Citation: CPT Pharmacometrics Syst. Pharmacol. (2014) 3, e129; doi:10.1038/psp.2014.27
© 2014 ASCPT All rights reserved 2163-8306/14
www.nature.com/psp

TUTORIAL

Modeling and Simulation of Count Data

EL Plan[1,2]

Count data, or number of events per time interval, are discrete data arising from repeated time to event observations. Their mean count, or piecewise constant event rate, can be evaluated by discrete probability distributions from the Poisson model family. Clinical trial data characterization often involves population count analysis. This tutorial presents the basics and diagnostics of count modeling and simulation in the context of pharmacometrics. Consideration is given to overdispersion, underdispersion, autocorrelation, and inhomogeneity.

CPT Pharmacometrics Syst. Pharmacol. (2014) **3**, e129; doi:10.1038/psp.2014.27; published online 13 August 2014

OVERVIEW

Pharmacometrics,[1] while having first been solely applied to continuous data due to its historical ties with pharmacokinetics[2] and its methodological complexities,[3] now commonly includes analysis of discrete type data.[4] This tutorial, intended for pharmacometricians familiar with basic concepts in population modeling, simulation, and model-based drug development,[5] aims at presenting the foundations of count data analyses and builds on a recent tutorial on time to event models.[6] After defining count data and alternative analysis approaches, the main count models will be described with an emphasis on their assumptions, which will be completed by considerations in the context of drug development.

COUNT DATA THEORY

Count data definition

Count data are generally defined as numbers of events per interval. For count data to be recorded, events need to be specified beforehand, and then recognized and counted during the experiment. The events should, in essence, be able to and, in practice, tend to occur several times in order for the gathered data to be qualified and identifiable as count. Mathematically, counts are non-negative integers, since an event cannot happen an incomplete or a negative number of times. Naturally, zero event/count is a possibility; certain entities may even display particularly many null observations, which phenomenon will be addressed later in this tutorial. Technically, there is no upper boundary to a count, because there can theoretically be close to an infinite number of events taking place.

Count data are not restricted to biomedicine; in fact, numerous statistical advances in the domain took place within other fields. Early applications ranged from the number of convictions pronounced by jury panels per year to the number of phone calls received at a call center per minute, through the number of cavalrymen killed by horse-kicks per year. In preclinical or clinical trials, counts are typically reported for each subject, e.g., number of epilepsy seizures per month,[7] number of incontinence episodes per week,[8] or number of gastrointestinal symptoms per day.[9] Such individual data are assumed in this tutorial, therefore finding their application within the context of population analysis.[10]

Underlying hazard process

Since count data are intrinsically event frequency measures, the inherent connection with repeated time to event is evident. Events are counted within time intervals essentially when their exact time of occurrence is ignored, usually for convenience and practical reasons. For simplicity, both for data record and for data analysis, time intervals are generally—although not systematically—of fixed length, e.g., months, weeks, or days. The shorter the time intervals, the less summarized the occurrences are. As in survival analyses, events are presumed instantaneous and assumed to arise only one at a time, i.e., the probability that an individual experiences two events at the exact same instant is zero.

In consonance with the analogous event data, count data are driven by an underlying hazard. Hazard is the instantaneous rate of the events.[6] Determining whether and how this hazard varies with covariates, including treatment effect, is typically the aim of the data analysis. When time affects the hazard, i.e., the hazard is not constant, summarizing the information in counts per interval may dilute the effect. The extent of this dilution depends on the pattern of the time effect and its coincidence with the time interval limits. Nevertheless, in studies where the goal is to most accurately characterize a time-varying hazard, a repeated time to event analysis, although slower, will be more appropriate than a count analysis. The exceptions to this principle are if the events are rare enough and if the time intervals are short enough for only one or zero count to be recorded.[11]

Noncount data

Some types of data are sometimes unsuitably identified as event count data, such as grouped binary data or ordinal data. Grouped binary data arise from counting the number of successes or failures resulting from a certain number of binary outcome processes, i.e., Bernoulli trials. A binomial model, handling the number of tests performed and therefore the implicit maximum, appears more adapted to analyze grouped binary data. A Poisson model can nevertheless

[1]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden; [2]Pharmetheus, Uppsala, Sweden. Correspondence: EL Plan (elodie.plan@farmbio.uu.se; elodie.plan@pharmetheus.com)
Received 3 February 2014; accepted 13 May 2014; published online 13 August 2014. doi:10.1038/psp.2014.27

be appropriate to describe binomial data if the number of experiments is large, since the two models are mathematically equivalent for an infinite number of repetitions. Regarding ordinal data, ordered categorical models are the suitable models for the evaluation.

On the other hand, true event count data are often handled with models for a different data type. For instance, they may be categorized, i.e., dichotomized (event, no event) or made ordinal (low, medium, and high frequency of events). This practice, intended to simplify the process, risks information and power loss. Another alternative approach applied to count data is to treat them as continuous, which does not involve data transformation. A continuous model, despite not addressing the integer nature of the data, may be successfully applied to count data, predominantly when the distribution is wide with high counts. These characteristics commonly apply to events in a volume or an area, e.g., number of neutrophils per microliter of blood or number of bacteria per square centimeter of a Petri dish. It should be noted that the classic count methods can also be used for these types of observations nonetheless.

Count data visualization

There are several—standard or not—ways to visualize count data, and a representative sample will be given in this tutorial. The familiar representation of the dependent variable—here the counts—vs. time (**Figure 1**, left panel) is inevitable, although it does not carry the notion of the integer nature of the observations. This type of graph will, as for other data, inform about the variability between subjects as well as a potential time course.

A histogram of the frequency of the counts (**Figure 1**, right panel) is highly recommended, because of the indication it gives about the shape of the distribution. This is referred to as probability mass function (pmf) owing to the discrete nature of the probability distribution of counts, as opposed to the probability density function associated with continuous data. In **Figure 1**, the histogram was intentionally rotated to exhibit how the two y-axes coincide, which is, however, not common practice. It can be highlighted that most variables included in the count model structure *in fine* govern the pmf.

COUNT MODEL STRUCTURE

The Poisson model

The most important count model is the Poisson model. It is also the model to which all other count models presented in this tutorial converge, when their extra parameter goes to zero or infinity. It is named after the French mathematician who reintroduced the concept[12] a century after de Moivre published it.[13]

The model describes the observations using the Poisson distribution. Accordingly, the random variable in the model is Poisson distributed. The term Poisson model used thereafter corresponds to this definition.

An observation Y during an interval j for an individual i is said to follow a Poisson distribution when the probability that it takes values $n = 0,1,2,\ldots$ can be expressed as in Eq. 1.

$$P\left(Y_{ij} = n\right) = \frac{\lambda_i{}^n}{n!} \cdot e^{-\lambda_i} \tag{1}$$

where λ_i is a parameter and ! the factorial function.

Mean count parameter

The parameter λ (lambda) is the single parameter of the Poisson model in its simplest form. The expectation for the Poisson distribution $E(Y)$ is equal to λ, and the parameter will therefore also correspond to the arithmetic mean of the counts occurring during a certain time. λ is by definition a positive real number, i.e., it is not necessarily an integer.

In pharmacometrics, parameters are often expressed as a mixed effect—a combination of a fixed (θ) and a random (η) effect. A common parameterization preventing negative values is given in Eq. 2. The introduction of interindividual

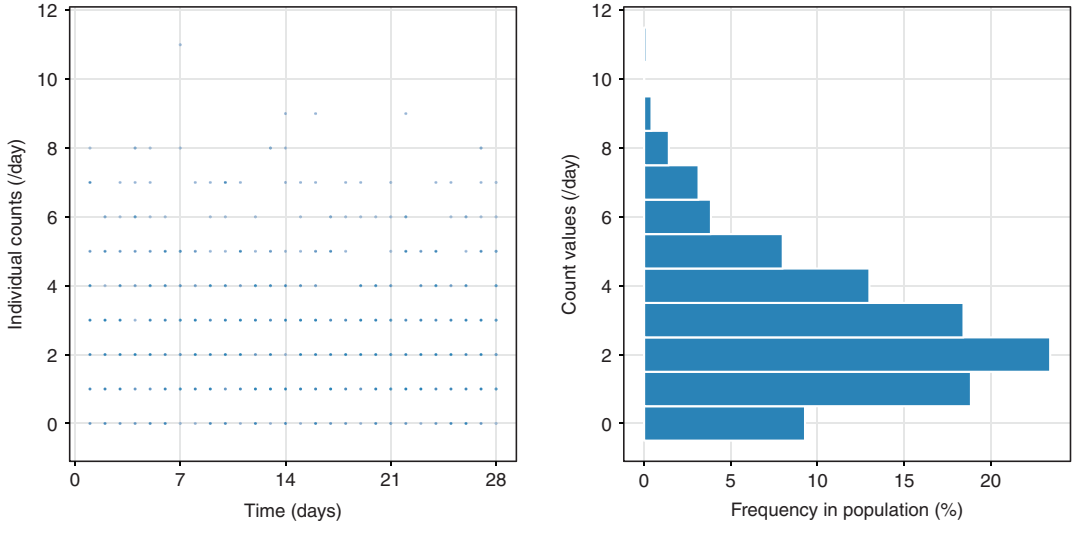

Figure 1 Representation of count data. The left panel displays the time course of the counts in a 25-individual population, whereas the right panel reveals the probability mass function of the counts in the same population.

