Joint modeling of hierarchically clustered and overdispersed non-gaussian continuous outcomes for comet assay data
Peer-reviewed author version

Joint Modeling of Hierarchically Clustered and Overdispersed Non-Gaussian Continuous Outcomes for Comet Assay Data

Aklilu Habteab Ghebretinsae[1]   **Christel Faes**[1]

Geert Molenberghs[1,2]   **Helena Geys**[3,1]   **Bas-Jan Van der Leede**[3]

[1] *I-BioStat, Universiteit Hasselt, B-3590 Diepenbeek, Belgium*

[2] *I-BioStat, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

[3] *Janssen Pharmaceutica, Turnhoutseweg 40, B-2430 Beerse, Belgium*

Abstract

Multivariate longitudinal or clustered data are commonly encountered in clinical trials and toxicological studies. Typically there is no single standard endpoint to assess the toxicity or efficacy of the compound of interest, but co-primary endpoints are available to assess the toxic effects or the working of the compound. Modeling the responses jointly is thus appealing to draw overall inferences using all responses and to capture the association among the responses. Non-Gaussian outcomes are often modeled univariately using exponential family models. To accommodate both the overdispersion and hierarchical structure in the data, Molenberghs *et al.* (2010) proposed using two separate sets of random effects. This paper considers a further extension to a multivariate setting with hierarchically clustered and overdispersed non-Gaussian outcomes. Gamma random effect for the over-dispersion and normal random effects for the clustering in the data are being used. The two outcomes are jointly analyzed by assuming that the normal random effects for both endpoints are correlated. The association structure between the response is analytically derived. The fit of the joint model to data from a so-called comet assay are compared with the univariate analysis of the two outcomes.

Some Keywords: Comet Assay; Gamma random effect; Hierarchical model; Joint model; Normal random effect.

1   Introduction

Multivariate longitudinal or clustered data are commonly encountered in clinical trials and toxicological studies. Typically, there is no single standard endpoint to assess the toxicity or efficacy of the compound of interest, but multiple endpoints, the so-called co-primary endpoints, are available to assess the toxic effects or the activity of the compound. In a comet assay, for instance, different

outcomes/co-primary endpoints (Lovell and Omori 2008, Wiklund and Agurell 2003) are used to assess the DNA damage of a cell as a result of an exposure: the tail length, tail intensity, and tail moment; these outcomes will formally be introduced in the next section. Most often, the tail length and tail intensity are used. Typically, univariate analyses are conducted to asses the treatment effect on each endpoint separately, leading to as many conclusions as there are endpoints regarding the same treatment effect. In particular, for the comet assay, one tends to focus primarily on tail intensity because of its discriminative power. Ideally though, one would prefer to reach a conclusion on the overall effect using all outcomes simultaneously, necessitating joint modeling. An added value of joint modeling is that inferences can be drawn about the association between outcomes as well.

Various modeling approaches for specifying a joint distribution are possible (Fitzmaurice *et al.*, 2009, Ch. 14; Fieuws and Verbeke, 2004). First, this can be effectuated by *specifying the full multivariate distribution* of the outcomes. This allows for drawing marginal inferences regarding the characteristics of the individual outcomes, but it requires many parameters and while the multivariate Gaussian distribution is well-known, there are many distributions for which no commonly accepted multivariate distribution is available. Second, by the use of *conditional models* where the joint distribution is expressed as the product of the conditional distribution of the first outcome conditional on the second outcome and the marginal distribution of the second outcome. However, factorization can be done in many ways, leading to different results, and it requires the specification of many parameters. Third, *shared-parameter models* can be entertained, where a pair of outcomes are associated by using a common latent variable, e.g., a common random effect. This is a simple but very strong assumption about the association between outcomes. Fourth, one can relax the latter assumption by using *multivariate random effects*, in which the two outcomes are associated via separate but correlated random effects. This is more flexible than shared-parameter models, but might still fail to fully capture the association structure and/or the variance function. Fifth, dimension reduction using principal component can be used, upon which the principal components are subjected to univariate analysis. While simple, the resulting inferences may not be about the parameters of direct scientific interest. In this manuscript, we focus on a flexible multivariate random effects approach.

The paper is organized as follows. In Section 2, the comet data are introduced. The joint model for

two hierarchical, overdispersed non-Gaussian outcomes is outlined in Section 3, and characteristics of the models are derived. It is then applied to comet data in Section 4.

2   Data Description

A comet assay, regularly encountered in pre-clinical research (Ejchart and Sadlej-Sosnowska 2003, Lovell and Omori 2008), is a sensitive method to assess DNA damage. During the last decade the assay gained widespread use in various areas and has emerged as a standard tool in the pharmaceutical for the assessment of the safety of potential new drugs. Typically, a comet assay is a single cell microgel electrophoresis method detecting DNA damage in any target tissue or organ of which a single cell suspension can be prepared. Cells are embedded in agarose, membranes are lysed and proteins extracted. Exposure to high alkali (pH>13.0) allows expression of single strand breaks and subsequent alkaline electrophoresis ensures migration of DNA fragments out of the nucleus. Visualization of this DNA migration (typical comet-like structures) is performed by a fluorescent dye. An image analysis system coupled to a microscope permits quantification of DNA damage at the single cell level. Three measures are commonly used: the tail migration (i.e., tail length), percentage tail intensity, and tail moment. Tail length is the distance from the perimeter of the comet head to the last visible point in the tail; percentage tail intensity is the percentage of DNA fragments present in the tail, while tail moment is the product of the amount of DNA in the tail and the mean distance of migration in the tail. In many experiments, the cells from a single animal are placed on a number of slides. Although there is no consensus among the experts as to the most appropriate statistical method and design (the number of slides and the replicates), some studies (Wiklund and Agurell 2003, Smith *et al.* 2008) indicate 3 slides and about 50 replicates/cells would be appropriate.

The statistical analysis of such a comet assay is complicated because of several issues in the data. The comet assay represents a hierarchical design with animals nested within doses, a number of slides per animals and several cells measured per slide. Comet measures from an animal are oftentimes not normally distributed but are rather asymmetric, skewed, bi- or multi-modal, a mixture of different distributions, etc. The complications that arise from the various non-normal distributions of comet endpoints are avoided in most standard analyses through the use of the central limit theorem. While

3

the original data at the cell level may not be normally distributed, mean (or median) summaries at slide or animal level will be approximately normally distributed (given the typically large sample sizes) and are thus amenable to standard statistical analyses. Hierarchical or multilevel models make use of information on the various levels of variability but may be quite complex in terms of the distribution between cells of the same animals and difficult to interpret and explain. Their advantage, however, is that they provide estimates of the variability at each level and make use of the information at the cell level thus increasing the power of the study especially if the between-animal variability is not too large. Variability is expected between slides because of the variability in the handling of the different slides, and also variability between animals is expected, because of the individual-specific differences. Previous (univariate) analyses on the same data indicate the importance of slide variability in contrast to the smaller rat variability (Ghebretinsae *et al.* 2011).

Here the data correspond to 24 male rats that received a daily oral dose of a compound. Rats were randomized to either one of the three dose levels (low, medium, and high) or the vehicle (control). On the day of necropsy, an extra group of three animals received a single dose of a positive control (Pos. C.). The animals were sacrificed 3 hours after the last dose administration, their livers were removed and processed for the comet assay. For each animal, a cell suspension is prepared. From each cell suspension, three replicate samples were prepared for scoring. Fifty randomly selected, non-overlapping cells per sample were then scored for DNA damage using a semi-automated scoring system. Thus, a total of 150 liver cells were scored per animal, on three slides. Data for tail intensity and tail length are represented in Figure 1. Previous work on the same data indicate dispersion is more pronounced for tail length, due in part to the occurrence of zeroes.

