



Research Article

A Minimal Continuous-Time Markov Pharmacometric Model

Emilie Schindler¹ and Mats O. Karlsson^{1,2}

Received 3 February 2017; accepted 1 June 2017; published online 20 June 2017

Abstract. In this work, an alternative model to discrete-time Markov model (DTMM) or standard continuous-time Markov model (CTMM) for analyzing ordered categorical data with Markov properties is presented: the minimal CTMM (mCTMM). Through a CTMM reparameterization and under the assumption that the transition rate between two consecutive states is independent on the state, the Markov property is expressed through a single parameter, the mean equilibration time, and the steady-state probabilities are described by a proportional odds (PO) model. The mCTMM performance was evaluated and compared to the PO model (ignoring Markov features) and to published Markov models using three real data examples: the four-state fatigue and hand-foot syndrome data in cancer patients initially described by DTMM and the 11-state Likert pain score data in diabetic patients previously analyzed with a count model including Markovian transition probability inflation. The mCTMM better described the data than the PO model, and adequately predicted the average number of transitions per patient and the maximum achieved scores in all examples. As expected, mCTMM could not describe the data as well as more flexible DTMM but required fewer estimated parameters. The mCTMM better fitted Likert data than the count model. The mCTMM enables to explore the effect of potential predictive factors such as drug exposure and covariates, on ordered categorical data, while accounting for Markov features, in cases where DTMM and/or standard CTMM is not applicable or conveniently implemented, e.g., non-uniform time intervals between observations or large number of categories.

KEY WORDS: NONMEM; non-linear mixed effects models; ordered categorical data; pharmacokinetic-pharmacodynamic; serial correlations.

INTRODUCTION

In clinical trials and practice, efficacy and safety evaluation often involves endpoints that are of categorical nature. In the simplest case, these endpoints are binary, e.g., responder versus non-responder and presence versus absence of a favorable or undesired event. In more complex cases, when the intensity of a variable that is not directly or easily quantifiable is of interest, ordered categorical variables are collected, for example, pain intensity, sedation grade, or adverse effect severity (e.g., none, mild, moderate, severe, or life threatening). Ordered categorical data may also arise from a categorization of a continuous variable, e.g., grade of neutropenia. In order to assess disease progression and/or

treatment effect, these variables are collected repeatedly in each patient over the course of a therapy. Depending on the frequency of collection, the outcome (state) of two consecutive assessments may not be independent but intercorrelated beyond what is predicted by taking standard predictors (dose, time, individual patient characteristics) into account. A way to handle such correlation is to make an outcome conditional on the observed state at the previous assessment in the individual (but not on the whole history). Such models are referred to as (first-order) Markov models.

Population pharmacokinetic-pharmacodynamic (PKPD) modeling has proved useful in elucidating exposure-response relationships for categorical outcomes by use of non-linear mixed effect (NLME) models. NLME models allow for simultaneous analysis of the data from all individuals in a study, description of the typical trend in the population, and quantification of between-subject variability (BSV). Ordered categorical data have mostly been characterized using proportional odds (PO) models, first introduced for NLME models by Sheiner, in which the probability of a response is described as a function of explanatory factors (drug exposure, covariates, etc.) (1). PO models are often appropriate for characterizing PKPD relationships in the absence of serial

Electronic supplementary material The online version of this article (doi:10.1208/s12248-017-0109-1) contains supplementary material, which is available to authorized users.

¹Department of Pharmaceutical Biosciences, Uppsala University, Box 591, SE-75124, Uppsala, Sweden.

²To whom correspondence should be addressed. (e-mail: mats.karlsson@farmbio.uu.se)

correlations. When potential serial correlations are identified or suspected, Markov models should be used to describe the transition between states (2).

Discrete-time Markov models (DTMM) describe the probability of transition between states from one observation to the next and the potential covariate influence, including drug effects, on these probabilities. DTMM combines the PO model with a transition model. DTMM was first introduced into PKPD models by Karlsson et al. to characterize sleep stages in insomnia patients treated with temazepam (2), and since then applied to various outcomes such as sedation scores in acute stroke (3), central nervous system side effects (4), improvement scores in rheumatoid arthritis (5), grades of proteinuria (6), fatigue, and hand-foot syndrome in cancer patients (7,8). In DTMM, the influence of the preceding state on the probability of the current score is assumed to be the same whatever the time interval is between the two assessments. DTMM is therefore appropriate when modeling data with uniform observation intervals. In DTMM, the number of transitions to model, and therefore the number of model parameters to estimate, increases dramatically with the number of categories. For computational reasons, application of DTMM for variables with large number of categories (typically more than six) may not always be feasible. In this case, although alternative models accounting for serial correlations have been proposed (9), modelers often treat the outcome as a continuous variable, thereby ignoring the true discrete nature of the data.

When observation intervals are non-uniform, continuous-time Markov models (CTMM) should be preferred. The first PKPD application of CTMM was to model the position of a tablet in the gastrointestinal tract (10–12), and it has since then been used to describe various ordered categorical outcomes, e.g., pain in animals (13), extrapyramidal side effects of antipsychotic drugs in schizophrenia (14), improvement scores in rheumatoid arthritis (15), and neuropsychiatric impairment in immunodeficiency virus (16). In CTMM models, the influence of the previous state on the probability of the current state typically decreases with increasing time between observations. For dichotomous variables, CTMM and DTMM should provide similar description of the data. When more than two categories are considered, CTMM models typically involve fewer model parameters compared to DTMM. The reason for this is the often-applied assumption that transitions only occur between neighboring states. Like a tablet moving from the stomach to colon is assumed to pass the small intestine in between, a patient reporting no side effect at one time and moderate the next time is assumed to have had mild side effect at some intermediate time. However, for categorical variables with a large number of states, CTMM may still encompass many parameters and result in complex introduction of predictors and long runtimes. We believe that a simplification of the model would help model building and exploration of exposure-response relationships.

In this work, we present a version of CTMM, the minimal CTMM (mCTMM), as an alternative to DTMM or standard CTMM. In the mCTMM, the assumption is made that the transition rate between two consecutive states is

independent on the state and governed by a single parameter, the mean equilibration time (MET). Through this assumption and a reparameterization, the Markov property can be expressed through the MET and the remaining parameters can be expressed as typically done for PO models. The application of mCTMM and its comparison to PO, DTMM, and count models are then illustrated by three real data examples: fatigue data and hand-foot syndrome (HFS) data from sunitinib-treated gastrointestinal stromal tumor (GIST) patients (7) and Likert pain score data in patients with neuropathic pain (9).

METHODS

Theory

Let Y_{ij} be the categorical response for the i^{th} individual at the j^{th} occasion. $P(Y_{ij} \geq k)$ is the probability that an observation Y_{ij} is greater than or equal to a score k , also known as the cumulative probability of score k . Figure 1 provides a schematic representation of the models described in the following for an example scale with four categories (0 to 3). In the following sections, k is assumed to take values of 0, 1, ..., K .

PO Model

In the PO model, the logit of $P(Y_{ij} \geq k)$, which is the log odds of falling into or above category k , is expressed as a function of a vector \mathbf{x} of p predictors x_{1ij}, \dots, x_{pij} (e.g., drug exposure, concomitant treatments, biomarkers) specific to the i^{th} individual at the j^{th} occasion (Eq. 1).

