Potential Biases Arising from Epidemic Dynamics in Observational Seroprotection Studies

Rebecca Kahn, Lee Kennedy-Shaffer, Yonatan H. Grad, James M. Robins, and Marc Lipsitch

Correspondence to Dr. Marc Lipsitch, Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115 (email: mlipsitc@hsph.harvard.edu)

Author affiliations: Center for Communicable Disease Dynamics, Harvard TH Chan School of Public Health, Boston, Massachusetts, USA (Rebecca Kahn, Lee Kennedy-Shaffer, Marc Lipsitch); Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, Massachusetts, USA (Rebecca Kahn, James M. Robins, Marc Lipsitch); Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, Massachusetts, USA (Lee Kennedy-Shaffer, James M. Robins); Department of Immunology and Infectious Diseases, Harvard TH Chan School of Public Health, Boston, Massachusetts, USA (Yonatan H. Grad, Marc Lipsitch).

This work was supported in part by Award Number U54GM088558 from the US National Institute Of General Medical Sciences and by Award Number U01IP001121 from the US Centers for Disease Control and Prevention. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

© The Author(s) 2020. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
Abstract

The extent and duration of immunity following SARS-CoV-2 infection are critical outstanding questions about the epidemiology of this novel virus, and studies are needed to evaluate the effects of serostatus on reinfection. Understanding the potential sources of bias and methods to alleviate biases in these studies is important for informing their design and analysis.

Confounding by individual-level risk factors in observational studies like these is relatively well appreciated. Here, we show how geographic structure and the underlying, natural dynamics of epidemics can also induce noncausal associations. We take the approach of simulating serologic studies in the context of an uncontrolled or a controlled epidemic, under different assumptions about whether prior infection does or does not protect an individual against subsequent infection, and using various designs and analytic approaches to analyze the simulated data. We find that in studies assessing whether seropositivity confers protection against future infection, comparing seropositive individuals to seronegative individuals with similar time-dependent patterns of exposure to infection, by stratifying or matching on geographic location and time of enrollment, is essential to prevent bias.

Keywords: bias, epidemic dynamics, SARS-CoV-2, seroprotection
The extent and duration of immunity following SARS-CoV-2 infection are critical outstanding questions about the epidemiology of this novel virus (1). Serologic tests, which detect the presence of antibodies, are becoming more widely available (2). However, the presence of antibodies, or seroconversion, does not guarantee immunity to reinfection, and experimental data with other coronaviruses raise concerns that antibodies could under some circumstances enhance future infections (3). Studies are needed to evaluate the short and long term effects of seropositivity. Understanding the potential sources of bias and methods to alleviate biases in these studies is important for informing their design and analysis.

Serologic studies may be useful for a variety of reasons, including to assess the cumulative incidence of infection within a community, to identify risk factors for transmission, and to determine the extent of clustering of infections within a community (4,5). While these types of studies are often cross-sectional and use seroconversion as the endpoint, we consider here longitudinal studies where seroconversion is the exposure of interest.

These seroprotection studies may be conducted by starting with a cross-sectional serological survey, where the tested individuals are then followed to identify future infections. To obtain a sufficient cohort of seropositive individuals, enrollment may need to occur on multiple days. The follow-up to identify future infections depends on regular monitoring of symptoms and/or PCR testing for the virus. Consistent case definitions across the study, as well as tracking individual enrollment and seroconversion dates, are key to reduce the risk of misclassification. If cases are defined based on symptom onset, the study outcome will be the association between seropositivity and progression to symptoms. If cases are based on virologic testing, the study
outcome will be the association between seropositivity and infection. These endpoints have different public health implications and the choice should depend on the scientific question of interest (6).

