Carcinogenesis vol.18 no.12 pp.2461–2462, 1997

LETTER TO THE EDITOR

Remarks on a model of labeling indices

Qi Zheng

Division of Biometry and Risk Assessment, National Center for Toxicological Research, Jefferson, AK 72079, USA

Dear Sir,

Moolgavkar and Luebeck (1) proposed a stochastic birth–death type model from which the cell division rate can be estimated using experimental labeling index (LI) data. Their paper (referred to as 'ML' in the following) made the following assumptions: (i) only continuous labeling (e.g. implantation of osmotic pumps) is considered; (ii) cell cycle phases are ignored; and (iii) the birth (division) and death rates for labeled and unlabeled cells are the same. Accepting all the above assumptions, this letter points out an error in the ML paper.

Following ML, let $X(t)$ denote the number of unlabeled cells at time t, and $Y(t)$ the number of labeled cells. Define $N(t) = X(t) + Y(t)$. Let the cell birth (division) and death rates be β and δ respectively for both labeled and unlabeled cells. The starting point of the ML paper is the following observation:

$$E[X(t)] = X(0)e^{-(\beta+\delta)t}, \quad [1]$$

$$E[Y(t)] = X(0)e^{(\beta-\delta)t}(1 - e^{-2\beta t}). \quad [2]$$

To help concentrate on the central issue, note that equations [1] and [2] can also be deducted from known results. By definition, the unlabeled cells will remain so (alive and unlabeled) by time t with probability $e^{-(\beta+\delta)t}$ and, hence, $X(t) \sim$ binomial $[X(0),e^{-(\beta+\delta)t}]$, or $E[X(t)] = X(0)e^{-(\beta+\delta)t}$. On the other hand, the process $N(t)$ is equivalent to a birth–death process with birth and death rates β and δ, respectively. So $E[N(t)] = X(0)e^{-(\beta+\delta)t}$ (e.g. Bailey (2), p. 94) and, hence, the expression for $E[Y(t)]$. Based on the fact that $E[Y(t)]/E[N(t)] = 1 - e^{-2\beta t}$, ML postulated that $Y(t) \sim$ binomial $[N(t), 1 - e^{-2\beta t}]$ and, consequently, the log likelihood function would be

$$\sum_{i=1}^{n} [Y(t_i) \log(1 - e^{-2\beta t_i}) - 2\beta t_i X(t_i)] \quad [3]$$

where $Y(t_i)$ and $X(t_i)$ are observed labeled and unlabeled cell counts at time t_i ($i = 1,\ldots,n$).

In fact, not only is the distribution of $Y(t)$ elusive, but even efforts to seek an analytic expression for $E[Y(t)/N(t)]$ seem futile. Fortunately, resorting to Taylor series approximation (first term) yields

$$E\left[\frac{Y(t)}{N(t)}\right] \top \frac{E[Y(t)]}{E[N(t)]} = 1 - e^{-2\beta t}, \quad [4]$$

which coincides with the 'binomial' mean given by the ML paper. Were $X(t)$ and $Y(t)$ independent processes, equation [4] should be exact. Furthermore, the following relation is easy to derive

$$\mathrm{Cov}[X(t),Y(t)] = X(0)e^{-2\delta t}(e^{-2\beta t} - 1). \quad [5]$$

Table I. Least squares fitting of the original data

Times (days)	Controls: $\hat{\beta} = 0.01370$		Treated: $\hat{\beta} = 0.02329$	
	Measured LI	Predicted LI	Measured LI	Predicted LI
$t_1 = 2$	8.71	5.33	13.42	8.90
$t_1 = 7$	17.31	17.45	30.94	27.82
$t_1 = 14$	31.26	31.85	44.62	47.91

Equation [5] reveals that $|\mathrm{Cov}[X(t),Y(t)]| \leqslant X(0)e^{-2\delta t}$. That is, $\mathrm{Cov}[X(t),Y(t)]$ is bounded for $t \in [0,\infty)$ and decays to zero exponentially fast if the cell death rate δ is positive. Therefore, one can conjecture that $1 - e^{-2\beta t}$ may approximate $E[Y(t)/N(t)]$ reasonably well. One can, hence, adopt the classic nonlinear least squares technique to fit experimental data. Specifically, the following function is to be minimized in estimating β:

$$\sum_{i=1}^{n} \left\{ \frac{Y(t_i)}{X(t_i)} - (1 - e^{-2\beta t_i}) \right\}^2. \quad [6]$$

Table I provides results using the least squares technique. Although the least squares approach shows some minor improvement (in terms of squared deviations), I am not jumping at any conclusions in this regard based solely on one or two small data sets. The focus is theoretical foundation.