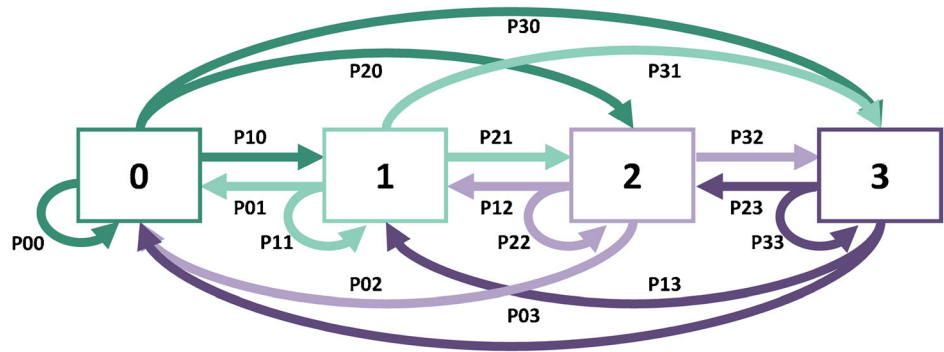
$$\text{logit}(P(Y_{ij} \geq k)) = \log\left(\frac{P(Y_{ij} \geq k)}{1 - P(Y_{ij} \geq k)}\right) = \alpha_k + f(\boldsymbol{\theta}, \mathbf{x}) \quad (1)$$

Reciprocally, $P(Y_{ij} \geq k)$ can be retrieved by the inverse of the logit function, known as the expit function ($\text{expit}(a) = \frac{e^a}{1 + e^a} = \frac{1}{1 + e^{-a}}$), as shown in Eq. 2.

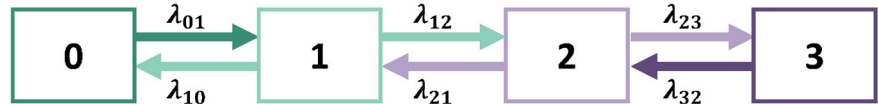
$$P(Y_{ij} \geq k) = \frac{1}{1 + e^{-(\alpha_k + f(\boldsymbol{\theta}, \mathbf{x}))}} \quad (2)$$

The intercept α_k corresponds to the log odds of scoring at least k when $f(\boldsymbol{\theta}, \mathbf{x})$ is 0. $f(\boldsymbol{\theta}, \mathbf{x})$ is a function of a vector of covariate coefficients $\boldsymbol{\theta}$ describing the effect of the predictors, i.e., covariate coefficients, on the log odds of falling into or above any category. Predictors' effects can be described as linear or non-linear functions. A major assumption of the PO model is that all categories are affected to the same extent by the predictors, i.e., the parameters $\boldsymbol{\theta}$ are not specific to any score. Note that for the lowest category ($k = 0$) $P(Y_{ij} \geq k) = 1$, and α_0 is not estimated. To ensure that $P(Y_{ij} \geq k) > P(Y_{ij} \geq k + 1)$, the constraint $\alpha_k > \alpha_{k+1}$ is required. To do so, α_k for the second lowest category ($k = 1$) is estimated with no constraint and the following parameterization is used for the higher categories: $\alpha_{k+1} = \alpha_k + b_{k+1}$, where b_{k+1} are the estimated parameters constrained to be negative. Moreover,

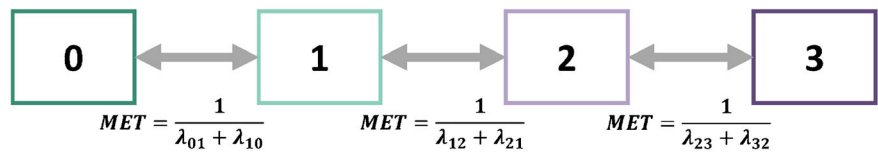
Discrete-time Markov model (DTMM)



Continuous-time Markov model (CTMM)



Minimal continuous-time Markov model (mCTMM)



Proportional odds model



Fig. 1. Schematic representation of a discrete-time Markov model (*top*), a continuous-time Markov model (*middle-top*), a minimal continuous-time Markov model (*middle-bottom*), and a proportional odds model (*bottom*), exemplified for a variable with four categories (0 to 3). In the DTMM, transition probabilities cover a finite although fixed time interval from one observation to the next; it is possible in that finite time interval to have gone from any state x to any state y . Conversely, in the CTMM, a differential transition probability (rate constant lambda) between non-neighboring states will be 0 in the infinitesimal change in time. This results in DTMM requiring a greater parameterization than CTMM. P_{yx} transition probability from state x to state y , λ_{xy} transfer rate constant from state x to state y , MET mean equilibration time

when investigating BSV on α_k , BSV terms are incorporated additively on α_1 so that the PO assumption is respected.

The probability of observing exactly score k can be derived as follows:

$$P(Y_{ij} = k) = \begin{cases} 1 - P(Y_{ij} \geq k + 1) & \text{if } k = 0 \\ P(Y_{ij} \geq k) - P(Y_{ij} \geq k + 1) & \text{if } 0 < k < K \\ P(Y_{ij} \geq k) & \text{if } k = K \end{cases} \quad (3)$$

Importantly, in the PO model, potential dependency of the current observation on the previous observation (Markov elements) is ignored.

Discrete-Time Markov Model

In the DTMM, time is viewed as a discrete variable; i.e., it takes a discrete set of values such as the observation times in a clinical study. DTMM equations are similar to the PO model but describe the transition probability from state k_{j-1} to state k_j , i.e., the probability of each state k_j at the current

occasion j given that a state k_{j-1} was observed at the previous occasion $j-1$ (Eqs. 4 and 5).

$$\begin{aligned} P(Y_{ij} = k_j | Y_{i(j-1)} = k_{j-1}) \\ = P(Y_{ij} \geq k_j | Y_{i(j-1)} = k_{j-1}) - P(Y_{ij} \geq k_j + 1 | Y_{i(j-1)} = k_{j-1}) \end{aligned} \quad (4)$$

Similarly to the PO model, the cumulative probability of score k_j , given that the previous observed score is k_{j-1} , is modeled using the expit function:

$$P(Y_{ij} \geq k_j | Y_{i(j-1)} = k_{j-1}) = \left(\frac{1}{1 + e^{-\left(\alpha_{k_j|k_{j-1}} + f(\theta, \mathbf{x}|k_{j-1})\right)}} \right) \quad (5)$$

$\alpha_{k_j|k_{j-1}}$ is the intercept of the logit of the cumulative probability of score k_j given that k_{j-1} was observed at the previous occasion. $f(\theta, \mathbf{x}|k_{j-1})$ is a function of a vector of parameters θ describing the effect of the predictors \mathbf{x} on the

log odds of falling into or above any category. $f(\boldsymbol{\theta}, \mathbf{x} | k_{j-1})$ is conditional on the previous observed score. For the lowest category ($k = 0$), $P(Y_{ij} \geq k | Y_{i(j-1)} = k_{j-1}) = 1$ and no $\alpha_{k_j | k_{j-1}}$ is therefore estimated. To ensure that $P(Y_{ij} \geq k | Y_{i(j-1)} = k_{j-1}) > P(Y_{ij} \geq k + 1 | Y_{i(j-1)} = k_{j-1})$, the constraint $\alpha_{k_j | k_{j-1}} > \alpha_{k_{j+1} | k_{j-1}}$ needs to be applied. This is achieved by estimating $\alpha_{k_j | k_{j-1}}$ for the second lowest category ($k = 1$) and by using the following parameterization for higher categories: $\alpha_{k_j+1 | k_{j-1}} = \alpha_{k_j | k_{j-1}} + b_{k_j+1 | k_{j-1}}$, where $b_{k_j+1 | k_{j-1}}$ are the negative estimated parameters. BSV terms are incorporated additively on $\alpha_{k_j | k_{j-1}}$ for the second lowest category ($k = 1$) so that the PO assumption is respected. One major assumption in the DTMM is that the transition probabilities are independent on the time between two subsequent occasions. Moreover, the score preceding the first observation is not known and is assumed to be the same as the current.

