

Age distribution of cancer in mice: the incidence turnover at old age

FRANCESCO POMPEI,^a MICHAEL POLKANOV^b AND RICHARD WILSON^b

^aDivision of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

^bDepartment of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

We have studied cancer incidence in mice as a function of age in those cohorts where the rodents are allowed to live very close to their full natural lifetime. We find that the incidence rises as a function of age, but then flattens and turns over at an age of about 800 days. This behaviour is similar to that which we observed (Pompei and Wilson, 2001) in the Surveillance, Epidemiology, and End Results (SEER) data where the age distribution of human cancer incidence turns over at about age 80. Although other fits are possible, the three-parameter beta function model fits both the mouse data and the human data well. The beta model implies, and the data do not deny, the interpretation that cancer is not a certainty and mice may also outlive their cancers, although high-dose cohort results suggest cancer might be certain if dose is sufficiently high. Limited data suggest that the cancer age distribution, including the turnover, may be time shifted by dietary restriction. *Toxicology and Industrial Health* 2001; 17, 7–16.

Key words: age; cancer cell senescence; cancer incidence; cancer model; dietary restriction

Introduction

In another paper (Pompei and Wilson, 2001) we discuss data on the incidence of cancer in humans as a function of age. These are primarily the Surveillance, Epidemiology, and End Results (SEER) data of the National Cancer Institute (Ries et al., 2000) but we also studied two cohorts from Hong Kong (Parkin et al., 1997) and from the Netherlands (de Rijke et al., 2000). We point out that the incidence increases steadily with age up to age 70, appears to flatten off for all cancers at about age 80, and falls thereafter. The age of maximum incidence is remarkably consistent for all adult cancers, considering incidences vary over two orders of magnitude. We point out that two well-known theories of cancer, the multistage model developed by Armitage and Doll (1954) and Armitage (1985) and the two-stage clonal expansion model discussed by Armitage and Doll (1957) and in more detail by Moolgavkar (1978) and Moolgavkar and Knudsen (1981) can easily be modified to allow a flattening of the age–incidence curve but cannot be easily modified to allow a turn over at high ages. The data, taken at face value, imply an interpretation different from the traditional view that ‘if you don’t die of anything else you will die of cancer’ may not be accurate and might be replaced by ‘if you live to 90, you will have beaten cancer’.

One explanation of the turnover data observed at old age was suggested to us by Sir Richard Sir Richard Doll (2001) who emphasized that at older ages, records of cancer are less reliable, since attending physicians often used the nebulous cause of death ‘old age’. This, indeed was a stated reason that in their seminal work Armitage and Doll (1954) stopped their analysis at age 74. While 50 years later the collectors of the SEER data claim that their data are more reliable, other verifications of the turnover seem highly desirable. In support of his suggested explanation, Sir Richard pointed out that all or most members of several cohorts of persons occupationally exposed to high levels of certain pollutants died of cancer (e.g., β -naphthylamine, Case, 1966). We note that while pathologists who examine the data on cancers in rodents may have biases, they are unlikely to have Doll’s suggested bias. We are therefore in the process of examining all data sets of cancers in animals to see whether an age turnover of incidence is present.

In many bioassays, including most of the bioassays of the National Toxicology Program (NTP), animals are killed before the end of the full normal lifetime in a ‘terminal sacrifice’ which makes most of these bioassays unsuited for this study. A few, however, remain. More importantly, Dr R.L. Kodell of National Center for Toxicological Research kindly made available to us the original data of the ED01 study where 24,000 mice were exposed to various amounts of 2-acetylaminoflourene (2-AAF). This paper shows the age turnover in the incidence rate of cancers, particularly fatal cancers, in that bioassay. This turnover tends to substantiate the idea that the age turnover in people is

1. Address all correspondence to: Francesco Pompei, Department of Physics, Harvard University, Jefferson Laboratories, Cambridge, MA 02138, USA
E-mail: fpompei@post.harvard.edu

unlikely to be an artifact as suggested by Doll. As we point out in our paper on human cancer turnover, our conclusions depend critically on the reliability of the data at elevated age. We should also note that neither of these animal studies was designed to examine the turnover in cancer incidence with age, and thus might be subject to unknown biases.

Methods

Data sources

Initial evaluation of cancer incidence turnover in rodents was performed with the Toxicology Data Management System (TDMS) database, obtained from NTP contractor Analytical Sciences. The files contain data for mice and rat bioassays published in technical reports TR-341 through TR-491, which were issued from September 1989 to July 1999. We subsequently edited the data to remove results from dietary restriction studies, which affect the age distribution of tumours differently than *ad libitum* feeding of all other studies (Haseman, 2001). The database's control animals were searched for high-prevalence tumour rates for animals classified as 'natural death' or 'moribund sacrifice' for each 100-day age interval. The tumour rates were calculated as the number of animals with a specific tumour divided by the total animals dying by natural death in each age group. Since nearly all *ad libitum* studies were terminated at 2 years (the exception being TR-440: ozone), only data for <800 days age was included.

A separate analysis of the dietary restricted data for scopolamine hydrobromide trihydrate (TR-460, TMDS TR-445 Study No. 0512108) controls was conducted, following the same method as above. These animals were fed a restricted diet to maintain 85% of the weight of the *ad libitum* group.

Since the TMDS database lacked statistical power, particularly for older animals in any one study for direct age-specific incidence measurement, a data set for a much larger single study, the ED01 study, (Cairns, 1980) was obtained from FDA National Center for Toxicology Research (courtesy RL Kodell) and used for the bulk of the data in this work. Designed to detect the effective dose of 2-AAF required to produce 1% tumour rate, the original study included 24,192 female BALB/c mice, and 23,419 were included in the database we obtained.

Importantly, the ED01 data included cause of death from neoplasms by type of neoplasm, as a pathology entry. This made it possible to produce an objective examination of age-specific mortality caused by each type of cancer, which was conventionally calculated as: $M(t) = \% \text{ mortality} / 100 \text{ animal-days} = 100 \times [(\text{No. of animals dying of tumour in the 100 day period}) / (\text{No. of animal-days at risk in the 100 day period})]$. Additionally, the survival rate for the animals

were such that the ED01 study was extended to 33 months, compared to the more typical 24 months, thus producing significant data for older animals. The data were searched both for deaths caused by neoplasms, and for morbidity caused by neoplasms. It was noted that for the groups of dose=30, 35, 60 ppm, large numbers of animals were apparently misclassified as dead or moribund from neoplasms about a month before final termination at day 1001, but ought to have been classified as terminally sacrificed. These data are assumed to be in error, and not included in natural death or morbidity from neoplasms. All dosed animals considered were dosed continuously over their lifetimes.

