

Assessing Poisson Variation of Intestinal Tumor Multiplicity in Min Mice Carrying a Robertsonian Translocation

Michael A. Newton* and David I. Hastie
University of Wisconsin and University of Bristol

May 2004

Abstract

Tumor multiplicity is a frequently measured phenotype in animal studies of cancer biology. Conditions under which this measurement has a Poisson distribution represent a biological and statistical reference point that is usually violated owing to uncontrolled sources of variation. A recent experiment on murine intestinal tumors presented conditions which seem to provide for Poisson-distributed tumor counts. In considering hypothesis testing strategies, model choice, and Bayesian approaches, we quantify the positive evidence favoring Poisson variation in this experiment. Statistical techniques used include likelihood-ratio testing, BIC and AIC, negative-binomial modeling, reversible-jump MCMC, and posterior predictive checking. The BIC-based posterior approximation is found to be quite accurate in this small- n case study.

1 Introduction

According to first principles, the Poisson distribution naturally models the number of cancerous tumors that appear in a tissue during a fixed time period (e.g., Moolgavkar and Knudson 1981; Kokoska 1987); however, extra-Poisson variation is widely observed in this sort of count data (e.g., Drinkwater and Klotz 1981; Moser *et al.* 1990; Moser *et al.* 1992; Gould *et al.* 1996; Nagase *et al.* 1999; Chulada *et al.* 2000; Ramachandran *et al.* 2002). The Poisson tumor multiplicity model corresponds to a notion of pure randomness in tumor formation. Yet in highly controlled laboratory experiments, both measured and unmeasured factors may fluctuate among the experimental units and thus infuse additional variation into the stochastic process of tumor formation. Cancer researchers are interested

*To whom correspondence should be addressed: newton@stat.wisc.edu

in the sources of variation affecting tumor growth and in the biological conditions that approximate the state of pure randomness.

Those who carry a defective copy of the adenomatous polyposis coli gene (Apc in mice, APC in humans) are predisposed to intestinal cancer (e.g., Hardy *et al.*, 2000). The Min (multiple intestinal neoplasia) strain of the laboratory mouse carries a nonsense Apc allele, and, consequently, presents with tumors throughout the intestinal tract (Moser *et al.* 1990; Su *et al.* 1992). Min mice thus provide an experimental system to study the biology and genetics of normal intestinal tissue and its neoplastic transformation (Dove *et al.* 1998). As in humans, the murine intestinal tract is comprised of a large number of organized groups of proliferating cells called crypts, and it is from within aberrant crypts that tumors begin to form (Li *et al.* 1994; Preston *et al.* 2003). A Poisson distribution of tumor counts would emerge, for example, if after the post-natal establishment of the crypt layer (Simon and Gordon 1995) each crypt becomes aberrant with some small probability and does so independently of other crypts. Other scenarios may also lead to the same final distribution (Arratia *et al.* 1990).

Haigis and Dove (2003) obtained data on tumor multiplicities from 4 groups of Min mice which were arranged to be genetically identical but which carried different forms of the Robertsonian translocation (Rb9); this is a chromosomal construct in which certain pairs of chromosomes are fused together. (The translocation alters the genome organization without changing its content.) The Haigis-Dove mice, which carried a fusion of chromosome 7 with the Apc-harboring chromosome 18, were used primarily to study factors by which the wild-type Apc allele is inactivated within tumors; an important observation was that animals carrying an Rb9 translocation had significantly reduced tumor multiplicity, indicating that non-genetic chromosomal factors play a role in tumor susceptibility. Another striking feature of the experimental data was that the tumor multiplicities in Rb9 animals appeared to exhibit purely Poisson variation. Being a biological and statistical reference point, the claim of Poisson-ness is of interest in its own right. The purpose of the present article is to look closely at the assessment of Poisson variation in the Haigis-Dove data.

Table 1 reports the full set of tumor multiplicity data analyzed by Haigis and Dove (2003). As in many mouse studies, the sample size here is modest; there are 4 groups of sizes 16, 17, 15, and 18. For notation we retain the group labels as in the source paper where the reader may find the precise definitions. Briefly, the +/+ group is a control group in which animals do not carry any Rb9 translocation, and the other groups correspond to different ways that the fusion of chromosomes 7 and 18 can occur. Following a standard

protocol, tumor counts were measured on the whole intestinal tract from animals sacrificed at 60 days of age. The significant reduction in tumor multiplicity in the Rb9 groups is evident from inspection of Table 1 and this can be confirmed by a variety of statistical tests. (Haigis and Dove, 2003, used the nonparametric Tukey-Kramer test.) The possible Poisson-ness of tumor multiplicity is suggested by the close agreement of sample means and sample variances in the Rb9 groups. Of interest here is how to quantify this close agreement.

2 Methods

2.1 Testing strategies

Experimentalists often adopt a hypothesis-testing posture when evaluating data. A simple testing approach is to assume that the data are Poisson, with group-specific means perhaps, and then to assess the evidence against this hypothesis. Conditioning on group sums to eliminate nuisance parameters, the Poisson counts become multinomials. We can consider the conditional distribution of some test statistic – for example each group provides a sample coefficient of dispersion (sample variance divided by sample mean) – and calibrate this test statistic using Monte Carlo (Barnard 1963).

A difficulty with this simple approach is that the inference does not really address the question of interest. For example, in the event that we do not reject Poisson-ness for one of the Rb9 groups, we can infer only that this group is consistent with Poisson variation – i.e., we have not compiled sufficient contradictory evidence. We would not have directly measured the Poisson-ness of the Rb9 data. Further, from historical data we expect extra-Poisson variation; it seems to be an undue contrivance to consider the Poisson hypothesis as a null by which our inference is gauged.

Professor Drinkwater (personal communication) suggested a different hypothesis-testing strategy for the Haigis-Dove problem based on his experience that the negative-binomial distribution often fits tumor-count data (e.g., Drinkwater and Klotz, 1981). Recall that this well-studied, two-parameter family has probability mass function, for $x \geq 0$,

$$p(x|\kappa, \lambda) = \frac{1}{\kappa^{1/\kappa}\Gamma(1/\kappa)} \frac{\Gamma(x + 1/\kappa)}{\Gamma(x + 1)} \frac{\lambda^x}{(\lambda + 1/\kappa)^{x+1/\kappa}}. \quad (1)$$

In this parameterization, the expected tumor count is λ and the variance is $\lambda(1 + \kappa\lambda)$ for $\kappa \geq 0$. The limit $\kappa \rightarrow 0$ corresponds to the Poisson distribution. Recall also that when viewing this distribution as a mixture of Poisson mixands, $1/\kappa$ is the shape parameter of

the Gamma mixing distribution (Greenwood and Yule 1920). Drinkwater’s proposal was to allow group-specific λ values and to perform a likelihood ratio test for a common shape parameter (i.e. common κ). The rationale is that the shape parameter characterizes the degree of overdispersion and so testing for equal shapes would assess the significance of variation in the observed coefficients of dispersion. An advantage in using Drinkwater’s approach is that we are not assuming the Poisson null hypothesis up front. However, again the inference does not directly address the scientific question of Poisson-ness – a limiting case which has both biological and statistical importance. A solution, which is not obvious to many experimentalists, is to eschew hypothesis testing and to frame the problem as one of model choice or model averaging.

2.2 Model Choice

Based on historical data, it is reasonable to adopt the negative-binomial model as a general description of tumor multiplicity. Various sub-models are obtained by restricting the parameters; for instance setting $\kappa = 0$ corresponds to the Poisson case. To deal properly with the multiple groups, we introduce group labels $i = 1, \dots, 4$, and allow the possibility that each group can have its own mean parameter λ_i and dispersion parameter κ_i . Therefore, the full model contains 8 free parameters. Drinkwater’s likelihood ratio test considered above compares this full model to the 5 parameter null model in which $\kappa_1 = \kappa_2 = \kappa_3 = \kappa_4 \geq 0$ and the 4 means are unrestricted. The simple Monte Carlo tests work on the Poisson null hypothesis that $\kappa_i = 0$.

