



Springer

Dear Author:

Please find attached the final pdf file of your contribution, which can be viewed using the Acrobat Reader, version 3.0 or higher. We would kindly like to draw your attention to the fact that copyright law is also valid for electronic products. This means especially that:

- You may not alter the pdf file, as changes to the published contribution are prohibited by copyright law.
- You may print the file and distribute it amongst your colleagues in the scientific community for scientific and/or personal use.
- You may make an article published by Springer-Verlag available on your personal home page provided the source of the published article is cited and Springer-Verlag is mentioned as copyright holder. You are requested to create a link to the published article in LINK, Springer's internet service. The link must be accompanied by the following text: The original publication is available on LINK **<http://link.springer.de>**. Please use the appropriate URL and/or DOI for the article in LINK. Articles disseminated via LINK are indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks and consortia.
- You are not allowed to make the pdf file accessible to the general public, e.g. your institute/your company is not allowed to place this file on its homepage.
- Please address any queries to the production editor of the journal in question, giving your name, the journal title, volume and first page number.

Yours sincerely,

Springer-Verlag Berlin Heidelberg

Environmental genotoxicity evaluation: Bayesian approach for a mixture statistical model

A. M. de S. Bueno, C. A. de B. Pereira, M. N. Rabello-Gay, J. M. Stern

Abstract. The data analyzed in this paper are part of the results described in Bueno et al. (2000). Three cytogenetics endpoints were analyzed in three populations of a species of wild rodent – *Akodon montensis* – living in an industrial, an agricultural, and a preservation area at the Itajaí Valley, State of Santa Catarina, Brazil. The polychromatic/normochromatic ratio, the mitotic index, and the frequency of micronucleated polychromatic erythrocytes were used in an attempt to establish a genotoxic profile of each area. It was assumed that the three populations were in the same conditions with respect to the influence of confounding factors such as animal age, health, nutrition status, presence of pathogens, and intra- and inter-population genetic variability. Therefore, any differences found in the endpoints analyzed could be attributed to the external agents present in each area. The statistical models used in this paper are mixtures of negative-binomials and Poisson variables. The Poisson variables are used as approximations of binomials for rare events. The mixing distributions are beta densities. The statistical analyzes are under the bayesian perspective, as opposed to the frequentist ones often considered in the literature, as for instance in Bueno et al. (2000).

Keywords: Cell proliferative indices, Micronucleated cells, Prior and posterior probabilities, Beta-(negative) Binomial distribution, Beta-Poisson distribution, Mixture of Beta distributions

1 Introduction

To evaluate the environmental genotoxicity in a certain area the investigator collects a certain number of living specimens to detect damage in their cells. The

A. M. de S. Bueno
Univ. Fed. Santa Catarina – CCB,
Florianópolis, SC, Brazil

C. A. de B. Pereira (✉)
Univ. São Paulo-IME-CEP05508-090, São Paulo, SP, Brazil
e-mail: cpereira@ime.usp.br

M. N. Rabello-Gay
Instituto Butantan, Laboratório de Genética, Brazil

J. M. Stern
Univ. São Paulo-IME-MAC,
São Paulo, SP, Brazil

data collected from each individual consists of a sample of cells where the frequency of cells with damage (failure) are recorded. Hence, there are two levels in our sample: the sample of individuals and the sample of cells from each individual. The investigator faces a dilemma: should a great number, n , of individuals be used to collect a small number, m , of cells or a smaller number of individuals with a larger number of cells? The ideal situation occurs when both samples are large. The first situation is easily obtained in the laboratory, whereas, when collecting data from exposed individuals in the field, the second situation prevails. The latter is the case analyzed in a real situation of environmental exposure reported in Bueno et al. (2000). In the present paper the data receive an alternative statistical treatment using mixtures of distributions and a Bayesian approach. In Bueno et al. (2000), the chance of a cell being damaged (probability of failure) is considered the same for each individual within a certain experimental group. Hence, independent Binomial (positive or negative) models with the same probability of failure were considered. It is not realistic, however, to consider that different individuals react in an identical way to the same external stimulus. Therefore, in this paper we assume that the failure rates (probabilities), π , are different for each individual within a group and are generated independently by the same beta distribution. Consequently, the number of damaged cells in each individual is a beta-(negative)binomial random variable. The beta-(negative) binomial is a mixture of (negative) binomials with the beta being the mixing distribution (Aitchison and Dunsmore 1975). For the case of rare events, we consider the approximation from binomial to Poisson distributions with parameter $m\pi$, where m is the number of individual sample of cells. Again, π is generated by a beta distribution. This mixture here is called beta-Poisson distribution.

