A stochastic model for cancer risk

Rinaldo B. Schinazi

Department of Mathematics

University of Colorado at Colorado Springs, CO 80933-7150, USA.

Running head: Cancer risk

Keywords: cancer risk, mathematical model, branching process, mutation

Corresponding author: Rinaldo B. Schinazi, Department of Mathematics, U.Colorado at Colorado Springs, 1420 Austin Bluffs, Colorado Springs CO 80933-7150, USA.

Phone: 719 262 3920

Fax: 719 262 3605

email: rschinaz@uccs.edu

ABSTRACT

We propose a simple stochastic model based on the two successive mutations hypothesis to compute cancer risks. Assume that only stem cells are susceptible to the first mutation and that there is a total of D stem cell divisions over the life time of the tissue with a first mutation probability μ_1 per division. Our model predicts that cancer risk will be low if $m = \mu_1 D$ is low even in the case of very advantageous mutations. Moreover, if $\mu_1 D$ is low the mutation probability of the second mutation is practically irrelevant to the cancer risk. These results are in contrast with existing models but in agreement with a conjecture of Cairns. In the case where m is large our model predicts that the cancer risk depends crucially on whether the first mutation is advantageous or not. A disadvantageous or neutral mutation makes the risk of cancer drop dramatically.

Cancer has long been thought as being provoked by successive somatic mutations, see Knudson (2001) for a history of this hypothesis

and Barrett (1986) for supporting experimental evidence. For instance, the age incidence of retinoblastoma is consistent with a two mutations scenario (Hethcote and Knudson (1978)) while the age incidence of colorectal cancer is consistent with five successive mutations (Knudson (2001)). However, while there seems to be general agreement that cancer has an early and a late stage there are biological doubts on the exact number of stages. Armitage (1985) argues that two-stages models explain several cancers.

In this paper we propose a simple stochastic model based on the two successive mutations hypothesis to compute cancer risks.

THE MODEL

We assume that cells susceptible to the first mutation undergo a fixed number D of divisions over the lifetime of the tissue and that there is a probability μ_1 per division of producing a cell with a type 1 mutation. Assuming that all divisions are stochastically independent, the (random) number of type 1 mutation cells produced over the lifetime of the tissue is Poisson distributed with mean $m = D\mu_1$. Once a type 1 cell appears it starts a branching process. More precisely, the process starts with a single type 1 cell and after a unit time it may die with probability $1 - p_1$ or divide in two type 1 cells with probability p_1. Successive generations of type 1 cells follow the same rules independently from each other. Hence, the mean number of daughter cells per cell is $2p_1$. It is well known, see for instance Section I.9 in Schinazi (1999), that the branching process survives forever with positive probability if and only if $p_1 > 1/2$.

We also assume that at each division of a type 1 cell there is a probability μ_2 for each daughter cell that it be a type 2 cell. The probability, denoted by $S(p_1, \mu_2)$, that a branching process started by a single type 1 cell eventually gives birth to at least one type 2 cell may be computed exactly (see Schinazi (2006)):

$$S(p_1, \mu_2) = 1 - \frac{1}{2p_1(1 - \mu_2)^2}\Big(1 - \sqrt{1 - 4p_1(1 - p_1)(1 - \mu_2)^2}\Big).$$

Assuming that branching processes started by different type 1 cells are independent of each other we get that the number of type 1 branching processes that eventually give birth to at least one 2 cell is also Poisson distributed with mean $D\mu_1 S(p_1, \mu_2) = mS(p_1, \mu_2)$. Hence, the probability of cancer over the lifetime of a particular tissue is

$$R(m, p_1, \mu_2) = 1 - \exp(-mS(p_1, \mu_2)).$$

As Figure 1 illustrates the parameter m (the mean number of first mutations over the lifetime of the tissue) has a dramatic effect on R. For $m = 0.01$ the risk of cancer is below 1% even for values of p_1 near 1, that is, even if the first mutation is extremely advantageous. On the other hand the parameter μ_2 is almost irrelevant for low values of m, see Figure 2.

If m is large then the crucial parameter is p_1. A neutral or slightly disadvantageous first mutation (that is, $p_1 \leq 0.5$) lowers the risk of cancer dramatically, see Figure 3. For $p_1 \leq 0.5$ the parameter μ_2 becomes important (as can be checked by direct computation).

DISCUSSION

Our model shows that even if mutated cells multiply exponentially (as they do in a branching process with $p > 1/2$) a two mutations cancer has a low risk provided $m = \mu_1 D$ is small. Moreover, for small m the risk of cancer, in this model, does not depend on the second mutation probability μ_2.

If we assume that only stem cells are susceptible to the first mutation then $m = \mu_1 D$ represents the mean of first mutations for D divisions of stem cells. Our model predicts that a good strategy to prevent cancer is a low m. For low m the risk of cancer is low even if the first mutation is advantageous (i.e. $p_1 > 1/2$) and if the second mutation rates μ_2 is high. This is consistent with the picture of Cairns (2002) regarding carcinogenesis. In particular, he conjectures that stem cell mutation rates are low and that this affords protection against cancer.

