An Extended Random-effects Approach to Analysing Repeated, Overdispersed Count Data

Clarice G. B. Demétrio

ESALQ/USP, Piracicaba, SP, Brasil

Clarice.demetrio@usp.br

joint work with
Geert Molenberghs, Hasselt University, Belgium
Geert Verbeke, Katholieke Universiteit Leuven, Belgium

VIII Encontro dos Alunos
Pós-graduação em Estatística e Experimentação Agronômica
Piracicaba, SP
November, 20, 2018

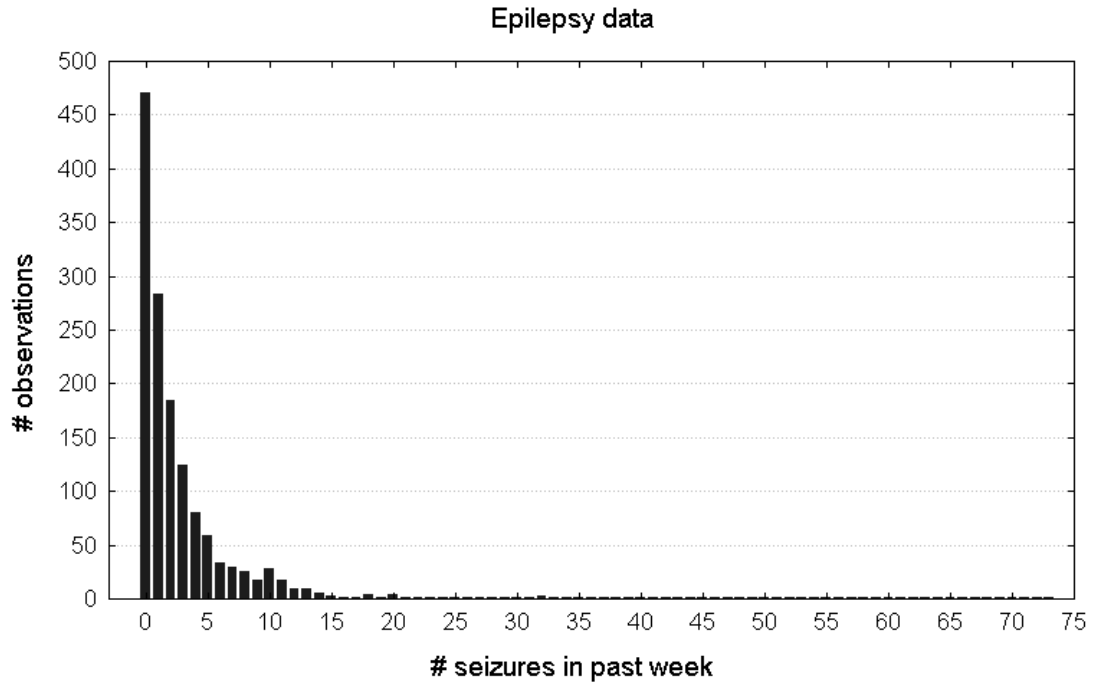
Outline

- Motivating application - A Clinical Trial in Epileptic Patients
- Generalized linear models
- Poisson regression models
- Overdispersion in GLM's
- Univariate overdispersed count data
- Longitudinal overdispersed count data
- Estimation
- Discussion of the example
- Final remarks

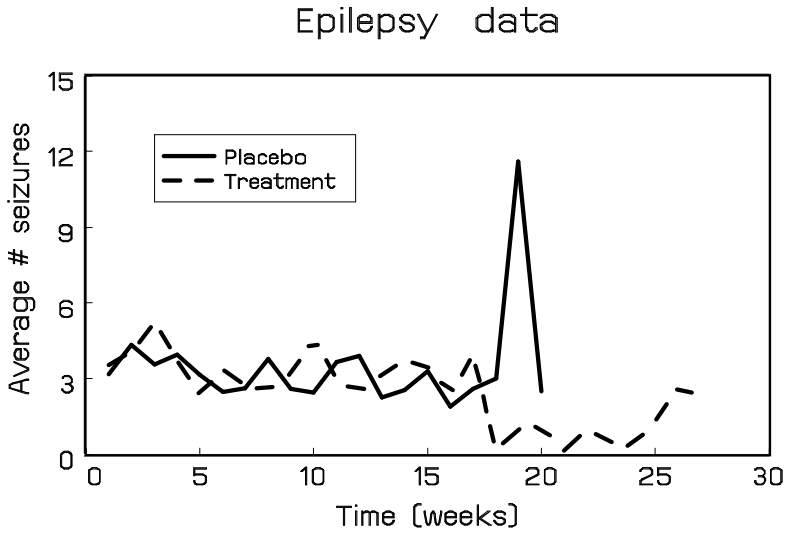
Motivation - A Clinical Trial in Epileptic Patients

