The impact of non-pharmaceutical interventions on early-stage COVID-19 epidemic dynamics in rural communities in the United States.

Punya Alahakoon[1, 2], Peter G. Taylor[1], James M. McCaw[*,3]

[1]School of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia.

[2]Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

[3]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia.

Abstract

COVID-19, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has affected millions of people around the globe. We study the spread of SARS-CoV-2 across six rural counties in North and South Dakota in the United States. The study period is from early March 2020 to mid-June 2021, from near the onset of the pandemic to just before the arrival of the *Delta* variant in the United States. We model the transmission dynamics in each county using a stochastic compartmental model and analyse the data within a Bayesian hierarchical statistical framework. We estimate the model parameters and characterise the effects of non-pharmaceutical interventions (NPIs) implemented at the time. Counterfactual analyses in which NPIs were lifted earlier indicate that notified cases may have increased by up to 51% in the case of low NPI stringency levels with potential additional deaths. Our study underscores the importance of timely public health measures and compliance with them. From a methodological perspective, our study demonstrates that despite the inherent high variability in epidemic behaviour in small rural communities, the combination of stochastic modelling and application of Bayesian hierarchical analyses can support the quantitative evaluation of the impact of public health measures in small low population density communities.

[*]jamesm@unimelb.edu.au

1 Introduction

The impact of early-stage COVID-19, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was far from uniform across the world. This impact was notably disproportionate across different regions within the United States (Kettl, 2020; Lurie & Sharfstein, 2023). Although most existing studies have focused on high-density populations (Bo et al., 2021; Kissler, Tedijanto, Goldstein, Grad, & Lipsitch, 2020; Kucharski et al., 2020), rural communities, in particular, remain understudied in this regard and require further investigation (Mueller et al., 2021). Inspired by a news article in the *New York Times* on this very topic (Leatherby, 2020), we focus on understanding the SARS-CoV-2 transmission dynamics across rural counties.

Our study considers infections that occurred in the rural counties Miner, Buffalo, and Faulk in South Dakota (SD) and Towner, Eddy, and Golden Valley in North Dakota (ND). We consider the period before the *Delta* variant gained dominance in the United States (US); that is, from early March 2020 to mid-June 2021. According to a study by the Centers for Disease Control and Prevention (CDC) (Paul et al., 2021), during April 11–24, 2021, the *Alpha* variant represented 66% of US infections, and 5% of infections were from the *Gamma* variant. During this early phase of the pandemic, vaccines were not yet available. Disease transmission was mitigated through non-pharmaceutical interventions such as international travel restrictions, social distancing measures, mask-wearing, and limiting mobility by imposing interventions such as restricting public transport availability, event cancellations, school closures, working from home requirements and stay-at-home orders(Banholzer et al., 2021; Liu et al., 2021; Reiner et al., 2021).

The main objectives of this study are two-fold. Firstly, we aimed to illustrate the importance of a hierarchical analysis for small populations. Given that the populations and final outbreak sizes of all the US counties we considered were small, we modelled each outbreak using a stochastic epidemic model. To obtain a comprehensive understanding of the epidemiological dynamics, we adopted a Bayesian hierarchical analysis that enabled simultaneous modelling of multiple counties. Through this framework, we obtained overall estimates for the entire group of outbreaks whilst still accounting for the inherent variability within each county.

Secondly, after conducting a hierarchical analysis and estimating the epidemiological parameters for these counties, we aimed to explore the impact of the non-pharmaceutical interventions that were deployed in these rural communities on COVID-19 cases and deaths. We quantified this impact through a counterfactual analysis, examining what might have occurred had certain public health measures not been implemented. Our analysis revealed that, compared to the actual interventions in place during the early phase of the pandemic, the timing of relaxing (say, after two weeks, one month, two months) and the change in the stringency of the non-pharmaceutical interventions (modelled as an increase in the force of infection by 20%, 60%, and 80%) dramatically influenced the epidemic dynamics, driving an increase in the number of cases and deaths.

2 Materials and methods

2.1 Data

We sourced publicly available datasets from the COVID-10 Data Hub (Guidotti & Ardia, 2020). These datasets contained daily COVID-19 case numbers, deaths, the presence and magnitude of the public health intervention measures, and the daily vaccine intake from mid-December 2020. We studied three rural counties each from South and North Dakota, all of which had populations of fewer than 2500 people in 2019.

We selected these six counties such that each was sufficiently distant from all others so that it is reasonable to assume that the disease dynamics in each of the selected counties evolved independently from the others (see Figure 1). See Supplementary Material S1 for a comprehensive understanding of the available data and our steps of data pre-processing.

Figure 2 depicts the daily infection trends and vaccine uptake from early March 2020 to mid-June 2021. Additionally, Table 1 summarises the six counties under study. We obtained values for the population sizes of these counties in 2019 from various sources. For further details, Supplementary Material S1 includes these details and a graphical representation of the counties' intervention measures.

