COVID-19: Impact on Death Counts by Age Groups

Written by Liman Wei in Nov. 2020 and Posted on https://github.com/weiliman/Course-Assignments

Introduction

In this report we are investigating the impact of covid-19 on death counts of different age groups in Quebec, Canada. It is believed by an imaginary government that the first wave (March - May) primarily affected people aged over 70, whereas the second wave (Sept - now) mainly influences people aged under 50. We are going to test this hypothesis based on Quebec's daily mortality data.

Abstract

An imaginary government announces that the first wave of covid-19 primarily affected older people, and the second wave are caused by irresponsible people who aged under 50. Through this report we are studying the death count data from Quebec to prove or disprove the government's statement. We fitted our models using data before covid-19 to predict the death rates without covid-19, and compared the resulted predictions with the real data points. As a result, we found the death counts of people aged over 70 are seriously impacted by the first wave of covid, however, in the second wave, the death intensity of over 70s is similar to the historical records. On the other hand, both wave 1 and wave 2 of covid increases the death rates of people aged under 50. Hence we conclude we found evidence supporting the statement from the government.

Methods

The sample data from Quebec contains 563 records of daily mortality counts from Jan 2010 to Oct 2020. We fitted the model using records before covid-19 (March 2020), and used this model to predict the number of deaths we will have if no covid-19 happens (pre-covid deaths). To contrast the difference between two age groups (under 50s and over 70s), we fitted the same model twice for each age group. Then we used these two pre-covid models to predict the death counts during wave 1 and wave 2 of covid, and compare the posterior predictions with the real death counts.

Overviews of the sample data are shown in Figure 1 and Figure 2 for both age groups. In these graphs we can clearly see there is a seasonal change in the death counts, which is possibly caused by flu seasons. To model the seasonal effect, I created four new variables by transforming the time variable using sin and cos with frequencies of 6 months and 12 months.

The models we are using are based on Bayesian semi-parametric inference. A math representation for the models is:

$$Y_i \sim Poisson(\lambda_i) \qquad log(\lambda_i) = X_i\beta + U(t_i) + W_i$$

with priors:

$$[U_1...U_T]^T \sim RW2(0, \sigma_1^2) \qquad W_i \sim N(0, \sigma_2^2)$$

$$P(\sigma_1 > 0.01) = 0.5 \qquad P(\sigma_2 > log(1.2)) = 0.5$$

where Y_i represents the number of deaths at date i (measured in years), X_i is the covariates containing four basis sinusoidal functions that models the seasonal effects. $U(t)$ is a second-order random walk, the prior states the slope of log CO2 concentration changes by 0.01 from one year to the next. W_i is the random effect of date i, and the IQR is set to 1.2, hence the odds ratio of IQR is a 20% increase.

The posterior quantiles of fixed and random effects are listed in Table 1, 2. The predicted death intensities and random effects are plotted in Figure 3 - 6 with 95% confidence interval.

To compare the pre-covid models and the real death counts, we draw 30 posterior samples from each model, and plotted them in Figure 7 - 12 with comparison to real data points. These plots also contain the 0.025 and 0.975 posterior distribution quantiles (95% confidence interval).

To calculate the excess death caused by covid-19, we substract the pre-covid deaths (posterior samples) from the real deaths, and plotted the results in Figure 13 and Figure 14. We also created tables of predicted total excess death counts during the first wave (Table 3), and during the second wave (Table 4).

Results

Comparing Figure 3 and Figure 4, we can see the posterior predictions (black lines) are close to real data points (the red points) for under 50s, however, the predicted values are far from the real data points for over 70s during covid period. Similar pattern is shown in Figure 7 and Figure 8 (plots of posterior samples).

Figure 9-12 gives us a more detailed view of the posterior samples. From these plots, we can see the actual deaths (red points) are slight higher than posterior samples (black lines) for under 50s in first waves of covid, however, the actual data points are much higher than posterior samples for over 70s in the first wave. Hence we conclude the first wave of covid primarily affects people aged over 70.

For the second wave, some real data points exceed the 0.025 quantile of posterior distribution for people aged under 50, whereas nearly all real data points stay between the 0.025 and 0.975 posterior distribution quantiles for over 70s. Hence we conclude that the actual number of deaths of people aged under 50 are higher than historical data in second wave, while the actual number of deaths of over 70s is close to historical data.

