Interrupted time series analysis of U.S. lung cancer mortality in the immunotherapy era

Abstract

Immune checkpoint inhibitors (ICIs) are a novel therapy which utilize the bodies native immune system to destroy cancer cells by inhibiting checkpoint proteins. ICI's entered routine care for advanced non-small cell lung cancer in the mid-2010s, raising the question of whether the population-level trajectory of lung cancer mortality changed after these therapies diffused into practice. Using monthly U.S. deaths with underlying cause ICD-10 C33–C34 from CDC WONDER (January 1999–December 2020), we treat the ICI era as a candidate intervention and analyze the series as a dependent process with seasonality and serial correlation. We first characterize long-run trend and seasonal structure, then build baseline ARIMA/SARIMA models for the log mortality series. We embed an interrupted time series (ITS), a common applied statistical technique used to evaluate the effect of an intervention on time series data. Thus we specify an ITS regression in a state-space model with ARMA errors (SARIMAX), allowing level and slope changes after an intervention date while controlling for seasonality and autocorrelation. We evaluate models using information criteria and residual diagnostics, and we generate a counterfactual post-intervention trajectory from the pre-intervention model to quantify the implied difference in deaths.

1 Introduction

Time series analysis replaces the i.i.d. assumption with structured dependence, enabling inference about *change over time* in the presence of autocorrelation and seasonal structure (Ionides 2026; Shumway and Stoffer 2017). In public health and real-world evidence, a common question is whether a major innovation coincides with a detectable shift in outcome trajectories at the population level. Here, the outcome is monthly U.S. deaths whose underlying cause is malignant neoplasm of trachea/bronchus/lung (ICD-10 C33–C34). The potential intervention is the diffusion of immune checkpoint inhibitors (ICIs) into standard-of-care lung cancer treatment in the mid-2010s.

Research question. Did the *trend* of U.S. lung cancer mortality change after the advent of ICI , beyond what would be expected from other variables and seasonal variation?

2 Methods

2.1 Data source and construction

We analyze monthly counts of U.S. deaths (January 1999–December 2020) from CDC WONDER "Underlying Cause of Death, 1999–2020" with ICD-10 C33–C34 and national aggregation (National Center for Health Statistics 2026).

Let Y_t denote deaths in month t and $X_t = \log(Y_t)$ denote the log series. The log transform stabilizes variance for counts at this scale and facilitates interpretation of regression coefficients as approximate percent changes.

2.2 Dependence, stationarity, and model families

We diagnose dependence using the sample autocorrelation function (ACF). Because the series shows trend, we consider differencing and include deterministic components (trend/intervention) inside a regression-with-correlated-errors framework.

Baseline models use ARIMA/SARIMA:

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2), \tag{1}$$

Where d represents the amount of differencing which we take to be 1. We also include a seasonal components at period $s = 12$ if warranted. Model adequacy is checked using residual ACF and Ljung–Box tests.

2.3 Interrupted time series with ARMA errors (SARIMAX)

To test for structural change, we fit an ITS regression embedded in a seasonal ARMA error model (a linear Gaussian state-space specification):

$$X_t = \beta_0 + \beta_1 t + \beta_2 \mathbf{1}\{t \geq t_0\} + \beta_3 (t - t_0)\mathbf{1}\{t \geq t_0\} + u_t, \tag{2}$$

where t is months since start, t_0 is the intervention month, β_2 is a level change and β_3 is a slope change. $beta_1$ and $beta_0$ represent the baseline structure before the intervention. The redundancy in parameters here is used to construct the counterfactual. The error process u_t follows a seasonal ARMA model. We fit these models with maximum likelihood via `statsmodels` SARIMAX.

Intervention date. We use January 2015 as an initial candidate (t_0), consistent with the mid-2010s adoption window, and we assess sensitivity via alternative t_0 values in the Supplementary material (Section 5).