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

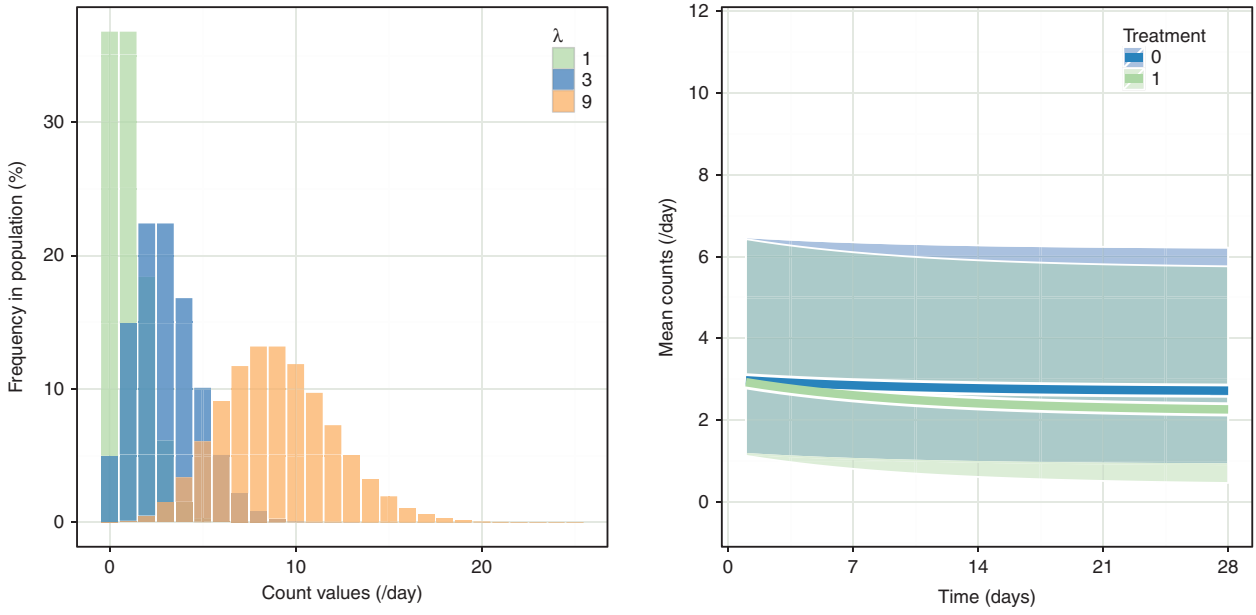

Figure 2 Properties of the parameter λ. The plots represent how varying λ values between individuals or groups of individuals result in shifted probability mass functions (left panel, with three distributions) and shifted mean counts (right panel, stratified on treatment). Varying λ within individuals can also be revealed in the right panel with fluctuating means over time, here represented with their 95% prediction intervals.

variability (IIV)—the random effect—results in λ to be interpretable as the mean number of events per time interval for individual *i*.

$$\lambda_i = \theta \cdot e^{\eta_i} \tag{2}$$

As opposed to, for example, a normal distribution, the variance of the Poisson distribution is not governed by an additional parameter, but by λ_i as well. As a result, an individual's λ_i value defines both the mean and the variance of the discrete probability distribution for the observations of this specific individual, as in Eq. 3.

$$\lambda_i = \mathrm{E}(Y_{ij}) = \mathrm{Var}(Y_{ij}) \tag{3}$$

This feature of the Poisson distribution is illustrated in **Figure 2** (left panel) with three pmfs corresponding to three λ values. It can be observed that as λ increases: (i) the mean of the distribution shifts and the probability of zeros decreases, (ii) the function is less skewed and approximates a normal distribution, and (iii) the variance increases and higher counts are encountered.

Poisson model properties

The equality between mean and variance of the counts experienced by an individual constitutes an important property of the model through its distribution. Named equidispersion, it should be investigated in the data exploration and disclosed in the model report. Though restrictive on the statistical level, this trait naturally applies to many *in vivo* phenomena.

Another characteristic intrinsic to the Poisson model is independence between events or memoryless process. Common for discrete probability functions, it means that the occurrence of an event does not alter the probability of another. While admissible in idealistic cases such as rolling a die and counting the number of occurrences a 6 is obtained, this assumption may generally not be fulfilled when dealing with biomedical symptoms.

The third Poisson property, presented here and discussed elsewhere in the tutorial, is the constance of the event rate within time intervals, also known as homogeneous process. On account of an evident lack of information at this level, this aspect is inherent to the data collection. While echoing the consequences of low-resolution time intervals, the feature can be perfectly acceptable in many cases.

Exponential model family

Considering λ as a piecewise constant rate further enables to establish the link between the Poisson model and the hazard model.

The similarity between the two exponential models can be depicted, for instance, by computing the probability of zero event to happen within an interval according to the Poisson model (Eq. 4) and the same probability given by the time-to-event approach when the hazard is known to be constant (Eq. 5).

$$P(Y_{ij} = 0) = \frac{{\lambda_i}^n}{n!} \cdot e^{-\lambda_i} = \frac{{\lambda_i}^0}{0!} \cdot e^{-\lambda_i} = e^{-\lambda_i} \tag{4}$$

$$P(Y_{ij} = 0) = S(t_j) = e^{-\int_{t_{j-1}}^{t_j} h_i(u)\, du} = e^{-h_i \cdot (t_j - t_{j-1})} \tag{5}$$

The analytic solution of the integral in the case of a constant hazard makes more apparent that λ, the variable describing the number of events within an interval in Eq. 4, corresponds to the product in Eq. 5 of the rate of the events and the length of the interval.

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The proof of the mathematical equivalence between a piecewise proportional hazards model and a Poisson regression model was produced, along with demonstration of identical estimates, independently by two teams in 1980.[14,15]

The advantages of acknowledging the direct relationship between λ and a constant hazard reside in the meaning of the methods, the understanding of the results, and the interpretation of the conclusions. Stating the existence of a clear correspondence (Eqs. 4 and 5) between the two models also suggests shared robustness and accuracy in terms of inferences.

Mean count function

If λ, in the standard Poisson model does not vary within time intervals, it can nevertheless vary between them. Since λ represents the mean counts, it can be susceptible to time changes and is the parameter on which a potential drug effect will primarily be explored. These elements place λ at the heart of disease progression and pharmacodynamics characterizations, which include *in silico* description of the time course (natural history or placebo effect) of the disease status assessment and the action of the drug (if achieving a measurable response). The disease status assessment is in this case the counted events chosen as end point during, e.g., a clinical trial. The assessment is expected to undergo a response qualified either as a desired effect or as an adverse event, both of which can take the form of a stimulation or an inhibition.

After establishment of the count structural model, model building can be carried out on λ the same way it is for a continuous model,[16] in terms of a variability model and a covariate model. In pharmacometric models, λ is typically a function that can be affected by time ($f(t_j)$), drug effect ($g(PK_i(t_j))$, and covariate effects ($\alpha_1 X_1, \alpha_2 X_2, \ldots \alpha_K X_K$). Eq. 2 may therefore be extended to, e.g., Eq. 6.

$$\lambda_i = \theta \cdot e^{\eta_i} \cdot (1 - f(t_j)) \cdot (1 - g(PK_i(t_j)) \cdot (1 + \alpha_1 X_1) \cdot (1 + \alpha_2 X_2) \cdots (1 + \alpha_K X_K) \quad (6)$$

The variable $PK_i(t_j)$ stands for a pertinent pharmacokinetic model-derived exposure metric[17] driving the response. It is apparent that the exposure variable will account for the entire time interval and, therefore, is likely to be a summary metric (e.g., steady-state concentration or average area-under-the-curve) as opposed to momentary concentration.

Note that setting the functions f and g (in Eq. 6) to asymptotically approach zero allows the mean count to tend to zero while keeping it from becoming negative. Prevalent exponential decay and sigmoid E_{max} functions, respectively, can be successfully used in this context. To the same end, others may prefer a parameterization on the natural logarithmic scale such as, e.g., Eq. 7. The fixed effect of the intercept θ in this case will not be directly comparable with the one from the previous equation.

$$\log(\lambda_i) = \theta + \eta_i + f(t_j) + g(PK_i(t_j)) + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_K X_K \quad (7)$$

Figure 2 (right panel) illustrates such a case with the mean response in a hypothetical population displaying a decline over the month, which is more pronounced in one of the two treatment groups. The ordinate axis clearly depicts λ as no longer only a parameter but a multidimensional function varying with time and exposure and furthermore, between individuals.