A host of different sub-models emerges by considering equality constraints among $\{\lambda_i\}$, equality constraints among $\{\kappa_i\}$ and the possibility that each κ_i may equal 0. Table 2 notes the 15 possible equality patterns among the mean parameters; these correspond to the Bell number of set partitions of 4 objects. We do not tabulate them all, but we observe that likewise there are 52 distinguishable constraint patterns for the dispersion parameters. To see this, note

$$52 = \sum_{j=0}^4 \binom{4}{j} a_j \tag{2}$$

where the a_j are also Bell numbers: $(a_0, a_1, a_2, a_3, a_4) = (1, 1, 2, 5, 15)$. There are $\binom{4}{j}$ ways to choose j groups that will have non-zero κ_i values; for each such choice there are a_j patterns of equality among the j non-zero values. Taken together, we have $780 = 15 \times 52$ distinct sub-models of the general negative-binomial model.

We label each sub-model by $m = (l, k)$ where $l \in \{1, \dots, 15\}$ is an index of the equality pattern among components of $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_4)$, and likewise $k \in \{1, \dots, 52\}$ is an index for the equality and zero pattern for $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_4)$. By our coding for example, $m = (12, 20)$ entails three free λ_i 's and two free non-zero κ_i 's. More specifically, the means satisfy: $\lambda_2 = \lambda_3$, $\lambda_1 \neq \lambda_3$, $\lambda_1 \neq \lambda_4$, and $\lambda_3 \neq \lambda_4$ (this is $l = 12$); the dispersion parameters satisfy: $\kappa_2 = \kappa_3 = 0$, $\kappa_1 > 0$, $\kappa_4 > 0$, and $\kappa_1 \neq \kappa_4$ (this is $k = 20$).

We can readily maximize the log-likelihood within each sub-model and note the number of free parameters. These quantities can be combined in various ways to score a balance between model fit and model complexity: the Akaike Information Criteria (AIC, Akaike 1983) and the Bayes Information Criteria (BIC, Schwarz 1978; Smith and Spiegelhalter 1980). Compared to testing strategies, the model choice approach more directly addresses the question of Poisson-ness of the Min tumor counts. The approach is limited, however, because it does not quantify the extent of evidence in favor of each hypothesis.

2.3 Bayesian approaches

Going further, we pursue posterior analysis using the negative-binomial likelihood and prior distributions over the multitude of sub-models. Our first calculation transforms the BIC values into approximate sub-model posterior probabilities (Kass and Wasserman 1995). This does not change the ranking of sub-models, because the approximate posterior mass for sub-model m is $p(m|\text{data}) \propto \exp(\text{BIC}/2)$, but it does allow us to calibrate the ranking and also to perform marginal posterior inference. For example, to obtain the marginal posterior probability of Poisson-ness for group i , we sum $p(m|\text{data})$ over the sub-models m for which $\kappa_i = 0$.

The BIC-based posterior computations rely on large-sample theory and ought to be viewed as a first-approximation to a more explicit Bayesian analysis. Unsure of the degree of this approximation, we implement a fully-specified Bayesian analysis via reversible jump (i.e., trans-dimensional) Markov chain Monte Carlo. An advantage of this approach is that the issues of model choice and parameter estimation are dealt with simultaneously and there are no large-sample approximations; a challenge is that some care is required in setting the prior and in implementing the posterior sampling.

We consider two priors which for convenience share the factorization

$$p(m, \boldsymbol{\lambda}, \boldsymbol{\kappa}) = p(m) p(\boldsymbol{\lambda}|l) p(\boldsymbol{\kappa}|k), \tag{3}$$

but which for robustness differ in how they assign mass $p(m)$ to the sub-models. The first

prior, denoted A , entails

$$p(m) = p(l)p(k) = (1/15) \times (1/52) = 1/780$$

for all sub-models m . This has the basic appeal of non-informativeness, but also it aims to mimic the previous BIC-based computations. The combinatorics of this uniform sub-model prior imply $P(\kappa_i = 0) = 15/52 \approx 0.29$ for each i , and thus global uniformity does not confer a balance in this marginal prior probability. The second prior, B , is set up specifically to focus on the question of Poisson-ness and entails $P(\kappa_i = 0) = 1/2$ for all i . To achieve this balance we retain independence of the pattern indices l and k and uniformity of l , but we modify $p(k)$ so that it boosts mass on the Poisson sub-models. If the pattern k entails j nonzero κ_i values, then we take $p(k) = 1/(2^4 a_j)$ where a_j is the j th Bell number, as in (2).

For priors A and B in (3) to be fully specified, we need to describe prior uncertainty in the free parameters conditional on m . We take vague independent Gamma priors and use the same set for both A and B . Specifically, free components of $\boldsymbol{\lambda}$ are regulated by a Gamma(2, 0.1) and the those of $\boldsymbol{\kappa}$ are regulated by a Gamma(1, 2). Posterior analysis indicates little sensitivity to these choices.

Posterior computation is enabled by MCMC. Standard Metropolis-Hastings updates are available to sample the free parameters in $\boldsymbol{\lambda}$ and $\boldsymbol{\kappa}$ conditional on the the data and the sub-model m (e.g. Robert and Casella 2002), however it is evident that different sub-models can have different dimensions, and thus these basic updates are not suitable for updating m . To accommodate the varying dimensionality, we invoke a trans-dimensional MCMC sampler that allows jumping between sub-models of different dimensions (Green 1995, 2003). In contrast to many other applications of trans-dimensional MCMC, this case is interesting because there is only a partial nesting of sub-models. Furthermore, the presence of two pattern variables l and k in m means the situation is slightly more general than the typical problem involving only a single model index. In updating m , we accommodate the partial nesting and cycle through two move classes: one proposes a new mean pattern l given the dispersion pattern k , and the other proposes a new k given l . The close linkage between m and $(\boldsymbol{\lambda}, \boldsymbol{\kappa})$ means that free parameter values need to be created and destroyed during these proposals. Further details are in the Appendix.

The output from the MCMC provides direct access to properties of marginal posterior distributions. We can look not only at the probabilities of the pattern indices l and k and various parameters, but we can evaluate the marginal posterior probabilities of Poisson-ness in each treatment group. Through MCMC output analysis we are confident that the reported posterior probabilities are accurate to two decimal places.

The full Bayesian analysis proposed here rests on the general negative binomial model. Although it is known to fit well in historical data (e.g. Drinkwater and Klotz 1981), its validity ought to be checked in the present case. Posterior predictive checks are suitable for this purpose (e.g. Gelman *et al.* 2003). For each 4-group data set simulated from the posterior predictive, we find the best fitting negative binomial distribution in each group, and then we measure the maximum distance between each fitted distribution function and the corresponding empirical distribution.

3 Results

3.1 Testing strategies

We find via Monte Carlo testing that each of the Rb9 groups is consistent with the Poisson hypothesis but the +/+ group is not (Table 3).

Following Professor Drinkwater’s proposal, we computed a generalized likelihood ratio statistic of 4.1 which, nominally, is distributed as a chi-square with 3 degrees of freedom (Table 4). This gives p -value 0.25 which, of course, is not significant. Although the sample coefficients of dispersion for the Rb9 groups appear to be quite different from the +/+ group, the data are consistent with the common-shape negative-binomial distribution allowing group-specific means.

One might object to our use of the chi-square calibration; bootstrapping or some technique to accommodate the boundary effects might be better, but in our view that would be placing effort in the wrong direction. One might also pool the Rb9 groups to improve sample size, but this also has little effect and anyway seems to involve arbitrary pre-test decisions.