Analytical methods

The method of developing age-specific mortality rate data was designed to emulate as precisely as possible the method of obtaining age-specific mortality in humans: natural deaths were tabulated with cause of death as determined by pathology. In the ED01 experiments, animal cages were examined twice daily, and those that died were immediately removed to be autopsied to establish cause of death. Moribund animals were treated in the same manner as dead animals, and listed as a separate removal category. Since the animals were obviously not treated for cancer, age-specific cancer mortality is believed to be a good estimate of age-specific cancer incidence, as it was in human cancer studies (Armitage and Doll 1954) before development of our current successful interventions for many human cancers.

Independent estimates of age-specific tumour incidence were developed with cumulative incidence data for spontaneous neoplasms in untreated groups for the ED01 study published by Sheldon and Greenman (1980). Animals were examined at the scheduled terminal sacrifice periods, and those dying of natural causes were combined with the nearest scheduled sacrifice group. The age-specific incidence was calculated as: $I(t) = [(\text{Cumulative incidence \% for the group sacrificed at age } t_2) - (\text{Cumulative incidence \% for the group sacrificed at age } t_1)] / [t_2 - t_1]$, where t is the average age at death. The cumulative incidence is defined as the proportion of animals with the tumour in the sacrifice group at age t_1, t_2, \dots, t_n . This method of calculation assumes that each group of mice is identical to the others, which is a fundamental result of the randomization of the animals and uniformity of facilities of the ED01 study design. This implies that cumulative incidence is different between age groups only due to age and random effects.

Statistical analysis of the turnover for the ED01 data was performed by comparing the mean of the incidence for a given age group to the mean of the incidence for the immediately younger age group with a standard two-sample z-test. By comparing the age-specific incidence for a

given age group to the age-specific incidence of a similar group at younger age, the test provides a method of evaluating the key elements of the age distribution statistically: are there tumour sites for which incidences clearly increase at middle age, and also clearly decrease at older age at the same sites? The test statistic is: $z = [(\text{Observed age-specific incidence of group in age range 2}) - (\text{Observed age-specific incidence of group in age range 1})] / \sqrt{[(\text{SD of group in age range 1})^2 + (\text{SD of group in age range 2})^2]}$.

The null hypothesis is that there is no difference in incidence between the two ages, with the alternate hypothesis that incidence increases or decreases as age increases, employing the one-tailed p -value as the test of significance. Significance is accepted at the $p < 0.05$ level.

Dividing the age-specific mortality data into four age groups for statistical analysis: 200–400 day, 400–600 day, 600–800 day, and 800–1001 day group, the test produces results of the form: $M(400-600) > M(200-400)$; $p = 0.01$, denoting that mortality due to tumours in the older age group exceeds that in the younger group, with probability 0.01 that the result was random. The notation $M(t)$ is used to denote age-specific mortality, $I(t)$ denotes age-specific incidence, and $CI(t)$ denotes cumulative incidence, throughout. Error bars indicate ± 1 SEM throughout.

Beta model

The beta model for fitting the mice data was used by the authors in Pompei and Wilson (2001) for age distribution of cancer data in humans including the turnover at ages > 80 . The derivation is presented in the Appendix of this work, as applied to modeling the animal data that is the subject of this study.

Since mortality removes the animals with tumours from the population at risk, the age-specific mortality data is properly interpreted as a hazard function, with the general form $h(t) = f(t) / [1 - F(t)]$, where $f(t)$ is the probability density function (pdf) of the modeled (assumed) cancer-causing mechanism, and $F(t)$ its time integral. Using the beta function as derived in the Appendix, $b(t) = (\alpha t)^{k-1} (1 - \beta t)$ as the pdf, we write the exact hazard function associated with the beta pdf as $b_h(t) = b(t) / [1 - B(t)]$, where $B(t)$ is the integral of $b(t)$. Then $b_h(t) = [(\alpha t)^{k-1} (1 - \beta t)] / [1 - (\alpha t)^k (1 - \beta t)]$; $0 \leq t \leq \beta^{-1}$, where $a = [\alpha k^{1/(k-1)}]^{(k-1)/k}$ and $b = k\beta / (k+1)$. Since $b < \beta$, then $b_h(t) \rightarrow 0$ as $b(t) \rightarrow 0$, thus predicting the identical age at zero incidence for both the hazard function and pdf, which is a critical feature of the beta model. Further, it can be shown that the shapes of the exact hazard function $b_h(t)$ and the beta function pdf $b(t)$ are the same, thus providing robustness when there is uncertainty to the composition of the animals at risk.

To compute lifetime cumulative probability of death or morbidity by cancer, the time integral of the implied pdf

$f(t)$ from the $M(t)$ hazard function data (not the beta pdf) was evaluated as: $F(t) = \int f(t) dt = 1 - \exp[-\sum M(t)]$, where the age-specific mortality are summed over assumed lifetime of 1001 days.

Results

Of 2093 tabulated combined male and female B6C3F1 mice TDMS controls removed for natural death or moribund sacrifice at age < 800 days, 621 mice had hepatocellular carcinoma and 432 had hepatocellular adenoma, the most prevalent neoplasms. The age distribution results are given in Figure 1(a, b).

