RT-MAE 2010-08

A FULLY BAYESIAN PARAMETRIC APPROACH FOR CYTOGENETIC DOSIMETRY

by

Carlos Daniel Paulino, Giovani L. Silva
and
Márcia D'Elia Branco

Palavras-Chave: Calibration, categorical data, continuation-ratio logits, Bayesian methods, nonlinear structural model.

Classificação AMS: 62F15, 62J12, 62P10 .

- Novembro de 2010 -

A fully Bayesian parametric approach for cytogenetic dosimetry

Carlos Daniel Paulino[1,‡], Giovani L. Silva[1] and Márcia D. Branco [2]

[1]Department of Mathematics, Universidade Técnica de Lisboa and CEAUL, Portugal.

[2]Department of Statistics, Universidade de São Paulo, Brazil.

[‡]E-mail dpaulino@math.ist.utl.pt

Abstract

This paper describes a new statistical analysis strategy to problems of cytogenetic dosimetry involving ordinal polythomous responses. Models relating the multivariate response to dose take the data ordinality into account and are analysed in a fully Bayesian fashion in the application here considered. In particular, these models are compared in order to select the best one for purposes of drawing inferences of interest and dose prediction is naturally addressed by its practical importance. This work was motivated by an *in vitro* experimental study on radiation exposure of human blood cell cultures, previously analysed in the literature by other methods, but its interest holds in many other applications of the biological and environmental field involving data sets yielded from the same type of assays for genetic damage.

1 Introduction

Cytogenetic dosimetry is a field of the dose-response studies dealing with the relationship between the level of exposure to radiation and some measure of genetic aberration, wherein a special interest is devoted to the calibration problem towards drawing inferences on unknown exposure doses for given observed responses. Bender *et al.* (1988) provides a comprehensive discussion of this topic and a general review of the statistical calibration problem can be found in Osborn (1991).

In this paper we confine ourselves to *in vitro* studies in which human blood samples are exposed to a range of doses of a given agent, and a polytomous response related to genetic

aberrations is recorded for each dose. Specifically we take the experimental study of radiosensitivity described in Ochi-Lohnmann *et al.* (1996) and Madruga *et al.* (1996) as an illustration of alternative procedures we propose to analyse data sets involving ordinal categorical responses in the framework of cytogenetic dosimetry problems. These problems are relevant in applications namely related to ecotoxicological studies and biomonitoring of human populations such as referred to in Fenech (2000).

This experimental study involved lymphocytes cultures obtained from a few individuals belonging to three groups to be compared in terms of chromosomal susceptibility to ionizing radiation. One of them consisted of untreated cancer patients with basocellular carcinoma and the remaining ones are two control groups consisting of healthy subjects differing each other in terms of age. Blood samples of each individual were irradiated in a ^{60}Co source with doses of 0, 20, 50, 100, 200, 300, 400 and 500 cGy (at an average dose rate of 1 Gy/min).

Some of these samples were subject to a cytokinesis-block micronucleus assay with cytochalasin-B originating cells that completed only one nuclear division (binucleated cells). The data here considered refer to frequencies of cells displaying zero, one and two or more micronuclei out of the total of cells for every individual exposed to each radiation dose. These micronuclei, resulting from chromosome break or loss or whole chromosome that fail to incorporate the main nuclei during the mitosis process, express the DNA damage induced spontaneously or by radiation. For details on the cytokinesis-block micronucleus technique see, e.g., Fenech (2000). The same type of data was also obtained for (mononucleated) cells that did not undergo the aforementioned cellular division process. Since all observed counts are summarized in Madruga *et al.* (1996), we refer interested readers to there for the sake of space.

Madruga *et al.* (1996) based their analysis upon a Dirichlet posterior distribution for the original parameters of a multinomial model for the cell frequency vector that corresponds to a log-Dirichlet distribution of second kind for the ordinary (baseline) logits. The latter is then appproximated by a bivariate Normal distribution after Aitchison and Shen (1980). The nonlinear preditor relating these logits to dose levels they consider with no further comparative analysis is fitted by classical methods. Kottas *et al.* (2002), taking just a subset of the same data, use a linear relationship between the ordinary logits and log-dose and perform a Bayesian analysis based on a noninformative prior for the parameters of this linear preditor. The ensuing results are compared with those associated with a fully nonparametric analysis on cumulative probabilities based on Dirichlet processes.

