Interpreting cancer incidence trends: challenges due to the COVID-19 pandemic

Angela B. Mariotto (iD), PhD,* Eric J. Feuer (iD), PhD, Nadia Howlader (iD), PhD, Huann-Sheng Chen (iD), PhD, Serban Negoita (iD), PhD, Kathleen A. Cronin (iD), PhD

Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA

*Correspondence to: Angela B. Mariotto, PhD, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland, USA, 9609 Medical Center Dr 4E602, Bethesda, MD 20892-9764, USA (e-mail: mariotta@mail.nih.gov).

Abstract

The considerable deficit in cancer diagnoses in 2020 due to COVID-19 pandemic disruptions in health care can pose challenges in the estimation and interpretation of long-term cancer trends. Using Surveillance, Epidemiology, and End Results (SEER) (2000-2020) data, we demonstrate that inclusion of the 2020 incidence rates in joinpoint models to estimate trends can result in a poorer fit to the data and less accurate or less precise trend estimates, providing challenges in the interpretation of the estimates as a cancer control measure. To measure the decline in 2020 relative to 2019 cancer incidence rates, we used the percent change of rates in 2020 compared with 2019. Overall, SEER cancer incidence rates dropped approximately 10% in 2020, but for thyroid cancer the decrease was as large as 18% after adjusting for reporting delay. The 2020 SEER incidence data are available in all SEER released products, except for joinpoint estimates of trends and lifetime risk of developing cancer.

One of the main missions of US cancer registries is to monitor efforts in cancer control by providing annually updated estimates of trends using the most recent data. In November 2022 the National Cancer Institute received data from the Surveillance, Epidemiology, and End Results (SEER) registries, including new cases diagnosed in 2020, the first year of the COVID-19 pandemic. These data will be used in annual reports to present cancer statistics and trends and represents the first opportunity to measure the impact of COVID-19 on cancer incidence rates.

Initial analyses of cancer registry data have shown a substantial decline in the number of cancer records reported to SEER and the National Cancer Database (1) during 2020. Both studies have also concluded that this decline was not due to changes in registry operations or data collection but reflected the impact of COVID on health services (2-5) and their consequential delays and reductions in cancer screening, diagnosis, and treatment.

The joinpoint model (6) is used in many cancer statistics reports, including the Annual Report to the Nation on the Status of Cancer (7), to describe cancer incidence trends. It has been developed to estimate long-term trends, including years at which the trends change (joinpoints) and measures summarizing trends in segments between joinpoints, that is, the annual percent change (APC) (6). The joinpoint model was not designed to accommodate a 1-time anomaly in the rates for the most recent year nor to estimate a change in 1 single calendar year. Including the 2020 incidence rates in the joinpoint model can bias long-term estimates and interpretation of their association with risk factors as well as cancer control efforts in cancer prevention and screening. We do not mean to dismiss the importance of the 2020 incidence data, especially because pandemic-associated delays in

screening, diagnosis, and treatment will likely have an impact on cancer mortality for many years to come. For these reasons we propose to exclude 2020 from the joinpoint analyses but include a supplementary more appropriate metric to measure the sudden decline in 2020 cancer incidence rates relative to 2019 rates, that is, the percent change of the rates in 2020 compared with 2019. Calculation details are included as Supplementary Methods (available online).

To demonstrate the challenges of including the 2020 incidence rates in joinpoint regression models, we provide examples for all cancer sites combined as well as melanoma and prostate cancer. We fit 3 joinpoint models to age- and delay-adjusted rates using 22 registries (SEER-22) (8) data from 1) November 2021 submission (2000-2019), 2) November 2022 submission (2000-2020), and 3) November 2022 submission excluding 2020. We used delay adjustment factors estimated from the November 2021 submission to inflate the incidence rates for the most recent years of data of both submissions (9). Delay adjustment factors are estimated by cancer site, registry, age group, gender, race, and year of diagnosis and multiplied to all cases based on these same factors to obtain delay-adjusted counts (10,11). We use the adjusted counts to estimate age-adjusted incidence rates, which serve as input in the joinpoint models and percent change estimates. Details are provided in the Supplementary Methods (available online).