3   Joint Model for Two Hierarchical, Overdispersed Positive Outcomes

In this section, a joint model for hierarchical, overdispersed positive outcomes is proposed. First, the setting of a single hierarchical, overdispersed outcome is introduced, which is then extended to the multivariate setting.
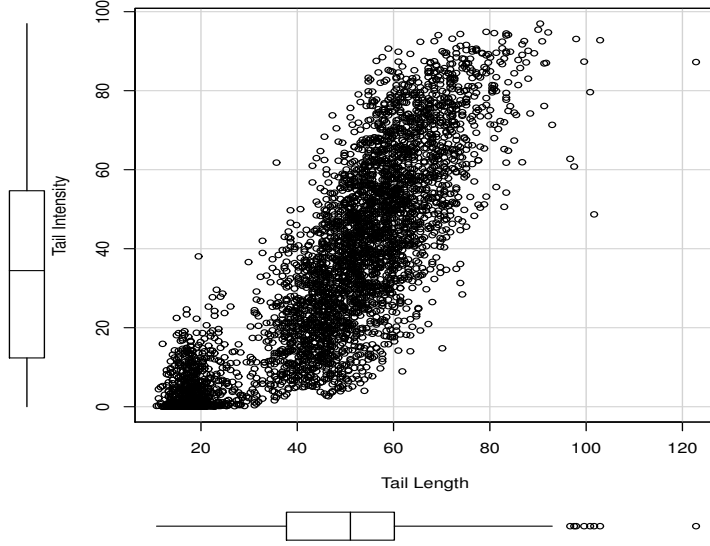
4

Figure 1: *Scatter plot and box plots of the tail length versus tail intensity*

3.1 Univariate Analysis

Because the primary outcomes, tail intensity and tail length, are skewed, non-negative and continuous, which is similar to many time-to-event data (Duchateau and Janssen 2007), an exponential or Weibull distribution is a natural choice. To start, we account for one level in the hierarchy of the data, namely the variability between slides. As proposed by Ghebretinsae *et al* (2011), we use a combined Weibull model with normal random effects to handle the hierarchy in the data and a gamma conjugate random effect to account for overdispersion in the response. This model falls into the model family as proposed by Molenberghs *et al.* (2010).

Let Y_{ij} be the jth cell of subject i measured for tail length or tail intensity, grouped in to $\boldsymbol{Y}_i$. The model can be expressed as:

$$f(\boldsymbol{y}_i|\boldsymbol{\theta}_i, \boldsymbol{b_i}) = \prod_{j=1}^{n_i} \lambda \rho \theta_{ij} y_{ij}^{\rho-1} e^{\eta_{ij}} e^{-\lambda y_{ij}^\rho \theta_{ij} e^{\eta_{ij}}},$$

$$\eta_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\xi} + \boldsymbol{z}_{ij}'\boldsymbol{b_i},$$

$$f(\boldsymbol{\theta}_i) = \prod_{j=1}^{n_i} \frac{1}{\beta_j^{\alpha_j}\Gamma(\alpha_j)} \theta_{ij}^{\alpha_j-1} e^{-\theta_{ij}/\beta_j},$$

$$f(\boldsymbol{b_i}) = \frac{1}{(2\pi)^{q/2}|D|^{1/2}}e^{-\frac{1}{2}\boldsymbol{b_i}'D^{-1}\boldsymbol{b_i}}.$$

with $\boldsymbol{b_i}$ the zero-mean normally-distributed slide-specific random effects, with variance-covariance D, to account for the clustering of observations and θ_{ij} the gamma-distributed measurement-specific random effects to accommodate for overdispersion. Further, λ and ρ are Weibull parameters, and α_j and β_j are gamma parameters. Here, η_{ij} is a linear predictor, with fixed-effects parameter $\boldsymbol{\xi}$ and design vectors $\boldsymbol{x}_{ij}$ and $\boldsymbol{z}_{ij}$ for the fixed effects and random effects, respectively. This model is called the Weibull-Gamma-Normal model.

3.2 Joint Analysis

The proposed joint model for tail length and tail intensity assumes a Weibull-Gamma-Normal model for both endpoints. The endpoints are associated by the use of bivariate normal random effects for the two endpoints, instead of the use of two separate (univariate) random effects, which we will call the Weibull-Gamma-Multivariate Normal model.

Let Y_{1ij} and Y_{2ij} be the jth measurements of subject i for the two outcomes, tail length and tail intensity. With notation similar to the above, the linear part for the two responses are assumed to be:

$$\eta_{1ij} = \boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1 + b_{1i},$$
$$\eta_{2ij} = \boldsymbol{x}'_{2ij}\boldsymbol{\xi}_2 + b_{2i},$$

with $\boldsymbol{x}_{1ij}$ and $\boldsymbol{x}_{2ij}$ design matrices, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ vectors of unknown fixed-effect parameters, and b_{1i} and b_{2i} the cluster-specific random intercepts for the first and second outcomes, respectively. These two random effects are assumed bivariate normally distributed:

$$\begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} d_1^2 & rd_1d_2 \\ rd_1d_2 & d_2^2 \end{pmatrix} \right],$$

with d_1^2 and d_2^2 the variances of the random intercepts and r the correlation between them. The association between the two endpoints is induced via the parameter r. More details on this are given in the next section. Conditionally on the normally distributed random effects (b_{1i} and b_{2i}),

6

it is assumed that the two outcomes are independent. Testing for treatment effect based on both endpoints simultaneously is conveniently done by way of a likelihood ratio test for the treatment effect parameters in both endpoints combined.

In terms of estimation, we opt for maximum likelihood using partial marginalization. This implies that the gamma random effects are analytically integrated out from the likelihood, while numerical integration, as implemented in the SAS procedure MIXED, is invoked to marginalize over the normally distributed random effects. The code is given in the Appendix.

3.3   Correlation Between Both Responses

The association between both outcomes is captured via the bivariate normal random effects. However, the correlation between the two random effects is not necessarily equal to the correlation between the two responses. Furthermore, a significant correlation at the cluster level does not necessarily imply a significant correlation between the two responses taken from the same cell. In this section, it is established how the correlation between the outcomes is related with the correlation between the random effects.

The correlation between two measurements from the same subject for a single response, also called the intraclass correlation (ICC), is equal to:

$$\text{Corr}(Y_{\ell i j}, Y_{\ell i k}) = \left[e^{\frac{d_\ell^2}{\rho^2}} - 1 \right] \bigg/ \left[\frac{2\rho B(\alpha - \frac{2}{\rho}, \frac{2}{\rho}) e^{\frac{d_\ell^2}{\rho^2}}}{B(\alpha - \frac{1}{\rho}, \frac{1}{\rho})^2} - 1 \right]$$

with $\ell = 1, 2$, ρ the shape parameter of the Weibull distribution, d_ℓ^2 the random-effects variance (equal to d_1^2 or d_2^2 for tail length and tail intensity, respectively), and α the shape parameter of the Gamma random effects distribution. $B(\cdot, \cdot)$ is the beta function. A large value for the shape parameter (α) indicates a small amount of overdispersion, which in the limit reduces to the Weibull-Normal model for a univariate outcome. In this case, the intraclass correlation reduces to:

$$\text{Corr}(Y_{\ell i j}, Y_{\ell i k}) = \left[e^{\frac{d_\ell^2}{\rho^2}} - 1 \right] \bigg/ \left[\frac{2\rho \Gamma(\frac{2}{\rho}) e^{\frac{d_\ell^2}{\rho^2}}}{\Gamma(\frac{1}{\rho})^2} - 1 \right],$$

with $\Gamma(\cdot)$ the gamma function.

