

Assortative mixing among vaccination groups and biased estimation of reproduction numbers

Colin Klaus^a PhD, Matthew Wascher^b PhD, Wasiur R. KhudaBukhsh^c PhD, Joseph H. Tien^d PhD, Grzegorz A. Rempala^e PhD, Eben Kenah^{f*} PhD

a. College of Public Health Division of Biostatistics and The Mathematical Biosciences Institute. The Ohio State University. Columbus, OH. USA.

b. College of Arts and Sciences: Mathematics. The University of Dayton. Dayton, OH. USA.

c. School of Mathematical Sciences. University of Nottingham. Nottingham, UK.

d. Department of Mathematics. The Ohio State University. Columbus, OH. USA.

e. College of Public Health Division of Biostatistics. The Ohio State University. Columbus, OH. USA.

f*. Corresponding Author. College of Public Health Division of Biostatistics. The Ohio State University. Cunz Hall. 1841 Neil Ave. Columbus, OH. 43210.

*telephone: 614-292-4409

*email: kenah.1@osu.edu.

Assortative mixing, where there is more mixing within infection risk groups than would be expected to occur at random, has long been known to affect epidemic dynamics. A classic example comes from sexually transmitted diseases, where assortative mixing within groups that have different levels of sexual activity was shown to increase the initial growth rate of the infection and the basic reproduction number (R_0) compared to the same population with more random choices of sexual partners.¹ Assortative mixing within age groups has also been shown to affect dynamics and statistical inference for diseases spread through respiratory droplets,² which motivates the widespread use of contact matrices for age groups in epidemic models. More recent studies have shown that assortative mixing with respect to vaccination status can affect outbreak sizes and estimates of vaccine efficacy in network-based epidemic models.^{3,4}

We hypothesized that assortative mixing among vaccination groups (vaccinated and unvaccinated) might be a source of bias in population-level estimates of the effective R_0 for the Delta variant of SARS-CoV-2. For example, and for a *fixed* amount of observed transmission, over-prescribed mixing between these groups may exaggerate R_0 due to counteraction by an increased proportion of contacts benefitting from at least one vaccination. The prevalence of vaccination varies greatly across rural and urban areas, and other social groupings within which assortative mixing is likely. According to Ohio Department of Health (ODH) data,⁵ the prevalence of vaccination among adults in Ohio counties ranges from slightly under 20% to slightly under 70%, with an overall prevalence of approximately 55%. To explore the potential impact of assortative mixing on estimation of R_0 , we modified an age-stratified Susceptible-Exposed-Infected-Removed (SEIR) model of SARS-CoV-2 transmission in the state of Ohio to allow for assortative mixing within vaccination groups. This model was parameterized and fit using data from the ODH,⁵ the Centers for Disease Control and Prevention (CDC),⁶ and the United States Census Bureau.⁷ The contact matrix for age groups and some other parameters were taken from Prem et al. and Bubar et al.^{8,9} The model R_0 is the spectral radius of the next-generation matrix.⁹

To make the rate of between-group contact ρ (≤ 1) times the rate of within-group contact, we multiply each within-group contact rate β_{ii} by a and each between-group contact rate β_{ij} by ρa . The factor a ensures that the overall rate of contact is not changed, and it is found by solving the following equation:

$$a \binom{n_0}{2} + a \binom{n_1}{2} + a \rho n_0 n_1 = \binom{n_0 + n_1}{2} \quad \#(1)$$

where n_0 is the number of unvaccinated individuals and n_1 is the number of vaccinated individuals. For n sufficiently large so that $\binom{n}{2} \approx \frac{n^2}{2}$, we get

$$a = \frac{(n_0 + n_1)^2}{n_0^2 + 2\rho n_0 n_1 + n_1^2} \cdot \#(2)$$

As intended, this gives us $a = 1$ when $\rho = 1$.

For several rates of assortative mixing, ρ , we fit ODH daily reported incidence cases using a Bayesian inference approach in which posterior distributions were sampled using a hybrid Markov chain Monte Carlo scheme (see online appendix). Figure 1 shows a histogram of the posterior distribution of R_0 and the fit to daily ODH incidence data. Although the estimates of R_0 differ considerably, there is almost no difference in the fit of the model to daily incident cases reported to ODH.

Despite the potential importance of assortative mixing among vaccination groups in understanding SARS-CoV-2 transmission, there is almost no quantitative empirical research available on this topic. A

Google search for phrases such as “covid19 + assortative mixing + vaccination” (at 18:22 BST on October 25, 2021) returned about 87 thousand results, of which the most relevant refer to age-assortative mixing and its potential impact on vaccination strategies. A search for the same terms on Google Scholar (at 18:25 BST on October 25, 2021) returned more than four hundred hits, with the most relevant emphasizing the interplay between age-assortative mixing and vaccination

Although the epidemic modeling community routinely incorporates age-structured mixing matrices, assortative mixing among groups defined by other risk factors for infection are potential sources of bias in estimating epidemic parameters and the impact of interventions, including R_0 . Vaccination is one of the most important determinants of the risk of infection with SARS-CoV-2 in regions where a vaccine is widely available. An overestimate of R_0 could lead to undue pessimism about our ability to control the COVID-19 pandemic through vaccination and social distancing.