COUNT MODEL COMPONENTS FOR DATA VIOLATING POISSON ASSUMPTIONS

Nonequidispersed events

Equidispersion can be verified numerically by computing and comparing the mean and the variance of the observations of each individual represented in the raw data. A statistical test consists in calculating the χ^2 distributed ratio between the variance and the mean, i.e., the index of dispersion,[18,19] and testing equality to one. However, in a context of time-varying mean count, only limited conclusions can be drawn from this practice. Therefore, such calculations should only be interpreted along with other diagnostics and must not preclude further investigations.

Graphically, when individual means and variances are plotted against each other, a linear relationship suggests that the condition is respected. Violations of the equidispersion assumption can be of two kinds: overdispersion or underdispersion. Overdispersion corresponds to cases of a variance greater than the mean for most of the population, and underdispersion to the opposite, a variance lower than the mean. In the mean–variance plot (**Figure 3**, left panel), these types of profiles are respectively situated above and below the identity line. Again, this may be impacted by a nonconstant λ, and thus, a stratification on time (e.g., by week in this daily record example) should be considered.

Another diagnostic designed to indicate whether the observed distribution approximately compares with a Poisson distribution is the Poissonness plot.[20,21] It is based on a quantity called count metameter, calculated from the frequency of each encountered count (f_n), the factorial of the count values ($n!$), and the sum of all frequencies (N). According to Eq. 8, if the data follow a Poisson distribution, the metric is linear with regard to the possible counts n, with an intercept $-\lambda$ and a slope log (λ).

$$f_n = \frac{{\lambda_i}^n}{n!} \cdot e^{-\lambda_i} \cdot N \Leftrightarrow \log(f_n) = n \cdot \log(\lambda_i) - \lambda_i - \log(n!) + \log(N) \Leftrightarrow \log\left(\frac{f_n + n!}{N}\right) = -\lambda_i + n \cdot \log(\lambda_i) \quad (8)$$

This diagnostic (**Figure 3**, right panel), similar in concept to the Q–Q plot, however, seems to be more adapted to detect overdispersed distributions—more points not falling on a straight line—than underdispersed ones, the double Poisson in this example. The diagnostic power of the available plots is limited due to the stochastic nature of the count distribution as well as the potentially complex λ profile. Therefore, one should not completely rely upon them and investigate alternative models described in this section.

Overdispersion or underdispersion circumstances may not always be explainable. The first case, denoting that the within-individual variability is especially large compared with the mean, i.e., that the same patient can experience few events during most time intervals and many at some other times in the same study, can be encountered in types of diseases presenting sudden outbreaks of symptoms. The second case,

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

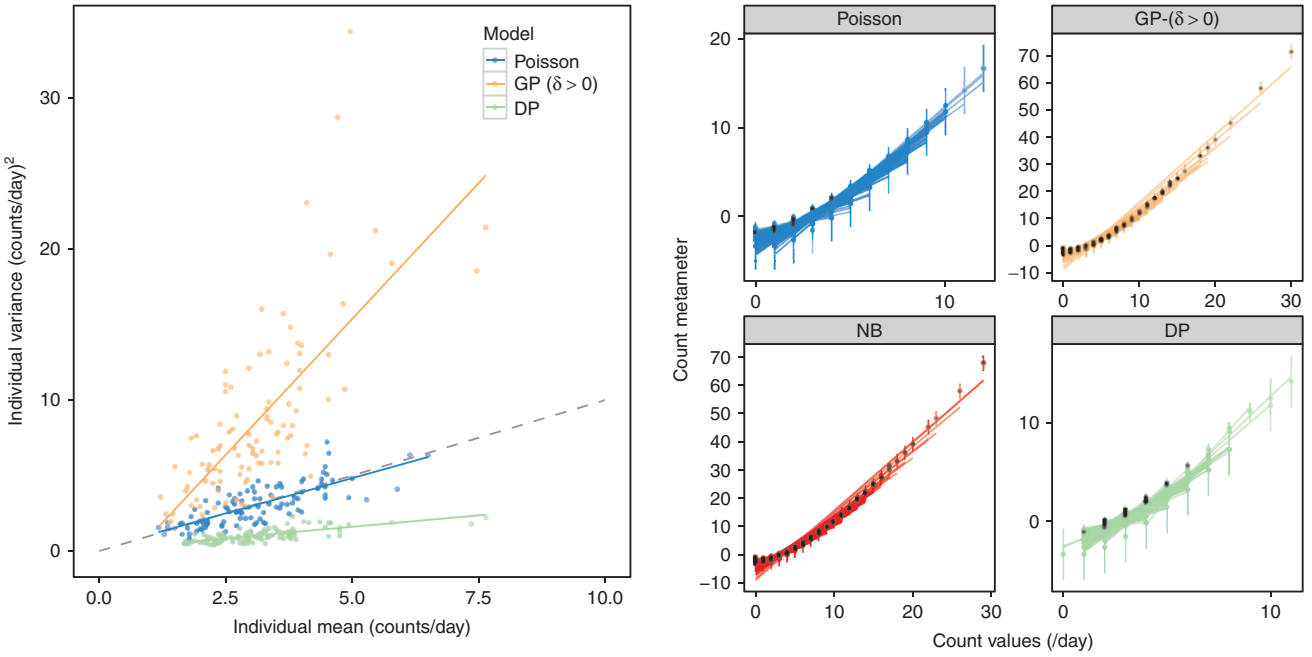

Figure 3 Dispersion plots. The left panel is a mean–variance plot and the right panel is a Poissonness plot. Each dot represents metrics of all counts for one individual on the left, and a metric for each possible count of one individual on the right. Each line consists in a linear regression of the dots belonging to the same model on the left and to the same individual on the right, and the additional dashed line on the left panel displays the identity. The dots on the right panel also have their 95% confidence interval represented and are colored black if the linear regression did not include this interval. The plots were produced from simulations with different models: Poisson, negative binomial (NB), generalized Poisson with a positive dispersion parameter (GP (δ >0)), and double Poisson (DP).

referring to notably low within-individual variability vs. mean, i.e., homogeneous records, maybe displayed by rather constant experiences and robust symptoms, like in diseases qualified as slowly progressive. The cause of the dispersion phenomenon may be identifiable, e.g., epilepsy counts may substantially differ between morning and evening and the bimodal distribution may be misinterpreted as overdispersed, or recurring migraines may be treated with rescue medication leading the count distribution to appear underdispersed. But in many cases, driving forces for the noise in the process are missing or unknown, and resulting overdispersion or underdispersion seen to occur are left to be characterized in the absence of explanatory factors.

Analyzing data not conforming to the equidispersion assumption related to the Poisson distribution may require the use of candidate models different from the Poisson model. Several models have to this end been proposed and a few are presented here.

Negative binomial model. The negative binomial (NB) model,[22,23] also sometimes referred to as the inverse binomial model, is the main alternative to the Poisson model in case of overdispersion (**Figure 4**, right panel). It uses a supplementary parameter, $\mathcal{O}$ (omicron), responsible for an increased variability compared with the mean counts. This overdispersion parameter belongs to the range $\mathcal{O} \in (0,\infty)$, but values below 1 are generally sufficient. The discrete probability distribution (Eq. 9) still has a mean of $E(Y_{ij}) = \lambda$, while the variance is now $\mathrm{Var}(Y_{ij}) = \lambda_i(1+ \mathcal{O}_i \cdot \lambda_i)$.

$$P(Y_{ij} = n) = \frac{\Gamma\left(n + \mathcal{O}_i^{-1}\right)}{\Gamma\left(\mathcal{O}_i^{-1}\right) \cdot n!} \cdot \left(\frac{1}{1+\mathcal{O}_i \cdot \lambda_i}\right)^{\mathcal{O}_i^{-1}} \cdot \left(\frac{\lambda_i}{\mathcal{O}_i^{-1} + \lambda_i}\right)^{n} \quad (9)$$

Although the distribution is not defined for $\mathcal{O} = 0$, it tends toward a Poisson distribution when approaching it. This means that by allowing the parameter to vary between individuals—by including a random effect—both NB-like and Poisson-like profiles can be captured. The NB model was found adequate when applied to epilepsy data in several analyses[24,25] and to infectious diseases.[26]

Generalized Poisson model. The generalized Poisson model[27,28] is another model used to account for deviation from equidispersion in data to be analyzed. The additional parameter, δ (delta), governs the dispersion of the pmf. The distribution can result in an overdispersed pmf if δ is positive ($\delta > 0$) and in an underdispersed one if δ is negative ($\delta < 0$) (**Figure 4**). The limits within which δ exists are $\delta \in$, $\max(-1, -\lambda_i / \Delta), 1)$, where Δ is the largest positive integer greater than or equal to 4 that satisfies the inequality $\lambda + \Delta\delta > 0$. The parameterization in Eq. 10 permits the mean count to be equal to $E(Y_{ij}) = \lambda_i$ and the variance to $\mathrm{Var}(Y_{ij}) = \lambda_i/(1-\delta_i)^2$.

$$P(Y_{ij} = n) = \frac{\lambda_i(1-\delta_i) \cdot \left(\lambda_i(1-\delta_i) + n\delta_i\right)^{n-1}}{n!} \cdot e^{-(\lambda_i(1-\delta_i)+n\delta_i)} \quad (10)$$

Clearly, the generalized Poisson model reduces to the Poisson model and therefore adjusts for equidispersion when $\delta = 0$.