A crude analysis of this longitudinal study would compare time from enrollment to infection between those that are seropositive and those that are seronegative at enrollment. However, because seroprotection studies are observational, as the exposure (i.e., seropositivity) is not assigned at random, potential confounders must be controlled for to obtain unbiased estimates. Studies of seropositivity and its effect on future infection are particularly prone to confounding because factors that affect someone’s risk of infection and therefore their serostatus prior to enrollment (the exposure) are likely similar to factors that affect someone’s risk of infection after enrollment (the outcome). For example, individuals in high-risk occupations (e.g., health care workers) are more likely to become seropositive and are more likely to be exposed again once they are seropositive.

Confounding by individual-level risk factors is relatively well appreciated. Less obvious perhaps is that geographic structure (7) or the underlying, natural dynamics of epidemics (8,9) can induce noncausal associations between an exposure and an outcome. For example, even when seropositivity confers no protection against future infection, if the overall size of an epidemic is very different in different communities, individuals in communities with small epidemics will have low prevalence of the exposure (seropositivity) and low incidence of the outcome (infection after enrollment), while individuals in communities with larger epidemics will have higher prevalence of the exposure and higher incidence of the outcome, biasing estimates of the effect
of seroprotection. Bias may also occur if individuals are enrolled at different times during an epidemic. If enrollment occurs during an upward trajectory (such as the early exponential phase of an epidemic), individuals enrolled early in the epidemic will be both less likely to be seropositive (exposure) and also less likely to become infected at a given point in time after enrollment (outcome) than those with a later date of enrollment. Moreover, in an epidemic that is controlled (thus with an up-then-down trajectory of incidence) the representation of seropositive individuals will increase with time, but the rate at which these individuals experience the outcome will increase then decrease, creating potential for confounding in either direction.

In this study we take the approach of simulating such studies in the context of an uncontrolled or a controlled epidemic, under different assumptions about whether prior infection does or does not protect an individual against subsequent infection, and using various designs and analytic approaches to analyze the simulated data. By identifying the direction and comparative magnitude of bias of the estimated degree of protection relative to a known true effect of prior infection (known because we have built it into the simulations), we identify means of designing and analyzing such studies that can render them less likely to show bias due to these confounding factors. This framework of simulating studies in the context of an epidemic has been widely used to understand experimental (10) and observational (8,11) studies of risk factors and prevention interventions for infectious disease.
METHODS

We simulate a stochastic outbreak of a disease in a network of people grouped into communities, with each community’s outbreak seeded by introductions over time (7,12). For each simulation, we generate a network graph, where individuals are grouped into either one community of 10,000 people or 10 communities of 1,000 people each. People are only connected to individuals in their own community, with the probability of such a connection based on an input parameter in the simulation. For “well mixed” communities, every individual is connected to every other individual within their community, while for simulations with “clustered” communities, individuals have a limited number of connections within their community, which creates smaller sub-communities, or “clusters”, by chance. In these latter simulations, individuals may have varying numbers of actual connections but all have the same expected number. The network graph of a “well mixed” community is a complete graph, while that of a “clustered” community is a random graph with uniform edge probability. In simulations with 10 communities, all communities are independent of one another, conditional on the introduction of infection from the outside. At each time step in the model, each susceptible individual has a daily probability of infection from each of their infectious contacts of $1 - e^{-\beta}$, where $\beta$ is the force of infection. Hence $e^{-\beta}$ is the conditional infection-free survival probability over a single day among those at risk at the start of the day. If a subject has $n$ infectious contacts on a given day, the force of infection is $n\beta$ and thus the day’s conditional probability of infection is $1 - e^{-n\beta}$. Since the number of contacts per individual varies by simulation, $\beta$ varies by simulation to keep $R$ fixed (see Web Appendix 1). The outbreak is seeded with stochastic introductions into the communities between days one and fifty based on an external force of infection (different from
\( \beta \), see Web Figure 1), which means in simulations with multiple communities, outbreaks may start at different times in each community, and some communities may avoid infection completely.