3.2 Model choice

Table 5 indicates the position of several sub-models by the BIC and AIC criteria. Both criteria concur that, in the context of the general negative-binomial model, the best explanation of the Haigis-Dove data entails Poisson variation and common mean in Rb9 cis and Rb9 trans groups, and common shape but different means in the other two groups. Poisson variation in all Rb9 groups is a close second place explanation for both criteria, though they balance fit and complexity differently for other sub-models. By this analysis we have some positive evidence in favor of the Poisson hypothesis for Rb9 groups.

3.3 Bayesian approaches

In the first calculation, we transform the BIC values into sub-model posterior probabilities. The top two sub-models from Table 5 each carry 10% of the posterior mass, the best 11 sub-models account for half the posterior mass, and 80% probability concentrates on the best 36 sub-models. Again we can conclude that evidence favors Poisson variation in the Rb9 cis and Rb9 trans groups.

The full Bayesian analysis via MCMC allows a range of inferences and measures the accuracy of the BIC-based computation. Part of the analysis is to study trace plots such as in Figure 1, which help to confirm that our MCMC sampler is performing sufficiently. Figures 2 and 3 show histograms of posterior samples for the group-specific parameters λ_i and κ_i . These distributions are marginal with respect to the sub-model index $m = (l, k)$. Though the parameter values themselves are directly linked to the data, the constraint pattern behind these parameters is of greater interest here. Table 6 records the (non-zero) marginal posterior probabilities of the mean-pattern index l . The modal pattern $l = 12$, which accounts for just over half the posterior mass under either prior, entails 3 different rates of tumor occurrence: a common rate for Rb9 trans and Rb9 cis, and distinct rates for the other groups. Significantly, there is no posterior support for any of the values of l in which the control group $+/+$ shares a mean parameter with any of the other groups. This is consistent with earlier calculations indicating the strong effect of the Rb9 translocation on expected tumor count.

A similar analysis of the dispersion pattern k indicates that two patterns contain most of the posterior mass under both priors. They are $k = 5$, in which only the $+/+$ group is not Poisson, and $k = 9$, in which $+/+$ and Rb9/Rb9 share a non-zero κ value and the other groups are Poisson. Other dispersion patterns are considerably less probable, although there is more mass outside the modal two values using prior A than using prior B .

Table 7 shows 11 sub-models $m = (l, k)$ which include the top 6 sub-models for each prior. Ranking is by marginal posterior probability. The top two sub-models, $m = (12, 5)$ and $m = (12, 9)$, are the same as those selected by BIC and AIC (Table 5), although the order is reversed; further differences emerge as we look down the ranking. These top two sub-models account for 18% of the total posterior mass under the flat prior A and 34% under the Poisson-balanced prior B .

Of primary interest is the posterior probability of Poisson-ness, i.e. $P(\kappa_i = 0|\text{data})$ for each group i . Table 8 shows these marginal probabilities for the full analysis via MCMC using both priors, and for the BIC approximation. The final evidence is relatively weak,

but all calculations support the Poisson model for Rb9 trans and Rb9 cis. The support is positive and in favor of the Poisson hypothesis rather than being a lack of evidence against this hypothesis. Further, the support is quantified probabilistically at between 0.6 and 0.8, depending on the prior. An interesting feature of marginalization versus selection is that although the most probable sub-model entails Poisson variation for Rb9/Rb9 (in the full analysis, prior *A*), the chances are better than 0.5 that this variation is not Poisson.

As expected by theory (e.g., O’Hagan 1994), the flat-prior calculations (prior *A*) are in closer agreement with the BIC approximation than those for prior *B*. Noting the small sample size, a striking feature in this case study is the quality of the asymptotic BIC approximation.

To assess the goodness-of-fit of the general negative binomial model we consider posterior predictive checks. Figure 4 compares the empirical distribution of tumor counts within each group to a best-fitting negative binomial for that group. To calibrate these fits, we compare this figure to similar results for 1000 data sets simulated from the posterior predictive. The 1000 parameter settings were taken by thinning one of the MCMC runs under prior *A*. The results in Figure 5 indicate an adequate fit of this general model, and thus support the model-based computations.

4 Discussion

The biological mechanisms that regulate the intestinal epithelium must coordinate the continual proliferation of tens of millions of cells. We know that tumors form when this regulation fails, and, though much is known about this cancer biology, very little is known about the early stages of tumor formation. A necessary early genetic step in the process is loss of the APC (Apc) gene product. One pathway to Apc inactivation was closed in the Rb9 animals of the Haigis-Dove experiment, and so their tumors were obliged to lose Apc by some other mechanism. By our statistical analysis we have measured the distribution of the resulting tumor counts and have found positive evidence supporting Poisson variation. Poisson variation represents an ideal biological reference point which has not been achieved in other studies of the Min mouse. The Poisson model is consistent with independent and identically distributed Bernoulli crypt failure within the mouse, and thus our findings convey some basic information about the possible mechanisms of Apc inactivation.

The statistical tools employed in this case study include hypothesis testing, model selection, model averaging, and posterior predictive checking. Though they vary in terms

of computational demands and the nature of their conclusions, all approaches support the hypothesis, so tantalizingly clear from inspection of Table 1, that Rb9 data are Poisson-distributed.

This case study provides a context for comparing different statistical approaches. We find limitations of classical hypothesis testing. We find that different model choice schemes can agree on the best explanation of data and can provide positive support for certain hypotheses. Although the sample size is small, we find that asymptotic relationships between BIC and posterior model probability can be quite accurate. We gain insights into the use and effectiveness of fully-specified Bayesian analysis fitted with trans-dimensional MCMC. Though it is much more computationally demanding than the other approaches and requires more detailed specification, it also provides the richest and most quantitative conclusions. This approach deals simultaneously with sub-model constraints and free-parameter values, and thus avoids problems inherent in other approaches. Aspects of the sub-model rankings are sensitive to the prior specification, but other features of the posterior are not, thus demonstrating the nature of information in the data.

References

- Akaike, H. (1983). Information measures and model selection. *Bulletin of the International Statistical Institute*, **50**, 277–290.
- Arratia, R., Goldstein, L., and Gordon, L. (1990). Poisson approximation and the Chen-Stein method. *Statistical Science*, **5**, 403–424.
- Barnard, G. (1963). Discussion of paper by M. S. Bartlett. *Journal of the Royal Statistical Society Series B*, **25**, 294.
- Chulada, P. C., Thompson, M. B., Mahler, J. F., Doyle, C. M., Gaul, B. W., Lee, C., Tiano, H. F., Morham, S. G., Smithies, O., and Langenbach, R. (2000). Genetic disruption of *Ptgs-1*, as well as of *Ptgs-2*, reduces intestinal tumorigenesis in *Min* mice. *Cancer Research*, **60**, 4705–4708.
- Dove, W. F., Cormier, R. T., Gould, K. A., Halberg, R. B., Merritt, A. J., Newton, M. A., and Shoemaker, A. R. (1998). The intestinal epithelium and its neoplasms: genetic, cellular, and tissue interactions. *Phil. Trans. R. Soc. Lond. B*, **353**, 915–923.
- Drinkwater, N. R. and Klotz, J. H. (1981). Statistical methods for the analysis of tumor multiplicity data. *Cancer Research*, **41**, 113–119.

- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman and Hall.
- Gould, K. A., Luongo, C., Moser, A. R., McNeley, M. K., Borenstein, N., Shedlovsky, A., Dove, W. F., Hong, K., Dietrich, W. F., and Lander, E. S. (1996). Genetic evaluation of candidate genes for the *Mom1* modifier of intestinal neoplasia in mice. *Genetics*, **144**, 1777–1785.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**, 711–732.
- Green, P. J. (2003). *Highly Structured Stochastic Systems*, Chapter 6, Trans-dimensional Markov chain Monte Carlo, pp. 179–196. Oxford University Press.
- Greenwood, M. and Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society*, **83**, 255–279.
- Haigis, K. and Dove, W. (2003). A Robertsonian translocation suppresses a somatic recombination pathway to loss of heterozygosity. *Nature Genetics*, **33**, 33–39.
- Hardy, R. G., Meltzer, S. J., and Jankowski, J. A. (2000). ABC of colorectal cancer: Molecular basis for risk factors. *British Medical Journal*, **321**, 886–889.
- Hastie, D. I. (2004). *Towards automatic reversible jump MCMC*. Ph. D. thesis, Department of Mathematics, University of Bristol.
- Kass, R. E. and Wasserman, L. (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criteria. *Journal of the American Statistical Association*, **90**, 928–934.
- Kokoska, S. M. (1987). The analysis of cancer chemoprevention experiments. *Biometrics*, **43**, 525–534.
- Li, Y. Q., Roberts, S. A., Paulus, U., Loeffler, M., and Potten, C. S. (1994). The crypt cycle in mouse small intestinal epithelium. *Journal of Cell Science*, **107**, 3271–3279.
- Moolgavkar, S. H. and Knudson, A. G. (1981). Mutation and cancer: a model for human carcinogenesis. *Journal of the National Cancer Institute*, **66**, 1037–1052.
- Moser, A. R., Dove, W. F., Roth, K. A., and Gordon, J. I. (1992). The *Min* (multiple intestinal neoplasia) mutation: its effect on gut epithelial cell differentiation and interaction with a modifier system. *Journal of Cell Biology*, **116**, 1517–1526.

- Moser, A. R., Pitot, H. C., and Dove, W. F. (1990). A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science*, **247**, 322–324.
- Nagase, H., Mao, J. H., and Balmain, A. (1999). A subset of skin tumor modifier loci determines survival time of tumor-bearing mice. *Proceedings of the National Academy of Sciences*, **96**, 15032–15037.
- O’Hagan, A. (1994). *Kendall’s Advanced Theory of Statistics Volume 2B Bayesian Inference*. Edward Arnold.
- Ramachandran, S., Fryer, A. A., Lovatt, T., Smith, A., Lear, J., Jones, P. W., and Strange, R. C. (2002). The rate of increase in the numbers of primary sporadic basal cell carcinomas during follow up is associated with age at first presentation. *Carcinogenesis*, **23**, 2051–2054.
- Robert, C. P. and Casella, G. (2002). *Monte Carlo Statistical Methods*. Springer.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.
- Simon, T. G. and Gordon, J. I. (1995). Intestinal epithelial cell differentiation: new insights from mice, flies and nematodes. *Current Opinion in Genetics & Development*, **5**, 577–586.
- Smith, A. F. M. and Spiegelhalter, D. J. (1980). Bayes factors and choice criteria for linear models. *Journal of the Royal Statistical Society Series B*, **42**, 213–220.
- Su, L. K., Kinzler, K. W., Vogelstein, B., Preisinger, A. C., Moser, A. R., Luongo, C., Gould, K. A., and Dove, W. F. (1992). A germline mutation of the murine homolog of the APC gene causes multiple intestinal neoplasia. *Science*, **256**, 668–670.

Appendix: MCMC Implementation

Our sampler is designed to take advantage of the partial nesting of the $\boldsymbol{\lambda}$ and $\boldsymbol{\kappa}$ constraint patterns. In particular, suppose the current mean pattern is indexed by l . All moves that propose a new state l' of different dimension (i.e. $q_{l'} \neq q_l$), are such that either the pattern of $\boldsymbol{\lambda}$ corresponding to l is nested within the pattern corresponding to l' or vice versa. At most a move that alters the dimension by one is attempted. Table 9 details the permissible models l' which can be proposed given that the current model is l . We emphasise again that for such moves, k remains constant between the current and new model. A similar table (not shown) can be created for the second move class which updates k while keeping l fixed.

Consider the dynamics of one particular trans-dimensional move type. Suppose we are in sub-model $m = (l, k)$. We propose a move to sub-model $m' = (l', k)$ as follows. With probability b_l we propose a move to l' such that $q_{l'} = q_l + 1$ (a birth move). Alternatively, with probability d_l we propose a move to l' such that $q_{l'} = q_l - 1$ (a death move). We set $b_0 = 1$, $d_0 = 0$ and $b_{15} = 0$, $d_{15} = 1$ and for all other values of l we take $b_l = d_l = \frac{1}{2}$. Having chosen between a birth and death move, we choose the particular value of l at random from the $r_{l, q_{l'}}$ particular candidates. With the proposed model in place, we need to propose new parameter values $\boldsymbol{\lambda}'$. Importantly, we note we are actually proposing a move from the q_l unique parameter values $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{q_l}$ to the $q_{l'}$ unique parameters $\tilde{\lambda}'_1, \dots, \tilde{\lambda}'_{q_{l'}}$. For the birth move, this is achieved by generating a $\text{LogNormal}(0, \sigma_\lambda)$ random variable, u , and then setting

$$\tilde{\lambda}'_{i_1} = \tilde{\lambda}_{i_1} u, \quad \tilde{\lambda}'_{i_2} = \frac{\tilde{\lambda}_{i_1}}{u}, \quad \text{and} \quad \tilde{\lambda}'_j = \begin{cases} \tilde{\lambda}_j & j \neq i_1, j < i_2 \\ \tilde{\lambda}_{j+1} & j \neq i_1, j > i_2. \end{cases}$$

The values of indices i_1 and i_2 are obvious from the current and proposed mean patterns l and l' as shown in the example below. The proposed parameter vector $\boldsymbol{\lambda}'$ is then immediate from $\tilde{\lambda}'_1, \dots, \tilde{\lambda}'_{q_{l'}}$ and $\boldsymbol{\kappa}$ remains unchanged.

To clarify this move, consider the example of moving from model $l = 8$ to model $l' = 12$. This is a birth move since $q_l = 2$ and $q_{l'} = 3$ and is allowed since $l' = 12$ is tabulated as an allowable candidate (Table 9) for $l = 8$. Suppose the value of $\boldsymbol{\lambda}$ is $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_4)$. The 2 unique parameters are $\tilde{\lambda}_1 = \lambda_1 = \lambda_4$ and $\tilde{\lambda}_2 = \lambda_2 = \lambda_3$. Here $i_1 = 1$ and $i_2 = 3$, which means we propose a move to $\tilde{\lambda}'_1, \dots, \tilde{\lambda}'_3$, by setting $\tilde{\lambda}'_1 = \tilde{\lambda}_1 u$, $\tilde{\lambda}'_2 = \tilde{\lambda}_2$ and $\tilde{\lambda}'_3 = \frac{\tilde{\lambda}_1}{u}$. This gives $\boldsymbol{\lambda}' = (\lambda'_1, \dots, \lambda'_4) = (\tilde{\lambda}'_1, \tilde{\lambda}'_2, \tilde{\lambda}'_2, \tilde{\lambda}'_3)$.

Calculating the acceptance probability of a birth move is straightforward once the dy-

namics of the reverse **death** move are considered. To do so, we suppose now that the chain is in model $m' = (l', k)$, with parameters $\boldsymbol{\lambda}'$ and $\boldsymbol{\kappa}$. Having chosen a **death** move (with probability $d_{l'}$) we propose a new mean structure l , at random from the r_{l', q_l} possible candidates. The proposed mean parameter $\boldsymbol{\lambda}$ is then given by

$$\tilde{\lambda}_{i_1} = \sqrt{\tilde{\lambda}'_{i_1} \tilde{\lambda}'_{i_2}}, \text{ and } \tilde{\lambda}_j = \begin{cases} \tilde{\lambda}'_j & j \neq i_1, j < i_2 \\ \tilde{\lambda}'_{j-1} & j \neq i_1, j \geq i_2, \end{cases}$$

and the dummy random variable u is given by

$$u = \sqrt{\frac{\tilde{\lambda}'_{i_1}}{\tilde{\lambda}'_{i_2}}},$$

where as above, i_1 and i_2 are apparent from the new and existing mean structures l and l' .

