM_{total} , the optimized doses and number of subjects/group were more efficient and provided less NLMEM parameter uncertainty than the reference designs. The constructs of the optimization was premised on theoretical linear and non-linear dose effects and optimization of the NLMEM with actual dose effect values based on an observed clinical trial will likely result in a

different design paradigm. Nevertheless, the methods presented provide an opportunity to evaluate different designs and optimize design variables to assess the impact on parameter uncertainty. For instance, optimizing both the dose and the number of subjects/group for a 5-way cross-over design based on the 'Equally Weighted FIM' approach, further increase the efficiency of the trial design relative to optimizing either the dose or the number of subjects/group alone, 103% or 107%, respectively. In addition, this optimized 5-way cross-over design is nearly as efficient as the dose optimized 6-way cross-over design (91%). These considerations can make a major impact in reducing the associated costs for these types of trials.

However, as noted previously, the time to optimize each design increased as more components (number of epochs/subject/arm or cross-over arms) are added. Fortunately, regardless of the number of epochs/subject/arm for each individual cross-over design, the optimized dose and number of subjects per group were consistent, and using a reduced number of epochs/subject/arm can decrease runtimes when examining more complex design variable optimization. Therefore, additional design variables that could be optimized to provide more robust designs, including covariates (smoking or alcohol habits, weight, and gender) that have demonstrated pronounced effect in sleep patterns^{19,34,35} or additional cross-over arms, can be examined without overtly increasing the runtimes. For example, the high uncertainty for the theoretical dose effect parameters D_{50} (TP1) in the 'From awake' sub-model and D_{50} (TP3) in

Fig 7: Observed parameter %RSE from SSE for the optimized 4-way crossover designs based on 'Average Probability Weighted FIM' and 'Equally Weighted FIM' and reference design.

the 'From asleep' sub-model could be merely a consequence of the limit of the number of crossovers explored, and the impact of additional cross-over arms could be explored with modest increase in runtime. For this situation, a dose and number of subjects/group optimized 9-way crossover design based on the 'Equally Weighted FIM' approach using 100 epochs/subject/arm was obtained reasonably quick, and SSEs demonstrated that the parameters still had poor precision (> 62 %RSE). The SSE revealed that including these additional cross-over arms might not be cost effective in relation to the minor decrease in parameter uncertainty.

In theory, the precision for the parameter estimates predicted by the FIM_{total} should be an asymptotic lower bound of the estimated parameter precision due to the Cramer-Rao inequality³⁰. However, this was generally not observed for most of the parameters estimated in the sleep model, a feature that is most evident after placebo dosing (Figure 3). The analytical approximation of the FIM_{total} and the lack of correlation between fixed and random effects could account for the higher predicted parameter uncertainty obtained via the FIM_{total} relative to SSE. During SSE, the conditional estimates for IIV during the computation of the parameter estimation are used and account for the correlation between fixed and random effects within each sub-model. Another factor that possibly contributed to higher predicted parameter uncertainty obtained via the FIM_{total} relative to SSE could be the FIM_{total} weighting schemes. As the average probability of being awake was $\sim 20\%$ over the night after placebo dosing, the 'From awake' sub-model made a relatively small contribution to the FIM_{total} based on the 'Average Probability Weighted FIM' approach. This small contribution to the FIM_{total}

led to higher predicted uncertainty for the parameters in the 'From awake' sub-model than predicted based on the 'Equally Weighted FIM' approach (Figure 3). However, the predicted parameter uncertainty based on the 'Equally Weighted FIM' approach more closely approximated the SSE parameter uncertainty computations. These results demonstrate the importance of assigning the weighting scheme to the FIM_{total} for design optimization. Finally, the Cramer-Rao inequality holds for an unbiased estimator, which is not always the case for the Laplace methods utilized by NONMEM³⁶ and was demonstrated for some of the parameters. This is especially evident in D_{50} (TP1) in the 'From awake' sub-model and D_{50} (TP3) in the 'From asleep' sub-model.