Figure 13 and Figure 14 shows the number of excess deaths (red real deaths - black posterior samples) of two age groups from March to October. The quantiles of predicted total excess deaths during two waves of covid-19 are listed in Table 3 and Table 4 for both age groups. From these plots and tables, we can see the impacts of covid on death intensity of people under 50 are similar in both waves. In contrast, people aged over 70 are seriously affected during the first wave of covid, but the second wave nearly has no influence on the death intensity of over 70s. Hence we conclude that first wave of covid mainly affected people aged over 70, while the second wave only has influence on people aged under 50.

Discussion

The results of this analysis agrees with the government's opinion. The first wave of covid from March to May has a much larger influence on death counts of older people than the deaths counts of younger people. The second wave of covid starting in September increases the death counts of people aged under 50, whereas the death counts of people aged over 70 is close to the historical values.

One possible reason for this phenomena is, older people are vulnerable to virus and hence are seriously affected by the first wave of covid. In the second wave, the government regulations protect majority of the older people, however, some younger people no longer cooperate with government regulations. They start to behave irresponsibly, as a result, the death counts of younger people increases in the second wave.

References

The Quebec death count dataset used in this report is available from
https://www.stat.gouv.qc.ca/statistiques/population-demographie/deces-mortalite/nombre-hebdomadaire-deces_an.html