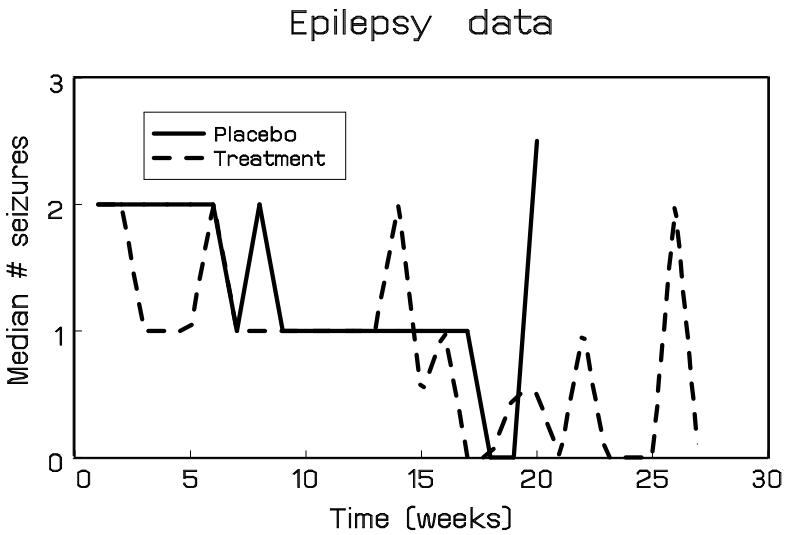
- a randomized, double-blind, parallel group multi-center study for the comparison of placebo with a new anti-epileptic drug (AED)
- after a 12-week baseline period, 45 epilepsy patients were assigned to the placebo group, 44 to the active (new) treatment group
- patients measured weekly during 16 weeks (double-blind) and some up to 27 weeks in a long-term open-extension study
- outcome of interest: the number of epileptic seizures experienced during the last week, i.e., since the last time the outcome was measured
- key research question: whether or not the additional new treatment reduces the number of epileptic seizures

Considerations about the data



- a very skewed distribution, with the largest observed value equal to 73 seizures in week





- unstable behavior explained by:
  - presence of extreme values,
  - very little observations available at some of the time-points, especially past week 20
- longitudinal count data:
  - discrete data
  - possible correlation between measurements for the same individual

	# Observations		
Week	Placebo	Treatment	Total
1	45	44	89
5	42	42	84
10	41	40	81
15	40	38	78
16	40	37	77
17	18	17	35
20	2	8	10
27	0	3	3

- serious drop in number of measurements past the end of the actual double-blind period, i.e., past week 16

Generalized Linear Models (GLM's)

– unifying framework for much statistical modelling (Nelder and Wedderburn, 1972)

– an extension to the standard normal theory linear model

– three components:

- independent random variables Y_i, $i = 1, \ldots, n$, from a linear exponential family distribution with means μ_i and constant scale parameter ϕ,

$$f(y) \quad \equiv \quad f(y|\theta, \phi) \;=\; \exp\left\{\phi^{-1}[y\theta - \psi(\theta)] + c(y, \phi)\right\},$$

  where $\mu = \mathsf{E}[Y] = \psi'(\theta)$ and $\mathsf{Var}(Y) = \phi\psi''(\theta)$.

- a linear predictor vector $\boldsymbol{\eta}$ given by

$$\boldsymbol{\eta} = X\boldsymbol{\beta}$$

  where $\boldsymbol{\beta}$ is a vector of p unknown parameters and $X = [\mathbf{x}_1, \ldots, \mathbf{x}_n]^T$, the design matrix;

- a link function $g(\cdot)$ relating the mean to the linear predictor, i.e.

$$g(\mu_i) = \eta_i = \mathbf{x}_i^T\boldsymbol{\beta}$$

Poisson regression models

If Y_i, $i = 1, \ldots, n$, are counts with means μ_i, the standard Poisson model assumes that $Y_i \sim \text{Pois}(\mu_i)$ with

$$f(y_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}$$

and

$$\text{E}(Y_i) = \mu_i \quad \text{and} \quad \text{Var}(Y_i) = \mu_i \quad \text{(too restrictive!)}$$

The canonical link function is the log

$$g(\mu_i) = \log(\mu_i) = \eta_i$$

and $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$.

