

Cancer Suppression at Old Age

Charles Harding, Francesco Pompei, Ellen E. Lee, and Richard Wilson

Jefferson Laboratories, Department of Physics, Harvard University, Cambridge, Massachusetts

Abstract

Increased age is regularly linked with heightened cancer risk, but recent research suggests a flattening around age 80. We report that, independent of cancer site or time period, most incidence rates decrease in the more elderly and drop to or toward zero near the ceiling of human life span. For all major organ sites, male and female, we use 1979 to 2003 Surveillance, Epidemiology, and End Results registry records (8–26% of the U.S. population) to construct three sequential cross-sections at 10-year intervals, totaling 129 sets of age-specific cancer data. To compute incidence rates, we estimate older populations at risk with census counts and NIH life tables. This article provides both a minimal and a more comprehensive extension of Surveillance, Epidemiology, and End Results cancer rates to those above 85. Almost all cancers peak at age ~ 80. Generally, it seems that centenarians are asymptomatic or untargeted by cancers. We suggest that the best available justification for this pattern of incidence is a link between increased senescence and decreased proliferative potential among cancers. Then, thus far, as senescence may be a carcinogen, it might also be considered an anticarcinogen in the elderly. We model rising and falling incidence rates with a β curve obtained by appending a linearly decreasing factor to the well-known Armitage-Doll multistage model of cancer. Taken at face value, the β model implies that medical, diet, or lifestyle interventions restricting carcinogenesis ought to be examined for possible effects on longevity. [Cancer Res 2008;68(11):4465–78]

Introduction

Age plays a strong role in carcinogenesis, but the exact nature of this role is unclear. For half a century, cancer investigators have reported that typical cancer rates accelerate with age. Our understanding of age-specific cancer rates traditionally comes from mortality data. More frequently collected than incidence records, mortality data have also been termed more reliable; in the past, autopsy was widely considered as a method of verifying or disaffirming many uncertain diagnoses (1). After late childhood or early adulthood, cancer mortality broadly rises in exponential fashion or by power law. Accordingly, age has been seen as precipitator, enabler, or cause of cancer. Neglecting mechanism, it has even been called a potent carcinogen (2). Yet, as early as 1954, rates were observed to flatten above 75 (3). At that time, many of the very old did not know their own age, and “old age” itself was an accepted cause of death. Consequently, cancer investigators questioned the reliability of old-age (above age 74) mortality

records and any associated analysis. Because modern medicine is able to treat many cancers successfully, mortality is now a less reliable tool for understanding cancer incidence. On the other hand, greatly improved cancer diagnoses and the establishment of well-supported cancer registries over the past 35 years have improved cancer incidence data to the point that it is now generally accepted as high quality (4). Here we have accepted the reliability of older age registry data, generally understudied but with samples positively assessed (5, 6).

It is self-evident that the number of new cancer cases tends to zero at ages near the end of a natural life span because the number of persons alive and at risk tends to zero. But investigators working with recent data from Surveillance, Epidemiology, and End Results (SEER; ref. 7) and other modern registries have confirmed that cancer incidence rates (new cases divided by the number of person-years at risk) after about age 80 also tend to sharply decrease (8–10). Pathology studies in autopsy have other systematic errors but tend to support the idea that cancer prevalence increases more slowly at old age and these cancers tend to be less virulent in these older patients (11, 12). This fall off is contrary to simple models of cancer incidence. The Armitage-Doll multistage model, in its usual approximate form $I(t) = (p_1 p_2 \dots p_k) t^{k-1} / (k-1)! = at^{k-1}$, and the well-known Moolgavkar-Vernon-Knudson two-stage clonal expansion model (13) both predict that cancer incidence will rise indefinitely with age.

We note three classes of explanation that have been offered for this observation. The first class of explanation is that the multistage theory is accurate, but the probabilities at each stage are dependent on environmental influences and these change with time. Such a change in probability would obviously be the case if younger people have chemical exposures, which cease at age 70. They would have to be exposures of late-stage carcinogens that dominate the probability. The probability would have to drop by at least a factor of four to achieve this. However, we have found no data that quantify this.

The second is there is some biological effect that must be added to the simple multistage model. A subset of this is that there is a variation in susceptibility. This was studied very early by Cook and colleagues (14) who showed that the fall off with age could be explained by a variation in susceptibility only if the number of persons who are susceptible considerably varies with cancer site, an idea they considered unlikely. Two other factors led Cook to reject the idea that variation in susceptibility is crucial. First, when groups of people were exposed to sufficiently strong carcinogenic chemicals, the susceptibility increases so that over 50% of the group develop cancer. It seems more plausible to assume that this is a gradual increase in susceptibility as the dose is increased rather than a sudden nonsusceptible/susceptible switch at a threshold dose. Second, it seems implausible for susceptibility to vary in just the necessary way to make incidence turnovers occur at the same age, independent of cancer site and rarity. The Cook result is confirmed by an exact multistage model by Ritter and colleagues

Requests for reprints: Francesco Pompei, Department of Physics, Harvard University, Oxford Street, Cambridge, MA 02138. Phone: 617-495-3387; Fax: 617-495-0416; E-mail: fpompei@post.harvard.edu.

©2008 American Association for Cancer Research.
doi:10.1158/0008-5472.CAN-07-1670

(15). The result for three cancers is shown in Fig. 1 from Ritter and colleagues. Good fits can be obtained only if the fraction of susceptibles in the population depends on the cancer. Eighteen percent of people would be susceptible to colon cancer, 5% susceptible to non-Hodgkin's lymphoma, 1% susceptible to brain cancer, and 82%, 95%, and 99%, respectively, are completely immune to these cancers. This deterministic model seems to us unlikely. Finkel (16), using a more realistic distribution of susceptibilities, showed that the inclusion only flattened the incidence curve with age. Several other biological processes have been suggested as candidates to explain the cancer fall off: changes in immune system activity with age (17) and the consequences of increased cellular senescence with age (18, 19). These are mentioned further in the Discussion.

Although laboratory animals, rats and mice, are not people and have significant biological differences, they have traditionally given guidance on aspects of cancer to be examined in people. Laboratory rats and mice are usually killed (terminal sacrifice) at 2 years or less. But a study shows that when rats or mice have been allowed to live to their natural lives up to 3 years or more, cancer mortality can fall. In particular, a study of 2,234 undosed control mice of the ED01 study suggested a similar effect, but these are now being reanalyzed more carefully because of problems in decoding the data from the 30-year-old study (20).

A third class is that incidence data might not be derived consistently for the young and the old. For example, cancer incidence data are less objective than mortality and might be underreported. There are many possible reasons for this. Therefore, in this study, three data cross-sections at 10-year intervals were studied to find if there were any significant cohort effects due to environmental, lifestyle effects or improvements in diagnoses or other factors affecting the reliability of the incidence data that have changed with time. In addition, we decided to study more cancer types in both genders. Animal studies can also help to elucidate this question because underreporting at older ages is less likely. Because some aspects of cancer may be racially dependent, from either differing levels of exposure to pollutants, unequal diagnosis, or inherent biological effects, we also examined a cross-section of the black population in the SEER database.

This is a subject worthy of investigation due to the possible clues to cancer suppression that might lead to prevention measures. Because one half of the U.S. population lives to age 80 and beyond, and that the oldest age groups are the fastest growing of all age groups worldwide, makes the results of studies of this phenomenon immediately applicable in many areas.

Materials and Methods

The SEER U.S. cancer registry presently records incidence data from ~26% of the U.S. population, with complete patient-by-patient information on age of diagnosis, year of diagnosis, and cancer type. We considered only malignant diagnoses (by SEER standards) and only the first diagnosis of each cancer type in any patient. For the period 1973 to 2003, we used SEER*Stat software (21) to obtain patient case listings for 20 major male organ site cancers and 21 major female organ site cancers in the broadest populations possible. We developed software to automate most of the procedure of computing incidence. With it, these case listings were separated into three 5-y cross-sections: 1979 to 1983, 1989 to 1993, and 1999 to 2003. Cancer cases were grouped by sex, cross-section, and twenty-three 5-y age groups ranging from 0 to 4 to 110 to 114. Computing incidence rate requires person-years at risk. SEER*Stat provides yearly populations by registry, sex, and eighteen 5-y age groups ranging from 0 to 4 to 85+. To

partition the 85+ population category into 5-y age groups, we developed a simple estimation procedure that draws from the U.S. census (for more recent years; ref. 22), NIH life tables (23–25) and postcensal estimates (for earlier years; refs. 26, 27). This procedure is given in the Appendix. For the most recent cross-section, the estimation procedure is effectively a direct application of Census 2000 regional population figures (which are partitioned by 1- or 5-y age group up to 110–115) to SEER-covered areas, county by county. In addition to our comprehensive extension of SEER data with 5-y age groups, we provide a more minimal extension, which splits those 85+ into 85 to 99 and 100+. An identical procedure was used to obtain populations in the minimal extension.