On the other hand, the correlation between the two outcomes (tail length and tail intensity) of the same cell is given by the following expression

$$
\mathsf{Corr}(y_{1ij}, y_{2ij}) = \left(e^{\frac{rd_1 d_2}{\rho_1 \rho_2}} - 1 \right)
$$

$$
\times \frac{B\left(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1}\right)}{\left[2\rho_1 B\left(\alpha_1 - \frac{2}{\rho_1}, \frac{2}{\rho_1}\right) e^{\frac{d_1^2}{\rho_1^2}} - B\left(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1}\right)^2 \right]^{1/2}}
$$

$$
\times \frac{B\left(\alpha_2 - \frac{1}{\rho_2}, \frac{1}{\rho_2}\right)}{\left[2\rho_2 B\left(\alpha_2 - \frac{2}{\rho_2}, \frac{2}{\rho_2}\right) e^{\frac{d_2^2}{\rho_2^2}} - B\left(\alpha_2 - \frac{1}{\rho_2}, \frac{1}{\rho_2}\right)^2 \right]^{1/2}}
$$

$$
= \frac{\left(e^{\frac{rd_1 d_2}{\rho_1 \rho_2}} - 1 \right)}{\left(e^{\frac{d_1^2}{\rho_1^2}} - 1 \right)^{\frac{1}{2}} \left(e^{\frac{d_2^2}{\rho_2^2}} - 1 \right)^{\frac{1}{2}}} \sqrt{\mathsf{ICC}_1} \sqrt{\mathsf{ICC}_2}, \tag{1}
$$

where ICC_1 and ICC_2 are the intracluster/class correlation for responses 1 and 2, respectively. The correlation between the two endpoints is proportional to the correlation between the two random effects, with the same sign. So, when two random effects are positively or negatively correlated, the correlation between endpoints follows accordingly and when the correlation between the two random effects is zero, then the correlation between the two endpoints is zero as well. In other words, the correlation is induced entirely by the correlation between the two random effects. This correlation also depends on the Weibull shape parameters ρ_1 and ρ_2.

For a joint model based on two linear mixed model, the bivariate correlation between the two endpoints is given by $\mathsf{Corr}(Y_{1ij}, Y_{2ij}) = r\sqrt{\mathsf{ICC}_1}\sqrt{\mathsf{ICC}_2}$ (Fitzmaurice *et al.*, 2009, Ch. 14). It is by definition smaller than or equal to the correlation between the two random intercepts. Only when both intra-class correlations are 1, equality holds. However, it is not straightforward in this case. Details on the calculations are given in Appendix B.

3.4 Simulation Study

A set of simulations was conducted to evaluate the performance of the different models in terms of the type I and II error rates, as well as bias of the parameter estimates. Two batches of simulation

were done for two levels of overdispersion. We considered two treatment groups: active and vehicle control. In the first set of simulations, we assume there are 3 animals, hence 9 slides, and a total of 450 cells in each treatment group. In contrast, in the second batch, 6 animals and 18 slides each with 10 cells, are considered. Two responses were generated, one with and the other without overdispersion. The first response $Y_{1ij} \sim$ Weibull$(\rho_1, \theta_{ij} e^{\eta_{1ij}})$ follows a Weibull-gamma-normal, while $Y_{2ij} \sim$ Weibull$(\rho_2, e^{\eta_{2ij}})$ follows a Weibull-normal, with further

$$\eta_{kij} = \beta_{k0} + \beta_{k1} T_{ij} + b_{ki}, \qquad (k = 1, 2)$$

$$\theta_{ij} \sim \text{Gamma}\left(\alpha, \frac{1}{\alpha}\right).$$

T_{ij} is the indicator for the treatment group. Random effects are correlated and follow:

$$\begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} d_1^2 & r d_1 d_2 \\ r d_1 d_2 & d_2^2 \end{pmatrix} \right].$$

Different correlation levels between the random effects as well as different overdispersion level were considered to gauge the impact of these characteristics. We set the random-effects standard deviations to $d_1 = d_2 = 0.2$ and the Weibull shape parameters to $\rho_1 = \rho_2 = 0.4$. The correlation r ranges over 0.9, 0.6, and 0.3 in both sets of simulations. Because interest lies in assessing the type I and II error rates, the data are generated under the null ($\beta_{10} = \beta_{11} = -1$ and $\beta_{21} = \beta_{21} = -1$) for the type I error rate, and under the alternative ($\beta_{10} = -1$, $\beta_{11} = -1.3$, $\beta_{20} = -1$, and $\beta_{21} = -1.3$) to assess the type II error rate. An overdispersion level of $\alpha = 0.8$ is used for the first set of simulation and $\alpha = 1.5$ for the second one. A total of 200 such datasets is generated per run. The two responses are analyzed separately using: (1) a traditional model, i.e., analysis of variance on the summary measure (mean) of the log-transformed response; (2) a classical Weibull model; (3) a Weibull-normal model; (4) a Weibull-gamma-normal model; and finally (5) Joint modeling.

4 Application to the Comet Data

4.1 Univariate Analyses

Univariate analyses for tail intensity and tail length are performed separately for the comet assay data. The endpoints are analyzed both with and without overdispersion, using the Weibull-Gamma-

Table 1: *Analysis for Tail Length*

Effect	Par.	Weibull-Gamma-Normal Estimate(s.e.)	Weibull-Gamma-Normal p-value	Weibull-Normal Estimate(s.e.)	Weibull-Normal p-value
Veh.	β_0	-30.9295(0.7264)	0.0001	-15.6378(0.2517)	0.0001
Low *vs.* veh.	β_1	-11.9378(0.4445)	0.0001	-4.4965(0.2243)	0.0001
Med.*vs.* veh.	β_2	-12.1552(0.4472)	0.0001	-4.5998(0.2245)	0.0001
High *vs.* veh.	β_3	-12.6026(0.4525)	0.0001	-4.8290(0.2251)	0.0001
Pos. C.*vs.* veh.	β_4	-9.6419(0.4762)	0.0001	-3.4808(0.2718)	0.0001
Low *vs.* Med.	β_5	-0.2174(0.3398)	0.5241	-0.1033(0.2206)	0.6410
Low *vs.* High	β_6	-0.6648(0.3403)	0.0543	-0.3325(0.2206)	0.1358
Med. *vs.* High	β_7	-0.4474(0.3402)	0.1923	-0.2292(0.2206)	0.3019
Weibull Par.	ρ	10.7072(0.2474)	0.0001	4.9585(0.0580)	0.0001
s.d. of RE	$\sqrt{d}$	0.9881(0.08592)	0.0001	0.6464(0.0543)	0.0001
OD par.	α	0.8932(0.0463)	0.0001	—	—
-2 loglik.		28069		29793	

Normal and Weibull-Normal models, respectively. Summary results are presented in Tables 1 and 2. For tail intensity, inclusion of the overdispersion random effect neither improved the likelihood, nor affected the estimates and precision of the estimate. On the other hand, for tail length, inclusion of the overdispersion random effect greatly improved the likelihood and also affected the parameter estimation and precision. If we consider the contrast between the low and high dose group for tail length, for instance, the p-value was 0.1358 based on the model without overdispersion and 0.0543 with overdispersion. For both endpoints, there is a major effect of the compound as compared to the vehicle group. However, the conclusion for the contrasts between the three dose level is different based on both responses. Based on tail intensity, there was a significant difference among the dose levels.