For a well fitting model (Hinde and Demétrio, 1998a,b):

$$\text{Residual Deviance} \approx \text{Residual d.f.}$$

Overdispersion in GLM's

What if Residual Deviance $\gg$ Residual d.f.?

(i) Badly fitting model
- omitted terms/variables
- incorrect relationship (link)
- outliers

(ii) variation greater than predicted by model: $\Longrightarrow$ **Overdispersion**
- count data: $\text{Var}(Y) > \mu$
- counted proportion data: $\text{Var}(Y) > m\pi(1-\pi)$

Univariate Overdispersed Count Data

Y_i – counts with means λ_i (Hinde and Demétrio, 1998a,b)

Negative Binomial Type Variance

$$Y_i|\lambda_i \sim \text{Pois}(\lambda_i) \quad \text{with} \quad \log \lambda_i = \mathbf{x}_i^T \boldsymbol{\beta}$$

$$\text{E}(Y_i|\lambda_i) = \lambda_i \qquad \text{Var}(Y_i|\lambda_i) = \lambda_i$$

- no particular distributional form: $\text{E}(\lambda_i) = \mu_i$ and $\text{Var}(\lambda_i) = \sigma_i^2$

$$\text{E}(Y_i) = \mu_i \qquad \text{Var}(Y_i) = \mu_i + \sigma_i^2$$

- $\lambda_i \sim \Gamma(\alpha, \beta_i)$

$$\text{E}[Y_i] = \mu_i = \alpha\beta_i \quad \text{Var}(Y_i) = \alpha\beta_i(1+\beta_i) = \mu_i + \frac{\mu_i^2}{\alpha} \quad (NegBinII)$$

- $\lambda_i \sim \Gamma(\alpha_i, \beta)$

$$\text{E}[Y_i] = \mu_i = \alpha_i\beta \qquad \text{Var}(Y_i) = \mu_i(1+\beta) = \phi\mu_i \quad (NegBinI)$$

Poisson-normal model

Individual level random effect in the linear predictor

$$Y_i|b_i \sim \text{Pois}(\lambda_i) \quad \text{with} \quad \log \lambda_i = \mathbf{x}_i^T \boldsymbol{\beta} + b_i$$

where $b_i \sim N(0, d)$, which gives

$$\begin{aligned} \text{E}[Y_i] &= e^{\mathbf{x}_i^T \boldsymbol{\beta} + \frac{1}{2}d} := \mu_i \\ \text{Var}(Y_i) &= e^{\mathbf{x}_i^T \boldsymbol{\beta} + \frac{1}{2}d} + e^{2\mathbf{x}_i^T \boldsymbol{\beta} + d}(e^d - 1) = \mu_i + \mu_i(e^d - 1)\mu_i \end{aligned}$$

i.e. a variance function of the form

$$\text{Var}(Y_i) = \mu_i + k\mu_i^2$$

Longitudinal Overdispersed Count Data

Y_{ij}: the jth outcome for subject i, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$

$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{in_i})'$: the vector of measurements for subject i

Negative Binomial Type Variance extension

$$Y_{ij}|\lambda_{ij} \sim \text{Poi}(\lambda_{ij}),$$

$\boldsymbol{\lambda}_i = (\lambda_{i1}, \ldots, \lambda_{in_i})'$, with $\text{E}(\boldsymbol{\lambda}_i) = \boldsymbol{\mu}_i$ and $\text{Var}(\boldsymbol{\lambda}_i) = \boldsymbol{\Sigma}_i$

Unconditionally,

$$\text{E}(\mathbf{Y}_i) = \boldsymbol{\mu}_i, \quad \text{and} \quad \text{Var}(\mathbf{Y}_i) = M_i + \Sigma_i$$

where M_i is a diagonal matrix with the vector $\boldsymbol{\mu}_i$ along the diagonal