Figure 1 shows the trends using the November 2021 submission (first column), the newest data from the November 2022 submission including 2020 (middle column) and excluding 2020 (last column). For all cancers combined, the model including the 2020 rate shows a poorer fit to the data, indicating that rates have
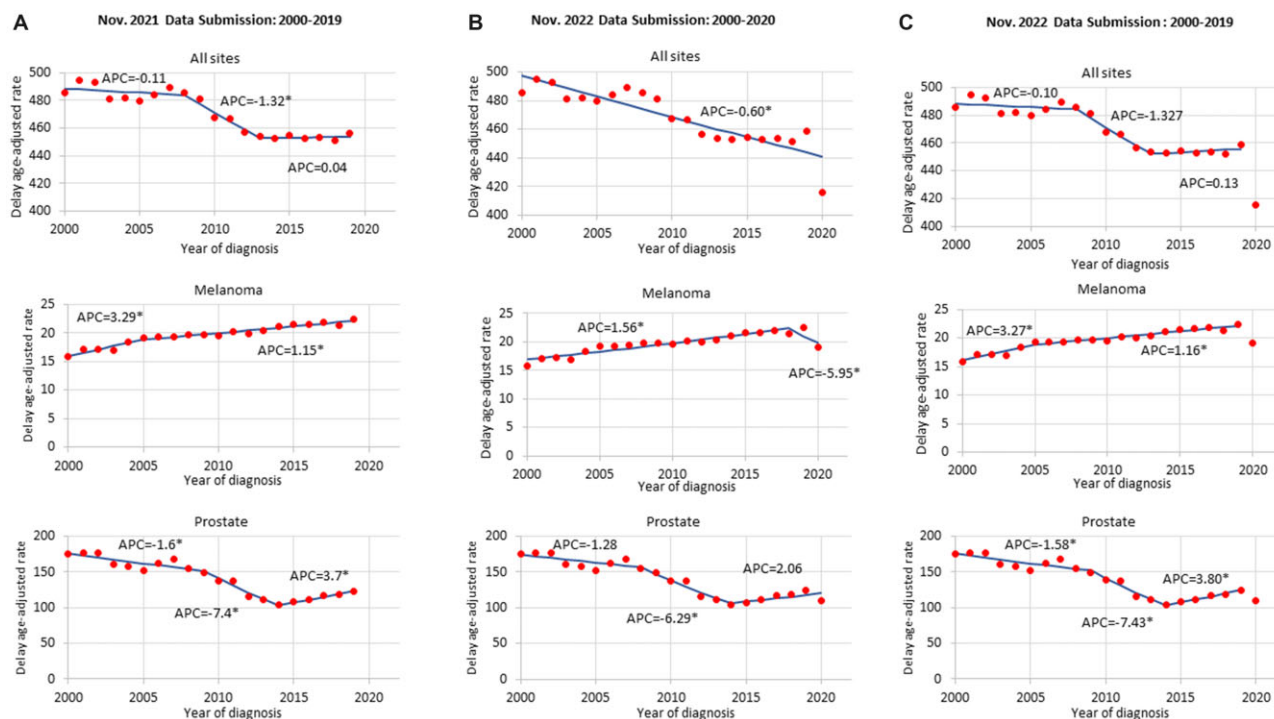
Figure 1. Trends in delay-adjusted age-standardized incidence rates for all cancer sites combined, melanoma and male prostate cancer using the Surveillance, Epidemiology, and End Results (SEER)-22. A) November 2021 data submission (2000-2019), B) November 2022 data submission (2000-2020), and C) November 2022 data submission (2000-2019), thus excluding 2020 data. An asterisk (*) indicates that the annual percent change (APC) is statistically significantly different from zero (P < .05).

Table 1. Percent change (PC)[a] in age-standardized, delay-adjusted incidence rates from 2019 to 2020 by cancer site and sex, using the November 2022 data submission