A conventional significance test for α would test the null hypothesis $H_0 : \alpha = 0$. However, this does

Table 2: *Analysis for Tail Intensity*

Effect	Par.	Weibull-Gamma-Normal		Weibull-Normal	
		Estimate(s.e.)	p-value	Estimate(s.e.)	p-value
Veh.	β_0	-2.4628(0.0774)	0.0001	-2.4628(0.0774)	0.0001
Low *vs.* veh.	β_1	-2.8125(0.0911)	0.0001	-2.8126(0.0911)	0.0001
Med.*vs.* veh.	β_2	-3.0565(0.0920)	0.0001	-3.0566(0.0920)	0.0001
High *vs.* veh.	β_3	-3.2777(0.0929)	0.0001	-3.2778(0.0929)	0.0001
Pos. C. *vs.* veh.	β_4	-1.7941(0.1079)	0.0001	-1.7941(0.1078)	0.0001
Low *vs.* Med.	β_5	-0.2440(0.0874)	0.0065	-0.2440(0.0874)	0.0065
Low *vs.* High	β_6	-0.4652(0.0875)	0.0001	-0.4652(0.0875)	0.0001
Med. *vs.* High	β_7	-0.2212(0.0874)	0.0133	-0.2212(0.0874)	0.0133
Weibull Par.	ρ	1.4158(0.0189)	0.0001	1.4158(0.0189)	0.0001
s.d. of RE	$\sqrt{d}$	0.2201(0.0248)	0.0001	0.2201(0.0248)	0.0001
log OD par.	$\log(\alpha)$	13.9715(2.0370)	0.0001	—	—
-2 loglik.		33769		33769	

not correspond to the absence of overdispersion. Rather, overdispersion vanishes as α approaches infinity.

4.2 Analysis Based on a Combined Endpoint

It is often desirable to opt for a summary analysis of both endpoints, at least to avoid multiple and perhaps conflicting inferences from the univariate analyses. To this end, define tail moment as the product of both the mean distance of migration in the tail and amount of DNA in the tail (intensity). Although not directly the product of the two responses, it indirectly combines information from both endpoints. Also here, inclusion of overdispersion improved the fit and had impact on the parameter estimates as well as on the standard errors; see Table 3. Remember that the contrasts between low, medium and high doses was significant based on the univariate analysis using tail intensity, but not

Table 3: *Analysis for Tail Moment*

Effect	Par.	Weibull-Gamma-Normal		Weibull-Normal	
		Estimate(s.e.)	p-value	Estimate(s.e.)	p-value
Veh.	β_0	1.0946(0.1023)	0.0001	0.8294(0.0781)	0.0001
Low *vs.* veh.	β_1	-3.8954(0.1515)	0.0001	-3.4995(0.1159)	0.0001
Med.*vs.* veh.	β_2	-4.2326(0.1558)	0.0001	-3.8116(0.1171)	0.0001
High *vs.* veh.	β_3	-4.5767(0.1599)	0.0001	-4.1341(0.1186)	0.0001
Pos. C. *vs.* veh.	β_4	-2.2365(0.1594)	0.0001	-1.9365(0.1360)	0.0001
Low *vs.* Med.	β_5	-0.3372(0.1174)	0.0052	-0.3121(0.1104)	0.0059
Low *vs.* High	β_6	-0.6813(0.1180)	0.0001	-0.6346(0.1106)	0.0001
Med. *vs.* High	β_7	-0.3441(0.1174)	0.0044	-0.3224(0.1104)	0.0046
Weibull Par.	ρ	1.3199(0.0237)	0.0001	1.2429(0.0159)	0.0001
s.d. of RE	$\sqrt{d}$	0.3174(0.0316)	0.0001	0.2990(0.0291)	0.0001
log OD par.	$\log(\alpha)$	10.9408(2.6443)	0.0001	$-$	$-$
-2 loglik.		19980		20004	

using the endpoint tail length. Using tail moment, none of the contrasts are significant. This shows that some effects might be lost by summarizing two endpoints by a single endpoint.

4.3 Joint Analysis

The univariate analyses on the two endpoints (Section 4.1) lead to multiple inferences. The univariate analysis on the combined endpoint (Section 4.2) uses a summary endpoint, but which may not always be interpretable. It also renders impossible assessment of the association between the endpoints. As a third and appealing alternative, a joint analysis models both endpoints simultaneously and accommodates association between them. Conveniently, a test for the overall treatment effect based on both endpoints can be done using likelihood ratio tests.

12

The two preferred models in the univariate analyses, Weibull-Normal for tail intensity and Weibull-Gamma-Normal for tail length are now combined into a joint model by assuming that the normal random effects are correlated. Table 4 presents the results. The contrasts of interest based on each endpoints separately as well as on the overall effect based on both endpoints is provided. The estimates are slightly different from the univariate analyses. The three contrasts (low *versus* medium, low *versus* high, and medium *versus* high) have p-values of (0.5242,0.0556,0.1959) and (0.0055,0.0001,0.0119) based on the first and second endpoints and (0.0302,0.0001,0.0302) based on the two endpoints combined. The correlation between the two random intercepts was highly significant, and estimated as 0.6049 (s.e. 0.098). The intraclass correlations were estimated as 0.1991 (s.e. 0.02704) and 0.04180 (s.e. 0.0089) for tail length and tail intensity, respectively. As a result, the pairwise correlation is estimated as 0.05499 (s.e. 0.0129).

4.4   Simulation Results

The first simulation run is summarized in Table 5, with the rest deferred to the Appendix C. Generally, the type I error rate for all models was approximately the nominal one, except for the classical Weibull model. This could be ascribed to the independence assumption between the outcomes in this model. Indeed, ignoring the correlation may underestimate the standard errors (see Table 5). This has an adverse impact on the assessment of treatment effect, in the sense that a compound can easily erroneously be declared toxic. The error rate is higher for the first response where the hierarchical structure and overdispersion are omitted, in contrast to the second response.

We now turn to the power of the test. Analyzing the two responses using the various appropriate models has higher power when compared to the traditional model. The discrepancy between the proper and traditional models increases with decreasing variance of the random effects (results not presented here). This is not surprising because, when the variability between clusters is high, then the measurements within a cluster are similar. In that case, summarizing the observations has little impact. The shape parameter has an impact as well. When it gets smaller, the density becomes more skewed and the traditional approach, relying on normality and hence symmetry, drifts apart. On the other hand, the underestimated standard error when the simple classical model was employed

Table 4: *Joint Model, Weibull-Normal Model for Tail Intensity and Weibull-Gamma-Normal for Tail Length*

Effect	Par.	Tail Length Estimate(s.e.)	p-value	Tail Intensity Estimate(s.e.)	p-value	Overall G^2	p-value
Veh.	β_0	-29.0574(0.6537)	0.0001	-2.4620(0.0763)	0.0001	—	—
Low *vs.* veh.	β_1	-11.1543(0.4083)	0.0001	-2.8064(0.0893)	0.0001	248	0.0001
Med.*vs.* veh.	β_2	-11.3575(0.4107)	0.0001	-3.0502(0.0902)	0.0001	256	0.0001
High *vs.* veh.	β_3	-11.7724(0.4154)	0.0001	-3.2701(0.0912)	0.0001	264	0.0001
Pos. C.*vs.* veh.	β_4	-8.9600(0.4400)	0.0001	-1.7826(0.1056)	0.0001	171	0.0001
Low *vs.* Med.	β_5	-0.2032(0.3177)	0.5242	-0.2438(0.0854)	0.0055	7	0.0302
Low *vs.* High	β_6	-0.6181(0.3181)	0.0556	-0.4637(0.0856)	0.0001	25	0.0001
Med. *vs.* High	β_7	-0.4149(0.3181)	0.1959	-0.2199(0.0854)	0.0119	7	0.0302
log of Weib.P	ρ	10.0336(0.2210)	0.0001	1.4152(0.0189)	0.0001	—	—
s.d. of RE	$\sqrt{d}$	0.9227(0.0797)	0.0001	0.2133(0.0239)	0.0001	—	—
log of OD par.	α	1.0052(0.0517)	0.0001	—	—	—	—
Correlation	r	0.6049(0.0979)				0.0001	
-2 loglik		61824					

not only inflated the type I error rate, it also exaggerates the power of the test.