- the diagonal structure of M_i reflects the conditional independence assumption

  – all dependence between measurements on the same unit stem from the random effects

- components of $\boldsymbol{\lambda}_i$ independent – pure overdispersion model, without correlation between the repeated measures

- $\lambda_{ij} = \lambda_i \Rightarrow \text{Var}(\mathbf{Y}_i) = M_i + \sigma_i^2 J_{n_i}$

  – a Poisson version of compound symmetry

- also possible to combine general correlation structures between the components of $\boldsymbol{\lambda}_i$

Poisson-normal model extension – a GLMM

$$\begin{aligned} Y_{ij}|\mathbf{b}_i &\sim \text{Poi}(\lambda_{ij}), \\ \ln(\lambda_{ij}) &= \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i, \\ \mathbf{b}_i &\sim N(\mathbf{0}, D) \end{aligned}$$

$\mathbf{x}_{ij}$ and $\mathbf{z}_{ij}$: p- and q-dimensional vectors of known covariate values

$\boldsymbol{\beta}$: a p-dimensional vector of unknown fixed regression coefficients

Then, unconditionally,

$\boldsymbol{\mu}_i = \text{E}(\mathbf{Y}_i)$ has components:

$$\mu_{ij} = \exp\left(\mathbf{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\mathbf{z}'_{ij}D\mathbf{z}_{ij}\right)$$

and the variance-covariance matrix is

$$\text{Var}(\mathbf{Y}_i) = M_i + M_i\left(e^{Z_i D Z'_i} - J_{n_i}\right)M_i$$

Models Combining Overdispersion With Normal Random Effects

$$\begin{aligned} Y_{ij}|\theta_{ij}, \mathbf{b}_i &\sim \text{Poi}(\lambda_{ij}) \\ \lambda_{ij} &= \theta_{ij} \exp\left(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i\right) \\ \mathbf{b}_i &\sim N(\mathbf{0}, D) \\ \text{E}(\boldsymbol{\theta}_i) = \text{E}[(\theta_{i1}, \ldots, \theta_{in_i})'] &= \Phi_i \\ \text{Var}(\boldsymbol{\theta}_i) &= \Sigma_i \end{aligned}$$

Then, $\boldsymbol{\mu}_i = \text{E}(\mathbf{Y}_i)$ has components:

$$\mu_{ij} = \phi_{ij} \exp\left(\mathbf{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\mathbf{z}'_{ij} D \mathbf{z}_{ij}\right)$$

The variance-covariance matrix is

$$\text{Var}(\mathbf{Y}_i) = M_i + M_i \left(P_i - J_{n_i}\right) M_i$$

where the $(j,k)^{\text{th}}$ element of P_i is

$$p_{i,jk} = \exp\left(\frac{1}{2}\mathbf{z}'_{ij} D \mathbf{z}_{ik}\right) \frac{\sigma_{i,jk} + \phi_{ij}\phi_{ik}}{\phi_{ij}\phi_{ik}} \exp\left(\frac{1}{2}\mathbf{z}'_{ik} D \mathbf{z}_{ij}\right)$$

Estimation for the Poisson-normal and Combined Models

- random-effects models fitted by maximization of the marginal likelihood, by integrating out the random effects from conditional densities

- likelihood contribution of subject i is from:

$$f_i(\mathbf{y}_i|\boldsymbol{\beta}, D, \phi) \quad = \quad \int \prod_{j=1}^{n_i} f_{ij}(y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \phi) \; f(\mathbf{b}_i|D) \; d\mathbf{b}_i$$

- likelihood for $\boldsymbol{\beta}$, D, and ϕ:

$$L(\boldsymbol{\beta}, D, \phi) \quad = \quad \prod_{i=1}^{N} \int \prod_{j=1}^{n_i} f_{ij}(y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \phi) \; f(\mathbf{b}_i|D) \; d\mathbf{b}_i.$$

- key problem: presence of N integrals – in general no closed-form solution exists (Verbeke and Molenberghs, 2000; Molenberghs and Verbeke, 2005). To solve the problem, use of
  - numerical integration – SAS procedure NLMIXED
  - series expansion methods (penalized quasi-likelihood, marginal quasi-likelihood), Laplace approximation, etc – SAS procedure GLIMMIX
  - hybrid between analytic and numerical integration
- in some special cases (linear mixed effects model, Poisson-normal model), these integrals can be worked out analytically – also true for the combined model
- Fully Bayesian inferences

