A New Method of Predicting US and State-Level Cancer Mortality Counts for the Current Calendar Year

Ram C. Tiwari, PhD; Kaushik Ghosh, PhD; Ahmedin Jemal, PhD, DVM; Mark Hachey, MS; Elizabeth Ward, PhD; Michael J. Thun, MD, MS; Eric J. Feuer, PhD

ABSTRACT Every January for more than 40 years, the American Cancer Society (ACS) has estimated the total number of cancer deaths that are expected to occur in the United States and individual states in the upcoming year. In a collaborative effort to improve the accuracy of the predictions, investigators from the National Cancer Institute and the ACS have developed and tested a new prediction method. The new method was used to create the mortality predictions for the first time in Cancer Statistics, 2004 and Cancer Facts & Figures 2004. The authors present a conceptual overview of the previous ACS method and the new state-space method (SSM), and they review the results of rigorous testing to determine which method provides more accurate predictions of the observed number of cancer deaths from the years 1997 to 1999. The accuracy of the methods was compared using squared deviations (the square of the predicted minus observed values) for each of the cancer sites for which predictions are published as well as for all cancer sites combined. At the national level, the squared deviations were not consistently lower for every cancer site for either method, but the average squared deviations (averaged across cancer sites, years, and sex) was substantially lower for the SSM than for the ACS method. During the period 1997 to 1999, the ACS estimates of deaths were usually greater than the observed numbers for all cancer sites combined and for several major individual cancer sites, probably because the ACS method was less sensitive to recent changes in cancer mortality rates (and associated counts) that occurred for several major cancer sites in the early and mid 1990s. The improved accuracy of the new method was particularly evident for prostate cancer, for which mortality rates changed dramatically in the late 1980s and early 1990s. At the state level, the accuracy of the two methods was comparable. Based on these results, the ACS has elected to use the new method for the annual prediction of the number of cancer deaths at the national and state levels. (CA Cancer J Clin 2004;54:30–40.) © American Cancer Society, 2004.

INTRODUCTION

Approximately one in every four deaths in the United States is due to cancer, making it the second-leading cause of death after heart disease. For 2003, the estimate of all cancer deaths for both sexes combined is 556,500.1,2 Billions of dollars are spent annually on research, treatment, prevention, and other costs related to the disease. Thus, for effective planning, resource allocation, and communication about cancer, it is critical to have accurate estimates of the number of cancer cases and deaths occurring in the current year.

In this issue of CA and in Cancer Facts & Figures, the American Cancer Society (ACS) publishes estimates of the number of cancer cases and deaths for the current year, based on projections from observed data that ended three
years in the past. For example, the published estimates of cancer deaths in Cancer Facts & Figures 2003 were based on projections from mortality data from 1979 to 2000.\textsuperscript{2} The estimated number of deaths projected to occur in individual states and in the nation is based on underlying cause-of-death data reported to the National Center for Health Statistics (NCHS). Since 1995, the ACS has used a method of estimating the number of deaths in the current year that applies statistical projections and subjective judgment to adjust for recent changes in death rates.

In this article, we briefly describe the ACS method and present an alternative method for projecting the number of cancer deaths in the current year. This new method is called the state-space method (SSM). To compare the predictive accuracy of the SSM and the ACS method, we used both methods to estimate deaths for the years 1997 to 1999, based on data that would have been available at the time. We compared these estimates to the actual number of deaths that occurred from specific cancers and from all cancers combined at the state and national levels. Although neither method was consistently more accurate for every cancer site, in general the SSM performed better. The time-varying coefficients give added flexibility to this model, allowing it to adjust to rapidly evolving trends. Based on the results of this collaboration with the National Cancer Institute, the ACS has elected to use the SSM model in predicting the number of cancer deaths in Cancer Facts & Figures 2004 and in Cancer Statistics, 2004.\textsuperscript{3,4}

DATA DESCRIPTION

Mortality data for the United States are compiled by the NCHS of the Centers for Disease Control and Prevention. The information is based on the cause of death reported on death certificates. Because of the multistep process of data collection, tabulation, and publication and the large number of records involved, the latest mortality data available to the public from the NCHS are three years old. Thus, an accurate method of projection is needed to estimate the cancer burden in the current year. Our analysis considers a novel statistical approach for making such projections. In the discussion, we note the additional possibility of using preliminary mortality estimates produced by the NCHS since 1995, which may allow for the use of two- rather than three-year predictions.

