The Effectiveness of Pfizer's Shot: Evidence from Israel's Vaccination Campaign Between December 2020 and May 2021

Nicholas K. Branigan

June 4, 2021

1 Introduction

In March 2020, the World Health Organization declared a pandemic over the novel coronavirus (COVID-19). By then, it was clear that if we were to avoid the calamity of the virus rampaging through the population, we had to develop impressive vaccines, and develop them fast. So, when eight months later Pfizer announced results from the Phase III trial of its candidate, BNT162b2, the news was almost too good to believe. Not only had the company developed a safe and effective jab in well under a year, an accomplishment without historical parallel, but its vaccine was also among the most effective to have ever been created against any disease [8][9][10]. BNT162b2 was a triumph—and a reminder of how scientific knowledge will quash human suffering.

In this project, I examine how BNT162b2 is performing outside of the pristine setting of clinical trials. I estimate its effectiveness at preventing symptomatic infection, hospitalization, and death by applying a Bayesian negative binomial linear regression to data from Israel between December 2020 and May 2021. I estimate the vaccine's effectiveness at preventing infection beginning 14 days after the second dose. I have chosen this threshold for consistency with guidance from national health authorities like the CDC regarding when the vaccine reaches full protective power [4].[1] I estimate effectiveness at preventing hospitalization and death beginning 7 days after the second dose.[2] For hospitalizations and deaths, I perform a naive and a lagged analysis. I run each of these estimations over two groups: those under 60 and those over and including 60.[3] Thus, in all, I fit 10 models.

2 Data

2.1 Overview

Data are from the Israeli government's COVID-19 hub and cover all vaccinations, infections, hospitalizations, and deaths from COVID-19 in the state between December 20, 2020 and May 22, 2021 [5]. The website is in Hebrew. I used Chrome's translation feature to translate the pages into English and identify the three dataset that are the basis of this project. Helpfully and somewhat surprisingly, the dataset comma separated values files are in English. One dataset consists of the number of vaccinations administered in each of the age groups 0-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, and 90+ for each day between December 20,

[1] An apparently little-known fact is that Pfizer's headline estimates from their trial (95% protection against infection with 95% credible interval of 90.3-97.6%) are actually for protection against infection beginning 7 days after the second dose [8]. Meanwhile, the Phase III trial for the other mRNA vaccine, mRNA-1273, set two weeks after the second dose as the threshold for analysis [2]. What is the CDC's justification for taking individuals to be fully vaccinated two weeks after the second dose of BNT162b2 or mRNA-1273 even though BNT162b2 is as effective at one week post dose 2 as mRNA-1273 is at two weeks post dose 2 (in fact, the point estimate for BNT162b2 is higher)? If the CDC has officially commented on this, I am not aware of it. I suspect that the CDC's decision was motivated by a desire to minimize confusion and simplify messaging.

[2] I would have preferred to estimate protection against all events at two weeks following dose 2, but I am constrained by my data for hospitalizations and deaths.

[3] From here on, I refer to the latter group as those over 60.

2020 and May 25, 2021. Another consists of infections recorded in each of the age groups by week, beginning with 12/20/2020-12/26/2020 and ending with 5/16/2021-5/22/2021, and by vaccination status at the time of infection. The 9 mutually exclusive statuses are no vaccination and for each dose: 1-6 days after, 7-13 days after, 14-20 days after, and more than 20 days after. The third dataset consists of hospitalizations and deaths recorded for each week between 12/20/2020-12/26/2020 and 5/16/2021-5/22/2021 for three vaccination statuses: not vaccinated, after first dose, and after second dose.

2.2 Pre-Processing Decisions

A key model variable is the exposure $\lambda_{t,v}$ (see the following section). I now describe how I estimate this for each week for my infections models.

Since infections are reported weekly, I need to estimate weekly exposure figures from the daily data. For a given week, I consider the number of people in some age group to have reached more than two weeks after their second shot to be the number on the Thursday of that week. During the early stages of the vaccination campaign, when the number of vaccinated individuals is growing rapidly, this choice is quite important. For example, during the week of 1/24/2021, the number of people under 60 who have reached more than 2 weeks after their second shot is 4,340 on Monday and 40,689 on Thursday. The difference in the over 60 group is even more pronounced: 1,635 on Monday and 76,949 on Thursday. Using the Monday of that week would likely lead to a substantial underestimate of the number of person days of exposure to infection among the vaccinated groups during that week. I calculate the vaccinated exposure for a week in 100,000 person days by multiplying the number of two-or-more-weeks-after-the-second-shot-vaccinated individuals during that week by 7 and dividing by 100,000. To calculate the unvaccinated exposure, I first determine the total number of individuals in one of the age categories of my analysis (under 60 or over 60). I estimate Israel's population as 9.19 million and the country's fraction of the population over the age of 60 as .1646 [1][11]. To get the vaccinated exposure in 100,000 person days for a given week, I then subtract from the number of individuals in one of the age categories the number in that category who have received at least one dose by the Thursday of the week, then multiply by 7 and divide by 100,000.

My estimation of exposure for hospitalizations and deaths is almost identical to the above procedure. In the preceding paragraph simply replace "infection" with "hospitalization" or "death". Then, more than two weeks after the second shot becomes more than one week after the second shot. Estimating the vaccine's protective effect against hospitalizations and deaths does introduce one subtly. Whereas a symptomatic infection is diagnosed reasonably shortly after an infection, a hospitalization tends to lag the time of infection by weeks, and a death usually lags the infection by even longer. When vaccinations are rapidly rising, as they were during most of the period I am studying, estimating the exposure for these events during week t as the 100,000 person days for the vaccinated group during week t would overestimate the vaccine's effectiveness. To illustrate this, suppose that deaths always lag confirmed infections by two weeks. Then, the deaths in week t should not be normalized by the total risk accumulated over a group during week t but by the risk accumulated during week $t-2$. The mean time from infection to hospitalization is roughly one week and from infection to death is roughly two [3][7]. Thus, in my lagged analysis I calculate week t lagged exposure for hospitalizations as the mean of the exposures for weeks t, $t-1$, and $t-2$, and for deaths as the mean of the exposures for weeks t, $t-1$, $t-2$, and $t-3$. I apply this smoothing over the exposures to account for the fact that there is substantial variability across individuals in the time from infection to hospitalization or death. Whereas some will die three weeks after infection, others will die the same week.