Full Marginal Density for the Combined Model

The joint probability of $\mathbf{Y}_i$ takes the form:

$$\mathsf{P}(\mathbf{Y}_i = \mathbf{y}_i) = \sum_{\mathbf{t}} \left[\prod_{j=1}^{n_i} \binom{y_{ij} + t_j}{y_{ij}} \binom{\alpha_j + y_{ij} + t_j - 1}{\alpha_j - 1} (-1)^{t_j} \beta_j^{y_{ij}+t_j} \right]$$

$$\times \exp\left(\sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{x}'_{ij} \boldsymbol{\beta} \right)$$

$$\times \exp\left\{ \frac{1}{2} \left[\sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{z}'_{ij} \right] D \left[\sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{z}_{ij} \right] \right\}$$

where $\mathbf{t} = (t_1, \ldots, t_{n_i})$ ranges over all non-negative integer vectors

– special cases can be obtained very easily

– usefully used to implement maximum likelihood estimation, with numerical accuracy governed by the number of terms included in the series

Partial marginalization

– integrate over the gamma random effects only, leaving the normal random effects untouched

The corresponding probability is:

$$\mathrm{P}(Y_{ij} = y_{ij}|\mathbf{b}_i) = \left(\begin{array}{c} \alpha_j + y_{ij} - 1 \\ \alpha_j - 1 \end{array} \right) \left(\frac{\beta_j}{1 + \kappa_{ij}\beta_j} \right)^{y_{ij}} \left(\frac{1}{1 + \kappa_{ij}\beta_j} \right)^{\alpha_j} \kappa_{ij}^{y_{ij}}$$

where $\kappa_{ij} = \exp[\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i]$

– we assume that the gamma random effects are independent within a subject – the correlation is induced by the normal random effects

– easy to obtain the fully marginalized probability by numerically integration the normal random effects out of $\mathrm{P}(Y_{ij} = y_{ij}|\mathbf{b}_i)$, using SAS procedure NLMIXED

Analysis of the Epilepsy Data

Y_{ij}: the number of epileptic seizures patient i experiences during week j of the follow-up period

t_{ij}: the time-point at which Y_{ij} has been measured, $t_{ij} = 1, 2, \ldots, 27$

models with random intercept

$$\ln(\lambda_{ij}) = \begin{cases} (\beta_{00} + b_i) + \beta_{01} t_{ij} & \text{if placebo} \\ (\beta_{10} + b_i) + \beta_{11} t_{ij} & \text{if treated} \end{cases}$$

where $b_i \sim N(0, d)$

or random intercept and random slope

$$\ln(\lambda_{ij}) = \begin{cases} (\beta_{00} + b_{1i}) + (\beta_{01} + b_{2i}) t_{ij} & \text{if placebo} \\ (\beta_{10} + b_{1i}) + (\beta_{11} + b_{2i}) t_{ij} & \text{if treated} \end{cases}$$

where $\mathbf{b}_i \sim N(0, D)$

- Formally comparing the models with random intercepts and random slopes in time with their counterparts with random intercepts only, produces likelihood ratio test statistics of
  - 205.5 in the Poisson-normal case and
  - 19.0 in the combined-model case.
- Thus, at the same time, the combined model strongly reduces the need for random slopes, but does not remove it.
- Indeed, when comparing the test statistics against their reference distribution, a 50:50 mixture of a χ^2_1 and χ^2_2 distribution, $p < 0.0001$ was obtained in both cases.
- Estimates of the parameters in the models with random effects versus those without random effects are rather different (random effects of a non-conjugate type).