Appendix: Graphs and Tables

Graphs

Overviews of sample data

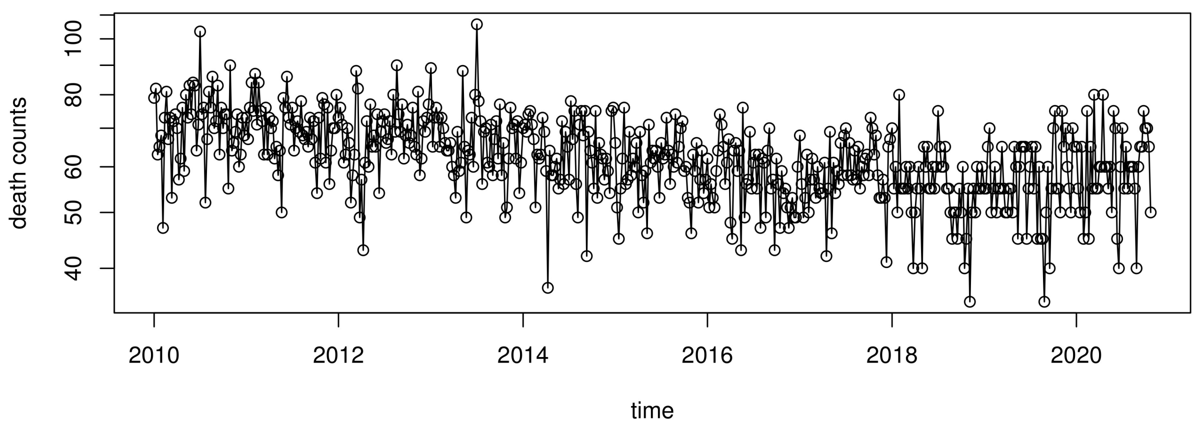

Figure 1: Death Counts of People Aged Under 50 from 2010 to 2020

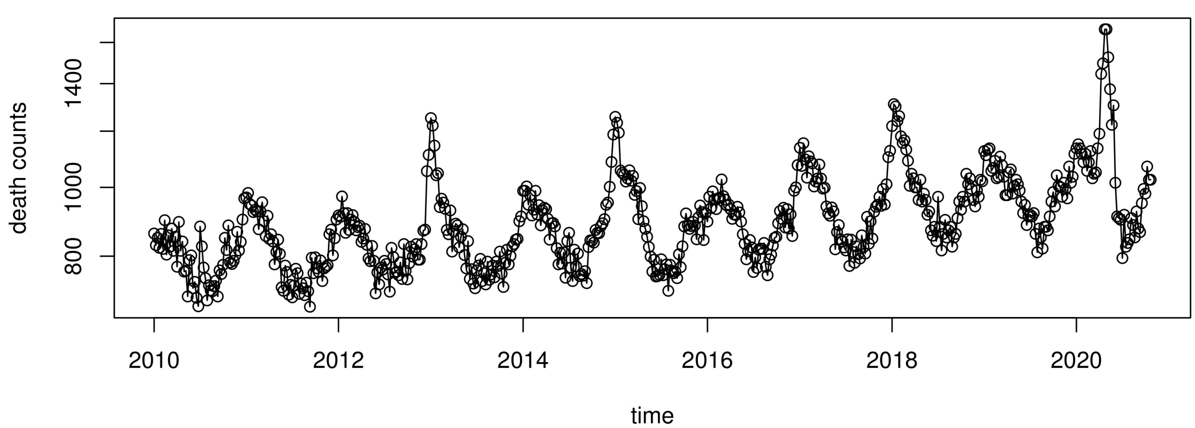

Figure 2: Death Counts of People Aged over 70 from 2010 to 2020

Graphs of Predicted Death Intensity

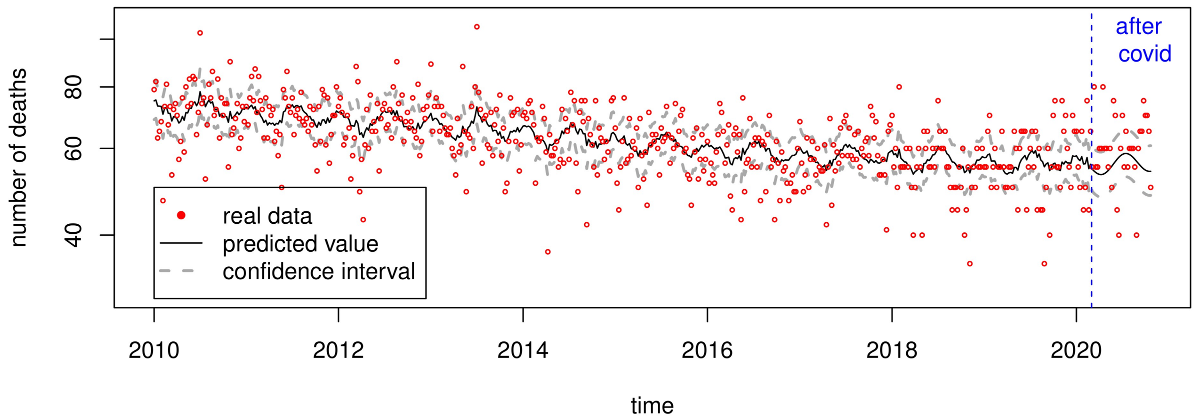

Figure 3: Predicted Death Intensity Over 10 Years (Under 50 Years Old)

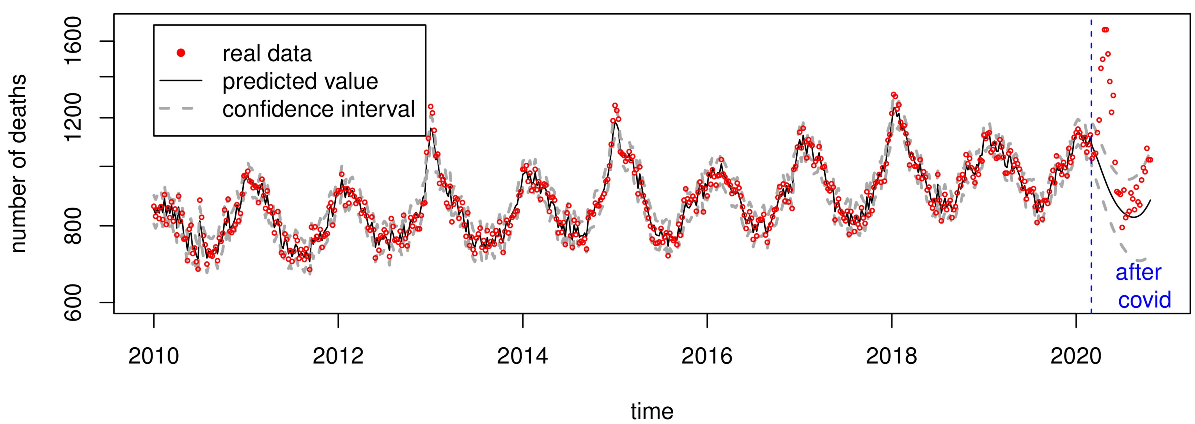

Figure 4: Predicted Death Intensity Over 10 Years (Over 70 Years Old)