The previous ACS method has been used since 1995.\textsuperscript{5} This approach considers cancer mortality data from 1979, the first year deaths were coded according to the ninth revision of the International Classification of Diseases (ICD–9), through the most recent year for which the mortality data is available, ie, three years before the current calendar year. Our analysis uses mortality data from 1969 to 1999, using ICD-8 through ICD-10 codes.\textsuperscript{5–8} These codes were made comparable using Surveillance Epidemiology and End Results (SEER) program recode.\textsuperscript{9}

MATERIALS AND METHODS

Mortality Prediction Methods

Wingo, et al.\textsuperscript{5} have described the ACS method in a previous publication, and the mathematical derivation of the SSM is explained in other publications.\textsuperscript{10}

Quadratic Time Series Model (the Previous ACS Method)

The statistical method previously used by the ACS to estimate the number of deaths from cancer during 1995 to 2003 is a multistep procedure.\textsuperscript{5} First, a quadratic time trend is fitted to the mortality counts, based on the mortality data from 1979 to the most recently available year. This models the long-term variation in trend as a function of time and the square of time. Second, the discrepancies between the fitted and observed values, called residuals, are calculated. An autoregressive process is then fitted to the residuals obtained from the first step. This process models the residuals at any point of time as a function of those at previous time points and accounts for the short-term
behavior of the mortality trend. In the third and final step, the combined model is projected three years into the future. The model fitting and prediction are implemented using PROC FORECAST in SAS software. This gives three-year-ahead predictions and 95% prediction intervals for the mortality counts. The default setting (which is the one used by the ACS) of PROC FORECAST requires at least seven years of data to successfully fit the model and then predict into the future.

Because the previous statistical approach does not account for recent changes in cancer death rates, it has been combined with subjective judgments about which of five alternative predictions seem most plausible. These include the three-year-ahead point prediction, the upper and lower 95% prediction limits and the midpoints between these prediction limits, and the point estimate. The value selected is then rounded to either the nearest 10s or the nearest 100s in the most recent projections.

The method described above is used to obtain both the state- and national-level predictions. For the state-level predictions, an additional restriction is imposed that the predictions for the 50 states and the District of Columbia must add up to the national-level prediction. Where there is a discrepancy between the two, the difference is proportionally allocated to the states and these revised numbers become the state-level predictions. It should be noted that the projection methods are applied at the national level for a large number of cancer sites. For site groupings (digestive or respiratory systems, for example, as in Cancer Facts & Figures 2003, page 4) and all sites combined, the predictions from individual sites are summed. At the state level, fewer individual cancer sites (see Cancer Facts & Figures 2003, page 5) and all the cancers combined are modeled directly. We follow the same procedures for our new proposed method.

State-Space Method

As an alternative to the previous method, we have used the SSM to predict mortality counts. The motivation for developing this method was to improve sensitivity to short-term trends and to eliminate subjectivity. The SSM model consists of two main parts: a measurement equation and the transition equation. In the measurement equation, the mortality counts follow a linear model with time-varying regression coefficients (also known as the state vector of parameters, because all the information about the current state of the process is assumed to be present in this vector). These time-varying regression coefficients also follow a linear model. The two equations combined give a quadratic trend over short time segments, in contrast to the ACS model, which assumes a quadratic trend over the entire time period.

In addition, a transition equation is used to model the dependence of the current state to its immediate past, using another linear model. Both the measurement and transition equations are assumed to incorporate random errors. When the transition error variances are zero, the SSM reduces to fitting a quadratic time trend with independent error terms, mimicking the ACS model. In this sense, the SSM represents a generalization of the previous ACS model. The error variances are estimated from the data, using the method of moments described in Ghosh, et al. The SSM model is more flexible than a standard regression model, because the former has time-varying coefficients, allowing the model to adjust to sudden changes in the observed data. In contrast, the standard (fixed coefficient) regression model tries to fit one curve through the entire data set, which may not provide a good fit if multiple rapid changes occur. Details of how the optimal estimate of the current state vector is computed are provided elsewhere.

The sensitivity of the SSM to sudden changes in the data can be disadvantageous when real or random variations in the observed series give rise to a zigzag curve. As a compromise between the accuracy of fit and projections, we have modified the SSM by introducing two tuning parameters. In the adjusted model, the error variances are rescaled by two constants, one each for the measurement error and transition error. These tuning parameters are estimated by minimizing the sum of squares of the differences between the observed
number of deaths and their three-year-ahead predictions.