3 Model

3.1 Specification

Let $y_{t,v}$ be the number of cases of COVID-19, hospitalizations from COVID-19, or deaths from COVID-19 for either those under or over 60.

$$y_{t,v} \sim \text{Negative Binomial}\left(\mu_{t,v}, \phi\right),$$

where

$$\mu_{t,v} = \lambda_{t,v} \exp\{\beta^\top x_{t,v}\}, \quad \beta \sim \mathcal{N}(\epsilon, I), \quad \text{and} \quad \frac{1}{\sqrt{\phi}} \sim \mathcal{N}(0, 9).$$

$t \in \{1, \ldots, T\}$ is the week. $v \in \{0, 1\}$ denotes vaccinated or unvaccinated. $\lambda_{t,v}$ is the exposure in 100,000 person days. $x_{t,v} \in \mathbb{R}^{T+1}$ consists of T indicators for the weeks $(x_1, \ldots, x_T)$, and an indicator for whether the group is vaccinated or unvaccinated (x_{T+1}). ϵ determines the location of the Gaussian prior on the regression coefficients. $I \in \mathbb{R}^{T+1 \times T+1}$ is the identity matrix. ϕ controls the dispersion. For the negative binomial distribution, I have used the parameterization

$$\text{Negative Binomial}(y \mid \mu, \phi) = \binom{y + \phi - 1}{y} \left(\frac{\mu}{\mu + \phi}\right)^y \left(\frac{\phi}{\mu + \phi}\right)^\phi.$$

For $Y \sim \text{Negative Binomial}(\mu, \phi)$,

$$\mathbb{E}[Y] = \mu, \quad \text{and} \quad \text{Var}[Y] = \mu + \frac{\mu^2}{\phi}. \tag{1}$$

The efficacy of the vaccine against some event is defined as

$$\left(1 - \frac{\text{events in the vaccinated group per exposure in this group}}{\text{events in the unvaccinated group per exposure in this group}}\right) 100\%.$$

Since

$$\log\left(\frac{\mu_{t,1}}{\lambda_{t,1}}\right) - \log\left(\frac{\mu_{t,0}}{\lambda_{t,0}}\right) = \log\left(\frac{\mu_{t,1}/\lambda_{t,1}}{\mu_{t,0}/\lambda_{t,0}}\right) = \beta_{T+1},$$

the efficacy of the vaccine is $\left(1 - e^{\beta_{T+1}}\right) 100\%$.

3.2 Priors

The random variable parameters of the model that demand priors are β and ϕ. I have chosen $\epsilon_1, \ldots, \epsilon_T$ (the prior means for the week effects) for each sub-model after examining the trend of events in the unvaccinated group over the T weeks. I have set ϵ_{T+1} (the prior for the vaccine's effectiveness) at -2. This translates to a prior belief that the vaccine's effectiveness is, with maximum likelihood, 86.47%, and with probability 0.95 between 63.21% and 95.02%. I have made this choice based on my awareness of Pfizer's Phase III results, as well as other studies that have examined the real-world effectiveness of BNT162b2 [6][8]. My choice also reflects my expectation that the results from the immaculate Phase III setting are roughly an upper bound on the performance in this setting. I have chosen to set my uninformative normal prior on $1/\sqrt{\phi}$ rather than ϕ since ϕ is inversely related to the variance of the negative binomial. Setting the same normal prior on ϕ would predispose the model toward overdispersion, but I have no reason to suspect that the data are especially overdispersed.

4 Model

4.1 Fit

I established the posterior on the model's parameters β and ϕ with Hamiltonian Monte Carlo using the No-U-Turns-Sampler in Stan. I ran 8 chains in parallel and gathered somewhat over 80,000 samples from the posterior.