The first challenge in modeling the problem is the choice of the prior distributions for the parameters of the mixing likelihood. We describe these choices in Sect. 4.

The final objective of our statistical analysis is to build the posterior “predictive” probability density of the failure rate for a new individual in each group. With these distributions, we will be able to evaluate the probability of a new individual from an specific group having a higher frequency of damaged cells than an individual from another group. These probabilities will allow us to analyze the genotoxicity differences among areas. In the following sections we describe the various aspects of the problem, both biological and statistical.

2

Background

The impact of toxic products introduced by humans in the environment results from complex interactions that can seldom be reproduced in laboratory. Such interactions, however, may be assessed when they occur, using plants and animals as sentinel surveillance systems. This *in situ* approach is a widely accepted method for identifying risks to ecosystems and human health (Sandhu and Serres 1989). The methodology was used by Bueno et al. (2000) to evaluate the impact of environmental compounds occurring in the Itajaí Valley, State of Santa Catarina, Southern Brazil. Lumber, textile, paper, cigarette, and pesticide industries, besides intense rice cultivation, are the region's main economic activities. Three areas were chosen for the study: a preservation area (PA) – to monitor animals not directly exposed to environmental pollution and for comparison with the other two areas; a rice field (RF) – to evaluate the genotoxic effects of pesticides used in

the rice culture; and an industrial area (IA) – to assess the effect of industrial, domestic and agricultural waste.

Two species of wild rodents naturally occurring in the three areas served as sentinels and genotoxic damage – and damage to DNA – was the kind of adverse biological effect looked for. To evaluate the genotoxic impact of the environmental agents prevailing in each exposed area, genotoxicity indices were assessed in animals from the exposed areas and compared to those from the preservation area.

In this paper we chose one of the species from Bueno et al. (2000) – *Akodon montensis* (Am) – and three cytogenetics endpoints analyzed to illustrate an alternative statistical approach to the same data. The endpoints analyzed here are:

- the PCE/NCE ratio: the frequency of PCE (polychromatic erythrocytes) among the first 100 NCE (normochromatic erythrocytes) observed in a slide;
- the mitotic index (MI): the frequency of metaphases among 2000 cells;
- the MNPCE frequency: the frequency of micronucleated polychromatic erythrocytes within 2000 polychromatic erythrocytes counted.

Both the PCE/NCE ratio and the MI reflect the proliferative capacity of the bone-marrow cells. The first being the relation between young, immature, red cells and adult, mature, red cells, indicates if this process is normal. The latter shows the proportion of dividing cells in the white series. In other words, these indices show if hemopoiesis is being affected or not by some factor. Besides revealing possible cytotoxic effects of environmental contaminants, they may also reflect the influence of confounding factors such as, age, health, nutritional status, presence of pathogens, and intra- and inter-population genetic variability. It was assumed that these factors may have been present in all areas of study but did not prevail in any of them (Bueno et al. 2000). Therefore, differences in the endpoints would be caused by the external agents present in each area.

Micronucleated cells present in their cytoplasm DNA structures called micronuclei, originated from acentric chromosome fragments or from whole chromosomes lagging behind in their movement towards the cell poles during cell division. These two types of damage to the cell DNA – chromosome breaks and malsegregation, that is, errors in the chromosome distribution when the cell divides – are associated with the etiology of cancer, abortion, perinatal death, malformation, and mental retardation in humans. Micronuclei and chromosome breaks can only be visualized after the cell suffers one division.

The analysis of control endpoints, such as PCE/NCE and MI, shows whether or not cell division is altered. Once we are certain that the sample of individuals is within an acceptable range, with respect to the proliferative capacity of the cell lineages where the effect is looked for, our attention will be focused on the sample of cells that are exposed to the environmental agent.