Consistent with SEER*Stat, yearly populations were used for person-years at risk without removing cancer type prevalence pools. The largest such pool is for prostate cancer, for which the January 2002 SEER prevalence statistics show peaks at 8.75% of the male population from ages 70 to 79. All other malignant cancer prevalences are each <3% of the population at any age. Correcting with a not-at-risk pool would not significantly change these results.

Where x is the count of new diagnoses and n is the person-years at risk, $\pm 34.1\%$ two-sided confidence intervals for incidence rates were calculated according to the normal distribution (for $x > 10$), according to the Poisson distribution (for $x < 10$ and $n > 1,000$), and by exact binomial proportion (for $x < 10$ and $n < 1,000$). In the latter two situations, software by John Pezzullo (28) was used. In all cases $x \ll (n - x)$. There has been much discussion on how to address this "inverse Poisson problem." Physicists using small numbers have in recent years used the prescription of Feldman and Cousins (29). This gives slightly smaller error bars than used here. Considering the 85+ data, it may seem counterintuitive for data points of larger/younger populations to have wider confidence intervals than some data points of smaller/older populations. However, this is indeed possible with exact binomial and Poisson confidence intervals. It should be noted that neither census and life table inaccuracies nor uncertainties in their application to SEER areas are accounted for in the confidence intervals. There is also error in the age value of our data, but this error is strictly limited to <2.5 y and is therefore not displayed. The two minimal extension data points are located at the age of mean population size within the 15-y age groups.

For the most recent cross-section (1999–2003), incidence rates of four main cancers were also calculated for black Americans. They were obtained through a method similar to that used for the general population. Details are in the Appendix.

For convenience of illustration, an empirical model was used to fit the data, which has the form of a β function $I(t) = at^{k-1}(1-bt)$, where $I(t)$ is the age-specific incidence, a is a combined rate constant for limiting stage transitions, k is the number of limiting (slow) stages to produce the cancer, and b is an empirical term to be described. This model was derived by adding an empirical term decreasing with age, $(1-bt)$, to the approximate multistage model $I(t) = at^{k-1}$ that assumes that the probability of proceeding from one stage to the next is independent of age. If we derive the values of a , k , and b by fitting all the data from age 0, there would be excessive emphasis on the more precise measurements at low ages. Accordingly, we ignore the data at early ages but minimize the sum of the least squares difference between the formula and the data for each cancer, beginning at the rising side inflection point of the data, about age 50. Each point is weighted by the square of the SE with appropriate upper limits when 0 is measured. Where susceptible fractions are discussed, these results are deduced from the raw incidence data computed by integrating the age-specific incidence results from age 0 to 110+. This method is effectively a retrospective calculation of the fraction of people who were susceptible to the cancer, if one assumes that there is a susceptible fraction and immune fraction for each cancer.

Results

Figures 2A to D and 3A to D show age-specific cancer incidence for all 20 of the major organ site cancers listed in SEER for males and all 21 major organ sites listed for females, along with the totals

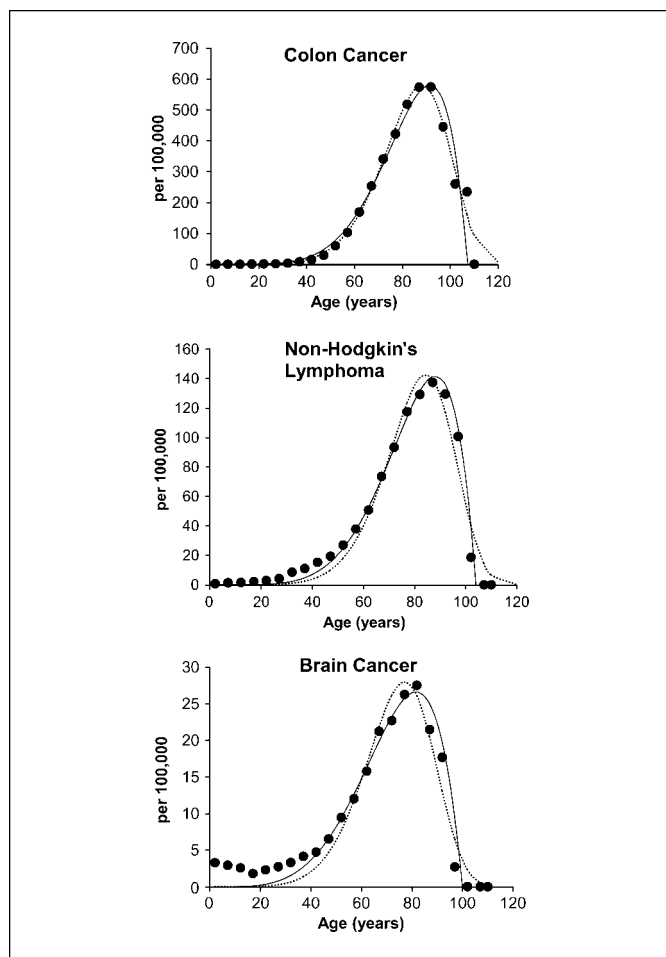


Figure 1. Comparison between SEER data (solid points), β -senescence model (solid line), and variable susceptibles model (dotted line). The dotted line shows a fit when 18.5% of people are assumed susceptible to colon cancer, 4.6% susceptible for non-Hodgkin lymphoma, and 0.95% susceptible for brain cancer [data from Ritter and colleagues (15)].

for all cancers, for the periods 1979 to 1983, 1989 to 1993, and 1999 to 2003, for a total of 129 cancer data sets. The well-known increase in cancer with age (exponential or power law) is seen up to about age 80 for nearly all cancers. Age 74 to 80 has been the limit for most prior investigations. On each of the graphs is shown the fitted values of a , b , and k . Each of the fitted values of b in the β model fit is close to $1/100$. An alternate illustration that the incidence for each cancer tends to fall in the same way would have been to plot a single fit to all cancers on each figure. To do both would overly complicate the plots. This is a consequence of the fact that the number of new cases diagnosed for each cancer falls more rapidly with age than the number of persons at risk falls at each age.

For the total of all cancers, the cancer incidence above age 100 is small, and determined with good statistical accuracy, especially for the more recent cross-section, whose population figures are the most accurate, and the number of persons living up to, and above, age 100 has increased. However, even for the rare cancers, where the statistical accuracy is less, the fall off is consistent with the same trend. Because there are surviving men and women who are cancer-free in age groups in which there are no new cases reported, the measured age-specific incidence is zero for these age groups.

The group of cancers among the black population shown in Fig. 4 shows a small absolute difference from the all-race database but a very similar fall off with age.

Two of the rarest cancers, testicular and Hodgkin lymphoma, show somewhat different patterns of rising cancer incidence with age, suggesting a different relationship with age than that derived from the two classic multistage assumptions of a transition probability independent of age and gradual accumulation of alterations over time. The prostate cancer results are not as "smooth" as the others, which may be a result of the rapid increase in diagnoses by introduction of the prostate-specific antigen (PSA) test in 1991. The data show that age-specific incidences of many cancers have changed little in 25 years, but there are some notable exceptions that are well known among oncologists. In males, leukemia, liver cancer, non-Hodgkin lymphoma, melanoma, stomach, and thyroid cancer rates have changed, and among females, cancers of the cervix, esophagus, liver, lung and bronchus, melanoma, non-Hodgkin lymphoma, stomach cancer, and thyroid cancer have changed appreciably during over the 25 years spanned by the three cohorts. Other cancers for both males and females have changed less. No attempt was made to relate these changes to differences in diagnostics or diet/lifestyle/environmental changes over the 25-year period, with the exception of the aforementioned PSA test for prostate cancer. The important fact is that, even in these cases, the fall off above age 80 was similar in all cohorts, suggesting that the cause of the fall off is not an error of data collection that diminished with time.

Discussion

We reiterate that the most important conclusion, confirmed by the more recent data, is that the cancer incidence falls rapidly to zero within the statistical uncertainty. This suggests a simple, common, reason for this fall off. But we must reexamine all the alternate explanations already listed in the Introduction.

The following specific reasons have been suggested at various times for questioning the reliability (underreporting) of the incidence data: older people are screened for cancer less frequently than younger people, and there is an often enormous amount of comorbidity in older people that may affect diagnosis. The fact that the fall off in the oldest age groups seems to be similar for all three cohorts suggests that these reasons are either not important or that time frames studied are too homogenous. The possibility of underreporting of cancer incidence in the oldest was also investigated by de Rijke and colleagues (30), who found that reporting in a group of elderly Dutch men of age over 95 was 87% accurate for cancer incidence registries 1989 to 1995.

An assumption in the multistage theory is that the probability of proceeding from one stage to the next is independent of age. But this can be questioned for those over age 80. Among the reasons for a possible reduction in the probabilities for the later stages, without assuming a new biological effect, are (a) older people change their diets with a possible reduction in dietary carcinogens, (b) older people lose weight, which may have an effect on several cancers, (c) older people may decrease substance use reducing exposure to carcinogens such as tobacco and alcohol, and (d) older people have less occupational or environmental exposures to occupational carcinogens.