In addition to the possible improvements mentioned above to provide more robust trial designs in phase-advance sleep studies, optimal design could be expanded to non-homogeneous, discrete sleep multinomial Markov-chain models^{18,19}. For this type of model, the sleep stages evaluated are partitioned into five different sleep states: awake, sleep stage 1, sleep stage 2, slow wave sleep and REM sleep employing a temporal dependence of transition probabilities via piecewise linear functions. Implementing this model for optimal design is complicated by the implementation of the FIM weighting. In the current work, one weighting method employed ('Average Probability Weighted FIM') was based on calculating the average probability of awake, $p(0)$, and computing '1- $p(0)$ ' to compute the probability of being asleep. For the five different sleep states, Bayes Theorem for multinomial discrete values could be used to compute the average probability of each sleep stage. Additionally, partitioning the FIM into smaller segments, e.g. based on each breakpoint within

Table 5. Observed (SSE) efficiency for D-optimal crossover designs relative to the reference designs.

Crossover	D _{eff,Obs} ^a (%)			
	D-optimal design for dose		D-optimal design for number of subject/group	
	'Average Probability Weighted FIM'	'Equal Weighted FIM'	'Average Probability Weighted FIM'	'Equal Weighted FIM'
2	122	136	115	122
3	116	124	111	116
4	113	119	107	112
5	109	117	104	108
6	105	115	--	--

^a Estimates were determined using Equation 19 where the $|\text{COV}^{-1}|$ from SSE replaced $|\text{FIM}(\vec{\theta}, \vec{X})|$ relative to the corresponding reference crossover design.

each sub-model, to impact the weighting might improve the precision of the parameter estimates and lead to more globally optimized design variables. These implementations may prove more useful as the multinomial Markov-chain model^{18,19} provide a better understanding of the underlying sleep architecture.

Conclusion

In conclusion, although the %RSE and efficiency based on $\text{FIM}_{\text{total}}$ were not equivalent to those obtained from SSE, the optimal experimental design methodologies presented for a non-homogeneous, dichotomous Markov-chain phase advance sleep NLMEM provide more efficient trial designs and decrease the NLMEM parameter uncertainty. These methodologies can be expanded to optimize trial designs for more granulated sleep models or applied to other models that include temporal dependence on event probabilities.

References

1. Pfister M and D'Argenio DZ. (2010) The emerging scientific discipline of pharmacometrics. *J Clin Pharmacol* 2010, 50(9 Suppl):6S.
2. Ette EI and Williams PJ. (2007) *Pharmacometrics: The Science of Quantitative Pharmacology*. Hoboken, NJ, USA: John Wiley and Sons Inc.
3. Karlsson KE. (2010) Benefits of pharmacometric Model-Based Design and Analysis of Clinical Trials, in *Acta Universitatis Upsaliensis*: Uppsala.
4. Wallin, J. (2009) Dose Adaptation Based on Pharmacometric Models, in *Acta Universitatis Upsaliensis*: Uppsala.
5. Retout S, Duffull S, and Mentre F. (2001) Development and implementation of the population Fisher information matrix for the evaluation of population pharmacokinetic designs. *Comput Methods Programs Biomed.* 65(2): p. 141-51.
6. Aarons L (1992) Population pharmacokinetics. *Int J Clin Pharmacol Ther Toxicol* 30:520–522.
7. Aarons L (1991) Population pharmacokinetics: theory and practice. *Brit J Clin Pharmacol* 32:669–670.
8. Sheiner LB, Rosenberg B, Marathe VV. (1977) Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *J Pharmacokinet Biopharm* 5:445–479.
9. Ancoli-Israel S, Roth T. (1999) Characteristics of insomnia in the United States: Results of the 1991 National Sleep Foundation Survey. I. *Sleep* 22(2): S347–S353.
10. Simon GE and VonKorff M. (1997) Prevalence, burden, and treatment of insomnia in primary care, *Am J Psychiatry* 154: 1417-1423.
11. Walsh JK and Engelhardt CL. (1999) The direct economic costs of insomnia in the United States for 1995, *Sleep* 22(2): S386-S393.
12. Stoller MK. (1994) Economic effects of insomnia. *Clin Ther* 16: 873-897 (1994).
13. Lavie P. (1986) Ultrashort sleep-waking schedule. III. 'Gates' and 'forbidden zones' for sleep. *Electroencephalogr Clin Neurophysiol.* 63(5):414-425.
14. Walsh JK, Deacon S, Dijk DJ, Lundahl J. (2007) The selective extrasynaptic GABAA agonist, gaboxadol, improves traditional hypnotic efficacy measures and enhances slow wave activity in a model of transient insomnia. *Sleep* 30:593–602.
15. Kemp B, Kamphuisen HA. (1986) Simulation of human hypnograms using a Markov chain. *Sleep* 9:405-414.
16. Karlsson MO, Schoemaker RC, Kemo B, Cohne AF, van Gerven JM, Tuk B, Peck CC, Danhof M. (2000) A pharmacodynamic Markov mixed-effects model for the effect of temazepam on sleep. *Clin Pharmacol Ther.* 68:175-88.