Every individual in a certain locality, from the most sensitive to the most resistant, will be subjected to the exposure. The same can be said about its cells. There are two ways of considering the sample of exposed cells.

The first is to assume that each cell among the totality of cells from all individuals in the sample have the same chance of being damaged by the environmental agent. Consequently, the sample of cells is the total of cells obtained from all individuals, as if all of them were from one individual.

The second is to assume that, since each individual reacts differently to the same agent, the chance of a cell being damaged is different among individuals. Hence, the group of cells from each individual has to be treated as a separate entity.

Statistical models

In order to identify appropriate statistical models to analyze our data we have taken into account the following facts:

- 1) The counts for the endpoint PCE/NCE are observations of negative – binomial variables. PCE and NCE cells are observed up to obtain the 100th NCE cell. The frequency of PCE is then recorded;
- 2) The counts for metaphases among $m = 2000$ cells, the MI, were considered as an observation of a Poisson variable. Note that they are counts of rare events; and
- 3) The frequency of micronucleated polychromatic erythrocytes within 2000 polychromatic erythrocytes counted, the MNPCE, were also considered as observation of a Poisson variable.

270

The statistical questions posed in this paper are comparisons of three groups (one species in three areas), where the observations in each individual, within an area, are either a negative binomial variable or counts of rare events, a Poisson variable. It should be clear that each individual, despite being in the same area, reacts differently to an environmental aggression and different individuals will have, with high probability, different failure rates. The model used here considers that, for each area i , the failure rates (unobservable) of the individuals are generated independently by the same beta distribution with mean ε_i and standard deviation σ_i , $i = 1, 2, 3$. The frequency (observable) of damaged cells within a sample of cells is the observation for each individual in the i -th area. The mean and the standard deviation of the beta distribution are the only common quantities among the individuals of a specific area. These quantities may differ from area to area.

The model just described is a mixture model, that is, the likelihood function for the PCE/NCE in the i -th area, $i = 1, 2, 3$, is a product of beta-(negative)binomial densities evaluated at the observed frequencies depending on the unknowns ε_i and σ_i . For the MI and MNPCE we will have beta–Poisson densities instead of the beta-(negative)binomial ones.

Below we define the probabilistic properties of our models. We consider the following notation:

n_i = number of individuals studied in area i , $i = 1, 2, 3$. The order for the areas is PA, RF, and IA.

m_{ij} = sample size (observable) of cells collected from the j -th individual in the i -th area, $j = 1, 2, \dots, n_i$.

x_{ij} = number of damaged cells (observable) and

$X_i = (x_{i1}, x_{i2}, \dots, x_{ini})$.

π_{ij} = failure rate (unobservable) for the j -th individual in the i -th area, and

$\Pi_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{ini})$.

ε_i = mean (unobservable) of the beta distribution that generates the π_{ij} 's in the i -th area.

σ_i = standard deviation (unobservable) of the beta distribution that generates the π_{ij} 's in the i -th area.

The following restrictions about conditional probabilities are considered:

- Conditional on the knowledge of the π_{ij} 's, the observations x_{ij} 's are mutually independent and the distribution of each one depends only on the respective π_{ij} .
- For each area i , $i = 1, 2, 3$, the (unobservable) random variables, π_{ij} , form a sample of n_i independent and identically distributed beta random variables

with mean ε_i and standard deviation σ_i . A set of random quantities is mutually independent if its joint probability density is the product of the marginal probability densities.

- The random vectors Π_1 , Π_2 and Π_3 are also mutually independent.