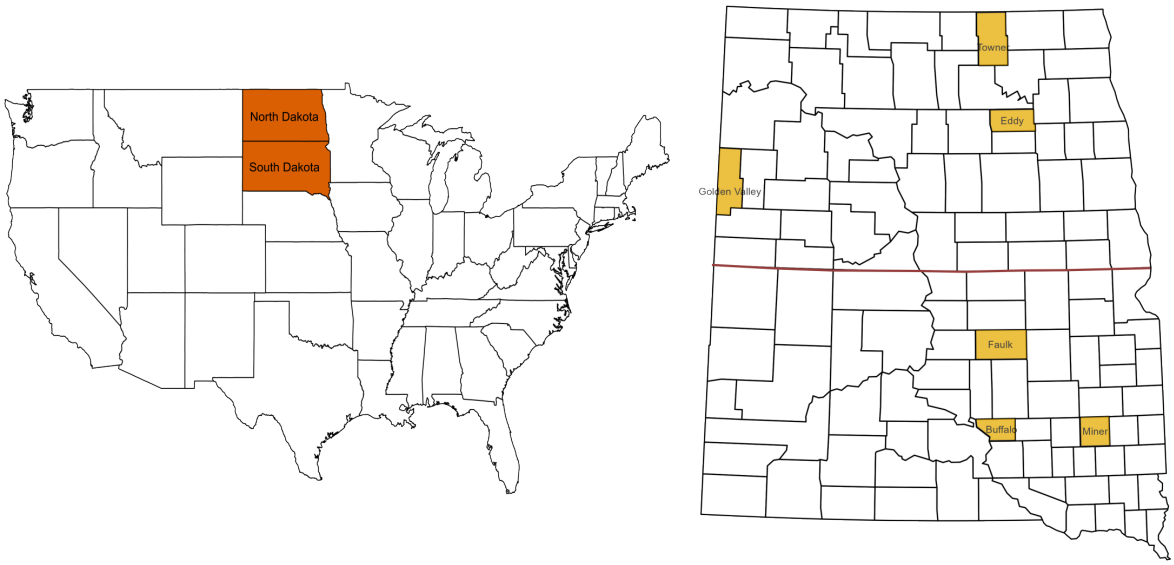

Figure 1: **Left:** Map of the United States with South and North Dakota in brown.
Right: The South and North Dakota counties used in this study are in yellow. Given the substantial distances between the six counties, it is reasonable to assume that outbreaks within each county evolved independently. However, the possibility of disease reintroduction from neighbouring counties (not included in this study) cannot be entirely disregarded.

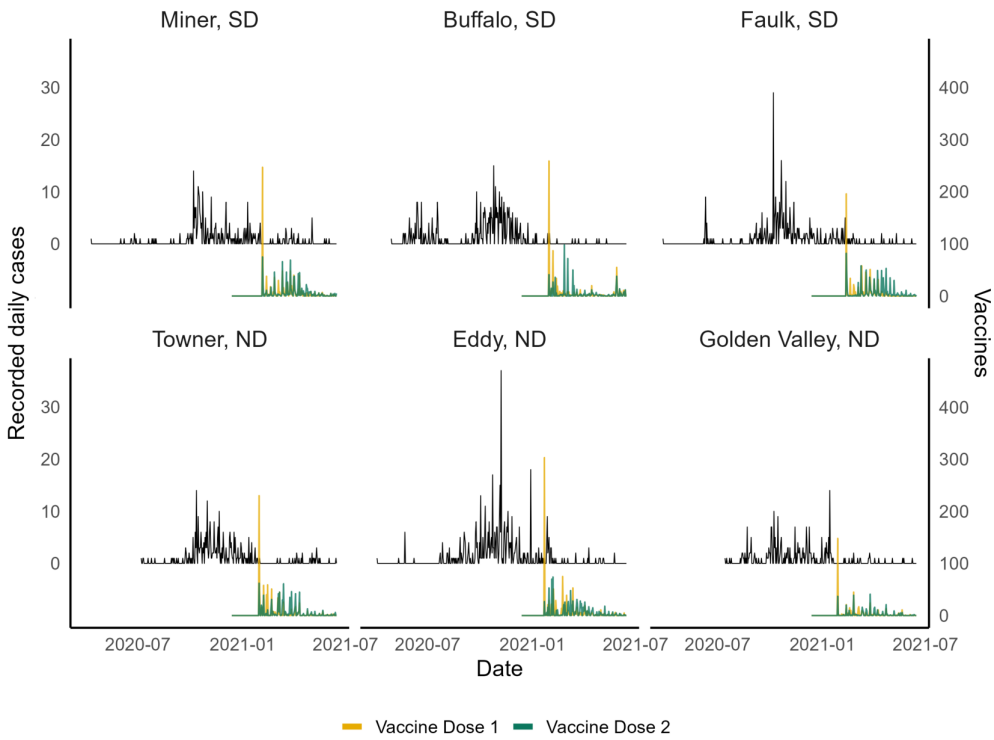

Figure 2: The number of cases and the number of vaccinations recorded each day in Miner, Buffalo and Faulk counties in South Dakota and Towner, Eddy, and Golden Valley counties in North Dakota.

Table 1: Summary of the details of COVID-19 cases from early March 2020 to mid-June 2021

County	**Population size in 2019**	**Total number of infections**	**Total number of deaths due to COVID-19**	**Total number of first doses of vaccine**	**Total number of second doses of vaccine**
Miner, SD	2211	308	10	913	882
Buffalo, SD	2026	436	13	794	659
Faulk, SD	2312	377	13	753	877
Towner, ND	2224	330	11	830	801
Eddy, ND	2301	488	6	984	927
Golden Valley, ND	1845	266	2	293	271

2.2 Modelling and assumptions

We modelled each outbreak using a stochastic epidemic model. Let $N(t)$ represent the initial population size of a county at time t. Our assumptions are as follows. A single exposed person in the population will initiate the epidemic in each county. Furthermore, additional exposed persons can enter the county at the rate of b. Before vaccine availability, all individuals were fully susceptible to the disease. Once a fully susceptible person (S) becomes exposed (E) to an infectious person, they can follow two paths: either become asymptomatically infectious (A_1 and A_2) or pre-symptomatically infectious (P). An asymptomatic person will eventually recover (R_A). A pre-symptomatically infectious person will then be symptomatically

infectious (I). Once an infectious person starts to show symptoms, they will either choose to get tested and notify the health authorities (C) or else they may decide neither to get tested nor notify (I_2). We assume that the observed time-series data of the daily infections consists of individuals in the C compartment. Furthermore, we assume that individuals in the C compartment isolate immediately and take no further part in the transmission process. A person in C and I_2 compartments can die (with death rate d) or else they will eventually recover (R). Given our focus on the early phase of the COVID-19 pandemic, before the emergence of immunity-avoiding variants of the SARS-CoV-2 virus, we do not allow for re-infection in our study period.

Once the vaccines became available in mid-December 2020, we assume that their protection effects would take hold after fourteen days; that is from early January 2021. Once a fully susceptible person is inoculated from the first dose, they enter the V_1 compartment after fourteen days and we assume they gain partial protection. A person in the V_1 compartment can be inoculated with the second dose of the vaccine and will enter the V_2 compartment after fourteen days with improved protection against disease acquisition compared to those in the V_1 compartment. Once exposed to the virus, a person who was previously in either the V_1 or V_2 compartments will follow the transition process of an asymptomatic or symptomatically infectious person as described above. We formulated this model as a continuous-time Markov chain with the transition rates described in Figure 3. We have described the parameters, all the other related assumptions, and information in Supplementary Material S2.

