

# Updated Methodology for Projecting U.S.- and State-Level Cancer Counts for the Current Calendar Year: Part II: Evaluation of Incidence and Mortality Projection Methods



Kimberly D. Miller<sup>1</sup>, Rebecca L. Siegel<sup>1</sup>, Benmei Liu<sup>2</sup>, Li Zhu<sup>2</sup>, Joe Zou<sup>3</sup>, Ahmedin Jemal<sup>1</sup>, Eric J. Feuer<sup>2</sup>, and Huann-Sheng Chen<sup>2</sup>

## ABSTRACT

**Background:** The American Cancer Society (ACS) and the NCI collaborate every 5 to 8 years to update the methods for estimating the numbers of new cancer cases and deaths in the current year for the U.S. and individual states. Herein, we compare our current projection methodology with the next generation of statistical models.

**Methods:** A validation study was conducted comparing current projection methods (vector autoregression for incidence; Joinpoint regression for mortality) with the Bayes state-space method and novel Joinpoint algorithms. Incidence data from 1996–2010 were projected to 2014 using two inputs: modeled data and observed data with modeled where observed were missing. For mortality, observed data from 1995 to 2009, 1996 to 2010, 1997 to 2011, and 1998 to 2012, each projected 3 years forward to 2012 to 2015. Projection methods were evaluated using the average absolute relative deviation

(AARD) between observed counts (2014 for incidence, 2012–2015 for mortality) and estimates for 47 cancer sites nationally and 21 sites by state.

**Results:** A novel Joinpoint model provided a good fit for both incidence and mortality, particularly for the most common cancers in the U.S. Notably, the AARD for cancers with cases in 2014 exceeding 49,000 for this model was 3.4%, nearly half that of the current method (6.3%).

**Conclusions:** A data-driven Joinpoint algorithm had versatile performance at the national and state levels and will replace the ACS's current methods.

**Impact:** This methodology provides estimates of cancer data that are not available for the current year, thus continuing to fill an important gap for advocacy, research, and public health planning.

## Introduction

In January of each year, the American Cancer Society (ACS) releases projected numbers of new cancer cases and deaths for the current calendar year for the total U.S. and individual states. These estimates are based on methods codeveloped by ACS and the NCI, which are updated every 5 to 8 years. This currently requires a projection of 4 years ahead for incidence and 3 years ahead for mortality from the most recent observed data, which lag behind the current year due to the time required for data collection, cleaning, and dissemination. Mortality can be forecast from complete long-term observed data, but gaps in the availability of incidence data require a two-step process in which complete cancer counts are first estimated and then projected ahead (1). In the first paper in this series (referred to as Paper I), we evaluated approaches for the spatiotemporal prediction of incidence

for the most recent 15 years of data in the United States (2). The second component of the methodology, evaluated herein, is a temporal projection of data 4 years ahead for incidence and 3 years ahead for mortality.

The ACS has published contemporary invasive cancer estimates since 1951, and over this time, the statistical models used to perform the time series projection have rapidly evolved as the coverage of population-based cancer registries has expanded and novel techniques have become available (3–7). The current methodologies based on the analyses of Chen and colleagues (2012) and Zhu and colleagues (2012; refs. 3, 7), hereafter collectively referred to as the *Cancer* 2012 methodology papers, are vector autoregression via Hilbert–Huang transform (HHT) for incidence and a Joinpoint regression (JP) model for mortality. These papers also established the use of 15 data years for the time series input, using observed data for mortality and modeled data for incidence. This is notable because invasive cancer incidence data in the last 15 available data years are increasingly complete, raising the question of whether a combination of observed and modeled data where observed data were missing would provide better results for the projection models over using only modeled data alone, as has been done in the past.

In this article, we reevaluate our current methodologies and compare them with novel approaches in time series projection. We also evaluate our incidence methodology for whether all modeled data or a combination of observed data with modeled data where observed are missing provide more accurate results. Finally, we select a temporal projection method to be used for the ACS's cancer case and death estimates for 2021 and beyond.

<sup>1</sup>Surveillance and Health Equity Science, American Cancer Society, Kennesaw, Georgia. <sup>2</sup>Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, Maryland. <sup>3</sup>Information Management Services, Inc., Rockville, Maryland.

**Note:** Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

**Corresponding Author:** Kimberly D. Miller, Surveillance and Health Equity Science, American Cancer Society, Inc, Atlanta, GA 30303. Phone: 404-327-6471; E-mail: kimberly.miller@cancer.org

Cancer Epidemiol Biomarkers Prev 2021;30:1993–2000

doi: 10.1158/1055-9965.EPI-20-1780

©2021 American Association for Cancer Research

## Materials and Methods

### Data sources

Data for the evaluation of the temporal projection for incidence were obtained by cancer type, sex, and state from the final model selected in Paper I (2) and included observed and predicted invasive cancer cases diagnosed in 1996 to 2014 for 50 states and the District of Columbia. However, four of these states (KS, MN, NM, NV) were excluded from the analysis for Paper II because they were missing observed data for 2014, the year used for the evaluation of the temporal projection results. Mortality data on cancer deaths occurring during 1995 to 2015 were provided by the National Center for Health Statistics, obtained via SEER\*Stat software, and were available for all 50 states and the District of Columbia for all years (8).

For analysis at the national level, we used all cancer sites combined and 47 individual cancer sites for incidence and 46 individual cancer sites for mortality (colon and rectal cancer deaths are combined in mortality data due to the large proportion of rectal deaths misclassified as colon). For the state-level data, we selected all cancer sites combined and the 21 most common cancers, corresponding to those sites for which ACS publishes annual state-level estimates (Supplementary Table S1). Cases were classified according to the *International Classification of Diseases for Oncology*, 3rd edition (recoded as necessary via the SEER site recode variable; ref. 8), whereas deaths were classified according to the *International Classification of Diseases*, 10th revision (9).