The carcinogens concerned would have to be dominating a late-stage probability. The factor $(1 - bt)$, which seems to be necessary to fit the data, could be interpreted as the reduction of probability of

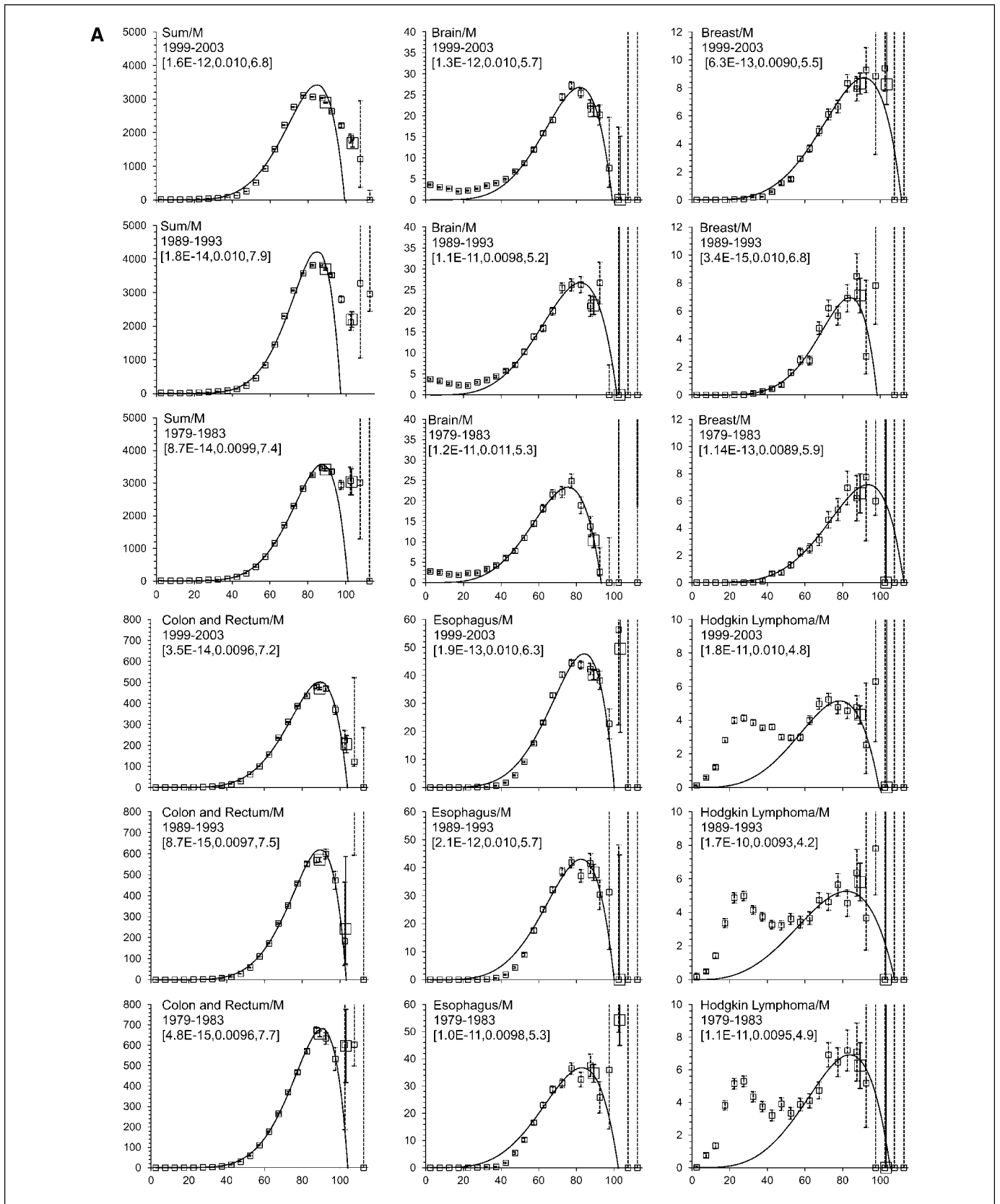


Figure 2. A, age-specific cancer incidence (per 100,000) for total cancers and each of five major organ sites for males, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - (1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

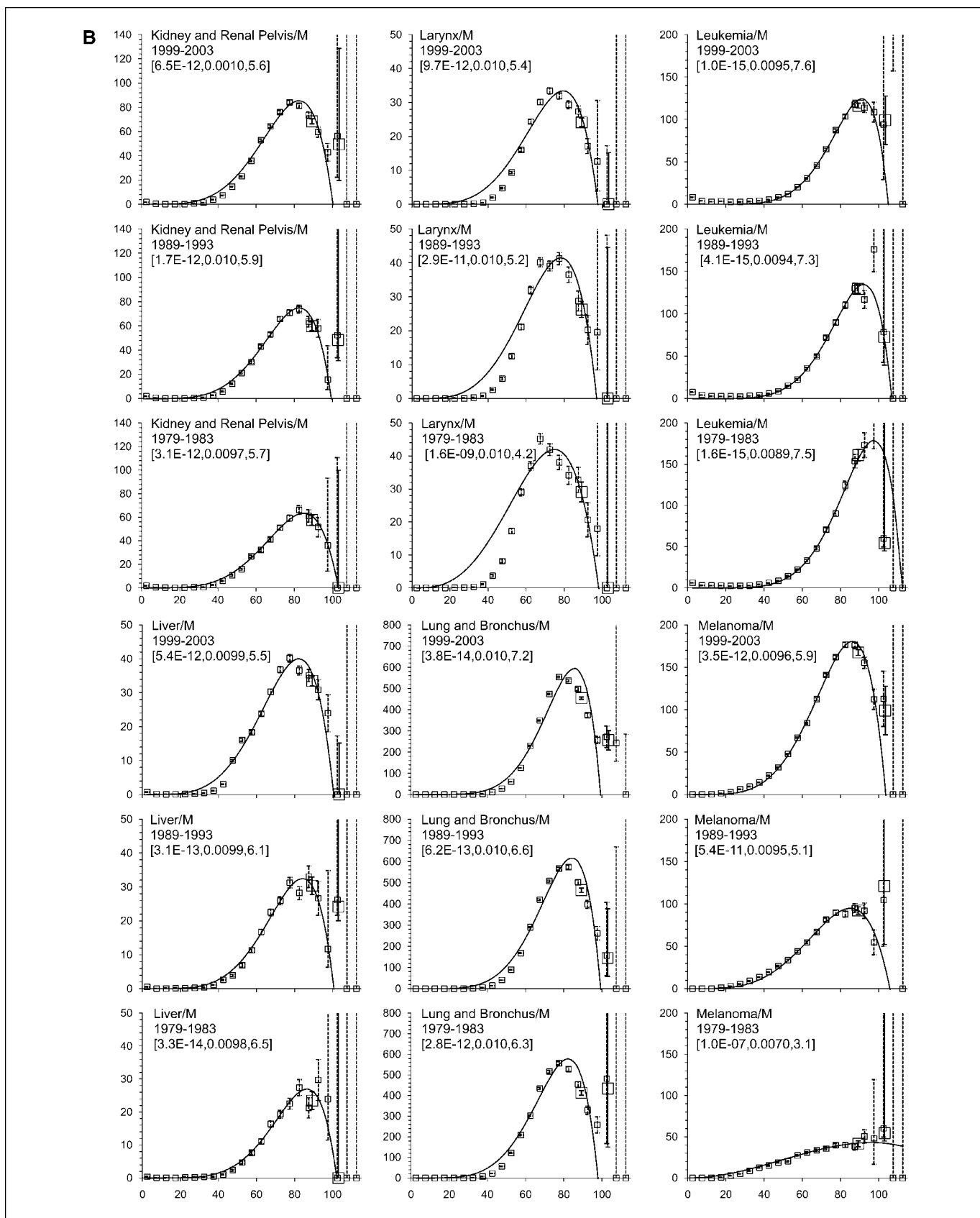


Figure 2 Continued. B. age-specific cancer incidence (per 100,000) for each of six major organ sites for males, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - 1(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