The parameter estimates are also biased and the bias was higher for the estimates of the first response with overdispersion. When the first response is analyzed with the Weibull-normal model that does not account for overdispersion, the power of the test was lower and the parameter estimates still biased, though they were slightly better than under the classical Weibull model. In fact, the power of the test for the traditional model was even better. On the other hand, analyzing the data coming from the Weibull-normal model by using the Weibull-gamma-normal, leads to the same results in terms of the type I error rate, the power of the test as well as the parameter estimates, underscoring

Table 5: *Type I error rate, power of the test and parameter estimate. Correlation between $r = 0.3$ and $\alpha = 0.8$. W: Weibull; G: gamma; N: normal; J: joint.*

	Response	Par.	Trad.	W	WN	WGN	JWGN
Power	1		0.38	0.57	0.26	0.435	0.448
	2		0.565	0.895	0.655	0.655	0.642
	Combined						0.743
Est.(s.e.)	1	β_{10}		-0.9635(0.0575)	-0.9977(0.0901)	-1.0049(0.1129)	-1.006(0.1171)
		β_{11}		-1.1066(0.0596)	-1.1443(0.09171)	-1.2977(0.1133)	-1.3073(0.1176)
		$\beta_{11} - \beta_{10}$		-0.1430(0.0671)	-0.1466(0.1173)	-0.2929(0.1528)	-0.3000(0.1586)
	2	β_{20}		-0.9965(0.0598)	-1.0085(0.0854)	-1.0085(0.0854)	-1.0079(0.0865)
		β_{21}		-1.2765(0.0646)	-1.2934(0.0890)	-1.2934(0.0890)	-1.3007(0.0892)
		$\beta_{21} - \beta_{20}$		-0.2801(0.0671)	-0.2849(0.1087)	-0.2849(0.1087)	-0.2918(0.1087)
Type I	1		0.05	0.335	0.065	0.06	0.065
	2		0.045	0.26	0.05	0.05	0.0365
	Combined						0.048
Est.(s.e.)	1	β_{10}		-0.9580(0.0574)	-0.9917(0.0894)	-0.9929(0.1116)	-0.9985(0.1145)
		β_{11}		-0.9728(0.0578)	-1.0050(0.0897)	-0.9941(0.1116)	-0.9911(0.1140)
		$\beta_{11} - \beta_{10}$		-0.0148(0.0670)	-0.0134(0.1162)	-0.0012(0.1499)	0.0101(0.1565)
	2	β_{20}		-0.9783(0.0595)	-0.9911(0.0859)	-0.9911(0.0859)	-0.9901(0.0856)
		β_{21}		-0.9927(0.0598)	-1.0039(0.0861)	-1.0039(0.0861)	-1.0053(0.0869)
		$\beta_{21} - \beta_{20}$		-0.0143(0.0667)	-0.0128(0.1096)	-0.0128(0.1096)	-0.0148(0.1120)

the importance of accommodating overdispersion. Given the elaborate nature of the joint analysis, it is not surprising that some convergence problems emerge. Its power is higher than that from the univariate analyses. The rise in power increases with decreasing correlation between the random effects. This is to be expected because lower correlation implies that a pair of outcomes is more informative. Finally, we also noted that as the cluster size grows larger, the power of the test is higher for all models (details not given).

5   Concluding Remarks

Co-primary endpoints are commonly used to assess the toxic effect of a certain compound in toxi-cological studies. Univariate analyses are often done on each endpoint separately; but this leads to multiple inferences. Joint modeling of the endpoints is appealing to make overall inferences as well as to capture the association among the outcomes. In this paper, joint model using a random-effect was presented in a bivariate setting with hierarchically clustered and overdispersed non-Gaussian continuous outcomes. Thus, the model accounts for: (1) overdispersion; (2) repeated measures over time; (3) and the multivariate nature of the outcomes.

Two Weibull-gamma-normal models were combined using bivariate normally distributed random effects. This is a simple and relatively less restrictive approach compared to a shared parameter model and it can be easily implemented in standard software like in the SAS procedure MIXED.

It was applied to the comet assay data which exhibit two outcomes namely tail length and tail intensity. Univariate analyses indicate that a model with overdisperion (Weibull-Gamma-Normal) is necessary for tail length and a model without overdispersion (Weibull-Normal) is sufficient for the tail intensity. The contrast between low, medium, and high dose level using the two endpoints leads to different conclusion.

References

Duchateau, L. and Janssen, P. (2007) *The Frailty Model*. New York: Springer.

Ejchart, A. and Sadlej-Sosnowska, N. (2003) Statistical evaluation and comparison of comet assay results. *Mutation Research*, **534**, 85–92.

Fieuws, S. and Verbeke, G. (2004) Joint modelling of multivariate longitudinal profiles: pitfalls of

the random effects approach. *Statistics in Medicine*, **23**, 3093–3104.

Fitzmaurice, G., Davidian, M. , Verbeke, G. and Molenberghs, G. (2009) *Longitudinal Data Analysis*. Chapman and Hall/CRC.

Ghebretinsae, A.H., Faes, C., Molenberghs, G., De Boeck, M., and Geys, H. (2011) A Bayesian, generalized frailty model for comet assays. *Submitted for publication.*

Lovell, D.P. and Omori, T. (2008) Statistical issues in the use of the comet assay. *Mutagenesis*, **23**, 171–182.

Molenberghs, G., Verbeke, G., Demétrio, C.G.B., and Vieira, A. (2010) A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science*, **25**, 325–347.

Smith, C.C., Adkins, D.J., Martin, E.A., and O'Donovan, M.R. (2008) Recommendations for the design of the rat comet assay. *Mutagenesis*, 1–8.

Wiklund, S.J. and Agurell, E. (2003) Aspects of design and statistical analysis in the Comet assay.*Mutagenesis*, **18**, 167–175.