The disease natural history follows a Susceptible-Exposed-Infectious-Susceptible’ (SEIS’) model, where under the null hypothesis (i.e., no immunity) those in the S and S’ compartments are equally susceptible, while under the alternative hypothesis, those in S’ are less susceptible (in principle, perhaps completely immune, but in keeping with prior evidence about coronaviruses, we assume partially immune) (13,14). In simulations with partial immunity, we make the simplifying assumption that susceptibility is immediately decreased following the infectious period and remains constant over time. Seroconversion is assumed to be detectable at the end of the infectious period. We simulate scenarios with limited control measures in place (\( R_E = 1.5 \)) and scenarios in which control measures that reduce the force of infection per infected individual (\( \beta \)) are implemented at day 120 of the study period, reducing \( R_E \) from 2 to 0.8. \( \beta \) is set to yield these values of \( R_E \). Table 1 shows the specific numbers corresponding to these parameters of the simulations, and Web Appendix 1 describes the generation of the network and outbreak in more detail.

For each simulation setting (one or ten communities, well mixed or clustered communities, control measures or not, and seroprotective efficacy), we consider three sampling designs: enrolling individuals on a single day without matching (day 100), enrolling individuals on multiple days (days 50, 100, 150) without matching, and enrolling individuals on multiple days with matching of enrolled seropositive and seronegative individuals. Enrollment on multiple
days may occur, for example, if different cross sectional surveys are conducted, and this study enrolls the participants in those surveys. A random sample of individuals are enrolled into the study at these specified time points over the course of the outbreak. We classify individuals as seropositive or seronegative based on their serostatus on day of enrollment into the study, and then we follow them up until they are infected or until the study period ends at day 200. In the unmatched designs, we enroll half of the individuals in each community into the study, with an equal number enrolled on each day of enrollment. In the matched designs, for every seropositive individual enrolled on each day of enrollment, we also enroll one seronegative individual on that day from the same community. This increases the balance between exposure arms but reduces the overall sample size.

For each simulation setting and sampling design, we conduct two analyses. First, we conduct an unstratified analysis in which we calculate the hazard ratio of infection comparing seropositive to seronegative individuals, using a Cox proportional hazards model with time starting from enrollment (i.e., possibly not the same calendar time if individuals enroll on different dates). Second, given the potential for stochasticity to generate heterogeneous outbreaks between communities (7), we also conduct an analysis stratified by community and day of enrollment to prevent confounding by these variables. In this analysis, a Cox proportional hazards model with time starting from enrollment is fit with a separate baseline hazard function for each community and day of enrollment combination, but a common hazard ratio due to seropositivity. R code for the simulations and analysis is available on Github (15), and additional analyses examined are described in Web Appendix 2, Web Figure 2, and Web Figure 3.
RESULTS

Figure 1 shows the results for 1,000 simulations for each of 36 combinations of parameters (see Table 1). Figures 1A–D summarize results from simulations with limited control measures in place (R_E=1.5). Figures 1A and 1C are under the null, meaning seropositivity provides no protection against reinfection (β⁺ = β⁻, where β⁺ is the force of infection for contact between an infectious individual and a seropositive individual and β⁻ is the force of infection for contact between an infectious individual and a seronegative individual). In Figures 1B and 1D, seropositivity reduces susceptibility by 50% (β⁺ = 0.5*β⁻) and 95% (β⁺ = 0.05*β⁻), respectively.

Simulations are in well mixed communities, meaning everyone within a community is connected to each other, except in Figure 1C which has random clustering within each community. This clustering leads to correlations between infection status of particular individuals close together in the network and may be understood as creating multiple smaller (albeit overlapping) “communities” within each discrete community.

