Time Series Models for Public Health Surveillance

Colorectal Cancer Incidence, Inequalities, and Prevention Priorities in Urban Texas

UTSouthwestern Medical Center

Connor Donegan, MA[1,2]; Amy E. Hughes, PhD[1]; Simon Craddock Lee, PhD, MPH[3]

1. O'Donnell School of Public Health, UTSW; 2. Geospatial Information Sciences, The University of Texas at Dallas; 3. Population Health, University of Kansas Medical Center

Advancing Health Surveillance Methodology

This research aims to advance public health surveillance research practices by providing an open-source software package, **surveil.** Joinpoint regression modeling is the method of choice in the field, but it has considerable shortcomings:

- Standard errors are over-confident
- Trends are forced to be piecewise-linear

Bayesian first-difference (random walk) models offer:

- More principled measures of uncertainty
- Greater flexibility

Using Markov chain Monte Carlo (MCMC), inference can also be undertaken on any quantities of interest that are derived from model estimates of risk.

Model Specifications

We model trends in disease risk **η** using the **Poisson likelihood for case counts** and the **first-difference prior**:

$$y_t \sim \text{Poisson}(P_t * \exp(\eta_t))$$

$$\eta_t \sim \text{Gauss}(\eta_{t-1}, \tau^2), t>1$$

where y_t is the observed number of cases at time *t*, P_t is the population at risk, and τ is a scale parameter.

Two parameters require prior probabilities be assigned: log-risk in year 1, η_1, and the scale parameter, τ. The **default priors are diffuse on the scale of log-risk**:

$$\eta_1 \sim \text{Gauss}(-5, 5^2), \eta_1 < 0$$

$$\tau \sim \text{Gauss}(0, 1), \tau > 0.$$

Users can easily set custom prior distributions. Binomial models and correlated time-series models are also implemented in **surveil**.

Evaluating Colorectal Cancer (CRC) Prevention Progress in Urban Texas

The *Texas Cancer Plan* identifies cancer equity as a priority but does not specify any equity-related targets. We modeled trends in CRC incidence by race-ethnicity in Texas' four largest metropolitan areas (centered on Austin, Dallas, Houston, and San Antonio) using **surveil**.

We found a period of robust risk reduction (2004—2012) was followed by a plateau (2013-2018), plus persistently greater risk for Black residents. **The findings indicate two priorities**: **1)** initiating a new period of robust prevention progress, and **2)** closing the gap between Black residents and other groups.