A series of 15 data years was used as the input for all projections because we previously found that a longer time series often did not provide superior projection results for mortality in the *Cancer* 2012 methodology (3). In addition, limiting the input to 15 years for incidence allows for the use of largely complete observed data for the underlying computation, as most registries in the United States were reporting to national programs by the late 1990s. For incidence, the input years for model validation included data from 1996 through 2010; the delay-adjusted observed data for 2014 were used for validation. Because the landscape for incidence data completeness has rapidly changed in recent years, incidence data were input in two different ways: (i) all modeled data using predicted counts provided by Paper I and (ii) a combination of observed and modeled data (hereafter referred to as “mixed”), where modeled data were only used when observed data were missing. To test the sensitivity of different projection methods to temporal differences in input data, we used four different series of data for the mortality validation (1995–2009 with projection to 2012, 1996–2010 with projection to 2013, 1997–2011 with projection to 2014, and 1998–2012 with projection to 2015).

### Statistical projection methods

For incidence, three different methods were compared: (i) vector autoregression with HHT (VAR); (ii) JP; and (iii) the Bayes state-space (BSS) method. For mortality, Joinpoint regression and the BSS were compared. A number of variations were used for the Joinpoint and VAR models that are described in detail below.

### Vector autoregressive model via HHT (VAR-None, VAR-Trend)

This approach first decomposes the input data via a method called the HHT using empirical model decomposition (3, 7), which translates incidence data into multiple data vectors (intrinsic mode functions, or IMF) that combined sum to the original input. These IMFs are projected individually via a time series projection method (vector autoregression) and are afterward summed to obtain the predicted count. The vector autoregressive (VAR) model is a multivariate

version of an autoregressive model for correlated individual time series. An R software package “vars” is used for obtaining the forecast counts. Within the “vars” package, it is possible to specify whether a stochastic or deterministic forecast should be produced, either by setting the type to “none” or to “trend,” respectively. Herein, we used VAR-None and VAR-Trend to represent these options, respectively. Notably, the VAR was not evaluated for mortality as a result of the method’s poor performance in the *Cancer* 2012 methodology, which was likely due to the lower variability in mortality patterns as compared with incidence (3).

### Joinpoint regression (JP-MBIC APC/AAPC, JP-WBIC APC/AAPC, JP-WBIC-alt APC/AAPC)

Joinpoint software version 4.7.0.0 (10) and SAS 9.4 (11) were used to perform all Joinpoint-based projections. Joinpoint regression summarizes trends in cancer incidence and mortality by modeling rates as a function of time using piecewise linear segments on a log scale. The model is fitted by the weighted least squares method up to a specified number of change points (Joinpoints). Predicting future rates can be done by extending the fitted model using the slope of the last segment. When modeled in log scale, the slope of a segment is equivalent to the annual percent change (APC) for a single segment. Alternatively, the average APC (AAPC) represents a weighted average of APCs for a specified time period potentially spanning over several Joinpoint segments, with the weights equal to the length of each segment (12).

In the *Cancer* 2012 methodology studies (3, 7), it was shown that projection accuracy depends on the model selection criteria used to select the number of Joinpoints, and it was determined that the current mortality projection method should be based on the modified Bayesian information criterion (MBIC). Since then, a few new model selection procedures have been added to the Joinpoint software, including a weighted Bayesian information criterion (WBIC) and WBIC-Alt, an alternate WBIC. These algorithms are described more in detail in the Supplementary Data. In this study, we evaluate the prediction accuracy based on these new procedures, as well as reevaluate the performance of the MBIC, for both incidence and mortality. Another variation we considered for the first time in this analysis is the use of AAPC of the last 4 years for projection, in addition to the current method of using the APC of the last segment of the Joinpoint model, with the hypothesis that a 4-year AAPC may make the projection more stable than the APC of a final short Joinpoint segment only 3 years’ duration, as is possible with the APC method.

The Joinpoint model projections considered in our comparison include three selection algorithms for the number of Joinpoints: MBIC, WBIC, and WBIC-Alt; and two types of slope: APC and AAPC. In total, we considered six combinations of Joinpoint methods for incidence and mortality: MBIC-APC, MBIC-AAPC, WBIC-APC, WBIC-AAPC, WBIC-Alt-APC, WBIC-Alt-AAPC.

### BSS method

In the *Cancer* 2012 methodology, we evaluated the BSS method, which is a generalized linear model fitted using the Bayesian paradigm, assuming that for any year, the observed incidence or mortality count by sex and geographic level is a realization of a random variable, following a Poisson distribution (13). The BSS method is flexible and can easily adapt to changing trends. It is a stochastic model-based method that can capture sudden changes in trends and perform better in shorter time series and noisy data (3). We again applied this same methodology, which is described in detail in *Cancer* 2012 methodology (3). Detailed discussion of the BSS methods can be found in Schmidt and Pereira (14). R was used to perform all BSS projections.