For simulations with one well mixed community with the same day of enrollment for all individuals (top lines of Figures 1A, 1B, and 1D), a crude analysis returns unbiased results. If enrollment occurs on different days (Figures 1A, 1B, and 1D, second and third lines), a crude analysis yields an upwardly biased estimate of the hazard ratio, making seropositivity appear harmful. However, matching on day of enrollment or stratifying the analysis by day of enrollment removes this bias.
With multiple communities (and thus multiple, unconnected epidemics, as in the bottom halves of Figures 1A, 1B, and 1D), an unadjusted analysis creates the same upward bias, regardless of whether enrollment is on the same or multiple calendar dates, as the same calendar date does not mean the same phase of the epidemic in each of the communities. Once again, the bias is upward because individuals in communities with larger or more advanced epidemics are exposed to higher hazards and are more likely to be seropositive at baseline (Figures 2A–D). As before, the bias can be removed by a matched design or stratified analysis, this time matching or stratifying on both community and day of enrollment. For analyses with a high number of infectious contacts for any enrolled individuals (e.g., Figures 1B and 1D with different days of enrollment), the estimated hazard ratio is between the ratio $\frac{\beta^+}{\beta^-}$ and the null HR=1. This occurs because an individual’s hazard is not simply the product of their number of contacts and the force of infection. This is not a bias in the conventional sense, but rather a difference between the ratio $\frac{\beta^+}{\beta^-}$ and the parameter that is estimated by the Cox model (see details in Web Appendix 1). For settings with a lower force of infection or fewer infectious contacts, this difference is imperceptible.

Clustering of contacts within communities (a departure from the assumption of a well mixed epidemic, Figure 1C) produces an upward bias even in the matched design and stratified analyses. As noted, this reflects that the different parts of the network have different local prevalence at any given time, resulting in a milder form of the same heterogeneity-induced bias seen when there are many discrete communities. Because these clusters of high and low prevalence areas overlap and arise during the study, there is no a priori way to adjust for them.
In the simulations summarized in Figures 1E and 1F, transmission is reduced partway through the outbreak in one or more well mixed communities, representing intensified control measures ($R_E=2 \Rightarrow 0.8$). In these simulations, there are fewer reinfections, as reflected in the wider interquartile ranges. As before, the single-community estimates are unbiased when all individuals enroll on the same day, but when enrollment occurs on different days or there are multiple communities, the estimates are biased. In the single-community simulations with two different days of enrollment, the unstratified, non-matched analysis estimates are slightly biased away from the null, making seropositivity look protective. This occurs because there are more seropositives at later enrollment dates when the average hazard over the rest of the study is lowest (Figures 2E and 2F).

Hence, with multiple communities or multiple enrollment dates, confounding can go in either direction depending on the dynamics of the epidemic at the times of enrollment. Matching on enrollment alleviates the different biases, as does stratification in cases where there are infections in both the seropositive and seronegative arms. If there are substantially fewer seropositive individuals than seronegative individuals and the risk of infection after enrollment is low (i.e., because of effective control measures), there can be settings with no infections among the seropositive enrollees in some or all strata. In these cases, stratified analyses can lead to unstable results because methods to account for one arm with zero cases (e.g., adding a case to each arm) can over-correct when the zero-case arm has far fewer individuals than the other. Matched designs are thus preferable because they remove this imbalance between the two exposure arms.
We note that in the simulations under the null with limited control measures (Figures 2A and 2C), the daily hazard (proportion in the S compartment moving to the E compartment) initially increases during the early spread of the virus and then begins to plateau. In simulations with controlled epidemics and/or immunity (Figures 2B, 2D–F), the daily hazard increases and then decreases.

DISCUSSION

We find that in studies assessing whether seropositivity confers protection against future infection, comparing seropositive individuals to seronegative individuals with similar time-dependent patterns of exposure to infection is essential, because otherwise confounding can bias results; accounting for differential exposure among seropositive individuals and seronegative individuals is necessary to prevent bias. This bias can arise from either having multiple days of enrollment over the course of the study by design or by having multiple communities where the outbreak stochastically starts at different times. Matching in the design or stratifying in the analysis on community and day of enrollment alleviates this bias in well mixed communities. When there is clustering within communities, a slight upward bias remains, suggesting the local network structure in a study is an important factor to consider.