This paper aims at developing a fully Bayesian analysis of the whole dataset, with no concession to a hybrid approach and without falling into the theoretical and computational

complexities of a nonparametric approach. Calibration models on parametric functions that allow for the data ordinal nature are considered and compared towards selecting the best one for purposes of drawing inferences of interest. Further analytic examination of selected models is taken in order to compare the three groups, what enables to get a more parsimonious overall model. The selected calibration model for each group is used to predict dose levels for given further observed frequencies.

The layout of the paper is as follows. Section 2 introduces the statistical modelling of this cytogenetic dosimetry problem. In Section 3 a Bayesian analysis of the illustrative dataset is described, with presentation and discussion of the results obtained. Section 4 is devoted to a summary and further comments on the approach followed in this work.

2 Statistical modelling

Focusing on the dataset previously described, for each radiation dose d_i, $i = 1, 2, \ldots, k$ ($k = 8$), which n_i cells were exposed to, the observable response vector is denoted by $Y_i = (Y_{i2}, Y_{i1}, Y_{i0})$, where Y_{i2} is the number of cells with two or more micronuclei, Y_{i1} is the number of cells with one micronucleus and Y_{i0} is the number of cells with no micronucleus. The probabilistic model considered is a product-multinomial family, with probability function

$$f(y_1, y_2, \ldots, y_k | \{n_i, \pi_i\}) = \prod_{i=1}^{k} n_i! \prod_{j=0}^{2} \frac{\theta_{ij}^{y_{ij}}}{y_{ij}!}, \tag{1}$$

where $\sum_{j=0}^{2} y_{ij} = n_i$ and $\sum_{j=0}^{2} \theta_{ij} \equiv 1'\pi_i = 1$, with the probabilities of each category stacked in the vector π_i depending somehow on the radiation dose.

The i-th trinomial probability function, taking response category ordering into account, can be factored into a product of two binomial distributions, the marginal distribution for Y_{i2} and the conditional distribution for Y_{i1} given $n_i - Y_{i2}$,

$$f(y_{i2}, y_{i1} | n_i, \pi_i) = f(y_{i2} | n_i, \theta_{i2}) f(y_{i1} | n_i - y_{i2}, \theta_{i1}/(\theta_{i1} + \theta_{i0})). \tag{2}$$

Reparametrizing these two binomial distributions to the corresponding ordinary logits one gets the so-called continuation-ratio logits

$$L_{i1} \equiv \ln\left(\frac{\theta_{i2}}{\theta_{i1} + \theta_{i0}}\right), \qquad L_{i2} \equiv \ln\left(\frac{\theta_{i1}}{\theta_{i0}}\right), \tag{3}$$

that contrast each category with a grouping of categories from lower levels of the response ordinal scale. The formulae (1)-(3) extend to the case of more than 3 response categories (e.g., Agresti, 2002).

One may contemplate other link functions for the (theoretical) proportions of the two binomial components of the probability model. For instance, if an asymmetric link such as the complementary log-log (or extremit) function was to be considered, one would get transformations of the proportions of cells with fewer than two micronuclei and the proportions of cells with none micronucleus within those with fewer than two micronuclei,

$$E_{i1} = \ln\{-\ln(\theta_{i1} + \theta_{i0})\}, \qquad E_{i2} = \ln\left\{-\ln\left(\frac{\theta_{i0}}{\theta_{i1} + \theta_{i0}}\right)\right\}. \tag{4}$$

Dependence of π_i on the radiation dose is expressed through modelling the continuation-ratio logits (or the corresponding alternative link functions). The structural models here considered were chosen taking account of the empirical calibration curves and comparative purposes towards the selection of the "best" model in order to draw the inferences of interest. These models include a simple linear, quadratic or non-linear structure on $\{L_{ij} \equiv L_j(d_i)\}$, $j = 1, 2$; $i = 1, \ldots, k$, where d_i denotes the i-th dose,

$$L_j(d_i) = \alpha_j + \beta_j d_i \tag{5}$$

$$L_j(d_i) = \alpha_j + \beta_j d_i + \gamma_j d_i^2 \tag{6}$$

$$L_j(d_i) = \gamma_j + \frac{\alpha_j}{\beta_j + d_i}. \tag{7}$$

The preditor functional structure of these models has often been used in the dose-response problem literature, even though applied to other probability models or parametric functions (see Madruga *et al.*, 1994).