**Method of evaluation**

Incidence and mortality data were evaluated separately. Cases and deaths were stratified by cancer type, sex, and state. For every cancer type/sex/state combination, we calculated the prediction error for each validation year (defined as 2014 for incidence and 2012–2015 for mortality) by computing the absolute value of the predicted minus the observed count. To account for the relative differences in observed case/mortality counts among different cancers and/or geographic areas, an absolute relative deviation (ARD, expressed as a percent) to the observed count is used:

$$ARD = \frac{|predicted\ count - observed\ count|}{observed\ count + c}$$

where *c* is a small, positive constant added to the denominator to avoid numerical error when the observed count was zero. We used the ARD as calculated in the *Cancer* 2012 methodology (with *c* = 0.5; refs. 3, 7) for each cancer type/sex/state combination to evaluate the performance of the projection methods. We used two metrics, first calculating the average ARD (AARD, expressed as a percent) and second assessing the dispersion of the ARDs. After evaluating the general performance of the models at the national and state levels using these metrics, we then closely examined the performance of the models for the most common cancers.

We identified the lowest AARD overall and for cancer types grouped by the number of observed cases/deaths in the validation year(s), both at the national and the state level. This was done to assess projection methods' performance by the magnitude of the observed case or death counts, a function of both the rarity of the cancer and, at the state level, population size. While also evaluating the overall fit, we considered methods within 10% of the model with the lowest AARD to be generally comparable. For incidence, AARDs were also assessed individually for modeled versus mixed input data for the projections. Mortality was additionally examined by validation year to confirm overall results were not unduly influenced by one projection year; due to the computational burden, only one validation year was used for incidence.

One goal of the current analysis was to select a model that produces less extreme aberrant results (among the many hundreds of cancer type/sex/state combinations) that need to be considered and adjusted on an individual basis. As such, we evaluated the variation of the ARDs and the magnitude of aberrant results for each method in box-and-whisker plots. The "box," or rectangular portion of the plot, indicates the interquartile range (IQR) of the ARDs, whereas the whiskers represent data points up to 1.5 IQR outside of the top and bottom quartiles. Outliers were identified as

falling outside of this range and were examined for frequency and magnitude of error.

As a final step, we examined the ARDs individually for the most common cancers for both incidence and mortality at the national level to assess the frequency at which each model produced the smallest ARD for these cancers. We gave overall preference in our final decision to models that had lower overall AARDs and the least extreme outliers for the most common cancers (i.e., the largest category of observed case or death counts) at the national level and that also produced reasonable results for other cancers at the national and state levels. The focus on national level results was primarily because these estimates are generally more stable than state-level estimates due to underlying variability in counts. More importantly, however, we focused specifically on the most common cancers because a large ARD for these types at the national level translates to a large absolute error in the number of projected cases or deaths, which would have important implications for the overall estimate of the cancer burden in the United States. Conversely, a larger relative error for rarer cancers and state-level estimates may still translate to a reasonable estimate in absolute terms.

**Results**

**Incidence**

**Table 1** shows the AARDs for all scenarios (cancer type–sex combinations) combined by method at the national and state level, organized by whether the data were all modeled versus mixed. At the national level, mixed and modeled input results were similar, with the exception of the JP-MBIC variations, for which the differences in the AARDs were approximately 10% in absolute terms. Overall, the JP-WBIC-AAPC model with mixed data input had the lowest overall AARD (4.9%); however, most methods, including the stochastic variation of the current methodology (VAR-None) were within 10% of this AARD, with the exception of the VAR-Trend, the BSS, and the JP-MBIC methods with mixed input data.

When the AARDs were stratified by observed 2014 case counts, notable differences emerged, particularly for the category including the most common cancers, that is, those with >49,000 reported cases in 2014 (**Table 2**). For this category, the JP-WBIC-AAPC using modeled input had the lowest AARD, largely as a result of most other models having a large ARD (>30%) for prostate cancer. However, the JP-WBIC-AAPC with modeled data continued to have then lowest AARD in this category even when prostate was excluded (2.5%), largely due to low ARDs for cancers of the male urinary bladder, female breast, and female lung (Supplementary

**Table 1.** AARDs (%) of the 4-year-ahead incidence projection for all cancer sites combined.

	VAR-None <sup>a</sup>	VAR-Trend <sup>a</sup>	JP-MBIC-AAPC	JP-MBIC-APC	JP-WBIC-AAPC	JP-WBIC-APC	JP-WBIC-AIt-AAPC	JP-WBIC-AIt-APC	BSS
National									
Modeled	5.2	6.0	5.2	5.0	5.4	5.3	5.4	5.4	7.5
Mixed	5.2	5.9	6.3	6.3	4.9 <sup>b</sup>	5.0	5.0	5.0	7.5
State									
Modeled	13.0 <sup>b</sup>	14.5	13.9	13.8	13.9	14.0	13.9	13.9	14.0
Mixed	16.6	21.2	16.4	16.1	20.4	19.7	20.7	19.6	19.8

Abbreviations: AARD, average absolute relative deviation; AAPC, average annual percent change; APC, annual percent change; BSS, Bayes state-space method; JP-MBIC, Joinpoint method with modified Bayesian information criterion; JP-WBIC, Joinpoint method with weighted BIC; VAR, vector autoregressive analysis method.

<sup>a</sup>Current method.

<sup>b</sup>Indicates the best fit for the national or state level. Methods within the 10% threshold of the best fit are indicated in gray highlight.