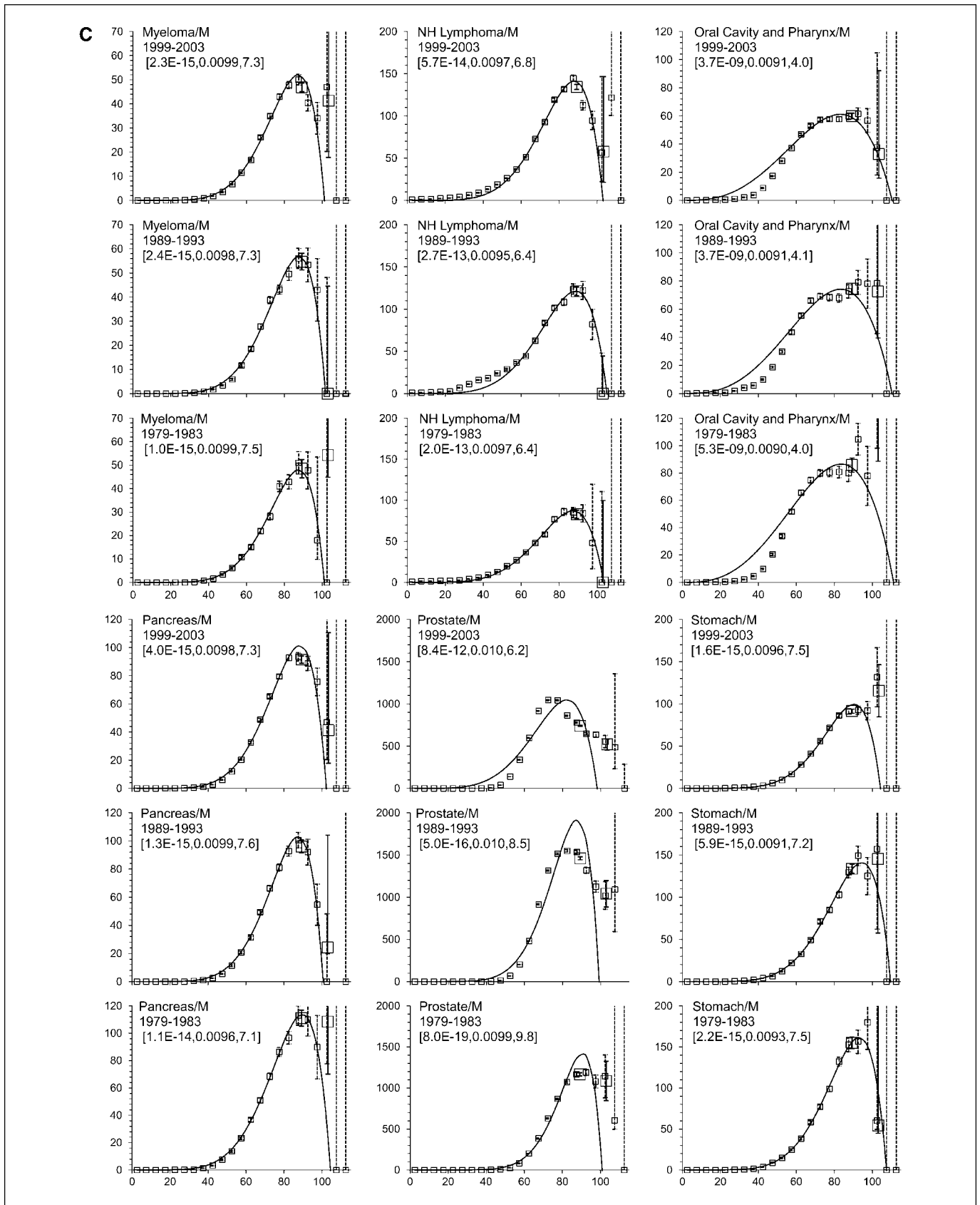


Figure 2 Continued. C. age-specific cancer incidence (per 100,000) for each of six major organ sites for males, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^{k-1}(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

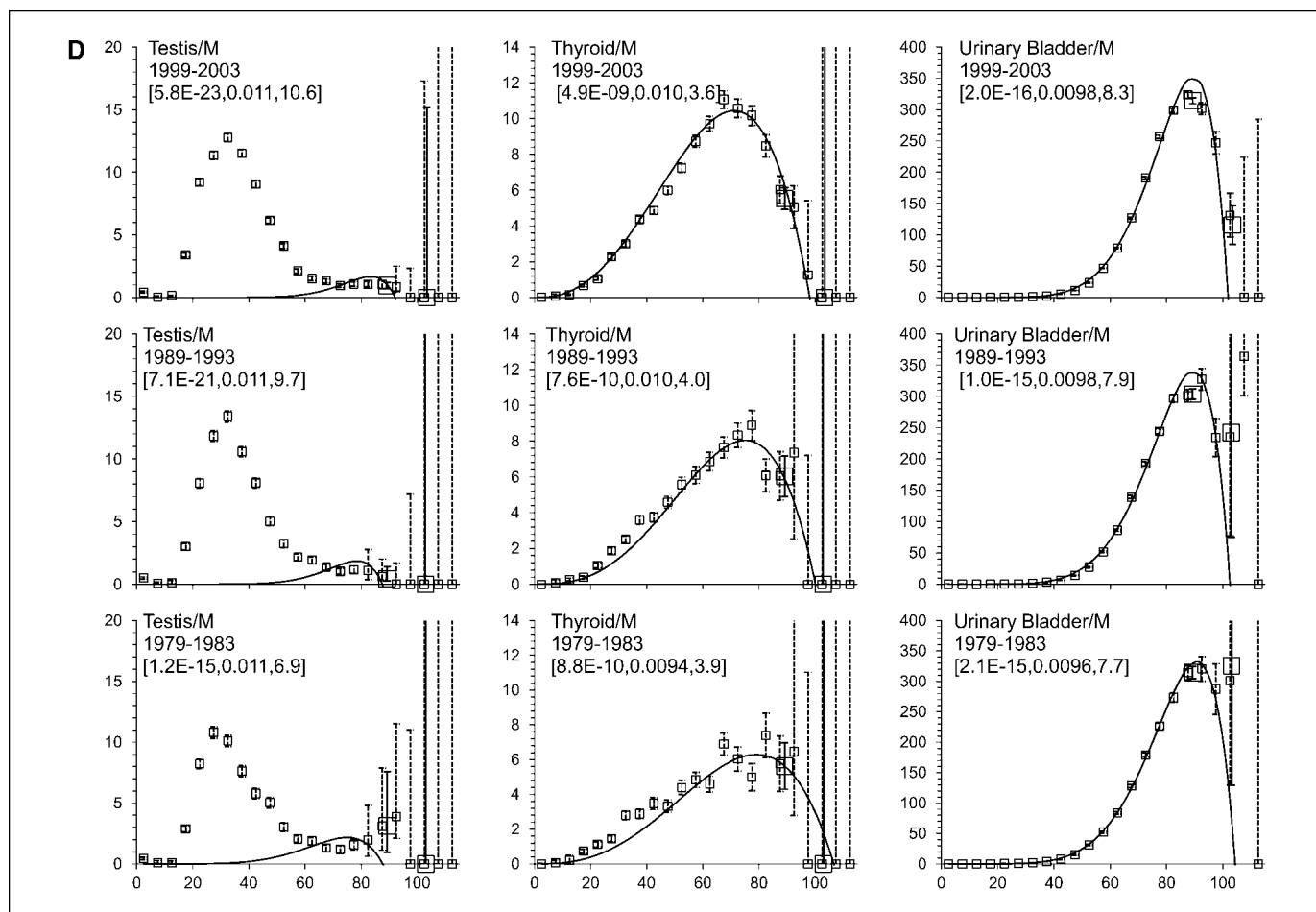


Figure 2 Continued. D, age-specific cancer incidence (per 100,000) for each of three major organ sites for males, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - 1(1 - bt)$. On each of the graphs, the fitted values are $[a, b, k]$. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

late-stage carcinogens at older ages. This would suggest a reduction in probability by $(1 - bt)$ or about a factor of 10 at age 90. Although this is possible, this factor would also have to be roughly constant over the 25 years of our study. As noted in the Introduction, Cook and colleagues and Ritter and colleagues considered, and believed unlikely, an explanation based on variation in sensitivity to cancers. Recently, Morgenthaler and colleagues (31) proposed a "fraction at risk" description of the susceptible population for mathematical modeling of cancer mortality reduction at old age, which included the suggestion that the nonsusceptible population might be simply assumed to be selected at random. This is effectively a stochasticity assumption, described as susceptibility. Svensson and colleagues (32) applied a variation in susceptibility model to colorectal cancer data in Norway, in which a distribution of high-risk and low-risk individuals was assumed. This resulted in age-specific cancer incidence leveling and reducing with age, as the high-risk individuals acquire the disease, leaving the healthier individuals cancer-free. Similarly, applying methods of randomly assigning variations in relative susceptibility at birth to model mortality, as proposed by Vaupel and colleagues (33), is similar. None of these three articles avoid the problems that Cook and Ritter pose. Although we cannot exclude the possibility that variations in susceptibility can explain our results, the requirements, noted in the Introduction, that the data impose are strict and we have found

no consistent model of such susceptible fraction variations that satisfies these strict requirements.

Our late colleague Sir Richard Doll wrote in the conclusion of a 2004 commentary (34) accompanying a commemorative republishing of his 1954 seminal article laying the foundation for the multistage interpretation of cancer formation: "The fact that only, say, 20% of heavy cigarette smokers would develop lung cancer by 75 years of age in the absence of other causes of death does not mean that 80% are genetically immune to the disease any more than the fact that usually only one cancer occurs in a given tissue implies that all the stem cells in the tissue that have not given rise to a malignant clone are also genetically immune. What it does mean is that whether an exposed subject does or does not develop a cancer is largely a matter of luck; bad luck if the several necessary changes all occur in the same stem cell when there are several thousand such cells at risk, good luck if they don't." Even if genetic or other information will eventually describe the unknown variables in the multistage model, they are now unknown and they must be considered stochastically.

