Routine Hospital-based SARS-CoV-2 Testing Outperforms State-based Data in Predicting Clinical Burden

Len Covello, MD

Community Hospital, Munster, Indiana

Andrew Gelman, PhD

Departments of Statistics and Political Science, Columbia University, New York, NY

Yajuan Si, PhD

Institute for Social Research, University of Michigan, Ann Arbor, MI

Siquan Wang, MS

Department of Biostatistics, Columbia University, New York, NY

12 Jan 2021

Corresponding author: Yajuan Si,

Email: yajuan@umich.edu, Telephone: 734-7646935

Survey Research Center, Institute for Social Research, University of Michigan,

ISR 4014, 426 Thompson St, Ann Arbor, MI 40104

ABSTRACT

Throughout the COVID-19 pandemic, government policy and healthcare implementation responses have been guided by reported positivity rates and counts of positive cases in the community. The selection bias of these data calls into question their validity as measures of the actual viral incidence in the community and as predictors of clinical burden. In the absence of any successful public or academic campaign for comprehensive or random testing, we have developed a proxy method for synthetic random sampling, based on viral RNA testing of patients who present for elective procedures within a hospital system. We present here an approach under multilevel regression and poststratification (MRP) to collecting and analyzing data on viral exposure among patients in a hospital system and performing publicly available statistical adjustment to estimate true viral incidence and trends in the community. We apply our MRP method to track viral behavior in a mixed urban-suburban-rural setting in Indiana. This method can be easily implemented in a wide variety of hospital settings. Finally, we provide evidence that this model predicts the clinical burden of SARS-CoV-2 earlier and more accurately than currently accepted metrics.

INTRODUCTION

Early knowledge of incidence and trends of viral transmission in communities is crucial, but in the absence of universal screening or random testing, interested parties have been left to extrapolate impressions of community viral behavior from nonrepresentative data. Public health professionals have relied on state-sourced positivity rates and raw numbers of positive tests in any given jurisdiction as proxies for the true viral burden. Unfortunately, these presumed proxies are subject to significant selection bias, as most testing protocols understandably target symptomatic and presumed-exposed populations. Further, tests have been applied with different criteria over time and geography according to test availability, perceived community viral burden, and disparate clinical or political testing norms. The uncontrolled nature of the data raises questions or criticism about their validity as determinants of policy that delimits clinical or economic behavior. Absent broad randomized testing—which has not occurred and does not seem to be forthcoming—we need a means of normalizing currently available data to better track trends in the true underlying incidence, either as a more reliable metric or as a reassurance of the validity of our current ones in predicting clinical behavior of SARS-CoV-2.

In the present article, we apply multilevel regression and poststratification (MRP), a standard adjustment method used in survey research that is particularly effective when sample sizes are small in some demographic or geographic slices of the data (1, 2). MRP has increasingly been shown to be useful in public health surveys, and has even been shown to be successful in highly unrepresentative probability or non-probability samples (3-5). We work with data from the Community Hospital group in Indiana, which serves an urban-suburban-rural mix of patients. COVID testing was already being performed for patients in this hospital system, and it was relatively costless to augment this data collection with the statistical analysis presented here. For

this reason, we believe that this method can be easily implemented in a wide variety of hospital settings.

METHODS

Study Data and Sample

Upon reopening to elective medical and surgical procedures after the early spring COVID-19 outbreak, clinical professionals in our hospital system were sufficiently concerned about asymptomatic viral shedding to test all patients for acute viral infection before performing any such procedure. All elective patients for invasive procedures are presumptively asymptomatic, as any potential surgical patient acknowledging symptoms or presenting a recent history of known viral exposure would have the procedure canceled or deferred. All prospective surgical (and other invasive procedure) patients in our hospital system were subjected to a preoperative evaluation of these issues and excluded if they showed evidence of symptoms or exposure.

This population presented a potentially valuable resource. There is a broad age, racial/ethnic, and economic diversity to this group, and its only overt correlation to disease status is that it is specifically selected for a lack of symptoms and a negative exposure history. Though not ideal, this group is a promising proxy for the general community. SARS-CoV-2 has clearly shown the ability to spread throughout the population via both asymptomatic and symptomatic infection. If we were to assume the as yet unverified but reasonable hypothesis that, for any uniform demographic, the ratio of asymptomatic-to-symptomatic viral infection is constant, then the asymptomatic population in a community would vary in a strict ratio with overall prevalence, and could therefore serve as an excellent proxy for true viral incidence. The trending of asymptomatic infection would be expected to be strictly proportional to clinical infection.

Of crucial importance, our sample group varies from a true random sample in predictable ways. It is selected rigorously for asymptomatic/non-exposed status, and age, racial/ethnic, and geographic demographics are well documented in the hospital electronic health records (EHR). It remains only to normalize our sample to the demographics of the larger community to represent the general population.

Measures

We subjected all patients to polymerase chain reaction (PCR) testing for viral RNA, 4 days before their intended procedure. Samples were submitted to LabCorp for analysis using the Roche cobas system. A 70% clinical sensitivity is presumed for this test, based on near 100% internal agreement with positive controls on in vitro analytics (6) and broadly observed clinical performance of PCR testing throughout the pandemic (7); however, asymptomatic and pre-symptomatic patients may be harder to detect than predicted by these analytic data, as dates of infection as well as symptom status/onset are known to have a large effect on sensitivity (8). These effects would need to be acknowledged and, to the degree possible, accounted for in the model. Specificity is near 100%, with false positives likely generated only by cross-contamination or switched samples. These false positives become important only when underlying prevalence is near zero (9), as was the case for our community this summer.

Statistical Analysis

We are interested in rates of SARS-CoV-2 infection in two populations: 1) Individuals undergoing care within the hospital system as patients, and 2) the community from which the hospital draws as a whole. In addition to adjusting for measurement error associated with PCR testing for SARS-CoV-2 infection, we need to generate standardized estimates that reflect

prevalence in the populations of interest rather than merely our sample of elective surgery patients we are drawing on. We anticipate that this sample of asymptomatic patients is a fairly representative group of the community-at-large, but also expect that poststratification to the target population with matching sociodemographics would help enhance the accuracy of our conclusions.

