What can we learn about SARS-CoV-2 prevalence from testing and hospital data?*

Daniel W. Sacks†

Nir Menachemi‡

Peter Embí§

Coady Wing¶

January 21, 2022

Abstract

Measuring the prevalence of active SARS-CoV-2 infections in the general population is difficult because tests are conducted on a small and non-random segment of the population. However, hospitalized patients are tested at very high rates, even those admitted for non-COVID reasons. We show how to use information on testing of non-COVID hospitalized patients to obtain tight bounds on population prevalence, under conditions weaker than those usually used. We apply our approach to the population of test and hospitalization data for Indiana, and validate our approach. Our bounds could be constructed at relatively low cost, and for other heavily tested populations.

Key words: COVID-19, SARS-CoV-2, prevalence, partial identification

*We are grateful to Justin Blackburn, Seth Freedman, Aaron Kolb, Charles Manski, and Kosali Simon for helpful comments on earlier drafts, to Evegenia Teal for assistance with the data, and to Agrayan Gupta for helpful research assistance. We acknowledge the Indiana University Pervasive Technology Institute for providing HPC resources that have contributed to the research results reported in this paper (`https://pti.iu.edu`). All errors are our own.

†Business Economics & Public Policy, Kelley School of Business, Indiana University (dansacks@indiana.edu).

‡Fairbanks School of Public Health and Kelley School of Business, Indiana University, and Regenstrief Institute, Inc. (nirmena@iu.edu).

§Biomedical Informatics, Vanderbilt University Medical Center, and Department of Medicine, Vanderbilt University (peter.embi@vumc.org).

¶O'Neill School of Public and Environmental Affairs, Indiana University (cwing@indiana.edu).

1 Introduction

Public and private responses to the spread of an infectious disease rely on accurate and timely estimates of disease prevalence. For example, individual precautionary behaviors respond to the prevalence of COVID-19, HIV/AIDS, Influenza, and Ebola (Allcott et al., 2020; Philipson, 2000, 1996), as do government-mandated shutdowns and other non-pharmaceutical interventions (Gupta et al., 2020). Despite its importance, prevalence is often measured poorly, especially during large outbreaks like the COVID-19 pandemic, when test capacity is scarce and tests are often allocated to symptomatic patients. Simple estimators of prevalence can be misleading in these situations. For example, the fraction of the *overall population* that tests positive for the disease severely understates prevalence because most infected people are not tested. On the other hand, the fraction of the *tested population* that tests positive likely overstates prevalence because tested people are more likely to be infected than untested people.

We propose a new approach to measuring prevalence under incomplete and nonrandom testing. Previous work by Manski and Molinari (2020) suggested estimating cumulative prevalence— the fraction of people who have ever been infected—from available data under a "test monotonicity" assumption. We adapt this assumption to the problem of estimating *current or point-in-time* prevalence. Applying test monotonicity to the general population produces bounds that are typically wide during the COVID-19 pandemic, because testing rates are low. Our primary contribution is to introduce a new approach resulting in tighter bounds. The overarching idea is to exploit data on testing and test results from institutional environments where people are tested at unusually high rates for

reasons that are unrelated to disease prevalence. We apply and validate the idea among patients hospitalized for reasons unrelated to COVID-19, such as labor and delivery (Sutton et al., 2020). But the basic strategy is more general. For example there may also be useful variation in disease testing rates associated with airports, workplaces, conferences, and other settings. Our methods could be used in these other situations to estimate COVID-19 prevalence or perhaps the prevalence of other diseases in a future outbreak.

Our analysis shows that the method is effective at reducing inferential uncertainty about point in time prevalence during the COVID-19 epidemic. In highly tested subpopulations such as hospitalized patients, test monotonicity implies relatively tight bounds on point-in-time prevalence. Of course, these tight bounds are only useful if the heavily tested population is in some sense representative of the general population.

We therefore introduce two representativeness assumptions. The first is "hospitalization monotonicity", which requires that patients hospitalized for non-COVID reasons are (if anything) weakly *more likely* to be infected with SARS-CoV-2 than the general population. This assumption implies that the upper bound on prevalence in the hospitalized subpopulation is a valid upper bound for the whole population. Because testing rates are higher in the non-COVID hospitalized population, the assumption typically implies a tighter population upper bound, albeit with less statistical precision due to lost sample size. We also consider a second assumption – "hospitalization independence" — which is more restrictive. It says that patients hospitalized for non-COVID reasons have the *same* risk as the general population. Using the independence assumption the hospitalized upper bound remains an upper bound on population prevalence. But it also turns out that – under independence – the lower bound for the non-COVID hospitalized subsample is

a lower bound for the general population.

These assumptions, though much weaker than the assumptions that are most often used to estimate prevalence—for example, that all positive cases are tested, or the population is tested at random—are nonetheless restrictive and their validity will likely vary across contexts and sub-populations. Either of these two assumptions permits some extrapolation from the non-COVID hospitalized population to the general population, but neither assumption is sufficient to point identify prevalence. We present results from both sets of assumptions and allow readers to reach their own conclusions.

We analyze data on the near-universe of COVID-19 tests and all-causes hospitalizations in Indiana in 2020. We start by using the test monotonicity condition developed by (Manski and Molinari, 2020) to estimate the weekly prevalence of active COVID-19 infections. These test monotonicity bounds are narrowest in the week of June 12, when they imply that between 0.05 percent and 4.5 percent of the Indiana population had an active infection. Using the hospital data and imposing our new assumptions narrows the bounds substantially. Under the hospital monotonicity assumption, the upper bound falls and we find COVID-19 prevalence in the week of June 12 is between 0.5 to 2.2 percent. Under hospital independence, the bound narrows further to 0.7 to 2.2 percent. In the average week, imposing hospitalization monotonicity on top of test monotonicity reduces the *width* of the prevalence bounds by about half. All three sets of bounds are transparent and simple to calculate. In the results section, we present a range of bounds to enable readers to choose the set of results that correspond to the assumptions they find most credible.

To assess the credibility of our assumptions in the Indiana context, we present three

pieces of evidence. First, we compute prevalence bounds for all COVID-unrelated hospitalizations as well as for specific categories of hospitalizations, such as vehicle accidents, appendicitis, labor and delivery, and heart attacks. The results are similar across these sub-populations, which suggests little selection across hospital conditions on the basis of COVID risk. Second, we show that hospitalized patients are drawn from areas with test rates similar to or lower than the general population, again suggesting similar COVID risk. Third, we show that our bounds contain the prevalence estimates from a random sample survey of the Indiana population (Menachemi et al., 2020; Richard M. Fairbanks School of Public Health, 2020).