Following the methods prescribed in Green (2003) it follows that the acceptance probabilities for the **birth** move and reverse **death** moves are given by $\min\{1, A\}$ and $\min\{1, A^{-1}\}$ respectively, where

$$A = \frac{\mathcal{L}(\mathbf{x}|m', \boldsymbol{\lambda}', \boldsymbol{\kappa})p(m', \boldsymbol{\lambda}', \boldsymbol{\kappa})}{\mathcal{L}(\mathbf{x}|m, \boldsymbol{\lambda}, \boldsymbol{\kappa})p(m, \boldsymbol{\lambda}, \boldsymbol{\kappa})} \times \frac{d_{l'} r_{l', q_l}}{b_l r_{l', q_l}} \frac{1}{g_{\sigma_\lambda}(u)} \times |J| \quad (4)$$

Here the first factor contains the ratio of posteriors, comprising the ratio of likelihoods \mathcal{L} and the ratio of priors which depends upon whether prior A or prior B is used. The second factor is the ratio of the proposal distributions, where g_{σ_λ} is the probability density function of the LogNormal(0, σ_λ) distribution. The final factor is the absolute value of the Jacobian J , which simplifies to

$$J = \begin{vmatrix} \frac{\partial \tilde{\lambda}'_{i_1}}{\partial \lambda_{i_1}} & \frac{\partial \tilde{\lambda}'_{i_2}}{\partial \lambda_{i_1}} \\ \frac{\partial \tilde{\lambda}'_{i_1}}{\partial u} & \frac{\partial \tilde{\lambda}'_{i_2}}{\partial u} \end{vmatrix} = \begin{vmatrix} u & \frac{1}{u} \\ \tilde{\lambda}_{i_1} & -\frac{\tilde{\lambda}_{i_1}}{u^2} \end{vmatrix} = -\frac{2\tilde{\lambda}_{i_1}}{u}.$$

In addition to the move detailed above, at each sweep of our sampler we include a **birth/death** move pair to update the dispersion structure k while keeping l fixed. The dynamics of this move are similar to the above and we omit details. In order to improve the mixing properties of our sampler, we also include a second model changing move type. Again, this move type, which we call a **switch** move, either proposes a new mean structure l while retaining the current dispersion index k , or vice versa. At each sweep, an attempt is made to update both l and k in this manner. Unlike the **birth/death** move type, rather than proposing a model with one more or one less free parameters, the **switch** move proposes a model with the same dimension. The new proposed values $\tilde{\boldsymbol{\lambda}}'$ (or $\tilde{\boldsymbol{\kappa}}'$) for the underlying parameters are then drawn from independent Log-Normal distributions, independently of

the current values $\tilde{\lambda}$ and $\tilde{\kappa}$. The acceptance probability is then easily derived following standard methods. Further details of the reversible jump moves along with all aspects of the MCMC sampler can be found in Chapter 3 of Hastie, 2004.

Beyond model changing moves, our sampler also includes standard Metropolis-Hastings moves to update each of the free parameters in λ and κ in turn. For the component being updated, the proposed new value is the product of the current value and a random Log-Normal increment, and is accepted with probability derived in the usual fashion.

For both priors we run the MCMC for 1,100,000 sweeps and discard the first 100,000 as burn-in. We sub-sample every 100 observations leaving 10,000 observations. Our sampler typically gives trans-dimensional acceptance rates of between 6% and 12% for prior A and 5% and 8% for prior B , depending on the move type. In both cases fixed dimensional acceptance rates vary between 10% and 47%.

Although some tuning has been carried out, the sampler has not been optimised and more desirable acceptance rates could no doubt be achieved. However, we believe that with the current parameter settings the mixing is adequate. This is supported by the trace plots in Figure 1 and the fact that multiple MCMC runs started from random initial states produce similar numerical results.

Reported posterior probabilities are averages across 2 independent MCMC runs. In all instances we estimate the Monte Carlo standard error to be less than 0.005.

Table 1: Data from the Haigis-Dove study. Right-most columns are sample means and variances.

Group	Tumor Multiplicities	Mean	Var.
+/+	80,103,112,121,121,121,131,140,140,150,166,169,194,199,199,262	150.5	2102.1
Rb9 trans	5, 7, 8, 8, 9, 9, 11, 12, 12, 13, 13, 13, 14, 15, 15, 16, 18	11.6	12.5
Rb9 cis	7, 7, 7, 8, 8, 8, 10, 10, 10, 10, 11, 11, 12, 12, 20	10.1	10.6
Rb9/Rb9	3, 4, 4, 5, 6, 6, 6, 6, 7, 7, 7, 9, 10, 10, 11, 11, 12, 15	7.7	10.3

Table 2: Possible patterns of equality among components of λ across the 4 groups. Within each row, two entries are equal if and only if the corresponding group mean values are equal. The first column indexes the pattern and the last column gives the number of free parameters.

index l	l pattern				dimension q_l
	$i = 1$	$i = 2$	$i = 3$	$i = 4$	
1	1	1	1	1	1
2	1	1	1	2	2
3	1	1	2	1	2
4	1	2	1	1	2
5	1	2	2	2	2
6	1	1	2	2	2
7	1	2	1	2	2
8	1	2	2	1	2
9	1	1	2	3	3
10	1	2	1	3	3
11	1	2	3	1	3
12	1	2	2	3	3
13	1	2	3	2	3
14	1	2	3	3	3
15	1	2	3	4	4

Table 3: P-values from exact conditional test of the Poisson assumption. The Monte Carlo test used $10^4 - 1$ simulated data sets and used the sample coefficients of dispersion as test statistics.

+/+	Rb9 trans	Rb9 cis	Rb9/Rb9
0.0001	0.37	0.38	0.15

Table 4: Drinkwater approach. Shown are results of the likelihood ratio test for common κ . In each case the maximum likelihood estimate of the group-specific λ is the sample mean (see Table 1). Shown here are MLEs of the group-specific κ values on the alternative hypothesis and the common estimate on the null hypothesis. The contribution to the log-likelihood is L for the unrestricted case and L_0 for the null. The generalized likelihood ratio statistic is $4.1 = 2 \times (212.55 - 210.50)$.

	Alternative		Null	
	$\hat{\kappa}$	L	$\hat{\kappa}$	L_0
+/+	0.076	-82.52	0.054	-82.94
Rb9 trans	0.001	-45.21	0.054	-45.99
Rb9 cis	0.000	-37.43	0.054	-38.21
Rb9/Rb9	0.033	-45.34	0.054	-45.41
Total		-210.50		-212.55

Table 5: Model choice. For a collection of 11 interesting sub-models (out of 780), shown are the sub-model index m and its components l , the mean-pattern index, and k , the dispersion-pattern index. The l -pattern is as in Table 2. The k -pattern is similar except that 0 indicates $\kappa_i = 0$ (Poisson variation). L is maximized loglikelihood, q is number of free parameters, $\text{BIC}(r)$ and $\text{AIC}(r)$ are, respectively, the rank of the sub-model in terms of the BIC and the AIC.

sub-model	l pattern				k pattern				L	q	$\text{BIC}(r)$	$\text{AIC}(r)$		
$m = (l, k)$	$i =$	1	2	3	4	$i =$	1	2	3	4				
(12,9)		1	2	2	3		1	0	0	1	-211.62	4	1	1
(12,5)		1	2	2	3		1	0	0	0	-211.68	4	2	2
(12,13)		1	2	2	3		1	0	1	1	-212.43	4	3	6
(12,14)		1	2	2	3		1	1	0	1	-212.46	4	4	7
(15,9)		1	2	3	4		1	0	0	1	-210.71	5	8	3
(15,5)		1	2	3	4		1	0	0	0	-210.76	5	9	4
(12,20)		1	2	2	3		1	0	0	2	-211.42	5	14	5
(14,9)		1	2	3	3		1	0	0	1	-212.80	4	7	31
(14,5)		1	2	3	3		1	0	0	0	-213.31	4	11	35
(15,16)		1	2	3	4		1	1	1	1	-212.55	5	40	52
(15,52)		1	2	3	4		1	2	3	4	-210.50	8	160	91

Table 6: Full Bayesian approach. Shown are mean patterns l with non-zero marginal posterior probabilities (see Table 2).