Denoting the parameter vectors of each structural model by δ_j, $j = 1, 2$, the statistical model for the observed data gets expressed by

$$f(\{y_i\}|\{n_i, d_i\}, \{\delta_j\}) = \prod_{i=1}^{k}\left\{\binom{n_i}{y_{i2}} \frac{e^{y_{i2}L_1(d_i;\delta_1)}}{(1 + e^{L_1(d_i;\delta_1)})^{n_i}} \binom{n_i - y_{i2}}{y_{i1}} \frac{e^{y_{i1}L_2(d_i;\delta_2)}}{(1 + e^{L_2(d_i;\delta_2)})^{n_i - y_{i2}}}\right\} \tag{8}$$

Due to absence of specific prior information on any model parameters, we adopted independent Normal distributions for each component of δ_j, $j = 1, 2$, centered on 0 and with a large variance (equal to 10^6). The analysis of this Bayesian model allows us to compare the diverse dose-response structures and draw parametric inferences of interest, as described in the following section.

When it is intended to predict an unknown dose which an individual with known response vector was exposed to, the sampling model (8) that was selected previously is augmented with the distributional factor corresponding to this further data. Denoting the additional response vector by $Y_0 = (Y_{0j}, j = 0, 1, 2)$, with $\sum_{j=0}^{2} Y_{0j} = n_0$, and its unknown dose by

d_0, the statistical model to be considered is

$$f(\{y_i\}, y_0|\{n_i, d_i\},, n_0, d_0, \{\delta_j\}) = \prod_{i=1}^{k}\{f(y_{i2}|n_i, d_i; \delta_1)f(y_{i1}|n_i - y_{i2}, d_i; \delta_2)\} \times f(y_{02}|n_0, d_0; \delta_1)f(y_{01}|n_0 - y_{02}, d_0; \delta_2). \quad (9)$$

The associated prior includes a further factor concerning the prior distribution assigned to the parameter of interest d_0. Here we used a flat Normal distribution centered on the average of the observed doses. It must be truncated from negative values when deemed necessary.

3 Bayesian analysis

The complexity of the Bayesian models previously described demands resorting to Markov Chain Monte Carlo (MCMC) methods so as to obtain the posterior density for the respective parameters by simulation. For each model considered, the convergence and autocorrelation analysis by the usual methods (Gilks *et al.*, 1996) of the simulated chain allowed us to retain a MCMC sample of size 10,000 by taking every 5th iteration of the sequence, after removing 5,000 burn-in iterations. For reasons that have to do with the chain convergence, the analysis of the linear and quadratic models started with a previous standardization of the dose levels and with appropriate flat Normal priors centered on 0 for the associated parameters. The MCMC analysis was implemented in WinBugs (Spiegelhalter *et al.*, 2007). Convergence diagnostics and determination of highest posterior density (HPD) credible intervals were carried out via BOA software (Smith, 2007).

Comparison of models was carried out by assessment of their goodness-of-fit and complexity through some measures as follows: posterior mean of the Pearson parametric function (obtained via appropriate transformation from the simulated values for $\delta_j, j = 1, 2$), deviance information criterion (DIC) (Spiegelhalter *et al.*, 2002), and Carlin-Louis' version of Bayesian information criterion (BIC) (Carlin and Louis, 2000). The more refined approximation of BIC due to Raftery *et al.* (2007) was also considered but its results (not shown) did not cause any change in the model ordering.

Table 1 displays the results obtained for the logistic models (5)-(7). The non-linear I model is a two-parameter version of the non-linear model (7) (now labelled II) by taking $\gamma_1 = \gamma_2 = 0$. The model which fits "best", as defined by each measure, is underlined in Table 1. The emphasis here placed on logit-based models is due to the fact that they showed a better behaviour than the corresponding models based upon the complementary log-log function (the respective results are omitted for the sake of space).

Table 1: Comparison of continuation-ratio logit models based on posterior mean of Pearson function (PF), BIC_{CL} and DIC.