To illustrate, suppose that mortality data from 1969 to 2000 are to be used to predict cancer deaths in 2003. The new method uses a procedure to estimate the tuning parameters as follows: Data to 1975 are used to obtain the three-year-ahead prediction for 1978, and data to 1976 are used to obtain the prediction for 1979, for example. Proceeding like this yields the prediction for 2000. For each year the prediction error is estimated by computing the predicted minus the observed value. The sum of the squares of the prediction errors is obtained by adding the squares of these prediction errors for the years 1978 to 2000, providing a measure of discrepancy of the observed values from the projected values. The tuning parameters are adjusted to produce the smallest sum of the squares of the prediction errors. This final model is then used to fit the 1979 to 2000 mortality counts data and to project three years ahead to obtain the prediction for 2003.

The expressions for estimates of the state vector and the tuning parameters are obtained using an iterative numeric procedure. We have used “R” software to code the procedure just described.13,10

In our comparison, the estimates from the previous ACS method, denoted CFF (for Cancer Facts & Figures), are based on data from 1979 onward and include the subjective choice among the five possible candidates. Estimates obtained directly from PROC FORECAST, denoted PF, are based on data from 1969 onward, as are the estimates made using the SSM model.

RESULTS

Model Comparison and Results

Figures 1 and 2 compare the predicted number of deaths from lung cancer in women and prostate cancer in men obtained from the two statistical methods, PF and SSM, without any subjective selection of the most plausible estimate from PF. In Figure 1A, we used the data on observed number of female lung cancer deaths 1969 to 1994 to fit the two models, and we extrapolated one-, two-, and three-year-ahead projections for 1995, 1996, and 1997. Both models fit the observed data (1969 to 1994) very well. However, the predictions from the SSM model for 1995 to 1997 are closer to the observed data than are the predictions from the PF. Furthermore, as expected, the predictions from both models veer further away from the observed as time progresses. Because we did not want our validation to depend entirely on what was happening in any particular calendar year, the analysis was repeated for subsequent years. In Figure 1B, we used observed data 1969 to 1995 to extrapolate mortality numbers for 1996, 1997, and 1998, and in Figure 1C we used observed data from 1969 to 1996 to extrapolate mortality numbers for 1997, 1998, and 1999. Finally, Figure 1D shows how the actual projections would occur in practice, projecting out to the future where we currently have no data to validate the results. This panel uses the most recent available data (1969 to 2000) at the time this report was written and projects through 2003. Data from 1969 to 2001 are used in CFF 2004 to project through 2004.

Figures 2A to D illustrate the comparison of the SSM and PF models for prostate cancer. Before 1990, the SSM fits are more erratic than are those from PF, because of sensitivity to random error. However, after 1990, the projections from the SSM model are much closer to the observed values than are those from the PF method.

Table 1 compares the projected mortality counts for 1999, based on three methods (PF, CFF, and SSM) with the observed counts for eight cancer sites each in men and women. All the methods use observed data through 1969 and project to 1999 using three-year-ahead projections. Among the 16 sex-specific and site-specific projections, the SSM-generated projections were closer to the observed for nine sites, CFF-generated projections for five sites, and PF-generated projection for one site (PF and SSM were tied for one site).

Figures 3 and 4 compare the predictions from the PF, CFF, and SSM models for the years 1995 to 1999 for selected cancer sites.
using data up to three years before the prediction year. To make these figures comparable, although on different scales, the vertical axes are all drawn approximately ±25% from the average observed values. On this relative scale, the SSM follows the observed trend closer than do the other methods for male and female lung cancer, female breast cancer, male colorectal cancer, and prostate cancer. The best prediction overall is for colorectal cancer, followed by lung, breast, and prostate. Note that the PF and CFF methods in female colorectal cancer consistently underestimate whereas the SSM estimates bounce over and under the observed values. The extra variability of the SSM model in female colorectal cancer is a reflection of variation in past observations.

Table 2 shows a similar comparison for all cancer sites combined. The SSM model gives better predictions compared with both the PF and CFF methods for the years 1997 to 1999.