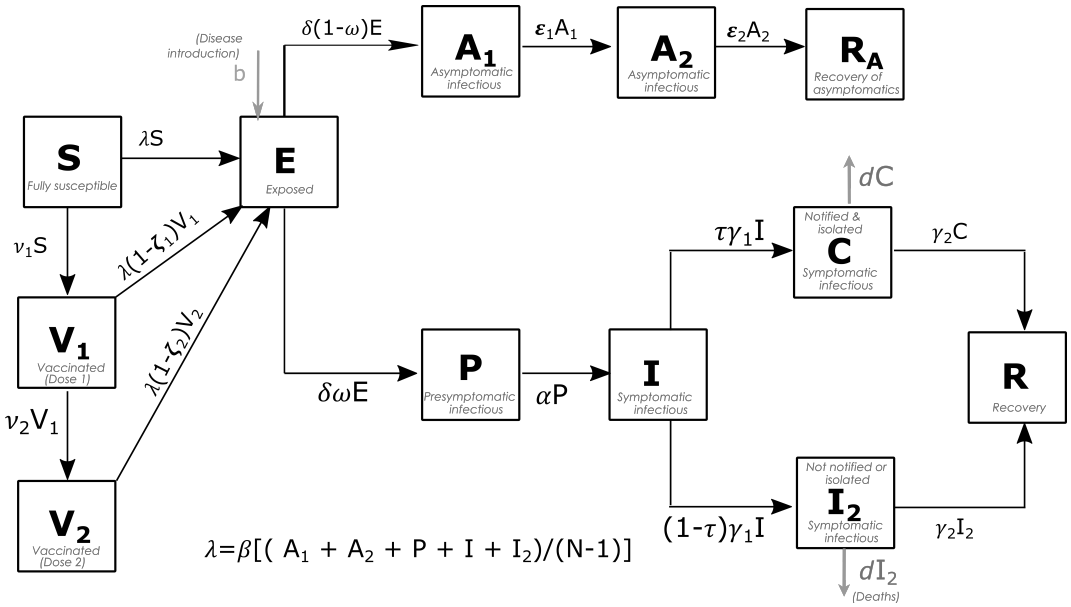

Figure 3: Our proposed model for the outbreaks in the US rural counties.
Here, a fully susceptible (S) person can become exposed (E) to SARS-CoV-2. They can either become asymptomatically infectious (A_1, A_2) and recover (R_A) eventually or become pre-symptomatically infectious (P). A pre-symptomatically infectious person will transition to the symptomatically infectious stage (I). They can either get tested, take part in notifying the health system and isolate (C), or not get tested, notify, or isolate (I_2). A person in C and I_2 stages will eventually recover (R). Deaths are possible in C and I_2 stages. Vaccinated individuals (V_1, V_2) will have reduced force of infection; but once exposed, they will follow the same stages of infection as a susceptible person.

2.3 Estimation framework

We adopted a Bayesian hierarchical analysis approach (Alahakoon, McCaw, & Taylor, 2022; Gelman & Hill, 2006; McGlothlin & Viele, 2018; Royle & Dorazio, 2008) to estimate the model parameters of each county. Under the hierarchical framework, we assumed that the transmission rate (β) and the proportion of individuals who get tested and notified (τ) were sampled from a multivariate truncated normal distribution. We assumed that all the other parameters would be county-specific. We conducted parameter estimation using the two-step algorithm introduced by Alahakoon et al. (2022). The details of the calibration of this algorithm can be found in Supplementary Material S3.

2.4 Policy measures implementation in South and North Dakota

Similar to many other States in the US, South and North Dakota implemented various policy measures designed to mitigate the spread of SARS-CoV-2. These measures varied in severity. The active policy measures during our analysis period were school closing, workplace closing, cancellation of events, stay-at-home restrictions, restricting internal and international movements, information campaigns, testing, contact tracing, facial covering mandates, and elderly people protection. Limitation on gatherings was not active most of the time in South Dakota but it was active for a while (four to five months, on and off) in North Dakota. There were no restrictions on transport in South Dakota whereas transport restrictions were recommended in North Dakota for the first three months of our analysis period. Further details and graphical illustrations regarding policy implementation in these states can be found in Supplementary Material S1.

During our analysis period, as most of these interventions were at least recommended, and by studying the mobility trends analysed from the CDC (see Supplementary Material S6), we assumed that for all the six counties, the percentage of people staying at home was significant, although it remained below 50%. There was also a noticeable reduction in travel to retail and recreation. Workplace mobility changes were negligible.

We were interested in investigating whether intervention measures were successful in mitigating the transmission in these counties. Our goal was to understand how these interventions impacted disease dynamics and to explore potential outcomes if they had been lifted. Therefore, we studied the impact of removing the interventions after two weeks, one month, and two months from the start of our analysis period. For each of these counterfactual scenarios, we studied the impact on the numbers of cases and deaths for each county if the transmission rate increased by 20%, 60%, and 80%. Motivated by studies such as Kissler et al. (2020) and Ferguson et al. (2020), we investigated the impact of non-pharmaceutical interventions across different stringency levels: high, moderate and low. In our counterfactual scenarios, we assumed that these three levels of stringency were represented by increasing the force of infection by 20%, 60%, and 80%. We computed the anticipated impact for each counterfactual as follows. For each sample from the joint posterior distribution of a county, we increased the force of infection by 20%, 60%, and 80% after two weeks,

one month, and two months. Under these nine scenarios, we then generated sample paths from our model explained in Figure 3. Further details can be found in Supplementary Material S7.

3 Results

Figure 4 shows the re-sampled paths of the six counties, generated by re-simulating sample paths from our model using the samples from the joint posterior distribution of each county. These simulated paths present posterior predictive checks (Gelman et al., 2013), which assess the predictive power of our proposed model. Figure 5 illustrates the simulated peak day vs. peak size for these counties. To mitigate the impact of the observed noise, we calculated the seven-day moving averages of the cases in each county and used these to calculate the observed peak time and size. Generally, the generated paths of the notified cases aligned well with the observed paths.