While most individuals are susceptible when they are enrolled into the study, it is possible for individuals to be exposed or infectious upon enrollment. Excluding individuals who are infected soon after enrollment (e.g., within the average latent period length) would remove many of these
cases. For potentially asymptomatic infections, these cases would not be able to be excluded in a study without viral testing for active infection. Small biases may occur if all individuals enrolled in the study are not susceptible at enrollment.

The results shown here assume perfect specificity of the serologic test. As expected (16), imperfect specificity causes bias towards the null (Web Appendix 3 and Web Figure 4). More complex interactions of immunity and infection, including immunity that wanes over the time scale of the study, viral-load dependent infection, and effects of repeated exposures, such as boosting of titers, may affect these biases as well, or introduce other potential biases. Further research is needed to understand the effects of these biological mechanisms in the specific context of SARS-CoV-2.

These simulations focus on the bias inherent in some study designs that may be considered, but do not address the feasibility of implementing these designs. In addition, we do not focus on the power of these studies; this may have important consequences in determining an adequate sample size. Sample size considerations will be particularly important in balancing the advantage of starting enrollment later, when the cumulative incidence is higher and thus the exposure arms are more likely to be balanced, and avoiding the tail of an outbreak or a setting after control measures have been implemented, which will reduce the infection risk for all participants. We have shown that matching can address these issues, but matching requires exposure status to be known at enrollment. This may be feasible if the study is designed following a serological survey, where individuals can be enrolled on the basis of their antibody presence from the survey. If the exposure needs to be measured for the seroprotection study, however, matching
may require far more serologic testing to be conducted, inflating the cost of the study.

Investigators will need to consider the relative sample size requirements and testing burden of these designs in the context of their specific study.

As serologic studies begin, understanding potential sources of bias and how to alleviate them are important for accurately estimating the extent and duration of immunity to SARS-CoV-2. Here we have focused on the impact of epidemic dynamics on estimation of seroprotection and have assumed all individuals in the model are exchangeable and differ only in whom they contact. Future work could examine additional heterogeneity, such as behaviors or factors that increase risk of infection, which might lead to further biases.

Acknowledgments: We thank Dr. Jan Vandenbroucke for helpful comments on the manuscript.
References

1. Lipsitch M. Opinion | Who Is Immune to the Coronavirus? 2020;(https://www.nytimes.com/2020/04/13/opinion/coronavirus-immunity.html). (Accessed April 13, 2020)

2. Branswell H, Mark Arsenault — Boston Globe, Begley S, et al. CDC begins studies for more precise count of undetected Covid-19 cases. STAT. 2020;(https://www.statnews.com/2020/04/04/cdc-launches-studies-to-get-more-precise-count-of-undetected-covid-19-cases/). (Accessed April 13, 2020)

3. Peeples L. News Feature: Avoiding pitfalls in the pursuit of a COVID-19 vaccine. Proc. Natl. Acad. Sci. U. S. A. [electronic article]. 2020;(https://www.pnas.org/content/early/2020/03/27/2005456117.abstract). (Accessed April 13, 2020)

4. Metcalf CJE, Farrar J, Cutts FT, et al. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. Lancet. 2016;388(10045):728–730.

5. Bryant JE, Azman AS, Ferrari MJ, et al. Serology for SARS-CoV-2: Apprehensions, opportunities, and the path forward. Sci Immunol [electronic article]. 2020;5(47).

6. World Health Organization. Correlates of vaccine-induced protection: methods and implications. Geneva, Switzerland: World Health Organization; 2013. https://www.who.int/immunization/documents/WHO_IVB_13.01/en/. (Accessed May 25, 2020)
7. Kahn R, Hitchings M, Bellan S, et al. Impact of stochastically generated heterogeneity in hazard rates on individually randomized vaccine efficacy trials. *Clin. Trials.* 2018;15(2):207–211.