4

Prior to posterior

The main interest in this article is to compare the three areas PA, RF, and IA. Under the Bayesian perspective, we can perform these comparisons focusing on two kinds of predictive distributions: the posterior distribution of the failure rate of a new individual or the predictive posterior distribution of the number of damaged cells within a sample of cells from this new individual. Let (π_1, π_2, π_3) and (x_1, x_2, x_3) be the vectors of failure rates and frequencies of damaged cells, respectively, for the three areas. Recall that we shall compare the three areas with respect to the three endpoints, PCE/NCE, MI, and MNPCE. Hence, we will calculate, for the failure rates, $\Pr\{\pi_1 < \pi_2|\text{data}\}$, $\Pr\{\pi_1 < \pi_3|\text{data}\}$, and $\Pr\{\pi_2 < \pi_3|\text{data}\}$ and for the damage cell frequencies, $\Pr\{x_1 < x_2|\text{data}\}$, $\Pr\{x_1 < x_3|\text{data}\}$, and $\Pr\{x_2 < x_3|\text{data}\}$. To compute these probabilities, the prior probabilities for the parameter vectors $(\varepsilon_1, \sigma_1)$, $(\varepsilon_2, \sigma_2)$, and $(\varepsilon_3, \sigma_3)$, must be accessed. The statistical model described considers that, among the three areas, the conditional distributions of the observations given these parameters are mutually independent. Consequently, assessing mutually independent priors for the parameters, the posterior distribution will preserve this independence property. That is, the vectors $(\varepsilon_1, \sigma_1)$, $(\varepsilon_2, \sigma_2)$ and $(\varepsilon_3, \sigma_3)$ are mutually independent a priori and a posteriori. Besides being convenient, this restriction on the choice of the priors is realistic. In addition, it allows the computation of the probabilities of interest listed above. As shown in Appendix A, even with this restriction, the computation of posteriors is intensive.

Usually the assessment of a prior distribution when the likelihoods are beta densities is controversial. If the choice is for the usual parameterization, beta (α, β) , one should consider bounded domains to define proper prior probability functions for (α, β) . Here, instead of (α, β) we believe that it is much more intuitive to have opinions about the mean, ε , and the standard deviation, σ , of the beta distribution. They are naturally bounded and the domain is given by

$$D = \{(\varepsilon, \sigma) : 0 < \varepsilon < 1 \text{ and } 0 < \sigma < \sqrt{\varepsilon(1 - \varepsilon)}\}$$

Recall that

$$\varepsilon = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \sigma^2 = \frac{\varepsilon(1 - \varepsilon)}{\alpha + \beta + 1}$$

Here, in order to permit fair comparisons with frequentist methods we use a uniform distribution over D . Other non-informative priors could be considered. For example, one could consider a uniform density for ε and, conditional on ε , a uniform density on $[0, \varepsilon(1-\varepsilon)]$ for σ . Since the sample of cells is large enough, the influence of the choice of such prior is minimum.

5

Results

Table 1 presents the observed values for the endpoints in all individuals from the three areas. The probability values comparing the three areas using π and x are

Table 1. Individual frequencies for each cytogenetic endpoint

| Unit | Sex | PCE/NCE | MI | MNPCE | Unit | Sex | PCE/NCE | MI | MNPCE |
|------|-----|---------|-----|-------|------|-----|---------|----|-------|
| PA1 | m | 19 | 31 | 7 | RF8 | m | 124 | 44 | 24 |
| PA2 | m | 52 | 19 | 9 | RF9 | m | 64 | 39 | 5 |
| PA3 | m | 50 | 26 | 1 | RF10 | m | 105 | 98 | 7 |
| PA4 | f | 31 | 54 | 8 | RF11 | f | 120 | 71 | 12 |
| PA5 | f | 68 | 22 | 6 | RF12 | f | 39 | 84 | 14 |
| PA6 | f | 53 | 35 | 6 | RF13 | f | 27 | 80 | 9 |
| PA7 | m | 45 | 30 | 12 | IA1 | f | 37 | 75 | 6 |
| PA8 | m | 28 | 46 | 3 | IA2 | m | 131 | 44 | 10 |
| PA9 | m | 99 | 21 | 6 | IA3 | f | 100 | 82 | 11 |
| PA10 | m | 27 | 27 | 5 | IA4 | m | 90 | 40 | 10 |
| PA11 | m | 31 | 45 | 8 | IA5 | f | 56 | 74 | 9 |
| PA12 | m | 57 | 32 | 14 | IA6 | f | 45 | 63 | 13 |
| RF1 | m | 86 | 45 | 23 | IA7 | m | 63 | 42 | 9 |
| RF2 | m | 31 | 112 | 8 | IA8 | m | 46 | 56 | 6 |
| RF3 | m | 83 | 80 | 13 | IA9 | m | 27 | 57 | 1 |
| RF4 | m | 58 | 53 | 17 | IA10 | m | 45 | 51 | 2 |
| RF5 | m | 31 | 58 | 6 | IA11 | m | 37 | 54 | 24 |
| RF6 | f | 90 | 69 | 14 | IA12 | f | 98 | 68 | 17 |
| RF7 | m | 175 | 17 | 12 | IA13 | m | 65 | 52 | 33 |