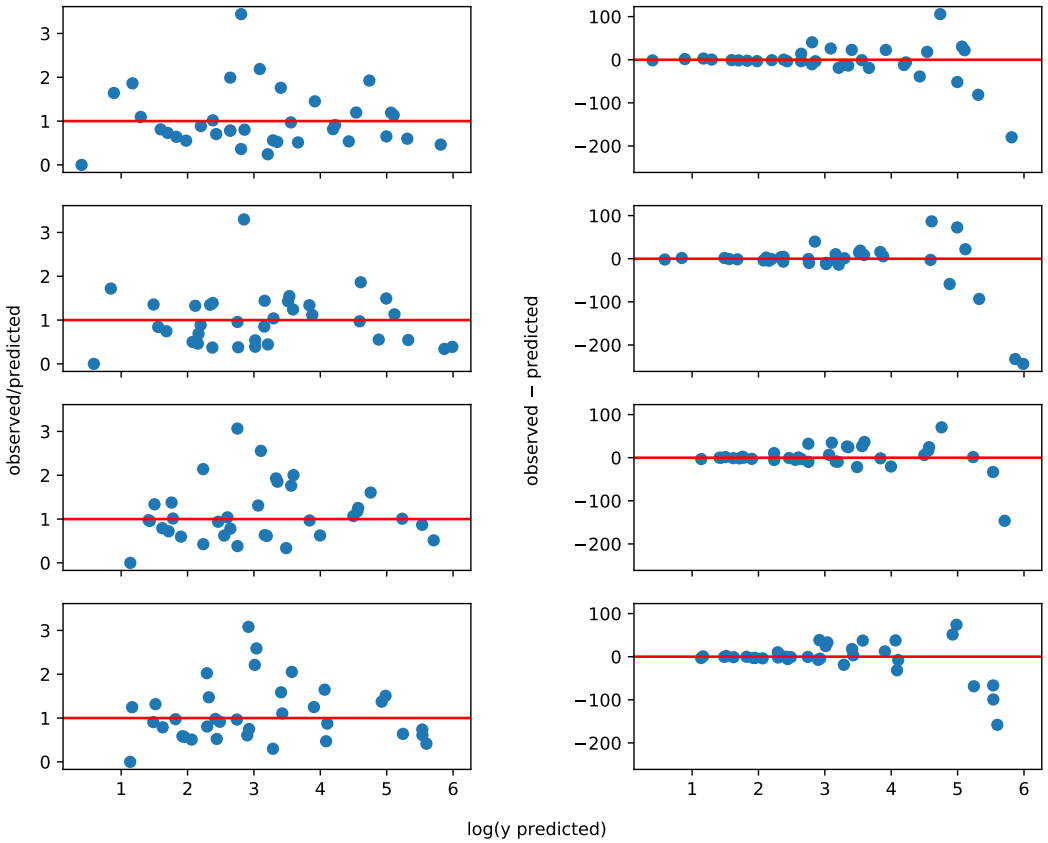

Figure 1: Residual plots for deaths over 60 lagged. Rows correspond to the same random draw from the simulated posterior.

The negative binomial distribution seems to honor the structure of the data, and the models overall seem to fit well. I include residual plots for all 10 of the models in the appendix. For each model, the residual plots reflect predictions made based on four random draws from that model's simulated posterior. I show both ratios of observed to predicted observations as well as the more standard observed minus predicted observations because of the vastly different scales of the observations. In December and January, Israel was experiencing a surge of the virus. But, in the last few months, the country seems to have tamed COVID-19. Case counts are low, deaths are near zero, and the country has almost fully returned to life as it was before March 2020.

Most of the residual plots closely resemble Figure 1. There is no clear pattern in the ratios of observations to predictions, and the model's predictions are (in percent terms) usually close to the observations. Moreover, the absolute value of the difference between observed and predicted grows with the size of the prediction. Though this would be worrying if the model assumed homoscedasticity, as, for example, ordinary least squares does, it would actually be worrying if the model did not exhibit this behavior. As a generalization of the Poisson distribution, the negative binomial's variance scales with the mean. (See equation 1.)

However, not all of the residual plots are unproblematic. In particular, the plots for deaths under 60 in Figure 8 and Figure 12 show a few nasty outliers. For some draws of β and ϕ from the posterior, the observed deaths are 5 or even 10 times the predictions. These misses are happening entirely for recent weeks and for the vaccinated group. I can think of two explanations for the model's failure on these points. The first is that though the model supposes that the vaccine's protective effect is constant over the study period, actually the vaccine's effectiveness has declined since December or January. This cannot be ruled out, but if this has happened, it is only pronounced for the under 60 group. We do not see the same misses, for

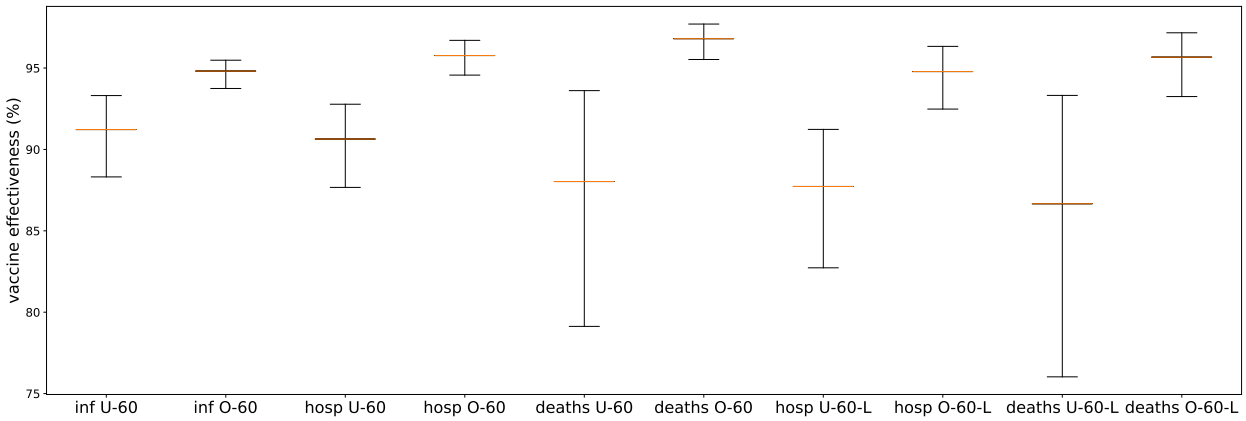

Figure 2: 95% credible intervals for effectiveness of the vaccine in the 10 models. The orange line denotes the estimate of the median of the posterior. "U" denotes "under", "O" denotes "over", and "L" denotes "lagged" (e.g., "deaths O-60-L" means "deaths over 60 lagged").

example, in FIGURE 1.

The second and more plausible explanation is that the degree of risk-aversion with respect to COVID-19 in the under 60 vaccinated group is diminishing over time compared to that in the under 60 unvaccinated group. My estimate of exposure based on person-days assumes that a person-day in the vaccinated and unvaccinated groups results in the same risk for a COVID-19 outcome. But if the members of the vaccinated group are considerably less concerned about the virus than those in the unvaccinated group, this assumption fails and the vaccine effectiveness estimate is biased. If there is a constant difference in risk attitudes over the time period of the analysis, the vaccine effectiveness estimate would be biased but the model would fit fine and we would not expect the undershooting predictions we have in FIGURE 12. So, what seems to be happening is that the difference in risk attitudes between the two under 60 groups is widening over time. This would not be surprising. In the early months, Israel imposed harsh restrictions to control surging case numbers, so exposure opportunities to COVID-19 were far fewer than they have been since late April. Moreover, those under 60 getting shots early on may have been able to reach the front of the line because they were at greater risk for bad COVID-19 outcomes.