Using the squared deviation as the measure of error between the observed and predicted death counts, we compared the accuracy of the PF, CFF, and SSM methods. These quantities are non-negative and become larger with increased discrepancy, giving a proportionally greater penalty to larger discrepancies. We believe that this measure of deviation is appropriate because a large error seems more serious than several smaller ones. We calculated squared deviations of PF, CFF, and SSM from the corresponding observed values for each of the three-year predictions for the years 1997, 1998, and 1999 for the comprehensive set of cancer sites reported in Cancer Facts & Figures. The
three-year predictions use data from 1969 until 1994, 1995, and 1996, respectively (with the exception of CFF, which uses data only from 1979). Then, for a particular site, the deviations so obtained are averaged for the three years. Table 3 shows the results for selected cancer site and sex combinations. In general, the average squared deviations for SSM are smaller than those for the PF and CFF methods.

Table 4 shows a summary of how the methods perform for a three-year period. The entries in the table are the squared deviations averaged over the comprehensive set of cancer sites reported in Cancer Facts & Figures.

In addition, to determine whether one method is better depending on the rarity of cancer deaths, we grouped all the sites into four categories according to the number of observed deaths in 1999. We averaged the squared deviations over all the cancers in a certain group for the years 1997 to 1999. Table 5 shows the results. In general, SSM performs better regardless of the rarity of the cancer.

Finally, Table 6 shows summary statistics for the comparison of the PF, CFF, and SSM methods when applied to the state-level data. The predicted values were adjusted so that the sum of the state predictions matches the corresponding national prediction. For each of the cancer sites listed, we averaged the squared deviations over all 50 states and the District of Columbia for the years 1997 to 1999.

Table 6 shows that the PF method performs better than the SSM at the state level (which in turn performs better than the CFF method), although the improvement is slight in most cases.
DISCUSSION

Our comparison of three methods of predicting deaths from cancer in the current year indicates that the SSM model performed better than any of the other methods at the national level. At the state level, the accuracy of the SSM and PF models were comparable, with a slight edge for PF over SSM. This is probably because the increased sensitivity of SSM to recent trends may sometimes cause it to react more quickly to random fluctuations in the smaller number of deaths in individual states. Although one approach could be to use SSM at the national level and PF at the state level, the decision was to use one single method, namely SSM, at both the state and national levels starting in Cancer Facts & Figures 2004 and Cancer Statistics, 2004.3,4

Despite the decrease in mortality rates that has occurred for several major cancer sites since the 1990s, a corresponding decline in the number of cancer deaths has not occurred.14 This is because the increase in the size and age structure of the population offsets the decrease in age-specific death rates. However, the number of deaths from cancer has increased more slowly than the size of the population because of the decline in age-specific rates. For example, the number of men in the United States increased 41% during the period of our analyses, from 97,884,292 in 1969 to 138,053,563 in 2000. During this interval, the number of deaths from colon cancer in men increased 29% from 22,044 in 1969 to 28,484 in 2000. However, the age-adjusted mortality rate (based on 2000 US standard) actually declined from 33.2 per 100,000 in 1969 to 25.2 per 100,000 in 2000.

The PF method is slower to adapt to the change in death counts, which occurred in the 1990s, because of the fixed regression coefficients. The subjective selection among the five data points generated by the PF, used in the CFF, did not adjust for recent trends as well as the SSM did. However, in some cases, such as female colorectal cancer, this sensitivity may cause results to fluctuate year to year more than is desired. This demonstrates the difficulty of deriving a method that is optimal in all situations.

| Cancer Sites / Sex Combinations          | Observed | 3-YR Predicted Values |
|-----------------------------------------|----------|-----------------------|
|                                         |          | PF       | CFF      | SSM      |
| Colon and Rectum (Females)              | 28,909   | 27,957   | 28,800   | 26,589   |
| Lung and Bronchus (Females)             | 62,662   | 69,421   | 68,000   | 62,658   |
| Melanoma of the Skin (Females)          | 2,686    | 2,821    | 2,700    | 2,807    |
| Breast (Females)                        | 41,144   | 45,345   | 43,300   | 42,746   |
| Ovary                                   | 13,627   | 14,317   | 14,500   | 13,284   |
| Hodgkin Lymphoma (Females)              | 618      | 715      | 600      | 656      |
| Non-Hodgkin Lymphoma (Females)          | 11,008   | 12,472   | 12,300   | 12,126   |
| Acute Lymphocytic Leukemia (Females)    | 582      | 707      | 600      | 591      |
| Colon & Rectum (Males)                  | 28,313   | 27,831   | 27,800   | 27,793   |
| Lung & Bronchus (Males)                 | 89,399   | 93,619   | 90,900   | 89,257   |
| Melanoma of the Skin (Males)            | 4,529    | 4,745    | 4,600    | 4,760    |
| Prostate                                | 31,728   | 39,027   | 37,000   | 32,733   |
| Testis                                  | 378      | 321      | 300      | 321      |
| Hodgkin Lymphoma (Males)                | 785      | 892      | 700      | 623      |
| Non-Hodgkin Lymphoma (Males)            | 11,794   | 13,692   | 13,400   | 13,163   |
| Acute Lymphocytic Leukemia (Males)      | 779      | 866      | 800      | 789      |