Graphs of Predicted Random Effects

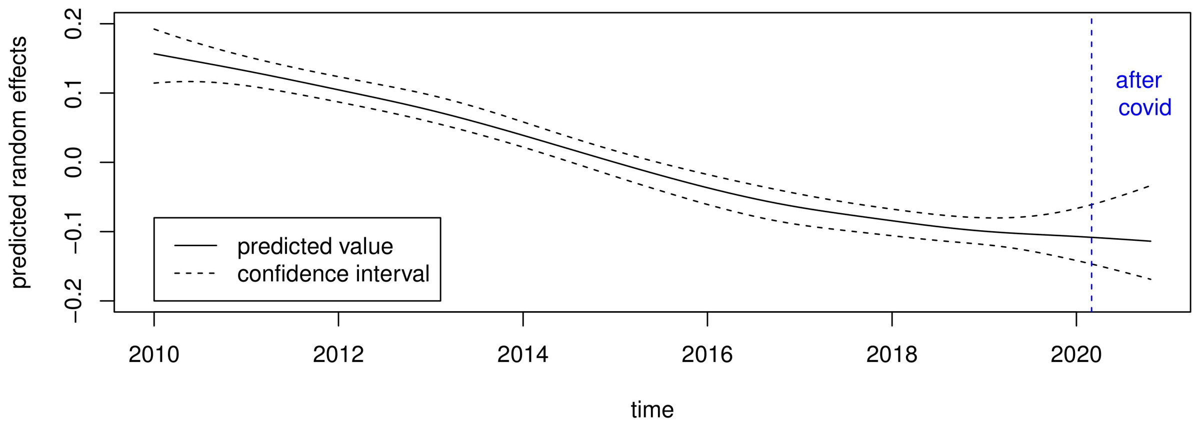

Figure 5: Predicted Random Effects (for Under 50 Years Old)

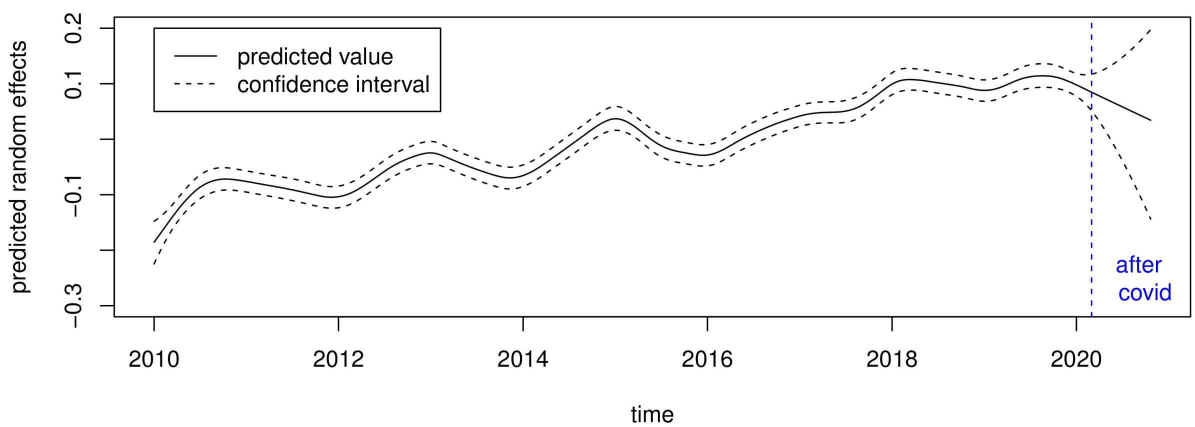

Figure 6: Predicted Random Effects (for Over 70 Years Old)