Although equation [3] is not the log likelihood function that the ML paper sought, why does maximizing it yield an estimate for β close to that produced by minimizing equation [6]? With large $X(0)$, $Y(t)$ is approximately normal with mean $E[Y(t)]$ as given in [2]. On the other hand, a binomial $[X(0)e^{(\beta-\delta)t}, 1–e^{-2\beta t}]$ random variable can also be approximated by a normal distribution with the same mean. Because both $X(0)$ and δ are nuisance parameters, the quantity $X(0)e^{(\beta-\delta)t}$ may be approximated by $N(t)$ on account of the fact that $E[N(t)] = X(0)e^{(\beta-\delta)t}$.

There exists another perspective. With large initial unlabeled cell pool, a deterministic model should be a reasonable approximation:

$$\frac{dX}{dt} = -(\beta + \delta)X,$$

$$\frac{dY}{dt} = (\beta - \delta)Y + 2\beta X.$$

Solving, one has

$$X(t) = X(0)e^{-(\beta+\delta)t},$$

$$Y(t) = X(0)e^{(\beta-\delta)t}(1 - e^{-2\beta t}).$$

This model also requires minimization of equation [6]. Assuming validity of assumptions (i)–(iii), it seems sensible to consider the least squares technique using equation [6] as a competitor against, if not substitute for, the likelihood approach using equation [3].

References

1. Moolgavkar,S.H. and Luebeck,E.G. (1992) Interpretation of labeling indices in the presence of cell death. *Carcinogenesis*, **13**, 1007–1010.
2. Bailey,N.T.J. (1964) *The Elements of Stochastic Processes with Applications to the Natural Sciences.* John Wiley & Sons, New York.

Received on November 18, 1996; accepted on August 14, 1997

Response

S.H.Moolgavkar and E.G.Luebeck

Dear Sir,

We are grateful to Dr Zheng for pointing out the error in our paper (1). We should have appealed to the central limit theorem and used non-linear least squares to estimate the parameters of the model. As Dr Zheng points out, however, whether the parameters are estimated via least squares or the likelihood based on the binomial distribution makes little difference to the results.

We would like to take this opportunity to discuss briefly some practical considerations in the interpretation of labeling indices. These arise from the fact that labeling indices in tissues are estimated from counts of labeled cells in two-dimensional sections through the tissues. The attendant stereological problems can make the interpretation of the counts difficult. For example, there is considerable interest in estimating rates of cell division in enzyme-altered foci arising in hepatocarcinogenesis experiments in rodents. There is evidence that cells in foci exhibit a positional gradient in cell proliferation rates, with cells near the surface dividing preferentially (2,3). For spherical foci, the ratio of surface area to volume decreases with increasing focal volume. Thus, the fraction of dividing cells decreases with increasing focal volume. A transection near the cap of a large spherical focus may show the same labeling index as a transection near the equator of smaller focus although the cells in the smaller focus may have, on average, a higher division rate than cells in the larger focus. Furthermore, a transection through the cap of a focus will have higher labeling index than a transection near the equator of the same focus. As a consequence, the labeling index measured on a two-dimensional section of a focus cannot be used to estimate the labeling index for the focus as a whole.

References

1. Moolgavkar,S.H. and Luebeck,E.G. (1992) Interpretation of labeling indices in the presence of cell death. *Carcinogenesis*, **13**, 1007–1010.
2. Buchmann,A., Stinchcombe,S., Kurner,W., Hagenmaier,H.P. and Bock,K.W. (1994) Effects of 2,3,7,8-tetrachloro- and 1,2,3,4,6,7,8-heptachlorodibenzo-p-dioxin on the proliferation of preneoplastic liver cells in the rat. *Carcinogenesis*, **15**, 1143–1150.
3. Pugh,P.D. and Goldfarb,S. (1992) Growth kinetics of microscopic hepatocellular neoplasms in carcinogen-resistant and carcinogen-responsive strains of mice. *Cancer Res*., **52**, 280–284.