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

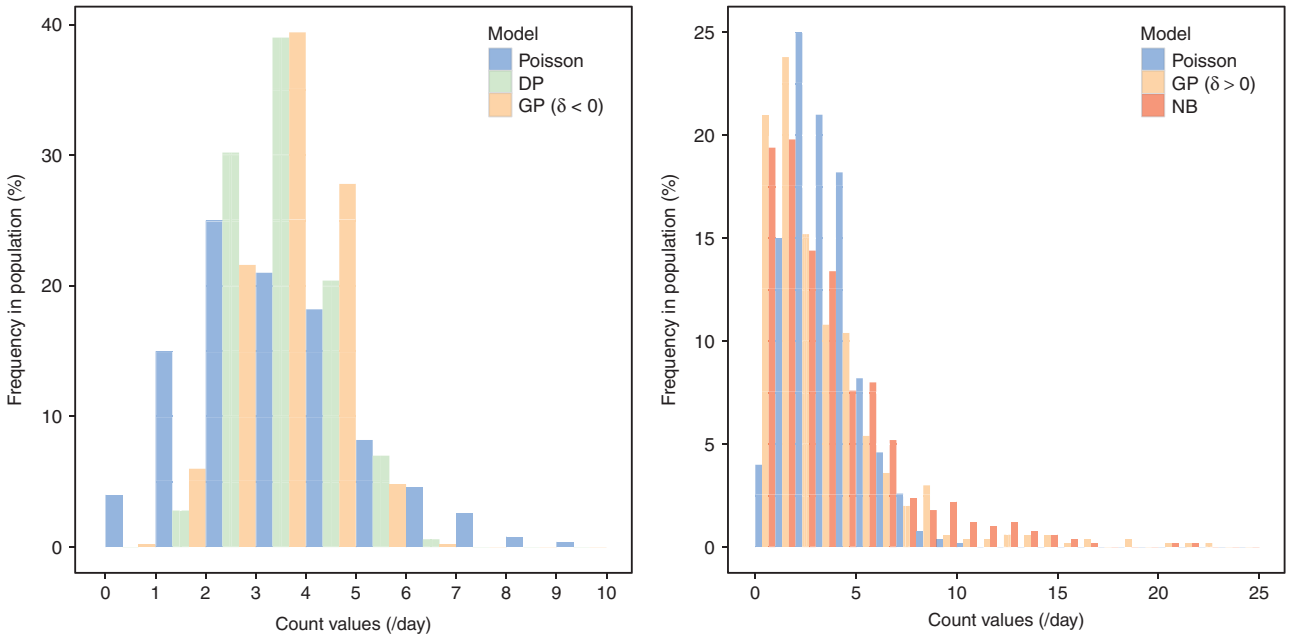

Figure 4 Nonequidispersed distributions. The same distribution density issued from a Poisson model ($\lambda = 3$) is plotted on both panels, together with two models for underdispersion (narrower distributions) on the left panel and two models displaying overdispersion (wider distributions) on the right panel. The explored models are double Poisson (DP) ($\upsilon = 3$), generalized Poisson (GP) ($\delta = -0.75$ and 0.50), and negative binomial (NB) ($\mathcal{O} = 0.75$).

The flexibility of the model to accommodate the twin properties of over- and underdispersion makes it an attractive candidate. Nonetheless, the dependence between the lower bound of δ and λ, two parameters to estimate, introduces an inconvenience. A step-wise procedure can be used to circumvent the difficulty by starting with $\delta = 0$ and then with the obtained fixed effect estimate λ setting $\max(-1, -\lambda/4) < \delta \leq 1$.

Double Poisson model. The double Poisson model[29,30] is generally promoted in the context of underdispersion (**Figure 4**, left panel), since in contrast to the previous models, the higher the extra parameter, the more underdispersed the distribution. Both forms of dispersion can be managed though, through the parameter υ (upsilon). The real number υ yields underdispersion when above 1 ($\upsilon > 1$) and overdispersion when below 1 ($\upsilon < 1$). The model gains in flexibility when IIV is associated with υ. The discrete probability distribution generated by Eq. 11 has a mean of approximately $\mathrm{E}(Y_{ij}) \approx \lambda_i$ and a variance of approximately $\mathrm{Var}(Y_{ij}) \approx \lambda_i / \upsilon^i$.

$$P(Y_{ij} = n) = \left(\frac{{\lambda_i}^n}{n!} \cdot e^{-\lambda_i}\right)^{\upsilon_i} \cdot \left(\frac{n^n}{n!} \cdot e^{-n}\right)^{1-\upsilon_i} \cdot \upsilon_i^{1/2} \cdot \left(1 + \frac{1-\upsilon_i}{12\upsilon_i\lambda_i}\left(1 + \frac{1}{\upsilon_i\lambda_i}\right)\right)^{-1} \quad (11)$$

As its name suggests, the model is based on the double exponential family. This extension of the Poisson model is obtained as an exponential combination of the Poisson density with mean λ and of the Poisson with mean n. In the case of $\upsilon = 1$, the double Poisson model collapses to the Poisson model. Technically, the double Poisson function only approximates a pmf. The last element of the equation is a normalizing constant specified to make the sum of the probabilities closer to unity.

Other nonequidispersion models. Several other names, notations, or parameterizations exist for the models described above that do not correspond to different distributions. The mentioned models were used to simulate density profiles to compare underdispersed (**Figure 4**, left panel) and overdispersed (**Figure 4**, right panel) data to a unique Poisson—equidispersed—set of data. While the means of all pmf were chosen to be the same, the dispersion parameters were set to illustrate the feature, which could be enhanced or diminished.

Other models were proposed to handle overdispersion, such as mixture models. This alternative is based on the suspicion of two (or more) underlying processes leading to an event. A mixture model can be developed with two (or more) λ values, where the observed sum of events has a larger variability than either of the two distributions.

Another frequent case of overdispersion is when it is due to an excess of a certain value. Zero counts are commonly observed in higher frequency in the data compared with the amount predicted by the model. A possible situation is a threshold needed to be crossed for positive counts to be observed, another being a poor compliance when adverse events are counted. The Hurdle model, built as a mixture of two different densities, may remedy to excess of zeros. The zero-inflated Poisson model is a straight-forward substitute, where the two mixture components, (i) zero and (ii) a Poisson pmf, are weighted by a factor corresponding *de facto* to the probability of not observing any event. The zero-inflated Poisson model can be combined with distribution functions other than the Poisson model, e.g., a zero-inflated NB model was explored in the analysis of antibody titers[31] to account for patients not developing immunity.