We use a Bayesian approach to account for unknown sensitivity and specificity and apply MRP to testing records for population representation, here using the following adjustment variables: sex, age (0-17, 18-34, 35-64, 65-74, and 75+), race (white, black, and other), and county (Lake and Porter). MRP has two key steps: 1) fit a multilevel model for the prevalence with the adjustment variables based on the testing data; and 2) poststratify using the population distribution of the adjustment variables, yielding prevalence estimates in the target population.

We poststratify to two different populations: patients in the hospital database (those who have historically and currently obtained care in our regional hospital system) and residents of Lake/Porter County, Indiana. For the hospital, we use the EHR database to represent the population of patients from three hospitals in the Community Health System (Community Hospital, St. Catherine Hospital, and St. Mary Medical Center). For the community, we use the American Community Survey 2014-2018 data from the two counties.

We particularly care about changes in viral incidence over time. Indeed, even if our demographic and geographic adjustment is suspect (given systematic differences between sample and populations), the greatest clinical utility lies in being able to predict how much the clinical burden present today is likely to change in the future. Here, the adjustment may be particularly important, as the mix of patients has changed somewhat during the study interval.

We denote the test result for individual i as y_i, where $y_i = 1$ indicating a positive result and $y_i = 0$ indicating negative. Let $p_i = \Pr(y_i = 1)$ be the probability that this person tests positive. The analytic incidence p_i is a function of the sensitivity, specificity, and the true viral incidence π_i for individual i: $p_i = (1 - \gamma)(1 - \pi_i) + \delta\pi_i$. We fit a logistic regression for π_i with covariates including sex, age, race, county, time, and the two-way interaction between sex and age. Using the model-predicted incidence $\hat{\pi}_i$, we apply the sociodemographic distributions in the hospital system and the community to generate the population level prevalence estimates, as the poststratification step in MRP. More statistical details are included in the Supplement.

We perform all computations in Stan and R; data and code are available at https://github.com/yajuansi-sophie/covid19-mrp.

Assumptions and Conjectures

We began the data collection with a few hypotheses or speculations. First, we expected that the ratio between asymptomatic and symptomatic patients would be relatively constant, for a uniform demographic distribution specific to age, gender, and race/ethnicity. Second, we anticipated that changes in PCR positivity among asymptomatic individuals would precede changes in symptomatic PCR-detected infections by several days, because of the known temporal relationship of viral shedding to the onset of clinical disease (10). Third, these hypotheses would imply that trends in our asymptomatic SARS-COV-2 infections would predict the behavior of the virus within the community as a whole. To this end, we wish to determine whether our model mirrors or predicts hospitalization rates in our area as a proxy for clinical viral burden.

In summary, we anticipated that appropriate modeling of the PCR dataset would allow us to measure changes in acute infection incidence as an early warning metric to grasp the developing trend of the disease, or at least in concert with any changes. Further, we aimed to evaluate the validity of positivity and counts of positive cases as metrics to predict clinical burden.

Our primary goal is to track SARS-CoV-2 infection prevalence over time, and to do so in a reliable way. We are assuming that the ratio of the number of asymptomatic patients to that of the symptomatic patients is fixed. We would then need to normalize each group from the sample demographics to true demographics via MRP. Once that is accomplished, we posit that our metric represents an appropriate approximation of a true random sample. In particular, it either is superior to what is being done now (tracking positivity and overall positive case numbers) or it provides a respectable metric to verify that those currently used data are reasonable approximations of fact. All of our data representations should either support one of those contentions, or provide evidence supporting or refuting critiques of our approach. To that end, we need to follow prevalence changes over time, but verify (through MRP and statistical analysis) that these trends are real changes in the community and not changes in the sample demographics. We then need to compare the MRP normalized trends in our model with currently employed metrics of viral prevalence trends: namely, positivity rates and numbers of positive tests in the community. Finally, we compare these metrics with hospitalization rates to determine how predictive our model and the current metrics may be of the community clinical burden of the virus.

RESULTS

Demographic Stability

We collect the preoperative PCR test time and results of patients in the hospital system, and demographic/geographic information including sex, age, race, and counties. As one of our study interests was to compare the analytic value of our method to established symptomatic testing metrics, we collected the records for both asymptomatic presurgical and symptomatic patients tested within our hospital system, where the asymptomatic patients are assumed as our proxy sample to the target population. Our data include daily records from Apr 28 to Nov 30, 2020, representing 23,412 asymptomatic and 9333 symptomatic patients who received PCR tests. We poststratified the patients with tests to the 35,838 hospital EHR records in 2019 and the 654,890 community residents in Lake and Porter counties. Table 1 summarizes the test results and sociodemographic distributions, as well as the sociodemographics in the hospital system and the community, thus illustrating the discrepancy between the sample and the population.

The observed incidence rates are quite naturally different between the PCR tests: 1.1% for asymptomatic patients and 24% for symptomatic patients. As compared to the hospital system patients, asymptomatic patients with PCR tests tend to be female, middle-aged (35-64) or old (65-74), and white. For this reason, neither the hospital patients nor the asymptomatic patients serve as a precise representation of the community population, in particular with an under-coverage for young, male, and nonwhite residents. These differences are not large (Table 1); nonetheless, they are potential sources of error if not accounted for in our statistical model, and can also interfere with estimates of trends if the demographic breakdown of hospital patients varies over time. Furthermore, the county representation is unbalanced. Some patients are from south Cook County, Illinois, and are grouped into the Lake County as a proxy. Fortunately for our analysis, these contiguous communities have similar socioeconomic and ethnic demographics.

Figure 1 presents the observed PCR test incidence over time for asymptomatic and symptomatic patients. The two groups present different prevalence magnitudes and trends. The prevalence changed over time with low values until September, then we see an increasing trend, with a spike at the end of October, and a decrease in November.