We suggest that there may be a dominant biological mechanism that reduces all cancer incidence to zero near the end of a human lifetime. Ideally, we would extend the Armitage-Doll model by adding a term derived from a biological model. Our first attempt is

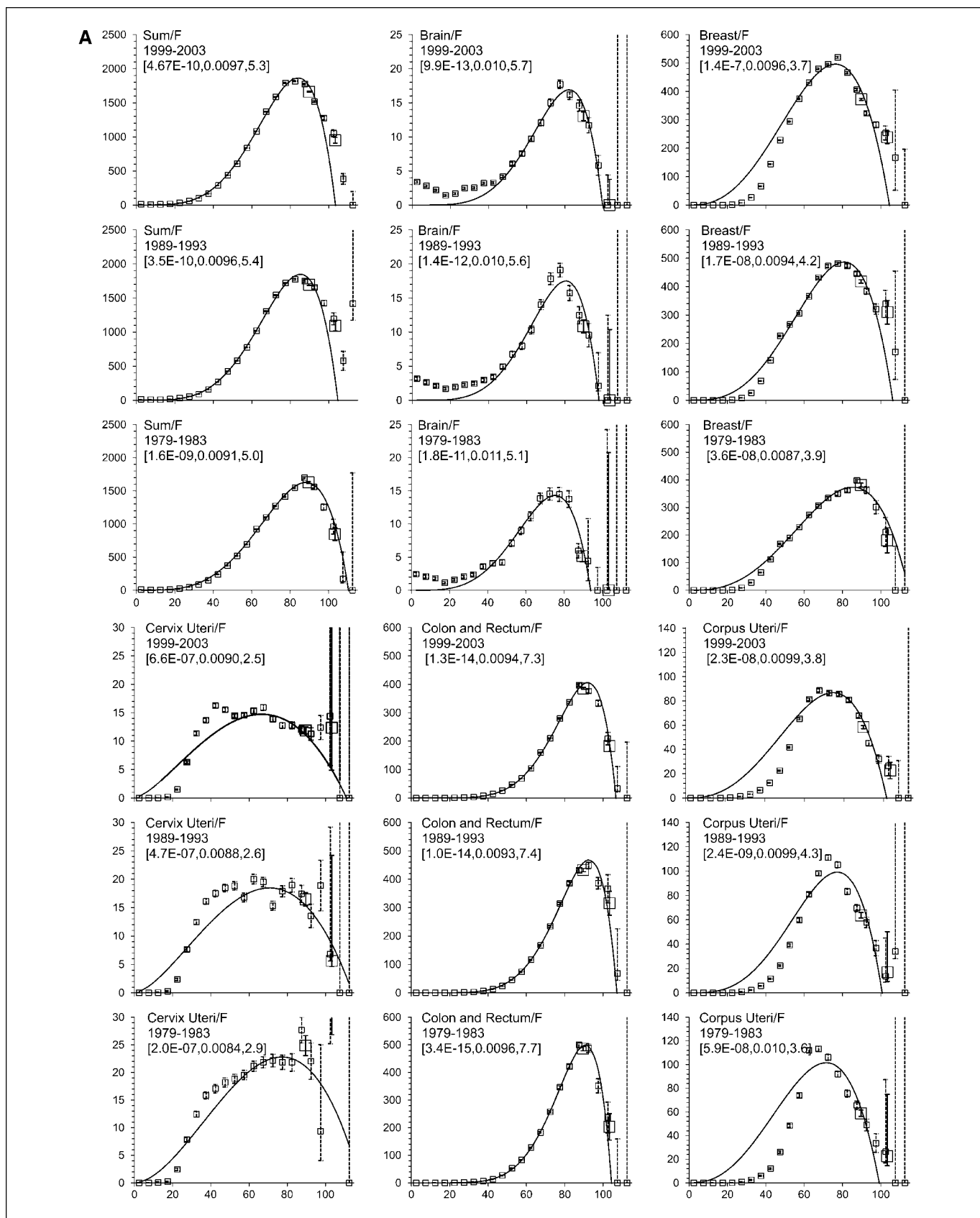


Figure 3. A, age-specific cancer incidence (per 100,000) for total cancers and each of five major organ sites for females, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^{k-1}(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

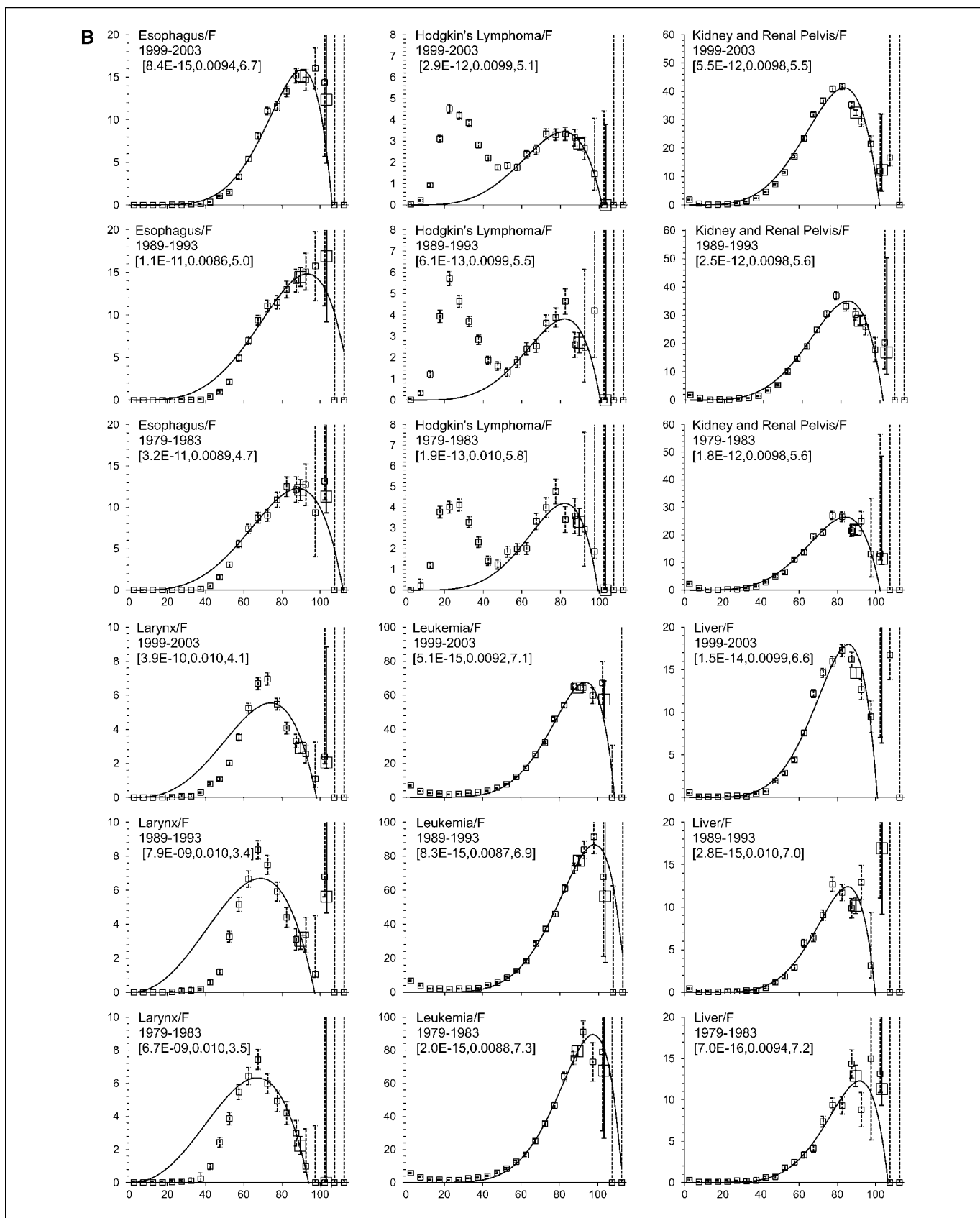


Figure 3 Continued. B. age-specific cancer incidence (per 100,000) for each of six major organ sites for females, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - 1(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