Downloaded from <http://aacrjournals.org/cebp/article-pdf/30/11/1993/3100430/1993.pdf> by guest on 25 June 2024

**Table 2.** National-level AARDs (%) of the 4-year-ahead incidence projection by 2014 observed case count.

	VAR-None <sup>a</sup>	VAR-Trend <sup>a</sup>	JP-MBIC-AAPC	JP-MBIC-APC	JP-WBIC-AAPC	JP-WBIC-APC	JP-WBIC-Alt-AAPC	JP-WBIC-Alt-APC	BSS
<3k (22)									
Modeled	6.5	6.2	7.6	7.3	8.5	7.8	8.3	8.2	5.7
Mixed	6.6	7.2	8.0	8.0	5.6	5.1 <sup>b</sup>	6.2	5.1 <sup>b</sup>	5.7
3k-<7k (20)									
Modeled	5.0	5.5	3.7	3.5	4.4	4.7	4.5	4.6	6.9
Mixed	4.8	6.1	3.8	3.8	3.5 <sup>b</sup>	3.6	3.5 <sup>b</sup>	3.6	6.9
7k-<2k (22)									
Modeled	4.2	6.5	4.0	3.7 <sup>b</sup>	4.3	4.1	4.4	4.2	7.1
Mixed	4.3	4.1	6.9	6.8	4.9	5.5	5.0	5.5	7.1
21k-<49k (14)									
Modeled	4.2 <sup>b</sup>	5.6	4.3	4.3	5.2	5.0	5.3	5.3	7.2
Mixed	4.8	6.0	6.4	6.4	4.6	4.8	4.4	4.8	7.2
49k or more (10)									
Modeled	6.3	6.0	6.4	6.4	3.4 <sup>b</sup>	3.7	3.5	3.6	6.9
Mixed	6.1	6.5	6.0	6.0	6.4	6.4	6.4	6.4	6.9

Note: Numbers in parentheses indicate the number of scenarios (cancer type–sex combinations) for that category.

Abbreviations: AARD, average absolute relative deviation; AAPC, average annual percent change; APC, annual percent change; BSS, Bayes state-space method; JP-MBIC, Joinpoint method with modified Bayesian information criterion; JP-WBIC, Joinpoint method with weighted BIC; VAR, vector autoregressive analysis method.

<sup>a</sup>Current method.

<sup>b</sup>Indicates the best fit by case count category across both mixed and modeled approaches. Methods within the 10% threshold of the best fit are indicated in gray highlight.

Table S2). However, some models provided better estimates for certain cancers, including the VAR-Trend with modeled data and the JP-WBIC-Alt-APC with modeled data.

The dispersion of the ARDs was similar for most categories, but the JP-WBIC-AAPC and the JP-WBIC-Alt-AAPC were the only models for which prostate cancer was not a major outlier. **Figure 1** shows the box-and-whisker plots of the ARDs for VAR-None (current methodology) and the JP-WBIC-AAPCs with modeled and mixed data inputs, which represents the models with the lowest overall AARD and the lowest AARD for the most common cancers at the national level.

At the state level, results differed substantially, with modeled data performing far superior to mixed data (**Table 1**). Of all approaches, the current methodology, VAR-None with modeled input, had the lowest overall AARD (13%). However, other methods with modeled input performed similarly, with almost all overall results (except the VAR-Trend and the BSS) within 10% of the lowest AARD. By the observed 2014 count, the VAR-None with modeled input had the lowest AARDs for cancers where the count was <750 cases (**Table 3**). The box plots by state for the VAR-None and JP-WBIC-AAPC methods are shown in the Supplementary Data (Supplementary Fig. S1).

### Mortality

**Table 4** shows the AARDs for scenarios at the national level for mortality, averaged over four individual 3-year ahead projections to 2012–2015. Overall, for all cancer types combined, the JP-WBIC-Alt-AAPC model had the lowest overall AARD (5.0%). However, all WBIC-based Joinpoint methods performed very similarly overall, with only the MBIC and the BSS methods falling outside the 10% threshold. The BSS approach performed the worst at the national level, with the largest AARD (6.9%).

When the AARDs were stratified by the magnitude of the observed death counts, the BSS provided the best fit for rare cancer types where the overall death count was <500, with an AARD of 9.2%. However, the AARDs for the Joinpoint methods were smaller