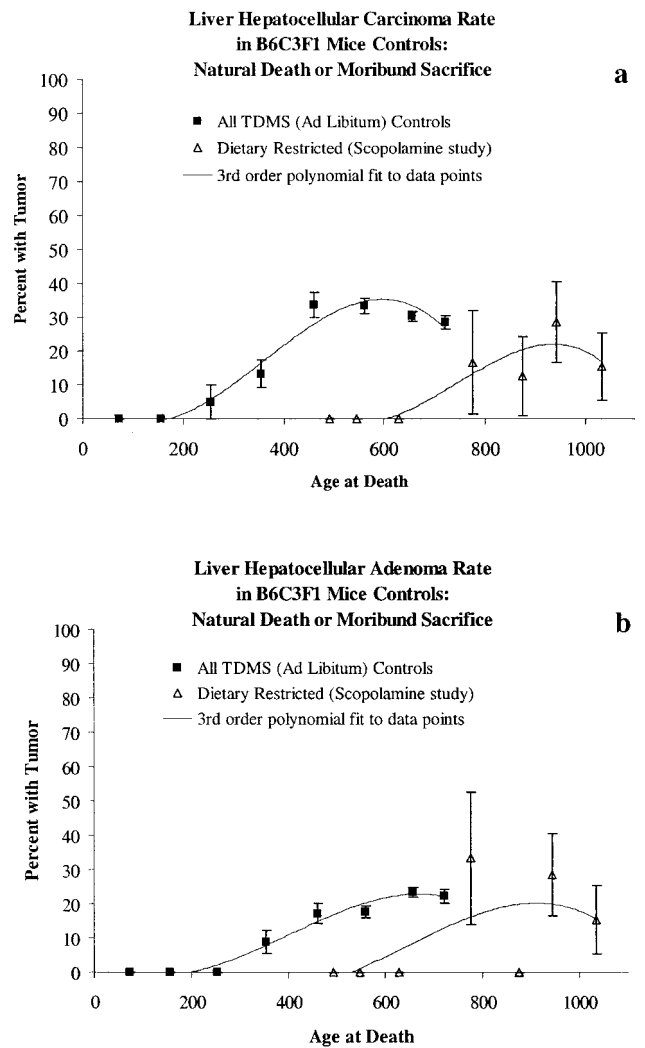


Figure 1. (a, b) Liver tumour rates for all TDMS *ad libitum* controls for mice removed for natural death or morbidity (solid symbols), and dietary restricted mice tumour rates of the TDMS scopolamine study controls (open symbols). A least-squares polynomial curve fit ($a_0 + a_1t + a_2t^2 + a_3t^3$) of the data points is fitted to each data set, for comparison purposes.

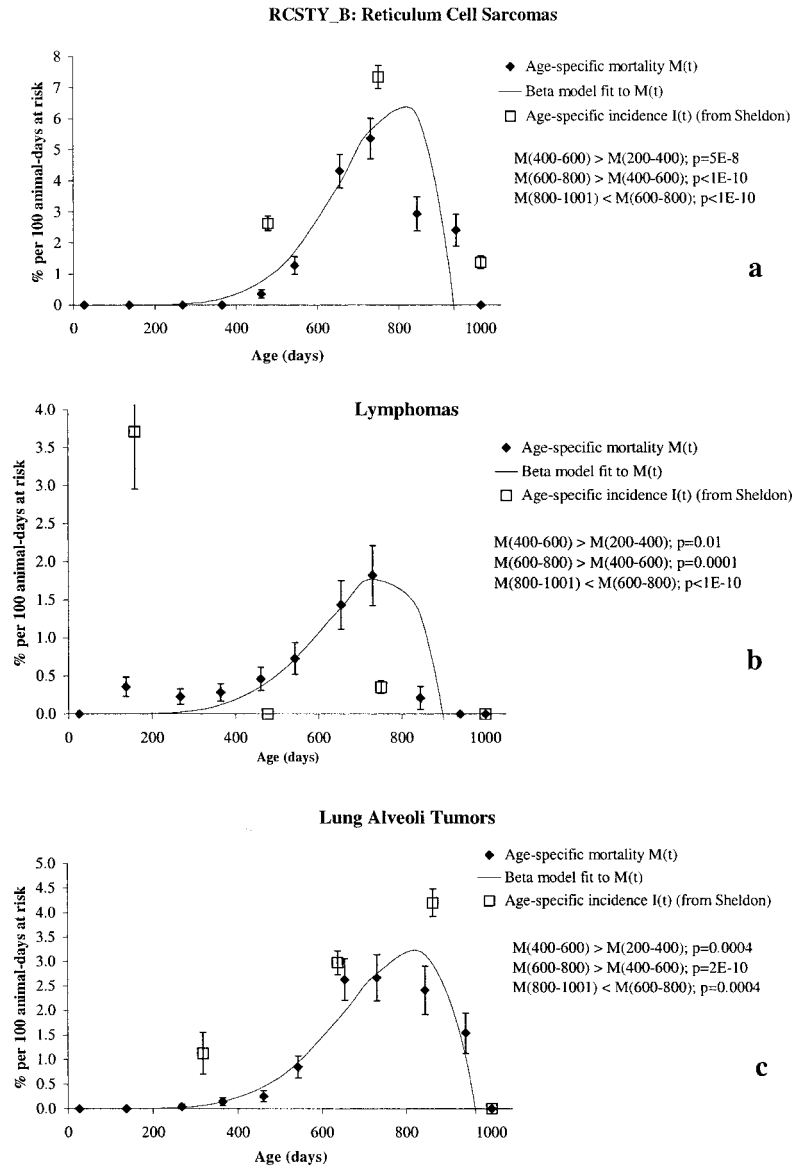


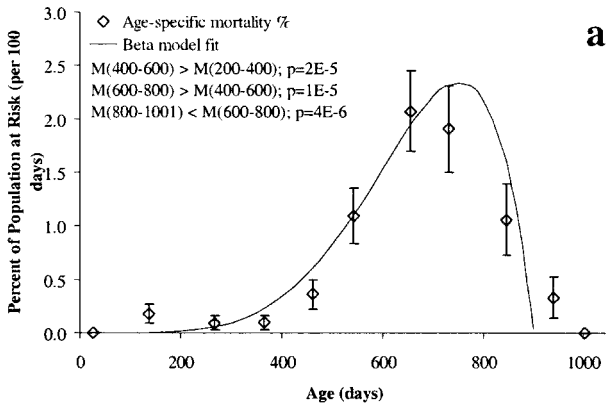
Figure 2. (a–c) Age-specific mortality (including morbidity) caused by the three most common causes of death by neoplasm for ED01 undosed control animals and data fit by the beta model. Tests of significant changes show in all cases that the oldest age group (800–1001 days) has significantly lower age-specific mortality than the 600–800 days group, which in turn has significantly higher age-specific mortality than both the 400–600 and the 200–400 days groups. Calculated age-specific incidence for the same tumour sites from data by Sheldon and Greenman (1980) are shown for comparison.