Our approach complements existing methods of estimating prevalence. Two common methods are biometric survey samples and backcalculation. In a biometric survey, a representative sample from the population is tested for the disease (Menachemi et al., 2020; Richard M. Fairbanks School of Public Health, 2020; Gudbjartsson et al., 2020). The method is slow and expensive, but likely produces accurate estimates. In backcalculation studies, data on observed hospitalizations or deaths are used to infer disease prevalence at earlier dates using assumptions about the progression of the disease, hospitalization rates, and case fatality rates (Brookmeyer and Gail, 1988; Egan and Hall, 2015; Flaxman et al., 2020; Salje et al., 2020). The approach we pursue combines passively collected administrative data with distributional assumptions that are less restrictive than those used in backcalculation (Manski, 1999; Wing, 2010; Stock et al., 2020; Manski and Molinari, 2020). An advantage of our approach is that, as states already report test rates, positivity rates, and COVID-related hospitalization rate, our bounds could be calculated at little additional cost in terms of data infrastructure, and they could also be calculated for

other heavily tested sub-populations. More generally, our results show that there is considerable value in reporting testing and positivity rates separately by the reason for test; tests that are not driven by symptoms or exposure may be particularly informative about population prevalence.

2 Inferring COVID Prevalence from Incomplete Testing

Our goal is to estimate the weekly prevalence of SARS-CoV-2 infections using the kind of administrative data that is available in most states, accounting for the fact that testing is likely to be unrepresentative. We use partial identification methods to construct upper and lower bounds on prevalence under alternative assumptions. Figure 1 gives a schematic representation of the data, assumptions, and results. Typically, more restrictive assumptions yield tighter bounds. We discuss arguments for and against key assumptions, which should help readers assess the credibility of the restrictions. However, we present estimated bounds under a variety of assumptions, so readers may focus on approach they find most convincing.

2.1 Notation and Worst Case Bounds

Use $i = 1...N$ to index the population of Indiana. C_{it} is a binary indicator that person i is infected with SARS-CoV-2 on date t. Leaving conditioning on the date implicit, the prevalence of SARS-CoV-2 in Indiana at date t is $Pr(C_{it} = 1) = \frac{1}{N}\sum_{i=1}^{N} C_{it}$. Next, let H_{it} be a binary indicator that the person was hospitalized with a specified diagnosis type. Then $Pr(C_{it} = 1|H_{it} = 1)$ is SARS-CoV-2 prevalence in the hospitalized sub-population

in Indiana on date t.

Let D_{it} indicate whether the person was tested. The share of the population tested on date t is $Pr(D_{it} = 1)$. Prevalence among the tested is $Pr(C_{it}|D_{it} = 1)$. $Pr(C_{it}|D_{it} = 0)$ represents prevalence among people who were not tested. By construction, the value of C_{it} is unknown for people with $D_{it} = 0$. This means that $Pr(C_{it}|D_{it} = 0)$ is not identified by the data on testing and test outcomes.

As a practical matter, testing is very rare in most of the country in any given week so C_{it} is unknown for most of the population. In the absence of any assumptions about selection into testing, the worst case bounds on prevalence are:

$$L_w = \underbrace{Pr(C_{it} = 1|D_{it} = 1)Pr(D_{it} = 1)}_{\text{Confirmed Positive Rate}}$$

$$U_w = \underbrace{Pr(C_{it} = 1|D_{it} = 1)Pr(D_{it} = 1)}_{\text{Confirmed Positive Rate}} + \underbrace{Pr(D_{it} = 0)}_{\text{Untested Rate}}.$$

These bounds define the range of prevalence that is compatible with the data. The lower and upper bounds for a given sub-population and time period can formed using relevant proportions from test and hospital data. For inference we use the bootstrap (Manski and Pepper, 2000; Kreider and Pepper, 2007; Manski and Pepper, 2009); see Appendix F for details.

2.2 Test monotonicity

To narrow the bounds, we impose a version of the "test monotonicity" assumption that Manski and Molinari (2020) proposed to compute cumulative prevalence:

Assumption 1. *(Test monotonicity)* $Pr(C_{it} = 1|D_{it} = 1) \geq Pr(C_{it} = 1|D_{it} = 0)$

When we work with specific subpopulations, such as hospitalized patients ($H_{it} = 1$), we use a conditional version of Assumption 1, which requires that $Pr(C_{it} = 1|D_{it} = 1, H_{it} = 1) \geq Pr(C_{it} = 1|D_{it} = 0, H_{it} = 1)$.

Assumption 1 requires that prevalence is at least as high among the tested as among the untested, conditional on any other covariates. Test monotonicity allows for a randomly chosen tested person to be arbitrarily more likely to test positive than a randomly chosen untested person. It also allows the possibility that the tested population may be neither positively nor negatively selected. Importantly, the test not need to hold at the individual level, only in aggregate. For example, it is possible that individual demand for testing responds heterogeneously to situational risk or institutional norms. For instance, some people might *avoid* testing when they experience additional risk factors, perhaps motivated by dread. Likewise some institutions might avoid ordering tests in cases where the test results would be redundant given observed symptoms. These individual level choices would tend to reduce positive selection into testing, and such behaviors are not ruled out by the test monotonicity assumption. The test monotonicity restriction implies only that, on net, these negative selection types do not outnumber the positive selection types.

Test monotonicity constrains the range of prevalence in the untested population. Specifically, Assumption 1 implies that $0 \leq Pr(C_{it} = 1|D_{it} = 0) \leq Pr(C_{it=1}|D_{it} = 1)$. Under test monotonicity the bounds on active prevalence are:

$$L_m = \underbrace{Pr(C_{it} = 1|D_{it} = 1)Pr(D_{it} = 1)}_{\text{Confirmed Positive Rate}},$$

$$U_m = \underbrace{Pr(C_{it} = 1|D_{it} = 1)}_{\text{Test Positivity Rate}}.$$

The new upper bound is the prevalence in the tested population, which is often called the *test positivity rate* and is widely reported. In our data, test rates are often less than 1 percent and positivity rates in the population are often 10 percent or less, so this assumption brings the upper bound down from 99 percent to 10 percent or less. These bounds address positive selection into testing because they are valid under test monotonicity, and hence allow for even potentially extreme selection.

2.3 Inferring Population Prevalence From Non-COVID Hospital Patients

It is straightforward to construct test monotonicity bounds on prevalence overall and in sub-populations, such as hospitalized patients. We emphasize that the test monotonicity assumption must apply in each subpopulation .[1] Testing rates are higher in hospital settings, which means the bounds are much tighter for hospital sub-populations than for the general population. Thus, assumptions that *link* hospital and population prevalence could substantially reduce uncertainty about population prevalence. We pursue two types of assumptions that enable extrapolation from non-COVID hospital popula-

[1] In support of test monotonicity in the hospitalized population, we find that patients with symptoms of COVID are about twice as likely as other patients to be tested.

tions to the general population: (i) hospitalization monotonicity and (ii) hospitalization independence.