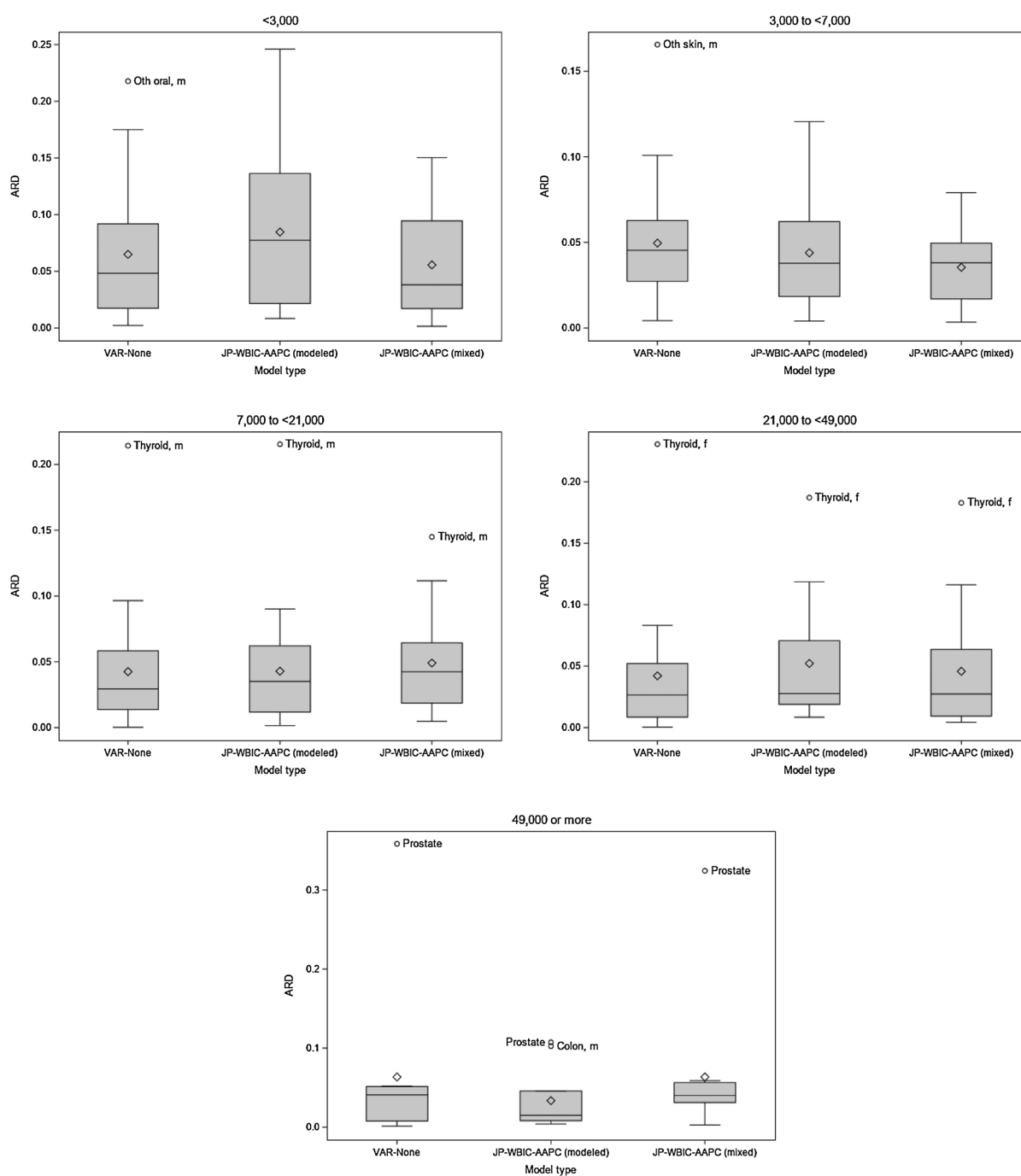
for more common cancer types and were generally within 10% of the smallest AARD by category. Notably, when examining the most common cancers (breast, lung, prostate, and colorectum), the WBIC methods consistently had the lowest AARDs or were within the 10% threshold. The WBIC-AAPC model performed the most consistently overall, falling within the 10% threshold for almost all categories examined. **Figure 2** shows the box-and-whisker plots of the national results for the current methodology, JP-MBIC-APC, and the JP-WBIC and JP-WBIC-Alt AAPC methods, which represent the models with the lowest AARD for the most common cancers and the lowest overall AARD, respectively.

At the state level, the BSS provided the lowest AARD (17.3%; Supplementary Table S3; Supplementary Fig. S2). The MBIC-based Joinpoint methods, including the current method, were also within the 10% threshold. However, the BSS method was only superior for the cancer types with fewer deaths overall and smaller states where the observed death count was <500. For all other categories, the current method (JP-MBIC-APC), as well as the AAPC variation, provided the lowest AARD or was within the 10% threshold.

## Discussion

We evaluated a number of time series projection methods for use by the ACS to estimate the number of new cancer cases and deaths in the current year and found that a single method, the JP-WBIC-AAPC, provided a good fit for both incidence and mortality at the national and state level. In addition, this method provided a very good fit for the most common cancers, for which reliable estimates are of the highest priority due to their importance for public health planning and resource allocation. When the models were compared in boxplot format, the JP-WBIC-AAPC also produced reasonably narrow IQRs as well as less extreme outliers.

However, each method has its own strengths and limitations, and it was impossible to select a single model that performed best for all cancer types and geographic levels. Several models provide valid alternatives in situations where the selected model underperforms.



**Figure 1.** National-level boxplots of incidence ARDs by model type. ARD, average relative deviation; F, female; m, male; oth skin, other nonepithelial skin. The diamond indicates the average ARD; the midline of each box indicates the median. The “whiskers” indicate the furthest ARD within 1.5 times the IQR. The small circles indicate outliers.

For example, several other Joinpoint models produced incidence estimates with smaller ARDs for male colon cancer than our selected model, particularly the JP-MBIC options with mixed data. The VAR-Trend and JP-WBIC-Alt-APC approaches with modeled data also provided lower ARDs for incidence of some common cancers at the national level.

Both of our previous models described in the *Cancer* 2012 methodology papers, VAR-None for incidence and JP-MBIC-APC for mortality, performed very well at the state level. In particular, the VAR-None had the lowest AARD overall at the state level for incidence and outperformed the other methodologies for all scenarios in which the observed case count in 2014 was <750. The VAR-None also