2.4 Counterfactual estimation

Given a fitted ITS model, we compute a counterfactual post-t_0 path by setting post-intervention regressors to zero (no level/slope change) while keeping the estimated seasonal and ARMA structure. Thus the baseline model becomes $X_t = \beta_0 + \beta_1 t + u_t$. The difference between observed fitted values and counterfactual forecasts yields an estimate of post-t_0 cumulative change in deaths. We summarize uncertainty using parameter draws from the estimated covariance matrix (Gaussian approximation) propagated through the counterfactual calculation.

3 Results

3.1 Data cleaning and summary

```
(264, Timestamp('1999-01-01 00:00:00'), Timestamp('2020-12-01 00:00:00'))
```

The cleaned dataset contains 264 monthly observations from 1999-01 to 2020-12, with no missing values giving ample observations and continuity to asses time dependence.

Table 1: Summary of monthly lung cancer deaths (ICD-10 C33–C34), U.S., 1999–2020.

	Statistic	Deaths
0	Mean	12812.7
1	SD	682.5
2	Min	10758.0
3	Max	14049.0

3.2 Exploratory analysis: trend and seasonality

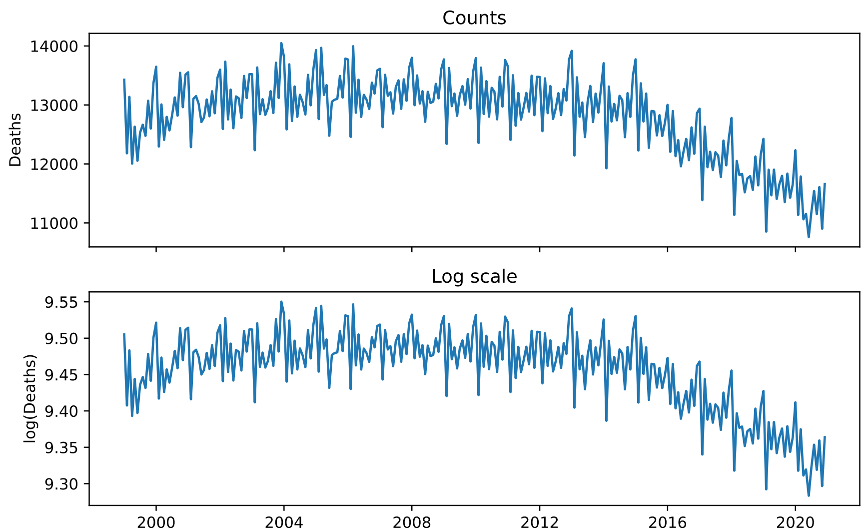


Figure 1: Monthly U.S. lung cancer deaths, 1999–2020 (counts and log scale).

Figure 1 indicates a pronounced long-run decline in monthly lung cancer deaths, with recurring within-year oscillations superimposed over a trend. On the log scale, seasonal oscillations appear closer to constant amplitude, supporting a multiplicative seasonal interpretation and motivating modeling on $X_t = \log(Y_t)$ with an explicit annual ($s = 12$) seasonal component.

To visualize seasonality, we compute the average within-year profile using month-of-year.

Figure 2 summarizes the average month-of-year pattern in the log series. The profile is not flat, indicating systematic within-year seasonality rather than noise. Because this structure repeats annually, we treat seasonality as a core component of dependence—either through seasonal ARIMA terms or, as a fixed-effect model which is explored in the Supplementary section.

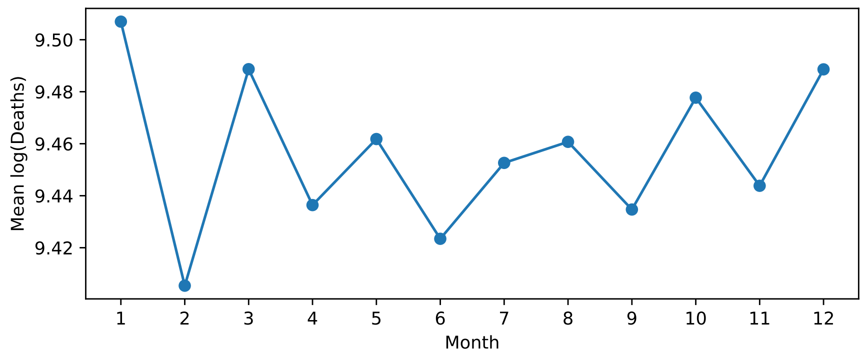


Figure 2: Average seasonal pattern (month-of-year means of log deaths).