Minimal Continuous-Time Markov Model

The mCTMM is a parsimonious version of the CTMM. In the CTMM, time is regarded as a continuous variable and the probability of each score is defined by an ordinary differential equation and corresponds to a compartment amount. At any time point, the sum of the probabilities of all scores equals to 1. At the time of observation, the amount of the compartment corresponding to the actual observed score is set to 1, whereas the amounts in the compartments for all other scores are reset to 0. A set of transfer rate constants λ governs the rate at which the probability amounts distribute between the different states during the time interval between two observations. $\lambda_{k, k+1}$ represents the transfer rate constant from state k to state $k + 1$. In the previously published CTMM, the probability amounts can only transfer to adjacent states (from k to $k + 1$ or $k - 1$, but not directly from, e.g., k to $k + 2$ or $k - 2$) (Eq. 6).

$$\frac{dP(Y_{ij} = k)}{dt} = \begin{cases} \lambda_{k+1, k} \cdot P(Y_{ij} = k + 1) - \lambda_{k, k+1} \cdot P(Y_{ij} = k) & \text{if } k = 0 \\ \lambda_{k-1, k} \cdot P(Y_{ij} = k - 1) + \lambda_{k+1, k} \cdot P(Y_{ij} = k + 1) - (\lambda_{k, k-1} + \lambda_{k, k+1}) \cdot P(Y_{ij} = k) & \text{if } 0 < k < K \\ \lambda_{k-1, k} \cdot P(Y_{ij} = k - 1) - \lambda_{k, k-1} \cdot P(Y_{ij} = k) & \text{if } k = K \end{cases} \quad (6)$$

In the mCTMM, the mean equilibration time (MET) between two consecutive states is assumed to be constant across states (Eq. 7).

$$MET = \frac{1}{\lambda_{k-1, k} + \lambda_{k, k-1}} = \frac{1}{\lambda_{k, k+1} + \lambda_{k+1, k}} \quad (7)$$

The transfer rate constants λ are expressed as functions of the MET and a set of steady-state probabilities $P_{ss, k}$ (Eqs. 8 and 9).

$$\lambda_{k-1, k} = \frac{1}{MET \cdot \left(1 + \frac{P_{ss, k-1}}{P_{ss, k}}\right)} \quad (8)$$

$$\lambda_{k, k-1} = \lambda_{k-1, k} \cdot \frac{P_{ss, k-1}}{P_{ss, k}} \quad (9)$$

The steady-state probabilities $P_{ss, k}$ are defined similarly to the PO model (Eqs. 10 and 11):

$$P_{ss}(Y_{ij} \geq k) = \frac{1}{1 + e^{-(\alpha_k + f(\boldsymbol{\theta}, \mathbf{x}))}} \quad (10)$$

$$P_{ss, k} = P_{ss}(Y_{ij} = k) = \begin{cases} 1 - P_{ss}(Y_{ij} \geq k + 1) & \text{if } k = 0 \\ P_{ss}(Y_{ij} \geq k) - P_{ss}(Y_{ij} \geq k + 1) & \text{if } 0 < k < K \\ P_{ss}(Y_{ij} \geq k) & \text{if } k = K \end{cases} \quad (11)$$

The effect of explanatory factors can be investigated on $P_{ss, k}$ as described by the function $f(\boldsymbol{\theta}, \mathbf{x})$ in Eq. 10 or on MET. In the mCTMM, estimated parameters include MET,

the intercepts α_k , and the parameters $\boldsymbol{\theta}$ related to the predictors' effect. Noteworthy, conversely to DTMM, the response probability for the first observation of each individual can be estimated to $P_{ss, k}$ without the need to make any assumption regarding the previous score.

An example of the code used in NONMEM software is provided as [Supplementary material](#).

Clinical Data and Published Models

To assess the predictive performance of the mCTMM and compare it to the PO model and models with Markov components, data from previously published clinical studies were reanalyzed using the PO model and the mCTMM. A summary of the available data and three published models is provided in the following for each example.

Example 1: Fatigue

The first example consisted of fatigue score data collected as five ordinal grades from 0 (no adverse effect) to 4 (life-threatening adverse effect) according to the National Cancer Institute-Common Terminology Criteria for Adverse Events (NCI-CTCAE) (version 3.0). Data were pooled from four clinical trials including a total of 303 patients with GIST treated with placebo or sunitinib, where fatigue status was assessed daily (7). As described in the original publication, since grade 4 was rare (1%), the categories 3 and 4 were lumped into a single category (grade 3+). A total of 55,027 fatigue scores were collected with a median follow-up duration of 157 days (range, 7–687). Fatigue data were originally described by a DTMM driven by the relative change from baseline over time in soluble vascular endothelial growth factor receptor 3 (sVEGFR-3), a pharmacodynamic biomarker related to sunitinib mechanism of action.

sVEGFR-3 time-course was predicted using a published indirect-response model (17).

Example 2: Hand-Foot Syndrome

In the second example, HFS data were collected from three clinical studies including 251 sunitinib-treated GIST patients and consisted of daily NCI-CTCAE grades ranging from 0 to 3 (grade 4 was never observed) (7). A total of 39,294 HFS scores were collected with a median follow-up duration of 147 days (range, 10–540). About 74% of the patients did not experience HFS, i.e., had scores of 0 for the entire study period. The original HFS DTMM model structure was the same as the fatigue model (7).

Example 3: Likert Pain Score

In the third example, Likert pain score data were available from 231 patients with painful distal diabetic neuropathy randomized into the placebo arm of three phase III studies. Daily pain scores were recorded in the evening on the 11-point Likert pain scale and ranged from 0 to 10. A total of 22,492 scores were collected with a median follow-up duration of 124.5 days (range, 2.5–125.5). Patients were allowed to take rescue medicine (acetaminophen) during the study, and the intake was recorded daily. Available data, demographic description, and covariate representation are presented elsewhere (9).

In the modeling analysis by Plan et al. (9), the Likert pain score data were treated as counts and described by a discrete probability distribution model based on a truncated generalized Poisson distribution accounting for under-dispersion. Discrete-time Markov properties were incorporated in the model through a probability inflation for transitions of 0, ± 1 , ± 2 , and ± 3 . The probability inflation of stable score (null transitions) was expressed as a time-dependent function. The placebo effect was accounted for by an exponential decay in the mean, with a maximal effect of 19% and a half-life of onset of 28 days. Moreover, acetaminophen use (yes/no) was found to increase the mean, denoting that rescue medication was associated with higher pain scores. This model will be referred to as the count model. Details about model implementation in NONMEM can be found on DDMoRe repository (<http://repository.ddmore.eu/model/DDMODEL00000194>).

Data Analysis and Model Evaluation

All data were analyzed and simulations performed using the NLME software NONMEM version 7.3 (18). Parameter estimation was carried out using Laplacian estimation method with the likelihood option (LIKE) in the estimation record. Data pre- and post-processing, graphical visualization, and model diagnostics were assisted by R software version 3.1.1, Perl-speaks-NONMEM (PsN) toolkit version 4, and Piraña version 2.9.0 (19).

The objective function value (OFV, $-2 \cdot \log$ -likelihood) was used as a goodness-of-fit criteria to compare models. No statistical test was performed for non-nested models. Of note, the OFV of the mCTMM, as presented above, could be compared to the OFV of the PO model but not to the one of

the DTMM model as in the latter, the previous score to the first observation was assumed the same as the current score. In order to be able to compare the OFVs of the mCTMM and DTMM, an alternative mCTMM model (mCTMM_{alt}) was implemented with the same assumption as in DTMM.

Parameter uncertainty was evaluated using the relative standard errors (RSE) obtained from the NONMEM R matrix following an evaluation step using the Monte-Carlo importance sampling estimation method. To evaluate the predictive performance of each model, 100 data sets were simulated according to the realized design of the original data set. For each simulated data set, the maximum achieved score was calculated for each individual and the frequency of each maximum achieved score across individuals was computed. The distribution of the latter from the 100 simulations was compared graphically to the maximum achieved score frequencies in the original data set. Moreover, the number of transitions in each simulated data set averaged by the number of individuals was calculated and its distribution was numerically compared to the corresponding observed value. Finally, visual predictive checks (VPCs) were generated, where the observed proportion of each score over time is compared to the corresponding 95% confidence intervals (CI) obtained from 100 simulations.