2.3.1 Hospitalization Monotonicity

For some situations or patient types, it is reasonable to assume that hospitalized patients are somewhat adversely selected on health. Applied to COVID-19, adverse selection suggests a hospitalization monotonicity assumption that SARS-CoV-2 prevalence is weakly higher in the hospitalized sub-population than the general population. Stated formally:

Assumption 2. *(Hospitalization Monotonicity)* $Pr(C_{it} = 1|H_{it} = 1) \geq Pr(C_{it} = 1)$

The hospitalization monotonicity assumption may be more credible for some types of hospital patients than others and such adjustments can be handled with additional conditioning. The important point is that layering the hospital monotonicity assumption on top of the test monotonicity assumption can help reduce the width of the bounds on prevalence in both the general population and the hospital population.

For example, suppose U_m^H and L_m^H are the upper and lower bounds on prevalence in the hospitalized sub-population under Assumption 1 (test monotonicity). And let U_m and L_m represent test monotonicity bounds in the general population. Adding Assumption 2 (hospitalization monotonicity) creates a cross-population restriction, which implies that the upper bound on *population prevalence* (U_m) cannot be larger than the upper bound on *hospital prevalence* (U_m^H). The bounds on population prevalence under both test monotonicity and hospitalization monotonicity are:

$$U_{m,h} = \min\left\{U_m, U_m^H\right\} = \min\left\{Pr(C_{it}=1|D_{it}=1), Pr(C_{it}=1|D_{it}=1, H_{it}=1)\right\}$$
$$= \min\left\{\text{Population test positivity}, \text{Hospital test positivity}\right\},$$

In practice, the upper bound on prevalence among non-COVID hospital patients is typically lower than the population upper bound. This means that the test positivity rate among non-COVID hospitalizations is an upper bound on population prevalence. Thus Assumption 2 typically tightens the bounds on prevalence. However, there are far fewer hospitalized patients than people in the population, so there is a trade-off between tighter identification and the statistical precision of the estimates of the bounds.

2.3.2 Hospitalization Independence

For some types of patients, it may be credible to assume that the risk of hospitalization is actually unrelated to the risk of SARS-CoV-2 infection. Formally, this type of hospitalization independence assumption can be written:

Assumption 3. *(Hospitalization Independence)* $Pr(C_{it}=1|H_{it}=1) = Pr(C_{it}=1)$

Assumption 3 (hospital independence) implies that people who are hospitalized for a specified non-COVID health condition have the same probability of being infected with the virus as the general population. An equivalent statement is that people with SARS-CoV-2 have the same probability of being hospitalized for a non-COVID condition as people without SARS-CoV-2.

Combining the test monotonicity and hospitalization independence assumptions can narrow the bounds on population prevalence. As before, let U_m and L_m represent test monotonicity bounds in the general population. And let U_m^H and L_m^H be test monotonicity bounds for the hospital population. Under both test monotonicity and hospital independence, the bounds on population prevalence are:

$$
\begin{aligned}
L_{m,ind} &= \max\left\{L_m, L_m^H\right\} \\
&= \max\left\{Pr(C_{it}=1|D_{it}=1)Pr(D_{it}), Pr(C_{it}=1|D_{it}=1,H_{it}=1)Pr(D_{it})\right\}, \\
U_{m,ind} &= \min\left\{U_m, U_m^H\right\} \\
&= \min\left\{Pr(C_{it}=1|D_{it}=1), Pr(C_{it}=1|D_{it}=1,H_{it}=1)\right\}.
\end{aligned}
$$

It turns out $U_{m,ind} = U_{m,h}$ so the upper bound is the same under hospitalization independence and hospitalization monotonicity. The hospital independence assumptions affects the population lower bound. In practice, the lower bound is higher among hospitalized sub-populations which means that – under the independence assumption – the lower bound on population prevalence is the confirmed positive rate among non-COVID hospitalizations.

2.4 Summary and data requirements

Figure 1 summarizes our methodological results and serves as a guide for interpreting our empirical findings. The overall approach requires data on the tested population

that can be linked to hospital inpatient records, which contain diagnosis information. We work with three main assumptions: two weak monotonicity assumptions, and one conditional independence assumption. The flow chart shows which assumptions yield which bounds on population prevalence. Using only data and no assumptions, we have worst-case bounds for prevalence in the general population and for hospitalized sub-populations. Under Assumption 1 (test monotonicity) the bounds tighten. The lower bound becomes the confirmed positive rate and the upper bound the test positivity rate.

Assumptions 2 and 3 let us extrapolate from the hospitalized sub-population to the general population. But these bounds also turn out to be fairly simple objects. Under Assumption 2 (hospitalization monotonicity) the upper bound on population prevalence tightens to the test positivity rate in the hospitalized population. Under Assumption 3 (hospitalization independence), the lower bound on population prevalence tightens to the confirmed positive rate in the hospitalized population and the upper bound is the same as it is under hospital monotonicity.

An appealing feature of these bounds is that they can be calculated without new data collection efforts. Every state already reports the number of tests and the number of positive tests, and many states report the number of COVID-related hospitalizations.[2] To release all of the bounds we report in the paper, states would only have to report test and positivity rates for non-COVID-related hospitalizations. This appears possible because many states already report "suspected" or "under investigation" COVID hospitalizations, defined as hospitalized patients exhibiting COVID-like illness (Arizona Department of

[2] See, e.g., The COVID Tracking Project (2020).

Health Services, 2020; Illinois Department of Public Health, 2020a,b)).[3]

3 Indiana Hospital and Testing Data

3.1 Test and hospitalization data sets

Our test data consist of all tests polymerase chain reaction (PCR) tests for SARS-CoV-2 conducted in Indiana between January 1, 2020 and December 18, 2020, reported to the Regenstrief Institute and available for research. This is the near-universe of PCR tests in the state. Appendix Figure A.1 shows that the number of cases in our tests data matches the state's reported number almost perfectly, until November, when they diverge somewhat, as our data appear to be missing some tests. The consequence of these missing tests, in our framework, is a reduced lower bound.

Our hospitalization data consist of all admissions to hospitals belonging to the Indiana Network for Patient Care (INPC), a health information exchange that centralizes and stores data from health providers across the state of Indiana.[4] The hospital data are derived from the same database that the state uses for reporting hospitalizations on its dashboard (Indiana State Department of Health, 2020). The hospital inpatient data contain separate observations for each admission. The same hospitalization can appear in the data set multiple times, for example because an insurer and a hospital both report it. To de-duplicate these records, we keep one observation per admission time (defined

[3] States reporting both confirmed SARS-CoV-2 hospitalizations and hospitalizations of suspected cases or cases under investigation include California (California Department of Health, 2020), Colorado (Colorado Department of Public Health and Environment, 2020), Mississippi (Mississippi State Department of Health, 2020), Tennessee (Tennessee Department of Health, 2020), and Vermont (Vermont Department of Health, 2020).

[4] See Grannis et al. (2005) for more details.

second-by-second), keeping the observation with the most diagnosis codes.[5] We link the testing and hospital data sets to each other and, for a subset of patients, to demographic information.