	Males					Females				
	2019 rate	2020 rate	PC	95% PC upper CI	95% PC lower CI	2019 rate	2020 rate	PC	95% PC upper CI	95% PC lower CI
All sites	499.1	451.2	−9.6%	−10.0%	−9.2%	432.6	392.5	−9.3%	−9.7%	−8.8%
Brain	7.5	7.3	−2.3%	−5.9%	1.4%	5.2	5.2	−0.1%	−4.2%	4.3%
Breast (females)	—	—	—	—	—	134.3	121.3	−9.7%	−10.4%	−8.9%
Cervix	—	—	—	—	—	7.9	7.1	−10.2%	−13.5%	−6.8%
Colorectal	43.5	38.4	−11.6%	−13.0%	−10.2%	32.5	29.3	−9.8%	−11.3%	−8.3%
Corpus and uterus	—	—	—	—	—	28.6	26.3	−7.9%	−9.5%	−6.3%
Esophagus	7.3	6.9	−5.4%	−8.8%	−1.9%	1.8	1.6	−7.8%	−13.9%	−1.2%
Hodgkin lymphoma	3.0	2.8	−7.2%	−12.7%	−1.5%	2.4	2.2	−7.3%	−13.3%	−0.9%
Kidney and RP	24.9	22.7	−8.7%	−10.6%	−6.9%	12.4	11.0	−11.5%	−13.9%	−9.1%
Larynx	4.7	4.2	−10.8%	−14.8%	−6.6%	1.0	0.9	−12.4%	−20.2%	−3.9%
Leukemia	19.4	17.9	−7.6%	−9.8%	−5.4%	11.6	11.2	−3.1%	−5.8%	−0.3%
Liver and IBD	14.7	13.2	−9.8%	−12.1%	−7.5%	5.5	5.2	−6.7%	−10.3%	−3.0%
Lung	56.3	50.0	−11.1%	−12.3%	−9.9%	46.4	40.4	−12.9%	−14.0%	−11.7%
Melanoma	28.7	24.1	−15.8%	−17.4%	−14.1%	18.0	15.4	−14.3%	−16.3%	−12.3%
Myeloma	9.4	8.7	−7.7%	−10.7%	−4.5%	6.4	5.9	−8.3%	−11.7%	−4.8%
NHL	23.3	21.5	−7.7%	−9.6%	−5.7%	16.1	14.7	−8.4%	−10.5%	−6.2%
Oral cavity and pharynx	17.8	16.6	−6.8%	−8.9%	−4.6%	6.5	6.2	−4.4%	−7.8%	−0.9%
Ovary	—	—	—	—	—	10.4	9.6	−7.4%	−10.1%	−4.6%
Pancreas	15.6	15.0	−4.3%	−6.7%	−1.9%	12.2	11.7	−3.9%	−6.3%	−1.4%
Prostate	124.2	109.9	−11.5%	−12.2%	−10.7%	—	—	—	—	—
Stomach	9.1	8.4	−8.3%	−11.3%	−5.1%	5.4	4.8	−10.5%	−14.1%	−6.8%
Testis	6.0	6.0	0.7%	−3.4%	5.0%	—	—	—	—	—
Thyroid	7.8	6.7	−14.0%	−17.1%	−10.7%	21.1	17.4	−17.9%	−19.7%	−16.0%
Urinary bladder	31.9	30.1	−5.7%	−7.4%	−4.0%	8.0	7.2	−9.4%	−12.2%	−6.4%

a Rates are per 100 000 population. We used SEER*Stat estimates of ratio = [rate (2020)/rate (2019)] and respective confidence interval (LCI, UCI). Percent change (PC) is calculated as (ratio − 1) × 100. The lower and upper limits of the confidence interval for PC are calculated as (UCI − 1) × 100 and (LCI − 1) × 100, respectively. CI = confidence interval; IBD = intrahepatic bile duct; NHL = non-Hodgkin lymphoma; RP = renal pelvis. Incidence rates were delay adjusted using the November 2021 submission delay adjustment factors (9). See site recode definition for details on cancer site codes: https://seer.cancer.gov/siterecode/icdo3_dwhoheme/index.html.

been declining at a constant rate since 2000 (APC = 0.60%) (Figure 1). The models excluding the 2020 rate are similar and suggest that rates were stable during 2013-2019. For melanoma the model including the 2020 incidence rates estimates a joinpoint in 2018 and suggests that rates began to decline from 2018 with APC = −5.12%. The models excluding 2020 rates indicate that the rates have been increasing at 1.15% per year since 2005. For prostate cancer including the 2020 incidence rate had little impact on the location of joinpoints, however; the increase in rates during 2014-2020 was not statistically significant (APC = 2.06). The models excluding the 2020 rate suggest that rates statistically significantly increased with APC = 3.7 and APC = 3.8 during 2014-2019. We used these few examples to illustrate the issues in estimating joinpoint trends. Among the top 10 cancer sites, the 2020 incidence rates changed the location of joinpoints in 8 cancer sites except for leukemia, prostate, and breast cancers (Supplementary Table 1, available online).