Model building, covariate evaluation and exposure-response investigation were not the focus of this work; the PO models and mCTMM presented here were therefore developed to reproduce as closely as possible the random effect structure (BSV) and the implementation of predictor effects on the score probabilities as in the published DTMM (fatigue, HFS) and count model (Likert). Additive and exponential BSV was evaluated as appropriate.

RESULTS

mCTMM was successfully fitted to all example data. A description of mCTMM features is provided below for each example. Table I compares the OFV, number of parameters, and average number of simulated transitions per individual obtained from the mCTMM, the PO model, and the DTMM (fatigue and HFS) or the count model (Likert). An improvement (decrease) in OFV was observed in all three examples when accounting for Markov elements; i.e., DTMM or count model and mCTMM always performed better than PO models. Other goodness-of-fit criteria, such as Akaike criteria or Schwarz criteria, may be calculated based on the information provided in Table I. Moreover, when simulating from the PO model, the number of transition per individual was consistently overpredicted compared to the observed data, by a factor of approximately 18, 8.8, and 1.4 in the fatigue, HFS, and Likert examples, respectively. This is also evident from Fig. 2, which exemplifies the observed and simulated HFS score versus time profiles. Figure 3 illustrates for each model and example the percentage of patients achieving a given maximum score. VPCs (Figs. 4, 5, and 6) showed a good predictive performance of the DTMM and count models, whereas PO models in all three examples had poor predictive performances of the trend in proportion of each score over time. VPCs of the mCTMM models are discussed in the following.

Table I. Data and model characteristics in the fatigue, hand-foot syndrome, and Likert pain score examples

	Observed	Proportional odds	mCTMM	DTMM	Count (Poisson)
Fatigue					
OFV	–	55,900	8382/7999 ^a	6766	–
No. of parameters (fixed/random effects)	–	5 (4/1)	7 (5/2)	20 (16/4)	–
Average no. of transitions per ID ^b	1.90	34.7 [28.7–40.0]	3.19 [2.63–3.94]	1.91 [1.68–2.15]	–
Hand-foot syndrome					
OFV	–	24,286	4784/4471 ^a	3802	–
No. of parameters (fixed/random effects)	–	5 (4/1)	7 (5/2)	19 (16/3)	–
Average no. of transitions per ID ^b	1.35	11.9 [8.20–15.7]	1.13 [0.770–1.48]	1.13 [0.821–1.50]	–
Likert pain score					
OFV	–	60,071	48,902	–	49,403
No. of parameters (fixed/random effects)	–	15 (13/2)	18 (14/4)	–	23 (12/11)
Average no. of transitions per ID ^b	41.3	58.7 [56.5–60.9]	41.9 [39.0–44.8]	–	66.8 [65.1–68.3]

DTMM discrete-time Markov model, ID individual, mCTMM minimal continuous-time Markov model, OFV objective function value
^aThe first OFV value corresponds to the mCTMM as presented in “Methods” section where the probability of the first score is estimated to the steady-state probability ($P_{ss,k}$) and can be compared to the PO model OFV. The second OFV value corresponds to the alternative mCTMM (mCTMM_{alt}), where the previous score to the first observation was assumed the same as the current score and can be compared to the DTMM OFV.
^bFor all models, mean [95% confidence interval] of the average number of transitions per individual, calculated based on 100 simulated data sets

Example 1: Fatigue

In the mCTMM for fatigue, the steady-state cumulative probabilities $P_{ss}(Y_{ij} \geq k)$ were a function of the model-predicted relative change in sVEGFR-3 from baseline over time ($sVEGFR3_{rel}(t)$) (Eq. 12).

$$P_{ss}(Y_{ij} \geq k) = \frac{1}{1 + e^{-(\alpha_k + \theta_{sVEGFR3} \cdot sVEGFR3_{rel}(t))}} \quad (12)$$

$\theta_{sVEGFR3}$ corresponds to the slope of the linear $sVEGFR3_{rel}(t)$ effect on the logit scale. α_k is as described in the “Methods” section. α_1 was estimated to be -1.85 , and the standard deviation ω_{α_1} of its additive BSV to 0.770 , denoting a large BSV in the score distribution in the absence of drug (as reflected by sVEGFR-3 changes). Consistent with the previously published DTMM (7), larger decreases in sVEGFR-3 were associated with increased steady-state probability of experiencing fatigue ($\theta_{sVEGFR3}$ estimated to

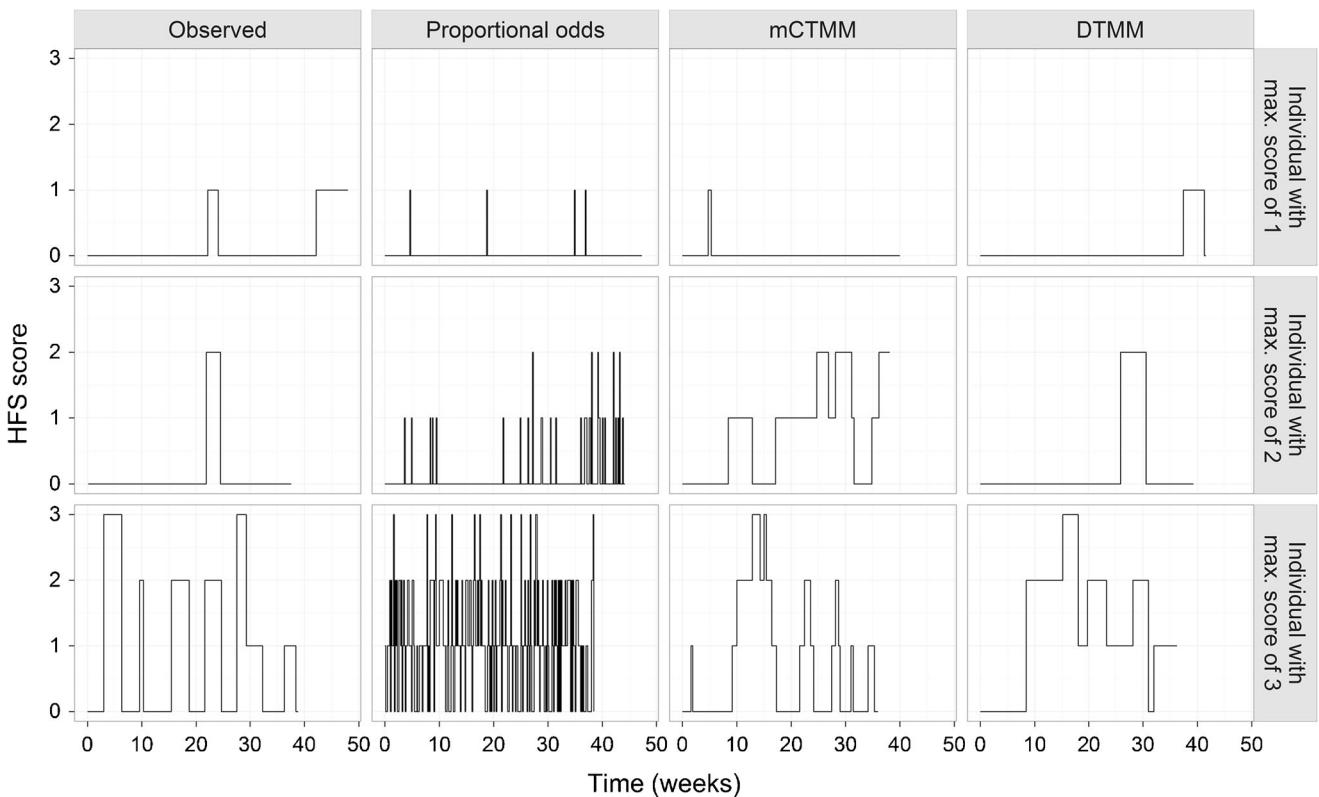


Fig. 2. Hand-foot syndrome (HFS) score versus time profiles with maximum scores of 1, 2, and 3, in the observed data set as compared to simulations from the proportional odds model, minimal continuous-time Markov model (mCTMM), and discrete-time Markov model (DTMM)