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

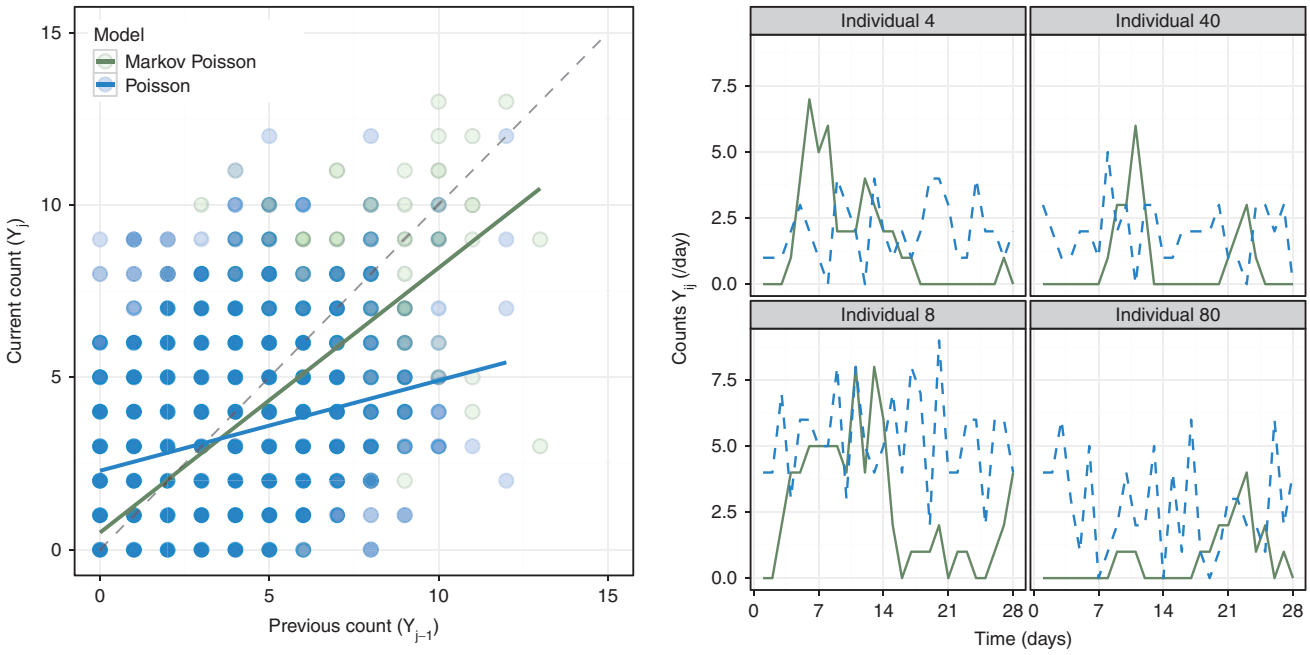

Figure 5 Serial correlated counts. A population displaying correlation between consecutive observations (Markov Poisson) is represented in comparison with a simulation from a Poisson model (assuming independence between counts). The left panel investigates the correlation between neighboring counts with the dots representing the counts and the solid lines the linear regressions (dash line is identity line). In the right panel four individuals are exhibited, with slowly varying profiles in the Markov population and highly fluctuating curves in the vanilla Poisson simulations.

More nonmixture models presenting dual properties of over- and underdispersion are: the compound generalized Poisson distribution,[32] the Katz System yielding the generalized event count model,[33,34] or the Conway–Maxwell Poisson model.[35,36]

A further reason for observing underdispersion is data truncation. Right truncation is observed when events above a fixed value (c_r) are not counted, whereas left truncation occurs when the censoring affects counts below a set value (c_l). The most common case of the latter is zero truncation. Models to be used in case of data truncation can be any of the above, yet their distribution must be truncated while retaining probabilities adding up to 1. Density truncation can be applied though division by the cumulative density function (cdf) within the range of possible counts, i.e., sum of probabilities $\sum_{n=0}^{c_r} P(Y=n)$ for right truncation or $1-\sum_{n=c_l}^{\infty} P(Y=n)$ for left truncation. This technique was also used to model score data.[37]

Nonindependence between events

Assuming that events happen in an independent manner may not be appropriate to describe disease symptoms. In epilepsy, the notion that seizures beget seizures is debated, but implies that seizures themselves are epileptogenic. Besides events directly triggering more events, events clustered in high-intensity periods or low-intensity periods can denote an underlying—indirect—dependence. Migraines are believed to follow such an unknown process leading low-activity epochs to intercalate with peaks. Micturition contains a psychology factor leading events to be artificially more regularly spaced compared with if truly driven by a random process.

Dependence between events can be diagnosed by scrutinizing the data. A simple way to visualize dependence is to plot consecutive observations against each other (**Figure 5**, left panel). A linear regression line can indicate in which fashion their dependence occurs. Individual profiles (**Figure 5**, right panel) may also support independence invalidation, since within-individual dependence in biomedical observations is likely to be expressed as trains of similar values. Correlations may be involved if the displayed time courses are smoother than when generated under an independence assumption.

In statistics, this property is named serial correlation and can be addressed in different ways. In pharmacometrics, approaches implemented to reflect dependence between observations include the first-order autoregressive (AR(1)) model,[38] correlating residual errors, and stochastic differential equations,[39] decomposing residual errors into system and measurement noise. However, both approaches have been described so far only for continuous processes.

For discrete models, published methods for serial correlation mainly relate to Markov models. Observations in Markov models are assumed dependent on previous states. The previous state in question is often the exact preceding one, referred to as first-order Markov model, but can instead or in addition implicate the observation before the last (i.e., second-order Markov), etc. Because the chain of states is discrete, i.e., time intervals are defined as a set of measurement occasions, the full terminology is discrete time Markov chain.

Note that the preceding observation is known and can therefore directly enter the model. This is a concept different from that of hidden Markov models,[40] where the state influencing the current observation is unknown. In a hidden Markov model, the latent system transitions between hidden

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

states, with the state where it is positioned at a certain time altering the probabilities for the observations at that time. This methodology was implemented for mixed-effects count models.[41]

Markov models are integrated as components to probability models through functions of a prior observation affecting the probability of a current observation (Eq. 12). In count models, most implementations are specified according to models considered by Zeger and Qaqish.[42] Markov components contribute to modify the pmf mainly via addition to λ, or through inflation-deflation probabilities (similar to the zero-inflated Poisson model).

$$P(Y_{ij} = n \mid Y_{ij-1}) = f(n; \lambda_i; Y_{ij-1}) \quad (12)$$

Markov element predictors can be of any type: binary, categorical, count, or continuous, since they can be the actual previous observation or a derived value. In pharmacodynamic count models, the two most encountered parameterizations are a continuous function of the previous count and a mixture function of the previous status no count/counts. Examples of the first case are a proportional E_{max} function of the preceding epilepsy seizure count[25] or a linear function of the past number of multiple sclerosis contrast enhancing lesions[43] (or a multiplicative power function of hypothetical events in **Figure 5**). Still applied to epilepsy, results were presented where the last observation information was dichotomized depending on whether the seizure was present or absent.[24,44]

Nonconstant event rate within time intervals

As expressed before, the concept of hazard[45] is key in this context, since this rate determines when the events, ultimately counted, occur. Applying a model from the Poisson family assumes that this rate is constant within the intervals where the counting is made. These stationary increments may not be suitable for all problems at hand. Time variations of event occurrence are likely to take place within all time intervals, but their capture may be more possible and crucial in some circumstances than others.

In the case of safety studies, when the count variable is a rare adverse event, this time variation is seldom of importance and often not feasible to capture, given the sparseness of the available data. Similarly, the description of cyclic fluctuations, like a circadian rhythm, where the length of the period coincides with the length of the interval, is likely not possible. However, prior or simultaneous characterization of a time-varying biomarker or a pharmacokinetic profile, suspected to influence the hazard, may be the ground for requiring nonconstant event rate within intervals.

Relaxing the stationarity assumption corresponds to converting the homogeneous Poisson process into an inhomogeneous (or nonhomogeneous) Poisson process. Ample research was performed in the branch of recurring events rate analysis. Inhomogeneous Poisson models couple (i) models with time-varying count intensity[46] to (ii) Poisson (or other) probability distribution models. In this case, the expectation of the sum of counts during a finite interval (Eq. 13) can be obtained from the sum of the expectations of each event.

$$\mathrm{E}(Y_{i(j-1,j)}) = \int_{t_{j-1}}^{t_j} \lambda_i(u)\, du \quad (13)$$

Where $\lambda_i(u)$ can be expressed as in Eq. 6 with the $PK_i(t_j)$ being for example the drug concentration (not averaged over the time interval any more) or a covariate X being time-varying.

Ordinary differential equations are, therefore, required most of the time in this family of models. Nevertheless, analytical solutions may be employed in some cases (e.g., the Weibull distribution). The parameter/function λ acts as hazard which accumulates over time. The likelihood of positive counts is set to the chosen pmf, while the likelihood of zero count is equal to the survivor function. The hazard, cumulative hazard, and survival were represented in **Figure 6**.

The literature reporting pharmacometric investigations using inhomogeneous Poisson models is relatively sparse. One of the studies is a sensitivity analysis made with regard to the data simplification to compare homogeneous models with an inhomogeneous model.[47] In other respects, a novel approach was proposed connecting a third type of model: an ordered categorical model, to estimate the maximal severity achieved from nonconstant repeated categorical events summarized per time intervals.[48]

One may note that the inhomogeneous approach relieves the considerations with regard to the regularity of the interval length, and an integration is done over elapsed time after the beginning of the study. Intervals of different lengths can nonetheless be efficiently handled in a homogeneous model. As indicated by Eq. 5, the interval duration can be taken into account such as Eq. 14. Hence, if symptoms are, for example, counted every day during stays at the clinic and every week when the patient is at home, the data can be analyzed without having to add daily counts into weekly ones.

$$\lambda_{ij} = \lambda_i \cdot (t_j - t_{j-1}) \quad (14)$$

COUNT MODELING PRACTICES

Model fit software

Model fit of count data can be performed in all major nonlinear mixed-effects modeling software, i.e., the pharmacometric software NONMEM, Monolix, Phoenix, and S-ADAPT, and the statistical software WinBUGS, R, SAS, and Matlab.