Table 1 shows the percent decline for 2020 compared with 2019 by cancer site and sex. Using this measure, we estimate that cancer incidence rates declined 9.6% for males and 9.3% for females after adjustment for reporting delay. The largest declines in 2020 were estimated for melanoma (15.8% males and 14.3% females) and thyroid (14.0% males and 17.9% females). The smallest decline was estimated for pancreatic cancer (4.3% males and 3.9% females), and there was no statistically significant decline for brain and testicular cancers. We estimated a smaller decline in all cancer sites compared with the 12.0% estimate using National Cancer Database data (1). In addition to differences in databases, our smaller decline reflects the use of the delay adjustment, which "inflates" the 2020 rates to reflect cases not yet reported but which will likely be reported in subsequent submissions.

In summary, the examples clearly demonstrate that the inclusion of 2020 incidence rates in joinpoint models can pose challenges in trends estimates and interpretation by providing a poorer fit to the data and less accurate and/or less precise long-term trend estimates. To continue providing accurate and precise long-term trends for monitoring the cancer burden, the National Cancer Institute proposes to exclude the 2020 cancer incidence rates from trends estimates in the 2023 annual cancer statistics reports. Importantly, the 2020 incidence data will be made available in all SEER released products, including SEER*Stat and all other cancer statistics. The decision to exclude 2020 rates from trend analysis in the future hinges on the 2021 rates. If the 2021 rates continue to decline, we may reconsider incorporating the 2020 rate. However, if the 2021 rates are like prepandemic levels, we will likely continue excluding the 2020 incidence rates from joinpoint analysis.

Data availability

This paper contains analysis of the November 2022 Surveillance Epidemiology and End Results (SEER) registries data submission to the National Cancer Institute (NCI). These data are available at https://seer.cancer.gov/data/access.html.

Author contributions

Angela B. Mariotto, PhD (Conceptualization; Formal analysis; Methodology; Visualization; Writing—original draft; Writing—review & editing), Eric J. Feuer, PhD (Conceptualization; Methodology; Writing—original draft; Writing—review & editing),

Nadia Howlader, PhD (Conceptualization; Methodology; Writing—original draft; Writing—review & editing), Huann-Sheng Chen, PhD (Conceptualization; Methodology; Writing—original draft; Writing—review & editing), Serban Negoita, PhD (Conceptualization; Writing—original draft; Writing—review & editing), Kathleen A. Cronin, PhD (Conceptualization; Methodology; Writing—original draft; Writing—review & editing).

References

1. Nogueira LM, Palis B, Boffa D, Lum S, Yabroff KR, Nelson H. Evaluation of the impact of the COVID-19 pandemic on reliability of cancer surveillance data in the National Cancer Database. *Ann Surg Oncol*. 2023;30(4):2087-2093.
2. London JW, Fazio-Eynullayeva E, Palchuk MB, Sankey P, McNair C. Effects of the COVID-19 pandemic on cancer-related patient encounters. *J Clin Oncol Clin Cancer Inform*. 2020;4:657-665.
3. Patt D, Gordan L, Diaz M, et al. Impact of COVID-19 on cancer care: how the pandemic is delaying cancer diagnosis and treatment for American seniors. *J Clin Oncol Clin Cancer Inform*. 2020;4:1059-1071.
4. Yabroff KR, Wu XC, Negoita S, et al. Association of the COVID-19 pandemic with patterns of statewide cancer services. *J Natl Cancer Inst*. 2022;114(6):907-909.
5. Kaufman HW, Chen Z, Niles J, Fesko Y. Changes in the number of US patients with newly identified cancer before and during the coronavirus disease 2019 (COVID-19) pandemic (vol 3, e2017267, 2020). *JAMA Netw Open*. 2020;3(8):e2017267.
6. Kim HJ, Fay MP, Feuer EJ, Medthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med*. 2001;19(4):655.
7. Nierengarten MB. Annual report to the nation on the status of cancer. *Cancer*. 2023;129(1):8-9.
8. Surveillance Research Program, National Cancer Institute. *Registry Groupings in SEER Data and Statistics*. https://seer.cancer.gov/registries/terms.html. Accessed March 15, 2023.
9. Surveillance Research Program, National Cancer Institute. *The 2022 Release of Delay Adjustment Factors and Rates*. https://surveillance.cancer.gov/delay/model.html. Accessed March 30, 2023.
10. Midthune DN, Fay MP, Clegg LX, Feuer EJ. Modeling reporting delays and reporting corrections in cancer registry data. *J Am Stat Assoc*. 2005;100(469):61-70.
11. Clegg LX, Feuer EJ, Midthune DN, Fay MP, Hankey BF. Impact of reporting delay and reporting error on cancer incidence rates and trends. *J Natl Cancer Inst*. 2002;94(20):1537-1545.