272

Table 2. Posterior probabilities used to compare areas

| Events | PCE/NCE | MI | MNPCE |
|-----------------|---------|-------|-------|
| $\pi_1 < \pi_2$ | 0.729 | 0.900 | 0.836 |
| $\pi_3 < \pi_2$ | 0.587 | 0.615 | 0.580 |
| $\pi_1 < \pi_3$ | 0.685 | 0.980 | 0.678 |
| $x_1 < x_2$ | 0.741 | 0.859 | 0.749 |
| $x_3 < x_2$ | 0.604 | 0.580 | 0.549 |
| $x_1 < x_3$ | 0.666 | 0.915 | 0.628 |

presented on Table 2. These figures indicate that, with relatively high probability, the preservation area presents lower failure rates and frequencies of damaged cells than the exposed areas. This conclusion is valid for all three endpoints.

We have chosen the PCE/NCE ratio to illustrate all posterior densities: the posterior densities of $(\varepsilon_1, \sigma_1)$, $(\varepsilon_2, \sigma_2)$, and $(\varepsilon_3, \sigma_3)$, the posterior predictive densities of π_1 , π_2 and π_3 and the posterior predictive probability functions of x_1 , x_2 , and x_3 .

For the PCE/NCE ratio, Fig. 1 presents the posterior densities of $(\varepsilon_1, \sigma_1)$ and $(\varepsilon_2, \sigma_2)$. Figure 2 shows the posterior predictive densities of π_1 , π_2 and π_3 and Fig. 3 the posterior predictive probability functions of x_1 , x_2 and x_3 . In an increasing order, all figures show that PA presents smaller values than IA, that presents smaller values than RF. This trend is valid for the three endpoints.

6

Discussion and conclusion

As observed by Bueno et al. (2000), the PCE/NCE ratio for the majority of the individuals of the samples from the three areas were within or very close to values reported in the literature that is, between 50 (Adler 1984) and 100 (Gollapudi and

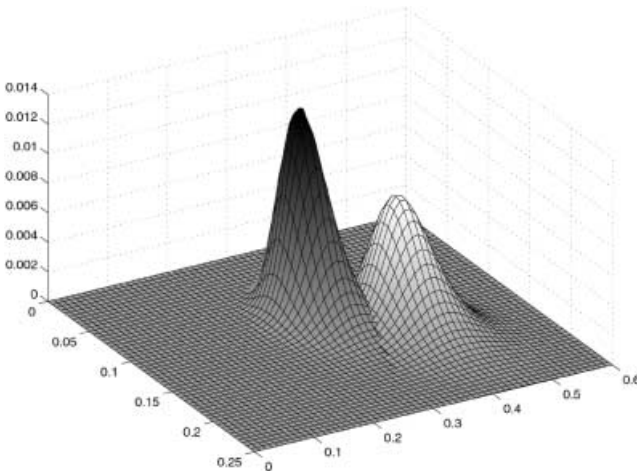


Fig. 1. Posterior distributions for $(\varepsilon_1, \sigma_1)$ and $(\varepsilon_2, \sigma_2)$ for PCE/NCE