PF: 3-year-ahead predicted value from PROC FORECAST.
CFF: Cancer Facts & Figures 1999, using data from 1979–1996.
SSM: 3-year-ahead predicted value from State Space Method.
BOLD numbers indicate best prediction.
Source: American Cancer Society. Cancer Facts & Figures, 2003.
Yet another method for estimating the number of cancer deaths is based on fitting a joinpoint (also known as changepoint) model to the available mortality data and then extrapolating it to future years. The Joinpoint regression software developed by the National Cancer Institute is based on the permutation test approach and is available at the Web site http://srab.cancer.gov/software/joinpoint.html. This method is used to characterize cancer trends in the United States. It fits a series of joined linear segments, usually on a log scale, to a data series. When fitted on a log scale, the slope of each segment can be characterized as an annual percentage change in the rates. The number and locations of the joinpoints are determined using a series of statistical tests, called permutation tests. The joinpoint method is useful to identify changes in the trend throughout the data series, although changes near the end of the series are usually of most interest.
Because this method is sensitive to changes in trend near the end of the series, it is a natural candidate for short-term extrapolation. Although the joinpoint method did not do as well overall as the SSM, it did reasonably well, especially for moderately rare cancer sites (i.e., those with 2,000 to 10,000 deaths), and for some cancer sites it performed better than any other method. In another report we compare in detail the joinpoint method with the others presented here.10

We continue to examine statistical methods that can reliably predict cancer incidence.
and death rates. Based on the results of this validation study, however, the ACS has elected to use the SSM to project mortality counts for Cancer Facts & Figures 2004. Another possible improvement for future estimates may be to use preliminary estimates of mortality, which are available from the NCHS approximately one year before the final estimates. This would allow two-year rather than three-year projections. These preliminary estimates (which have been available since 1995) have been shown to be generally within a few percentage points of the final estimates for most cancer sites at the national level. Studies to validate the use of these preliminary estimates are ongoing.

### TABLE 4 Average Squared Deviations Summarized by Year and Sex Across Cancer Sites*

| Category                        | Average Squared Deviations |
|---------------------------------|---------------------------|
|                                 | PF           | CFF         | SSM          |
| 1997—1999 (Females and Males combined) | 4,158,563 | 2,623,368 | 479,901       |
| Females 1997—1999               | 1,422,126 | 723,889    | 213,342       |
| Females 1997                    | 1,109,912 | 612,099    | 161,468       |
| Females 1998                    | 1,244,664 | 531,973    | 194,617       |
| Females 1999                    | 1,911,773 | 1,027,594  | 283,942       |
| Males 1997—1999                 | 1,880,298 | 1,588,336  | 213,528       |
| Males 1997                      | 1,629,660 | 2,364,223  | 273,434       |
| Males 1998                      | 1,744,912 | 1,390,087  | 226,434       |
| Males 1999                      | 2,266,322 | 1,010,699  | 140,715       |

PF: 3-year-ahead predicted value from PROC FORECAST.
CFF: Cancer Facts & Figures.
SSM: 3-year-ahead predicted value from State Space Method.
BOLD numbers indicate best prediction.
* Incorporates the comprehensive list used in Cancer Facts & Figures.
Source: American Cancer Society. Cancer Facts & Figures, 1997—1999.