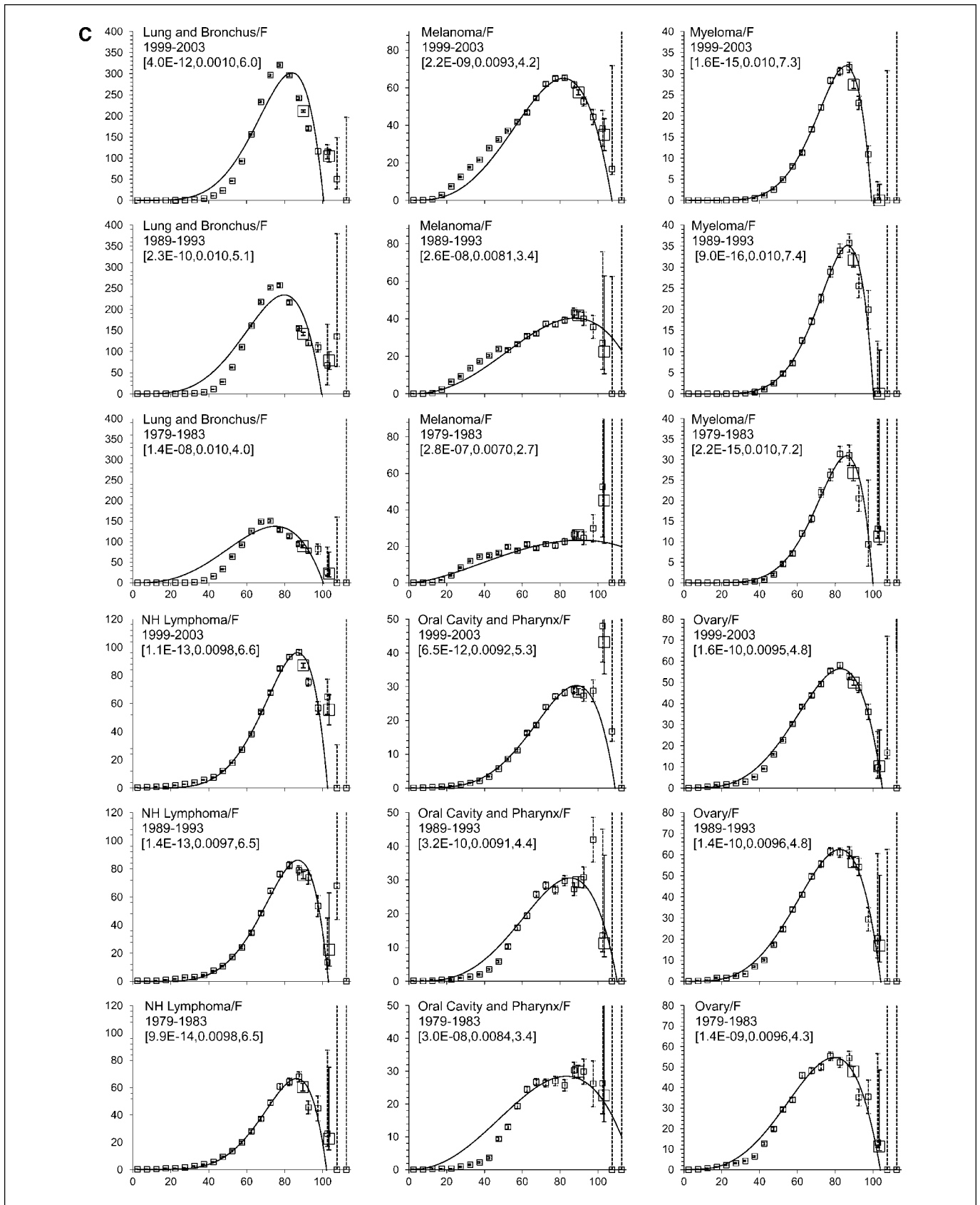


Figure 3 Continued. C. age-specific cancer incidence (per 100,000) for each of six major organ sites for females, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - 1(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

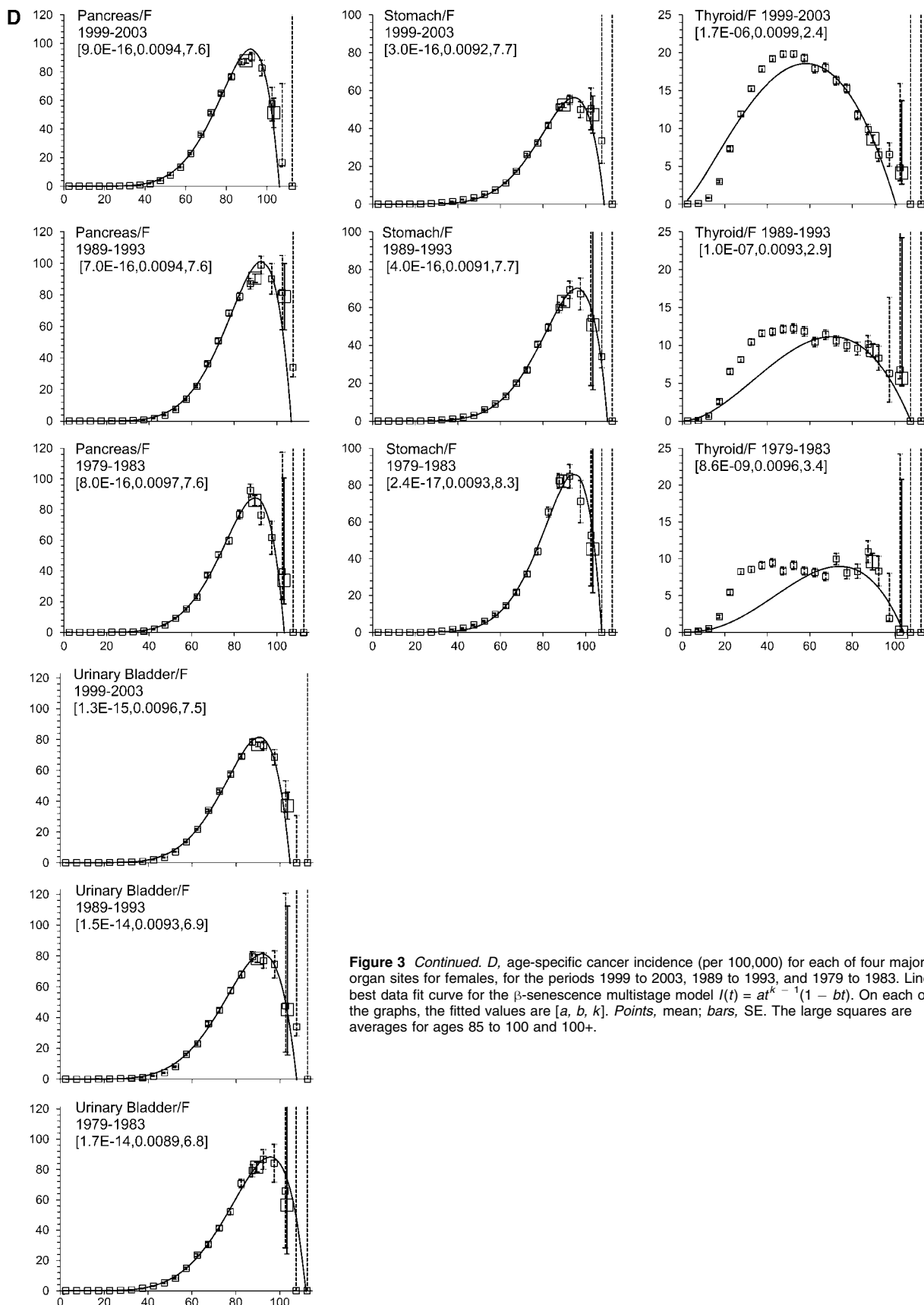


Figure 3 Continued. D, age-specific cancer incidence (per 100,000) for each of four major organ sites for females, for the periods 1999 to 2003, 1989 to 1993, and 1979 to 1983. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - 1(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