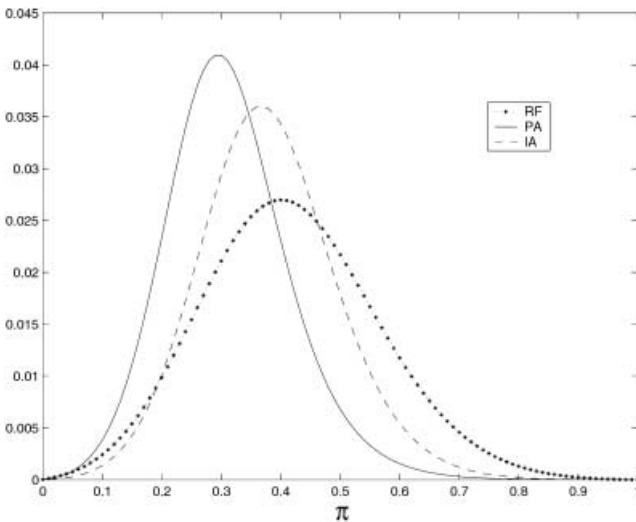


Fig. 2. Failure rate posterior densities for the mixture model for PCE/NCE

McFadden 1995). The values of the posterior probabilities for the PCE/NCE ratio, presented on Table 2, indicate that the hemopoiesis is slightly accelerated in the exposed areas RF and IA in relation to the preservation area PA. For the MI, the posterior probability values on Table 2 show a stronger trend in the same direction. The tendency of having smaller values in the preservation area suggests that the individuals in the exposed areas are reacting to environmental agents by increasing the proliferative capacity of their blood cells.

The frequency of damaged cells in the bone marrow is expressed by the MNPCE frequency. The posterior probabilities for this endpoint (Table 1) show that the frequency of damaged cells in the preservation area is smaller than in the exposed areas. Note that, for the three endpoints, the posterior probabilities suggest the existence of a risk order: $PA < IA < RF$.

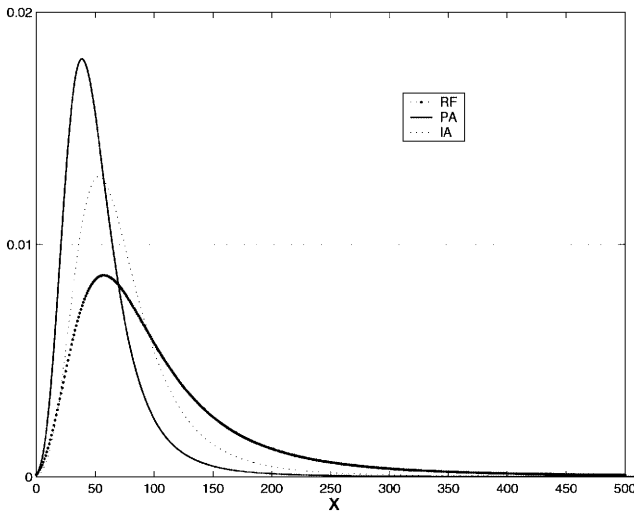


Fig. 3. Predictive probabilities for the mixture model for PCE/NCE

The results obtained with the evaluation of the posterior probabilities agree with the general conclusions reported by Bueno et al. (2000). The values for the differences among the areas, however, are much higher in Bueno et al. (2000) than the ones obtained in the present paper. The statistical models used in each case may be responsible for such disagreement. This disagreement does not result, however, as one might think, from the Bayesian approach used here. The use of a Bayesian statistical model without mixture would lead to the same results obtained by Bueno et al. (2000).

The statistical model considered here is the mixture model described in Sect. 3. The model without mixture considers the sufficient statistics – the sample total – as the observation in a specific area. This model ignores the variation among individuals and considers the cells of all individuals to be permutable, within and among individuals. That is, it considers the total of cells as if was sampled from a single animal.

By using the statistical model without mixture for the PCE/NCE ratio and considering a uniform prior for the failure rate, we obtain the posterior densities of the failure rates as illustrated on Fig. 4. Figure 5 contrasts the posterior densities of π_1 for the model without mixture with the observed failure rates.

It should be clear that for the model without mixture, most of the observed values of the failure rate, $x/(100+x)$, would have almost null density. This fact suggests that the model without mixture is not appropriate, resulting in an incoherence, since the observations are real and cannot have almost null densities. As seen on Fig. 6, all sample observation values have positive posterior densities under the model with mixture.