The language (e.g., NM-TRAN) used to code the probability functions naturally depends on the software of choice, with the Poisson (and oftentimes some other) distribution(s) being predefined in a few of the programs (e.g., WinBUGS, where one can write “dpois”).

Furthermore, the factorial function $(\cdot!)$, which is part of all count models, is not systematically embedded in all programs. Approximations such as a refinement of the Stirling’s formula or the one derived by Ramanujan are a solution to this. The gamma function $(\Gamma\cdot)$ can be handled the same way since it is directly related to the factorial function. For most models though, the approximation of the factorial does not affect parameter estimation as it is constant in the likelihood.

Another potential difficulty taking place during count modeling is numerical issues while dealing with high counts. This risk can be prevented and model stability increased by log-transforming the model or its components, e.g., the gamma function of the NB model.

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

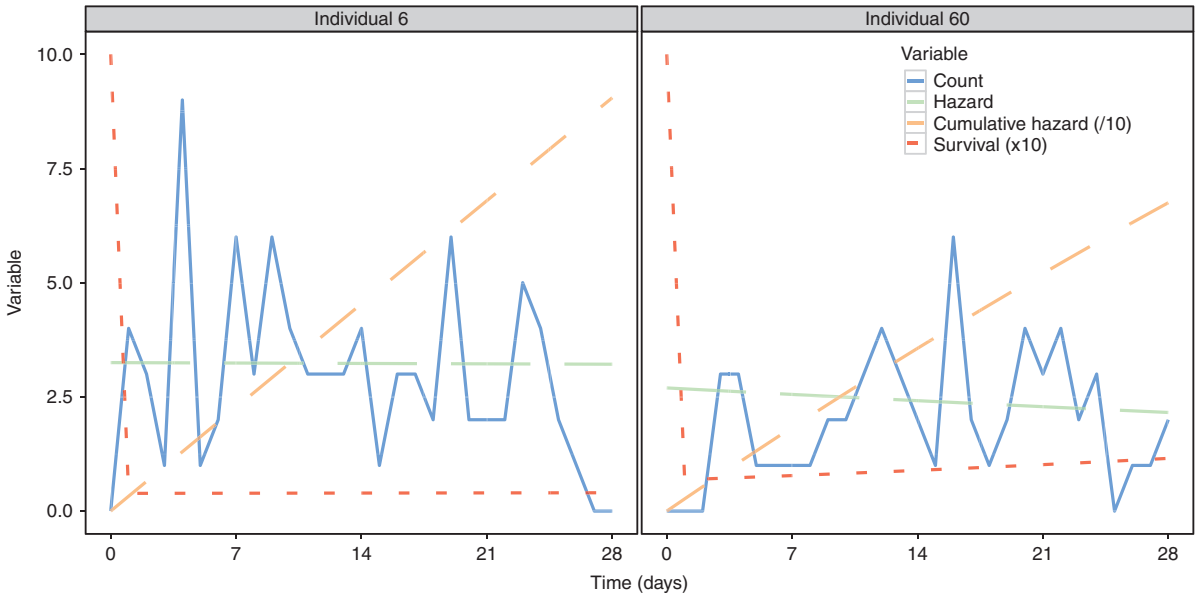

Figure 6 Inhomogeneous model variables. The plot displays variables for two individuals. The "count" variable corresponds to observed data and the "hazard" to the function λ. The "cumulative hazard" and the "survival" function are quantities involved in the estimation, and scaled here to share the same *y*-axis.

Estimation method choice

Maximum likelihood estimation is the prevalent parameter estimation method in pharmacometrics. Since count models consist in probabilities, they are easily interpretable as sets of likelihoods, in the same way the other types of discrete models are.[49] Estimation in population count models is, therefore, completed through likelihood-based approaches, requiring approximation of the integral over the random effects. As in other cases, estimating parameters through minimization of −2 times the log-likelihood presents advantages over maximization of the likelihood.

Maximum likelihood estimation algorithms may not perform equally well due to the challenging nonlinearity of the models. Accordingly, the use of the primary NONMEM estimation algorithm FOCE is prohibited and the alternative Laplace and sampling-based methods recommended. The evaluation[50,51] of the performance—accuracy and precision—of the algorithms Laplace in NONMEM, stochastic approximation of expectation maximization in Monolix, and Gaussian quadrature in SAS indicated that these estimation methods are appropriate for count analyses. Bayesian methods are also successfully used to estimate count model parameters as seen on available training material.[52]

Model development workflow

After appropriate graphical exploration of the data, knowledge about the data structure, profile, and trends is undeniably gained. This information can highlight features to reflect in the model structure as well as to use in guiding model building.

The model development workflow nevertheless follows a relatively consistent series of steps: (i) fit Poisson model, (ii) test other densities, (iii) develop structural model on λ, and (iv) refine interindividual variability model. In a case of a parallel arms placebo-controlled study, steps (i) and (ii) may be performed only on nontreated patients. Step (iii) includes the investigation of time effect, previous observation effect, covariate effect, drug effect, etc. Before step (iv), most parameters may already consist in an association between a fixed and a random effect, so correlations (i.e., covariance matrices) and semiparametric distributions (e.g., box-cox transformation) can be investigated.

Model evaluation, as in any other model exercise, takes place at each model modification to quantify and qualify the fit improvement.

An example in the form of a hands-on case study is provided as **Supplementary Material** online.

Model evaluation diagnostics

Numerical evaluation in count models is facilitated by the fact that most models are nested with the Poisson model. Hence, the likelihood ratio test can be applied between the hierarchical models of a workflow following the steps previously described.

Graphical evaluation does not include the same battery of goodness of fit plots as for continuous models. Regardless, residuals adapted to count models are applicable.[53] The Pearson residual, based on the estimated mean and variance compared with the observations, and the Anscombe residual, built only with the estimated mean and the observations, may be inspected vs. λ or or time to detect misspecifications.

Simulation-based evaluation remains nonetheless the gold-standard due to its ability to reliably depict several aspects—variability and structure—of the model efficiently. Based on a preferably high number of simulations potentially including uncertainty, a visual predictive check can, for instance, be produced. A visual predictive check typically displays percentiles of a dependent variable plotted against an independent variable both for observations and simulations, overlap of fitted data and

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

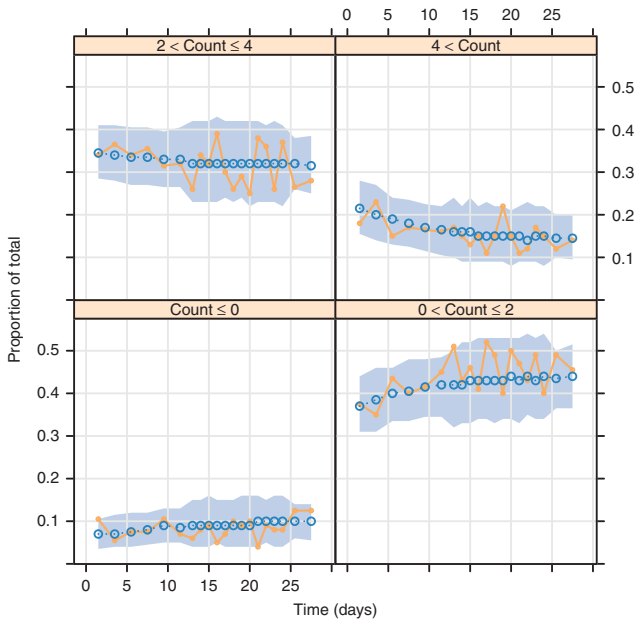

Figure 7 Visual predictive check for count data. Each panel represents the probability of observing a range of counts against time ($P(Y = 0), P(1 \le Y \le 2), P(3 \le Y \le 4), P(Y \ge 5)$). The solid line is the median of the observed data and the ribbon the 95% confidence interval around the median (dotted line) of 500 simulations from a count model fitted to count data.

simulated data is the goal. In the case of count data, the ordinate can be the number of events although its integer nature may alter the plot. A favored representation consists in plotting the proportions of counts within certain ranges. In **Figure 7**, these proportions are featured with regard to time. Other graphs to consider to complete the assessment include: a stratification on treatment and other important categorical covariates, a switch to dose on the *x*-axis if enough dose levels and to important continuous covariates, an adjustment of the count proportions to proportions of transition values between consecutive counts, and a replacement of the *y*-axis by the variance-to-mean ratio. Some software, like PsN,[54] feature an automation tool generating the visual predictive check with simulations, autobinning, and percentile calculation being done with one command.[55]

Count data simulation

Whether intended for model evaluation or model inferences, simulations are an important part of count data analysis. In software where the distribution function is predefined (e.g., R), it can be used both for estimation and for simulation, whereas in software where a user-defined function is necessary (e.g., NONMEM), the code should be adapted between estimation and simulation.