		Mononucleated cells			Binucleated cells		
Group	Model	PF	BIC_{CL}	DIC	PF	BIC_{CL}	DIC
Basocellular	Linear	386	554	511	184	344	311
carcinoma	Quadratic	193	357	293	135	288	238
	Nonlinear I	195	329	286	114	241	207
	Nonlinear II	183	341	275	99	244	192
Healthy young	Linear	1101	1396	1352	284	463	428
	Quadratic	264	481	414	135	298	245
	Nonlinear I	250	429	384	82	221	186
	Nonlinear II	85	264	196	83	242	188
Healthy older	Linear	314	452	411	131	264	232
	Quadratic	61	233	171	37	185	137
	Nonlinear I	111	258	217	25	154	122
	Nonlinear II	72	242	178	16	162	113

According to the criteria used, the best model depends on the group of subjects and type of cells. The non-linear II model may be taken as our choice for the carcinoma group, regardless of the cell type, as well as for mononucleated cells of the young healthy group and binucleated cells of the older healthy group. Note that in some cases BIC tends to penalize it more than DIC does in favour of the simpler non-linear model. For binucleated cells of the young healthy group the non-linear II model appears to be a little worse than its simpler counterpart whereas the quadratic model presents the best performance for mononucleated cells related to the older healthy group.

Figure 1 portrays calibration curves for the three groups, drawn from the posterior means of δ_j, $j = 1, 2$, parameters wherein the symbol o represents the values of the empirical continuation-ratio logits, for the binucleated cells (the others are not shown for reasons

of space saving). They show the fitting superiority in general of the non-linear models over the quadratic one. The exception occurs for mononucleated cells of the older healthy group. (figures left out for the above reasons).

The dose-response curves computed from the posterior means of $\{\theta_{ij}\}$, denoted by $\{\bar{\theta}_{ij}\}$, for the selected model for each group (according to the criteria pointed above) are displayed in Figure 2. As expected, the estimated proportions of damaged (unaffected) cells tend to increase (decrease) in general with the dose levels. A noticeable exception is the case of the older healthy group for mononucleated cells, as consequence of using the quadratic model. From a given high dose level the decrease of $\{\bar{\theta}_{i0}\}$ is reversed, in correspondence with an opposite monotony behaviour of $\{\bar{\theta}_{i1}\}$, as well as of $\{\bar{\theta}_{i2}\}$ though this latter feature is not captured over the dosage range of Figure 2. This unsatisfactory behaviour disappears if for the case at issue we adopt the non-linear II model (the second best one in the class considered) in that the curves will follow a predictable and very similar pattern to the other two groups (results not shown for the sake of space). This illustrates how a very good model in the light of observed data may lead to unwise extrapolations.

The curves concerning binucleated cells already show noticeable differences among the three groups over the range of high doses. The most pronounced decrease (increase) of $\{\bar{\theta}_{i0}\}$ ($\{\bar{\theta}_{i2}\}$) belongs to healthy young subjects. As opposed to Madruga *et al.* (1996), the curves for both $\{\bar{\theta}_{i0}\}$ and $\{\bar{\theta}_{i2}\}$ concerning both groups of healthy subjects are distinct. The sharpest descent of $\{\bar{\theta}_{i0}\}$ belongs to healthy young group, which displays a similar ascent for $\{\bar{\theta}_{i2}\}$ as that for the carcinoma patients up to the highest observed dose. From here this latter group shows a more marked increase in agreement with findings of other studies (see, e.g., Fenech,2002) wherein the same type of micronucleus assays has shown that individuals who develop some types of cancer and their relatives exhibit elevated sensitivity to the DNA-damaging effects of ionising radiation.

Table 2 displays parameter estimates regarding the model chosen by the criteria of Table 1 for each group $\times$ cell type setting. We may add that fitting the non-linear II model for binucleated cells of the young healthy group, the 95% HPD credible intervals for the parameters $\gamma_1([-0.405, 1.052])$ and $\gamma_2([-0.275, 0.337])$ suggest that the nested model without these parameters can be a better alternative in accordance with the model comparison results in Table 1.

The non-linear II model can still be reduced in the light of the data by elimination of just one parameter γ_j for the group of carcinoma patients irrespective of the cell type ($\gamma_1 = 0$) and for binucleated cells of the older healthy subjects ($\gamma_2 = 0$). The evidence of γ_2, actually related to an ordinary logit (recall (3)) being statistically significant for

mononucleated cells of the carcinoma and young healthy subjects points out against the choice of the non-linear I model by Madruga *et al.* (1996), reinforcing the comparative outcomes in Table 1.

Comparisons among the three groups can be made by integrating the corresponding product-trinomial distributions when parameterized by the same kind of structural model. For instance, for binucleated cells there is evidence that carcinoma and young healthy groups share the same parameters γ_1 and α_1. An analogous conclusion regarding γ_2, α_2 and β_2 holds within the two healthy groups. See Table 3.