Examining the posterior predictive distributions provides another check on the model's fit. In FIGURE 3, I plot these for each week in my deaths over 60 lagged model based on a Monte Carlo estimate using 1000 simulated draws from the posterior. Since the model is supported over all non-negative integers, I am showing only the area that the model assigns 90% probability to beginning with 0, which for this model is the mode. The red circles are the observed deaths. To keep the plots uncluttered, I do not also plot the model's predictions, but these are not equal to the yellow boxes. For the negative binomial distribution, the mean is not necessarily the mode. So, I am not predicting the yellow area only to be consistently missing low. In FIGURE 3 each observation appears on the corresponding week's posterior bar, so each observation's p-value is at least 0.1. And, most of the observations are located high on the bars in areas to which the model assigns high probability. Overall, deaths for those over 60 are reasonably well-accounted for by this model for each week.

4.2 Results

FIGURE 2 displays the estimates for the vaccine's protective effect in each of the 10 models. The vaccine appears to be uniformly more effective among those over 60 compared to under 60. This is likely an artifact of how I am able to estimate exposure, as discussed previously. (There are likely different differences in risk-attitudes toward COVID-19 between the vaccinated and unvaccinated groups under 60 and over 60.) Still, I recover similar effectiveness estimates for the two age brackets. The estimates for protection against

infection line up nicely with Pfizer's Phase III results. This is a pleasant surprise. Vaccines tend to work best in the highly controlled setting of a clinical trial and there has been concern that new variants of COVID-19 are diminishing vaccine effectiveness.

Pfizer's study was under-powered to detect protection against hospitalizations or deaths, but my estimates for protection against these events align with those from other studies that have examined BNT162b2 during its rollout [6]. The estimates from the lagged models are uniformly lower than from the corresponding unlagged models. With rapidly rising vaccinations during much of the analysis time, this is expected. The tightest credible intervals are for over 60, the group that seems to best satisfy the model's assumptions regarding exposure. Credible intervals are wider for the under 60 models, which, per the preceding discussion and the residual plots in the appendix, seem to conform to the modeling assumptions somewhat less well. The widest intervals are for the deaths under 60 models. This is consistent with the residual plots' suggestion that those models fit less well.

5 Conclusion

So far, BNT162b2 has been highly effective in Israel during the state's mass vaccination campaign of its population. The vaccine is highly protective against infection, hospitalization, and death among those under 60 and over 60. This evidence bodes well for the United States and other countries that are widely deploying this vaccine.

6 References

[1] Ofer Aderet. *Israel's population reaches 9.2 million as it celebrates 72 years*. en. URL: https://www.haaretz.com/israel-news/.premium-israel-s-population-reaches-9-2-million-as-it-prepares-to-celebrate-72-years-1.8800599 (visited on 06/04/2021).

[2] Lindsey R. Baden et al. "Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine". In: *New England Journal of Medicine* 384.5 (Feb. 2021), pp. 403–416. ISSN: 0028-4793. DOI: 10.1056/NEJMoa2035389. URL: https://doi.org/10.1056/NEJMoa2035389 (visited on 06/04/2021).

[3] David Baud et al. "Real estimates of mortality following COVID-19 infection". eng. In: *The Lancet. Infectious Diseases* 20.7 (July 2020), p. 773. ISSN: 1474-4457. DOI: 10.1016/S1473-3099(20)30195-X.

[4] CDC. *COVID-19 Vaccination*. en-us. Feb. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/fully-vaccinated.html (visited on 06/04/2021).

[5] *COVID-19 reservoir*. URL: https://data.gov.il/dataset/covid-19 (visited on 06/04/2021).

[6] Eric J. Haas et al. "Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data". English. In: *The Lancet* 397.10287 (May 2021), pp. 1819–1829. ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(21)00947-8. URL: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)00947-8/abstract (visited on 06/04/2021).

[7] Sung-mok Jung et al. "Real-Time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases". en. In: *Journal of Clinical Medicine* 9.2 (Feb. 2020), p. 523. DOI: 10.3390/jcm9020523. URL: https://www.mdpi.com/2077-0383/9/2/523 (visited on 06/04/2021).

[8] Fernando P. Polack et al. "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine". en. In: *New England Journal of Medicine* 383.27 (Dec. 2020), pp. 2603–2615. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa2034577. URL: http://www.nejm.org/doi/10.1056/NEJMoa2034577 (visited on 06/04/2021).

[9] Katie Thomas. "New Pfizer Results: Coronavirus Vaccine Is Safe and 95% Effective". en-US. In: *The New York Times* (Nov. 2020). ISSN: 0362-4331. URL: https://www.nytimes.com/2020/11/18/health/pfizer-covid-vaccine.html (visited on 06/04/2021).

[10] Katie Thomas, David Gelles, and Carl Zimmer. "Pfizer's Early Data Shows Vaccine Is More Than 90% Effective". en-US. In: *The New York Times* (Nov. 2020). ISSN: 0362-4331. URL: https://www.nytimes.com/2020/11/09/health/covid-vaccine-pfizer.html (visited on 06/04/2021).

[11] *World Population Prospects 2019*. Tech. rep. United Nations Department of Economic and Social Affairs Population Dynamics, June 2019.