The simulation code needed for Poisson distributed data is once again borrowed from the proportional hazard model (Eq. 15). It follows Algorithm 1, according to which the number of times n that r needs to be drawn from a uniform distribution for the sum $\sum_{k=0}^{n} \log(r_k) / -\lambda_i$ to reach 1 corresponds to the simulated number of events.

$$P(Y_{ij} = 0) = e^{-\lambda_i} \Leftrightarrow \log(r) = -\lambda_i \Leftrightarrow \frac{log\ (r)}{-\lambda_i} = 1 \qquad (15)$$

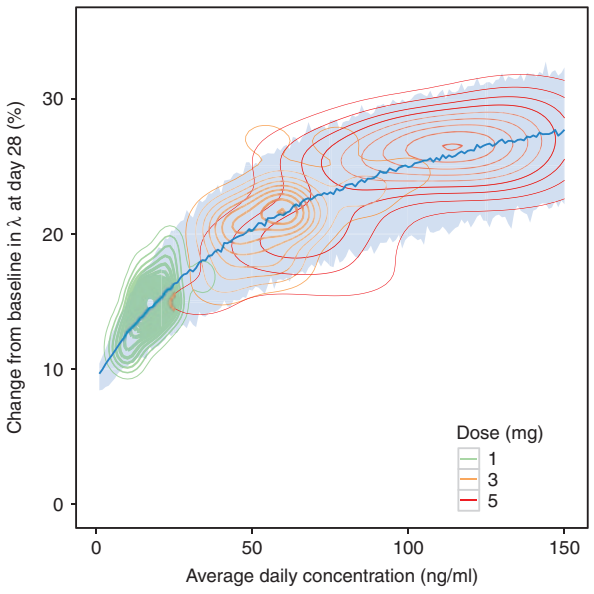

Figure 8 Dose selection plot. The exposure–response is depicted as change from baseline in λ at day 28 against average daily concentration. The blue elements are the model-based mean and 95% prediction interval generated from 500 simulations on a fine grid. The two-dimensional densities, built from patients observations, indicate the dose groups from the last conducted studies. A placebo group had been studied and a placebo effect observed.

```
t = 0; n=0;
while t < 1 do
    r ~ U(1);
    t = t + log(r)/ − λ_i ;
    if t < 1 then
        n=n+1
    end
end
n
```

Algorithm 1: Simulation of Poisson random variable

```
p = 0; n = 0;
r ~ U(1);
while p < r do
    p = p + P(Y_ij = n) ;
    if p < r then
        n=n+1
    end
end
n
```

Algorithm 2: Simulation of count model random variable

For pmfs different from the Poisson model, a more general simulation code is necessary. This code is described in Algorithm 2 and is suitable for all models presented in this tutorial.

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Count model-based drug development

Count models are inevitably part of decision-making within the drug development landscape, which is why solid knowledge about the approach is necessary. A fundamental aspect is that count modeling is a flexible approach detached from any particular therapeutic area. As more data are analyzed and more knowledge acquired, it is nevertheless imaginable and desirable that platform models are built in specific diseases. As of now, predominance of application of mixed-effects count models seems to be situated in epilepsy[24,25] and oncology.[8,9] Knowledge about the biological processes involved is invaluable and enables the development of mechanistic components into the models, but insight into the physiological variations is at times insufficient, justifying the choice of empirical functions.

The impact of count modeling can take place at all phases of drug development, as shown by analyzed studies ranging from animal experiments[56] to patient trials.[9] Decisions that can be informed include go-nogo or dose selection. The function λ is where the drug effect, subject of the evaluation, is most likely to be implemented and a target response to be defined. Graphical[57] representations should be sought (e.g., **Figure 8**) and parameter uncertainty considered when relevant. Drug development strategic questions generally involve several aspects such as efficacy and safety, in which case Gupta *et al.*[8] combined a count model and a categorical model to predict the therapeutic index of a new formulation. When the different aspects concern the same end point, like frequency and severity of a unique type of events, models can be combined on another level,[48] but this has not been done for count models yet.

A determinant factor in clinical development is trial efficiency. Underpowered or poorly designed studies lead to nonconclusive and expensive trials. Important decisions in study design can be supported by stochastic simulations and estimations from a previously developed model. Next to sample size and study duration, the length of the time interval during which events are counted may be investigated in the case of count studies. Optimal design,[58] recognized to significantly impact the degree of conclusiveness of an experiment through parameter precision, has been shown to be able to handle discrete/count data.[59–61] Hence, from a vantage point of a drug developer, count modeling and simulation can enable more precise, more certain, faster, and cheaper decisions.

CONCLUSIONS

Modeling and Simulation is a powerful tool to characterize, quantify, understand, predict, power, optimize, and rationalize (pre)clinical trial data and studies. As in the case of any tool, its impact is maximized if it is shaped to fit its purpose. In the case of count outcome, tools in the form of pmfs are available. They come with assumptions as well as extensions, which were the topic of this tutorial. The Poisson model, a close relative of the survival model, is the basis for all count data models. Adaptations for handling overdispersion, underdispersion, autocorrelation, or inhomogeneity were proposed in the literature and presented here. Other sources of information may be useful to the reader, such as Cameron *et al.*[62] or Coxe *et al.*.[63] All in all, count modeling is an integral part of the dynamic science of pharmacometrics alongside count data being possible end points collected during drug development.

Acknowledgments. We thank Mats Karlsson for encouraging data type-informed modeling analyses and Sebastian Ueckert for valuable comments.

Conflict of Interest. The author declared no conflict of interest.

1. Barrett, J.S., Fossler, M.J., Cadieu, K.D. & Gastonguay, M.R. Pharmacometrics: a multidisciplinary field to facilitate critical thinking in drug development and translational research settings. *J. Clin. Pharmacol.* **48**, 632–649 (2008).
2. Rowland, M., Benet, L.Z. & Riegelman, S. Two-compartment model for a drug and its metabolite: application to acetylsalicylic acid pharmacokinetics. *J. Pharm. Sci.* **59**, 364–367 (1970).
3. Sheiner, L.B. & Beal, S.L. Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. *J. Pharmacokinet. Biopharm.* **8**, 553–571 (1980).
4. Paule, I., Girard, P., Freyer, G. & Tod, M. Pharmacodynamic models for discrete data. *Clin. Pharmacokinet.* **51**, 767–786 (2012).
5. Mould, D.R. & Upton, R.N. Basic concepts in population modeling, simulation, and model-based drug development. *CPT Pharmacometrics Syst. Pharmacol.* **1**, e6 (2012).
6. Holford, N. A time to event tutorial for pharmacometricians. *CPT: Pharmacometrics Syst. Pharmacol.* **2**, e43 (2013).
7. Frame, B., Miller, R. & Lalonde, R.L. Evaluation of mixture modeling with count data using NONMEM. *J. Pharmacokinet. Pharmacodynamics* **30**, 167–183 (2003).
8. Gupta, S.K., Sathyan, G., Lindemulder, E.A., Ho, P.L., Sheiner, L.B. & Aarons, L. Quantitative characterization of therapeutic index: application of mixed-effects modeling to evaluate oxybutynin dose-efficacy and dose-side effect relationships. *Clin. Pharmacol. Ther.* **65**, 672–684 (1999).
9. Wirz, S., Wartenberg, H.C. & Nadstawek, J. Less nausea, emesis, and constipation comparing hydromorphone and morphine? A prospective open-labeled investigation on cancer pain. *Support. Care Cancer* **16**, 999–1009 (2008).
10. Sheiner, L.B. The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. *Drug Metabolism Rev.* **15**, 153–171 (1984).
11. Plan, E.L., Ma, G., Någård, M., Jensen, J. & Karlsson, M.O. Transient lower esophageal sphincter relaxation pharmacokinetic-pharmacodynamic modeling: count model and repeated time-to-event model. *J. Pharmacol. Experiment. Ther.* **339**, 878–885 (2011).
12. Siméon Denis Poisson and Christian Heinrich Schnuse. Recherches sur la Probabilité des Jugements en matière criminelle et en matière civile. (Meyer, Parris, France, 1841).
13. de Moivre, A. De mensura sorits seu, de probabilitate eventuum in ludis a casu fortuito pendentibus. (1711).
14. Holford, T.R. The analysis of rates and of survivorship using log-linear models. *Biometrics* **36**, 299–305 (1980).
15. Laird, N. & Olivier, D. Covariance analysis of censored survival data using log-linear analysis techniques. *J. Am. Stat. Association* **76**, 231–240 (1981).
16. Upton, R.N. & Mould, D.R. Basic concepts in population modeling, simulation, and model-based drug development: part 3-introduction to pharmacodynamic modeling methods. *CPT Pharmacometrics Syst. Pharmacol.* **3**, e88 (2014).
17. Mould, D.R. & Upton, R.N. Basic concepts in population modeling, simulation, and model-based drug development-part 2: introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e38 (2013).
18. Hoel, P.G. On indices of dispersion. *Ann. Math. Stat.* **14**, 155–162 (1943).
19. Selby, B. The index of dispersion as a test statistic. *Biometrika* **52**, 627–629 (1965).
20. Hoaglin, D.C. A Poissonness plot. *The American Statistician* **34**, 146–149 (1980).
21. Hoaglin, D.C., Mosteller, F. & Tukey, J.W. *Exploring Data Tables, Trends, and Shapes.* (John Wiley & Sons, Hoboken, NJ, 2011).
22. Gardner, W., Mulvey, E.P. & Shaw, E.C. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychol. Bull.* **118**, 392–404 (1995).
23. Land, K.C., McCall, P.L. & Nagin, D.S. A comparison of Poisson, negative binomial, and semiparametric mixed Poisson regression models with empirical applications to criminal careers data. *Sociological Methods & Research* **24**, 387–442 (1996).
24. Trocóniz, I.F., Plan, E.L., Miller, R. & Karlsson, M.O. Modelling overdispersion and Markovian features in count data. *J. Pharmacokinet. Pharmacodyn.* **36**, 461–477 (2009).
25. Ahn, J.E., Plan, E.L., Karlsson, M.O. & Miller, R. Modeling longitudinal daily seizure frequency data from pregabalin add-on treatment. *J. Clin. Pharmacol.* **52**, 880–892 (2012).
26. Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E. & Getz, W.M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