We examine the observed sociodemographic distributions of asymptomatic and symptomatic patients receiving PCR tests over time and find that the asymptomatic patients' profiling is stable, while the sample decomposition of the symptomatic patients changes over time. Details are presented in eFigure 1 of the Supplement. This discrepancy provides supporting evidence for our pre-study hypothesis that we should treat the asymptomatic samples as a substantially better proxy sample of the target hospital or community population than the corresponding symptomatic data. The variation of prevalence could be due to various sample decompositions across time, but variation in thresholds for testing symptomatic people over time and demographic is certainly a likely factor. Overall, our analysis here calls into question relying upon symptomatic data trends—as is currently the norm—in understanding the underlying true viral trends in the community and argues that asymptomatic testing is likely to be a superior proxy.

To correct for discrepancies between the sample demographics and those of the community at large, as necessitated by the above observations, we next apply MRP to model the incidence and poststratify to the hospital and community population for representative prevalence estimates. The outputs are given in Figure 2. For asymptomatic patients, the estimated positive PCR test prevalence is lower than the raw value after a spike between May 19 and May 25 and generally lower than 0.5% through September 28. These findings reflect a low observed clinical burden of COVID-19 in our community after the initial March-April outbreak; see District 1

hospitalization (11) in Figure 3. We observe an increasing trend in October and then decreasing throughout November, with MRP adjusted rates inflated significantly.

Prediction metrics of clinical burden

Figure 3 compares the MRP estimates of asymptomatic patients with the publicly released prevalence data in District 1, where our hospital system is located, and the state hospital bed occupancy rate. The district-level prevalence is higher than that of MRP estimates, as expected, since the state data largely involves the testing of symptomatic patients. The two prevalence rates both have an increasing trend since September, but deviate in November, where MRP estimates start decreasing. The occupancy rate of hospital beds in the data generally follows a similar trend but presents a lower increasing rate in November than the prevalence in District 1. We observe that MRP asymptomatic data are able to predict clinical behavior a week or two earlier than the District 1 prevalence data.

To test this conjecture, we focus on the September through November interval as that timeframe encompasses all of the observed growth in viral burden after very low levels throughout late spring and the summer. We are interested in evaluating whether MRP-adjusted asymptomatic patients could track SARS-CoV-2 related hospitalization rates—as measured by counts of hospitalizations and emergency department (ED) visits—better than the currently applied metrics within our counties: positivity rate and counts of positive cases. As noted, our expectation was that hospitalization census would lag viral incidence by a week or more. Further, we anticipated that ED COVID-19 visits would track actual viral incidence, perhaps with a few days' lag. These inferences follow from known lag times from exposure to symptoms to serious illness (10).

Our side-by-side analysis is illustrated in Figure 4. Each plot shows the week-to-week trend of the available metrics (MRP asymptomatic, positivity rate, and the number positive cases) and those of hospitalization rates (the number of hospitalizations and ED visits) within Lake and Porter Counties. All three metrics parallel hospitalization through September up until mid-October, after which the growth in positivity and counts of positive cases far outstrip the growth in hospitalization while the MRP data remains in strict parallel throughout. Further, the MRP estimates track even better with the ED visits. The hospitalization census data, on the other hand, show a 1-week lag. Indeed, we begin to see some decrease in the MRP adjusted asymptomatic positives in November that parallels a decrease in hospitalization while the generally accepted state metrics continue to increase and even accelerate. These data suggest that ongoing increases in District 1 positive testing metrics may simply be artifacts of the test selection process, rather than actual growth in the viral spread. Overall, a comparison of trends shows that symptomatic positive cases only begin to decrease at the time hospital census does, and fully a week after ED visits do. Positivity rate does not identify the apparent decrease in clinical burden and has in fact accelerated through that decrease.

DISCUSSION

Our analysis indicates that applying our MRP normalization to data on the prevalence of asymptomatic SARS-CoV-2 infection produces a valuable leading indicator of hospital and community risk. When we set out to create a model for tracking viral incidence, we recognized substantial shortcomings in the available testing and its interpretation. While state data has become much richer and testing protocols more uniform since we started applying our model, selection bias is still a substantial concern. To that end, our goal for this study was to develop an easily implemented testing strategy—available to any hospital system—that, after demographic

and geographic adjustment, could reasonably approximate a representative sample. In so doing, we hoped to assess the reliability of currently accepted metrics in their prediction of virus trends and, if possible, to improve our ability to anticipate those trends.

The asymptomatic preoperative patients we identify with our protocol are a favorable group to build upon. All sizable hospital systems have a ready-made group of such patients who can produce a large number of data points quite rapidly. As patients continue to seek medical procedures, the population continues to naturally expand over time, and lends itself to trending data. In our nearly 900-bed hospital system, we have thus far generated over 20,000 data points over 31 weeks, representing a community of approximately 700,000 residents. The weekly number of data points has been fairly stable over time, and that observation is likely similar to many hospital systems. We have demonstrated that this sample population is fairly representative of the community demographics as a whole and that there has been minimal sample decomposition over time. That this population stability is not matched by similar demographic stability in the symptomatic population and that we are able to employ MRP to account for any demographic skew and instability in our own protocol both strongly argue that our model is far more representative of random sampling than the currently employed positive case and positivity data. We argue that hospital-based asymptomatic testing with MRP is a more reliably random metric than any currently available and is easily generated from the routine testing of patients prior to their scheduled procedures.

Having established a reasonable statistical validity for our model, we wished to use it to measure the reliability of current state-based metrics. In a general sense, our analysis finds that in our community, all of the metrics trend similarly during viral surges. Broadly speaking, we would support the current view that numbers of positive cases and positivity both remain relatively

stable during periods that our pseudorandom proxy method predicts to be stable and have increased during periods that our proxy predicts show true viral increases. Since the beginning of our study in early May, there have been significant changes in test availability and certainly anecdotal evidence that the indications for testing have changed quite a bit as well. Consequently, the number of tests and clinical indications for testing have almost certainly both increased considerably over that interval, but the patterns cited above have remained stable. For that reason, we feel that there are good reasons to believe that the validity of positive case counts and positivity as metrics for viral spread is, in the event, relatively insensitive to test numbers, test availability, and clinical thresholds for testing, at least in our community.