The main aspect of this paper is that one must consider the intrinsic natural differences among sample (individuals) units. The aggregate observations of several individuals may not be sufficient to define an adequate statistical model. Usually, one defines an unknown parameter of interest and considers observations varying randomly around the parameter, but often a more complex hierarchical model is needed as in the case presented here. For example, one wants to compare two or more surgical techniques used in

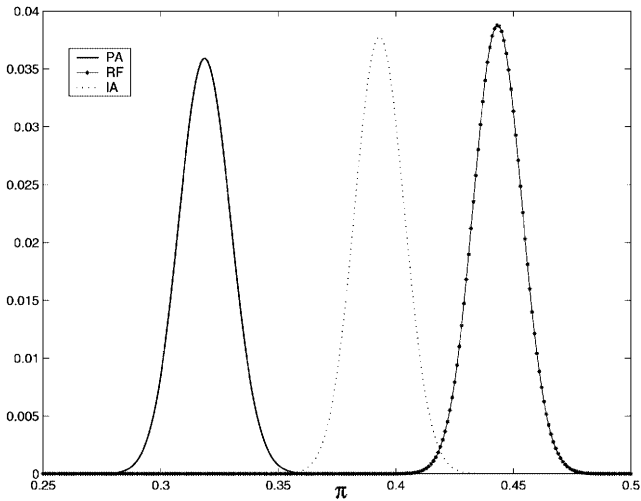


Fig. 4. Failure rate posterior densities for PCE/NCE model without mixture

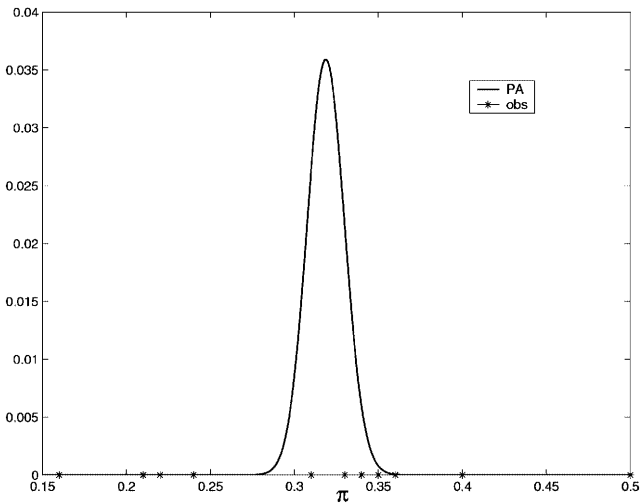


Fig. 5. Failure rates in PA – observed values and posterior density for PCE/NCE model without mixture

hospitals. For the hospitals using the same surgical techniques, to consider the frequency of failure surgeries as the only source of variation may not be appropriate. The conditions of the hospitals and their infrastructure are, in most cases, a large additional source of variability. Hierarchical meta-analysis procedures seem to be the correct techniques to be used (Berry 1990, 1993).

The choice of an adequate hierarchical statistical model may be a difficult task. In the case presented here, the beta-(negative)binomial and the beta-Poisson models are very appropriate, as indicated on Fig. 6, whereas the simple non-hierarchical model is inadequate, as indicated on Fig. 5.

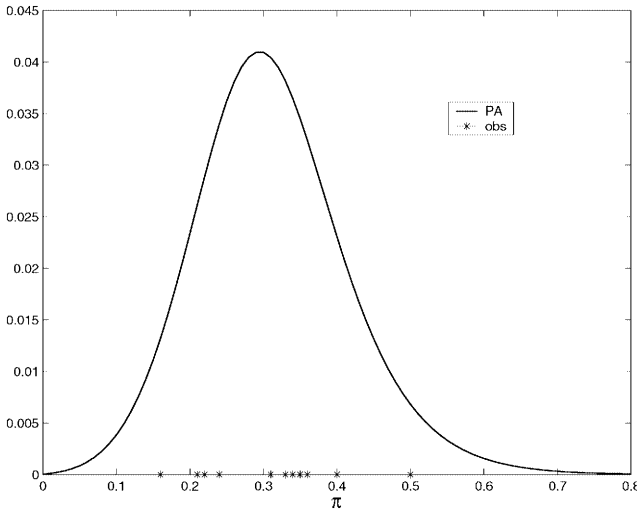


Fig. 6. Failures rates in PA – observed values and posterior densities for PCE/NCE model with mixture

The general conclusion that individuals suffer from environmental modifications is the same as in Bueno et al. (2000). We believe, however, that in this paper we present a sounder statistical model to support it.