Graphs of 30 Posterior Samples

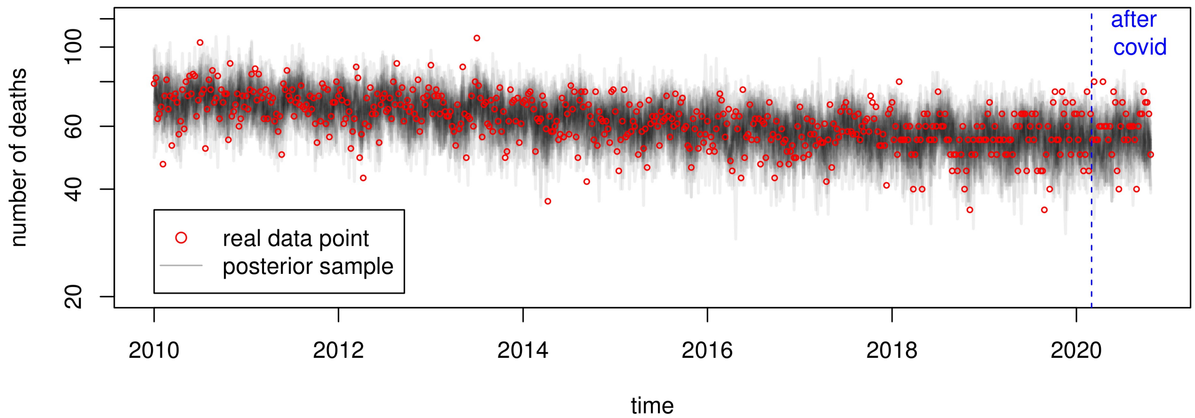

Figure 7: Posterior Samples over 10 years (Under 50 Years Old)

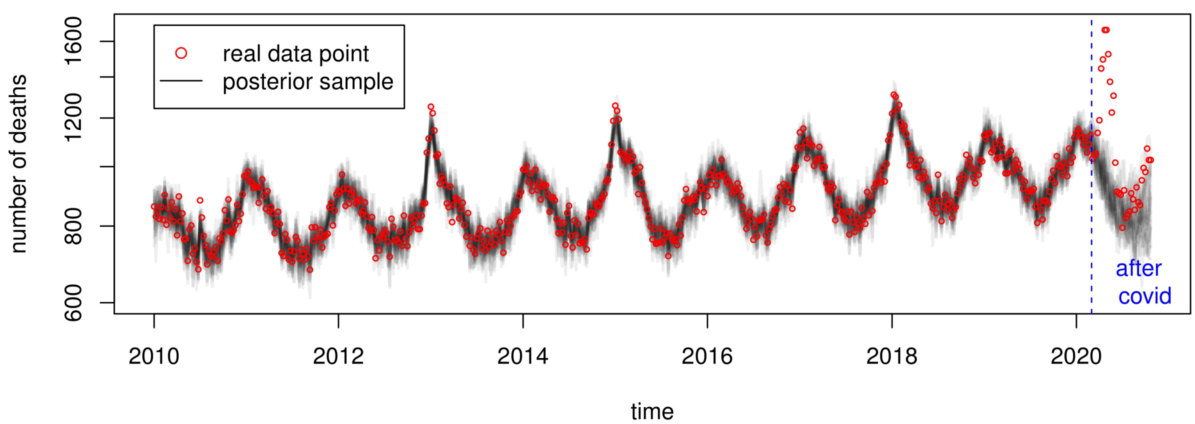

Figure 8: Posterior Samples over 10 years (Over 70 Years Old)

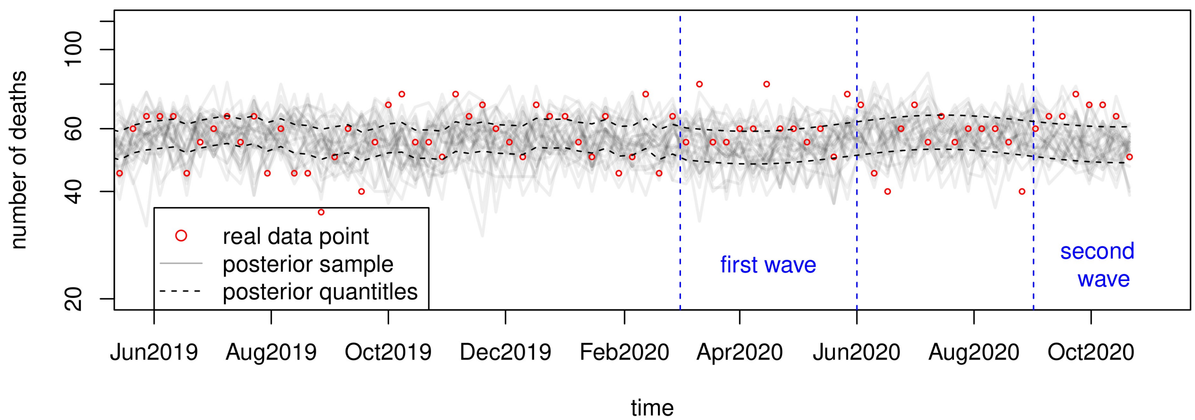

Figure 9: Posterior Samples over past year (Under 50 Years Old)