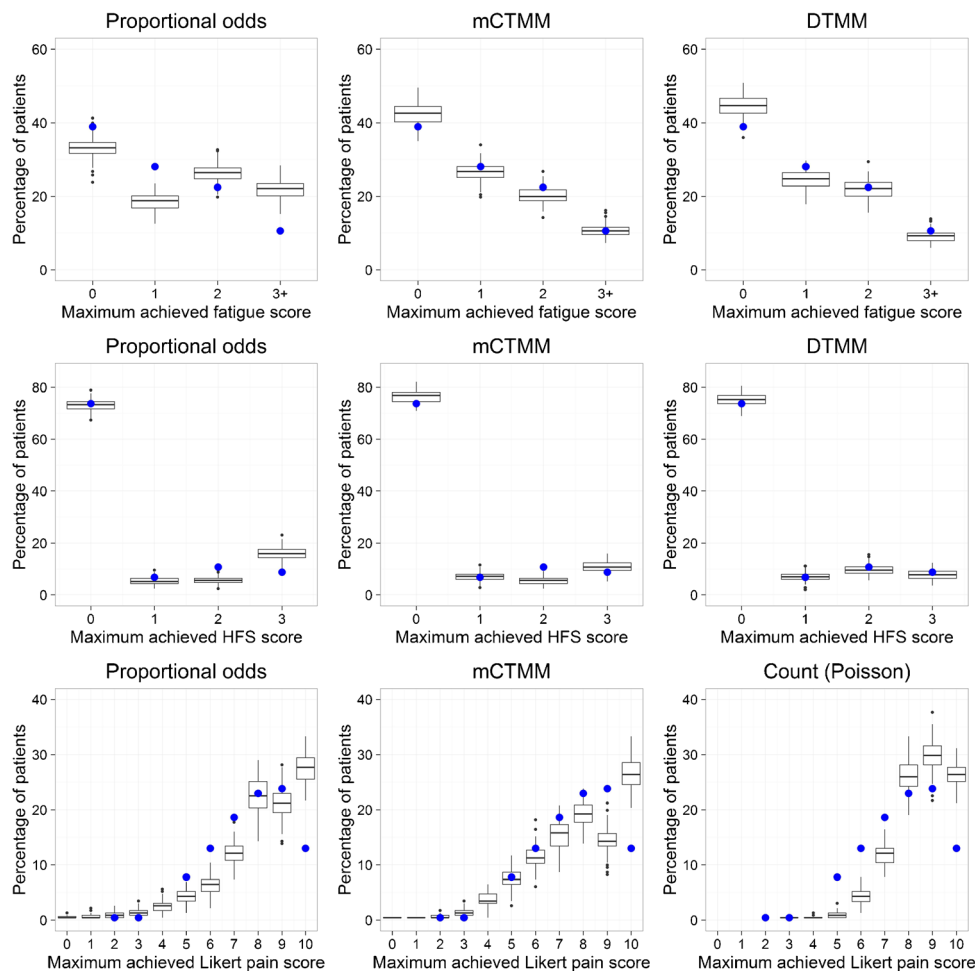


Fig. 3. Simulation-based diagnostic plots of the proportional odds model (*left*), minimal continuous-time Markov model (mCTMM, *center*), and discrete-time Markov (DTMM) or count models (*right*) for the fatigue (*top*), hand-foot syndrome (HFS, *middle*), and Likert pain score (*bottom*) examples. For each possible score, the percentage of patients achieving this score in the observed data set (*blue dots*) and in the simulated data sets (*box plots* based on 100 samples) are compared

be -1.87). MET was estimated to be 37.9 days. All CTMM parameters were estimated with reasonable uncertainty ($\leq 19\%$, Table II).

DTMM provided a better fit to the data (lower OFV) than mCTMM. However, mCTMM was more parsimonious, with seven estimated parameters versus 20 for DTMM. The average number of transitions per individual was best predicted by the DTMM and slightly overpredicted by the mCTMM. Both mCTMM and DTMM well described the proportion of patients with maximum achieved scores, whereas the PO model tended to underpredict the proportion of patients with maximum achieved scores of 0 and 1 and overpredict the proportion of patients with maximum achieved scores of 2 and 3+ (Fig. 3). VPCs (Fig. 4) of the mCTMM model showed at early time points a slight misprediction of the proportion of scores 0 and 1 and at late time points an underprediction of the proportion of scores of 1 (as for DTMM), but overall an appropriate description of the trend over time.

Example 2: Hand-Foot Syndrome

The mCTMM model for HFS was similar to the fatigue example. As for fatigue, more pronounced reduction in

sVEGFR-3 levels was associated with larger steady-state probabilities of developing HFS, which is consistent with previous finding with DTMM (7). MET was estimated to be 16.6 days. The uncertainty was below 30%RSE for all parameters except α_1 (45%RSE) and the associated BSV (90%RSE) (Table II).

HFS data were best fitted by the DTMM, which gave the lowest OFV but required 19 parameters versus seven for mCTMM. Both mCTMM and DTMM adequately predicted the average number of transitions per individual (Table I). This is also illustrated in Fig. 2, where for three example individuals with maximum achieved scores of 1, 2, and 3, the transition patterns simulated from DTMM and mCTMM were much more similar to the observed as compared to the PO model. Figure 3 shows that DTMM provided the best description of the maximum achieved scores. mCTMM tended to overpredict the proportion of patients with a maximum score of 0, and slightly underpredicted the maximum scores of 2, whereas the PO model accurately predicted the low scores of 0 and 1, but the proportions of patients with maximum scores of 2 and 3 were underpredicted and overpredicted, respectively. VPCs (Fig. 5) of the mCTMM