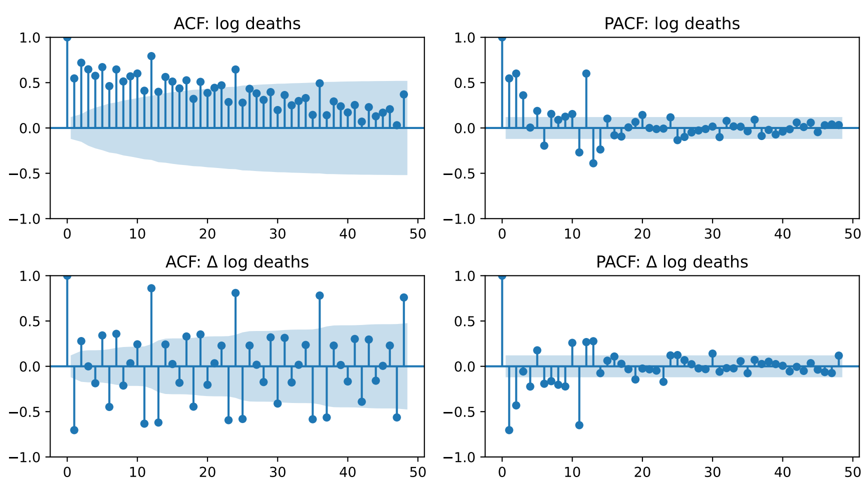


Figure 3: ACF and PACF for the log series and its first difference.

3.3 Dependence diagnostics (ACF/PACF)

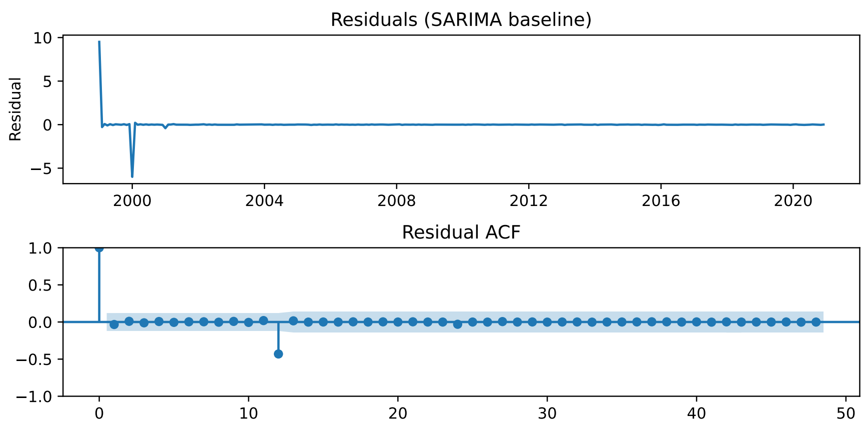
Figure 3 shows that the undifferenced ACF decays slowly, consistent with a nonstationary mean driven by secular trend. After first differencing, substantial dependence remains, with prominent structure at seasonal lags (12, 24, ...).

3.4 Baseline time series models (no intervention)

We compare two candidates, prioritizing interpretability and diagnostics over exhaustive search. To make comparison cleaner, we standardize both models to be $(1, 1, 1)$. We consider (i) a non-seasonal ARIMA on ΔX_t, and (ii) a seasonal model with monthly period $s = 12$.