3.2 Measuring tests and cases

In-hospital testing, positivity rate, and confirmed positives: Because our data do not record in-hospital testing, we match tests to hospitalizations based on date. We say a hospitalized patient is tested in-hospital if she had at least one SARS-CoV-2 test dated between 5 days prior to admission to 1 day after admission, and she had a positive if she had at least one positive test in that window. We look in a week-long window so that we can compare hospital and population testing rates. This window includes tests that happen a few days prior to admission, important for for patients with planned procedures, as well as tests a day after admission, important for admissions from the emergency department. Appendix Figure A.2, shows that among non-ICLI hospitalizations, test rates are especially elevated from 5 days before admission to 4 days after. We keep the post-admission window short to ensure that we do not pick up hospital-acquired SARS-CoV-2 infections (however Rhee et al. (2020) indicate that hospital-acquired SARS-CoV-2 infections are quite rare).

Population Testing and Positivity Rates: We examine population-wide test rates and positivity rates on a week-by-week basis. We define test rates as the share of the population tested at least once in a given week, and the positive rate as the share of people with

[5] In 0.02 percent of cases where there is ambiguity about which record to keep; in these cases we choose at random.

at least one positive test in a given week, among people tested that week.

3.3 Sample construction

Throughout, a patient is in the "test sample" if they are tested at least once, and in the "inpatient sample" if they are hospitalized at least once. We construct three hospitalization samples, as follows.

ICLI and non-ICLI Hospitalizations: We start by defining hospitalizations for influenza- and COVID-like illness (ICLI) using ICD-10 codes. We collect diagnoses codes for influenza-like illness from (Armed Forces Health Surveillance Center, 2015), and codes for COVID-like illness from (Center for Disease Control and Prevention, 2020). Appendix B lists these ICD-10 codes used to define the analytic samples. These diagnoses include include general symptoms such as cough or fever, as well as more specific diagnoses like acute pneumonia, viral influenza, or COVID-19. A hospitalizations is ICLI-related if it has any influenza- or COVID-like illness (ICLI) diagnoses. Non-ICLI hospitalizations are ones that have a diagnosis but are no ICLI-related diagnoses.

We start our analysis with the non-ICLI sample for two reasons. First, our hospital independence assumption is most plausible for hospitalizations that are not obviously COVID-related, and this sample meets that criteria. Second many states already classify hospitalizations as ICLI-related; thus non-ICLI hospitalizations are identifiable and measurable in near-real time, so this sample can be studied more broadly.

However, the non-ICLI sample may not satisfy the hospital representativeness assumptions. First, inclusion in this sample implicitly risks conditioning on COVID itself.[6]

[6] We show in Appendix E, however, that the magnitude of the resulting sample selection bias is small,

Second, COVID is a new disease with heterogeneous symptoms, so even if a patient is hospitalized because of COVID, she may not have one of our flagged diagnoses, and we may incorrectly call her hospitalization non-ICLI (Yang et al., 2020).

Clear-cause hospitalizations: To avoid these problems, we study a third sample, which we call the "clear cause" sample. These are hospitalizations with a clear cause that is not obviously COVID-related. We define clear-cause hospitalizations as hospitalizations with a diagnosis code for labor and delivery, AMI, stroke, fractures, crushes, open wounds, appendicitis, vehicle accidents, other accidents, or cancer. For all of these conditions except cancer, we flag hospitalizations with a diagnosis at any priority. For cancer, we flag hospitalizations with a cancer diagnosis code as the admitting diagnosis, the primary final diagnosis, or any chemotherapy diagnosis. Although the clear causes do not include COVID-like diagnoses (such as respiratory distress), a clear cause admission can nonetheless also be an ICLI hospitalization. This is because if a patient is admitted for, say, cancer, but has respiratory distress as a morbidity, we would treat it as both a cancer admission and an ICLI admission.

Summary statistics and test rates: We show summary statistics for all of our samples in Table 1, as well as for the state as a whole (from Census Fact Finder and United States Census Bureau (2019)). The average tested and hospitalized patient is substantially older than the population as a whole, and also more likely to be female. Because the tested and hospitalized samples are not age representative of the general population, in what follows we reweight all samples to match the population age distribution.[7] The tested

because ICLI hospitalizations are rare.

[7] Specifically, we calculate test rates and positivity rates in week-by-age-group cells, for age groups 0-17, 18-30, 30-50, 50-64, 65-74, and 75 and older. Then we average these age-specific rates across the age groups,

and hospitalized samples are fairly similar to the general population in terms of racial composition. Limiting the inpatient sample to admissions with diagnoses reduces our sample size substantially, but it does not appear to change its demographic profile. About one-in-three Hoosiers has ever had a COVID test, whereas about half of hospitalized Hoosiers have had a test.

Although hospitalized patients are about 44 percent (i.e. 49%/34%) more likely to have ever been tested than the general public, during the period of their actual hospitalization they are at least ten times more likely to be tested, as we show in Figure 2, which plots weekly age-adjusted for each sample.[8] The testing rate in the general population grew from less than 1 percent in May and June to a peak of 3 percent in mid-November. Test rates for ICLI hopsitalizations varied between 60 and 75 percent in most weeks. Testing rates among non-ICLI hospital patients and among the clear-cause non-COVID hospital patients were 25-40 percent in May and later months.

Despite their high rates, hospitalized patients are not always tested. Even ICLI patients are tested only about two-thirds of the times. Several factors explain incomplete testing. Highly symptomatic patients may not be tested because a test would not necessarily influence care, and could generate a false negative, and testing capacity was sometimes limited. They also might not receive a SARS-CoV-2 test if they had a positive influenza test, as that provides an alternative explanation for the symptoms. For asymptomatic patients, hospital policy encourages testing, but not always require it. The Chief Medical Officer of a large Indiana hospital system indicted that asymptomatic patients

weighting each group by its population share.

[8] We report the exact values of each of the test rates and the weekly number of admissions in Appendix Table A.1. and age-unadjusted rates in Appendix Table A.2.

would typically be tested at admission, but this might vary across hospitals depending on their capacity to isolate patients in private or semi-private rooms (Weaver, 2020). At another large hospital system, the Chief Medical Officer reported that testing was at times based on capacity, but patients coming into particular divisions were more likely to be tested, as were patients coming in for operations (Crabb, 2020). Our personal experience was that hospitals encouraged testing but did not strictly require it.[9] This anecdotal evidence indicates that non-ICLI hospitalizations provides a strong encouragement but not a mandate for testing among asymptomatic or mildly symptomatic patients. Greater testing of mildly symptomatic patients would be consistent with our test monotonicity assumption applied to non-ICLI hospitalizations.