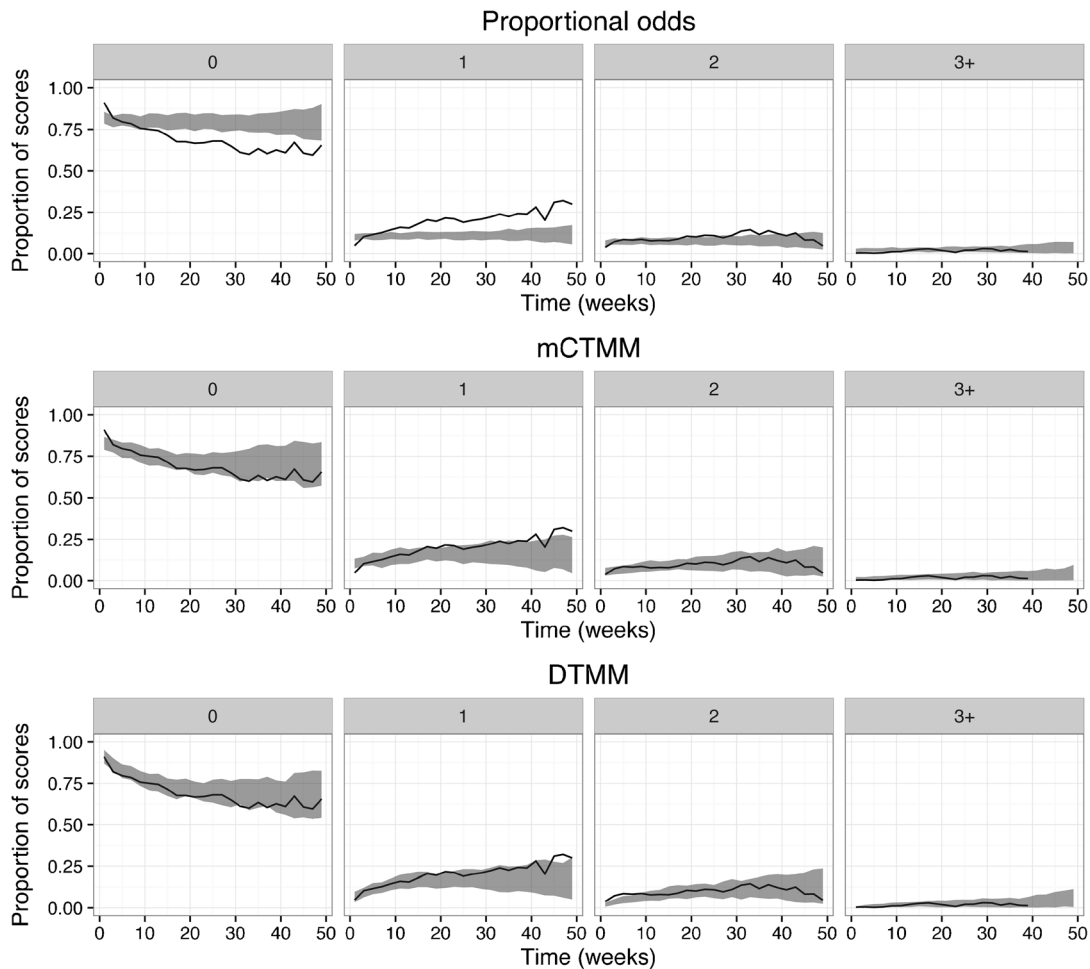


Fig. 4. Visual predictive checks of the proportional odds model, minimal continuous-time Markov model (mCTMM), and discrete-time Markov model (DTMM) for the fatigue example. The *solid lines* represent the observed time-course of the proportion of each severity grade, and the *shaded areas* are the 95% confidence intervals generated from simulations

model evidenced an underprediction of the proportion of scores of 1 and an overprediction of the proportion of scores of 3.

Example 3: Likert Pain Score

The steady-state probabilities of Likert scores were described as a function of a linear effect of time (on logit scale) with slope parameter θ_{time} , accounting for the placebo effect, and of an effect of acetaminophen intake (ACE coded as 0 if absence and 1 if presence) with coefficient θ_{ace} (Eq. 13).

$$P_{ss}(Y_{ij} \geq k) = \frac{1}{1 + e^{-(\alpha_k + \theta_{time} \cdot t + \theta_{ace} \cdot ACE)}} \quad (13)$$

A non-linear effect of time (infusion-like model) was also investigated but not supported by the data. The baseline MET (MET_0) was estimated to be 10.8 h; a linear increase in MET with a typical slope ($Slope_{MET}$) estimated to be 0.740 week^{-1} significantly improved model fit (dOFV = -539) and was consistent with the time-dependent increase in score stability (9). The probability of lower pain

scores was found to increase with time ($\theta_{time} = -0.141 \text{ week}^{-1}$) and decrease with acetaminophen intake ($\theta_{ace} = 1.84$). As shown in Table III, all parameters were estimated with reasonable uncertainty.

In the Likert pain score example, mCTMM gave the lowest OFV and an accurate description of the average number of transitions per individual, whereas the count model overpredicted it (Table I). mCTMM also best described the maximum achieved scores although the proportion of patients with the maximum score of 10 was overestimated (Fig. 3). All models could describe the low proportion of patients with maximum scores less than 5. Both the DTMM and the PO model underestimated the proportion of patients with maximum achieved intermediate scores of 5 to 7, and overestimated the proportion for largest scores. VPCs (Fig. 6) showed appropriate predictive performance of the mCTMM.

DISCUSSION

The mCTMM, a simplification of the CTMM, allows accounting for Markov elements with a minimal number of estimated parameters when analyzing ordered categorical

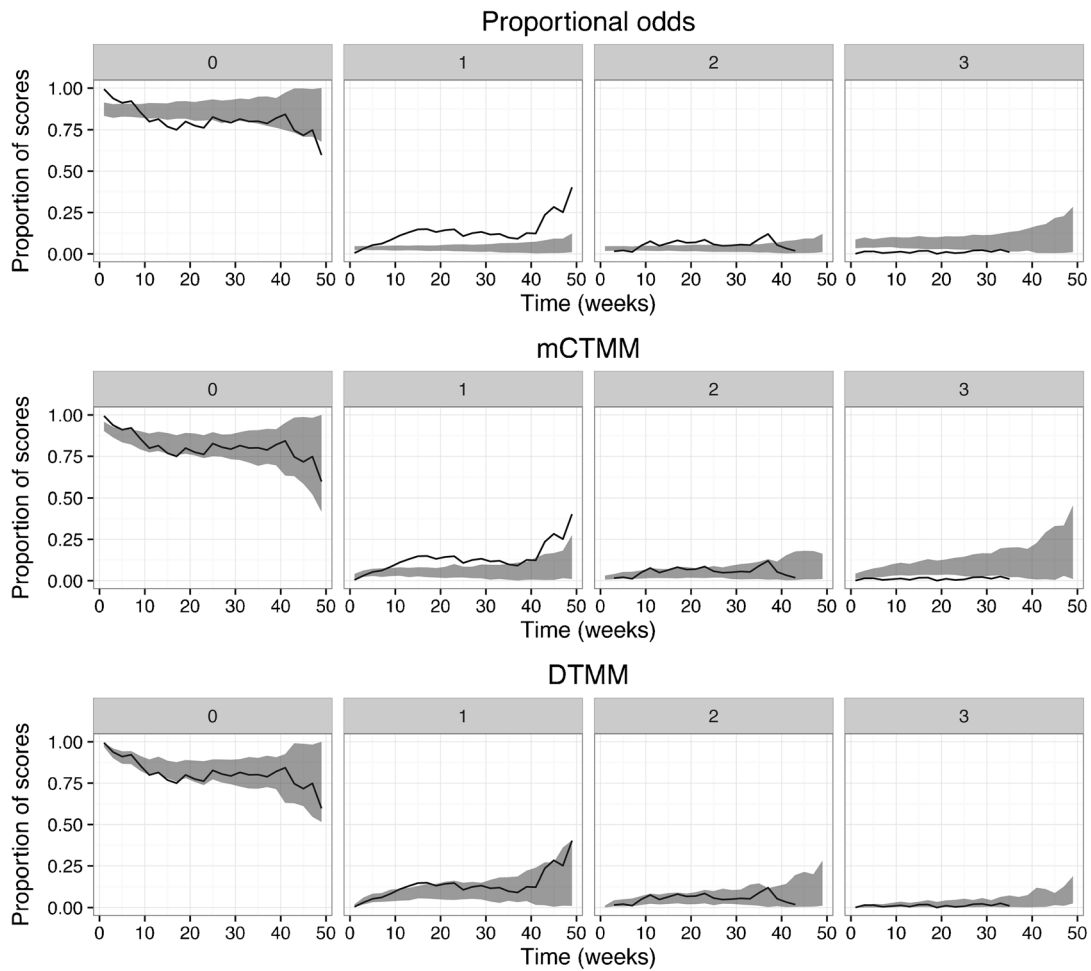


Fig. 5. Visual predictive checks of the proportional odds model, minimal continuous-time Markov model (mCTMM), and discrete-time Markov model (DTMM) for the hand-foot syndrome example. The *solid lines* represent the observed time-course of the proportion of each severity grade, and the *shaded areas* are the 95% confidence intervals generated from simulations