Once selected a dose-response model in the light of the experimental calibration data, we can use it towards estimation of unknown doses which blood cells of further individuals had been exposed to. With illustrative purposes we consider two new individuals by each group whose responses concerning binucleated cells are displayed in Table 4. The results were obtained by using the non-linear model that was selected previously for each group.

The dose estimation is relatively precise *a posteriori* for the first three subjects whose data suggest their blood cells would have been exposed to moderate doses. There is evidence that the remaining ones would have had their cells submitted to higher doses, possibly beyond the larger dose used in the calibration experiment, what accounts for the fact that their HPD credible intervals tend to be substantially wider. This was obviously expected on the grounds that dose prediction for the latter cases correspond to an extrapolation. The higher are the doses implied by the observed proportions, the larger is the predictive variability and wider are the credible intervals for the predicted dose. This is what one would obtain had we taken the more extreme cases exemplified in Madruga *et al.* (1996).

4 Concluding Remarks

This paper offers a new modelling approach to problems of cytogenetic dosimetry involving ordinal polythomous responses based on an appealing factorization of the product-multinomial probability function into binomial factors related to appropriate ratios of category probabilities. This allows us to contemplate several types of models for functions of these conditional probabilities, such as logits, probits and extremits, that take the data ordinal nature into account. In the application here revisited, involving the cytokinesis-block micronucleus assay, non-linear parametric models in continuation-ratio logits have played an important role, namely in assessing their fit and reduction and performing inverse prediction.

Of course distinct models might be entertained. In a preliminary study, specification of a

cubic spline structure with a knot succeeded for some settings in terms of aforementioned model comparison criteria. Specifically, these new models behaved better than the non-linear parametric models for binucleated cells of carcinoma and young healthy groups. Also, other asymmetric functions of those conditional probabilities can still be addressed with the purposes of fit comparison and possible simplification of the preditor form.

The analysis of the model followed a fully Bayesian route based on usual noninformative priors for the model preditor parameters, on the grounds that prior information was unavailable. It enabled us to make additional inferences and get some distinguishing outcomes from those concerning a hybrid analysis of the illustrative data set previously performed by Madruga *et al.* (1996). We believe that in such dosimetry problems there may be experts with prior beliefs on the original category probabilities, which may be elicited and accommodated in some convenient prior distribution (*e.g.*, Dirichlet). In such cases, it may be possible to convert this to the corresponding prior for the preditor parameters through Bedrick *et al.* (1996) approach, following a procedure analogous to that used by Paulino *et al.* (2003) for binary data.

The kind of analysis performed and the aforementioned suggestions are useful for many other applications of the cytokinesis-block micronucleus technique yielding data of a similar nature, among which we emphasize radiation sensitivity testing both for cancer risk assessment and optimisation of radiotherapy, testing of new pharmaceuticals and agrichemicals, problems in ecotoxicology and nutrition and biomonitoring of human populations (see Fenech, 2000 and references therein).

References

[1] Agresti, A. (2002). *Categorical Data Analysis*, 2nd. Wiley-Interscience, Hoboken, NJ.

[2] Aitchison, J. and Shen, S.M. (1980). Logistic-normal distributions: some properties and uses. *Biometrika* **67**, 261–272.

[3] Bedrick, E.J., Christensen, R. and Johnson, W. (1996). A new perspective on priors for generalized linear models. *Journal of the American Statistical Association* **91**, 1450–1460.

[4] Bender, M.A, Awa, A.A., Brooks, A.L., Evans, H.J., Groer, P.G., Littlefield, L.G., Pereira, C.A.B., Preston, F.J. and Wachholz, B.W. (1988). Current status of cytogenetic procedures to detect and quantify previous exposures to radiation. *Mutation Research* **196**, 103–159.

[5] Carlin, B.P. and Louis, T.A. (2000). *Bayes and Empirical Bayes Methods for Data*

Analysis, 2nd ed., Chapman and Hall/CRC Press, Boca Raton, FL.

[6] Fenech, M. (2000). The *in vitro* micronucleus technique. *Mutation Research* **455**, 81–95.

[7] Fenech, M. (2002). Chromosomal biomarkers of genomic instability relevant to cancer. *Drug Discovery Today* **7**(22), 1128–1137.

[8] Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.

[9] Kottas, A., Branco, M.D. and Gelfand, A.E. (2002). A nonparametric Bayesian modelling approach for cytogenetic dosimetry. *Biometrics* **58**, 593–600.

[10] Lunn, D.J., Thomas, A., Best, N.G. and Spiegelhalter, D.J. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337.