**Table 3.** State-level AARDs (%) of the 4-year-ahead incidence projection by 2014 observed case count.

	VAR-None <sup>a</sup>	VAR-Trend <sup>a</sup>	JP-MBIC-AAPC	JP-MBIC-APC	JP-WBIC-AAPC	JP-WBIC-APC	JP-WBIC-Alt-AAPC	JP-WBIC-Alt-APC	BSS
<25 (1,296)									
Modeled	<b>39.4<sup>b</sup></b>	43.6	<b>40.6</b>	<b>40.1</b>	46.9	46.2	46.9	46.4	<b>40.4</b>
Mixed	50.8	68.3	60.0	56.8	84.3	80.0	87.3	79.8	53.2
25-<75 (2,262)									
Modeled	<b>15.6<sup>b</sup></b>	17.3	<b>16.7</b>	<b>16.8</b>	<b>16.8</b>	<b>16.6</b>	<b>16.7</b>	<b>16.6</b>	<b>16.5</b>
Mixed	21.1	26.5	18.1	18.0	22.4	21.1	22.5	21.1	20.9
75-<250 (3,486)									
Modeled	<b>10.5<sup>b</sup></b>	<b>11.6</b>	<b>11.3</b>	<b>11.2</b>	<b>11.3</b>	<b>11.3</b>	<b>11.2</b>	<b>11.2</b>	<b>11.5</b>
Mixed	13.1	17.8	12.0	12.0	14.7	14.6	14.8	14.6	14.2
250-<750 (3,474)									
Modeled	<b>8.7<sup>b</sup></b>	9.8	9.7	<b>9.2</b>	<b>9.0</b>	<b>9.3</b>	<b>9.0</b>	<b>9.2</b>	<b>9.1</b>
Mixed	11.1	13.4	10.2	10.2	11.9	11.7	12.0	11.7	11.5
750-<2.5k (2,064)									
Modeled	<b>7.9</b>	10.0	8.6	8.5	<b>7.5<sup>b</sup></b>	<b>7.5</b>	<b>7.0</b>	<b>7.5</b>	<b>7.9</b>
Mixed	9.7	11.5	9.6	9.5	9.7	9.4	9.7	9.4	9.9
2.5k-<5k (564)									
Modeled	9.0	9.6	9.7	9.6	<b>7.3</b>	8.4	<b>7.2<sup>b</sup></b>	8.4	8.9
Mixed	11.2	12.3	9.8	10.2	8.7	8.8	8.9	8.6	11.4
5k or more (672)									
Modeled	7.2	6.6	7.3	7.5	<b>5.2</b>	5.7	<b>4.9<sup>b</sup></b>	5.7	6.7
Mixed	7.8	8.3	7.9	7.9	6.6	7.0	6.7	7.0	8.0

Note: Numbers in parentheses indicate the number of scenarios (cancer type–sex combinations) for that category. Abbreviations: AARD, average absolute relative deviation; AAPC, average annual percent change; APC, annual percent change; BSS, Bayesian state-space method; JP-MBIC, Joinpoint method with modified Bayesian information criterion; JP-WBIC, Joinpoint method with weighted BIC; VAR, vector autoregressive analysis method.

<sup>a</sup>Current method.

<sup>b</sup>Indicates the best fit by case count category across both mixed and modeled. Methods within the 10% threshold of the best fit are indicated in gray highlight.

performed comparatively well, on average, for incidence at the national level, with the lowest AARDs for 2014 case counts between 21,000 and 49,000. Similarly, for mortality, the JP-MBIC-APC performed well for all observed cancer death counts greater than 500.

As the collection of observed cancer incidence data become more complete (with fewer years missing because registries were not able to submit data deemed to meet quality standards), it is useful to know whether projections perform better when input data are all modeled or

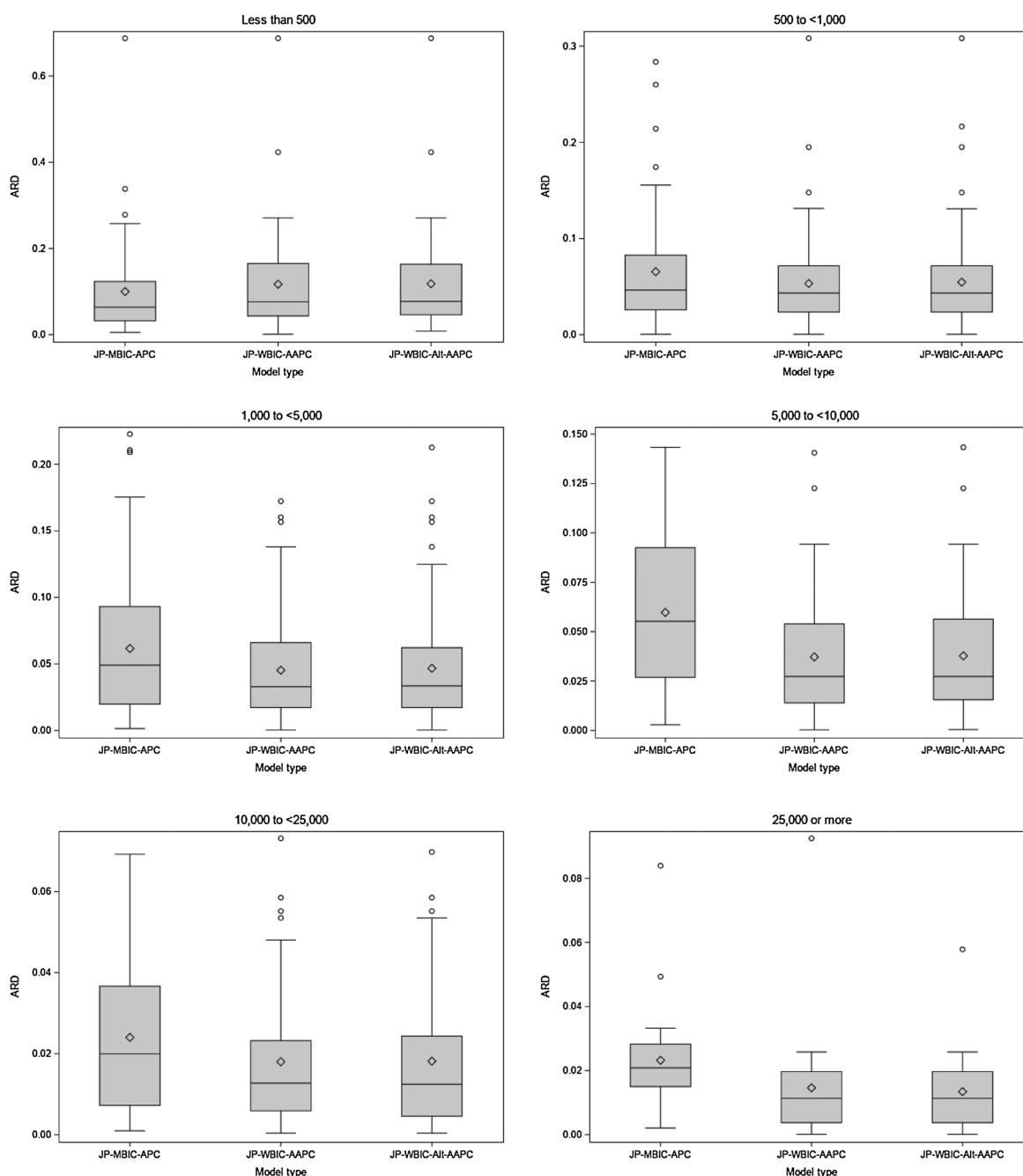
**Table 4.** National-level AARDs (%) of the 3-year-ahead mortality projection.