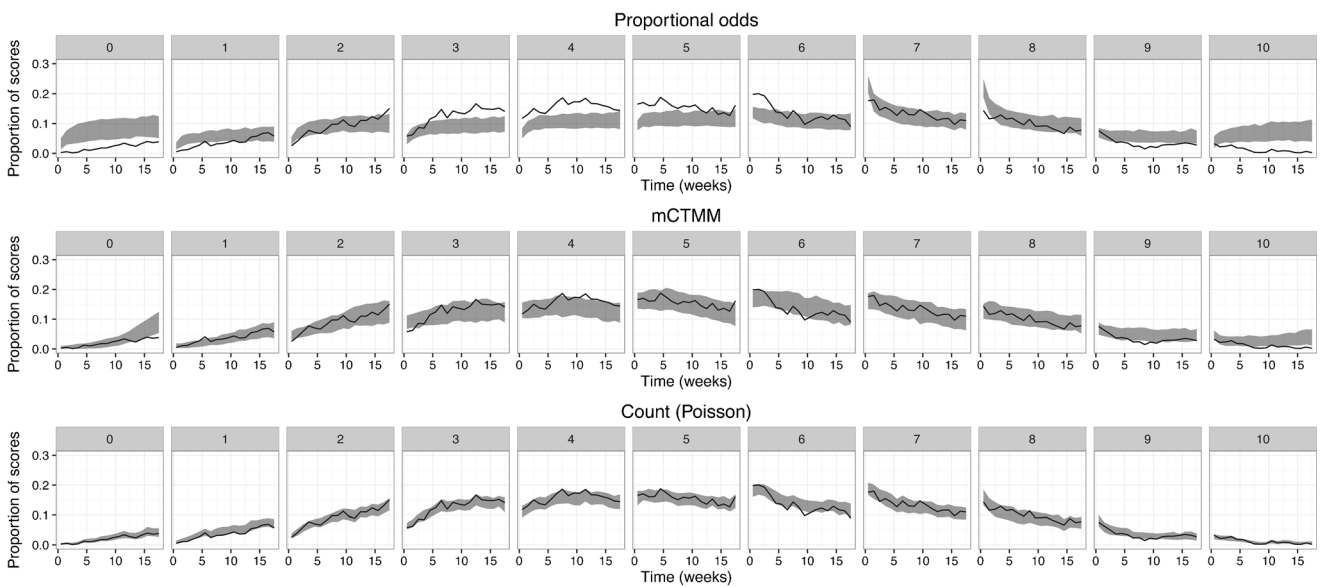


Fig. 6. Visual predictive checks of the proportional odds model, minimal continuous-time Markov model (mCTMM), and count model for the Likert example. The *solid lines* represent the observed time-course of the proportion of each severity grade, and the *shaded areas* are the 95% confidence intervals generated from simulations

Table II. Final parameter estimates and their uncertainty for the minimal continuous-time Markov model in the fatigue and hand-foot syndrome examples

Parameter	Fatigue estimate (RSE%)	Hand-foot syndrome estimate (RSE%)
α_1	-1.85 (7.5)	-17.5 (45)
b_2	-1.08 (7.9)	-1.99 (9.1)
b_3	-1.92 (10)	-2.66 (12)
MET (h)	909 (11)	399 (7.3)
$\theta_{sVEGFR3}$	-1.87 (15)	-20.5 (8.5)
$\omega_{\alpha 1}^a$	0.770 (19)	11.2 (90)
$\omega_{\theta sVEGFR3}^b$	1.07 (7.7)	0.475 (28)

RSE relative standard error, α_1 intercept parameter on the logit scale for score $k = 1$, b_k parameter for score k such that $\alpha_k = \alpha_{k-1} + b_k$, MET mean equilibration time, $\theta_{sVEGFR3}$ slope of the linear effect (on logit scale) of relative change from baseline over time in soluble vascular endothelial growth factor receptor 3 (sVEGFR-3), ω_p standard deviation of the between-subject variability associated with parameter P

^a Incorporated as additive

^b Incorporated as exponential

data with serial correlations. As expected, when applied to data with uniform assessment times previously analyzed with DTMM, as in the fatigue and HFS examples, the mCTMM did not provide any improvement in goodness of fit over DTMM; however, mCTMM could well predict the average number of transitions per individual as well as the maximum achieved scores while requiring much fewer estimated parameters than DTMM. In the HFS example, as opposed to DTMM, the mCTMM could not appropriately predict the

time-course of the proportion of scores of 1 and 3, as evidenced by the VPCs (Fig. 5). This may partly be explained by the single sVEGFR-3_{rel} effect applied in the mCTMM model on the steady-state probabilities, whereas the DTMM included four sVEGFR-3_{rel} effects (one for each $k_j - 1$) (7). When applied to the Likert pain scale, which comprises 11 ordered categories, mCTMM provided a better description of the data and had better simulation properties with regard to transitions than the previously published count model, which was based on a truncated Poisson distribution with Markov components. This last example evidenced that mCTMM provides a means for modeling discrete ordinal scales with large number of categories and Markovian features while keeping the number of estimated parameters reasonable.

Table III. Final parameter estimates and their uncertainty for the minimal continuous-time Markov model in the Likert pain score example

Parameter	Estimate (%RSE)
α_1	9.56 (2.5)
b_2	-1.90 (6.8)
b_3	-2.48 (3.5)
b_4	-1.90 (2.7)
b_5	-1.68 (2.5)
b_6	-1.66 (2.4)
b_7	-1.64 (2.3)
b_8	-1.76 (2.4)
b_9	-2.24 (3.1)
b_{10}	-2.07 (5.4)
MET ₀ (h)	10.8 (7.6)
θ_{time} (week ⁻¹)	-0.141 (13)
θ_{ace}	1.84 (4.7)
Slope _{MET} (week ⁻¹)	0.740 (22)
$\omega_{\alpha 1}^a$	3.19 (4.6)
ω_{MET_0} (h) ^b	0.881 (6.8)
$\omega_{\theta_{time}}$ (week ⁻¹) ^a	0.241 (4.9)
$\omega_{Slope_{MET}}$ (week ⁻¹) ^b	2.15 (7.7)

RSE relative standard error, α_1 intercept parameter on the logit scale for score $k = 1$, b_k parameter for score k such that $\alpha_k = \alpha_{k-1} + b_k$, MET₀ baseline mean equilibration time, θ_{time} slope of the time effect on the logit scale, θ_{ace} coefficient for acetaminophen effect on the logit scale, Slope_{MET} slope of the time effect on MET, ω_p standard deviation of the between-subject variability associated with parameter p

^a Incorporated as additive

^b Incorporated as exponential

Overall, mCTMM parameters were estimated with reasonable precision in all three real data examples, except the parameter α_1 and its BSV in the HFS model. This may be explained by the absence of HFS scores greater than zero in the absence of drug effect, as untreated patients do not spontaneously show HFS symptoms. This would translate into a steady-state probability of 0 scores ($P_{ss,0}$) that approaches 1 when the relative change from baseline in sVEGFR-3 (representative of the drug effect) is zero, which together with the asymptotic properties makes the estimation of α_1 difficult and explains the large uncertainties. The same issue may arise from data where some extreme score is predominant at baseline due to the inclusion criteria. In case this causes model instability, α_1 may be fixed to low/high value as appropriate. Sensitivity analysis may be performed to evaluate the impact of the choice of this value on the overall model fit and on other parameter estimates. Noteworthy, due to the high correlation between α_1 and its BSV, α_1 value may impact the BSV variance estimate.

As expected, PO models applied to all three examples performed poorly compared to Markov models and overpredicted the average number of transitions per individual. It has previously been shown that applying PO models to data with Markov properties may lead to potential structural model misspecification leading to erroneous conclusions about the exposure-response relationships and poor predictive performance with inflated number of transition. When Markovian features are evident from the data, the use of PO models should be avoided in favor of DTMM, CTMM, or

mCTMM. In case that the observation frequencies are not uniform throughout the study or across patients, either due to study protocol or to missing observations, DTMM should not be used as it does not account for the time since last observation. Instead, CTMM or mCTMM should be preferred. In the absence of univocal evidence of serial correlations from data exploration, mCTMM may be used instead of PO models to explore Markov components and relax the assumption that serial observations are independent. In the mCTMM, information about the equilibration time of the system is provided by a single parameter, the MET. The transit time from the lowest to the highest category can be obtained by multiplying the MET by the number of categories minus one. If the MET is estimated to a very small value as compared to the observation frequency, it means that the system is memoryless; hence, Markovian dependencies are unlikely and may be ignored by using PO models. Conversely, as seen in all three real data examples presented here, a MET estimate that is in the same order or greater than the observation frequency suggests that Markov elements are not negligible and should be accounted for (PO models should not be used). In this case, the mCTMM may be used as a starting point for model development and for exploring exposure-response relationships. Noteworthy, mCTMM can easily be extended to mimic a CTMM by investigating different MET estimates between the different states. In addition, as shown with the Likert pain scale example, the effect of time can be investigated on the MET to account for the fact that stable scores may become more or less frequent over time. The effect of other explanatory factors such as drug exposure or patient-specific characteristics may also be investigated on MET. When assessment intervals are uniform and predictor effect is constant, mCTMM with constant MET is expected to well predict individual score time-courses with homogeneous transition magnitude within a patient. Conversely, mCTMM may perform more poorly if patients experience transitions of heterogeneous magnitude, e.g., if daily observations are recorded and a single patient with constant exposure effect presents a combination of transitions of magnitudes 1 and 2.