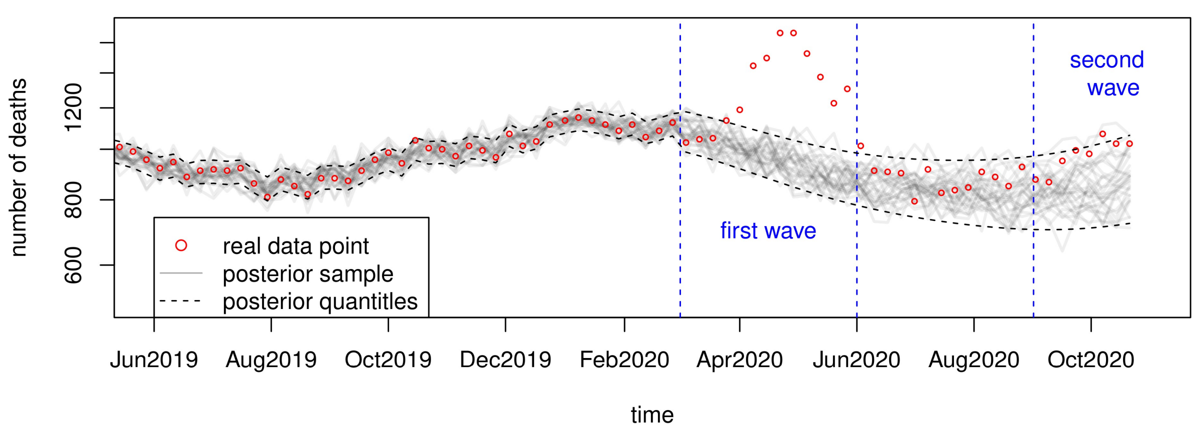

Figure 10: Posterior Samples over past year (Over 70 Years Old)

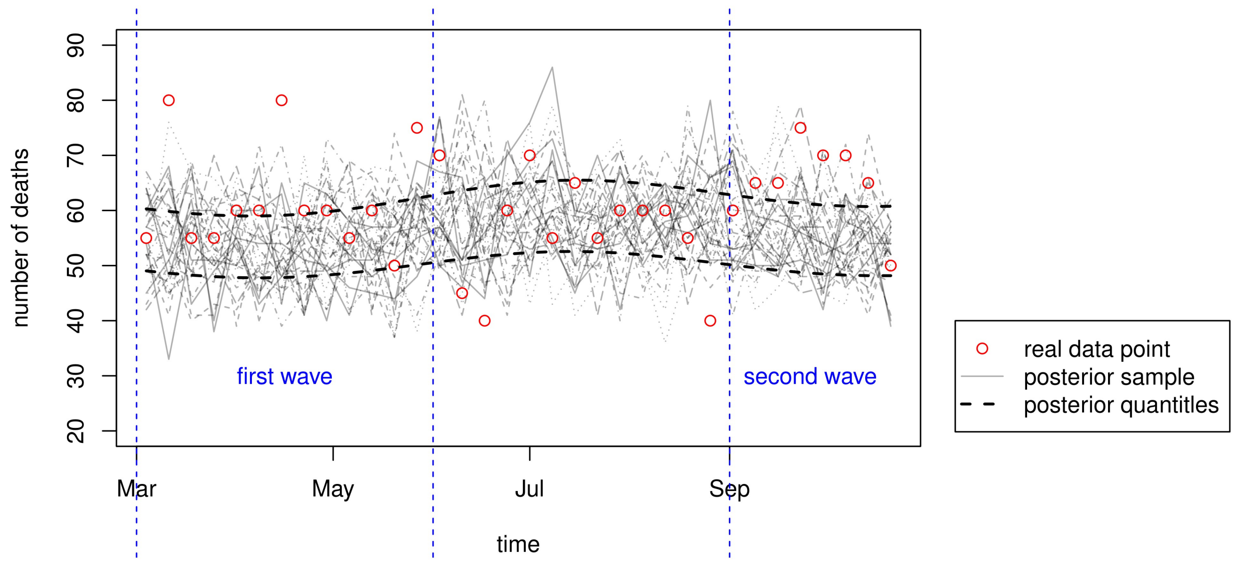

Figure 11: Posterior Samples post covid (Under 50 Years Old)