### TABLE 5 Squared Deviations, Averaged Over 1997 to 1999 and Selected Cancer Sites

| Category                                           | Average Squared Deviations |
|----------------------------------------------------|---------------------------|
|                                                    | PF           | CFF         | SSM          |
| Both Sexes                                         |               |             |              |
| Cancer sites with less than 2,000 deaths in 1999 (17 sites) | 27,956       | 28,032      | 16,078       |
| Cancer sites with 2,000 – 10,000 deaths in 1999 (12 sites) | 189,037      | 114,777     | 76,261       |
| Cancer sites with 10,000 – 40,000 deaths in 1999 (11 sites) | 5,621,810    | 5,495,801   | 680,773      |
| Cancer sites with more than 40,000 deaths in 1999 (4 sites) | 38,045,561   | 18,188,504  | 3,263,855    |

PF: 3-year-ahead predicted value from PROC FORECAST.
CFF: Cancer Facts & Figures.
SSM: 3-year-ahead predicted value from State Space Method.
BOLD numbers indicate best prediction.
Source: American Cancer Society. Cancer Facts & Figures, 1997—1999.
To predict incidence counts for the nation, incidence must first be spatially projected from SEER to the nation, and then projected forward in time to the current calendar year. An improved method for spatial projections of incidence has been developed and work is in progress to incorporate the SSM model to add a temporal component to this spatial model.  

Acknowledgment

The authors thank Ray Chang of IMS, Inc. for his help in preparing the tables for this article.

Table 6: Squared Deviations Averaged Over 1997 to 1999 and States for Selected Cancers

| Cancer Sites          | 3-YR Average Squared Deviations |
|-----------------------|---------------------------------|
|                       | PF     | CFF    | SSM    |
| All Sites             | 105,072| 530,196| 146,897|
| Female Breast         | 4,837  | 9,347  | 7,331  |
| Colon and Rectum      | 5,973  | 10,543 | 6,235  |
| Lung and Bronchus      | 10,863 | 77,525 | 21,563 |
| Non-Hodgkin Lymphoma  | 4,008  | 6,605  | 3,953  |
| Pancreas               | 1,278  | 4,248  | 1,314  |
| Prostate               | 6,797  | 46,196 | 5,461  |

PF: 3-year-ahead prediction from PROC FORECAST.
CFF: Cancer Facts & Figures.
SSM: 3-year-ahead predicted values from State Space method.
BOLD numbers indicate best prediction.

Source: American Cancer Society. Cancer Facts & Figures, 1997—1999.

References

1. American Cancer Society. Cancer Facts & Figures 2003. Atlanta, GA: American Cancer Society, 2003.
2. Jemal A, Murray T, Samuels A, et al. Cancer statistics, 2003. CA Cancer J Clin 2003;53:5–26.
3. American Cancer Society. Cancer Facts & Figures 2004. Atlanta, GA: American Cancer Society, 2004. In press.
4. Jemal A, Tiwari RC, Murray T, et al. Cancer statistics, 2004. CA Cancer J Clin 2004;54:8–29.
5. Wingo PA, Lands S, Parker S, et al. Using cancer registry and vital statistics data to estimate the number of new cancer cases and deaths in the United States for upcoming year. J Reg Management 1998;25:43–51.
6. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Vol. 1, 8th revision. Geneva, Switzerland: World Health Organization; 1965.
7. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Vol. 1, 9th revision. Geneva, Switzerland: World Health Organization; 1975.
8. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Vol. 1, 10th revision. Geneva, Switzerland: World Health Organization; 1992.
9. Surveillance, Epidemiology and End Results (SEER) Program. Available at: http://seer.cancer.gov/codrecode. Accessed October 2003.
10. Ghosh K, Tiwari RC, Feuer EJ, et al. Predicting U.S. cancer mortality using state space models 2003. Manuscript in preparation.
11. SAS/ETS User’s Guide, version 6, 2nd ed. Cary, NC: SAS Institute; 1993.
12. Harvey AC. Time Series Models, 2nd ed. Cambridge, MA: MIT Press; 1994.
13. Ihaka R, Gentleman R. R: a language for data analysis and graphics. J Comp Graph Stat 1996;5:299–314.
14. Ries LAG, Eisner MP, Kosary CL, et al. (eds). SEER Cancer Statistics Review, 1975–2000. Bethesda, MD: National Cancer Institute; 2003.
15. Kim HJ, Fay MP, Feuer EJ, Midhune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335–351.
16. Joinpoint software. Available at http://srab.cancer.gov/software/joinpoint.html. Accessed October 2003.
17. Arias E, Smith BL. Deaths: preliminary data for 2001. National Vital Statistics Report 2003;51:1–44.
18. Pickle LW, Feuer EJ, Edwards BK. U.S. predicted cancer incidence, 1999: complete maps by county and state from spatial projection models. NCI Cancer Surveillance Monograph Series, number 5. Bethesda, MD: National Cancer Institute; 2003.