It must be noted that we do not know when the cancers actually occurred, only that they occurred before the time of death. If it is assumed that the neoplasms did not cause the natural deaths (as specifically stated in TR-421), the natural death data might represent cumulative incidence of the neoplasm. In principle, the time derivative of the cumulative incidence curve would yield the desired age-specific incidence, which clearly tends to zero as implied by the flattening observed. However, in other studies the neoplasms were a direct cause of removal for morbidity (e.g., TR-390), or the effect of neoplasms on death or morbidity was uncertain (e.g., TR-391), which suggests that the

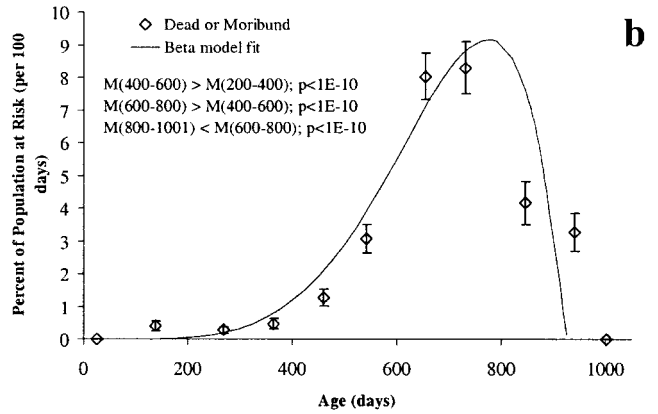
results of Figure 1 are an uncertain mixture of hazard function and cumulative incidence, thus making interpretation and modeling *via* the beta function equivocal. For purposes of interpretation of the general trends of the data, a least-squares polynomial fit of the data points is shown.

For the liver neoplasms, it seems clear that tumour rates as defined for autopsied mice dying from natural causes appear to level off in the range of about 500 to 800 days. The polynomial fits leave some impression of a turnover, but without data at elevated age and clearer definition of the death process relation to tumour incidence, it is difficult to place significant further weight on this evidence.

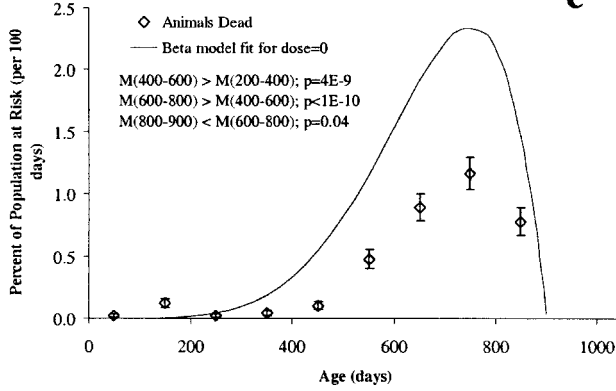
**Age-Specific Mortality for Dose = 0:
 Death Caused by Neoplasms**



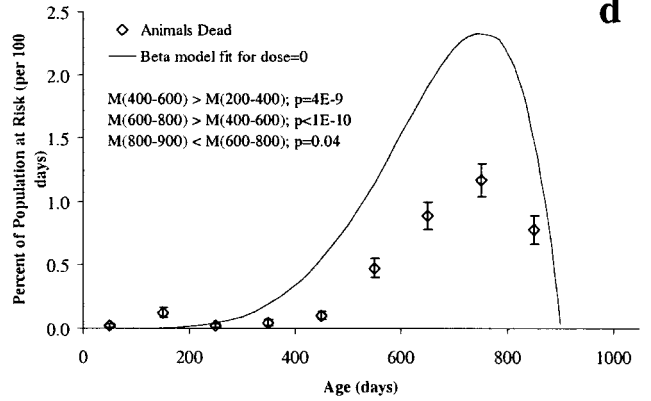
**Age-Specific Mortality for Dose = 0:
 Death or Moribund Caused by Neoplasms**



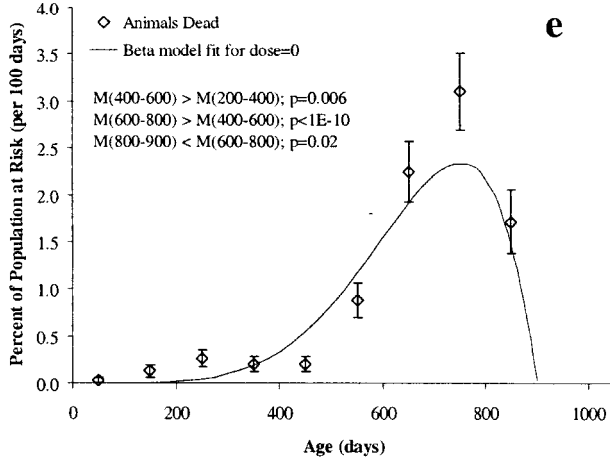
**Age-Specific Mortality for Dose = 30 ppm:
 Death Caused by Neoplasms**



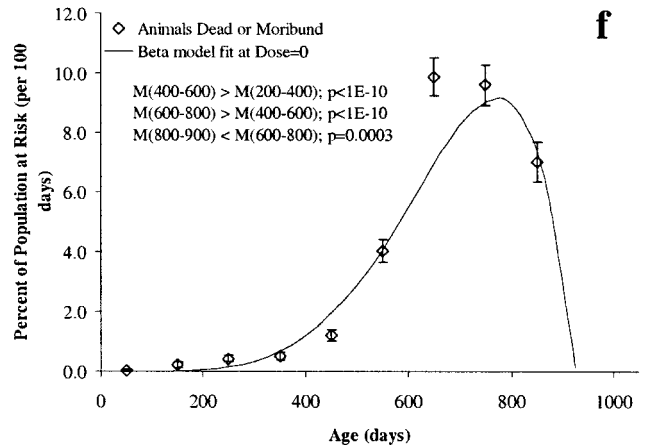
**Age-Specific Mortality for Dose = 30 ppm:
 Death Caused by Neoplasms**