In the case where m is large the cancer risk depends crucially on whether the first mutation is advantageous or not. If the first mutation is disadvantageous or neutral the risk of cancer drop dramatically.

Branching processes, such as the ones used here, have long been used in biology. For a recent example, see for instance Johnson and Barton (2002). Mathematical modeling of the successive mutations hypothesis goes back to at least Armitage and Doll (1954) who proposed a pure birth process model. They successfully modeled the increase in the number of cancers as a power law of age. This power law is observed in a number of countries and for a number of cancers. In contrast, the model proposed here is suitable to test hypotheses at the cell level rather than at the population level.

There are at least two recent papers that propose mathematical models for the role of stem cells in the appearance of cancer. Frank et al. (2003) assume that each cell (mutated or not) must undergo a certain deterministic number of divisions before being killed. They find the number of divisions for stem cells and transit cells that minimize the risk of cancer. For their model, unlike ours, the mutation rate for stem cells seems to have little influence on the risk of cancer (see Figure 3 in their paper). Michor et al. (2003) are interested in the same type of question with the particular goal of finding the proportion of stem cells in a tissue that minimizes the risk of cancer. There too the mutation rate of stem cells seems to have a role comparable to the subsequent mutation rates, see (2.6) and (2.7) in Michor et al. (2003). Hence, ours seems to be the first model that predicts a possible preponderant role for the mutation rate of stem cells.

LITERATURE CITED

Armitage P. (1985) Multistage models of carcimogenesis. Environmental Health Perspectives **63**, 195-201.

Armitage P. and Doll R. (1954) The age distribution of cancer and a multistage theory of carcinogenesis. British Journal of cancer **8:** 1-12.

Barrett J.C. (1986) Cell culture models of multistep carcinogenesis. In: *Age related factors in carcinogenesis* (ed. by A. Likhachev, V. Anisimov and R. Montesano), 181-202. International agency for research on cancer scientific publications no. 58, Lyon, France.

Cairns J. (2002) Somatic stem cells and the kinetics of mutagenesis and carcinogenesis. Proceedings of the National Academy of Sciences **99:** 10567-10570.

Frank S.A., Y. Iwasa and M.A. Nowak (2003) Patterns of Cell Division and risk of cancer. Genetics **163:** 1527-1532.

Hethcote H.W. and A.G. Knudson Jr (1978) Model for the incidence of embryonal cancers: Application to retinoblastoma. Proceedings of the National Academy of Sciences **75:** 2453-2457.

Johnson T. and Barton N.H. (2002) The Effect of Deleterious Alleles on Adaptation in Asexual Populations **162:** 395-411.

A.G.Knudson (2001) Two genetic hits (more or less) to cancer. Nature reviews Cancer **1:** 157-162.

Michor F., M.A. Nowak, S.A. Frank and Y. Iwasa (2003) Stochastic elimination of cancer cells. Proceedings of the Royal Society of London B **270**, 2017-2024.

Schinazi R.B. (2006) The probability of treatment induced drug resistance. Acta Biotheoretica, in press.

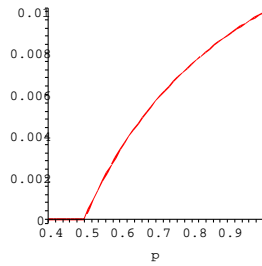Schinazi R.B. (1999) *Classical and spatial stochastic processes.* Birkhauser, Boston.

FIGURE 1. This is the plot of the risk R of a two muta-
tions cancer as a of function of p_1 for $m = \mu_1 D = 0.01$
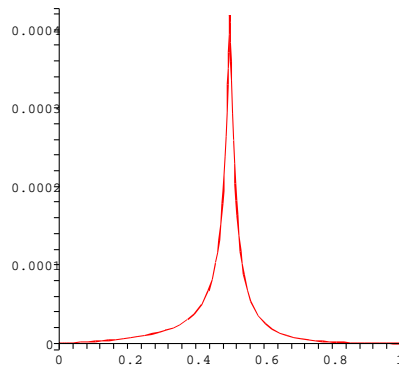and $\mu_2 = 10^{-5}$.

FIGURE 2. This plot illustrates our claim that for low
m the mutation rate μ_2 is practically irrelevant in a two
mutations cancer. We plot the difference between the
cancer risks for the models with $\mu_2 = 10^{-3}$ and $\mu_2 = 10^{-8}$
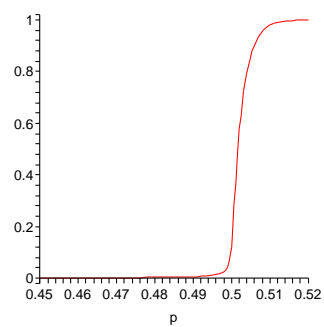as a function of p_1. For both models we set $m = \mu_1 D = 0.01$.

FIGURE 3. This is the plot of the risk R of a two mutations cancer as a of function of p_1 for $m = 100$ and $\mu_2 = 10^{-6}$.