Finally, we wanted to test each of these metrics as predictors of clinical burden. During the entire study period, we have used our model to predict clinical needs: staffing, bed and ventilator availability, personal protection equipment supplies, and so forth. Our general observation was that this proxy provided us some useful lead time to prepare for the virus. When we compare our model's behavior to that of the standard metrics, we find it to be generally a better predictor of clinical burden. The effect is best seen in our November data. During the week of November 3-10, we were able to predict that viral transmission was decreasing and that our hospitalization was likely to be at or near its peak. Comparison of our model with ED COVID presentations in our area demonstrates quite a precise correlation. It is clear that new acute presentations correlate quite precisely with our metric, and that these changes occur about a week before positive cases and hospitalization census data change. Further, we see positivity rates continue to rise in our area well past the time that our metric and numbers of positive cases have declined. Given that ED visits and hospitalization census rates have also declined in that interval, we find that our model and the number of positive cases appear to be much better and more current

predictors of the true viral clinical burden than positivity rates are. Our best evidence is that our model of asymptomatic testing provides about a one-week lead time on predicting hospitalization than numbers of positive cases, currently the best available broadly-used metric and that both substantially improve upon the utility of positivity rates.

In dealing with a case surge, the extra week of preparedness has been useful and nontrivial. The great benefit may also accrue from recognizing decreased transmission earlier, as we feel the model is able to do, as it may allow for opening up of needed clinical services and socioeconomic commerce in a community earlier than might otherwise be contemplated. In this sense, adherence to positivity rates may be particularly damaging.

We believe our model to be easily generalizable to many hospital systems. As discussed, the sample population and testing regime are readily available, likely to be reasonably representative and stable demographically over time, and easily normalized to true community demographics using the MRP code that we have made available (https://github.com/yajuansi-sophie/covid19-mrp). We feel that this approach represents a simple proxy for random sampling for any community that chooses to employ it. Further benefits might be gained by combining information from different hospital systems. The best way forward might be for individual hospitals and medical groups to gather and analyze their data as we propose in this article, with all the (de-identified) data shared in a common public repository, so that it would be possible for researchers to learn more by analyzing trends as they develop in the pooled dataset. This could be similar to other national data pooling efforts such as in the United States and Israel (12-13).

In addition, though, we demonstrate the clinical utility of less rigorous approaches as well. Should a system choose to track its patients according to our testing protocol, but not incorporate the MRP adjustments, the relative stability of the population demographics suggests that the

trends remain quite valid. Our internal data have shown potentially strong effects of age and racial/ethnic status on our metric (data not shown) so that one would need to ensure at least reasonable stability of those particular traits to trust observed raw trends without formal MRP adjustment. We also find that while results depend strongly on the sensitivity of the test being employed, the trends in the results do not (details in the Supplement).

This finding is encouraging for longer-term monitoring. Very inexpensive antigen testing is now becoming broadly available. These tests may be less sensitive or more time-specific than the PCR-based RNA testing we have been using and therefore less able to verify the true magnitude of viral spread. Nonetheless, our data show that they will likely function perfectly well to follow viral transmission and clinical burden trends, especially if normalized by MRP. Practically speaking, these trends are the prime concern of most healthcare entities.

ACKNOWLEDGEMENTS

References

1. Gelman, Andrew, and Thomas C. Little. Poststratification into Many Categories Using Hierarchical Logistic Regression. Survey Methodology. 1997;23:127–135.

2. Si, Yajuan, Rob Trangucci, Jonah Sol Gabry, and Andrew Gelman. Bayesian Hierarchical Weighting Adjustment and Survey Inference. Survey Methodology. 2020;46 (2):181-214.

3. Downes M, Gurrin LC, English DR, Pirkis J, Currier D, Spittal MJ, Carlin JB. Multilevel Regression and Poststratification: A Modeling Approach to Estimating Population Quantities From Highly Selected Survey Samples. Am J Epidemiol. 2018 Aug 1;187(8):1780-1790.

4. Zhang X, Holt JB, Lu H, Wheaton AG, Ford ES, Greenlund KJ, Croft JB. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. Am J Epidemiol. 2014 Apr 15;179(8):1025-33.

5. Wang, W., Rothschild, D., Goel, S., and Gelman, A. Forecasting Elections with Non-Representative Polls. International Journal of Forecasting. 2015;31:980-991.

6. Roche cobas 6800/8800. Test performance in individual samples. https://diagnostics.roche.com/global/en/products/params/cobas-sars-cov-2-test.html#productSpecs. Accessed November 1, 2020.

7. Woloshin, S, Patel, N, Kesselheim, AS. False Negative Tests for SARS-CoV-2 Infection—Challenges and Implications. N Engl J Med. 2020;383:e38. DOI: 10.1056/NEJMp2015897.

8. Kucirka, LM, Lauer, SA, Laeyendecker, O, Boon, D, Lessler, J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. Ann Int Med.2020;173(4):262-267.

9. Gelman, Andrew, and Bob Carpenter. Bayesian Analysis of Tests with Unknown Specificity and Sensitivity. Journal of the Royal Statistical Society Series C (Applied Statistics).2020; 69(5):1269-1283.

10. Lauer, SA, Grantz, KH, Bi, Q, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) from Publicly Reported Confirmed Cases: Estimation and Application. Ann Int Med. 2020;172(9):577-582.

11. Indiana State Department of Health. COVID-19 Region-Wide Test, Case, and Death Trends. https://hub.mph.in.gov/dataset/covid-19-region-wide-test-case-and-death-trends. Accessed December 21, 2020.

12. Rosenfeld, R., Tibshirani, R., Brooks, L., et al. COVIDcast. https://covidcast.cmu.edu/index.html. Accessed December 21, 2020.