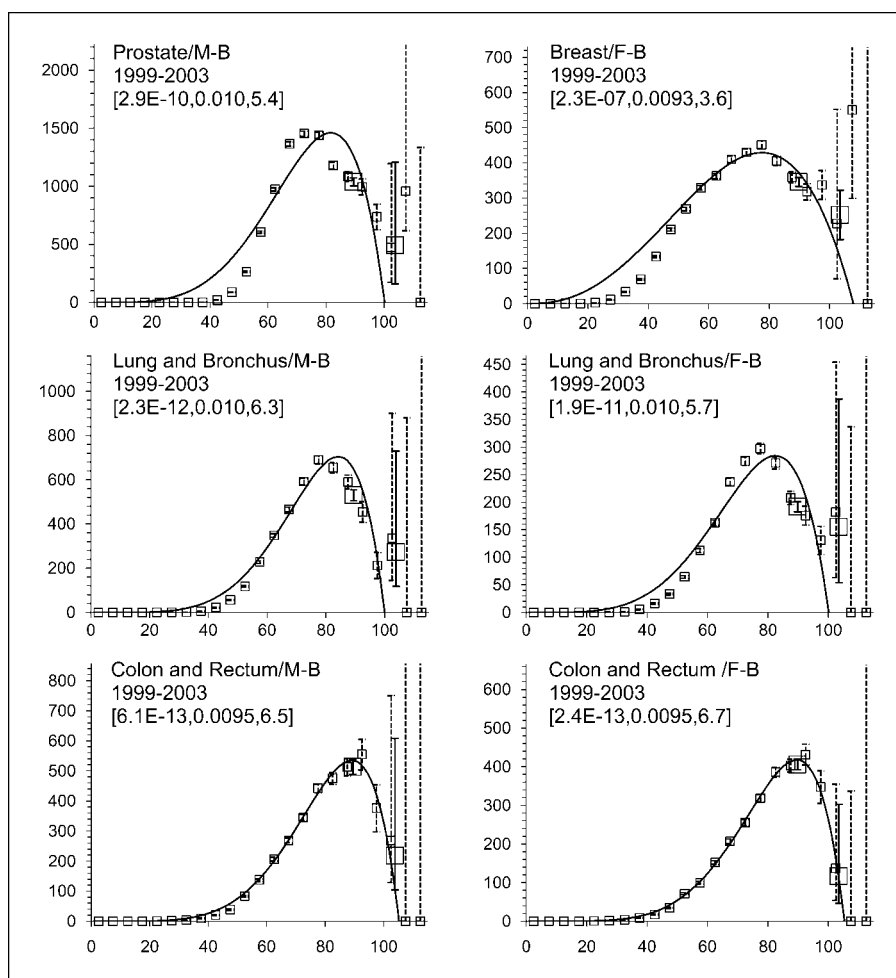


Figure 4. Age-specific cancer incidence (per 100,000) for six major organ sites for black males and females for the period 1999 to 2003. Line is best data fit curve for the β -senescence multistage model $I(t) = at^k - 1(1 - bt)$. On each of the graphs, the fitted values are [a, b, k]. Points, mean; bars, SE. The large squares are averages for ages 85 to 100 and 100+.

by the term $(1 - bt)$ derived by our speculative theory of the way cellular senescence might operate in people. That the formula fits the data reasonably well only states that it cannot be excluded. Other biological models have been suggested but we have located none that has yet been put into a mathematical form that can be tested in the same way.

Macieira-Coelho (35) proposes that cancer incidences are sensitive to the developmental stage of the organism. In the context of the mathematical multistage model, this corresponds to one or more of the probabilities of the transition from one stage to the next varying with developmental stage (or age). A strong age dependence at age above 80 would be required. The data suggest that, for 36 of 41 cancers studied over 25 years, there seems to be a common and highly uniform final stage that suppresses cancer from a peak at approximately age 80 to zero at age ~ 100 . Accordingly, we do suggest that all age-related cancer incidences fall to zero near the end of a human life span is not coincidental, but that both are related to the same cause, which in turn is related to age.

Discussions of the effect of interventions (i.e., negative from pollutants or positive from medical treatment) inherently involve time (age) dependence of transition probabilities and thus are similar to the idea of developmental stage effects. This has not been extensively studied in the context of the multistage model, although the expressions “late-stage” cancer-causing pollutants

and “early-stage” cancer-causing pollutants have been common. That cellular transitions occur with higher likelihood at certain times is consistent with the idea of several stages occurring in a specific order with age. Although this is not required by the multistage theory of cancer, it is consistent with it. It is important to note that epidemiologic cancer incidence data do not provide precision for comparing with cellular theories of cancer because the cancer is recorded when a person has symptoms and is motivated to seek medical care or appear during routine screening tests. The inherent assumption is that the cancers of concern are those that become medically relevant. Accordingly, the mathematical modeling inherently assumes that the time to creation of a medically relevant cancer from the enabling cellular alterations is either (a) short compared with the causal cellular events or (b) is itself a stage.

There is a practical importance in attempting to understand cancer incidence in the oldest. The main observation of the earlier studies and of this study is that cancer incidence falls to zero at approximately the end of a human life span. This suggests that there is some coupling between the fall off and life span. According to the simple senescence theory we proposed, cancer is reduced by increasing senescence. At the end of a life span, most cells become senescent, which prevents both cancer and the repair processes that support organ function. If true, interventions that have been observed to alter longevity and cancer, such as *p53* gene alterations

(36) and dietary restrictions, must be considered with both objectives in mind. It would seem appropriate to consider whether other biological explanations would also couple cancer and longevity. Reducing cancer by an intervention that reduces life span by increasing senescence or by other causes may be counterproductive. It becomes important, when attempting to reduce cancers by chemical intervention, to ensure that the life span is not thereby reduced. Recent investigators considering targeted therapies that “age” tumors into senescence are now aware of this potential problem (17, 37). Conversely, we suggest that, if it is possible to reduce the senescence rate to extend longevity, any concomitant increase in cancer might be limited in other ways such as dietary, environmental, or lifestyle improvements or other intervention of modern medical science. Ideally, we would find some intervention that both reduces cancer and extends life span, such as ensuring adequate mitochondria throughout life, being actively pursued by Yang and colleagues (38).

Historically, Armitage and Doll, Moolgavkar, and many others were able to contribute important ideas to the biology of cancer by modeling of the rising incidence of cancer with age. We suggest that studying the falling of cancer incidence with age above age 80 may similarly contribute to our basic understanding.

Appendix

In each year under study, we used all available SEER*Stat regions/registries except Alaskan Natives, employing 9 registries for 1979 to 1991, 12 registries for 1992 to 1999, and 16 registries for 2000 to 2003. The SEER 9 registries are Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, and Utah. SEER 12 additionally includes Los Angeles, San Jose-Monterey, and Rural Georgia. SEER 16 adds Greater California, Kentucky, Louisiana, and New Jersey. Sources used to partition the SEER-provided 85+ population are described as follows.

Given weaknesses in the accuracy of previous census data for the oldest ages, the census bureau has been especially careful in compiling the Census 2000 age counts (22), which are given by single year of age until 100, and then in two 5-year age groups and a 110+ category. “Postcensal residential population estimates” (26, 27) for 1980 to 2003 terminate with a 100+ age grouping. Those covering 1973 to 1979 aggregate everyone 85 and over. However, the National Center for Health Statistics’ Decennial Life Tables (for 1969–1971, 1979–1981, and 1989–1991) provide stationary (closed) population survival estimates up to 110+ (23–25). These we have extended to all years 1970 to 1990 by locating estimates at the decade and filling in interim year figures through linear trend from preceding to succeeding life table figures. Life table figures were unavailable for 1999 to 2001. For each applicable data source, all 1-year age categories above 85 were summed into 5-year age groups 85 to 89, 90 to 94, and so on.

Through the following methods, we reexpressed figures as proportions of the SEER 85+ populations: Let $C_n(2000)$ be the census 2000 total U.S. resident population of age group n ; let $P_n(t)$ be the postcensal U.S. resident population estimate of age group n in year t ; let $L_n(t)$ be the life table stationary population survival of age group n in year t . Considering each year, for each of $C_n(2000)$, $P_n(t)$, and $L_n(t)$ with n 85 to 89 and over, we took the proportion of the age group population n divided by the sum

(for that data source and year) over all n 85 and over. Respectively, call the resulting proportions $C_n(2000)$, $P_n(t)$, and $L_n(t)$. SEER registries were matched to analogous Census 2000 areas variously of state type, county type, and “urban area” type. Registry by registry, census populations were normalized to corresponding SEER populations. Registry by registry, discrepancies between total census populations and total SEER populations were always negligible (<1%).

Let $A_{n,r}(2000)$ be the normalized year 2000 area population of age group n and registry r ; let $C'_{n,r}(2000)$ be the census 2000 total U.S. resident population of age group n and registry r . For each age group and registry, a proportion $A_{n,r}(2000)$ was obtained by dividing $A_{n,r}(2000)$ by $C'_{n,r}(2000)$. Let $D_n(t)$ be the sum of $A_{n,r}(2000)$ over all used r in the t in question. For each $D_n(t)$, a proportion $d_n(t)$ was obtained by dividing $D_n(t)$ by the sum of $D_n(t)$ over all groups n 85 and over. Multiplying each $d_n(t)$ by each $C_n(2000)$, each $d_n(t)$ by each $P_n(t)$, and each $d_n(t)$ by each $L_n(t)$, and multiplying each resulting product by $S(t)$, the SEER-given 85+ population figures of year t give a host of population estimates. Counts using $C_n(2000)$ were applied to all years 2000 to 2003. Counts using $L_n(t)$ were applied to all years 1973 to 1979. Counts using $P_n(t)$ were applied to all years 1980 to 1999, and 100+ categories of these years were themselves separated according to renormalized [to the $P_n(t)$ -derived 100+ populations] $l(100-104)(t)$, $l(105-109)(t)$, and $l(110+)(t)$. Note that $L_n(t)$ for 1991 to 1999 were obtained by filling in interim year figures through linear trend from the preceding $l_n(1990)$ to succeeding $C_n(2000)$. In this manner, we obtained population estimates from 0 to 110+ by 5-year age group.