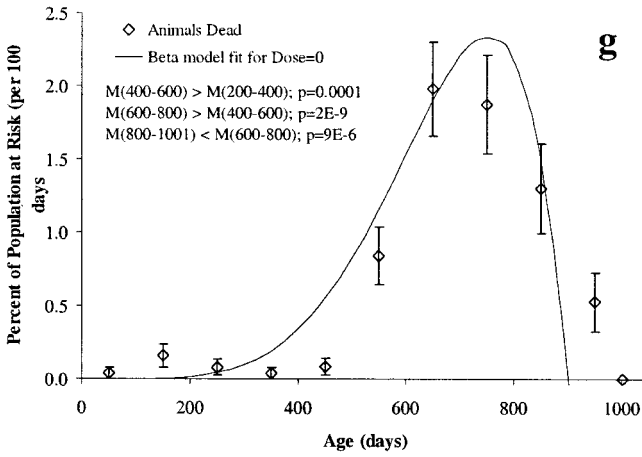
**Age-Specific Mortality for Dose = 35 ppm:
 Death Caused by Neoplasms**



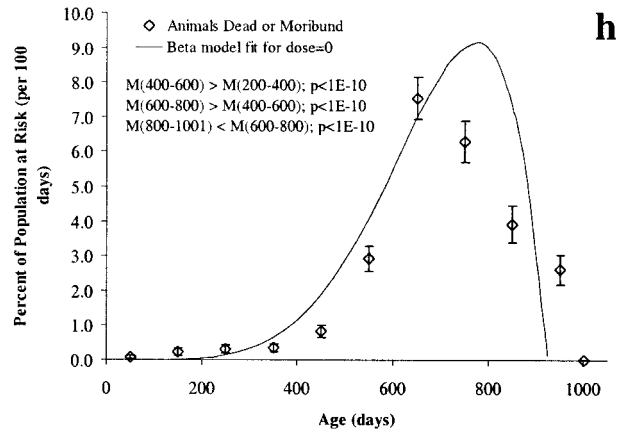
**Age-Specific Mortality for Dose = 35 ppm:
 Death or Moribund Caused by Neoplasms**



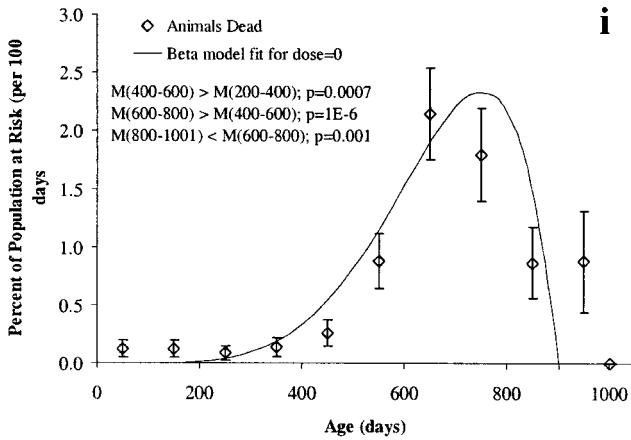
**Age-Specific Mortality for Dose = 45 ppm:
 Death Caused by Neoplasms**



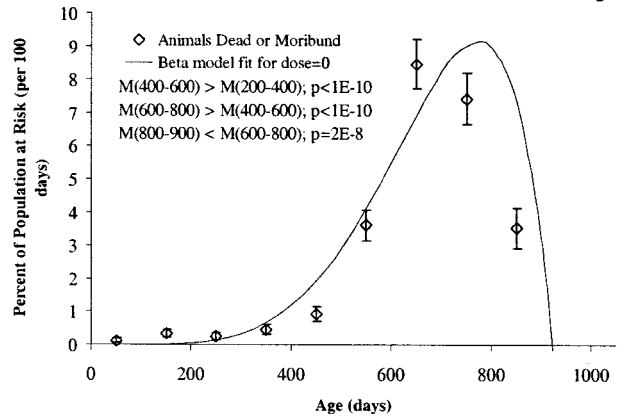
**Age-Specific Mortality for Dose = 45 ppm:
 Death or Morbidity Caused by Neoplasms**



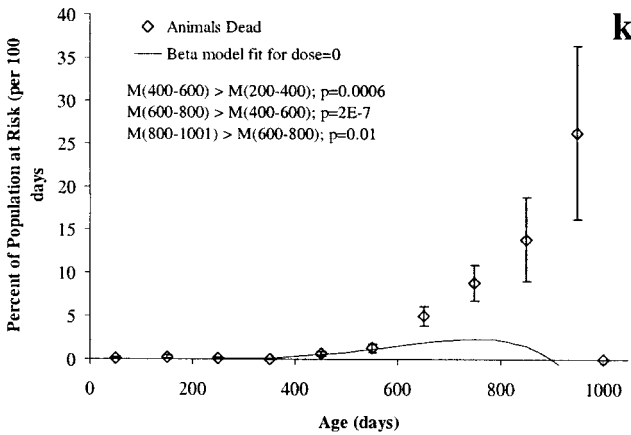
**Age-Specific Mortality for Dose = 60 ppm:
 Death Caused by Neoplasms**



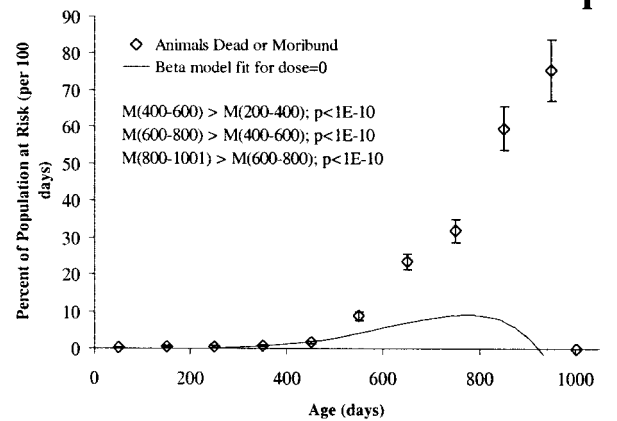
**Age-Specific Mortality for Dose = 60 ppm:
 Death or Morbidity Caused by Neoplasms**