13. Rossman, H., Keshet, A., Shilo, S. et al. A framework for identifying regional outbreak and spread of COVID-19 from one-minute population-wide surveys. Nat Med. 2020;26; 634–638. https://doi.org/10.1038/s41591-020-0857-9.

Table 1. Descriptive summary of test results and sociodemographic distributions.

	Asymptomatic PCR	Symptomatic PCR	Hospital	Community
Size	23412	9333	35838	654890
Incidence(%)	1.1	24	NA	NA
Female(%)	59	60	57	51
Male(%)	41	40	43	49
Age0-17(%)	2.9	16	8.7	24
Age18-34(%)	10	19	12	21
Age35-64(%)	46	44	30	40
Age65-74(%)	24	12	20	9
Age75+(%)	18	8.7	29	6.6
White(%)	72	75	65	69
Black(%)	14	10	19	19
Other(%)	14	15	16	12
Lake(%)	84	83	88	74
Porter(%)	16	17	12	26

Figure 1. Observed weekly PCR test incidence for asymptomatic and symptomatic patients in the Community Hospital system. Note the different scales on the two graphs.
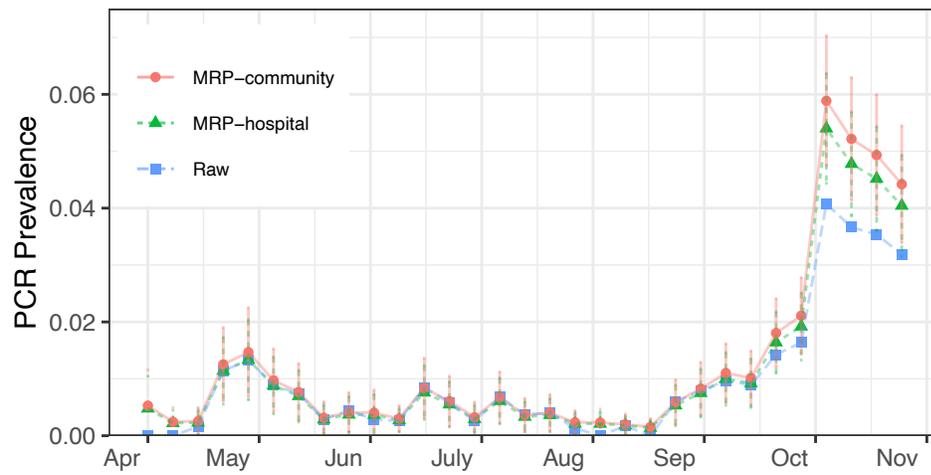
Figure 2: Estimated prevalence of the hospital system and community based on asymptomatic patients. The error bars represent one standard deviation of uncertainty.
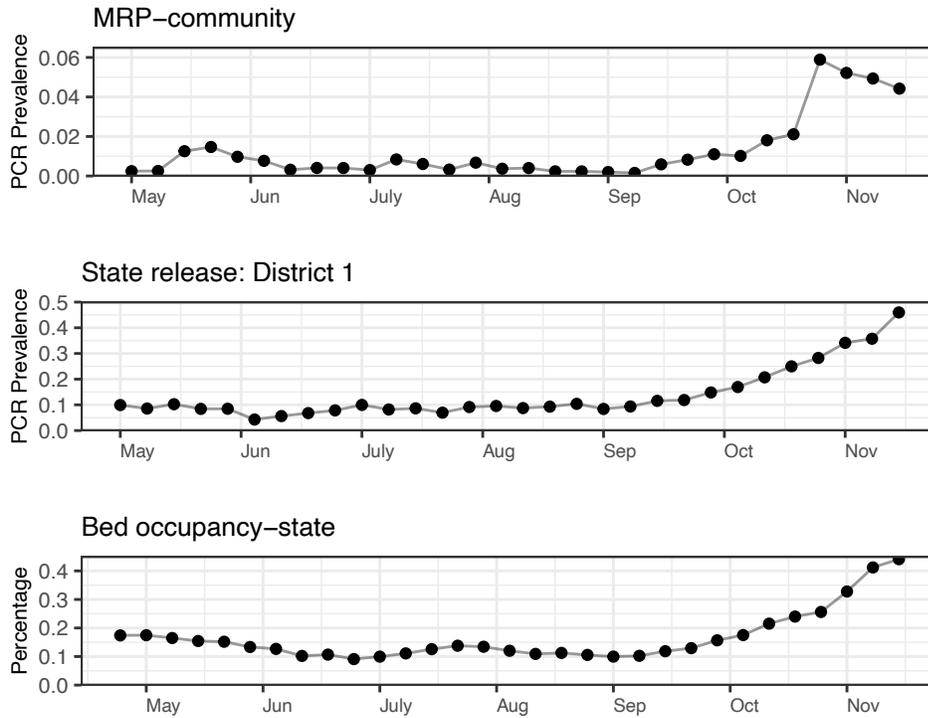
Figure 3: Comparison of MRP estimates with reported prevalence in District 1 and hospital bed occupancy rates in the state, Indiana.[10] Note the different scales on the three graphs.
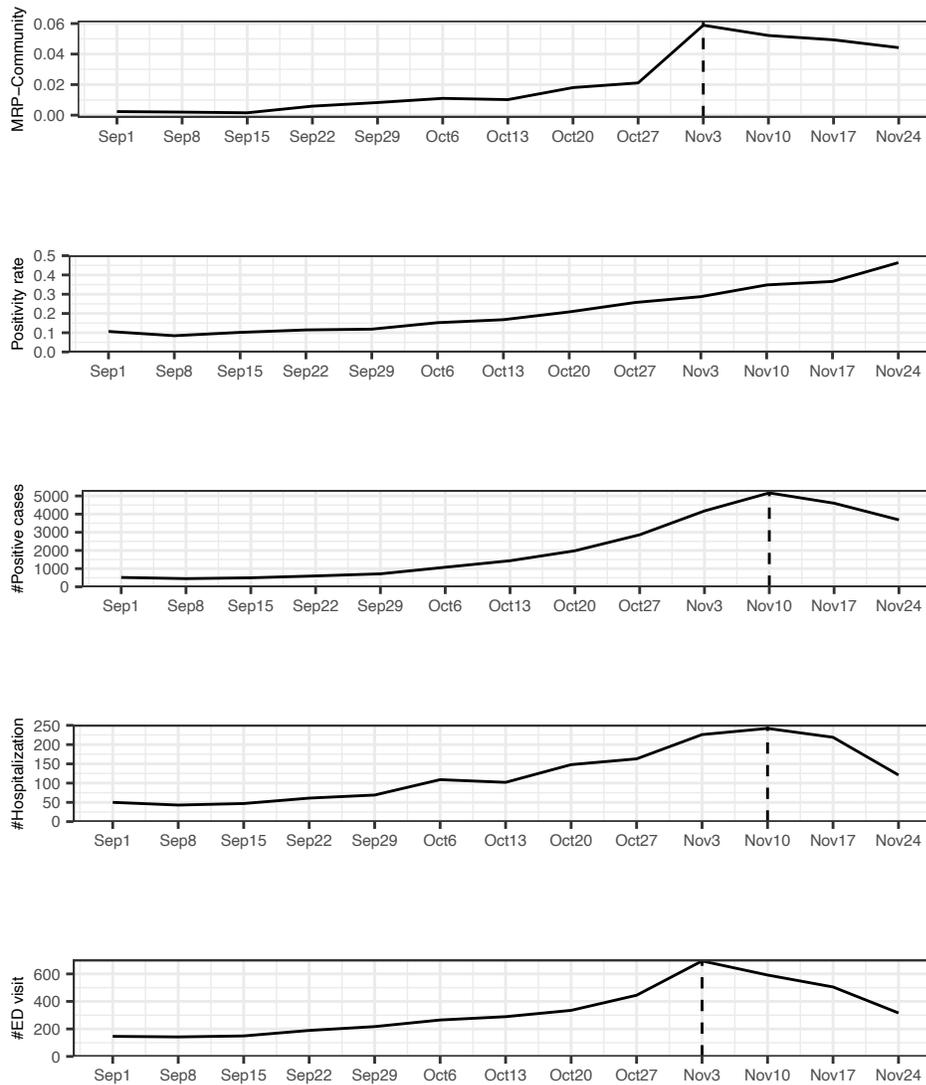
Figure 4: Comparison of MRP estimates with the reported hospitalization counts, ED visits, positivity rate, and the number of positive cases in Lake and Porter counties. The vertical dashed lines indicate the peak values. Note the different scales on the five graphs.