3.4 Justification of hospital representativeness assumptions

Our goal is to use the hospitalization data and hospital representativeness assumptions to tighten bounds on population COVID-19 prevalence. Here briefly provide three justifications for the hospital representativeness assumptions in our data; Appendix G provides more detail. First, we benchmark our bounds against prevalence estimates obtained from the Indiana COVID-19 Random Sample Study (Menachemi et al., 2020; Richard M. Fairbanks School of Public Health, 2020).[10] Our bounds under hospital monotonicity contain the random sample estimates, and we cannot reject the hypothesis that our bounds hospital independence do as well. Second, we show that non-COVID hos-

[9] One of us had a child born in August at a hospital in our sample. The mother was encouraged to obtain a SARS-CoV-2 test prior to admission, but the father (who attended the birth) was not. The mother's test result was not available until after admission. Happily, it was negative.

[10] Our data do not contain the test results from the Random Sample Study, so we compare our bounds to the published results.

pitalized patients are not significantly different from the general population in the likelihood of having been tested for SARS-CoV-2 *prior* to their hospitalization. Interpreting prior testing as a proxy for risk, this suggests that non-COVID hospitalized patients are at a similar risk, consistent with hospital monotonicity and independence. Third, we show below that bounds obtained for a variety of detailed hospitalization subsamples (such as labor and delivery, vehicle accidents, appendicitis, or cancer) are fairly similar. Under the assumption that some of these conditions (e.g. appendicitis) occur independently of COVID risk, this result is consistent with hospitalization monotonicity and independence. It also suggests robustness to the exact conditions used to define non-COVID hospitalization.

4 Bounds on COVID-19 prevalence

4.1 Bounds by broad samples

The high test rates among the hospitalized populations shown in Figure 2 imply relatively tight bounds on in-hospital prevalence under our test monotonicity, and on population prevalence under our representativness assumptions. We plot these population bounds in Figure 3.[11]. The top panel plots population bounds using ICLI-hospitalizations under test monotonicity. The remaining panels plot population bounds for non-ICLI hospitalizations and clear cause hospitalizations, under each representativeness assumption. We also plot 95% confidence intervals for the bounds in dashed lines.

Several patterns are clear in the figure. First, the ICLI hospitalized population has

[11] See Appendix Tables A.3 -A.6 for exact values and age unadjusted bounds

higher upper and lower bounds on prevalence than the other groups (under test monotonicity). For the ICLI patients, the prevalence bounds begin at 6-18 percent in the first week of our sample, increase to 33-39 percent in the last week of March, decline steadily to roughly 12-18 percent in the summer and 8-15 percent in early fall, and increase dramatically in November. Although high, these bounds rule out the possibility that even a majority of symptomatic patients are infected with SARS-CoV-2 in almost every week. The bounds for ICLI-hospitalized patients separate from the population wide bounds, implying rejection of representativeness conditions for the ICLI sub-population.

Second, the hospitalized populations have tight prevalence prevalence bounds; in fact the bounds for the COVID-unrelated hospitalizations sample are always contained within the population-based bounds. Under (conditional) test monotonicity and hospitalization monotonicity, the width of the bounds falls (relative to just test monotonicity) by 60 percent on average. Under hospitalization independence, the width falls by 70 percent. Accounting for the greater statistical uncertainty of the hospital-based estimates, the width still falls by 45-55 percent on average. Indeed, despite some sampling error, in all but three weeks, the 95% confidence interval for bounds under hospital independence are contained in the point estimate for the bounds under test monotonicity alone.

Third, the bounds for the non-ICLI hospitalization sample and for the clear-cause hospitalization sample are nearly indistinguishable, although the confidence intervals are somewhat wider for clear-cause hospitalizations. The main differences are that the upper bound for non-ICLI hospitalization is perhaps slightly higher, and the confidence intervals wider. This fact is important because non-ICLI hospitalizations are potentially easier to measure, but they may be negatively selected in the sense that by construction they

may exclude COVID-likely cases. (However Appendix E shows the bias is likely small in our application.) Empirically, the similarity of the clear-cause and non-ICLI bounds provides some evidence in support of using non-ICLI hospitalizations to measure general prevalence.

Fourth, meaningful trends are evident from the bounds under hospitalization independence. The upper bound for all samples shows a U-shaped pattern, with lower and upper bounds high in the spring, falling in the summer, and rising rapidly through the fall. This pattern does not necessarily indicate that prevalence follows a U-shaped trend, because the lower bound for the population as a whole remains fixed at essentially zero. However the non-ICLI hospitalization bounds are sufficiently tight to confirm that prevalence is lower in mid-summer than late fall. For example, the upper bound the week of September 18 is 1.6 percent; this is lower than the lower bound in any week after October 30. Our 95% confidence intervals also allow us reject that the bounds overlap between September 18 and any week in November. Thus under hospital independence and test monotonicity, our hospital based bounds are tight enough to show that prevalence unambiguously rose from summer to fall.

4.2 Bounds by cause of admission

Our overall clear-cause hospitalization sample pools many distinct causes, including among others labor and delivery, vehicle accidents, and other accidents, including falls. In principle these hospitalizations may differ in their SARS-CoV-2 infection risk. One might worry, for example, that pregnant women are especially cautious and careful not to

become infected. In contrast, people who get into vehicle accidents during the epidemic might be a less cautious group either because they are not careful drivers, or because they are out of the house at all.

Since the credibility of key assumptions may vary across different clear causes, we estimated test rates and bounds separately for each of our clear causes of hospitalizations. Because each individual cause has relatively few hospitalizations, we aggregate across all time periods to form these estimates. We focus on nine sets of causes: AMI (i.e. heart attack), appendicitis, cancer, fractures, labor/delivery, non-vehicles accidents, stroke, vehicle accidents, and wounds. These six groups have between 2,000 and 14,000 hospitalizations each. The age profile varies considerably across groups by cause of admission (Appendix Table A.7). All ages are represented in the cancer sample. Appendicitis and vehicle accidents both afflict more young people. AMI, stroke, and other accidents afflict older people; and labor and delivery is limited, of course, to women of childbearing age.[12] Because not all age groups are represented in every category, we do not age-weight these results.

We report bounds by clear cause of admission in Table 2. Labor and delivery is tested at the lowest rate, about 20 percent; the other groups are tested 25-40 percent of the time. The bounds are similar across all groups. The upper bound ranges from 2.1 percent for cancer to 8.7, and the lower bound (under hospitalization Independence) from .6 to 3.2 percent. The bounds are mutually consistent in the sense that their 95 percent confidence intervals overlap.

[12] These groups are not necessarily mutually exclusive, and in particular there is overlap between injury and accidents.

Patients admitted to the hospital for different reasons and with different demographic profiles are all nonetheless tested at a high rate and with similar bounds on prevalence. We might have worried that cancer patients, with pre-existing conditions, would be more cautious than patients from vehicle accidents. Instead we see both groups are on the low end. Thus overall there is no clear evidence that prevalence bounds differ substantially across types of admissions. This is perhaps reassuring for the view that pooling many distinct causes of admissions can nonetheless generate meaningful bounds on prevalence.