[11] Madruga, M.R., Pereira, C.A. and Rabello-Gay, M.N. (1994). Bayesian dosimetry: radiation dose versus frequencies of cells with aberrations. *Environmetrics* **5**, 47–56.

[12] Madruga, M.R., Ochi-Lohmann, T.H., Okazaki, K., Pereira, C.A. and Rabello-Gay, M.N. (1996). Bayesian dosimetry II: credibility intervals for radiation dose. *Environmetrics* **7**, 325–331.

[13] Ochi-Lohmann, T.H., Okazaki, K., Madruga, M.R., Pereira, C.A. and Rabello-Gay, M.N. (1996). Radiosensitivity of blood lymphocytes from basocellular carcinoma patients, as detected by micronucleus assay. *Mutation Research* **357**, 97–106.

[14] Osborn, C. (1991). Statistical calibration: a review. *International Statistical Review* **59**, 309–336.

[15] Paulino, C.D., Soares, P. and Neuhaus, J. (2003). Binomial regression with misclassification. *Biometrics* **59**, 670–675.

[16] Raftery, A., Newton, M., Satagopan, J. and Krivitsky, P. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In *Bayesian Statistics* 8 (J. Bernardo et al., eds.). Oxford University Press, Oxford, pp. 371-416. (with discussion)

[17] Silva, G.L., Dean, C.B., Niyonsenga, T. and Vanasse, A. (2008). Hierarchical Bayesian spatiotemporal analysis of revascularization odds using smoothing splines. *Statistics in Medicine* **27**, 2381–2401.

[18] Smith, B.J. (2007). *Bayesian Output Analysis Program (BOA), version 1.1.6*. The University of Iowa. Available from http://www.public-health.uiowa.edu/boa.

[19] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society B* **64**, 583–616.

Table 2: Posterior estimates for selected model parameters.

		Mononucleated cells			Binucleated cells		
Group & Parameter		Mean	S.D.	95% HPD CI	Mean	S.D.	95% HPD CI
		(nonlinear II model)			(nonlinear II model)		
Basal	α_1	-3437	568.4	(-4554, -2358)	-2240	532.7	(-3315, -1276)
cellular	α_2	-1081	143.6	(-1370, -819.4)	-285.8	67.44	(-421.7, -173.4)
carcinoma	β_1	436.8	52.27	(333.2, 536.8)	365.9	66.77	(238.8, 498.2)
	β_2	218.7	23.86	(174.4, 266.9)	80.74	19.91	(45.68, 119.8)
	γ_1	0.281	0.438	(-0.537, 1.173)	0.636	0.447	(-0.212, 1.528)
	γ_2	-0.698	0.177	(-1.032, -0.347)	-0.600	0.135	(-0.849, -0.33?
		(nonlinear II model)			(nonlinear I model)		
Healthy	α_1	-2066	368.6	(-2814, -1406)	-1220	45.02	(-1313, -1137)
young	α_2	-501.9	33.17	(-567.8, -438.3)	-567.3	19.52	(-605.3, -529.2)
	β_1	290.6	41.54	(211.4, 371.9)	197.1	13.6	(170.5, 223.5)
	β_2	88.14	6.323	(76.01, 100.6)	133.7	7.261	(119.2, 147.6)
	γ_1	-0.862	0.357	(-1.532, -0.144)	-	-	-
	γ_2	-1.398	0.068	(-1.529, -1.263)	-	-	-
		(quadratic model)			(nonlinear II model)		
Healthy	α_1	-7.036	0.144	(-7.326, -6.761)	-515.1	192.5	(-900.1, -229.6)
older	α_2	-4.988	0.054	(-5.095, -4.884)	-587.9	149.2	(-885.5, -335.6)
	β_1	0.011	0.0012	(0.009, 0.014)	118.7	44.24	(48.31, 207.7)
	β_2	0.012	0.0005	(0.011, 0.013)	166.1	36.43	(102.7, 240.1)
	γ_1	-9×10^{-6}	2×10^{-6}	$(-1\times10^{-5}, -5\times10^{-6})$	-1.267	0.282	(-1.764, -0.698)
	γ_2	-1×10^{-5}	9×10^{-7}	$(-2\times10^{-5}, -1\times10^{-5})$	0.018	0.204	(-0.368, 0.420)

Table 3: Posterior interval estimates for comparison among group parameters for binucleated cells concerning nonlinear II model.