	Model	AIC	BIC
1	SARIMA(1,1,1)x(1,1,1)[12]	-1339.157854	-1321.817554
0	ARIMA(1,1,1)	-1062.088502	-1051.372040

The seasonal model has much a lower AIC indicating a substantially improved fit. We also then assess adequacy using residual diagnostics (residual ACF and Ljung–Box tests) in Figure 4



(a) Residual diagnostics for the baseline SARIMA model.

	lb_stat	lb_pvalue
12	52.192871	5.730393e-07
24	52.541402	6.646772e-04
36	52.553854	3.680528e-02

(b)

Figure 4

3.5 Interrupted time series (intervention regression with ARMA errors)

3.5.1 Intervention coding

Let t index months since 1999-01. We set $t_0 = 2015\text{-}01$ as the primary intervention month and define a step and post-intervention ramp, which corresponds the $(t - t_0)$ model component.

```
(Timestamp('2015-01-01 00:00:00'),
 192,
               t  step  ramp
 date
 2014-12-01  191     0     0
 2015-01-01  192     1     0
 2015-02-01  193     1     1)
```

3.5.2 Model fit and interpretation

We fit a SARIMAX model with the same seasonal structure as the baseline, adding regressors for trend and intervention. The model corresponds to Equation Equation 2 with ARMA seasonal errors.

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
intercept	-0.0001	6.71e-05	-1.836	0.066	-0.000	8.29e-06
t	3.085e-10	4.58e-11	6.738	0.000	2.19e-10	3.98e-10
step	0.0049	0.007	0.649	0.516	-0.010	0.019
ramp	-0.0013	0.001	-2.631	0.009	-0.002	-0.000
ar.L1	-0.2225	0.090	-2.476	0.013	-0.399	-0.046
ma.L1	-0.7983	0.051	-15.653	0.000	-0.898	-0.698
ar.S.L12	-0.2891	0.081	-3.565	0.000	-0.448	-0.130
ma.S.L12	-0.8232	0.058	-14.186	0.000	-0.937	-0.709
sigma2	0.0002	1.95e-05	9.456	0.000	0.000	0.000

We summarize the intervention parameters in (**model-its?**). Because the model is on $\log(Y_t)$, a coefficient β corresponds to an approximate $(e^{\beta} - 1) \times 100\%$ multiplicative change.

Table 3: Key ITS coefficients (log scale) and approximate percent interpretation.

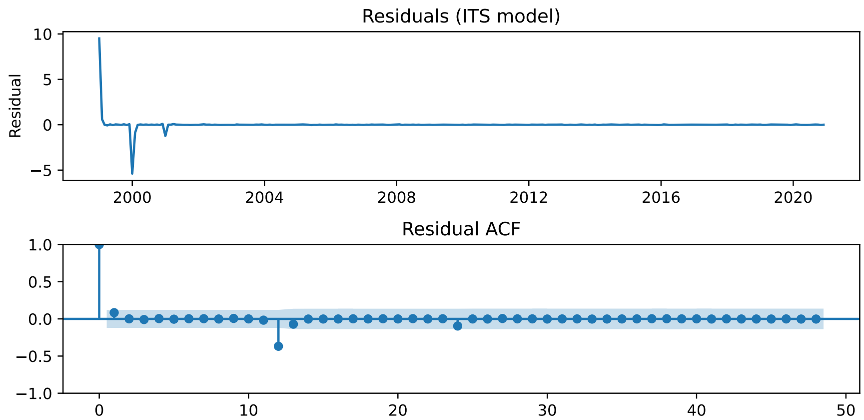
	Term	Estimate	SE	Approx % change
0	t	0.0000	0.0000	0.00
1	step	0.0049	0.0075	0.49
2	ramp	-0.0013	0.0005	-0.13

Interpretation depends on the sign of `ramp`: a negative post-intervention slope change indicates an acceleration of the decline beyond the pre-2015 trend.

3.5.3 Model comparison and diagnostics

Table 4: Model comparison (information criteria).