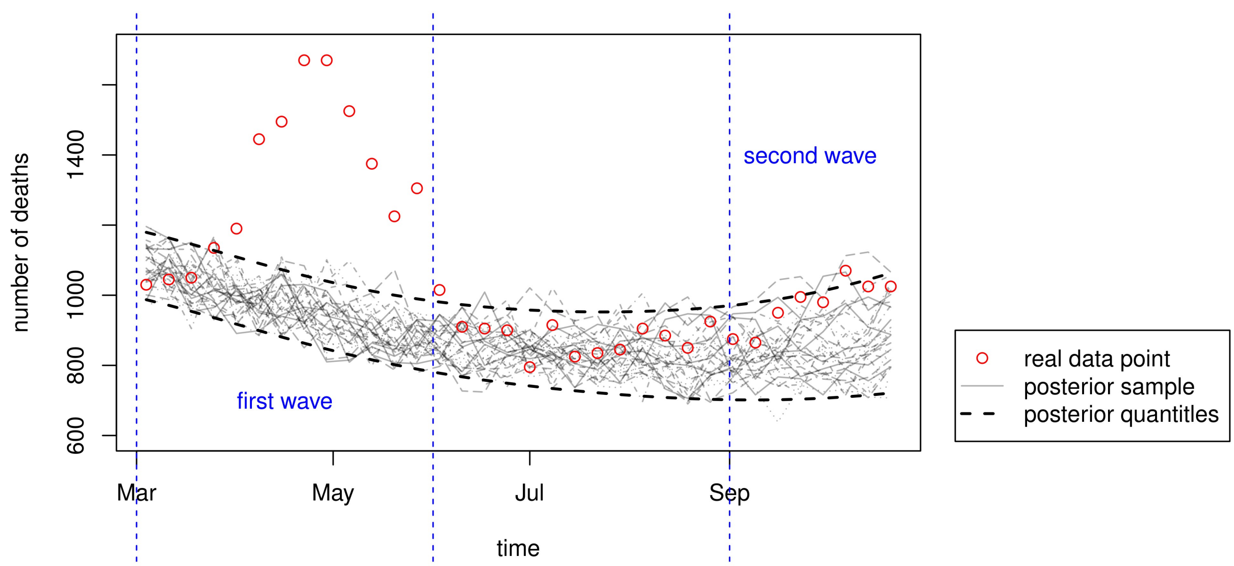

Figure 12: Posterior Samples post covid (Over 70 Years Old)

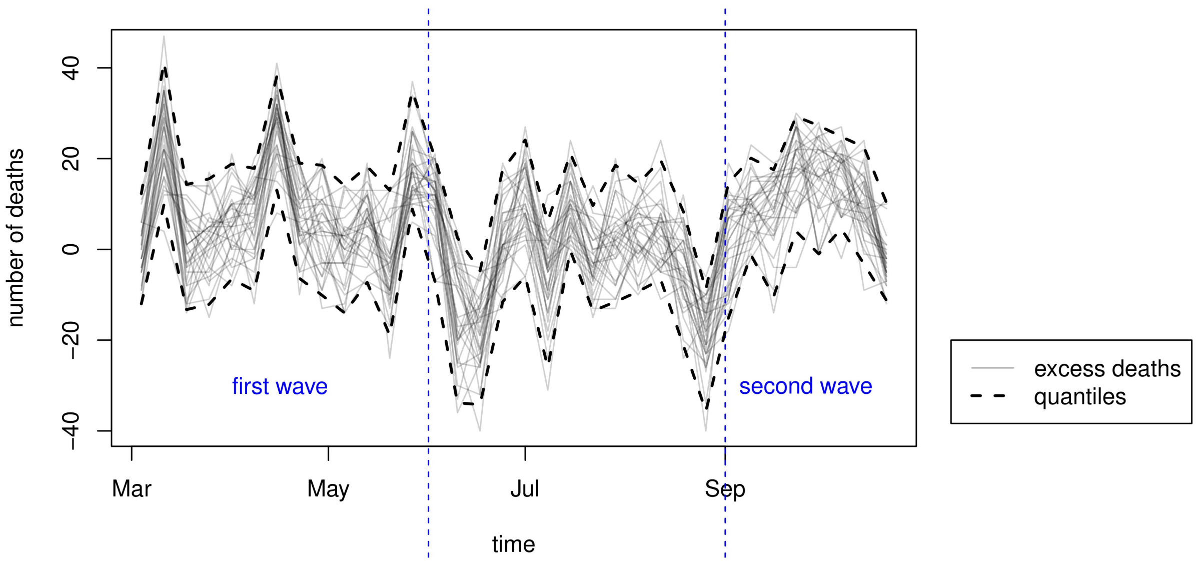

Figure 13: Posterior Samples of Excess Deaths since covid (Under 50 Years Old)