	JP-MBIC-APC <sup>a</sup>	JP-MBIC-AAPC	JP-WBIC-APC	JP-WBIC-AAPC	JP-WBIC-Alt-APC	JP-WBIC-Alt-AAPC	BSS
All scenarios combined	6.0	6.0	<b>5.1</b>	<b>5.1</b>	<b>5.1</b>	<b>5.0<sup>b</sup></b>	6.9
By data input years:							
1995–2009 → 2012	7.1	7.1	<b>5.7<sup>b</sup></b>	5.9	5.8	6.0	7.5
1996–2010 → 2013	4.8	4.8	4.3	3.9	4.0	<b>3.6<sup>b</sup></b>	5.8
1997–2011 → 2014	6.0	6.1	<b>4.7<sup>b</sup></b>	<b>4.7<sup>b</sup></b>	5.0	4.8	7.1
1998–2012 → 2015	6.0	6.0	<b>5.5<sup>b</sup></b>	5.7	5.7	5.7	7.1
Most common cancers:							
Breast (female)	2.4	2.1	<b>0.8<sup>b</sup></b>	0.9	1.3	1.4	2.7
Colorectum	2.8	2.8	2.1	2.1	<b>2.0<sup>b</sup></b>	<b>2.0<sup>b</sup></b>	2.9
Lung and bronchus	1.9	1.9	<b>1.6<sup>b</sup></b>	<b>1.6<sup>b</sup></b>	<b>1.6<sup>b</sup></b>	<b>1.6<sup>b</sup></b>	2.9
Prostate	3.8	3.8	<b>2.9<sup>b</sup></b>	<b>2.9<sup>b</sup></b>	<b>2.9<sup>b</sup></b>	3.8	4.9
By observed deaths (# scenarios)							
<500 (52)	<b>10.0</b>	<b>10.0</b>	11.5	11.2	11.9	11.7	<b>9.2<sup>b</sup></b>
500–<1k (70)	6.5	6.5	5.5	5.7	<b>5.2<sup>b</sup></b>	5.3	7.7
1,000–<5k (106)	6.2	6.2	4.6	<b>4.5<sup>b</sup></b>	4.7	<b>4.5<sup>b</sup></b>	6.8
5,000–<10k (50)	6.0	6.0	4.0	3.8	3.9	<b>3.7<sup>b</sup></b>	6.8
10k–<25k (46)	2.4	2.5	1.9	2.0	1.9	<b>1.8<sup>b</sup></b>	5.5
25k or more (28)	2.3	2.3	<b>1.3<sup>b</sup></b>	<b>1.3<sup>b</sup></b>	<b>1.3<sup>b</sup></b>	1.5	3.4

Abbreviations: AARD, average absolute relative deviation; AAPC, average annual percent change; APC, annual percent change; BSS, Bayes state-space method; JP-MBIC, Joinpoint method with modified Bayesian information criterion; JP-WBIC, Joinpoint method with weighted BIC; VAR, vector autoregressive analysis method.

<sup>a</sup>Current method.

<sup>b</sup>Indicates the best fit by row. Methods within the 10% threshold of the best fit are indicated in gray highlight.



**Figure 2.** National-level boxplots of mortality ARDs by model type. ARD, average relative deviation. The diamond indicates the average ARD; the midline of each box indicates the median. The “whiskers” indicate the furthest ARD within 1.5 times the IQR. The small circles indicate outliers.

a combination of both observed and modeled (where observed are missing) data. Most methods performed similarly at the national level regardless of whether modeled or mixed input was used, with the exception of the MBIC-based methods and some notable differences for the most common cancers. However, at the state level, performance was substantially better with all modeled input. Modeled data may perform better because they are largely “smoothed,” thus removing noise in the data that might be overexaggerated or underexaggerated

by certain projection methodologies. However, in some instances, this “smoothing” may attenuate and mask emerging trends. As such, while the modeled input generally provided the lowest overall AARDs, we believe using mixed input is a reasonable alternative approach, particularly with the JP-WBIC-AAPC method.