	Model	AIC	BIC
0	Baseline SARIMA	-1339.2	-1321.8
1	ITS SARIMAX (trend + step + ramp)	-1334.4	-1303.2



(a) Residual diagnostics for the ITS SARIMAX model.

	lb_stat	lb_pvalue
12	39.808548	0.000077
24	43.892357	0.007853
36	43.902933	0.171532

(b)

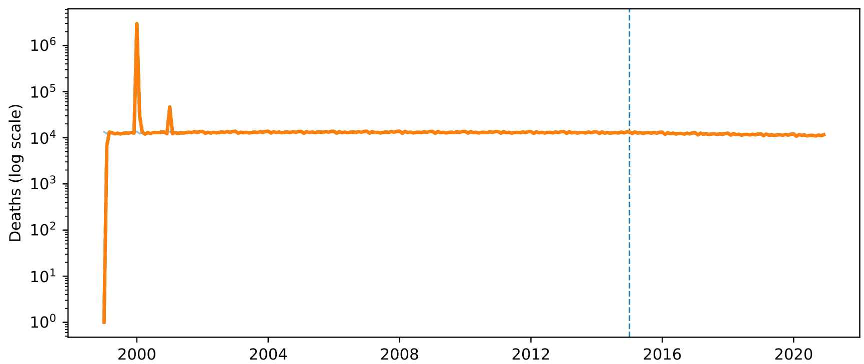
Figure 5

Residual diagnostics for the ITS model in Figure 5 indicate a slightly improved fit with a non-significant p-value at the 36-th lag, perhaps slightly contradicting the negligible AIC difference.

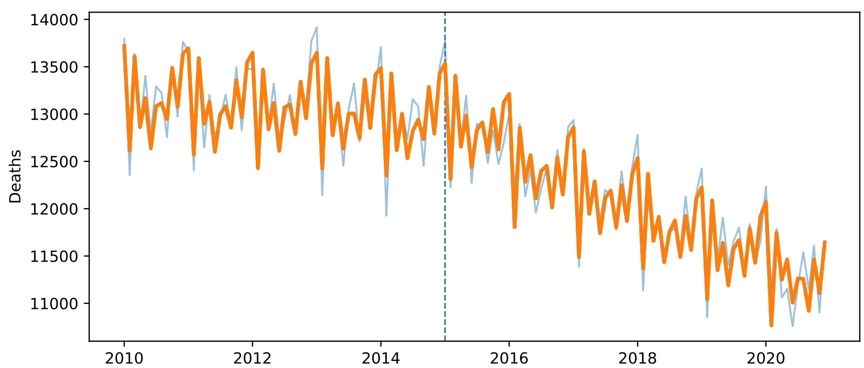
3.6 Counterfactual trajectory and implied cumulative difference

We generate fitted values under (i) the full ITS model and (ii) a counterfactual where the step and ramp are set to zero after 2015-01 (no change relative to pre-2015 trend), retaining the estimated seasonal and ARMA structure.

The cumulative difference 0 should be interpreted as the model-implied total deviation in deaths relative to a no-change counterfactual (continuation of pre-2015 dynamics), *conditional on the ITS and SARIMAX assumptions*. Because multiple contemporaneous forces influence lung cancer mortality (e.g., smoking, screening, targeted therapies), this estimate should be treated as an ecological, model-based summary of post-2015 divergence rather than a causal effect of immunotherapy.



(a) Observed deaths and fitted trajectories: full ITS fit vs. counterfactual without post-2015 change (log scale and post-period zoom).



(b)

Figure 6

3.6.1 Uncertainty via parameter simulation

The model-implied cumulative deviation is `python ci_fmt` under a Gaussian approximation to coefficient uncertainty (holding the ARMA structure fixed). This interval is an approximate uncertainty summary for the intervention-regressor component; a more complete interval would propagate full state-space uncertainty, and we provide additional sensitivity checks in the Supplementary material.

