Surveillance Research Program, NCI, Technical Report #2024-01

How to Handle 2020 and 2021 Incidence Rates in the Joinpoint Trend Model?

Nadia Howlader[1], Huann-Sheng Cheng[1], Danny Miller, Jeff Byrne, Annie Noone[1], Serban Negoita[1], Kathleen Cronin[1], Angela Mariotto[1]

[1] Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, MD, USA

[2] Information Management Services, Calverton, MD, USA

Due to COVID-19 pandemic-related delays in cancer diagnosis and screening, cancer incidence rates for all cancers combined fell 10% in 2020 relative to 2019 rates (1). In particular, decline in incidence rates started in March 2020, coinciding with the implementation of stay-at-home orders (2). Because the pandemic caused temporary anomalies in 2020 data, estimates of substantive interest, such as cancer incidence trends, may be biased. The SEER Program made some changes in the reporting of cancer trends last year (2023). For example, the 2020 incidence data were **excluded** from the estimation of Joinpoint trends and the reason for the exclusion are detailed in Mariotto et al. (2023) (1).

This year's SEER data release for April 17th, 2024, will include new cancer cases diagnosed in 2021, the second year of the COVID-19 pandemic. To decide whether to include or exclude the latest two data points from Joinpoint, we considered three model fits:
- (**Model 1**): Joinpoint fit based on 2000-2019 data (latest data points **2020** and **2021 excluded** from the model fit)
- (**Model 2**): Joinpoint fit based on 2000-2021 data (**all** data points **included** in the model fit)
- (**Model 3**): Joinpoint fit based on 2000-2021 data (**2020** data point **excluded** as outlier from the model fit)

In considering the three options, it is important to note that the 2020 and 2021 data points are valid measures, meaning that we have analyzed data to assess whether the low incidence of cancer cases in 2020 and the very slightly decrease in 2021 could be due to reporting delays, changes in registry operation, or changes in data collection. For example, this year we are still seeing similar 2020 drop in the incidence rates which support the conclusion that 2020 data was not an error in data collection by the registry system. From these analyses, we conclude that the data cannot be explained by new patterns of missing data and likely reflect the actual counts of cancer cases during the pandemic years (3).

We considered the strengths and limitations of each model.

 (**Model 1**): Joinpoint fit based on 2000-2019 data (**exclude** latest data points 2020 and 2021)
- **Strength:** Exclude the 2020 data point due to the pandemic, treating it as an outlier, and exclude the subsequent 2021 data point, which may have been similarly influenced by the pandemic.
- **Limitation:** The estimation of the most recent 5-year trend measure will not include the 2020 and 2021 data points.

(**Model 2**): Joinpoint fit based on 2000-2021 data (**include all** data points)
- **Strength:** Incorporates all data points, including the most recent data points (2020 and 2021), when estimating trends in cancer incidence.

- **Limitation:** Inclusion of the 2020 data point leads to a loss of power in the Joinpoint models, which has a few consequences:
    - A) The models will detect fewer joinpoints (e.g., all cancer sites combined) and/or
    - B) The models will provide less precise (larger variance, more estimates being statistically nonsignificant) recent 5-year annual percent change (APC) estimates. The loss of power is due to the added variance introduced into the model by the 2020 data point, which is considerably lower than the predicted trend seen in many cancer sites.
    - C) Data products that report only the trend, and that do not show the trendline as well as the underlying data, are highly susceptible to misinterpretation.
    - D) The gross patterns (i.e., whether the trend is increasing, stable, or decreasing) wildly fluctuate from year to year for some cancer sites.

(**Model 3**): Joinpoint fit based on 2000-2021 data (**exclude 2020** data point as outlier)
- **Strength**: The model exhibits better fit than Model 2, without compromising statistical power, and it incorporates the latest data point of interest, i.e., 2021.
- **Limitation**: The estimation of the most recent 5-year trend measure will not include the 2020 data point.