7 Appendix

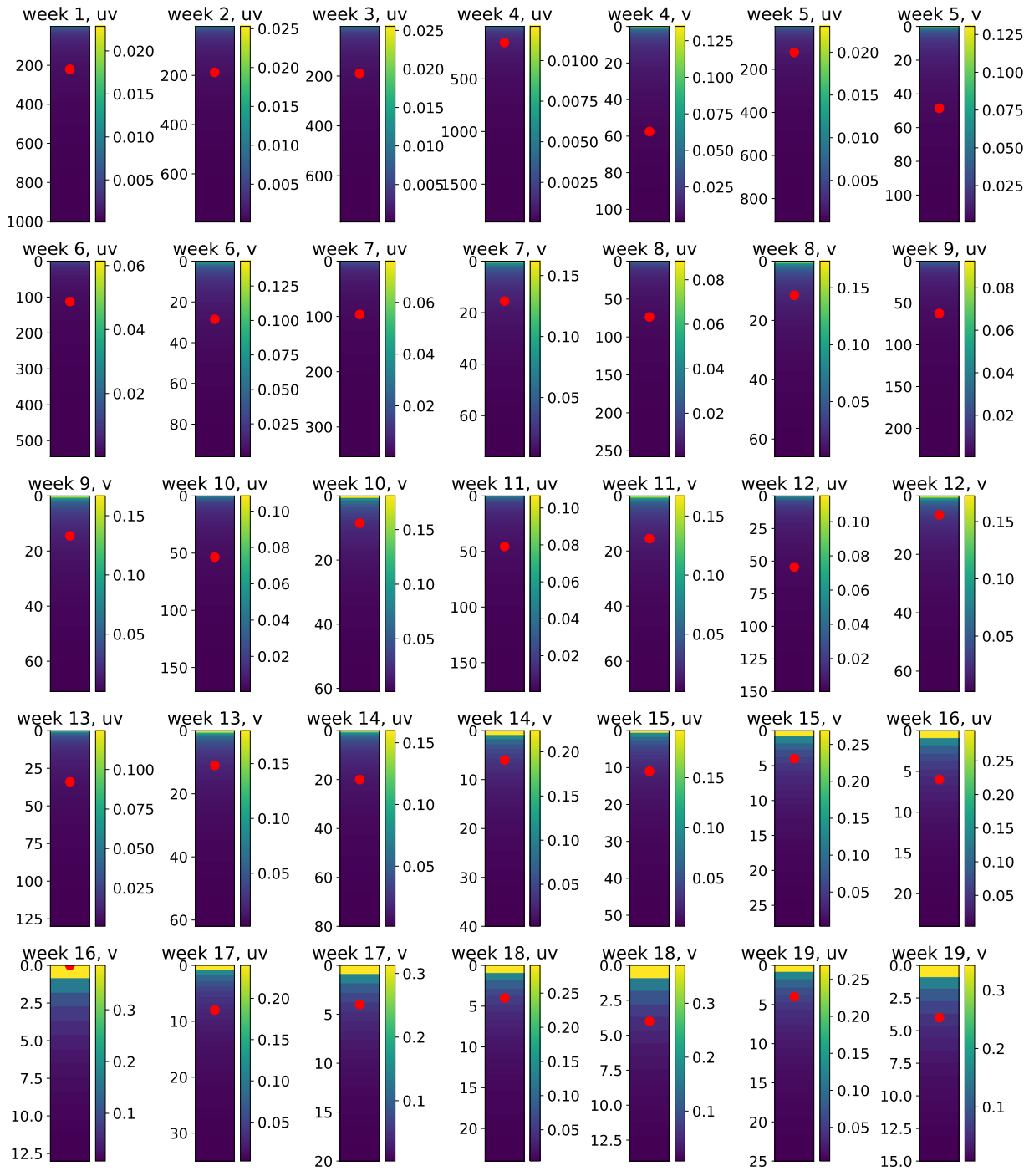

Figure 3: Posterior predictive distribution for the model deaths over 60 lagged. For each week, I am showing the region that is bounded below by the mode and that the model assigns a 90% probability to. The red circle marks the actual number of deaths observed for the week. "uv" denotes "unvaccinated", and "v" denotes "vaccinated."

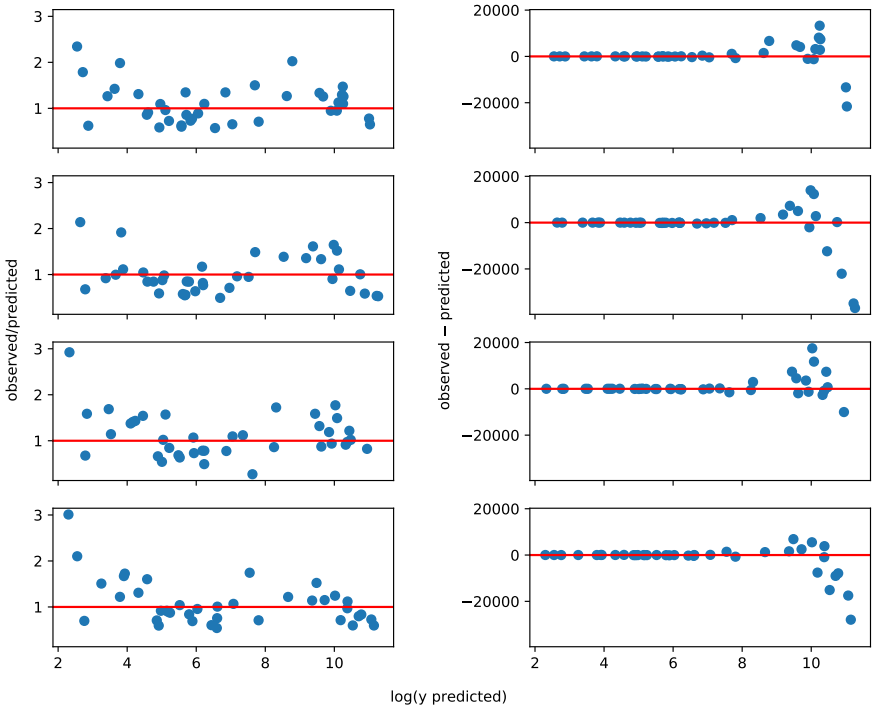

Figure 4: Infections under 60 model residual plots.