SUPPLEMENT

Here we present the supplemental materials on the modeling details and figures for demographic stability and posterior predictive check, for the paper entitled "Routine Hospital-based SARS-CoV-2 Testing Outperforms State-based Data in Predicting Clinical Burden" by Covello, Gelman, Si and Wang.

Modeling details

We use the following logistic regression in (1) to allow time variation of prevalence over time in the multilevel model parameters.

$$\text{logit}(\pi_i) = \beta_1 + \beta_2 \text{male}_i + \alpha^{\text{age}}_{\text{age}[i]} + \alpha^{\text{race}}_{\text{race}[i]} + \alpha^{\text{county}}_{\text{county}[i]} + \alpha^{\text{time}}_{\text{time}[i]} + \alpha^{\text{age}*\text{male}}_{\text{age}*\text{male}[i]},$$

where male_i is an indicator taking on the value 0.5 for men and -0.5 for women; $\text{age}[i], \text{race}[i]$, and $\text{county}[i]$ represent age, race, and county categories, with a two-way interaction term $\text{age} * \text{male}[i]$; $\text{time}[i]$ indices the time in weeks when the test result is observed for individual i; and the α parameters are vectors of varying intercepts to which we assign hierarchical priors:

$$\alpha^{\text{name}} \sim \text{normal}(0, \sigma^{\text{name}}), \qquad \sigma^{\text{name}} \sim \text{normal}_+(0, 2.5),$$

for $\text{name} \in \{\text{age}, \text{race}, \text{county}, \text{age} * \text{male}\}$. And we set the time-varying effect: $\alpha^{\text{time}} \sim \text{normal}(0, \sigma^{\text{time}})$, $\sigma^{\text{time}} \sim \text{normal}_+(0, 5)$, to allow for the possibility of large variations across time. The larger the estimated variation, the larger effects of the predictors.

Assume the prior information for the unknown sensitivity δ and specificity γ includes: y_γ negative results in n_γ tests of known negative subjects and y_α positive results from n_α tests of known positive subjects. The model for the number of positive results y out of n tests is specified as

$$y_\gamma \sim \text{Binomial}(n_\gamma, \gamma), \quad y_\delta \sim \text{Binomial}(n_\delta, \delta).$$

According to the test protocol, the sensitivity is around 70%, and the specificity is around 100%. We solicit prior information from previous testing results (2). For the sensitivity, the prior data y_δ/n_δ are: 70/100, 78/85, 27/37, and 25/35; and the prior data for the specificity y_γ/n_γ are: 0/0, 368/371, 30/30, 70/70, 1102/1102, 300/300, 311/311, 500/500, 198/200, 99/99, 29/31, 146/150, 105/108, and 50/52.