Index	l pattern				Posterior probability			
	l	$i =$	1	2	3	4	prior A	prior B
5			1	2	2	2	0.09	0.07
12			1	2	2	3	0.57	0.59
13			1	2	3	2	0.01	0.01
14			1	2	3	3	0.19	0.18
15			1	2	3	4	0.15	0.15

Table 7: Full Bayesian approach. Shown are posterior probabilities and rankings for 11 sub-models containing the 6 most highly ranked sub-models for both priors.

sub-model index	l pattern				k pattern				posterior probability		ranking				
	$m = (l, k)$	$i =$	1	2	3	4	$i =$	1	2	3	4	prior A	prior B	prior A	prior B
(12 , 5)			1	2	2	3		1	0	0	0	0.09	0.23	1	1
(12 , 9)			1	2	2	3		1	0	0	1	0.09	0.11	2	2
(14 , 9)			1	2	3	3		1	0	0	1	0.04	0.05	3	5
(15 , 5)			1	2	3	4		1	0	0	0	0.03	0.06	8	3
(12 , 20)			1	2	2	3		1	0	0	2	0.04	0.04	5	6
(14 , 5)			1	2	3	3		1	0	0	0	0.02	0.05	11	4
(12 , 13)			1	2	2	3		1	0	1	1	0.04	0.02	4	12
(15 , 9)			1	2	3	4		1	0	0	1	0.03	0.03	9	9
(12 , 14)			1	2	2	3		1	1	0	1	0.03	0.02	6	15
(15 , 16)			1	2	3	4		1	1	1	1	0.00	0.00	65	90
(15 , 52)			1	2	3	4		1	2	3	4	0.00	0.00	162	—

Table 8: Bayesian approach. Shown are posterior probabilities that tumor multiplicities follow a Poisson distribution for each group. Recall that prior B has $P(\kappa_i = 0) = 1/2$ and in other cases $P(\kappa_i = 0) = 15/52$.

Method	$P(\kappa_i = 0 \text{data})$			
	+/+	Rb9 trans	Rb9 cis	Rb9/Rb9
BIC-approx	0.00	0.58	0.57	0.37
Full, prior A	0.00	0.61	0.60	0.32
Full, prior B	0.00	0.80	0.80	0.53

Table 9: MCMC. Shown are the values of the mean-pattern index l' that can be proposed when the current mean-pattern index is l (labels as in Table 2).

Current mean structure l	Permitted birth proposals l'	Permitted death proposals l'	Permitted switch proposals l'
1	2, 3, 4, 5, 6, 7, 8	-	-
2	9, 10, 12	1	3, 4, 5, 6, 7, 8
3	9, 11, 13	1	2, 4, 5, 6, 7, 8
4	10, 11, 14	1	2, 3, 5, 6, 7, 8
5	12, 13, 14	1	2, 3, 4, 6, 7, 8
6	9, 14	1	2, 3, 4, 5, 7, 8
7	10, 13	1	2, 3, 4, 5, 6, 8
8	11, 12	1	2, 3, 4, 5, 6, 7
9	15	2, 3, 6	10, 11, 12, 13, 14
10	15	2, 4, 7	9, 11, 12, 13, 14
11	15	3, 4, 8	9, 10, 12, 13, 14
12	15	2, 5, 8	9, 10, 11, 13, 14
13	15	3, 5, 7	9, 10, 11, 12, 14
14	15	4, 5, 6	9, 10, 11, 12, 13
15	-	9, 10, 11, 12, 13, 14	-

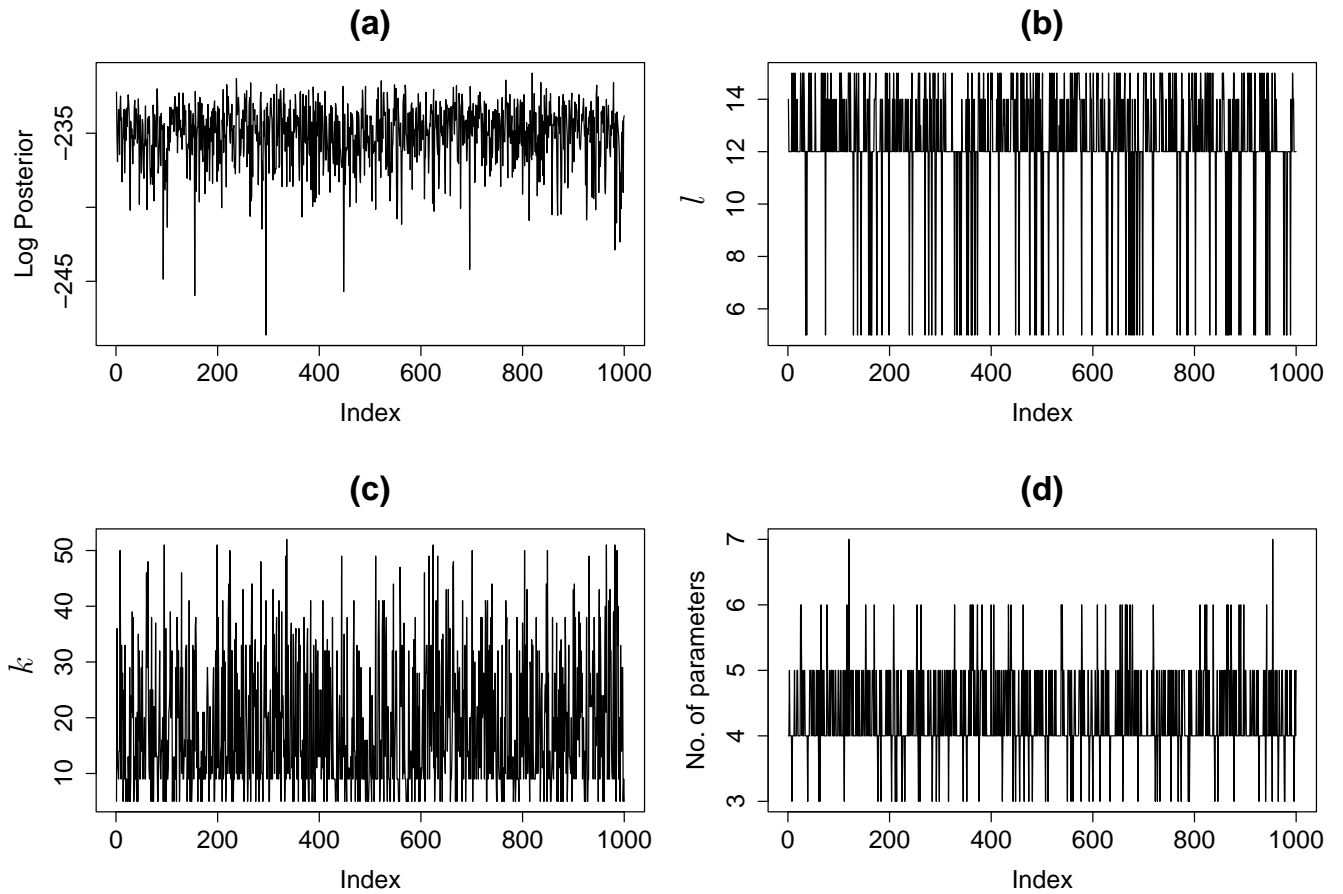


Figure 1: Trace plots from MCMC runs (chains thinned to 1000 observations): (a) log-posterior value (prior A); (b) mean-pattern index l (prior B); (c) dispersion-pattern index k (prior A) and (d) number of free parameters, $q_l + q_k$ (prior B).