Effect	Parameter	Poisson	Negative-binomial
		Estimate (s.e.)	Estimate (s.e.)
Intercept placebo	β_{00}	1.2662 (0.0424)	1.2594 (0.1119)
Slope placebo	β_{01}	−0.0134 (0.0043)	−0.0126 (0.0111)
Intercept treatment	β_{10}	1.4531 (0.0383)	1.4750 (0.1093)
Slope treatment	β_{11}	−0.0328 (0.0038)	−0.0352 (0.0101)
Negative-binomial par.	α_1	—	0.5274 (0.0255)
Negative-binomial par.	$\alpha_2 = 1/\alpha_1$	—	1.8961 (0.0918)
Var. of random int.	d	—	—
Difference in slopes	$\beta_{11} - \beta_{01}$	−0.0195 (0.0058;p =0.0008)	−0.0227 (0.0150;p =0.1310)
Ratio of slopes	β_{11}/β_{01}	2.4576 (0.8481;p =0.0038)	2.8085 (2.6070;p =0.2815)

Effect	Parameter	Poisson-normal (RI)	Combined (RI)
		Estimate (s.e.)	Estimate (s.e.)
Intercept placebo	β_{00}	0.8179 (0.1677)	0.9112 (0.1755)
Slope placebo	β_{01}	−0.0143 (0.0044)	−0.0248 (0.0077)
Intercept treatment	β_{10}	0.6475 (0.1701)	0.6555 (0.1782)
Slope treatment	β_{11}	−0.0120 (0.0043)	−0.0118 (0.0074)
Negative-binomial par.	α_1	—	2.4640 (0.2113)
Negative-binomial par.	$\alpha_2 = 1/\alpha_1$	—	0.4059 (0.0348)
Var. of random int.	d	1.1568 (0.1844)	1.1289 (0.1850)
Difference in slopes	$\beta_{11} - \beta_{01}$	0.0023 (0.0062;p =0.7107)	0.0130 (0.0107;p =0.2260)
Ratio of slopes	β_{11}/β_{01}	0.8398 (0.3979;p =0.0376)	0.4751 (0.3445;p =0.1591)

Effect	Parameter	Poisson-normal (RI+RS)	Combined (RI+RS)
		Estimate (s.e.)	Estimate (s.e.)
Intercept placebo	β_0	0.8943 (0.1789)	0.9233 (0.1795)
Slope placebo	β_1	−0.0272 (0.0099)	−0.0286 (0.0102)
Intercept treatment	β_0	0.6498 (0.1835)	0.6679 (0.1835)
Slope treatment	β_2	−0.0165 (0.0102)	−0.0161 (0.0103)
Negative-binomial par.	α_1	—	2.7913 (0.2604)
Negative-binomial par.	$\alpha_2 = 1/\alpha_1$	—	0.3583 (0.3343)
Var. of random int.	d_{00}	1.2752 (0.2208)	1.1799 (0.2212)
Corr. random int. and slopes	ρ_{01}	−0.3341 (0.1312)	−0.2480 (0.1786)
Var. of random slopes	d_{11}	0.0024 (0.0006)	0.0016 (0.0006)
Difference in slopes	$\beta_{11} - \beta_{01}$	0.0107 (0.0140;p =0.4460)	0.0125 (0.0142;p =0.3834)
Ratio of slopes	β_{11}/β_{01}	0.6065 (0.4286;p =0.1607)	0.5645 (0.4054;p =0.1673)

- negative-binomial improvement over standard Poisson
- Poisson-normal improvement over standard Poisson
- combined model further improvement
- impact on point and precision estimates as the slope difference and the slope ratio
- Poisson: $p < 0.01$ for the slope difference, $p < 0.01$ for the slope ratio
- negative-binomial: $p = 0.13$ for the slope difference, $p = 0.28$ for the slope ratio
- Poisson-normal: $p = 0.71$ (RI) and $p = 0.44$ (RI +RS) for the slope difference, $p = 0.04$ (RI) and $p = 0.16$ (RI +RS) for the slope ratio
- combined model: $p = 0.22$ (RI) and $p = 0.38$ (RI +RS) for the difference, $p = 0.16$ (RI) and $p = 0.17$ (RI +RS) for the slope ratio

Correlation functions (Vangeneugden et al, 2011)