8. Goldstein E, Pitzer VE, O’Hagan JJ, et al. Temporally Varying Relative Risks for Infectious Diseases: Implications for Infectious Disease Control. *Epidemiology.* 2017;28(1):136–144.

9. Koopman JS, Longini IM Jr. The ecological effects of individual exposures and nonlinear disease dynamics in populations. *Am. J. Public Health.* 1994;84(5):836–842.

10. Halloran ME, Auranen K, Baird S, et al. Simulations for designing and interpreting intervention trials in infectious diseases. *BMC Med.* 2017;15(1):223.

11. Ray GT, Lewis N, Klein NP, et al. Depletion-of-susceptibles Bias in Analyses of Intra-season Waning of Influenza Vaccine Effectiveness. *Clin. Infect. Dis.* 2020;70(7):1484–1486.

12. Hitchings MDT, Lipsitch M, Wang R, et al. Competing Effects of Indirect Protection and Clustering on the Power of Cluster-Randomized Controlled Vaccine Trials. *Am. J. Epidemiol.* 2018;187(8):1763.

13. Reed SE. The behaviour of recent isolates of human respiratory coronavirus in vitro and in volunteers: Evidence of heterogeneity among 229E-related strains. *J. Med. Virol.* 1984;13(2):179–192.

14. Callow KA, Parry HF, Sergeant M, et al. The time course of the immune response to experimental coronavirus infection of man. *Epidemiol. Infect.* 1990;105(2):435–446.
15. Kahn R. Serologic Studies. *GitHub*. (https://github.com/rek160/serologic-studies).
   (Accessed May 18, 2020)

16. Copeland KT, Checkoway H, McMICHAEL AJ, et al. Bias due to misclassification in the estimation of relative risk. *American Journal of Epidemiology*. 1977;105(5):488–495.

17. Meyers LA, Pourbohloul B, Newman MEJ, et al. Network theory and SARS: predicting outbreak diversity. *Journal of Theoretical Biology*. 2005;232(1):71–81.
Table 1. Parameters

| Parameter                                                                 | Values                                                                                                                                 |
|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Number of communities                                                    | 1, 10                                                                                                                                 |
| Average community size                                                   | 1 community simulations: 10,000 10 community simulations: 1,000                                                                     |
| Probability of connection with someone within the same community         | Well mixed: 1 (everyone is connected to everyone in their community) Clustered: 0.002 probability per edge for 1 community and 0.02 probability per edge for 10 communities |
| Probability of connection with someone in another community              | 0                                                                                                                                 |
| $R_E$ (17)                                                               | Controlled: 2.0 → 0.8 Uncontrolled: 1.5                                                                                               |
| Latent period                                                            | 5.6 days (gamma distribution with shape = 5, rate = 0.9)                                                                               |
| Infectious period                                                        | 10 days (gamma distribution with shape = 3, rate = 0.3)                                                                               |
| Days of simulation                                                       | 200                                                                                                                                     |
| Day control begins                                                       | Controlled: 120 Uncontrolled: Never                                                                                                    |
| Reduction in $\beta$ after control                                       | 60%                                                                                                                                     |
| Days of enrollment                                                       | Same day: 100 Different days (uncontrolled): 50, 100, 150 Different days (controlled): 100, 150                                    |
| % of individuals enrolled (unmatched)                                    | 50%                                                                                                                                     |
| Seropositivity protection                                                | 0 (null) 50% 95%                                                                                                                        |

$R_E$ = effective reproductive number
Table 2. Bias Summary

| Cause of bias                                      | Direction of bias | Ways to correct                                                                 |
|--------------------------------------------------|-------------------|---------------------------------------------------------------------------------|
| Multiple communities with different timing of epidemics | Upward            | Matched design or stratified analysis (matching works better when both number of seropositives and risk of infection are low) |
| Different days of enrollment                     | Upward or downward| Matched design or stratified analysis (matching works better when both number of seropositives and risk of infection are low) |
| Clustered communities                             | Upward            | Cannot correct a priori but could consider matching on household or neighborhood |
Figure 1. Hazard Ratios