**Age-Specific Mortality for Dose = 75 ppm:
 Death Caused by Neoplasms**



**Age-Specific Mortality for Dose = 75 ppm:
 Death or Morbidity Caused by Neoplasms**



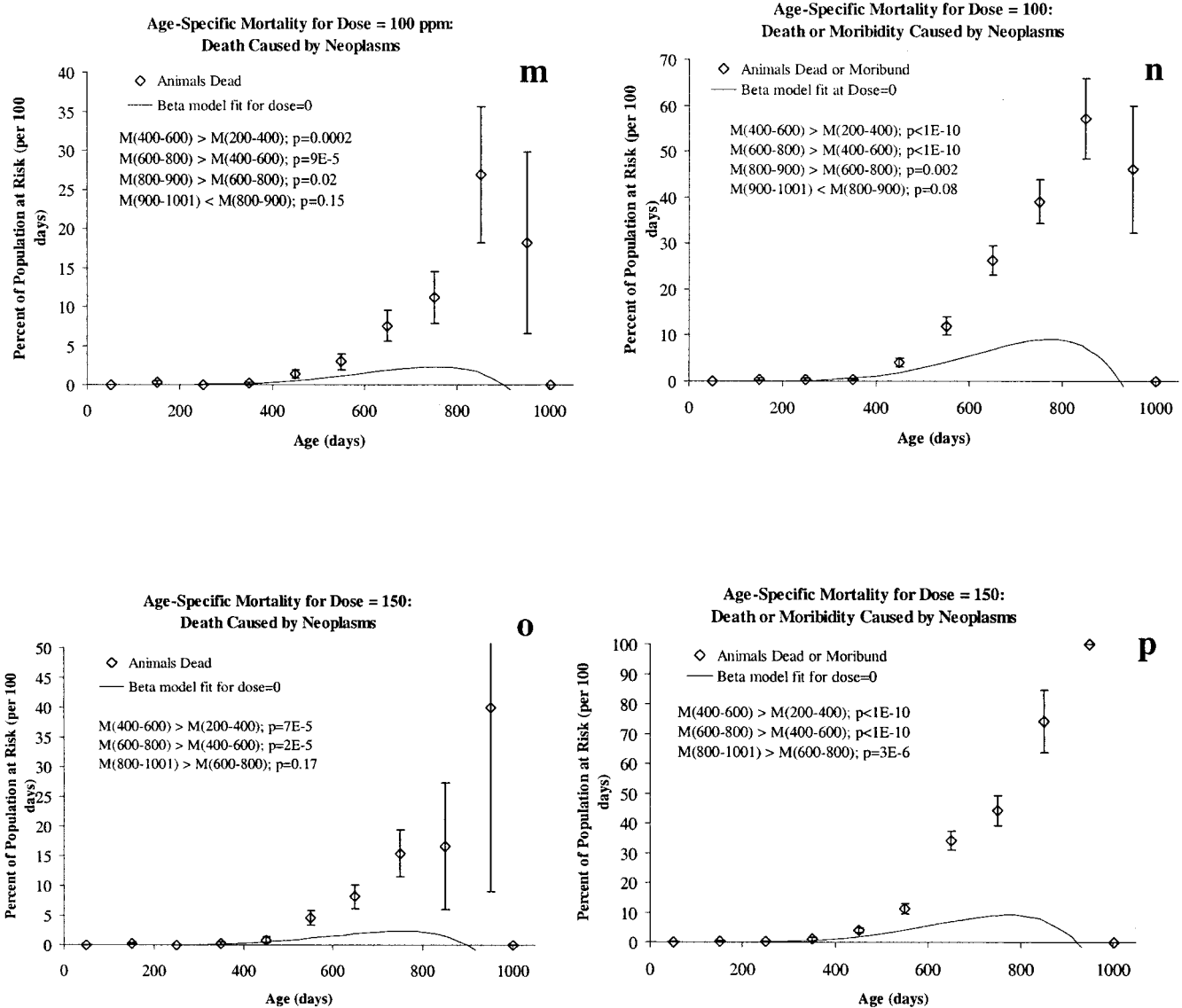


Figure 3. (a–p) ED01 age-specific mortality for causes of death (left) and death and morbidity (right) by all neoplasms versus dose of 2-AAF. For comparison, the beta model fit for the dose=0 data is shown in all curves. Tests of significant changes show at all doses up to 60 ppm, the oldest age group (800–1001 days) has significantly lower age-specific mortality than the 600–800 days group, which in turn has significantly higher age-specific mortality than both the 400–600 and the 200–400 days groups. For the dose=75, 100 and 150 ppm groups, age-specific mortality continues to increase beyond the age of turnover observed for the low-dose groups.

The dietary restriction data, which included 51 animals, 8 with liver carcinoma and 8 with liver adenoma, is also plotted in Figure 1(a, b). The data and curve fit are suggestive that the main effect of such restriction is to shift (or possibly stretch) the time scale, without appreciably influencing the peak value of the cancer rates.

Figures 2(a–c) and 3(a–p) present the age-specific mortality of the ED01 mice study. In contrast to the TDMS data, natural deaths were known to be caused by the neoplasms, ranging from 84% to 92% for controls and dosed cohorts (Kodell *et al.*, 1980), and the pathology data identified each animal for which death or morbidity was

caused by neoplasm. The reticulum cell sarcoma, lymphoma, and lung alveoli tumour mortality data of Figure 2 show age-specific mortality falling to near zero at age between 900 and 1000 days, as predicted by the beta model fit shown for reference. In all cases, the turnover is highly significant ($p<0.05$). The age-specific incidence data computed from Sheldon and Greenman (1980) data show turnover at about the same age as the mortality data, confirming that incidence as well as mortality tends toward zero.

The age-specific mortality for all sites combined for all dose groups of Figure 3(a–p) show good agreement with

Table 1. Lifetime^a cumulative probability^b of mortality from cancer.

| Dose | Site(s) of fatal tumours | Cumulative lifetime mortality from fatal tumours (%) | Cumulative lifetime mortality or morbidity from fatal tumour (%) |
|----------|--------------------------|--|--|
| Controls | All | 7 | 25 |
| 30 ppm | All | 4 | 16 |
| 35 ppm | All | 8 | 29 |
| 45 ppm | All | 7 | 22 |
| 60 ppm | All | 7 | 23 |
| 75 ppm | All | 43 | 87 |
| 100 ppm | All | 50 | 84 |
| 150 ppm | All | 58 | 93 |
| Controls | Reticulum cell sarcomas | | 15 |
| Controls | Lymphomas | | 5 |
| Controls | Lung alveoli tumors | | 10 |

^aAssuming natural lifetime of 1001 days.