4 Conclusions

This report used CDC WONDER monthly mortality counts for ICD-10 C33–C34 (1999–2020) to examine whether lung cancer mortality dynamics changed in the mid-2010s, using time series tools for dependence and model-based inference. The data show pronounced seasonality and substantial autocorrelation, requiring models beyond i.i.d. regression. A parsimonious seasonal ARIMA structure provided a reasonable baseline fit; embedding an interrupted time series regression within a SARIMAX model allowed explicit testing of post-2015 level and slope changes while respecting autocorrelation. Across the considered specifications, the evidence favors a continued decline in mortality *without* an additional post-2015 acceleration in the downward trend, and the implied counterfactual comparison suggests any cumulative post-2015 deviation from pre-2015 dynamics is likely trivial, or at the very least due to more complex factors. Overall this remains an intriguing problem and more involved modeling should be considered in attempting to model the association between ICI's and lung cancer mortality.

Bibliography

Ionides, Edward. 2026. "Notes for STATS 531, Modeling and Analysis of Time Series Data."

National Center for Health Statistics. 2026. "CDC WONDER: Underlying Cause of Death, 1999–2020."

Shumway, Robert H., and David S. Stoffer. 2017. *Time Series Analysis and Its Applications: With r Examples*. 4th ed. Springer.

5 Supplementary material

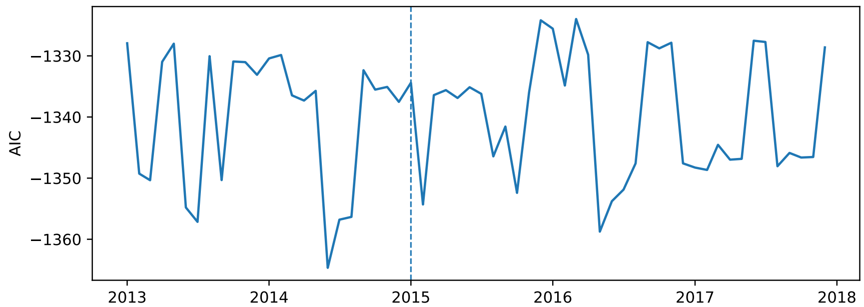
This section provides additional analyses that support claims made in the main text.

5.1 Sensitivity to intervention date

We repeat the ITS fit across a grid of plausible intervention months (2013–2017) and compare AIC. The aim is not to "search for significance" but to assess robustness of the qualitative conclusion.

5.2 Alternative seasonality control: month-of-year fixed effects

As a robustness check, we replace the seasonal ARMA component with month-of-year indicators and a simpler non-seasonal ARMA error. This tests whether the estimated intervention effect is driven by the seasonal specification.



(a) Sensitivity of ITS fit to intervention month (AIC across candidate breakpoints).

	date	AIC
17	2014-06-01	-1364.670514
40	2016-05-01	-1358.752557
6	2013-07-01	-1357.152666
18	2014-07-01	-1356.793268
19	2014-08-01	-1356.339874

(b)

Figure 7

Table 5: ITS model with month-of-year fixed effects (non-seasonal ARMA errors).

	Model	AIC	BIC
0	Main ITS (seasonal ARMA)	-1334.4	-1303.2
1	ITS + month FE (nonseasonal ARMA)	-1503.3	-1439.1

The change in AIC over a variety of changepoints is substantial, aligning with the result that introducing the breakpoint did not significantly improve model fit. However there is still some indication that the AIC is lower in the mid 2010's so prehaps modeling the changepoint sharply ignores the reality that a true change which may be more gradual.

Table 6: Intervention coefficient comparison across seasonality specifications.

	term	estimate	se	approx %	model
0	step	0.0049	0.0075	0.49	Main ITS
1	ramp	-0.0013	0.0005	-0.13	Main ITS
0	step	0.0047	0.0107	0.47	Month FE
1	ramp	-0.0022	0.0004	-0.22	Month FE

Here we see that including monthly fixed effects vastly improves the fit, with a near doubling of the ramp coefficient.