After fitting the Bayesian model, we adjust for the selection bias by applying the sociodemographic distributions in the hospital system and the community to generate the population level prevalence estimates, as the poststratification step in MRP. For each of the $2 * 5 * 3 * 2$ cells in the cross-tabulation table of sex (2 levels), age (5 levels), race (3 levels) and county (2 levels), we have the cell-wise incidence estimate $\hat{\pi}_j$, and population count N_j, where j is the cell index, and calculate the weekly prevalence estimate in the population,

$$\pi_{avg} = \sum_j N_j \, \hat{\pi}_j / \sum_j N_j.$$

Demographic stability

We examine the observed sociodemographic (sex, race and age) distributions of asymptomatic and symptomatic patients receiving PCR tests over time.
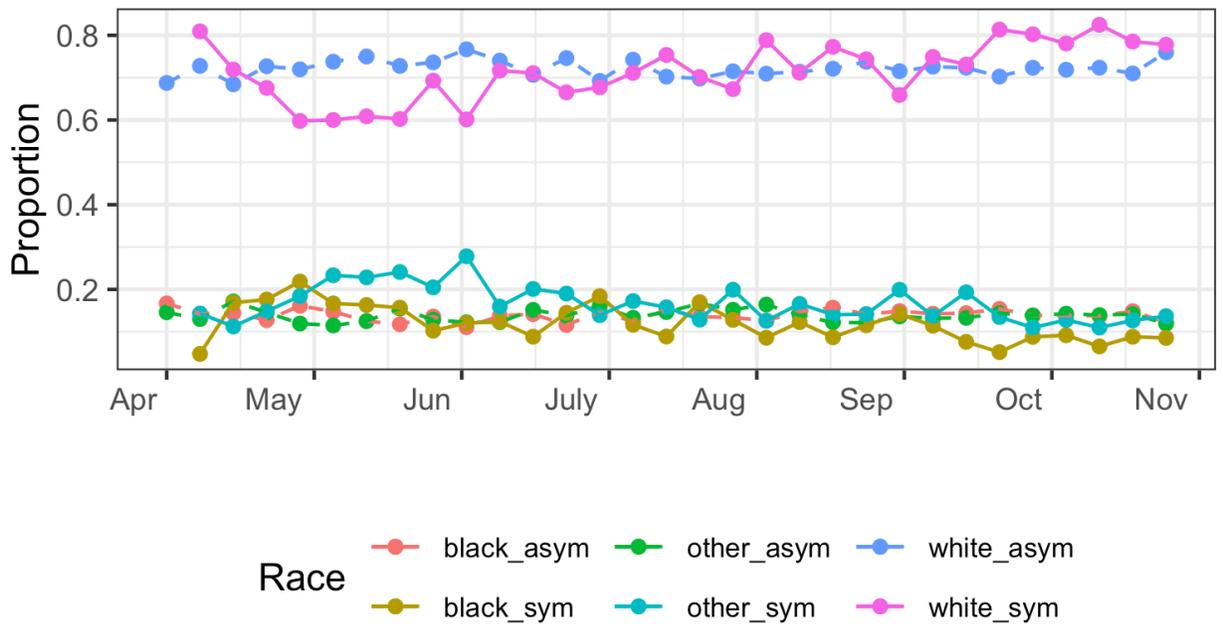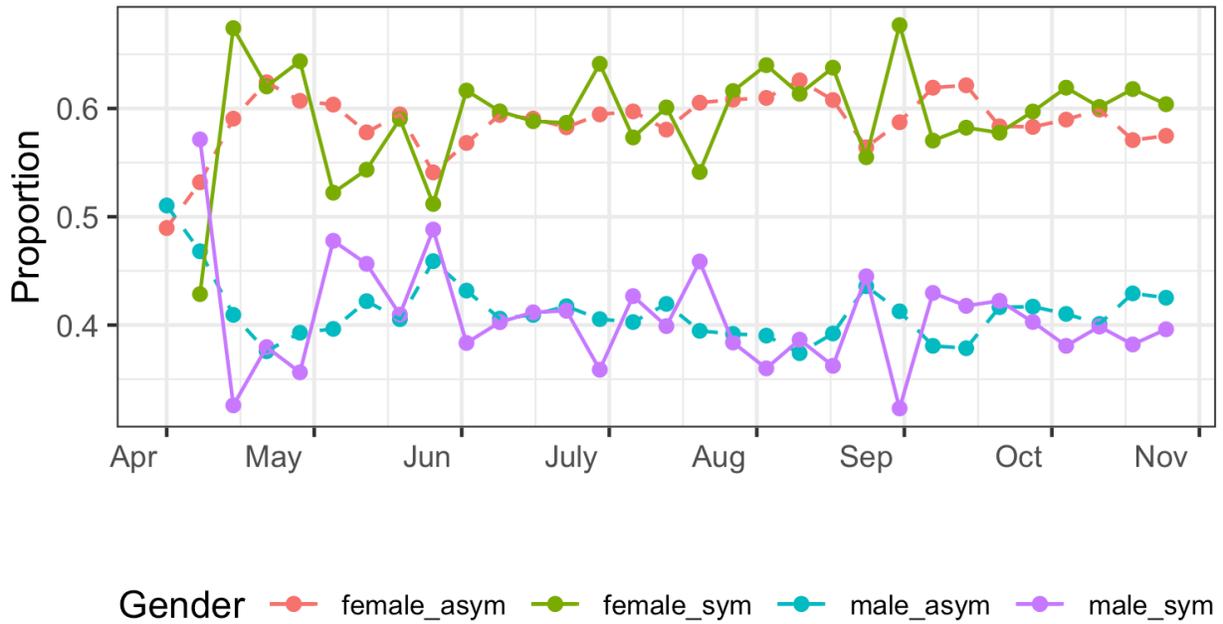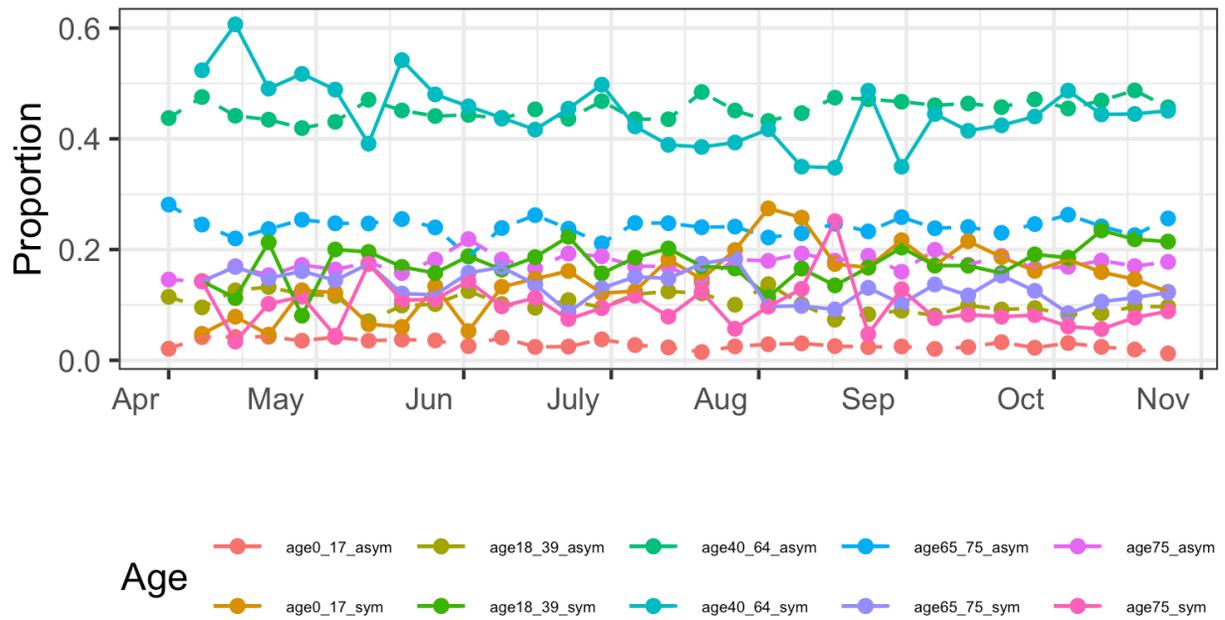
eFigure 1: Demographic distributions of asymptomatic and symptomatic patients across time.