The POLYMOD study shows how social survey methods could be used to better understand mixing patterns in an epidemic.¹⁰ Our simple experiment shows that such surveys could address an important gap in our ability to analyze the population-level transmission of disease - and, by extension, to design and evaluate public health interventions in future epidemics.

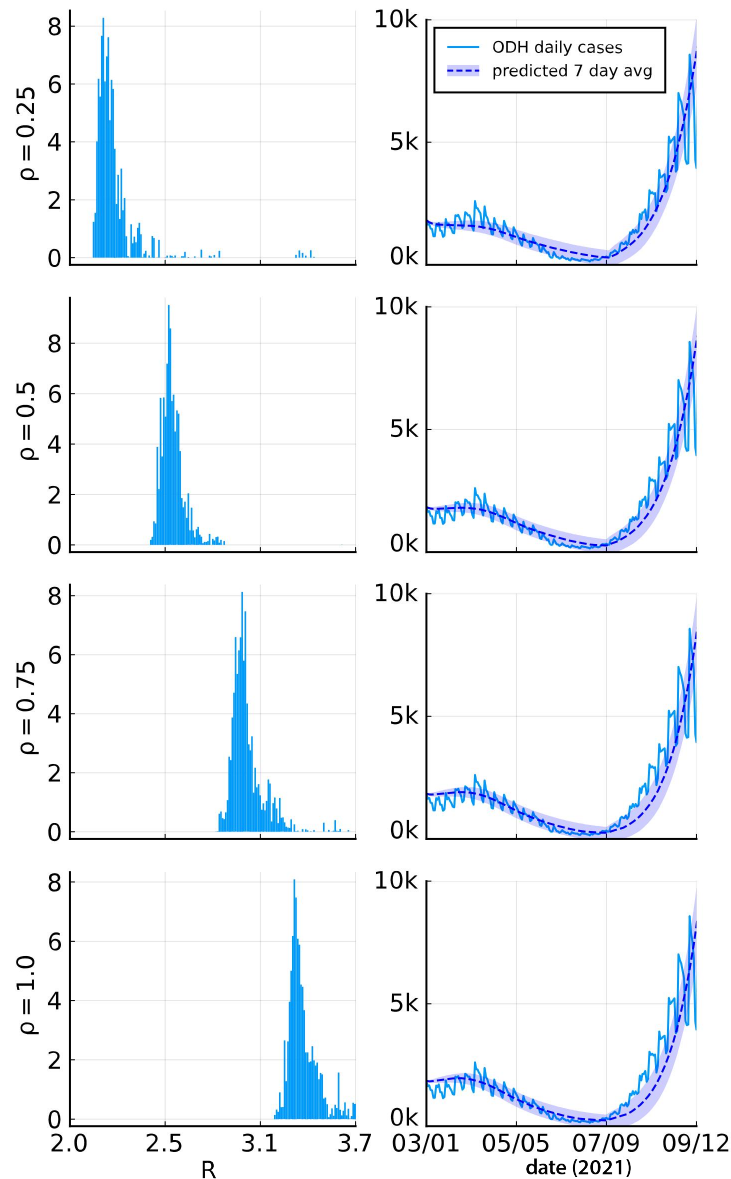


Figure 1: Left: Posterior distributions of R_0 under different degrees of assortative mixing within vaccination groups. Right: Posterior means and pointwise 90% credible intervals for the predicted 7-day moving average of daily COVID-19 cases plotted over daily ODH data.

Conflict of Interest Statement: The authors declared no conflicts of interest.

References

- 1 Gupta S, Anderson RM, May RM. Networks of sexual contacts: implications for the pattern of spread of hiv. *AIDS* 1989; **3**(12): 807 – 17.
- 2 Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol.* 2006; **64**(10): 936 – 44.

- 3 Barclay VC, Smieszek T, He Jianping et al. Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity. *PLoS ONE* 2014; **9**(2): 1 – 11.
- 4 Zivich PN, Volfovsky A, Moody J, Aiello AE. Assortativity and bias in epidemiologic studies of contagious outcomes: a simulated example in the context of vaccination. *Am J Epidemiol* 2021; **190**(11): 2442 – 52.
- 5 Ohio Department of Health. COVID dashboard [INTERNET]. Ohio; 2021 [updated 2021 Oct 29; cited 2021 Oct 29]. Available from: <https://coronavirus.ohio.gov/wps/portal/gov/covid-19/dashboards/overview>.
- 6 US Centers for Disease Control and Prevention. COVID-19 vaccinations in the United States, county [INTERNET]. US; 2021 [cited 2021 Sept 8]. Available from <https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh>.
- 7 US Census. County populations totals 2010-2019 [INTERNET]. US; 2021 [cited 2021 Sept 13]. Available from <https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html>.
- 8 Prem K, van Zandvoort K, Klepac P et al. Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era. *PLoS Comput Biol* 2021; **17**(7): e1009098.
- 9 Bubar KM, Reinholt K, Kissler SM et al. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* 2021; **371**(6532): 916 – 21.
- 10 Mossong J, Hens N, Jit M et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 2008; **5**(3): e74.