Parameter	95% HPD CI	Parameter	95% HPD CI	Evidence
$\gamma_1^C - \gamma_1^{HY}$	(-0.476, 1.148)	$\gamma_2^C - \gamma_2^{HY}$	(-0.953, -0.299)	$\gamma_1^C = \gamma_1^{HY} \neq \gamma_1^{HO}$
$\gamma_1^C - \gamma_1^{HO}$	(0.998, 2.803)	$\gamma_2^C - \gamma_2^{HO}$	(-1.037, -0.204)	$\gamma_2^C \neq \gamma_2^{HY} = \gamma_2^{HO}$
$\gamma_1^{HY} - \gamma_1^{HO}$	(0.794, 2.333)	$\gamma_2^{HY} - \gamma_2^{HO}$	(-0.406, 0.412)	
$\alpha_1^C - \alpha_1^{HY}$	(-1662, 113.0)	$\alpha_2^C - \alpha_2^{HY}$	(117.9, 497.2)	$\alpha_1^C = \alpha_1^{HY} \neq \alpha_1^{HO}$
$\alpha_1^C - \alpha_1^{HO}$	(-2772, -790.6)	$\alpha_2^C - \alpha_2^{HO}$	(42.5, 595.8)	$\alpha_2^C \neq \alpha_2^{HY} = \alpha_2^{HO}$
$\alpha_1^{HY} - \alpha_1^{HO}$	(-1678, -346.3)	$\alpha_2^{HY} - \alpha_2^{HO}$	(-272.4, 291.1)	
$\beta_1^C - \beta_1^{HY}$	(19.20, 247.90)	$\beta_2^C - \beta_2^{HY}$	(-102.07, -11.00)	$\beta_1^C \neq \beta_1^{HY} \neq \beta_1^{HO}$
$\beta_1^C - \beta_1^{HO}$	(112.00, 383.61)	$\beta_2^C - \beta_2^{HO}$	(-157.90, -17.60)	$\beta_2^C \neq \beta_2^{HY} = \beta_2^{HO}$
$\beta_1^{HY} - \beta_1^{HO}$	(9.60, 221.56)	$\beta_2^{HY} - \beta_2^{HO}$	(-99.50, 35.00)	

Table 4: Dose posterior estimates for the calibration problem under the chosen non-linear models for binucleated cells of two subjects per group.

	Posterior dose estimates		
Observed responses	mean	S.D.	95% HPD CI
$y_{0C}^{(1)} = (76, 240, 1186)$	232.8	17.58	(198.7, 267.5)
$y_{0HY}^{(1)} = (176, 401, 1930)$	252.2	10.95	(230.3, 273.2)
$y_{0HO}^{(1)} = (72, 241, 890)$	255.7	21.44	(214.3, 298.2)
$y_{0C}^{(2)} = (270, 451, 1083)$	606.6	47.8	(520.7, 701.8)
$y_{0HY}^{(2)} = (362, 725, 1329)$	582.8	26.43	(532.1, 635.3)
$y_{0HO}^{(2)} = (160, 319, 660)$	727.4	138.3	(515.0, 996.5)