Figure 1. Trends in Delay-adjusted Age-Standardized Incidence Rates for All Cancer Sites Combined, Male Prostate Cancer, and Colon & Rectum Cancer Using the Surveillance, Epidemiology, and End Results (SEER) 22 Registries. Model 1 uses data from 2000-2019 (first column). Model 2 includes all the data from 2000-2021 (middle column). Model 3 excludes 2020 data (last column). An asterisk indicates that the annual percent change (APC) is statistically significantly different from zero (P < .05).
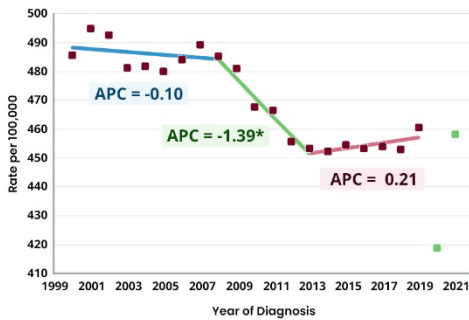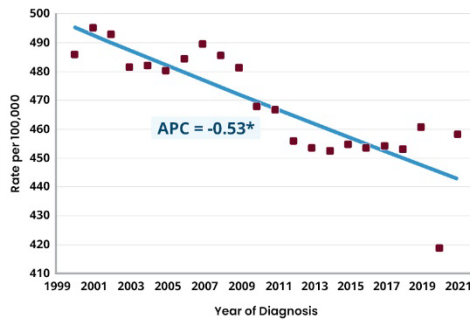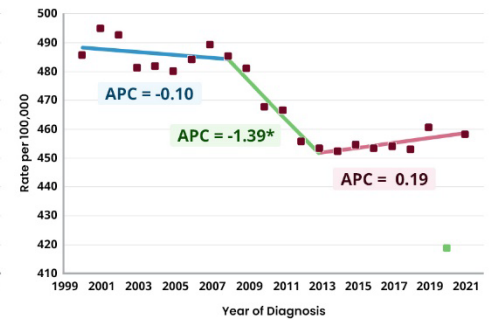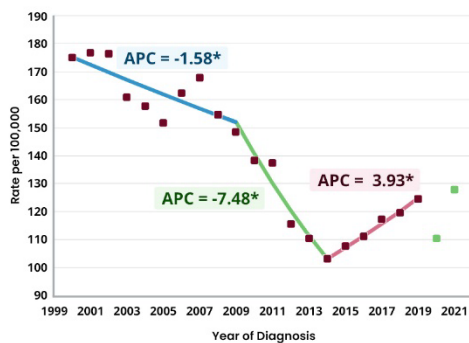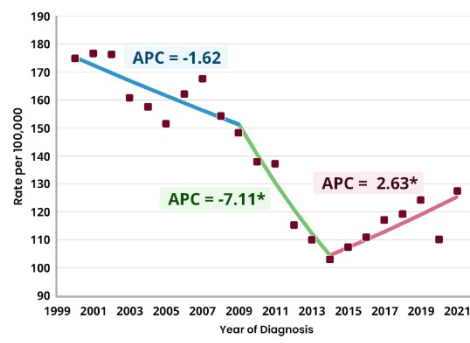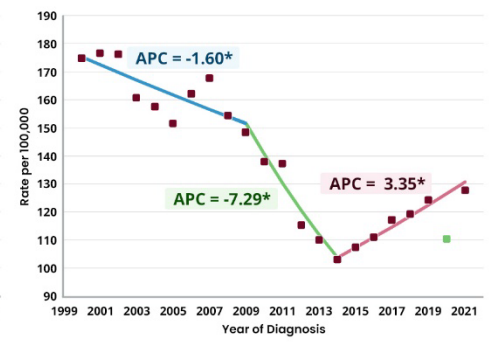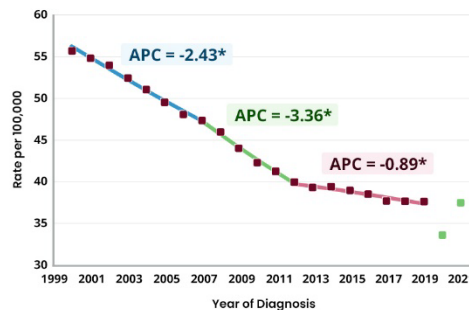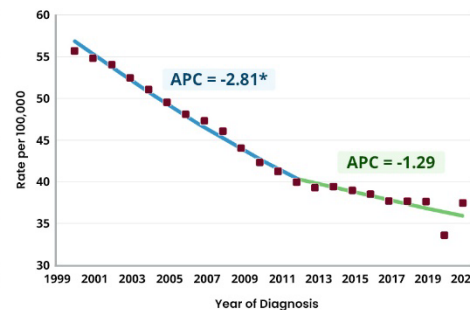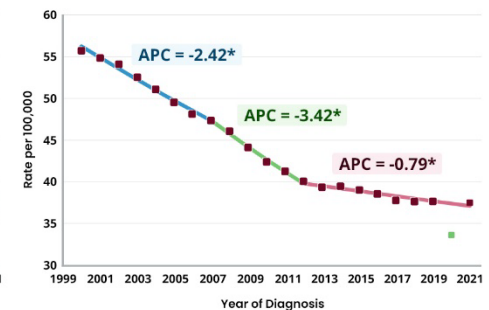
Recommendation:

We recommend fitting Joipoint models to be reported in SEER*Explorer using Model 3 (i.e., excluding the 2020 data point from the model but including the 2021 incidence data) as the default setting. SEER*Explorer will still display the 2020 incidence point on the graph. The main reasons for this approach are:

A) It aligns with how we presented SEER data and trends last year.
B) Trends from Model 3 were more closely aligned with pre-pandemic trends using the pre-pandemic data (2000-2019).
C) The model including all data points (Model 2) showed in some cases similar loss of power in the Joinpoint estimation that we had observed last year which could lead to estimation of less joinpoints and less accurate and/or less precise trend estimates.

Note that the 2020 incidence data would still be included in all SEER incidence databases and most cancer statistics in SEER*Explorer, Stat Facts, and other online reports. The 2020 incidence data will only be excluded when estimating trends and Devcan databases. This applies to incidence only, as mortality rates did not show a similar outlier in 2020.

1. Mariotto AB, Feuer EJ, Howlader N, Chen HS, Negoita S, Cronin KA. Interpreting cancer incidence trends: challenges due to the COVID-19 pandemic. J Natl Cancer Inst. 2023;115(9):1109-11.
2. Howlader N, Bhattacharya M, Scoppa S, Miller D, Noone AM, Negoita S, et al. Cancer and COVID-19: US cancer incidence rates during the first year of the pandemic. J Natl Cancer Inst. 2024;116(2):208-15.
3. Negoita S, Chen H-S, Sanchez PV, Sherman RL, Henley SJ, Siegel RL, et al. Annual Report to the Nation on the Status of Cancer, part 2: Early assessment of the COVID-19 pandemic's impact on cancer diagnosis. Cancer. 2024;130(1):117-27.