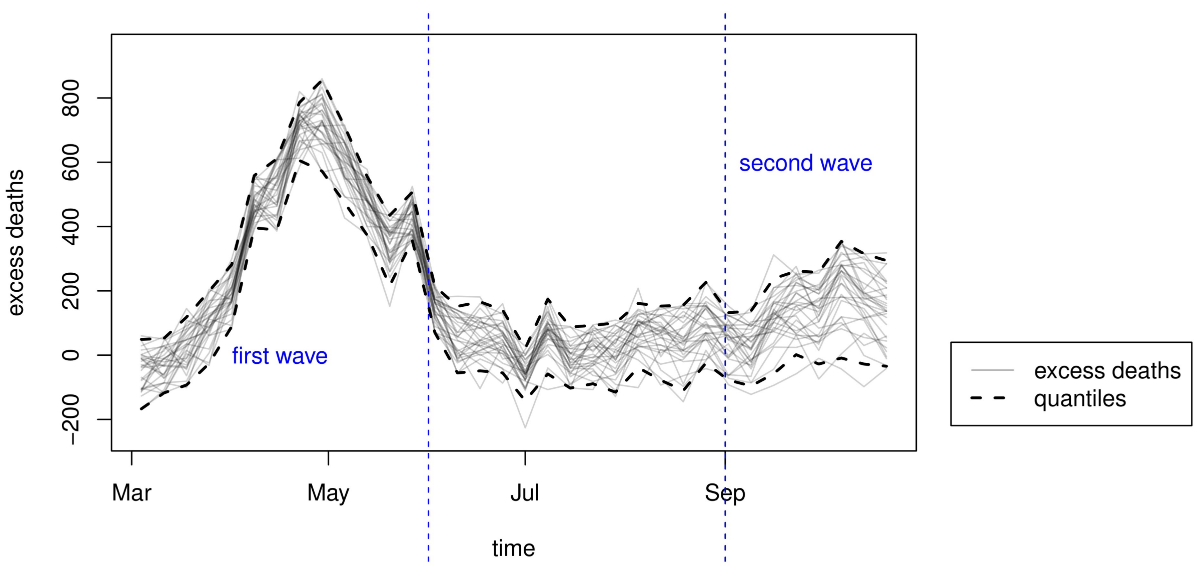

Figure 14: Posterior Samples of Excess Deaths since covid (Over 70 Years Old)

Tables

Tables of posterior fixed effects and random effects

Table 1: Posterior quantiles of fixed effects and random effects (under 50 years old)

	0.5quant	0.025quant	0.975quant
(Intercept)	4.129	4.116	4.141
sin12	-0.011	-0.028	0.005
sin6	0.013	-0.004	0.029
cos12	-0.011	-0.027	0.005
cos6	0.037	0.021	0.053
SD for timeIid	0.043	0.024	0.064
SD for timeForInla	0.012	0.003	0.037

Table 2: Posterior quantiles of fixed effects and random effects (over 70 years old)

	0.5quant	0.025quant	0.975quant
(Intercept)	6.785	6.778	6.793
sin12	0.064	0.056	0.071
sin6	0.011	0.005	0.018
cos12	0.117	0.109	0.125
cos6	0.012	0.005	0.018
SD for timeIid	0.041	0.036	0.045
SD for timeForInla	0.161	0.107	0.240

Tables of excess deaths (real data - predicted samples)

Table 3: Sample Quantiles of Excess Deaths from March to May Inclusively

	under.50.years.old	over.70.years.old
2.5%	36	3718
50%	99	4591
97.5%	176	5001

Table 4: Sample Quantiles of Excess Deaths from Sept 1 to Oct 21

	under.50.years.old	over.70.years.old
2.5%	30	-240
50%	76	1048
97.5%	124	1838