For Miner, the median simulated peak size was 6 (observed peak size was 7), occurring between September and October 2020 (observed peak day: October 17, 2020). For Buffalo, the median simulated peak size was 9 (observed peak size was 7), occurring between September and November 2020 (the observed peak day: October 28, 2020). For Faulk, the median of the simulated peak size was 8 (the observed peak size was 8), occurring between September and November 2020 (the observed peak day: October 10, 2020). For Towner, the median of the simulated peak size was 8 (the observed peak size was 5), occurring between September and October 2020 (the observed peak day: October 13, 2020). For Eddy, the median of the simulated peak size was 9 (the observed size was 10), occurring between September and November 2020 (the observed peak day: November 4, 2020). For Golden Valley, the median of the simulated peak size was 5 (the observed simulated size was 4), occurring between September and November 2020 (the observed peak day: October 6, 2020).

Refer to supplementary Material S5 for other diagnostic figures, such as vaccine intake.

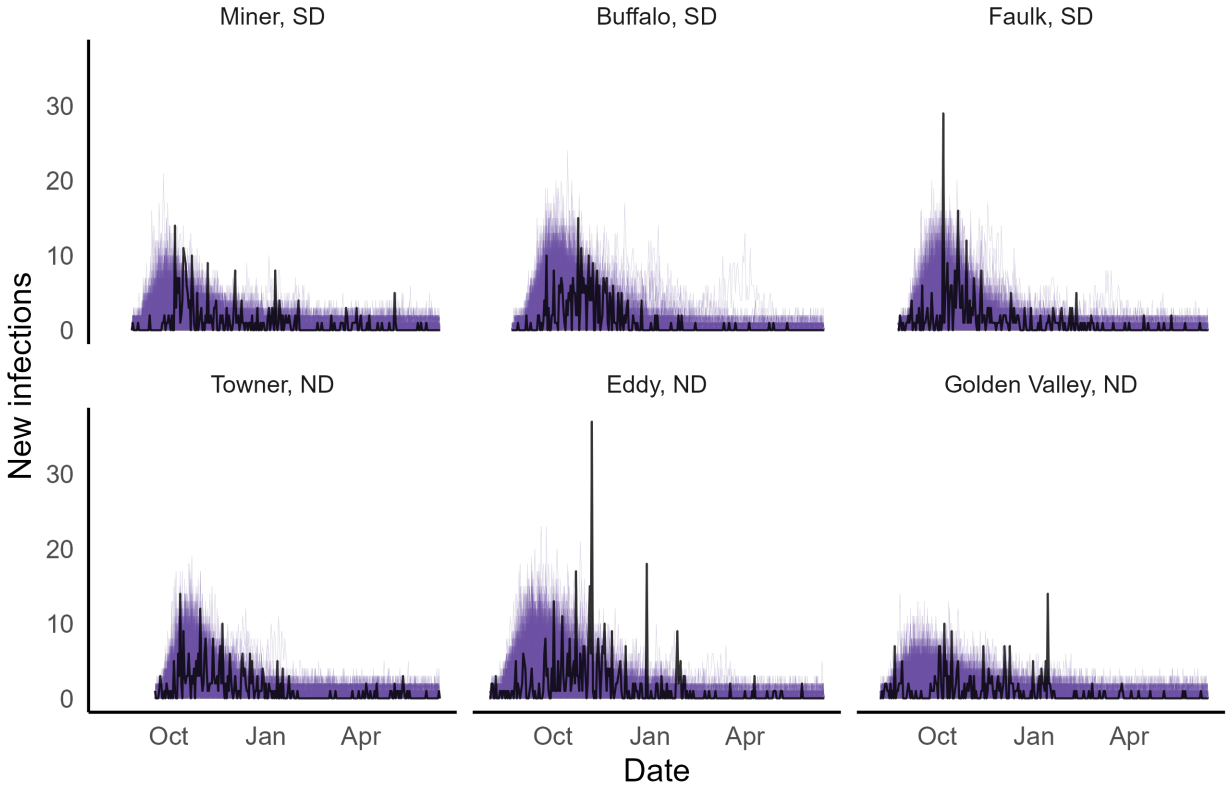

Figure 4: Re-sampled paths for the counties in purple. The observed cases are represented in black.

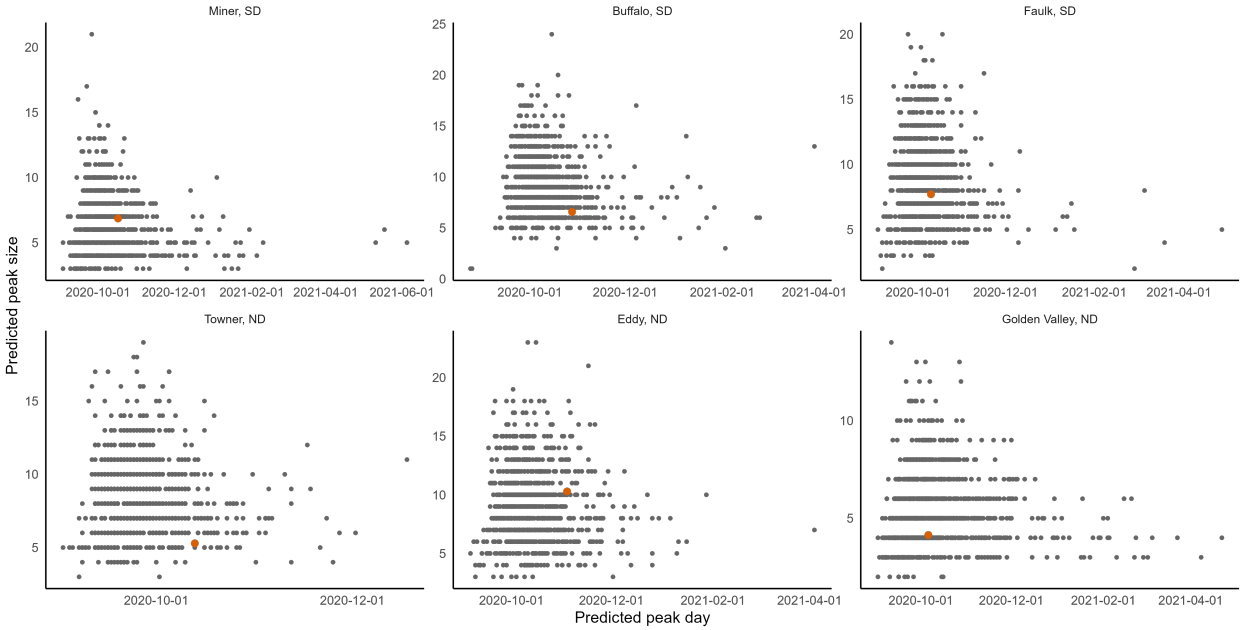

Figure 5: Plots of simulated peak day vs. peak size for the six counties. Observed peak days and sizes (calculated from the seven-day moving average path) are plotted in orange.