5 Conclusion

In fast-moving pandemics, testing is often limited and rationed to the most symptomatic, making it difficult to infer population prevalence. We propose examining populations that are heavily tested for reasons plausibly unrelated to their underlying disease prevalence, focusing on patients hospitalized during the COVID-19 pandemic for reasons unrelated to COVID-19. A test monotonicity monotonicity assumption yields tight bounds on prevalence for this population. To extrapolate from this population to the general population, we introduce and validate a relatively weak hospitalization monotonicity assumption, and a stronger hospitalization independence assumption. Under either assumption, the hospitalized population yields useful bounds on population prevalence, tighter than those obtained with population wide testing data.

Our bounds could be calculated by states or other health agencies with little additional data infrastructure. Similar bounds could be constructed for other groups that are tested often and for reasons unrelated to COVID-19 risk, such as students, international

travelers, and some workers. Of course, validating the representativeness assumptions (monotonicity or independence) would be important. Overall, our results show the value of reporting test rates and test results separately by reason for testing.

References

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Y. Yang (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *NBER Working Paper* (w26946).

Arizona Department of Health Services (2020). Data dashbaord. https://www.azdhs.gov/preparedness/epidemiology-disease-control/infectious-disease-epidemiology/covid-19/dashboards/index.php. Last accessed July 15, 2020.

Armed Forces Health Surveillance Center (2015, October). Influenza-like illness. AFHSC Standard Case Definitions, available at https://health.mil/Reference-Center/Publications/2015/10/01/Influenza-Like-Illness, last accessed June 28, 2020.

Brookmeyer, R. and M. H. Gail (1988). A method for obtaining short-term projections and lower bounds on the size of the aids epidemic. *Journal of the American Statistical Association 83*(402), 301–308.

California Department of Health (2020). Covid-19. https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/Immunization/ncov2019.aspx. Last accessed July 15, 2020.

Center for Disease Control and Prevention (2020, April). Icd-10-cm official coding and reporting guidelines april 1, 2020 through september 30, 2020. Available at https:

//www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf, last accessed June 28, 2020.

Colorado Department of Public Health and Environment (2020). Case data. https://covid19.colorado.gov/data/case-data. Last accessed July 15, 2020.

Crabb, D. (2020). Email correspondence with Nir Menachemi, December, 22, 2020.

Egan, J. R. and I. M. Hall (2015). A review of back-calculation techniques and their potential to inform mitigation strategies with application to non-transmissible acute infectious diseases. *Journal of The Royal Society Interface 12*(106), 20150096.

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, et al. (2020). Estimating the effects of non-pharmaceutical interventions on covid-19 in europe. *Nature*, 1–8.

Grannis, S. J., P. G. Biondich, B. W. Mamlin, M. G. Wilson, L. Jones, and J. M. Overhage (2005). How disease surveillance systems can serve as practical building blocks for a health information infrastructure: the indiana experience. In *AMIA Annual Symposium Proceedings*, Volume 2005, pp. 286. American Medical Informatics Association.

Gudbjartsson, D. F., A. Helgason, H. Jonsson, O. T. Magnusson, P. Melsted, G. L. Norddahl, J. Saemundsdottir, A. Sigurdsson, P. Sulem, A. B. Agustsdottir, et al. (2020). Spread of sars-cov-2 in the icelandic population. *New England Journal of Medicine*.

Gupta, S., K. I. Simon, and C. Wing (2020). Mandated and voluntary social distancing during the covid-19 epidemic. *Brookings Papers on Economic Activity (Conference Draft)*.

Illinois Department of Public Health (2020a). Covid-19 hospital resource utilization. https://www.dph.illinois.gov/covid19/hospitalization-utilization. Last accessed July 15, 2020.

Illinois Department of Public Health (2020b). Covid-19 syndromic surveillance. https://www.dph.illinois.gov/covid19/syndromic-surveillance. Last accessed July 15, 2020.

Indiana State Department of Health (2020). Novel coronavirus (covid-19). coronavirus.in.gov, last accessed July 24, 2020.

Kreider, B. and J. V. Pepper (2007). Disability and employment: Reevaluating the evidence in light of reporting errors. *Journal of the American Statistical Association 102*(478), 432–441.

Mahajan, S., C. Caraballo, S.-X. Li, Y. Dong, L. Chen, S. K. Huston, R. Srinivasan, C. A. Redlich, A. I. Ko, J. S. Faust, et al. (2021). Sars-cov-2 infection hospitalization rate and infection fatality rate among the non-congregate population in connecticut. *The American Journal of Medicine 134*(6), 812–816.

Manski, C. F. (1999). *Identification problems in the social sciences*. Harvard University Press.

Manski, C. F. and F. Molinari (2020). Estimating the covid-19 infection rate: Anatomy of an inference problem. *Journal of Econometrics*.

Manski, C. F. and J. V. Pepper (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica 68*(4), 997–1010.

Manski, C. F. and J. V. Pepper (2009). More on monotone instrumental variables. *The Econometrics Journal 12*, S200–S216.

Menachemi, N., B. E. Dixon, K. K. Wools-Kaloustian, C. T. Yiannoutsos, and P. K. Halverson (2021). How many sars-cov-2–infected people require hospitalization? using random sample testing to better inform preparedness efforts. *Journal of Public Health Management and Practice 27*(3), 246–250.

Menachemi, N., C. Yiannoutsos, B. Dixon, and et al. (2020). Population point prevalence of sars-cov-2 infection based on a statewide random sample — indiana, april 25–29, 2020. *MMWR Morbidity and Mortality Weekly Report*, 960–964. doi: `http://dx.doi.org/10.15585/mmwr.mm6929e1external`.

Mississippi State Department of Health (2020). Interactive chart: Mississippi covid-19 hospitalizations. `https://msdh.ms.gov/msdhsite/_static/14,21994,420,873.html`. Last accessed July 15, 2020.

Peci, A., A.-L. Winter, E.-C. King, J. Blair, and J. B. Gubbay (2014). Performance of rapid influenza diagnostic testing in outbreak settings. *Journal of clinical microbiology 52*(12), 4309–4317.

Philipson, T. (1996). Private vaccination and public health: an empirical examination for us measles. *Journal of Human Resources*, 611–630.

Philipson, T. (2000). Economic epidemiology and infectious diseases. In *Handbook of Health Economics*, Volume 1, pp. 1761–1799. Elsevier.

Rhee, C., M. Baker, V. Vaidya, R. Tucker, A. Resnick, C. A. Morris, M. Klompas, C. P. E. Program, et al. (2020). Incidence of nosocomial covid-19 in patients hospitalized at a large us academic medical center. *JAMA network open 3*(9), e2020498–e2020498.

Richard M. Fairbanks School of Public Health (2020). Iupui, isdh release findings from phase 2 of covid-19 testing in indiana. Available at https://fsph.iupui.edu/news-events/news/iupui-isdh-release-findings-from-phase-2-of-covid-19-testing-in-indiana.html, last accessed July 5, 2020.