The median and IQR of estimated hazard ratios, comparing seropositives to seronegatives, for each set of simulation settings: A) well mixed communities, uncontrolled, null seroprotection; B) well mixed, uncontrolled, 50% seroprotection; C) clustered communities, uncontrolled, null seroprotection; D) well mixed, uncontrolled, 95% seroprotection; E) well mixed, controlled, null seroprotection; F) well mixed, controlled, 50% seroprotection. Note the different x-axis scales.

We consider three sampling designs for each simulation setting: enrolling individuals on a single day without matching, enrolling individuals on multiple days without matching, and enrolling individuals on multiple days with matching. In the matched designs, for each seropositive individual enrolled on each enrollment day, a seronegative individual from the same community is also enrolled on that day. We compare analyses stratified by enrollment day and community (black) to unstratified analyses (grey). Simulations with zero events in either the seropositive or seronegative arm were excluded (percent of simulations excluded in each figure: A: 0.85%, B: 1.6%, C: 0.28%, D: 22.1%, E: 4.7%, F: 6.3%). For analyses with a high infection hazard for any enrolled individuals (e.g., Figures 1B, 1D, and 1F with different days of enrollment), the estimated hazard ratio is between the ratio of the force of infection between seropositive and seronegatives ($\beta^+ / \beta^-$) and the null HR=1. This occurs because an individual’s hazard is not simply the product of their number of contacts and the force of infection. This is not a bias in the conventional sense, but rather a difference between the ratio $\beta^+ / \beta^-$ and the parameter that is estimated by the Cox model (see Web Appendix 1 for more details).
Figure 2. Daily Hazards

The average simulated daily hazard of infection for those in the initial susceptible compartment (i.e. never infected) to move to the exposed compartment in the simulations with one community: A) well mixed communities, uncontrolled, null seroprotection; B) well mixed, uncontrolled, 50% seroprotection; C) clustered communities, uncontrolled, null seroprotection; D) well mixed, uncontrolled, 95% seroprotection; E) well mixed, controlled, null seroprotection; F) well mixed, controlled, 50% seroprotection. Note the different y-axis scales. Horizontal bars show lengths of follow-up for each day of enrollment. The height of the bars indicates the average hazard for that duration of follow-up. In A–D, follow-up begins on days 50, 100, and 150, while in E and F, follow-up begins on days 100 and 150 only. Vertical grey lines denote the day control measures are implemented, which reduce the force of infection by 60% (E and F). The number of infectious individuals continues to grow beyond the day of control for approximately the average length of the latent period (5.6 days) due to those infected in the days just before control. This causes the hazard to increase again after its initial drop before declining again.
B) Simulation Setting

1 Community
- Same day of enrollment, no match
  - Unstratified: 0.50 (0.42–0.60)
  - Stratified: 0.50 (0.42–0.60)
- Different day of enrollment, no match
  - Unstratified: 1.23 (1.08–1.39)
  - Stratified: 0.50 (0.44–0.57)
- Different day of enrollment, match
  - Unstratified: 0.51 (0.46–0.57)
  - Stratified: 0.50 (0.45–0.56)

10 Communities
- Same day of enrollment, no match
  - Unstratified: 1.62 (1.20–2.29)
  - Stratified: 0.50 (0.42–0.59)
- Different day of enrollment, no match
  - Unstratified: 3.52 (2.75–4.76)
  - Stratified: 0.53 (0.46–0.59)
- Different day of enrollment, match
  - Unstratified: 0.54 (0.49–0.59)
  - Stratified: 0.53 (0.48–0.58)
C) Simulation Setting