Appendix

A Software Code

```
proc nlmixed data=comet2r maxiter=5000 qpoints=30;

/*specification of initial values*/

parms int0=-28  int125=-40  int25=-40.2 int5=-41 int200=-36 in0=-2.9

in125=-5.5  in25=-6.2 in5=-6.5 in200=-4.9

logrho1=1.4 logrho2=2 logalpha1=1 sigma1=0.6 sigma2=0.117 r=0.2;


/*specification of the linear part of the model for the first response*/

if (dose=0) then eta1=int0+b;

else if (dose=1.25) then eta1=int125+b;

else if (dose=2.5) then eta1=int25+b;

else if (dose=5) then eta1=int5+b;

else if (dose=200) then eta1=int200+b;


/*specification of the linear part of the model for the second response*/

if (dose=0) then eta2=in0+c;

else if (dose=1.25) then eta2=in125+c;

else if (dose=2.5) then eta2=in25+c;

else if (dose=5) then eta2=in5+c;

else if (dose=200) then eta2=in200+c;


/*use appropriate link functions*/

k1=exp(eta1);

k2=exp(eta2);

rho1=exp(logrho1);

rho2=exp(logrho2);
```

```
alpha1=exp(logalpha1);


/*likelihood specification */

if res=1 then

loglik=logrho1+log(response)*(rho1-1)+(alpha1+1)*logalpha1+eta1-

(alpha1+1)*log(alpha1+(response**(rho1))*k1);

else if

res=2 then loglik=logrho2+log(response)*(rho2-1)+eta2-((response**rho2)*k2);

model response~ general(loglik);


/*normal random effect specification*/

random b c ~normal([0,0],[sigma1**2,r*sigma1*sigma2,sigma2**2]) subject=slide_id;

/*some estimate statements of interest for first response*/

estimate 'Veh vs Low' int125 - int0;

estimate 'Veh vs Med' int25 - int0;

estimate 'Veh vs High' int5 - int0;

estimate 'Veh vs P. Control' int200-int0;

estimate 'low vs Med' int25 - int125;

estimate 'low vs High' int5- int125;

estimate 'Med vs high' int5-int25;


/*some estimate statements of interest for second response*/

estimate 'Veh vs Low' in125 - in0;

estimate 'Veh vs Med' in25 - in0;

estimate 'Veh vs High' in5 - in0;

estimate 'Veh vs P. Control' in200-in0;

estimate 'low vs Med' in25 - in125;

estimate 'low vs High' in5- in125;

estimate 'Med vs high' in5-in25;
```

```
/*Intraclass correlation and correlation between two responses estimate statements*/

betta1=beta(alpha1-(2/rho1),(2/rho1));

betta2=beta(alpha1-(1/rho1),(1/rho1));

ex1=exp((sigma1/rho1)**2);

estimate 'ICC1'   (ex1-1)/((2*rho1*betta1*ex1/(betta2**2))-1);

gamma1=gamma(2/rho2);

gamma2=gamma(1/rho2);

ex2=exp((sigma2/rho2)**2);

estimate 'ICC2' (ex2-1)/((2*rho2*gamma1*ex2/(gamma2**2))-1);

ex3=exp((r*sigma1*sigma2)/(rho1*rho2));

estimate 'Corr' (ex3-1)/sqrt(((2*rho1*betta1*ex1/(betta2**2))-1)*

((2*rho2*gamma1*ex2/(gamma2**2))-1));

run;
```

B   Derivation of the Correlation Between Both Endpoints

Let Y_{1ij} and Y_{2ij} be the jth measurements of subject i for outcome 1 and 2 respectively. The linear part for the two responses are:

$$\eta_{1ij} = \boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1 + b_{1i},$$

$$\eta_{2ij} = \boldsymbol{x}'_{2ij}\boldsymbol{\xi}_2 + b_{2i},$$

with

$$\begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} d_1^2 & rd_1d_2 \\ rd_1d_2 & d_2^2 \end{pmatrix} \right].$$

The correlation between the two endpoints is, by definition:

$$\text{Corr}(Y_{1ij}, Y_{2ij}) = \frac{\text{Cov}(Y_{1ij}, Y_{2ij})}{\sqrt{\text{Var}(Y_{1ij})}\sqrt{\text{Var}(Y_{2ij})}}. \tag{2}$$

We will re-write this as

$$\mathsf{Corr}(Y_{ij}, Y_{ik}) = \frac{T1}{\sqrt{T2}\sqrt{T3}}. \tag{3}$$

It is also known that

$$\mathsf{Cov}(Y_{ij}, Y_{ik}) = \mathsf{E}\left(\mathsf{Cov}(Y_{ij}, Y_{ik}|b_{1i}, b_{2i})\right) + \mathsf{Cov}\left[\mathsf{E}(Y_{ij}|b_{1i}, b_{2i}), \mathsf{E}(Y_{ik}|b_{1i}, b_{2i})\right], \tag{4}$$

which we denote as $T1 = L1 + L2$.

Given the random effect, the two measurements are independent. Therefore

$$\mathsf{Cov}(Y_{ij}, Y_{ik}|b_{1i}, b_{2i}) = 0.$$

By integrating the gamma random effect, we have:

$$f(y_{1ij}|b_{1i}) = \frac{\lambda_1\rho_1 y_{1ij}^{\rho_1-1} e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1 + b_{1i}}\alpha_1^{\alpha_1+1}}{(\alpha_1 + \lambda_1 y_{1ij}^{\rho_1} e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1 + b_{1i}})^{\alpha_1+1}}, \tag{5}$$

with similar formula for $f(y_{2ij}|b_{2i})$.

Further, we have,

$$f(y_{1ij}|b_{1i}, b_{2i}) = f(y_{1ij}|b_{1i}), \tag{6}$$

$$f(y_{2ij}|b_{1i}, b_{2i}) = f(y_{2ij}|b_{2i}). \tag{7}$$

Therefore, the conditional expectation is given by

$$\mathsf{E}(Y_{1ij}|b_{1i}) = \int y_{1ij} f(y_{1ij}) dy_{1ij} = \frac{\alpha_1^{\frac{1}{\rho_1}} B(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1})}{\rho_1(\lambda_1 e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1 + b_{1i}})^{\frac{1}{\rho_1}}}, \tag{8}$$

with a similar formula for $\mathsf{E}(Y_{2ij}|b_{2i})$.

The covariance between $\mathsf{E}(Y_{1ij}|b_{1i})$ and $\mathsf{E}(Y_{2ij}|b_{2i})$ is

$$\mathsf{Cov}[\mathsf{E}(Y_{1ij}|b_{1i}), \mathsf{E}(Y_{2ij}|b_{2i})] = \frac{\alpha_1^{\frac{1}{\rho_1}}\alpha_2^{\frac{1}{\rho_2}} B(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1}) B(\alpha_2 - \frac{1}{\rho_2}, \frac{1}{\rho_2})}{\rho_1(\lambda_1 e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1})^{\frac{1}{\rho_1}}\rho_2(\lambda_2 e^{\boldsymbol{x}'_{2ij}\boldsymbol{\xi}_2})^{\frac{1}{\rho_2}}}\mathsf{Cov}(e^{-\frac{b_{1i}}{\rho_1}}, e^{-\frac{b_{2i}}{\rho_2}}). \tag{9}$$

As a result:

$$T1 = L2 = \frac{\alpha_1^{\frac{1}{\rho_1}}\alpha_2^{\frac{1}{\rho_2}} B(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1}) B(\alpha_2 - \frac{1}{\rho_2}, \frac{1}{\rho_2})}{\rho_1(\lambda_1 e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1})^{\frac{1}{\rho_1}}\rho_2(\lambda_2 e^{\boldsymbol{x}'_{2ij}\boldsymbol{\xi}_2})^{\frac{1}{\rho_2}}} e^{\frac{1}{2}(\frac{d_1^2}{\rho_1^2} + \frac{d_2^2}{\rho_2^2})}[e^{\frac{rd_1d_2}{\rho_1\rho_2}} - 1], \tag{10}$$

$$T2 = \mathsf{Var}(Y_{1ij}) = \frac{\alpha_1^{\frac{2}{\rho_1}} e^{\frac{d_1^2}{\rho_1^2}} \left[B\left(\alpha_1 - \frac{2}{\rho_1}, \frac{2}{\rho_1}\right) e^{\frac{d_1^2}{\rho_1^2}} - B\left(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1}\right)^2 \right]}{\rho_1 \left(\lambda_1 e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1}\right)^{\frac{1}{\rho_1}}}, \tag{11}$$

and

$$T3 = \mathsf{Var}(Y_{2ij}) = \frac{\alpha_2^{\frac{2}{\rho_2}} e^{\frac{d_2^2}{\rho_2^2}} \left[B\left(\alpha_2 - \frac{2}{\rho_2}, \frac{2}{\rho_2}\right) e^{\frac{d_2^2}{\rho_2^2}} - B\left(\alpha_2 - \frac{1}{\rho_2}, \frac{1}{\rho_2}\right)^2 \right]}{\rho_2 \left(\lambda_2 e^{\boldsymbol{x}'_{2ij}\boldsymbol{\xi}_2}\right)^{\frac{1}{\rho_2}}}. \tag{12}$$

Substituting $T1$, $T2$ and $T3$ in (3) gives a formula for the correlation between the two endpoints in (1).