Salje, H., C. T. Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hozé, J. Richet, C.-L. Dubost, et al. (2020). Estimating the burden of sars-cov-2 in france. *Science*.

Stock, J. H., K. M. Aspelund, M. Droste, and C. D. Walker (2020). Identification and estimation of undetected covid-19 cases using testing data from iceland. *medRxiv*.

Sutton, D., K. Fuchs, M. D'alton, and D. Goffman (2020). Universal screening for sars-cov-2 in women admitted for delivery. *New England Journal of Medicine 382*(22), 2163–2164.

Tennessee Department of Health (2020). Current covid-19 hospitalizations. https://www.tn.gov/health/cedep/ncov/data/hospitalization-data/current-covid-hospitalizations.html. Last accessed July 15, 2020.

The COVID Tracking Project (2020). The covid tracking project. https://covidtracking.com/. Last accessed July 15,2020.

UCSF Health Hospital Epidemiology and Infection Prevention (2020). Covid-19 diagnostic testing. https://infectioncontrol.ucsfmedicalcenter.org/sites/g/files/tkssra4681/f/COVID-19-testing-reference-sheet.pdf, last accessed July 31, 2020.

United States Census Bureau (2019, July). Quick facts: Indiana. Available at https://www.census.gov/quickfacts/IN, last accessed July 4, 2020.

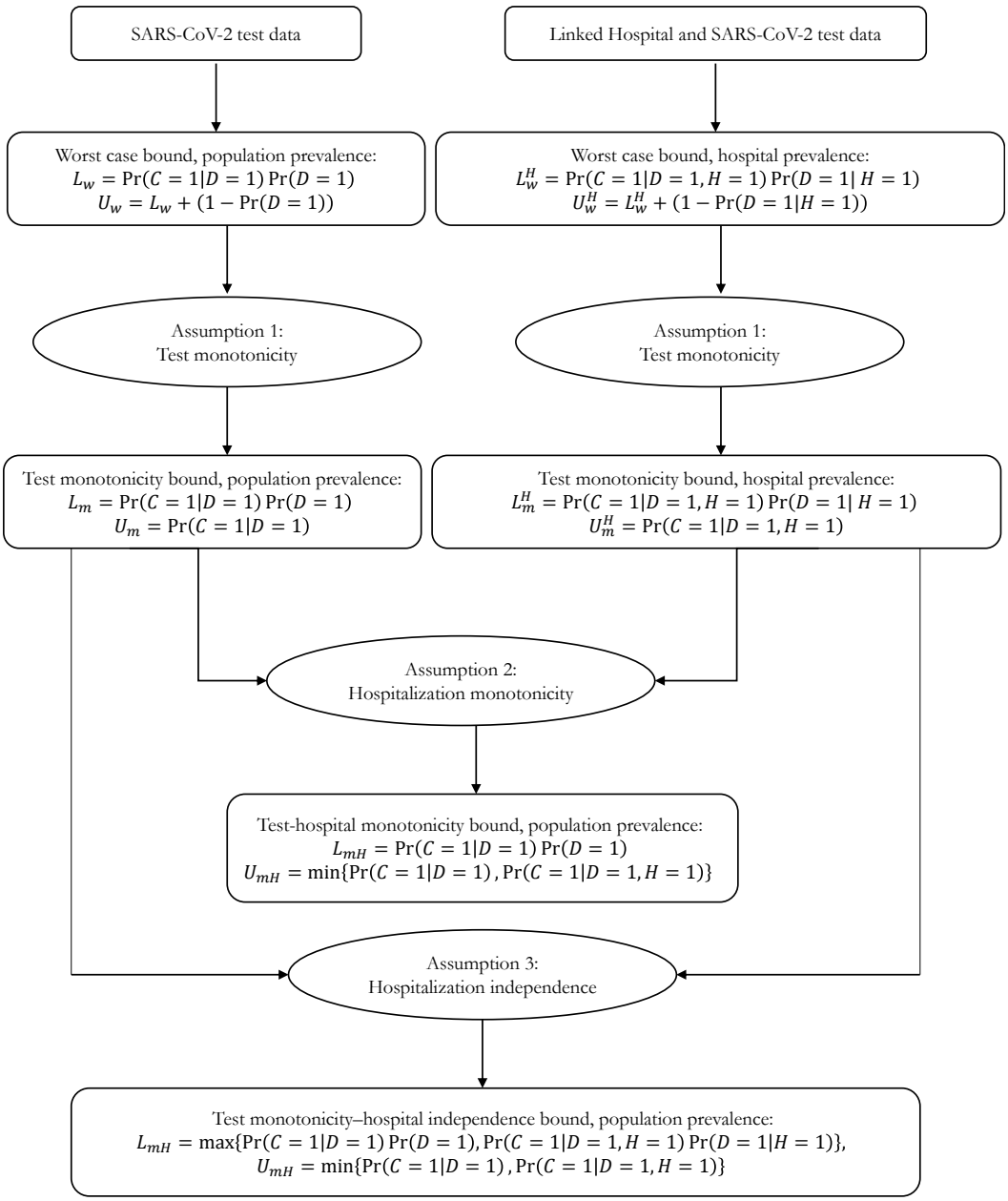
Vermont Department of Health (2020). Current activity in vermont. https://www.healthvermont.gov/response/coronavirus-covid-19/current-activity-vermont. Last accessed July 15, 2020.

Weaver, C. (2020). Email correspondence with Nir Menachemi, December, 22, 2020.

Wing, C. (2010). *Three essays on voluntary HIV testing and the HIV epidemic*. Syracuse University.