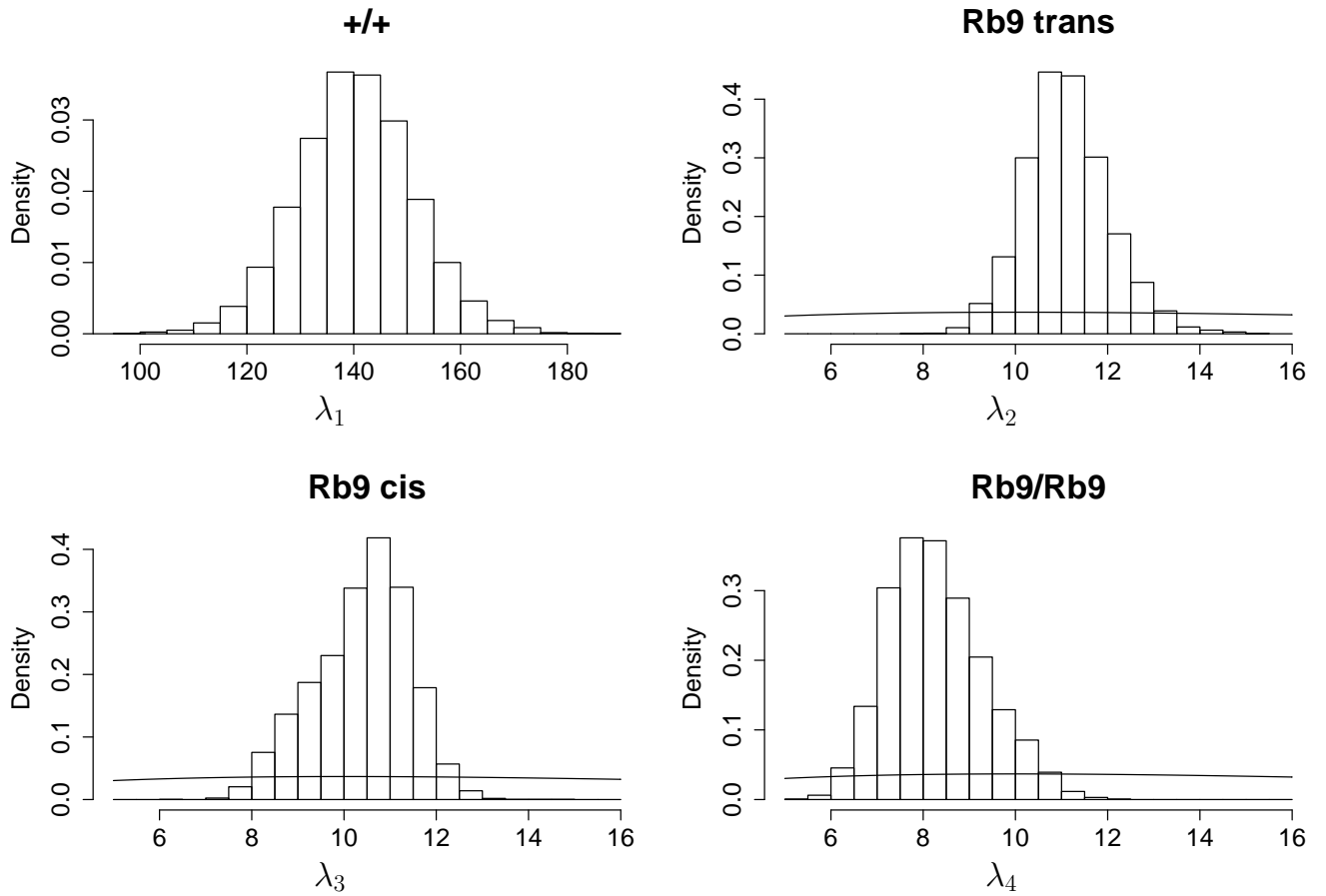


Figure 2: Histograms of marginal posterior samples of the components of λ (prior A): Solid lines denote the prior.

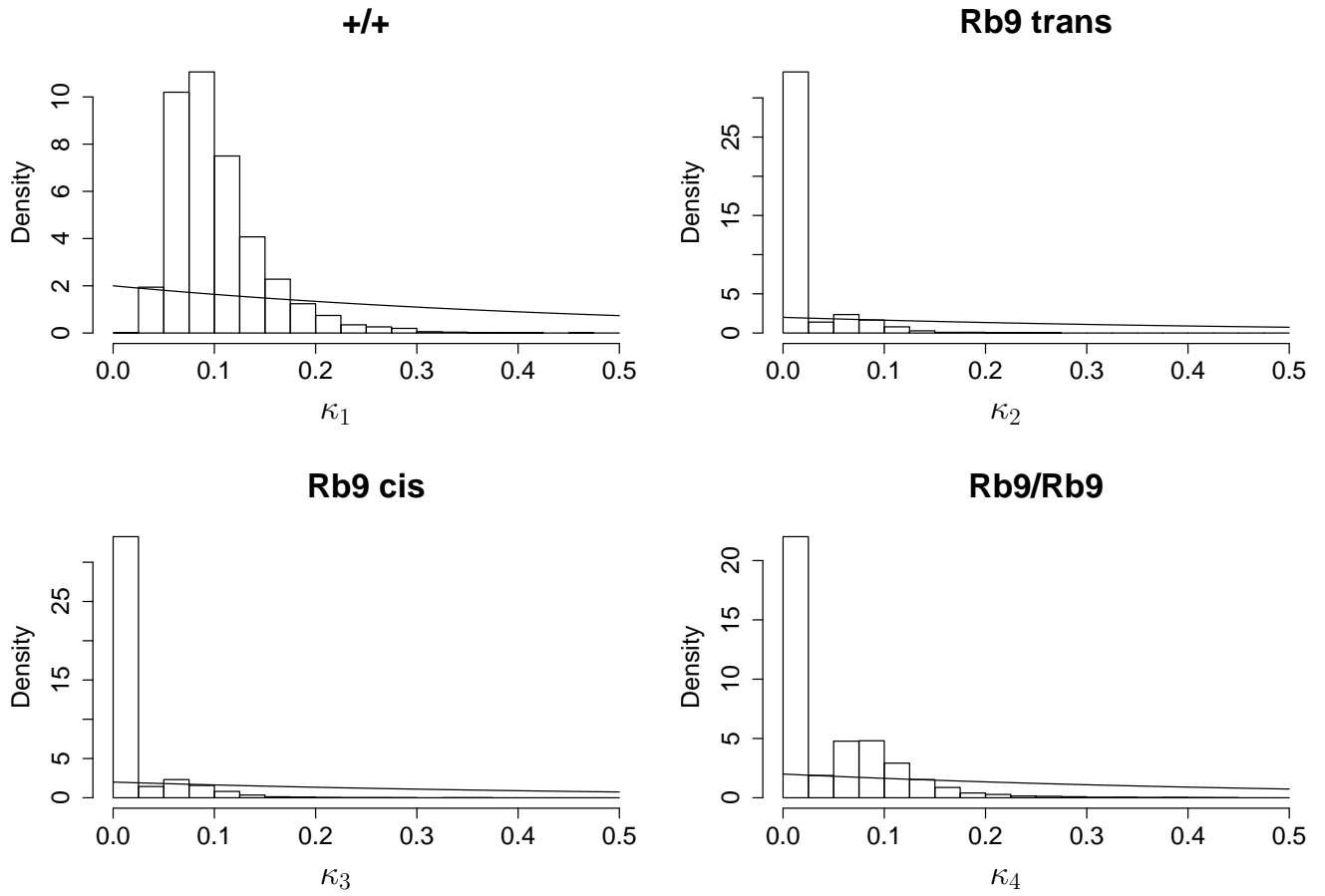


Figure 3: Histograms of marginal posterior samples of the components of κ (prior B): Solid lines denote the prior.