Figure 6 displays the posterior distributions of τ, the proportion of symptomatic individuals who got tested, notified and isolated immediately. Table 2 summarises these distributions for all the six counties with posterior median and the 95% Highest Posterior Density (HPD) intervals. Overall, τ for all the counties ranged from 0.27 to 1 with a posterior median closer to 0.6. Among the six counties, Eddy had the highest proportion of symptomatic individuals who got tested, notified and isolated immediately [median 0.637 and (0.320, 0.999) 95% HPD interval] while Towner had the lowest proportion [median 0.585 and (0.260, 0.968) 95% HPD interval].

Refer to Supplementary Material S5.1 for the posterior distributions of other parameters. Additionally, the marginal posterior distributions of the population-level estimate (hyper-mean) for τ had a median of 0.606 [(0.141, 0.998) 95% HPD interval]. Supplementary Material S4 provides further details on this and other hyper-parameters.

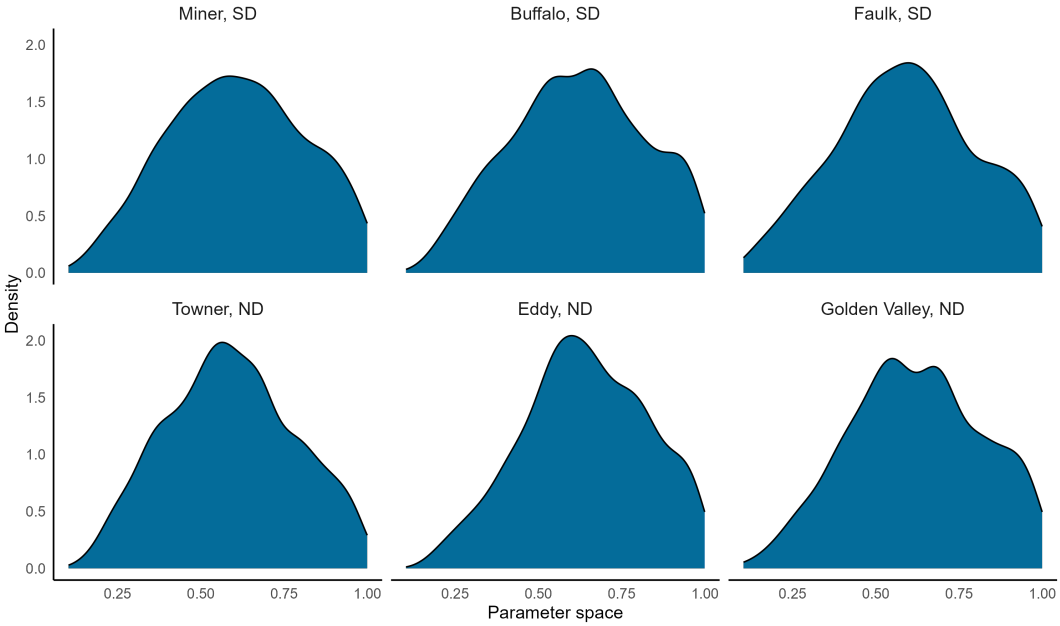

Figure 6: Posterior distributions of τ, the proportion of symptomatic individuals who were tested, notified, and isolated immediately. Each panel corresponds to one of the six counties examined in our study.

Table 2: Posterior medians and 95% HPD intervals for τ, the proportion of symptomatic individuals who were tested, notified, and isolated immediately

County	τ **Posterior median (95% HPD interval)**
Miner, SD	0.604 (0.277, 0.999)
Buffalo, SD	0.626 (0.281, 0.992)
Faulk, SD	0.590 (0.217, 0.971)
Towner, ND	0.585 (0.260, 0.968)
Eddy, ND	0.637 (0.320, 0.999)
Golden Valley, ND	0.617 (0.286, 0.998)

We calculated the basic reproduction number from the estimated parameters using the expression for $R_0 = \beta[(1-\omega)(\frac{1}{\epsilon_1}+\frac{1}{\epsilon_2})+\omega(\frac{1}{\alpha}+\frac{1}{\gamma_1}+((1-\tau)\frac{1}{\gamma_2+d})]$ (see Supplementary material for details). Figure 7 illustrates the distribution of the basic reproduction number (R_0), while Table 3 summarises the values with medians and 25% and 75% quantiles. Overall, R_0 values for all the counties ranged from 1.1 to 1.7 with the median closer to 1.5. Buffalo had the highest R_0 compared to other counties [median 1.415 and (1.306, 1.527) 25% and 75% quantiles] and Miner had the lowest R_0 [median 1.204 and (1.100, 1.305) 25% and 75% quantiles]. Additionally, we observed a high variability for R_0 for Eddy compared to all the other counties.