Limitations of our analysis for incidence include the use of a single validation year due to the time and computationally intensive process of producing the modeled input data in Paper I (2), as compared with

mortality, for which complete, observed data are readily available. This was a particularly important consideration for prostate, as case counts fluctuated substantially between 2010 and 2014 due to the U.S. Preventive Services Task Force recommendation against the routine use of the prostate specific antigen test for prostate cancer screening in 2012. Following this recommendation, prostate cancer rates dropped precipitously in a way that would have been impossible to predict in 2010 (15). Therefore, we also evaluated the relative errors for other common cancers individually to ensure our overall decision was not solely based on performance for prostate. This example likewise highlights the difficulty in predicting cancer occurrence for which abrupt fluctuations may occur or for which trends are emerging, which is a challenge regardless of the input and projection method selected. Another limitation in our selection of incidence input includes the use of pre-1999 data, for which there are several gaps in available data compared with later data years. However, this is unlikely to have adversely affected the performance of incidence projections that used modeled data only where observed data were missing, as many of the projection methods, particularly the Joinpoint models, emphasize recent trends where observed data predominate. Finally, there are a number of additional projection methods available that were not feasible or optimal to include in the current study. For example, we did not consider age-period cohort models (3, 16, 17) for both cases and deaths projection methods as they are generally optimized for longer time series data than we predetermined for this study (15 years).

In conclusion, we have demonstrated that it is reasonable to conservatively predict cancer occurrence in the short-term using a single method despite frequent fluctuations that may not follow deterministic patterns. However, while these models are robust, there continue to be challenges in projecting cases and deaths for cancers for which trends are emerging or for which patterns in occurrence may rapidly shift. Still, these estimates provide a reasonably reliable profile of the contemporary cancer burden and continue to be among the

most widely used estimates of cancer occurrence in the literature. As such, these estimates will continue to have broad applications in public health planning, advocacy, and patient and provider education. The JP-WBIC-AAPC method using the most recent 15 years of available data will be utilized for projections of incidence (with modeled input) and mortality starting with the 2021 edition of *Cancer Facts and Figures*. We feel this method was quite robust, performing well across the wide range of situations that are encountered, and should reduce the number of results that need to be considered and adjusted on an individual basis.

### Authors' Disclosures

R.L. Siegel reports employment with the American Cancer Society, which receives grants from private and corporate foundations, including foundations associated with companies in the health sector for research outside of the submitted work. However, R.L. Siegel is not funded by or key personnel for any of these grants and their salary is solely funded through American Cancer Society funds. J. Zou reports personal fees from NCI during the conduct of the study; personal fees from NCI outside the submitted work. No disclosures were reported by the other authors.

### Authors' Contributions

**K.D. Miller:** Formal analysis, visualization, methodology, writing—original draft. **R.L. Siegel:** Conceptualization, supervision, writing—review and editing. **B. Liu:** Data curation, writing—review and editing. **L. Zhu:** Data curation, writing—review and editing. **J. Zou:** Data curation, methodology, writing—review and editing. **A. Jemal:** Conceptualization, supervision, methodology, writing—review and editing. **E.J. Feuer:** Conceptualization, supervision, methodology, writing—review and editing. **H.-S. Chen:** Conceptualization, supervision, methodology, writing—original draft, writing—review and editing.

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.

Received January 4, 2021; revised May 3, 2021; accepted August 3, 2021; published first August 17, 2021.

### References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020;70:7–30.
2. Liu B, Zhu L, Zou J, Chen HS, Miller KD, Jemal A, et al. Updated methodology for projecting US and state-level cancer counts for the current calendar year: Part I: Spatio-temporal small area modeling for cancer incidence. *Cancer Epidemiol Biomarkers Prev* 2021 [Online ahead of print].
3. Chen HS, Portier K, Ghosh K, Naishadham D, Kim HJ, Zhu L, et al. Predicting US- and state-level cancer counts for the current calendar year: Part I: evaluation of temporal projection methods for mortality. *Cancer* 2012; 118:1091–9.
4. Pickle LW, Hao Y, Jemal A, Zou Z, Tiwari RC, Ward E, et al. A new method of estimating United States and state-level cancer incidence counts for the current calendar year. *CA Cancer J Clin* 2007;57:30–42.
5. Tiwari RC, Ghosh K, Jemal A, Hachey M, Ward E, Thun MJ, et al. A new method of predicting US and state-level cancer mortality counts for the current calendar year. *CA Cancer J Clin* 2004;54:30–40.
6. Wingo PA, Landis 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.
7. Zhu L, Pickle LW, Ghosh K, Naishadham D, Portier K, Chen H-S, et al. Predicting US- and state-level cancer counts for the current calendar year: Part II: evaluation of spatiotemporal projection methods for incidence. *Cancer* 2012; 118:1100–9.
8. Percy C, Fritz A, Jack A, Shanmugarathan S, Sobin L, Parkin DM, et al. International classification of diseases for oncology. 3rd ed. World Health Organization; 2000.
9. World Health Organization. International statistical classification of diseases and related health problems. World Health Organization; 2011.
10. NCI. Joinpoint trend analysis software; 2020. Available from: <https://surveillance.cancer.gov/joinpoint/>.
11. SAS Institute Inc. SAS 9.4/STAT 15.1 User's Guide Carry, NC: SAS Institute Inc; 2018.
12. Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK. Estimating average annual per cent change in trend analysis. *Stat Med* 2009;28:3670–82.
13. West M, Harrison J. Bayesian forecasting and dynamic models. New York, NY: Springer; 1997.
14. Schmidt AM, Pereira JBM. Modelling time series of counts in epidemiology. *Int Stat Rev* 2011;79:48–69.
15. Jemal A, Fedewa SA, Ma J, Siegel R, Lin CC, Brawley O, et al. Prostate cancer incidence and PSA testing patterns in relation to USPSTF screening recommendations. *JAMA* 2015;314:2054–61.
16. Møller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talbäck M, et al. Prediction of cancer incidence in the nordic countries: empirical comparison of different approaches. *Stat Med* 2003;22:2751–66.
17. Rosenberg PS, Anderson WF. Age-period-cohort models in cancer surveillance research: ready for prime time? *Cancer Epidemiol Biomarkers Prev* 2011;20:1263–8.