In the mCTMM, the steady-state probabilities are affected to the same extent by predictors regardless of the previous score. This differs from the DTMM where predictors' effects are commonly implemented as conditional on the previous score (4,7). Despite being less flexible than DTMM, the mCTMM provides a framework where the effect of predictors (placebo, drug exposure, or any other explanatory factor) on the score probabilities can be summarized by a single relationship. Moreover, the mCTMM allows for the data collected at the first occasion in each patient to be used and to inform parameter estimates: the probability of each score at the first visit is assumed to be at equilibrium. In the DTMM, observations from the first occasion cannot be used in an optimal manner as information on the previous score is unknown and the model is too complex to estimate the equilibrium probability; assumptions regarding the previous score have therefore to be made, e.g., that it is the same as the current score or based on some inclusion criteria. Additionally, with mCTMM, the steady-state probability at any time regardless of the previous score can easily be derived for each score, which is not possible with DTMM. Finally, the mCTMM presented here were all based on the PO assumption for steady-state probabilities; i.e., the predictors

affected all categories to the same extent on the log odds scale. This assumption may be relaxed by using the differential odds model as presented by Kjellsson et al. (20).

In some cases, study subjects may transit to a state that is not part of the rating scale being modeled and from which it is not possible to revert back. Examples of such states include dropout and death. Although not illustrated here, the mCTMM can easily be extended to other Markov models where transition rates from all or some of the studied scale states to a dropout/death state are estimated and predictor effects on these rates can be tested.

CONCLUSION

In conclusion, the mCTMM offers an alternative to existing pharmacometric Markov models to explore exposure-response relationships and predictor effects for ordered categorical data with suspected serial correlations. The mCTMM uses a restricted number of parameters, including the easily interpretable mean equilibration time, MET. Moreover, it can be used when observation intervals are not uniform and for scales with a large number of categories.

ACKNOWLEDGEMENTS

The authors thank Joakim Nyberg for his valuable input to this manuscript. This work was supported by the Swedish Cancer Society

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

1. Sheiner LB. A new approach to the analysis of analgesic drug trials, illustrated with bromfenac data. *Clin Pharmacol Ther.* 1994;56(3):309–22.
2. Karlsson MO, Schoemaker RC, Kemp B, Cohen AF, van Gerven JM, Tuk B, et al. A pharmacodynamic Markov mixed-effects model for the effect of temazepam on sleep. *Clin Pharmacol Ther.* 2000;68(2):175–88.
3. Zingmark PH, Ekblom M, Odergren T, Ashwood T, Lyden P, Karlsson MO, et al. Population pharmacokinetics of clomethiazole and its effect on the natural course of sedation in acute stroke patients. *Br J Clin Pharmacol.* 2003;56(2):173–83.
4. Zingmark PH, Kagedal M, Karlsson MO. Modelling a spontaneously reported side effect by use of a Markov mixed-effects model. *J Pharmacokinetic Pharmacodyn.* 2005;32(2):261–81.
5. Lacroix BD, Lovern MR, Stockis A, Sargentini-Maier ML, Karlsson MO, Friberg LE. A pharmacodynamic Markov mixed-effects model for determining the effect of exposure to certolizumab pegol on the ACR20 score in patients with rheumatoid arthritis. *Clin Pharmacol Ther.* 2009;86(4):387–95. Epub 2009/07/25

6. Keizer RJ, Gupta A, Mac Gillavry MR, Jansen M, Wanders J, Beijnen JH, et al. A model of hypertension and proteinuria in cancer patients treated with the anti-angiogenic drug E7080. *J Pharmacokinet Pharmacodyn*. 2010;37(4):347–63.
7. Hansson EK, Ma G, Amantea MA, French J, Milligan PA, Friberg LE, et al. PKPD modeling of predictors for adverse effects and overall survival in sunitinib-treated patients with GIST. *CPT: Pharmacometrics Syst Pharmacol*. 2013;2:e85.
8. Henin E, You B, VanCutsem E, Hoff PM, Cassidy J, Twelves C, et al. A dynamic model of hand-and-foot syndrome in patients receiving capecitabine. *Clin Pharmacol Ther*. 2009;85(4):418–25.
9. Plan EL, Elshoff JP, Stockis A, Sargentini-Maier ML, Karlsson MO. Likert pain score modeling: a Markov integer model and an autoregressive continuous model. *Clin Pharmacol Ther*. 2012;91(5):820–8.
10. Bergstrand M, Soderlind E, Weitschies W, Karlsson MO. Mechanistic modeling of a magnetic marker monitoring study linking gastrointestinal tablet transit, in vivo drug release, and pharmacokinetics. *Clin Pharmacol Ther*. 2009;86(1):77–83.
11. Henin E, Bergstrand M, Standing JF, Karlsson MO. A mechanism-based approach for absorption modeling: the gastro-intestinal transit time (GITT) model. *AAPS J*. 2012;14(2):155–63.
12. Bergstrand M, Soderlind E, Eriksson UG, Weitschies W, Karlsson MO. A semi-mechanistic modeling strategy for characterization of regional absorption properties and prospective prediction of plasma concentrations following administration of new modified release formulations. *Pharm Res*. 2012;29(2):574–84.
13. Snelder N, Diack C, Benson N, Ploeger B, van der Graaf P. A proposal for implementation of the Markov property into a continuous time transition state model in NONMEM. [www.page-meeting.org/?abstract=1536]. PAGE 18 2009.
14. Pilla Reddy V, Petersson KJ, Suleiman AA, Vermeulen A, Proost JH, Friberg LE. Pharmacokinetic-pharmacodynamic modeling of severity levels of extrapyramidal side effects with Markov elements. *CPT: Pharmacometrics Syst Pharmacol*. 2012;1:e1.
15. Lacroix BD, Karlsson MO, Friberg LE. Simultaneous exposure-response modeling of ACR20, ACR50, and ACR70 improvement scores in rheumatoid arthritis patients treated with certolizumab pegol. *CPT: Pharmacometrics Syst Pharmacol*. 2014;3:e143.
16. Bisaso KR, Mukonzo JK, Ette EI. Transition modeling of neuropsychiatric impairment in HIV. *Comput Biol Med*. 2016;73:141–6.
17. Hansson EK, Amantea MA, Westwood P, Milligan PA, Houk BE, French J, et al. PKPD modeling of VEGF, sVEGFR-2, sVEGFR-3, and sKIT as predictors of tumor dynamics and overall survival following sunitinib treatment in GIST. *CPT: Pharmacometrics Syst Pharmacol*. 2013;2:e84.
18. Beal SL, Sheiner LB, Boeckmann AJ, Bauer RJ. NONMEM 7.3.0 users guides. (1989–2013). Hanover: ICON Development Solutions.
19. Keizer RJ, Karlsson MO, Hooker A. Modeling and simulation workbench for NONMEM: tutorial on Pirana, PsN, and Xpose. *CPT: Pharmacometrics Syst Pharmacol*. 2013;2:e50.
20. Kjellsson MC, Zingmark PH, Jonsson EN, Karlsson MO. Comparison of proportional and differential odds models for mixed-effects analysis of categorical data. *J Pharmacokinet Pharmacodyn*. 2008;35(5):483–501.