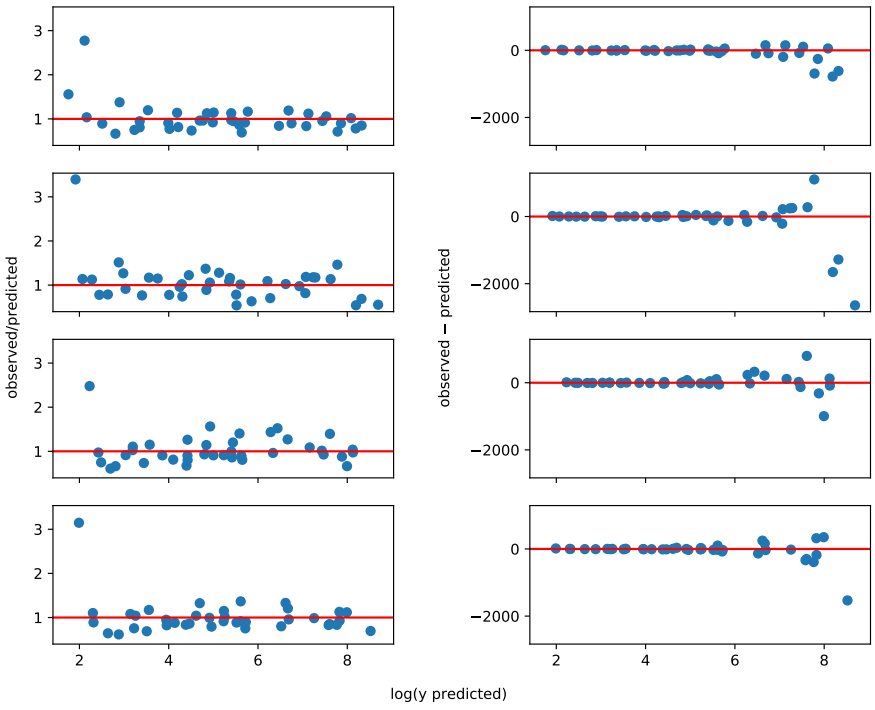

Figure 5: Infections over 60 model residual plots.

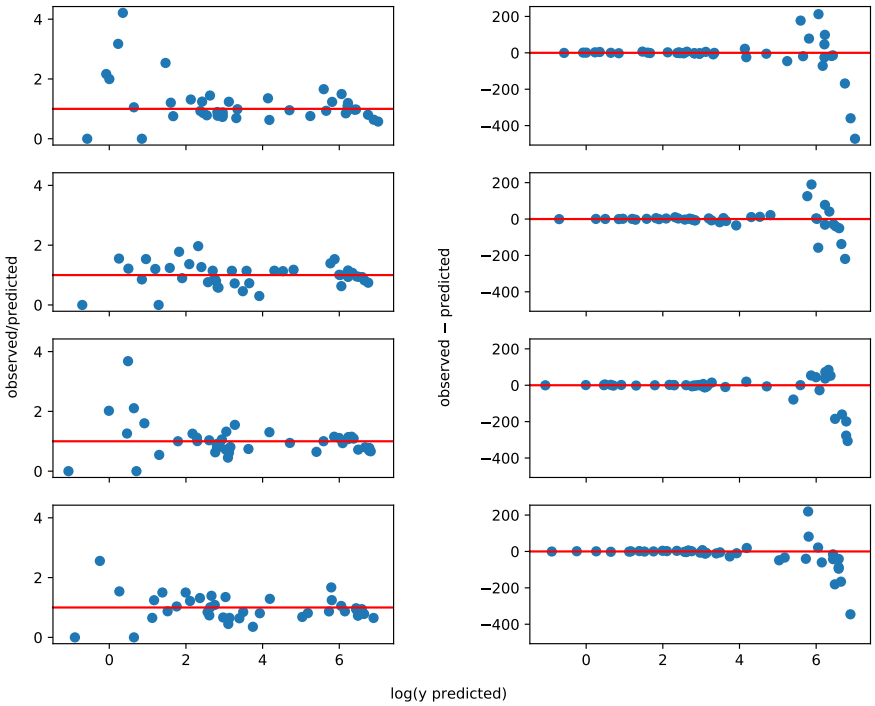

Figure 6: Hospitalizations under 60 model residual plots.