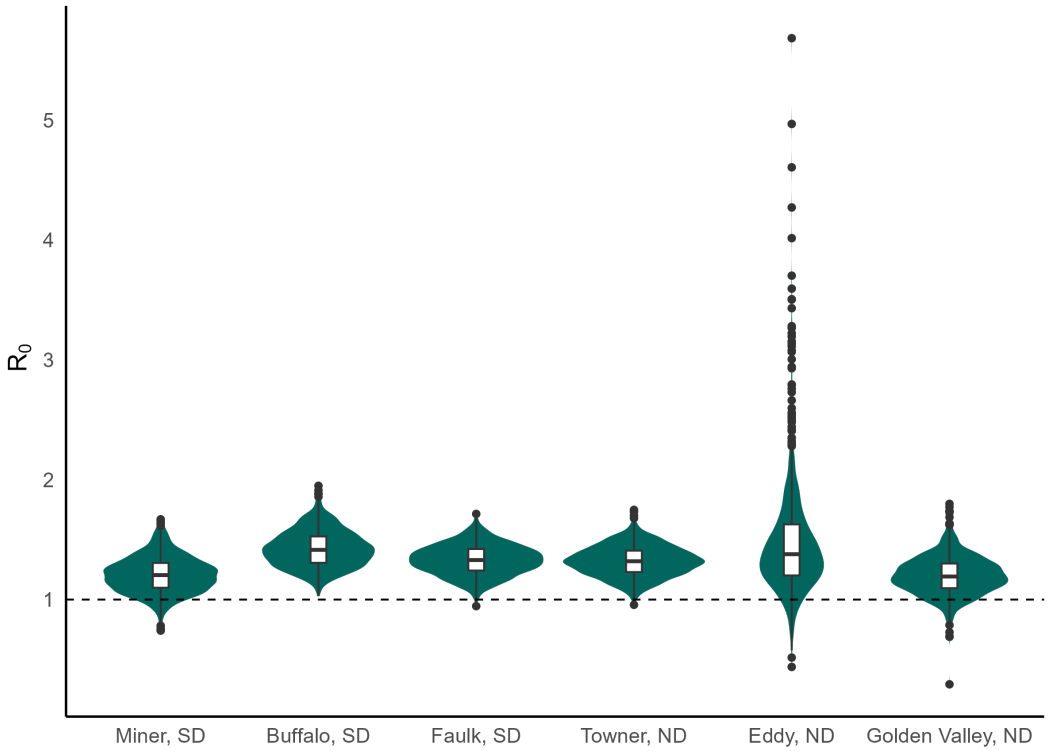

Figure 7: Violin plots of R_0 for the rural counties

Table 3: Estimated R_0s of the counties

County	**Median R_0 (25, 75)% quantile interval**
Miner, SD	1.204 (1.100, 1.305)
Buffalo, SD	1.415 (1.306, 1.527)
Faulk, SD	1.330 (1.243, 1.305)
Towner, ND	1.320 (1.230, 1.408)
Eddy, ND	1.379 (1.202, 1.627)
Golden Valley, ND	1.192 (1.095, 1.300)

Figure 8 shows the impact on the number of COVID-19 cases (that would have been notified) had the force of infection (FoI) increased by 20%, 60%, and 80% after 2 weeks, 1 month, and 2 months for Faulk,

South Dakota. Similar effects were observed in all other counties studied. For a more detailed analysis of the total number of cases and deaths, refer to the Supplementary material.

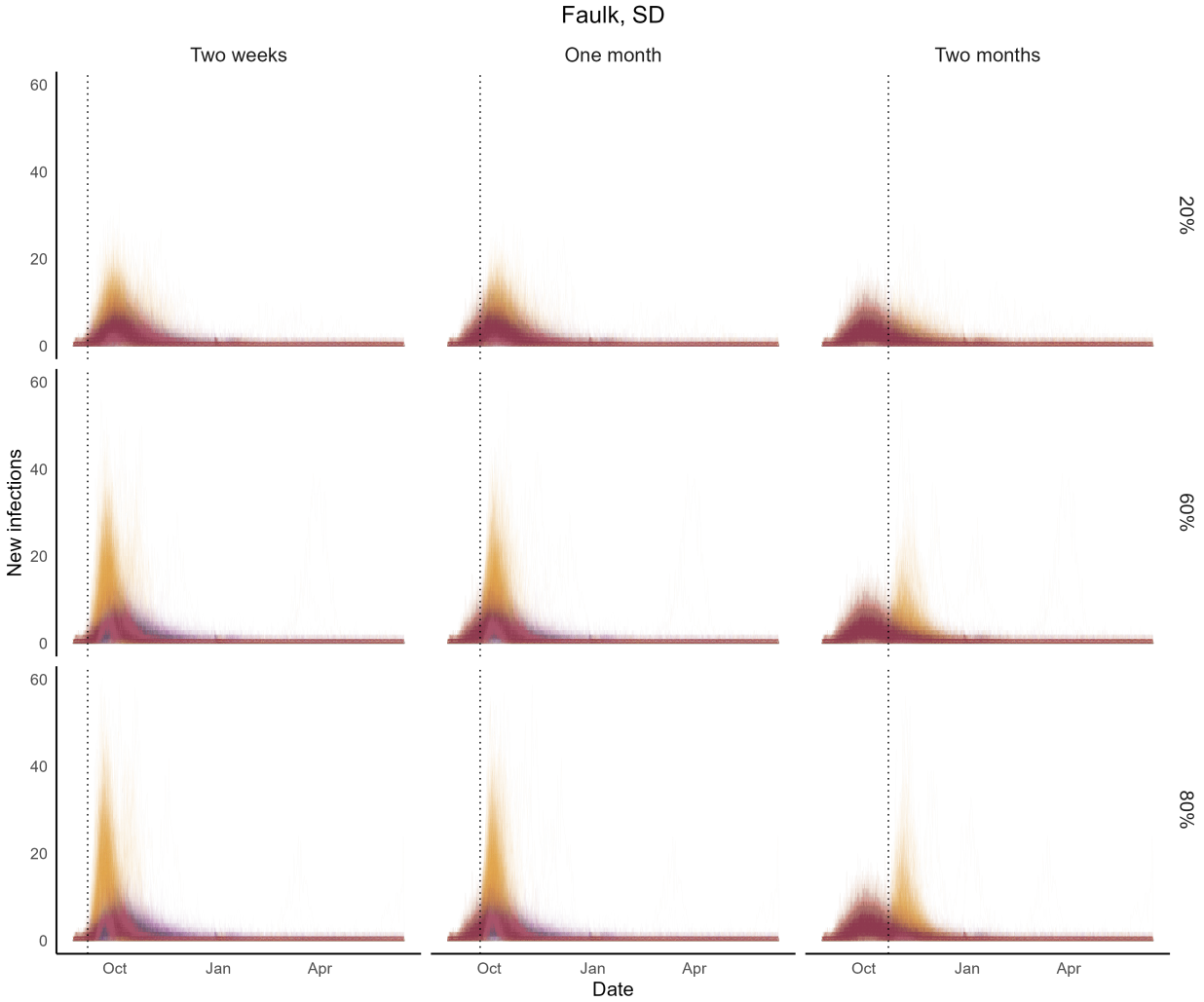

Figure 8: Impact of COVID-19 on Faulk, SD had the FoI increased by 20% (first row), 60% (second row), and 80% (third row)after 2 weeks (first column), one month (second column), and two months (third column). Purple paths are simulated paths with estimated FoI. Yellow sample paths are counterfactual paths.