27. Consul, P.C. & Jain, G.C. A generalization of the Poisson distribution. *Technometrics* **15**, 791 (1973).
28. Consul, P.C. & Shoukri, M.M. The generalized poisson distribution when the sample mean is larger than the sample variance. *Communications in Statistics - Simulation and Computation* **14**, 667–681 (1985).
29. Efron, B. Double exponential families and their use in generalized linear regression. *J. Am. Stat. Association* **81**, 709 (1986).
30. Ridout, M.S. & Besbeas, P. An empirical model for underdispersed count data. *Statistical Modelling* **4**, 77–89 (2004).
31. Bonate, P.L., Sung, C., Welch, K. & Richards, S. Conditional modeling of antibody titers using a zero-inflated poisson random effects model: application to Fabrazyme. *J. Pharmacokinet Pharmacodyn.* **36**, 443–459 (2009).
32. Ambagaspitiya, R.S. & Balakrishnan, N. On the compound generalized Poisson distributions. *Astin Bulletin* **24**, 255–263 (1994).
33. King, G. Variance specification in event count models: from restrictive assumptions to a generalized estimator. *Am. J. Political Sci.* **33**, 762–784 (1989).
34. Winkelmann, R. & Zimmermann, K.F. Count data models for demographic data. *Math. Popul. Stud.* **4**, 205–221, 223 (1994).
35. Zou, Y., Lord, D. & Geedipally, S.R. Over- and Under-Dispersed Count Data: Comparing Conway-Maxwell-Poisson and Double-Poisson Distributions. Transportation Research Board 91st Annual Meeting, Washington, DC, USA, 2012.
36. Zou, Y., Geedipally, S.R. & Lord, D. Evaluating the Double Poisson Generalized Linear Model. Transportation Research Board 92nd Annual Meeting, Washington, DC, USA, 2013.
37. Plan, E.L., Elshoff, J.P., Stockis, A., Sargentini-Maier, M.L. & Karlsson, M.O. Likert pain score modeling: a Markov integer model and an autoregressive continuous model. *Clin. Pharmacol. Ther.* **91**, 820–828 (2012).
38. Karlsson, M.O., Beal, S.L. & Sheiner, L.B. Three new residual error models for population PK/PD analyses. *J. Pharmacokinet. Biopharmaceutics* **23**, 651–672 (1995).
39. Tornøe, C.W., Overgaard, R.V., Agersø, H., Nielsen, H.A., Madsen, H. & Jonsson, E.N. Stochastic differential equations in NONMEM®: implementation, application, and comparison with ordinary differential equations. *Pharmaceutical Res.* **22**, 1247–1258 (2005).
40. Baum, L.E. & Petrie, T. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.* **37**, 1554–1563 (1966).
41. Delattre, M., Savic, R.M., Miller, R., Karlsson, M.O. & Lavielle, M. Analysis of exposure–response of CI-945 in patients with epilepsy: application of novel mixed hidden Markov modeling methodology. *J. Pharmacokinet. Pharmacodyn.* **39**, 263–271 (2012).
42. Zeger, S.L. & Qaqish, B. Markov regression models for time series: a quasi-likelihood approach. *Biometrics* **44**, 1019 (1988).
43. Velez de Mendizabal, N., Trocóniz, I.F., Hutmacher, M.M. & Bies, R. Discrete distribution models for relapsing-remitting dynamics observed in multiple sclerosis. *PAGE* **21** (2012).
44. Schoemaker, R., Snoeck, E. & Stockis, A. Exposure-response modeling of daily seizure counts in focal epilepsy trials. *PAGE* **19** (2010).
45. Singpurwalla, N.D. The hazard potential. *J. Am. Stat. Association* **101**, 1705–1717 (2006).
46. Cox, E.H., Veyrat-Follet, C., Beal, S.L., Fuseau, E., Kenkare, S. & Sheiner, L.B. A population pharmacokinetic–pharmacodynamic analysis of repeated measures time-to-event pharmacodynamic responses: the antiemetic effect of ondansetron. *J. Pharmacokinet. Biopharmaceutics* **27**, 625–644 (1999).
47. Simeoni, M. *et al.* Counting events: population approaches using NONMEM. *PAGE* **14**, 2005.
48. Plan, E.L., Karlsson, K.E. & Karlsson, M.O. Approaches to simultaneous analysis of frequency and severity of symptoms. *Clin. Pharmacol. Ther.* **88**, 255–259 (2010).
49. Yano, I., Beal, S.L. & Sheiner, L.B. The need for mixed-effects modeling with population dichotomous data. *J. Pharmacokinet. Pharmacodyn.* **28**, 389–412 (2001).
50. Plan, E.L., Maloney, A., Trocóniz, I.F. & Karlsson, M.O. Performance in population models for count data, part I: maximum likelihood approximations. *J. Pharmacokinet. Pharmacodyn.* **36**, 353–366 (2009).
51. Savic, R. & Lavielle, M. Performance in population models for count data, part II: a new SAEM algorithm. *J. Pharmacokinet. Pharmacodyn.* **36**, 367–379 (2009).
52. Metrum Institute. MI255: Exposure-Response Modeling of Categorical, Count, and Time-to-Event Data Using Bayesian Methods. <http://www.youtube.com/metruminst>. (2011). Lecture 6.
53. Pierce, D.A. & Schafer, D.W. Residuals in generalized linear models. *J. Am. Statistical Association* **81**, 977–986 (1986).
54. Lindbom, L., Pihlgren, P., Jonsson, E.N. & Jonsson, N. PsN-Toolkit–a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Computer Methods and Programs in Biomedicine* **79**, 241–257 (2005).
55. Keizer, R.J., Karlsson, M.O. & Hooker, A. Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e50 (2013).
56. Jonker, D.M., van de Mheen, C., Eilers, P.H., Kruk, M.R., Voskuyl, R.A. & Danhof, M. Anticonvulsant drugs differentially suppress individual ictal signs: a pharmacokinetic/pharmacodynamic analysis in the cortical stimulation model in the rat. *Behav. Neurosci.* **117**, 1076–1085 (2003).
57. Ito, K. & Murphy, D. Application of ggplot2 to Pharmacometric Graphics. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e79 (2013).
58. Box, J.F. R.A. Fisher and the design of experiments, 1922–1926. *The American Statistician* **34**, 1–7 (1980).
59. Nestorov, I., Graham, G., Duffull, S., Aarons, L., Fuseau, E. & Coates, P. Modeling and stimulation for clinical trial design involving a categorical response: a phase II case study with naratriptan. *Pharm. Res.* **18**, 1210–1219 (2001).
60. Ogungbenro, K. & Aarons, L. Sample size calculations for population pharmacodynamic experiments involving repeated dichotomous observations. *J. Biopharmaceutical Stat.* **18**, 1212–1227 (2008).
61. Nyberg, J., Karlsson, M.O. & Hooker, A. Population optimal experimental design for discrete type data. *PAGE* **18** (2009).
62. Cameron, A.C. & Trivedi, P.K. *Regression Analysis of Count Data*. (Cambridge University Press, Cambridge, England, 1998).
63. Coxe, S., West, S.G. & Aiken, L.S. The analysis of count data: a gentle introduction to poisson regression and its alternatives. *J. Personality Assessment* **91**, 121–136 (2009).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://www.nature.com/psp)

21638306, 2014, 8, Downloaded from https://ascpt.onlinelibrary.wiley.com/doi/10.1038/psp.2014.27 by Sri Lanka National Access, Wiley Online Library on [02/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License