Appendix A: Computational Aspects

The following expressions are the functions used in the text:

- Gamma function:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \text{for } \alpha > 0$$

- Beta function:

$$B(\alpha, \beta) = \int_0^1 p^{\alpha-1} (1-p)^{\beta-1} dp = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad \text{for } \alpha, \beta > 0$$

- Negative binomial probability function:

$$Bi(x|\pi, 100) = \binom{99+x}{x} \pi^x (1-\pi)^{100}, \quad 0 < \pi < 1$$

- Poisson probability function:

$$Po(x|m\pi) = \frac{e^{-m\pi} (m\pi)^x}{x!}, \quad \pi > 0$$

where m = cell's sample size.

- Beta probability density function:

$$Be(\pi|\alpha, \beta) = \frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(\alpha, \beta)} \quad \text{for } \alpha, \beta > 0$$

- Beta-(negative)binominal probability function:

$$\begin{aligned} \text{BeBi}(x|\alpha, \beta, 100) &= \int_0^1 \text{Be}(\pi|\alpha, \beta)\text{Bi}(x|\pi, 100)d\pi \\ &= \binom{99+x}{x} \frac{B(\alpha+x, \beta+100)}{B(\alpha, \beta)} \end{aligned}$$

- Beta-Poisson probability function:

$$\begin{aligned} \text{BePo}(x/t, \alpha, \beta) &= \int_0^1 \text{Be}(\pi/\alpha, \beta)\text{Po}(x/t\pi)d\pi \\ &= \frac{m^x M(\alpha+x, \alpha+\beta+x, -t)\Gamma(\alpha+\beta)\Gamma(\alpha+x)}{\Gamma(\alpha)\Gamma(x+1)\Gamma(\alpha+\beta+x)} \end{aligned}$$

$$M(a, c, z) = \sum_{j=0}^{\infty} \frac{(a)_j z^j}{(c)_j (1)_j}$$

is the Kummer function, also known as the confluent hypergeometric function $F_1(a, c, z)$.

- $(a)_j$ is the Pochhammer polynomial of argument a and degree j . That is, $(a)_0 = 1$ and $(a)_j = a(a+1) \cdots (a+j-1)$.
- The interval $[0, t]$ is the domain of the Poisson observations.

The prior probability function used here is a uniform distribution on the parameter space D of the new parameterization (ϵ, σ) as defined in Sect. 4.

The computation of the posterior densities for the beta-binomial and the beta-Poisson are greatly simplified by integrals BeBi and BePo having analytic expressions. Those expressions are written in terms of common transcendental functions like Γ and not so common ones like the Kummer or confluent hypergeometric functions, M . Efficient computer codes for all those transcendental functions are available (Spanier and Oldham 1987; Zhang and Jin 1996).

In order to compute the posterior for the mixing distribution we built a grid on the parameter space D . At every point on the grid the corresponding integral is evaluated.

From the mixing distribution we can compute the density for the failure rate π and from that we generate the predictive distribution of x , a future observation. All the grid sizes are made small enough so that the discretization effect in the final distributions can be neglected.

References

Adler I-D (1984) Cytogenetic tests in mammals. In: Venitt JPS, (ed., Mutagenicity Testing, IRL Press, Oxford, pp. 275-306
 Aitchison J, Dunsmore IR (1975) Statistical Prediction Analysis. Cambridge University Press
 Berry DA (1990) A Bayesian approach to meta-analysis and multicenter trials. In: Proceedings: Pharmaceutical Section of the American Statistical Association, pp. 1-10
 Berry DA (1993) A case of bayesianism in clinical trials (with discussion). Statistics in Medicine 12: 1377-1404

- Bueno AMS, de B. Pereira CA, Rabello-Gay MN** (2000) Environmental genotoxicity evaluation using cytogenetic end-points in wild rodents. *Environ. Health Perspectives* 108: 1165-1169
- Gollapudi BB, McFadden LG** (1995) Sample size for the estimation of polychromatic to normochromatic erythrocyte ratio in the bone marrow micronucleus test. *Mutation Res.* 347: 97-99
- Sandhu SS, de Serres FJ** (1989) In situ evaluation of biological hazard of environmental pollutants. *Mutation Res.* 216: 341-352
- Spanier J, Oldham KB** (1987) *An Atlas of Functions*, Hemisphere Pub. Corp, New York
- Zhang S, Jin J** (1996) *Computation of Special Functions*, J. Wiley, New York