^bCalculated as $1 - \exp[-\sum M(t)]$.

the general shape of the beta model shown for reference, for all doses except 75, 100 and 150 ppm, where dose-related effects become dominant. The beta curve shown for reference is the fit for the dose=0 cohort, providing a graphical indication of the effect of dose on the age distribution of cancer mortality. The data for death by neoplasms, and death or morbidity by neoplasms is shown separately, confirming that the age distribution of morbidity is very similar in shape to the mortality. In all cases except the high doses, the turnover is statistically significant ($p < 0.05$). Of interest is the observation that cancer mortality and morbidity are lower for the 30-ppm dose cohort than for undosed controls at all ages.

Table 1 tabulates the cumulative lifetime mortality caused by neoplasms for each of the ED01 cohorts examined. Except for the high-dose groups, age-specific mortality and morbidity falls well short of certainty, suggesting that incidence also falls well short of certainty.

Discussion

The data from the ED01 study show unequivocally that cancer age-specific incidence in mice, assumed to be well correlated with age-specific cancer mortality and morbidity, turns over after about 800 days for all tumour sites and doses discussed except for the highest doses. Because of the reliance on mortality and morbidity data, the data are limited to those sites where the cancer is fatal and near fatal. The age (about 800 days) of maximum incidence is about 80% of the maximum age. This may be compared to the age at peak incidence of people of about 85 years, which is also about 80% of maximum lifetime.

The 75-, 100- and 150-ppm results of Figure 3 and Table 1 support the paradigm (Doll, 2001) that sufficiently high doses of carcinogens must produce near certainty of cancer. However, as importantly, this data also suggests that the medical pathology bias toward underreporting cancer as

cause of death in the oldest humans (also Doll, 2001) is absent in the ED01 data. For the lower doses, the data support the similar conclusion drawn from human data, that cancer is not inevitable as the animal ages.

As shown in Figures 2 and 3, the beta function, which provides an accurate fit to human incidence data (Pompei and Wilson, 2001), also well represents the features of the ED01 data: non-linear increase to a peak incidence value during the first three-fourths of a lifetime, followed by leveling and sharp decrease during the final one-fourth of lifetime. This clearly suggests a similar biological mechanism in both species.

As mentioned in the Appendix, and discussed in Pompei and Wilson (2001) one interpretation of the beta function is the probability of achieving $k-1$ stages in any order for cancer creation, before achieving the one step that would prevent the precancerous cell from becoming a malignancy. The cancer creation process may be an exponential as opposed to the power law expression, but the resultant fit and interpretation are unchanged: the data supports the existence of an important cancer extinction process which dominates near end of life. Such biological mechanisms discussed in the human work include increasing apoptosis with age and increasing cell senescence with age. More recently, a slowing of microscopic tumour angiogenesis with age might be linked to a possible explanation (Folkman, 2001).

A promising biologically and mathematically consistent model of the underlying cancer extinction process may be proposed by considering the demonstration by Hart and Setlow (1976) that DNA synthesis in human cells, both scheduled (normal) synthesis and unscheduled (to repair damage from UV radiation) synthesis, markedly reduce with age. Young cells were found to undergo normal synthesis with near 100% probability, with the proportion reducing roughly linearly to about only 10% of the oldest cells able to synthesize DNA. Since repair synthesis is also reduced with age at about the same rate

as normal synthesis, the authors suggest that lack of DNA repair is not a determining characteristic with age, but rather that cells lose the capacity for any DNA synthesis as they age, and thus cannot replicate. As further discussed by Hart *et al.* (1979), lack of replicative ability defines senescence, which allows the cell to function normally, but inability to maintain genome integrity eventually leads to its death.

Recently, Rubelj and Vondracek (1999) and Rubelj *et al.* (2000) have proposed that cell senescence may be produced by a stochastic process which abruptly shortens DNA telomeres, thus causing immediate (within one cell cycle) loss in replicative ability, instead of gradual loss of telomere length. This causes cells at any age to suddenly switch from replicative to senescent, thus arresting DNA synthesis and proliferation by those cells. When, as proposed, the probability of this sudden senescence is uniformly distributed, and its cumulative probability approaches certainty of senescence for the oldest cells, this process may well model the results observed by Hart and Setlow (1976), and might mathematically be similar to the cancer extinction factor $(1-\beta t)$ of the beta model.

Accordingly, cell senescence might be a significant causal factor in the incidence turnover by a process which may approximate the beta model derivation assumption of uniformly distributed loss of proliferative ability (see Appendix), *i.e.*, linearly increasing probability of loss of proliferative ability with age, reaching certainty at approximately maximum life span. Since we believe that an explanation of the cancer incidence turnover necessarily involves the inclusion of some biological process or processes not included in the historically dominant models of cancer induction — the multistage and clonal expansion models, further work is required to determine if the candidate process is as easily added as the above implies.

Future work

This work supports the possibility that if we live long enough we outlive our cancers, a possibility directly proposed in our work with human data. We suggest that future work should include:

1. Examination of more animal data sets for the turnover in incidence.

2. Explorations of additions of the hypothesized cancer extinction or loss of proliferative ability factor to exact multistage and clonal expansion models, exploring the fits and biological implications.

Searching for evidence of a variation in the apparent cancer extinction effect with a) carcinogen dose, and b) dietary restriction. There is some evidence of this in the ED01 data for high doses of 2-AAF for the former, and in the TDMS data for the latter.

Acknowledgments

We express our sincere appreciation to Drs R.W. Hart of FDA NCTR and J.K. Haseman of the NIEHS, and Sir Richard Doll of the University of Oxford for their valuable comments and suggestions.