- Gamma random effects are assumed independent, need to consider the Poisson-normal and combined cases
- The fixed-effects structure is not constant, depends on time, need for the general correlation function
- For the Poisson-normal case, and for the placebo group

$$\text{Corr}(Y(t), Y(s)) = \frac{35.58 \cdot 0.99^{t+s}}{\sqrt{(4.04 \cdot 0.99^{t} + 35.58 \cdot 0.97^{t}) \cdot (4.04 \cdot 0.99^{s} + 35.58 \cdot 0.97^{s})}}$$

where $Y(t)$ represents the outcome for an arbitrary subject at time t

		Smallest value		Largest value	
Model	Arm	ρ	time pair	ρ	time pair
Poisson-normal, RI	placebo	0.8577	26 & 27	0.8960	1 & 2
Poisson-normal, RI	treatment	0.8438	26 & 27	0.8794	1 & 2
Combined, RI	placebo	0.8259	26 & 27	0.8981	1 & 2
Combined, RI	treatment	0.8383	26 & 27	0.8744	1 & 2
Poisson-normal, RI+RS	placebo	0.2966	1 & 27	0.9512	26 & 27
Poisson-normal, RI+RS	treatment	0.2936	1 & 27	0.9530	26 & 27
Combined, RI+RS	placebo	0.4268	1 & 27	0.9281	26 & 27
Combined, RI+RS	treatment	0.4225	1 & 27	0.9329	26 & 27

- For models with only random intercepts, the correlations range over a narrow interval; they are rather high and there is little difference between the Poisson-normal and combined models.
- For models with random intercepts and random slopes, several differences become apparent.
  - the values exhibit a much broader range between their smallest and largest values.
  - the range is somewhat over-estimated by the Poisson-normal model, which then narrows for the combined model, incorporating overdispersion effects, random intercepts, and random slopes.
- The random slope allows for the correlation to range over a considerable interval, while the overdispersion effect avoids the range to be overly wide.

- The Poisson-normal model forces the correlation and overdispersion effects to stem from a single additional parameter, the random-intercept variance d. Thus, considerable overdispersion also forces the correlation to increase, arguably beyond what is consistent with the data.

- In the combined model, there are *two* additional parameters, giving proper justice to both correlation and overdispersion effects.

- Half the subjects have missing measurements after the 16th week. This provides an additional motivation for the proposed model and its likelihood-based estimation, because, under the assumption of missingness at random inferences are valid.

- A corresponding analysis for a fully marginal model poses complex challenges (Molenberghs and Verbeke 2005).

Concluding Remarks

- normal random effects, to induce association between repeated Poisson data, and gamma random effects in the log-linear predictor for the overdispersion, integrated in the combined model

- special cases: standard negative-binomial and Poisson-normal models for repeated measures and univariate outcomes

- explicit expressions for means, variances, covariances and corelations of the combined model were derived

- closed form solutions obtained for the joint marginal probability of the outcome vector

- maximum likelihood estimation by integrating over the random effects, analytically, implemented in SAS procedure NLMIXED, or by combining analytic and numeric techniques

- epileptic seizures data analysis
  - impact on the conclusions about key scientific parameters
  - the correlations derived from the more conventional but also more restricted Poisson-normal model can be highly misleading – suggests a high within-patient correlation among any two time points within any of the two treatment arms
  - the correlations from the combined model are small to moderate

- general framework, encompassing the binary and Poisson types, in Molenberghs et al (2010)

Why to use a mixed-model approach when marginal correlation is of interest?

- non-likelihood based methods such as GEE treat correlation as nuisance parameters, they cannot be used for inferential purposes
- full likelihood methods may be highly prohibitive in terms of computational requirements
- the correlation ranges attainable in marginal models may be highly restricted,
- in a number of special but important cases, such as exchangeable clustered data, the entire range of positive correlations can be reached

References

Hinde, J. and Demétrio, C.G.B. (1998a) Overdispersion: Models and estimation. *Computational Statistics and Data Analysis*, **27**, 151–170.

Hinde, J. and Demétrio, C.G.B. (1998b) *Overdispersion: Models and Estimation.* São Paulo: XIII Sinape.

Molenberghs, G. and Verbeke, G. (2005) *Models for Discrete Longitudinal Data.* New York: Springer.

Molenberghs, G., Verbeke, G., and Demétrio, C.G.B. (2007) An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis* **13**, 513–531.

Molenberghs, G., Verbeke, G., Demétrio, C.G.B., Vieira, A. (2010) A Family of Generalized Linear Models for Repeated Measures With Normal and Conjugate Random Effects. *Statistical Science*, **25**, 325–347.

Vangeneugden, T., Molenberghs, G., Verbeke, G., and Demétrio, C.G.B. (2011) Marginal correlation in longitudinal binary data based on generalized linear mixed models. *Journal of Applied Statistics*, **38**: 215-232.