1 Community
   Same day of enrollment, no match
      Unstratified 1.20 (1.05–1.38)
      Stratified  1.20 (1.05–1.38)
   Different day of enrollment, no match
      Unstratified 3.12 (2.89–3.39)
      Stratified  1.15 (1.07–1.24)
   Different day of enrollment, match
      Unstratified 1.15 (1.07–1.24)
      Stratified  1.15 (1.07–1.24)

10 Communities
   Same day of enrollment, no match
      Unstratified 3.72 (2.85–5.03)
      Stratified  1.15 (1.02–1.31)
   Different day of enrollment, no match
      Unstratified 8.59 (6.81–11.31)
      Stratified  1.11 (1.02–1.20)
   Different day of enrollment, match
      Unstratified 1.10 (1.03–1.17)
      Stratified  1.11 (1.03–1.19)
D)

| Simulation Setting                                  | Median (IQR)   |
|-----------------------------------------------------|---------------|
| 1 Community                                         |               |
| Same day of enrollment, no match                    |               |
| Unstratified                                        | 0.06 (0.04–0.09) |
| Stratified                                          | 0.06 (0.04–0.09) |
| Different day of enrollment, no match                |               |
| Unstratified                                        | 0.12 (0.09–0.17) |
| Stratified                                          | 0.05 (0.04–0.07) |
| Different day of enrollment, match                   |               |
| Unstratified                                        | 0.05 (0.04–0.07) |
| Stratified                                          | 0.05 (0.04–0.07) |
| 10 Communities                                      |               |
| Same day of enrollment, no match                    |               |
| Unstratified                                        | 0.18 (0.10–0.29) |
| Stratified                                          | 0.06 (0.04–0.08) |
| Different day of enrollment, no match                |               |
| Unstratified                                        | 0.30 (0.19–0.45) |
| Stratified                                          | 0.06 (0.04–0.08) |
| Different day of enrollment, match                   |               |
| Unstratified                                        | 0.06 (0.04–0.07) |
| Stratified                                          | 0.05 (0.04–0.07) |
E) Simulation Setting

1 Community
Same day of enrollment, no match
Unstratified  0.99 (0.89–1.10)
Stratified  0.99 (0.89–1.10)
Different day of enrollment, no match
Unstratified  0.77 (0.72–0.85)
Stratified  0.99 (0.92–1.06)
Different day of enrollment, match
Unstratified  0.99 (0.93–1.06)
Stratified  0.99 (0.93–1.05)

10 Communities
Same day of enrollment, no match
Unstratified  3.00 (2.31–3.99)
Stratified  0.98 (0.89–1.08)
Different day of enrollment, no match
Unstratified  2.08 (1.62–2.76)
Stratified  0.99 (0.90–1.06)
Different day of enrollment, match
Unstratified  0.99 (0.93–1.05)
Stratified  0.99 (0.93–1.06)
F) Simulation Setting

1 Community
- Same day of enrollment, no match
  - Unstratified
    - Median (IQR): 0.50 (0.42–0.58)
  - Stratified
    - Median (IQR): 0.50 (0.42–0.58)
- Different day of enrollment, no match
  - Unstratified
    - Median (IQR): 0.34 (0.30–0.40)
  - Stratified
    - Median (IQR): 0.51 (0.45–0.56)
- Different day of enrollment, match
  - Unstratified
    - Median (IQR): 0.51 (0.46–0.55)
  - Stratified
    - Median (IQR): 0.50 (0.46–0.55)

10 Communities
- Same day of enrollment, no match
  - Unstratified
    - Median (IQR): 1.50 (1.14–2.08)
  - Stratified
    - Median (IQR): 0.51 (0.44–0.58)
- Different day of enrollment, no match
  - Unstratified
    - Median (IQR): 0.83 (0.64–1.14)
  - Stratified
    - Median (IQR): 0.52 (0.46–0.57)
- Different day of enrollment, match
  - Unstratified
    - Median (IQR): 0.53 (0.49–0.58)
  - Stratified
    - Median (IQR): 0.52 (0.47–0.56)