Derivation of the Intraclass Correlation

The correlation between the jth and k^{th} measurements of subject i for outcome 1, Y_{1ij} and Y_{1ik} is

$$\mathsf{Corr}(Y_{1ij}, Y_{1ik}) = \frac{\mathsf{Cov}(Y_{1ij}, Y_{1ik})}{\sqrt{\mathsf{Var}(Y_{1ij})}\sqrt{\mathsf{Var}(Y_{1ik})}} \tag{13}$$

with

$$\mathsf{Cov}(Y_{ij}, Y_{ik}) = \mathsf{E}(\mathsf{Cov}(Y_{1ij}, Y_{1ik}|b_{1i})) + \mathsf{Cov}\mathsf{E}(Y_{1ij}|b_{1i}), \mathsf{E}(Y_{1ik}|b_{1i}) \tag{14}$$

Given that the random effect between the measurements are independent, $\mathsf{E}(\mathsf{Cov}(Y_{1ij}, Y_{1ik}|b_{1i})) = 0$ and

$$\mathsf{Cov}(Y_{ij}, Y_{ik}) = \frac{\alpha_1^{\frac{2}{\rho_1}} B(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1})^2}{\rho_1^2 \left(\lambda_1^{\frac{2}{\rho_1}} e^{\boldsymbol{x}'_{1ij}\boldsymbol{\xi}_1 + \boldsymbol{x}'_{1ik}\boldsymbol{\xi}_1}\right)^{\frac{1}{\rho_1}}} e^{\frac{d_1^2}{\rho_1^2}} \left[e^{\frac{d_1^2}{\rho_1^2}} - 1 \right] \tag{15}$$

$\mathsf{Var}(Y_{1ij})$ is just the variance given in (12). Solving this leads to the intraclass correlation

$$\mathsf{Corr}(Y_{1ij}, Y_{1ik}) = \frac{\left[e^{\frac{d}{\rho_1^2}} - 1 \right]}{\left[\frac{2\rho_1 B(\alpha_1 - \frac{2}{\rho_1}, \frac{2}{\rho_1}) e^{\frac{d}{\rho_1^2}}}{B(\alpha_1 - \frac{1}{\rho_1}, \frac{1}{\rho_1})^2} - 1 \right]} \tag{16}$$

C   Simulation Results

Table 6: *Type I error rate, power of the test and parameter estimate. Correlation between $r = 0.9$ and $\alpha = 0.8$. W: Weibull; G: gamma; N: normal; J: joint.*

	Response	Parameter	Traditional	W	WN	WGN	JWGN
Power	1		0.445	0.595	0.255	0.505	0.4305
	2		0.565	0.915	0.675	0.675	0.6619
	Combined						0.6056
Estimate(s.e.)	1	β_{10}		-0.9616(0.0574)	-0.9959(0.0897)	-1.0016(0.1126)	-1.011(0.1185)
		β_{11}		-1.1117(0.0596)	-1.1486(0.0913)	-1.3082(0.1131)	-1.302(0.1171)
		$\beta_{11} - \beta_{10}$		-0.1501(0.0671)	-0.1527(0.1166)	-0.3066(0.1525)	-0.2997(0.1600)
	2	β_{20}		-0.9902(0.0597)	-1.0029(0.0861)	-1.0029(0.0861)	-1.0105(0.0883)
		β_{21}		-1.2838(0.0647)	-1.3016(0.0899)	-1.3016(0.0899)	-1.2990(0.0905)
		$\beta_{21} - \beta_{20}$		-0.2936(0.0672)	-0.2988(0.1099)	-0.2988(0.1099)	-0.2860(0.1125)
Type I	1		0.08	0.33	0.08	0.085	0.0909
	2		0.06	0.245	0.08	0.08	0.0666
	Combined						0.0512
Estimate(s.e.)	1	β_{10}		-0.9593(0.0574)	-0.9936(0.0896)	-0.997(0.1117)	-1.0019(0.1095)
		β_{11}		-0.977(0.0579)	-1.0101(0.0899)	-1.005(0.1116)	-1.0064(0.1083)
		$\beta_{11} - \beta_{10}$		-0.0180(0.0670)	-0.0164(0.1166)	-0.0081(0.1501)	-0.0083(0.1555)
	2	β_{20}		-0.9867(0.05965)	-0.99948(0.0859)	-0.9995(0.0859)	-1.013(0.0853)
		β_{21}		-0.9889(0.05976)	-1.0011(0.08598)	-1.0011(0.08598)	-0.9902(0.0825)
		$\beta_{21} - \beta_{20}$		-0.00214(0.0667)	-0.00164(0.1094)	-0.00165(0.1094)	0.0139(0.1093)

Table 7: *Type I error rate, power of the test and parameter estimate. Correlation between $r = 0.6$ and $\alpha = 0.8$. W: Weibull; G: gamma; N: normal; J: joint.*

	Response	Parameter	Traditional	W	WN	WGN	JWGN
Power	1		0.465	0.58	0.25	0.48	0.424
	2		0.555(111/200)	0.915	0.69	0.69	0.633
	Combined						0.696
Estimate(s.e.)	1	β_{10}		-0.9595(0.0574)	-0.9940(0.0901)	-0.9978(0.1116)	-1.002(0.1136)
		β_{11}		-1.108(0.0596)	-1.1462(0.0917)	-1.3035(0.11203)	-1.2979(0.1132)
		$\beta_{11} - \beta_{10}$		-0.1485(0.0671)	-0.1522(0.1173)	-0.3057(0.1510)	-0.2945(0.1533)
	2	β_{20}		-0.9903(0.05969)	-1.0028(0.0835)	-1.0028(0.0835)	-1.011(0.08384)
		β_{21}		-1.2823(0.0647)	-1.2970(0.0874)	-1.2970(0.0874)	-1.2967(0.0885)
		$\beta_{21} - \beta_{20}$		-0.2920(0.06719)	-0.2942(0.1060)	-0.2942(0.1060)	-0.2811(0.1073)
Type I	1		0.04	0.295	0.05	0.07	0.0666
	2		0.045	0.265	0.07	0.07	0.0588
	Combined						0.0504
Estimate(s.e.)	1	β_{10}		-0.9594(0.0574)	-0.9933(0.0898)	-0.9950(0.1106)	-0.994(0.1122)
		β_{11}		-0.9775(0.0578)	-1.0114(0.0901)	-1.0075(0.1105)	-1.002(0.1110)
		$\beta_{11} - \beta_{10}$		-0.0180(0.0670)	-0.0181(0.1168)	-0.0125(0.1485)	-0.0073(0.1537)
	2	β_{20}		-0.9855(0.0596)	-0.9980(0.0856)	-0.9980(0.0856)	-0.998(0.0848)
		β_{21}		-0.9920(0.0598)	-1.0037(0.0858)	-1.0037(0.0858)	-0.999(0.0851)
		$\beta_{21} - \beta_{20}$		-0.0065(0.0667)	-0.0057(0.1090)	-0.0065(0.1089)	-0.0005(0.1080)