Figure 9 summarises the overall changes in the total number of cases across all county sizes, for different stringency levels of the non-pharmaceutical interventions. As expected, increasing the force of infection (FoI) resulted in a proportional increase in case numbers. The greatest increase in the total number of cases occurred when the FoI reached its maximum (80% in this case) and interventions were relaxed early (within two weeks). Overall, for these six counties, the maximum median increase in the total number of cases was 50.7% [(43.6%, 57.6%) 25% and 75% quantiles] when the FoI increased by 80% after two weeks. However, we observed a negligible change in the total number of cases when interventions were relaxed after two weeks or one month, despite the increase in FoI. Conversely, relaxing interventions after two months resulted in a notably smaller increase in the total number of cases. The median change in the total cases between relaxing

interventions from one month to two months was 10.8%, 9%, and 11.8% with 20%, 60%, and 80% increases in FoI, respectively.

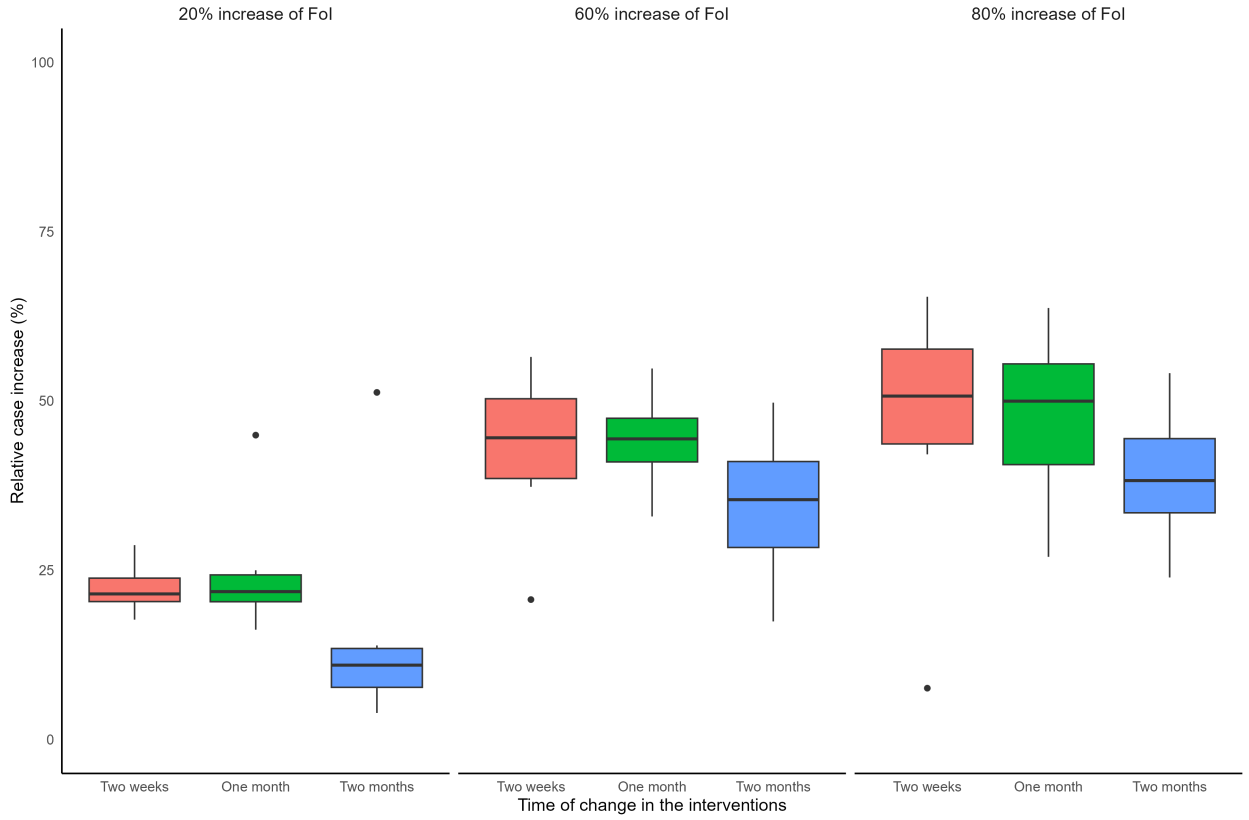

Figure 9: Relative total number of case increase (in percentages)of the counterfactual scenarios with comparison to the observed data. Each panel illustrates the total number of case increase if the NPI stringency changed from high (20% FoI increase), moderate (60% FoI increase), and low (80% FoI increase) after two weeks, one month and two months. Each box plot represents the distribution of relative case increases across all counties under a given scenario.

4 Discussion

In this study, we delved into the dynamics of COVID-19 transmission in rural settings, where population characteristics and occupational patterns stand in stark contrast to the urban environments typically studied in COVID-19 modelling research.

In particular, we studied the transmission dynamics of SARS-CoV-2 within six rural counties of North and South Dakota in the United States. These counties, characterised by small populations of 2500 and low population density, primarily rely on outdoor occupations such as farming, fishing, and forestry (DataUSA, 2022). We estimated that for this group of counties, the proportion of individuals who immediately got tested, notified, and isolated on average ranged from 0.105 to 0.998. We estimated that the reproduction number

ranged from 1.1 to 1.7 (quartiles) for all these counties and Buffalo, SD had a slightly high reproduction number [median 1.415 and (1.306, 1.527) quartile].