Posterior predictive check

To evaluate the model fitting, we apply a posterior predictive check by generating replicated data from the posterior model distributions with the same sample size as the raw data. We use the collected sample decomposition records every week and estimated prevalence rates of poststratification cells, defined by the cross-tabulation of age, gender, race/ethnicity, and county information, to generate replicated test results. We compare the weekly prevalence rates between the replicated data and observed data. eFigure 2 shows that the model of asymptomatic patients can capture the raw data structure, implying that this aspect of the data is captured well by the fitted model.

We have performed extensive sensitivity analyses of the estimates by changing the modeling mean structure and prior specifications, for example, using spline functions of time, assigning a flexible Gaussian process regression model as the prior distribution of time-varying effects, and changing hyperparameter values. The findings of the hospital- and community-level prevalence estimates are robust without changing conclusions.

We account for the uncertainty of sensitivity and specificity in a Bayesian framework and use the meta-analysis study findings as the prior specification (1). The presented results above are based on the prior information concentrated on the sensitivity value of 70% and specificity 100%. When we set the prior sensitivity data as 70/100, the MRP estimates are similar under the current prior setting. We also compare the PCR results when the prior sensitivity value is set at 65%, 60% and 55%; the results are approximately inflated by the reciprocal of the value of sensitivity, suggesting that the magnitude of our estimates is sensitive to the quality of PCR tests. Nevertheless, the trends within a given sensitivity remain stable.
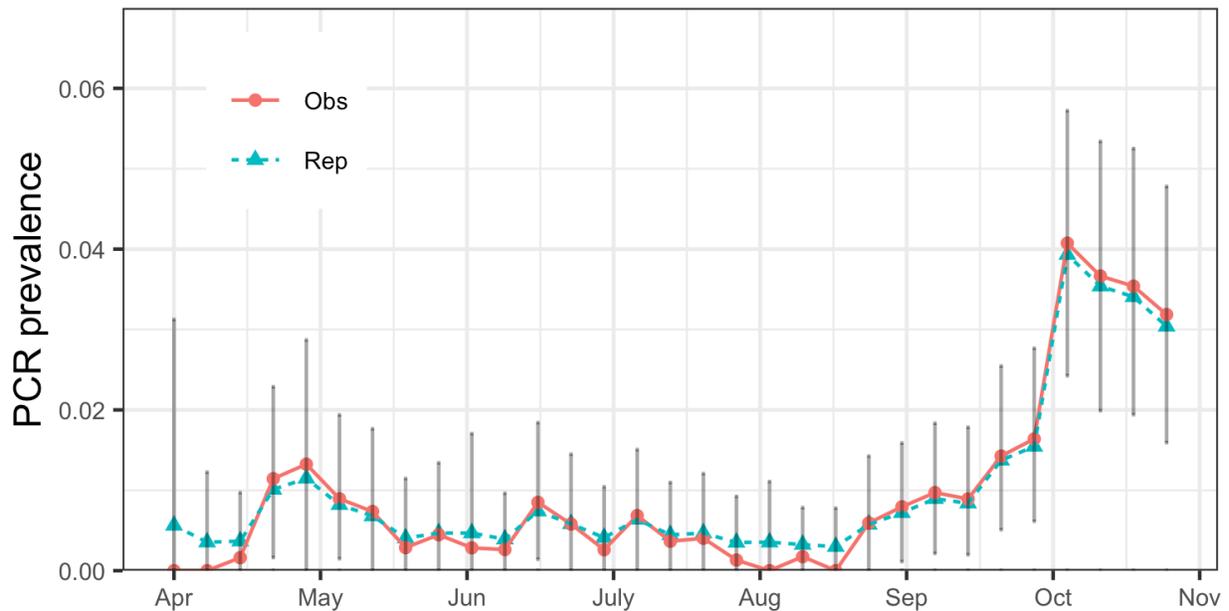
eFigure 2: Posterior predictive check: comparison of replicated and observed prevalence. The error bars represent the 95% credible intervals.

Reference

1. Gelman, Andrew, and Bob Carpenter. Bayesian Analysis of Tests with Unknown Specificity and Sensitivity. Journal of the Royal Statistical Society Series C (Applied Statistics).2020; 69(5):1269-1283.

2. Bendavid, E, B Mulaney, N Sood, et al. COVID- 19 Antibody Seroprevalence in Santa Clara County, California, Version 1. https://www.medrxiv.org/content/10.1101/2020.04.14.20062463v2.full.pdf. Accessed November 21, 2020.