Table 8: *Type I error rate, power of the test and parameter estimate. Correlation between $r = 0.9$ and $\alpha = 1.5$. W: Weibull; G: gamma; N: normal; J: joint.*

	Response	Parameter	Traditional	W	WN	WGN	JWGN
Power	1		0.34	0.43	0.23	0.33	0.262
	2		0.44	0.725	0.61	0.61	0.651
	Combined						0.648
Estimate(s.e.)	1	β_{10}		-1.0258(0.0939)	-1.076(0.1185)	-1.0182(0.1367)	-1.0216(0.1296)
		β_{11}		-1.2015(0.0977)	-1.2649(0.1229)	-1.3044(0.1395)	-1.2946(0.1312)
		$\beta_{11} - \beta_{10}$		-0.1757(0.1061)	-0.1889(0.1436)	-0.2862(0.1774)	-0.2654(0.1794)
	2	β_{20}		-0.9892(0.0943)	-1.0027(0.1049)	-1.0027(0.1049)	-1.0049(0.1008)
		β_{21}		-1.2850(0.1026)	-1.3014(0.1127)	-1.3014(0.1127)	-1.3080(0.1065)
		$\beta_{21} - \beta_{20}$		-0.2958(0.1063)	-0.2987(0.1236)	-0.2987(0.1236)	-0.3120(0.1236)
Type I	1		0.06	0.17	0.045	0.04	0.0394
	2		0.06	0.125	0.07	0.07	0.054
	Combined						0.0259
Estimate(s.e.)	1	β_{10}		-1.0235(0.0939)	-1.0731(0.1184)	-1.0141(0.1367)	-1.022(0.1341)
		β_{11}		-1.0106(0.0933)	-1.0609(0.1181)	-1.0052(0.1359)	-1.0138(0.1333)
		$\beta_{11} - \beta_{10}$		0.0129(0.1059)	0.0121(0.1432)	0.0089(0.1757)	0.0047(0.1771)
	2	β_{20}		-0.9886(0.0943)	-1.0024(0.1054)	-1.0024(0.1054)	-1.0096(0.1052)
		β_{21}		-0.9883(0.0946)	-0.9997(0.1055)	-0.9997(0.1055)	-1.0146(0.1034)
		$\beta_{21} - \beta_{20}$		0.00039(0.1056)	0.00276(0.1238)	0.00276(0.1238)	-0.0035(0.1227)

Table 9: *Type I error rate, power of the test and parameter estimate. Correlation between $r = 0.6$ and $\alpha = 1.5$. W: Weibull; G: gamma; N: normal; J: joint.*

	Response	Parameter	Traditional	W	WN	WGN	JWGN
Power	1		0.305	0.435	0.23	0.335	0.276
	2		0.46	0.73	0.64	0.64	0.6
	Combined						0.612
Estimate(s.e.)	1	β_{10}		-1.0248(0.0939)	-1.0754(0.1186)	-1.0170(0.1366)	-1.0359(0.1368)
		β_{11}		-1.2044(0.0979)	-1.2667(0.1230)	-1.3052(0.1393)	-1.3176(0.1387)
		$\beta_{11} - \beta_{10}$		-0.1795(0.1061)	-0.1913(0.1437)	-0.2882(0.1773)	-0.2746(0.1829)
	2	β_{20}		-0.9983(0.0946)	-1.0099(0.10417)	-1.0099(0.10417)	-1.0024(0.1022)
		β_{21}		-1.2803(0.10238)	-1.2961(0.1118)	-1.2961(0.1118)	-1.2923(0.1111)
		$\beta_{21} - \beta_{20}$		-0.2820(0.1063)	-0.2862(0.1220)	-0.2862(0.1220)	-0.2857(0.1280)
Type I	1		0.04	0.165	0.035	0.035	0.041
	2		0.06	0.095	0.07	0.07	0.062
	Combined						0.0396
Estimate(s.e.)	1	β_{20}		-1.0241(0.0938)	-1.0726(0.1181)	-1.0133(0.1368)	-1.0194(0.1282)
		β_{11}		-1.008(0.0933)	-1.0572(0.1177)	-0.9988(0.1361)	-0.9893(0.1266)
		$\beta_{11} - \beta_{10}$		0.0161(0.1059)	0.0154(0.1427)	0.0144(0.1759)	0.0307(0.1777)
	2	β_{20}		-0.9835(0.0942)	-0.9961(0.1045)	-0.9961(0.1045)	-1.0072(0.1031)
		β_{21}		-0.9823(0.0945)	-0.9929(0.1046)	-0.9929(0.1046)	-1.0011(0.0989)
		$\beta_{21} - \beta_{20}$		0.0012(0.1056)	0.0032(0.1225)	0.0032(0.1225)	0.0115(0.1223)

Table 10: *Type I error rate, power of the test and parameter estimate. Correlation between $r = 0.3$ and $\alpha = 1.5$. W: Weibull; G: gamma; N: normal; J: joint.*

	Response	Parameter	Traditional	W	WN	WGN	JWGN
Power	1		0.305	0.405	0.24	0.305	0.290
	2		0.445	0.75	0.65	0.65	0.667
	Combined						0.642
Estimate(s.e.)	1	β_{20}		-1.0314(0.0941)	-1.0789(0.1177)	-1.0223(0.1360)	-1.0124(0.1340)
		β_{11}		-1.2070(0.0979)	-1.2675(0.122)	-1.3077(0.1386)	-1.3005(0.1373)
		$\beta_{11} - \beta_{10}$		-0.1756(0.1061)	-0.1886(0.1422)	-0.2853(0.1761)	-0.2818(0.1775)
	2	β_{20}		-0.9955(0.0945)	-1.008(0.1048)	-1.008(0.1048)	-1.0132(0.1041)
		β_{21}		-1.2866(0.1026)	-1.303(0.1126)	-1.303(0.1126)	-1.3084(0.1095)
		$\beta_{21} - \beta_{20}$		-0.2911(0.1063)	-0.2954(0.1232)	-0.2954(0.1232)	-0.3063(0.1215)
Type I	1		0.03	0.16	0.04	0.03	0.055
	2		0.07	0.095	0.055	0.055	0.077
	Combined						0.070
Estimate(s.e.)	1	β_{20}		-1.0257(0.0939)	-1.0758(0.1183)	-1.0178(0.13647)	-1.0296(0.1342)
		β_{11}		-1.0107(0.09337)	-1.0611(0.1179)	-1.0040(0.13568)	-1.0127(0.1321)
		$\beta_{11} - \beta_{10}$		0.0150(0.1059)	0.1179(0.1429)	0.0138(0.1754)	0.0169(0.1719)
	2	β_{20}		-0.9983(0.0946)	-1.0102(0.1046)	-1.0102(0.1046)	-1.0152(0.1031)
		β_{21}		-0.9888(0.0946)	-0.999(0.1045)	-0.999(0.1045)	-1.0107(0.1031)
		$\beta_{21} - \beta_{20}$		0.0095(0.1056)	0.0103(0.1221)	0.0103(0.1221)	0.0020(0.1220)