17. Kjellsson MC, Ouellet D, Corrigan B, Karlsson MO. (2011) Modeling Sleep Data for a New Drug in Development using Markov Mixed-Effects Models. *Pharm Res.* 28(10):2610-2627.
18. Bizzotto R, Zamuner S, De Nicolao G, Karlsson MO, Gomeni R. (2010) Multinomial logistic estimation of Markov-chain models for modeling sleep architecture in primary insomnia patients. *J Pharmacokinet Pharmacodyn.* 7:137–155.
19. Bizzotto R, Zamuner S, Mezzalana E, De Nicolao G, Gomeni R, Hooker AC, Karlsson MO. (2011) Multinomial Logistic Functions in Markov Chain Models of Sleep Architecture Internal and External Validation and Covariate Analysis. *The AAPS Journal* 13(3):445-463.
20. Mentre F, Mallet A, Baccar D. (1997) Optimal design in random-effects regression models. *Biometrika* 84(2):429–442.
21. Ogungbenro, K. and L. Aarons. (2011) Population Fisher information matrix and optimal design of discrete data responses in population pharmacodynamic experiments. *J Pharmacokinet Pharmacodyn.* 38(4): p. 449-69.
22. Nyberg J, Karlsson MO, Hooker AC. (2009) Population optimal experimental design for discrete type data. *PAGE (Population Approach Group Europe)* 18, Abstr 1468, St. Petersburg, Russia. www.page-meeting.org/?abstract=1468.
23. Nyberg J, Hooker AC. (2011) Population optimal design with correlation using Markov Models. *PAGE (Population Approach Group Europe)* 20, Abstr 2233, Athens, Greece. www.page-meeting.org/?abstract=2233.
24. Beal S, Sheiner LB, Boeckmann A, Bauer RJ (1989–2009) *NONMEM user's guides*. Icon Development Solutions, Ellicott City.
25. Breslow, N.E. and D.G. Clayton, (1993) Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association.* 88(421): p. 9-25.
26. Longford, N. (1994) Logistic regression with random coefficients. *Comput Stat Data Anal.* 17: p. 1-15.
27. Wand, M.P. (2007) Fisher information for generalised linear mixed models. *Journal of Multivariate Analysis.* 98(7): p. 1412-1416.
28. Rao, C.R. (1945) Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* 37(3): p. 81-91.
29. Cramér H. (1946) *Mathematical methods of statistics*: Princeton University Press.
30. Larsen RJ, Marx ML. (1986) *An introduction to mathematical statistics and its application*. Prentice-Hall, Upper Saddle River, p 248.
31. PopED, version 2.13 (2013) <http://poped.sf.net>.
32. Nyberg J, Ueckert S, Strömberg E, Hennig S, Karlsson MO, Hooker AC. (2012). PopED: An extended, parallelized, nonlinear mixed effects models optimal design tool. *Computer Methods and Programs in Biomedicine.* 108(2): 789-805.
33. Foracchia M, Hooker AC, Vicini P, Ruggeri A.. (2004) POPED, a software for optimal experiment design in population kinetics. *Comput Methods Prog Biomed* 74(1):29–46.
34. Nakade M, Takeuchi H, Kurotani M, Harada T. (2009) Effects of meal habits and alcohol/cigarette consumption on morningness-eveningness preference and sleep habits by Japanese female students aged 18-29. *J Physiol Anthropol.* 28(2):83-90.
35. Ma J, Dijk DJ, Svetnik V, Tymofyeyev Y, Ray S, Walsh JK, Deacon S. (2011) EEG power spectra response to a 4-h phase advance and gaboxadol treatment in 822 men and women. *Clin Sleep Med.* 7(5):493-501A.
36. Plan EL, Maloney A, Trocóniz IF, Karlsson MO. (2009) Performance in population models for count data, part I: maximum likelihood approximations. *J Pharmacokinet Pharmacodyn.* 36(4): p. 353-66.