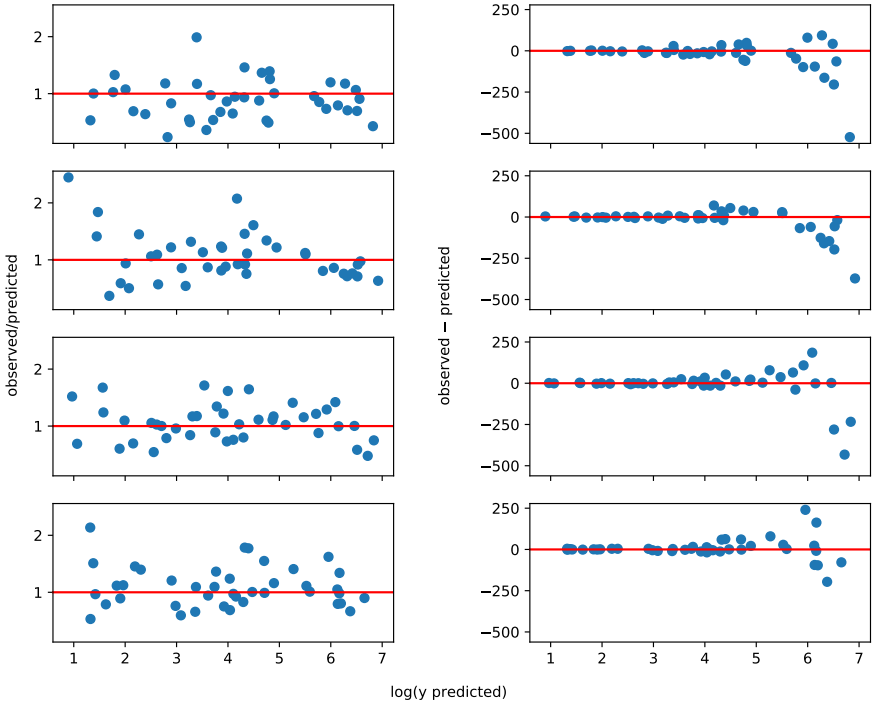

Figure 7: Hospitalizations over 60 model residual plots.

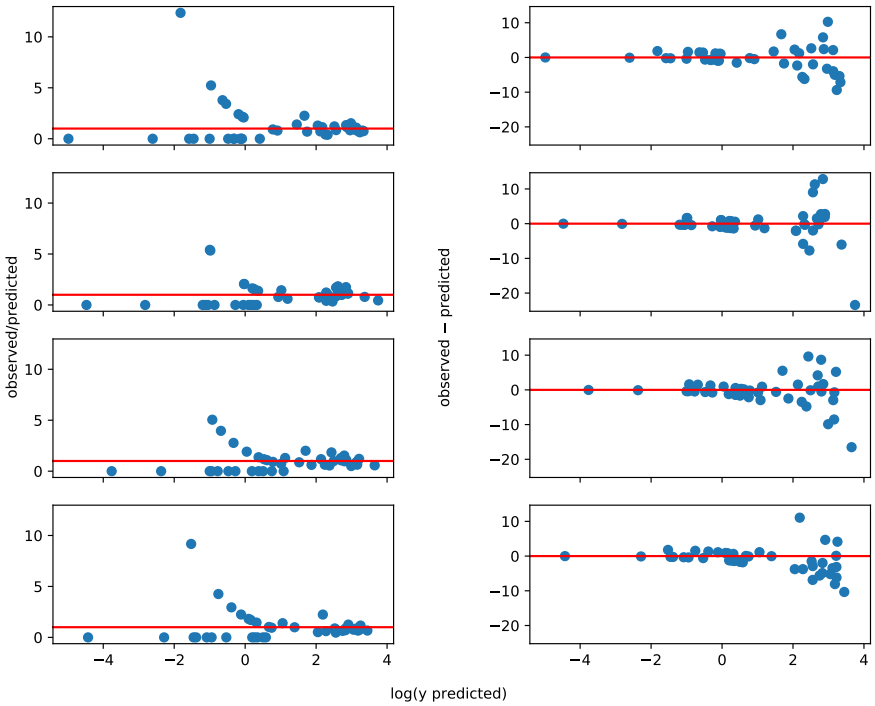

Figure 8: Deaths under 60 model residual plots.

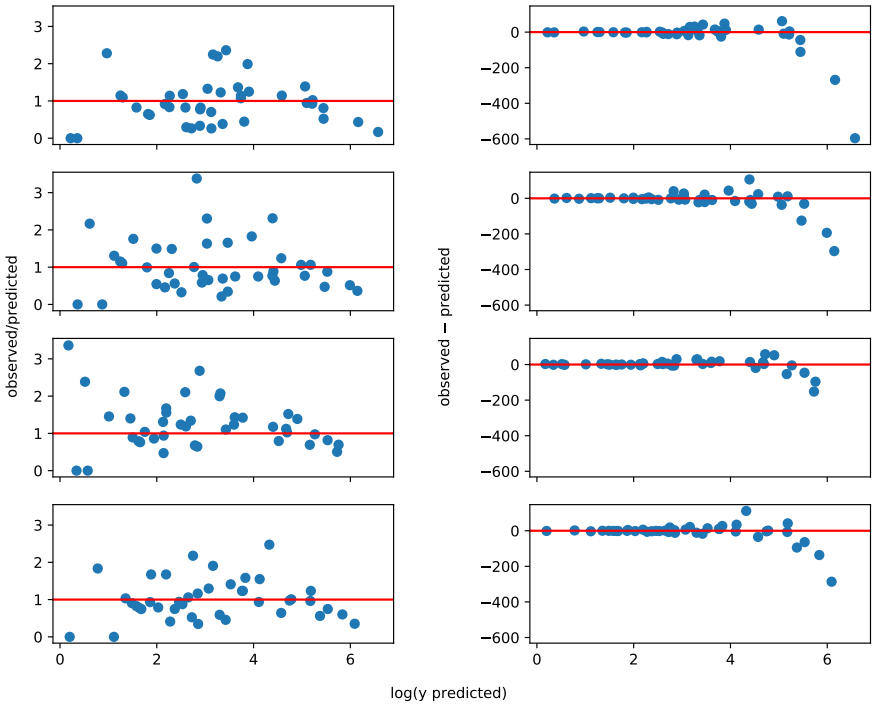

Figure 9: Deaths over 60 model residual plots.