The estimated partition of population from 85 to 110+ is a potential weakness due to census reporting problems for the oldest. Yet, note that for all ages, 1999 to 2003 populations are effectively a direct application of Census 2000’s published regional population figures to corresponding SEER-given regional populations. Consider any year in this cross-section. Take the ratio of any one under 85 age group population to another. Take this same ratio in any other year of the cross-section. The discrepancy between such ratios is less than $\sim 15\%$. Therefore, thus far, as this lack of variation under 85 denies extreme variation in the population over 85, 1999 to 2003 population figures should be well founded from 85 to 110+.

For the population of black Americans in the 1999 to 2003 cross-section, an identical procedure was used with the exception that 1999 population values were also obtained via Census 2000 proportions (i.e., without linear trend life table or postcensal estimates). Additionally, it is important to note that SEER race recodes differ from census 2000 values. SEER data provided white, black, or other. Census data included a greater division of “other” and also a biracial category. However, as the total “black alone” census population was never <90% of the total “black” SEER populations, census populations were normalized to SEER populations without further recoding/correction. This may introduce some bias into the population figures. Patient identity is not disclosed by SEER, only a (coded) patient ID number.

Acknowledgments

Received 5/6/2007; revised 11/29/2007; accepted 3/31/2008.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank the SEER staff for their full cooperation, Dr. Dmitri Burmistrov for his help in the early stages of this work, and the cancer experts, too numerous to name, who have commented on earlier versions.

References

1. Hartveit F. Clinical and post-mortem assessment of the cause of death. *J Pathol* 1977;123:193-210.
2. DePinho RA. The age of cancer. *Nature* 2000;408:248-54.
3. Armitage P, Doll R. The age distribution of cancer and a multistage theory of carcinogenesis. *Br J Cancer* 1954; 8:1-12.
4. Storm HH. Cancer registries in epidemiological research. *Cancer Causes Control* 1996;7:299-301.
5. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care* 2002;40:IV-3-IV-18.
6. Bach PB, Guadagnoli E, Schrag D, et al. Patient demographic and socioeconomic characteristics in the SEER-Medicare database. *Med Care* 2002;40:IV-19-IV-25.
7. Surveillance, Epidemiology, and End Results (SEER) Program [internet]. Version 6.2.4. Bethesda (MD): National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch c2006-. [Database], Incidence - SEER 17 Regs Limited-Use, Nov 2005 Sub (1973-2003 varying) - Linked To County Attributes - Total U.S., 1969-2003; April 2006, based on November 2005 submission [cited 2007 Jan]. Available from: <http://seer.cancer.gov/seerstat/> by downloading SEER*Stat software.
8. Saltzstein SL, Behling CA, Baergen RN. Features of cancer in nonagenarians and centenarians. *J Am Geriatr Soc* 1998;46:994-8.
9. Pompei F, Wilson R. Age distribution of cancer: the incidence turnover at old age. *Hum Ecol Risk Assess* 2001;7:1619-50.
10. Pompei F, Lee EE, Wilson R. Cancer turnover at old age. *Nat Rev Cancer* [internet]. 2004 [cited 2007 Jan];3. Available from: <http://www.nature.com/nrc/journal/v3/n5/full/nrc1073-c2.html>.
11. Stanta G, Campagner L, Cavallieri F, Giarelli L. Cancer of the oldest old. What we have learned from autopsy studies. *Clin Geriatr Med* 1997;13:55-68.
12. Imaida K, Hasegawa R, Kato T, et al. Clinicopathological analysis on cancers of autopsy cases in a geriatric hospital. *Pathol Int* 1997;47:293-300.
13. Moolgavkar SH, Knudson AG. Mutation and cancer: a model for human carcinogenesis. *J Natl Cancer Inst* 1981;66:1037-52.
14. Cook PJ, Doll R, Fellingham SA. A mathematical model for the age distribution of cancer in man. *Int J Cancer* 1969;4:93-112.
15. Ritter G, Wilson R, Pompei F, Burmistrov D. The multistage model of cancer development: some implications. *Toxicol Ind Health* 2003;19:125-45.
16. Finkel AM. A quantitative estimate of the variations in human susceptibility to cancer and its implications for risk management. In: Olin S, Farland W, Park C, Rhomberg L, Scheuplein R, Starr T, Wilson J, editors. *Low dose extrapolation of cancer risks: issues and perspectives*. Washington (DC): ILSI Press; 1995. p. 297-328.
17. Bonafe M, Barbi C, Storci G, et al. What studies on human longevity tell us about the risk for cancer in the oldest old: data and hypotheses on the genetics and immunology of centenarians. *Exp Gerontol* 2002;37:1263-71.
18. Pompei F, Wilson R. A quantitative model of cellular senescence influence on cancer and longevity. *Toxicol Ind Health* 2002;18:365-76.
19. Macieira-Coelho A. The decline of the clinical incidence of cancers during human senescence. *Gerontology* 2003;49:341-9.
20. Pompei F, Polkanov M, Wilson R. The age distribution of cancer in mice: the incidence turnover at old age. *Toxicol Ind Health* 2001;17:7-16.
21. Surveillance Research Program, National Cancer Institute. SEER*Stat [software on the internet]. Version 6.2.4. Bethesda (MD): National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch; c2006. Available from: <http://seer.cancer.gov/seerstat/>.
22. U.S. Census Bureau. Census 2000 summary file 1: 100 percent data, detailed tables [database on the internet]. Washington (DC): U.S. Census Bureau; 2001-. [Tables], PCT12 and PCT12b; [cited 2006 Dec]. Data generated by Charles Harding using American FactFinder in Dec 2006. Available from: <http://factfinder.census.gov>.
23. Centers for Disease Control and Prevention; National Center for Health Statistics. U.S. decennial life tables for 1969-71, vol. 1, no. 1. Hyattsville (MD): Centers for Disease Control and Prevention; 1975. Available from: <http://www.cdc.gov/nchs/products/pubs/pubd/lftbls/lftbls.htm>.
24. Centers for Disease Control and Prevention; National Center for Health Statistics. U.S. decennial life tables for 1979-81, vol. 1, no. 1. Hyattsville (MD): Centers for Disease Control and Prevention; 1985. Available from: <http://www.cdc.gov/nchs/products/pubs/pubd/lftbls/lftbls.htm>.
25. Centers for Disease Control and Prevention; National Center for Health Statistics. U.S. decennial life tables for 1989-91, vol. 1, no. 1. Hyattsville (MD): Centers for Disease Control and Prevention; 1997. Available from: <http://www.cdc.gov/nchs/products/pubs/pubd/lftbls/lftbls.htm>.
26. Population Estimates Program, Population Division; U.S. Census Bureau. Monthly resident population estimates of the United States: July 1, 1990 to July 1, 1999, with short-term projection to July 1, 2000; July estimates [internet]. Washington (DC): U.S. Census Bureau [internet release date 2000 April 11; cited 2006 Dec]. Available from: <http://www.census.gov/population/estimates/nation/intfile2-1.txt>.
27. Population Estimates Program, Population Division; U.S. Census Bureau. Quarterly intercensal resident population estimates of the United States: 1980 to 1989; July estimates [internet]. Washington (DC): U.S. Census Bureau [internet release date 2000 April 11; cited 2006 Dec]. Available from: http://www.census.gov/popest/archives/1980s/80s_nat_detail.html.
28. Pezzullo JC. Exact binomial and Poisson confidence intervals, confint.xls [software on the internet]. [place unknown]; Pezzullo, John C; c2007 [updated 2007 Aug 11; cited 2007 Sept]. Available from: <http://statpages.org/confint.html>.
29. Feldman GJ, Cousins RD. Unified approach to the classical statistical analysis of small signals. *Phys Rev D* 1998;57:3873-89.
30. de Rijke JM, Schouten LJ, Hillen HF, Kiemeny LA, Coebergh JW, van den Brandt PA. Cancer in the very elderly Dutch population. *Cancer* 2000;89:1121-33.
31. Morgenthaler S, Herrero P, Thilly WG. Multistage carcinogenesis and the fraction at risk. *J Math Biol* 2004; 49:455-67.
32. Svensson E, Moger TA, Tretli S, Aalen OO, Grotmol T. Frailty modelling of colorectal cancer incidence in Norway: indications that individual heterogeneity in risk is related to birth cohort. *Eur J Epidemiol* 2006;21: 587-93.
33. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 1979;16:439-54.
34. Doll R. Commentary: the age distribution of cancer and a multistage theory of carcinogenesis. *Int J Epidemiol* 2004;33:1183-4.
35. Macieira-Coelho A. Neoplastic disease through the human life span. *Biogerontology* 2001;2:179-92.
36. Tyner SD, Venkatchalam S, Choi J, et al. p53 mutant mice that display early ageing associated phenotypes. *Nature* 2002;415:45-53.
37. Sage J. Making young tumors old: a new weapon against cancer? *Sci Aging Knowledge Environ* 2005;2005: pe25.
38. Yang H, Yang T, Baur JA, et al. Nutrient-sensitive mitochondrial NAD⁺ levels dictate cell survival. *Cell* 2007;130:1095-107.