We conducted a counterfactual analysis to explore the impact of increasing the force of infection by 20%, 60%, and 80% after 2 weeks, 1 month and 2 months. Among the various FoI scenarios considered, we found that a surge of 80% (indicating minimal public health measures) two weeks after the implementation of interventions resulted, as expected, in the maximum relative increase in reported cases. On the other hand, we calculated that if the force of infection had increased by 20% after two months of implementing public health measures, the relative increase in the total cases (that get notified) would not have been substantial.

We used a Bayesian hierarchical analysis to estimate some of the parameters in this study, an essential aspect for improved parameter estimation. Specifically, we assumed that the transmission rate (β) and the proportion of symptomatic individuals who got tested, notified, and isolated immediately (τ) followed a hierarchical model and all the other parameters were county-specific. However, we note that this analysis could have been improved if we had conducted a hierarchical analysis 1) that allowed for some parameters to be sampled from a common distribution, 2) some parameters to be county-specific, and 3) other parameters to be the same across all the counties.

We only studied three counties each from South and North Dakota. It may be of interest to study more counties from each state and then conduct a hierarchical analysis with four levels where the additional level is taken for information sharing within the states. Furthermore, it is interesting to see if similar transmission dynamics are observed in other rural counties in other states.

Another possible future direction for this study would be to analyse the transmission dynamics when the counties are closer to each other, that is, the counties interact with each other through movements so, the epidemics do not evolve in isolation. In such a setting a hierarchical analysis with metapopulation model structures may be of interest.

Additionally, it will be advantageous to extend this framework to study the dynamics of the *Omicron* variant where re-infection is possible, the majority of the population is vaccinated, and the non-pharmaceutical intervention measures are no longer applied.

5 Acknowledgements

Punya Alahakoon was supported by a PhD Research Scholarship from the University of Melbourne. P.G. Taylor would like to acknowledge the support of the Australian Research Council via the Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS).

Unless otherwise mentioned, computations were carried out in MATLAB or R across 32 clusters (32 virtual computers). All the computations were carried out by the use of the Nectar Research Cloud (project Infectious

Diseases), a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy (NCRIS). All the plots were generated with ggplot2 Wickham (2016) in R. The codes and necessary data are publicly available (see Supplementary Material).

References

Alahakoon, P., McCaw, J. M., & Taylor, P. G. (2022). Estimation of the probability of epidemic fade-out from multiple outbreak data. *Epidemics*, 100539.

Banholzer, N., Van Weenen, E., Lison, A., Cenedese, A., Seeliger, A., Kratzwald, B., ... others (2021). Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLoS one*, *16*(6), e0252827. (Publisher: Public Library of Science San Francisco, CA USA)

Bo, Y., Guo, C., Lin, C., Zeng, Y., Li, H. B., Zhang, Y., ... others (2021). Effectiveness of non-pharmaceutical interventions on covid-19 transmission in 190 countries from 23 january to 13 april 2020. *International Journal of Infectious Diseases*, *102*, 247–253.

DataUSA. (2022). Retrieved from `https://datausa.io/` (Publication Title: Data USA)

Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., ... others (2020). *Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand* (Vol. 16). Imperial College London London.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Guidotti, E., & Ardia, D. (2020). COVID-19 Data Hub. *Journal of Open Source Software*, *5*(51), 2376. doi: 10.21105/joss.02376

Kettl, D. F. (2020). States Divided: The Implications of American Federalism for COVID-19. *Public Administration Review*, *80*(4), 595–602. Retrieved 2024-10-02, from `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7280573/` doi: 10.1111/puar.13243

Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H., & Lipsitch, M. (2020). Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. *Science*, *368*(6493), 860–868.

Kucharski, A. J., Klepac, P., Conlan, A. J., Kissler, S. M., Tang, M. L., Fry, H., ... others (2020). Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of sars-cov-2 in different settings: a mathematical modelling study. *The Lancet infectious diseases*, *20*(10), 1151–1160.

Leatherby, L. (2020, October). *The worst virus outbreaks in the U.S. are now in rural areas (published 2020)*. Retrieved from `https://nyti.ms/3dLYJIv` (Publication Title: Why The Worst Coronavirus Outbreaks Are Now in the Rural U.S. - The New York Times)

Liu, X., Xu, X., Li, G., Xu, X., Sun, Y., Wang, F., ... Zhang, L. (2021). Differential impact of non-pharmaceutical public health interventions on COVID-19 epidemics in the United States. *BMC Public Health*, *21*(1), 1–7. (Publisher: Springer)

Lurie, N., & Sharfstein, J. M. (2023). State-to-state differences in US COVID-19 outcomes: searching for explanations. *Lancet (London, England)*, *401*(10385), 1314–1315. Retrieved 2024-10-02, from `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10119555/` doi: 10.1016/S0140-6736(23)00726-2

McGlothlin, A. E., & Viele, K. (2018). Bayesian hierarchical models. *Jama*, *320*(22), 2365–2366. (Publisher: American Medical Association)

Mueller, J. T., McConnell, K., Burow, P. B., Pofahl, K., Merdjanoff, A. A., & Farrell, J. (2021). Impacts of the COVID-19 pandemic on rural America. *Proceedings of the National Academy of Sciences*, *118*(1), 2019378118. (Publisher: National Acad Sciences)

Paul, P., France, A. M., Aoki, Y., Batra, D., Biggerstaff, M., Dugan, V., ... others (2021). Genomic surveillance for SARS-CoV-2 variants circulating in the United States, December 2020–May 2021. *Morbidity and Mortality Weekly Report*, *70*(23), 846. (Publisher: Centers for Disease Control and Prevention)

Reiner, R. C., Barber, R. M., Collins, J. K., Zheng, P., Adolph, C., Albright, J., ... IHME COVID-19 Forecasting Team (2021, January). Modeling COVID-19 scenarios for the United States. *Nature Medicine*, *27*(1), 94–105. Retrieved 2022-12-20, from `https://www.nature.com/articles/s41591-020-1132-9` (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/s41591-020-1132-9

Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Elsevier.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from `https://ggplot2.tidyverse.org`