Appendix

The selection of the beta distribution for the data fits arises from the observation that the power law equation $I(t)=at^{k-1}$ (Armitage and Doll, 1954) well fits many cancer-site incidence data up to about age 70 (ignoring childhood cancers) At older ages, the incidence data markedly flatten, and show reduction at sufficiently elevated age. Accepting the validity of the power law fits at younger ages (but not necessarily the biological validity of the power law model itself), we add the hypothesis that a ‘cancer extinction’ term is influencing the carcinogenesis process, eventually becoming dominant at sufficiently elevated age.

Adding this cancer extinction term to the power law is accomplished directly by forming the probability statement: probability of cancer=the probability of reaching k stages and the cancerous cell does not die (or lose its proliferative ability). We write this probability and expand as: $P_c=P(bt^k \cap \text{not death})=P(bt^k | \text{not death})P(\text{not death})=bt^k \times P(\text{not death})$. The simplest assumption for a probability density for cancerous cell death is a uniform distribution over the time interval 0 to ct , leading to $P_c=bt^k(1-ct)$, where c is a constant. Taking the time derivative to convert the probability to a probability density function for a single cell: $f(t)=\alpha t^{k-1}(1-\beta t)$, where α and β are constants. We immediately recognize the beta distribution $f(x)=\lambda t^{r-1}(1-x)$ over the interval $0 \leq x \leq 1$, where $x=\beta t$. A textbook interpretation of $f(x)$ is the density for the $(r-1)$ th largest of r uniform $(0,1)$ random variables (Larson, 1982), which can be restated as the probability density function for achieving $(r-1)$ stages (cancer creation) without achieving the r th stage (cancer extinction).

Expanding from consideration of a single cell to N cells in an organ, and denoting $f(t)=F'(t)$, the probability of cancer is $G(t)=1-[1-F(t)]^N$. For large N , this simplifies to $G(t)=1-e^{-NF(t)}$, which is accurate to 10^{-10} for $N=10^8$ cells. As discussed by both Moolgavkar (1978) and Armitage (1985), the age-specific incidence function for the organ tissue is not the density function $G'(t)$ itself, but the associated hazard function, given by $h_c(t)=G'(t)/[1-G(t)]$, which represents the incremental risk of cancer per unit time given that the tissue has been cancer-free to time t . Completing the derivation, $h_c(t)=e^{-NF(t)}Nf(t)/e^{-NF(t)}=Nf(t)$. We note that the age-specific cancer incidence for a site tissue is related to the probability density function for one

cell by the constant N , thus leaving the beta model as $f(t)=\alpha t^{k-1}(1-\beta t)$, modified by only by a constant (absorbed into α) to apply to a multicellular organ site. The final expression chosen immerses the α constant into the $k-1$ power in order to preserve the historical view of $k-1$ stages, each with its own transition rate (assumed to be equal in this case), thus denoting the final form as $b(t)=(\alpha t)^{k-1}(1-\beta t)$.

References

- Armitage, P. 1985 Nov: Multistage models of carcinogenesis. *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.
- Armitage, P. and Doll, R. 1957: A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *British Journal of Cancer* 11, 2.
- Cairns, T. 1980: The ED01 study: introduction, objectives, and experimental design. *Journal of Environmental Pathology and Toxicology* 3, 1–7.
- Case R.A.M. 1966: Tumours of the urinary tract as an occupational disease in several industries. *Ann R Coll Surg Engl* 39, 213–35.
- de Rijke, J.M., Schouten, L.J., Hillen, H.F., Kiemeny, L.A., Coebergh, J.W., van den Brandt, P.A. 2000 Sep 1: Cancer in the very elderly Dutch population. *Cancer* 89, 1121–33.
- Doll, R. 2001: Private communication.
- Folkman, J. Feb 26, 2001: DEAS lecture, Harvard University.
- Hart, R.W., D'Ambrosio, S.M., Ng, K.J. and Modak, S.P. 1979 Feb: Longevity, stability and DNA repair. *Mechanisms of Ageing and Development* 9, 203–23.
- Hart, R.W. and Setlow, R.B. 1976 Jan–Feb: DNA repair in late-passage human cells. *Mechanisms of Ageing and Development* 5, 67–77.
- Haseman, J. 2001: Personal communication.
- Kodell, R.L., Farmer, J.H., Littlefield, N.A. and Frith, C.H. 1980: Analysis of life-shortening effects in female BALB/C mice fed 2-acetylaminofluorene. *Journal of Environmental Pathology and Toxicology* 3, 69–87.
- Larson, H.J. 1982: *Introduction to probability theory and statistical inference*, third edition. New York, NY: Wiley.
- Moolgavkar, S.H. 1978 Jul: The multistage theory of carcinogenesis and the age distribution of cancer in man. *Journal of the National Cancer Institute* 61, 49–52.
- Moolgavkar, S.H. and Knudsen, A.G. June 1981: Mutation and cancer: a model for human carcinogenesis. *Journal of the National Cancer Institute* 66, 1037–52.
- Parkin, D.M., Whelan, S.L., Ferlay, J., Raymond, L. and Young, J. 1997: *Cancer incidence in five continents*, Volume VII. IARC Scientific Publications No. 143, Lyon.
- Pompei, F. and Wilson, R. 2001 Nov: Age distribution of cancer: the incidence turnover at old age. *Human and Ecological Risk Assessment*, 7, 1619–50.
- Ries, L.A.G., Eisner, M.P., Kosary, C.L., Hankey, B.F., Miller, B.A., Clegg, L. and Edwards, B.K., editors, 2000: *SEER Cancer Statistics Review, 1973–1997*. Bethesda, MD: National Cancer Institute. http://seer.cancer.gov/Publications/CSR1973_1997/
- Rubelj, I. and Vondracek, Z. 1999 Apr 21: Stochastic mechanism of cellular aging — abrupt telomere shortening as a model for stochastic nature of cellular aging. *Journal of Theoretical Biology* 197, 425–38.
- Rubelj, I., Huzak, M. and Brdar, B. 2000 Jan 10: Sudden senescence syndrome plays a major role in cell culture proliferation. *Mechanisms of Ageing and Development* 112, 233–41.
- Sheldon, W.G. and Greenman, D.L. 1980: Spontaneous lesions in control BALB/C female mice. *Journal of Environmental Pathology and Toxicology* 3, 155–67.