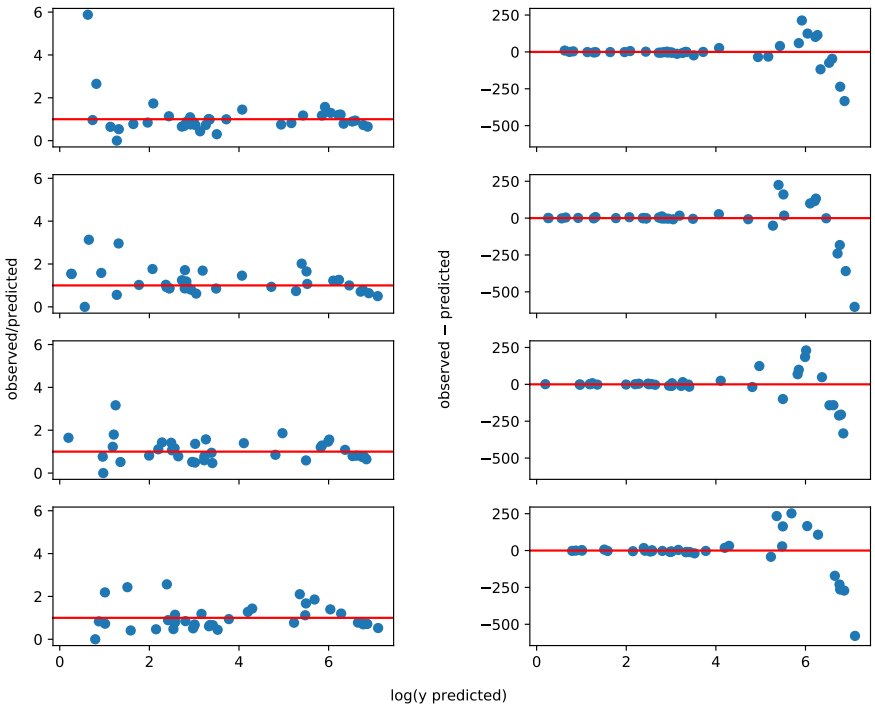

Figure 10: Hospitalizations under 60 lagged model residual plots.

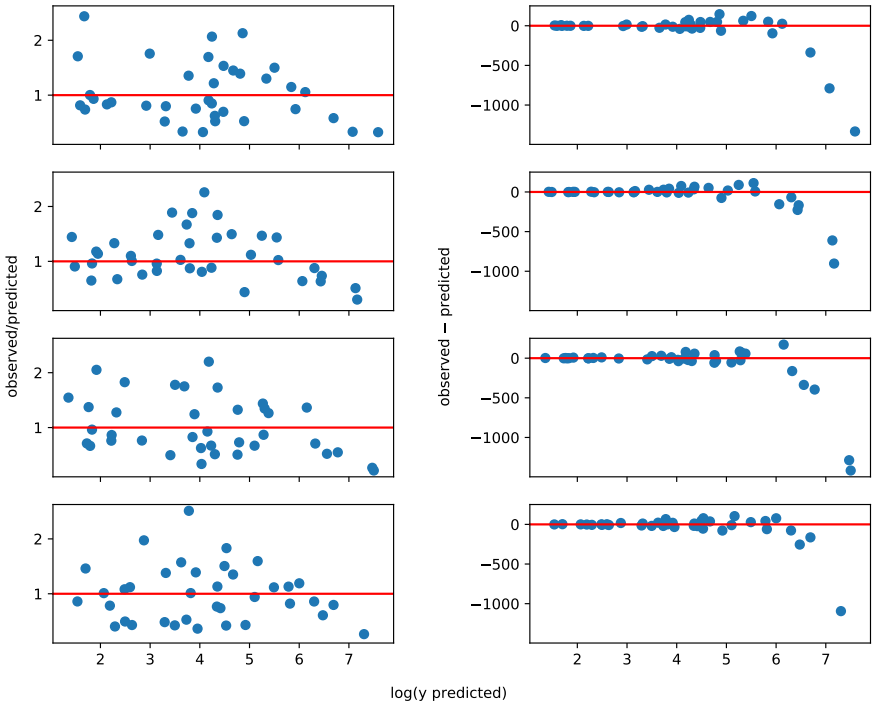

Figure 11: Hospitalizations over 60 lagged model residual plots.

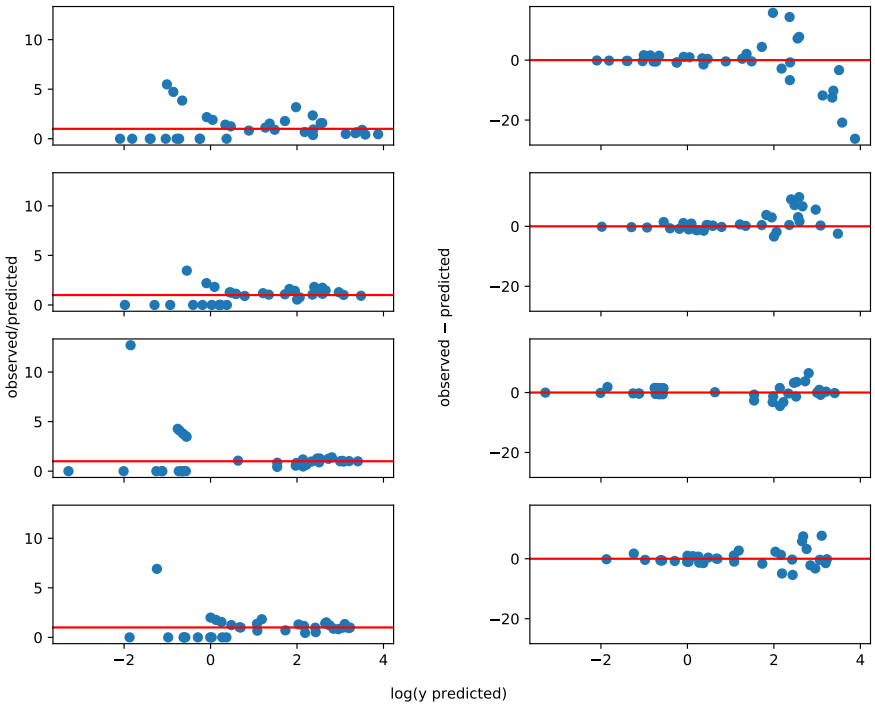

Figure 12: Deaths under 60 lagged model residual plots.