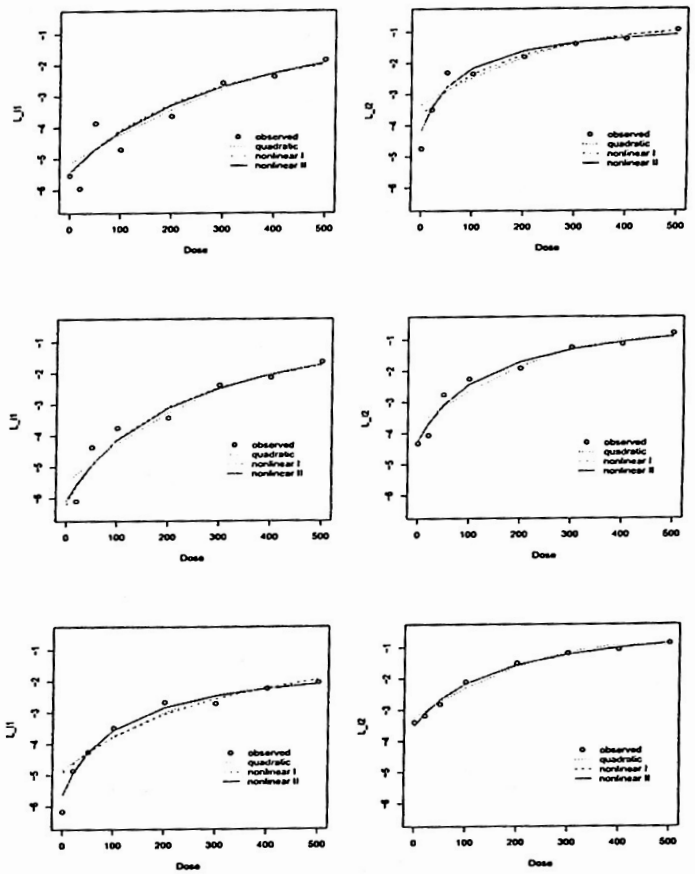

Figure 1: Observed and fitted continuation-ratio logits, L_{i1} (left) and L_{i2} (right), of binucleated cells for basocellular carcinoma (top), young healthy (middle) and older healthy (bottom) subjects.

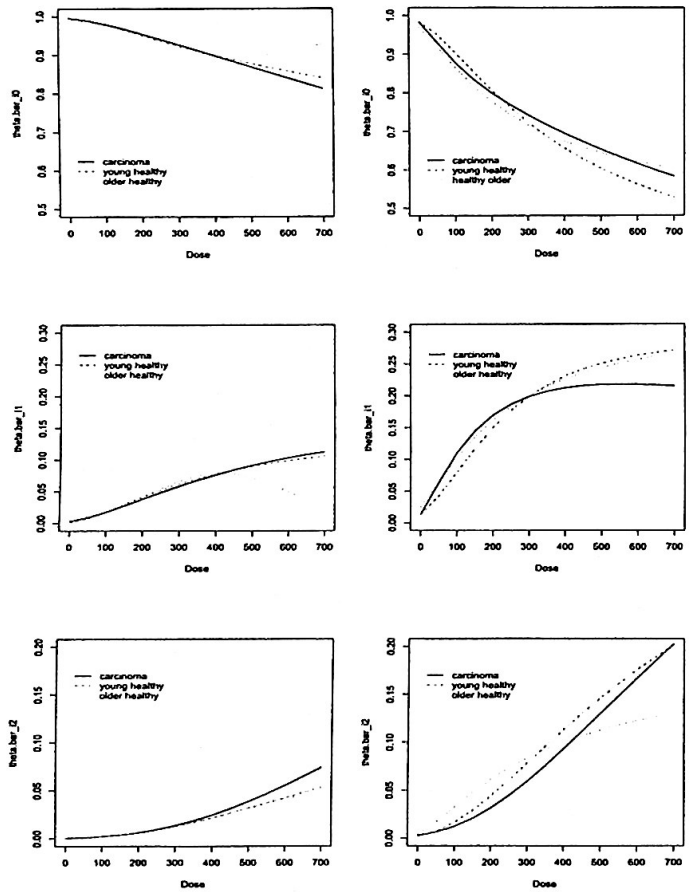

Figure 2: Posterior expected proportions of mono- (left) and binucleated (right) cells with none (top), one (middle) and two or more (bottom) micronuclei adjusted for each group, according to selected model (NL II: carcinoma + YH-mono + OH-bi; NL I: YH-bi; Quad: OH-mono).

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

2010-01 - BUENO, V.C. Component importance for a coherent system under a hyperminimal distribution. 09p. (RT-2010-01)

2010-02 - ABADI, M., ESTEVES, L.G., SIMONIS, A. Pointwise approximationos for sums of non-identically distributed Bernoulli trials. 12p. (RT-MAE-2010-02)

2010-03 - BUENO, V.C. A Series Representation of a Coherent System. 16p. (RT-MAE-2010-03)

***2010-04* - POLETO, F.Z., PAULINO, C.D., MOLENBERGHS, G. SINGER, J.M.** *Inferential implications of over-parameterization: a case study in incomplete categorical data. 33p. (RT-MAE-2010-04)*

2010-05 - TSUNEMI, M.H., ESTEVES, L.G., LEITE, J.G., WECHSLER, S. A Bayesian nonparametric model for Taguchi's on-line quality monitoring procedure for attributes. 23P. (RT-MAE-2010-05)

2010-06 - BUENO, V.C. Dynamic signatures of a coherent system. 18p. (RT-MAE-2010-06)

2010-07 - OIKAWA, K.F., ELIAN, S.N. Análise de Influência na Regressão em Cristas. 17p. (RT-MAE-2010-07)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05314-970 - São Paulo, Brasil