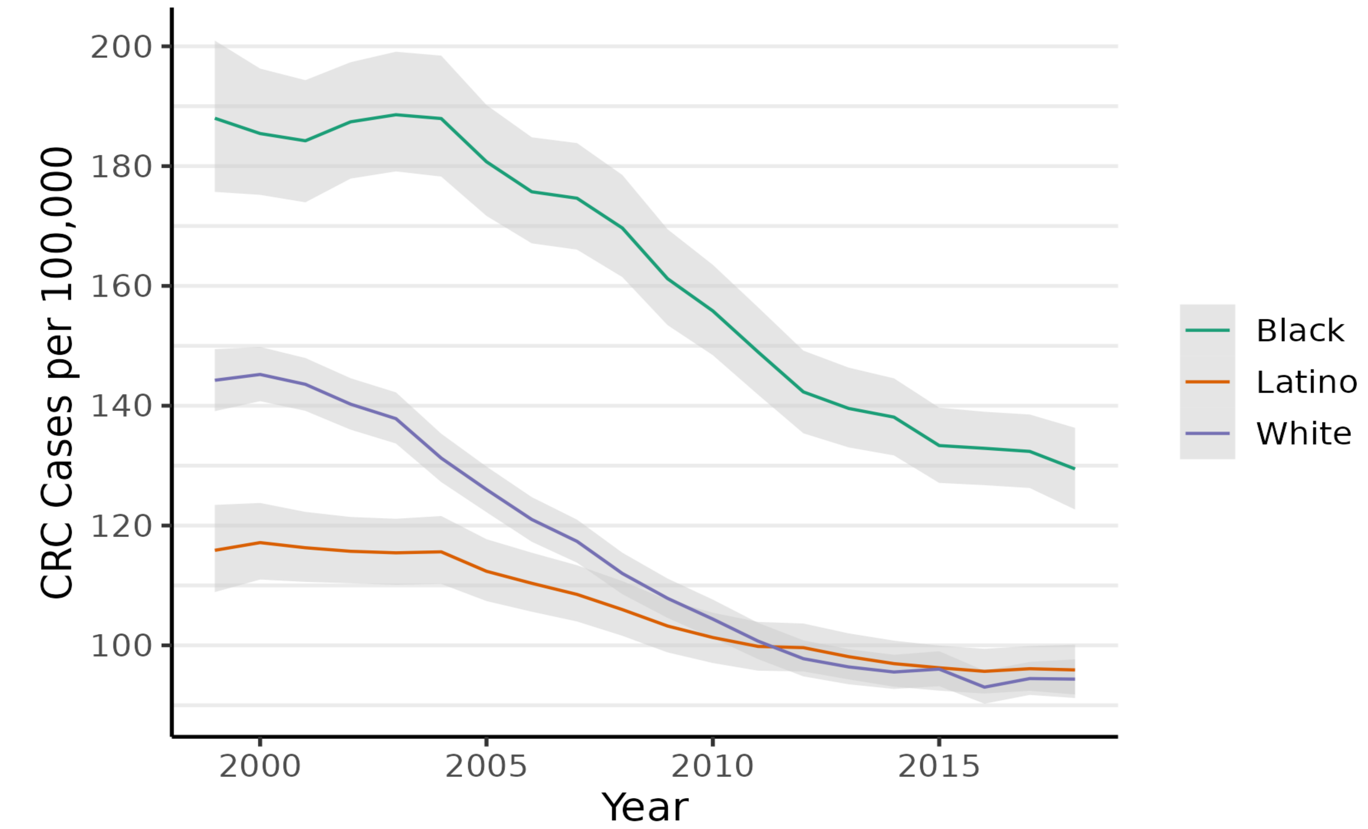

Figure 1. Age-standardized CRC incidence rates, ages 50-79, Texas MSAs.

The Black-White rate difference increased then fell to near its starting level, while the corresponding number of excess cases nearly doubled as a result of population growth.

Black-White risk inequality:

- **Cumulative Excess Cases**: 3,983 (95% CI 3746-4219)
- **100 x Cumulative Excess Risk/Cumulative Risk**: 31% (95% CI 29%-32%)

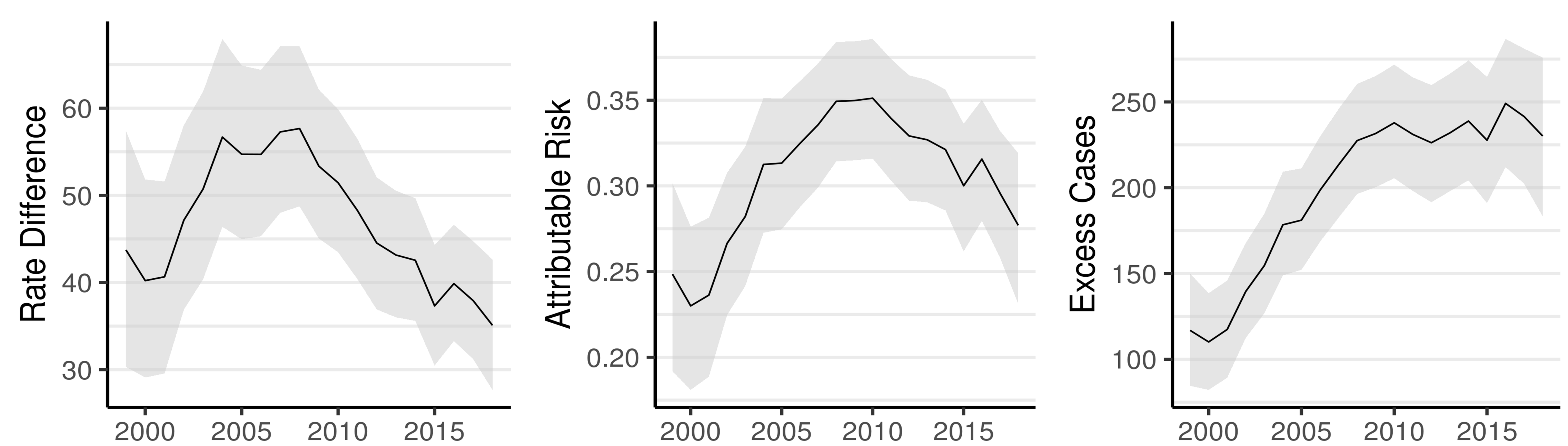

Figure 2. Black-White CRC risk inequality, ages 50-79, Texas MSAs.

The surveil R Package

surveil supports a streamlined workflow for analyzing mortality and disease incidence data, powered by the Stan modeling language.

Users with basic R programming skills get:

- Publication-quality visualizations
- Percent change analysis (annual, cumulative)
- Direct age-standardization
- A suite of health inequality measures
  - Theil's index
  - Rate ratio, rate difference, excess cases, proportion attributable risk
  - Proper adjustments for age-stratified populations

Experienced R users can analyze custom quantities of interest using MCMC analysis, such as percent change over custom time periods.

Install from R with:

```
install.packages("surveil")
```

Connor.Donegan@UTSouthwestern.edu

References

Cancer Prevention and Research Institute of Texas. *2018 Texas Cancer Plan.* Accessed from https://www.cprit.state.tx.us/about-us/texas-cancer-plan

CDC WONDER Online Database (2021). US Cancer Statistics—Incidence: 1999-2018. Accessed from http://wonder.cdc.gov/cancer-v2018.html

Clayton DG. Generalized Linear Mixed Models. In: Gilks, WR, Richardson, S, Spiegelhalter, DJ, eds. *Markov Chain Monte Carlo in Practice*. CRC Press, 1996: 275-302.

Donegan C, Hughes AE, Lee SJC (2022) Colorectal Cancer Incidence, Inequality, and Prevention Priorities in Urban Texas: Surveillance Study with the "surveil" Software Package. *JMIR Pub Health Surveill 8(8):e34589* DOI:10.2196/34589 PMID:35972778

Kim HJ, Fay MP, Feuer EJ, Midthune DN (2000). Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 19(3): 335-351.

Stan Development Team (2022). Stan Modeling Language Users Guide and Reference Manual, 2.3. https://mc-stan.org/.