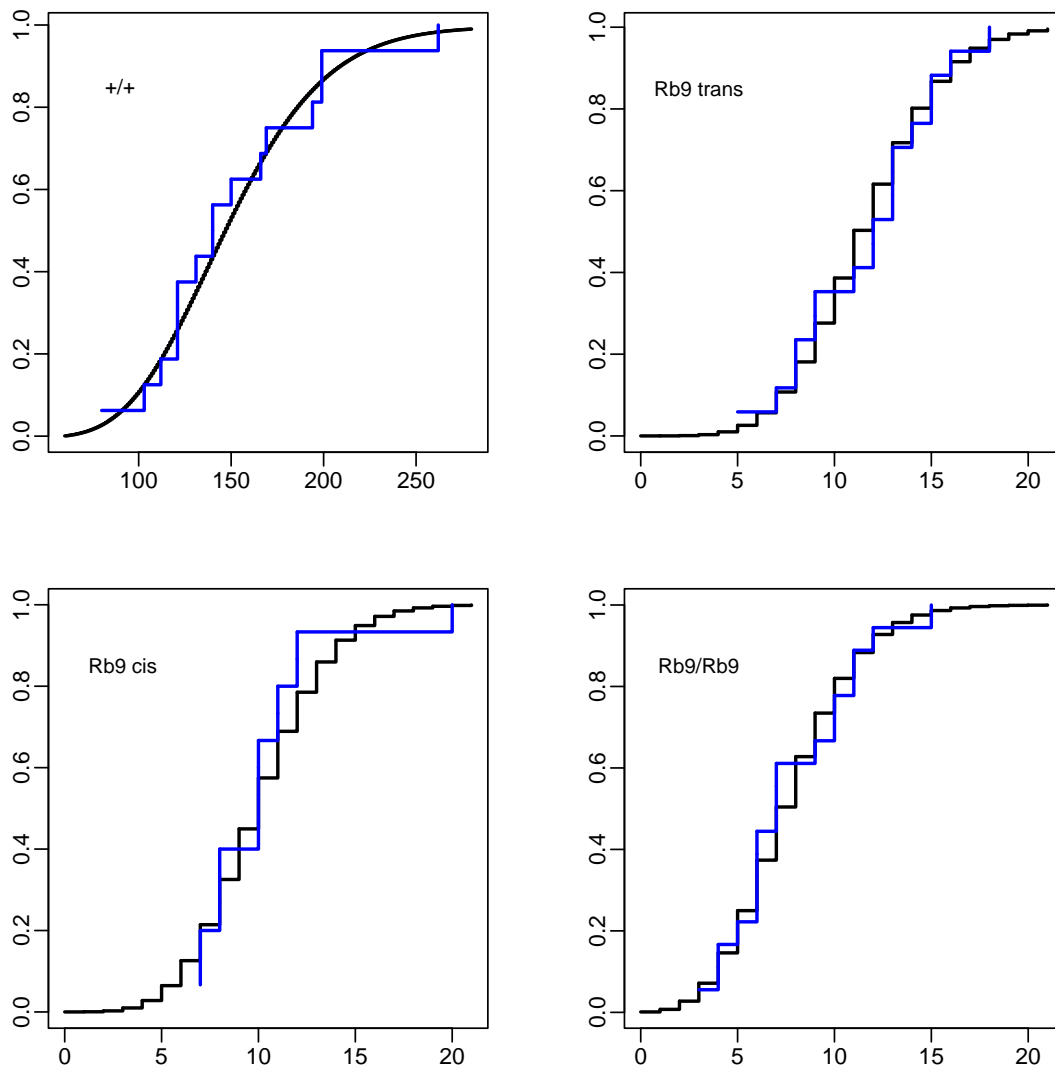


Figure 4: Empirical distributions (blue) of tumor counts from Table 1 and best fitting negative binomial distributions (black): Fitting was done by maximum likelihood separately in each group.

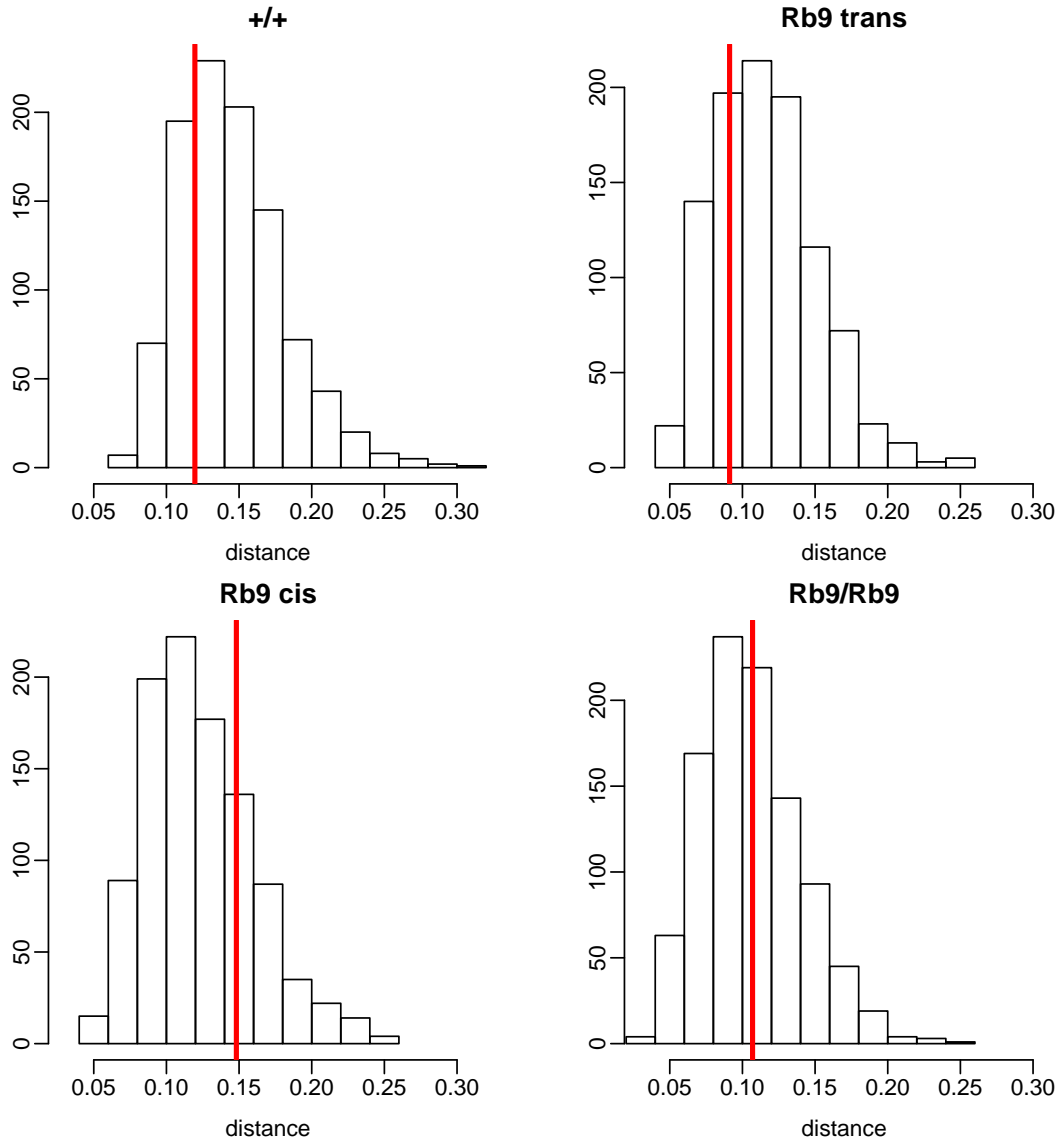


Figure 5: Posterior predictive checks using as test statistic the maximum difference between predicted-sample empirical and best fitting negative binomial, as in Figure 4: Red vertical lines indicate values for observed data, and thus we have no evidence against the general negative binomial model.