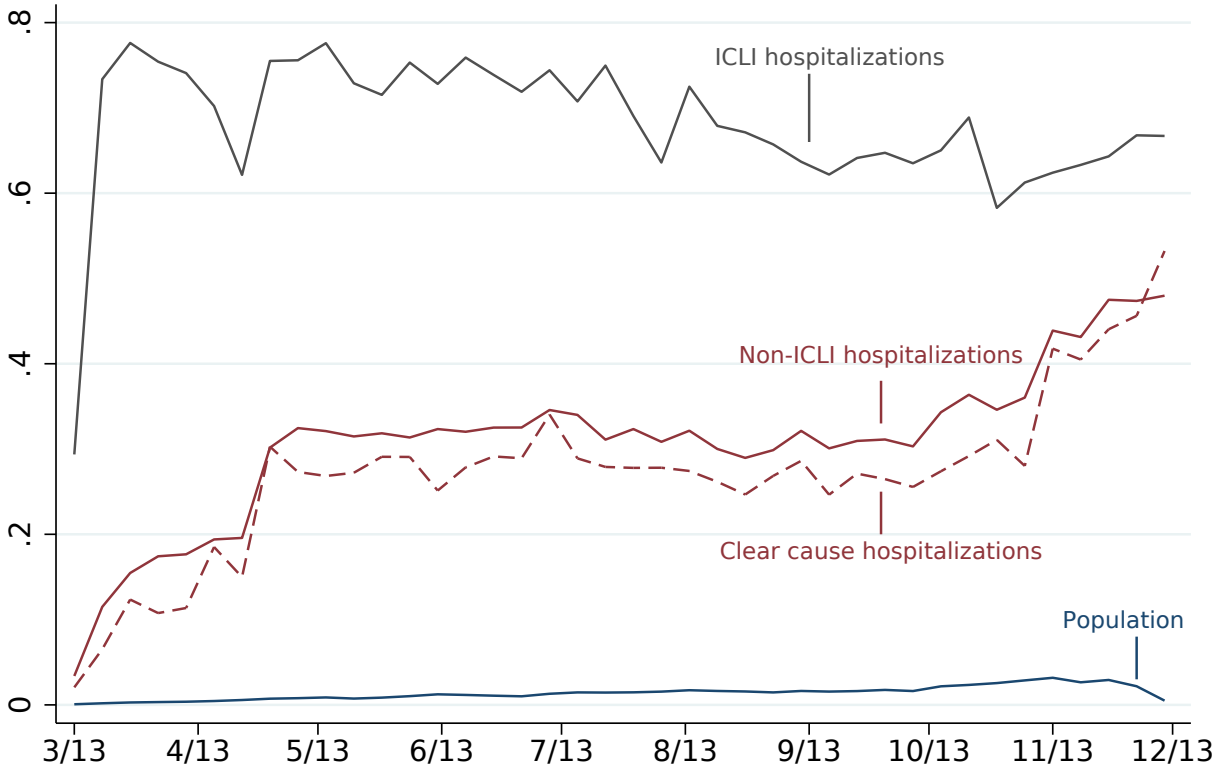
Yang, B. Y., L. M. Barnard, J. M. Emert, C. Drucker, L. Schwarcz, C. R. Counts, D. L. Murphy, S. Guan, K. Kume, K. Rodriquez, et al. (2020). Clinical characteristics of patients with coronavirus disease 2019 (covid-19) receiving emergency medical services in king county, washington. *JAMA Network Open 3*(7), e2014549–e2014549.

Figure 1: Stronger assumptions generate tighter prevalence bounds



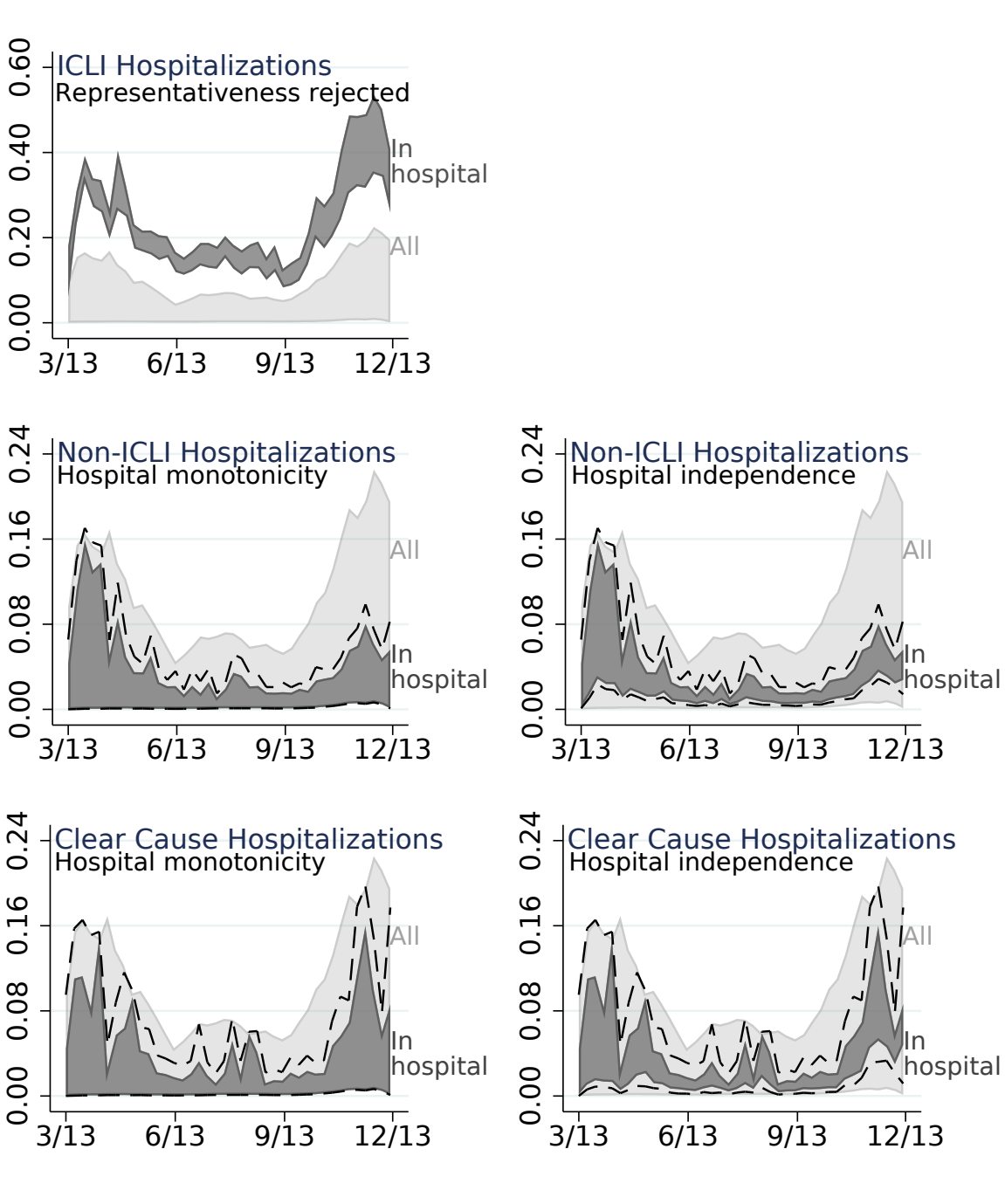
Notes: Figure illustrates how increasingly strong assumptions generate successively tighter bounds on prevalence. Without any assumptions the data yield the worst case bounds. With a test monotonicity assumption, we bound population prevalence and hospital prevalence. With hospitalization monotonicity, the hospital bounds are informative for population prevalence. With hospital independence, the bounds tighten further. The flowchart maintains the assumption of no measurement error throughout, for reasons described in Appendix D.

Figure 2: Weekly test rates by sample



Notes: Figure plots the age-standardized test rate in each seven-day period of our data, for four samples: the general population, ICLI hospitalizations, non-ICLI hospitalizations, and clear cause hospitalizations. ICLI hospitalizations have at least one diagnosis for influenza-like or COVID-like illness. Clear cause hospitalizations are hospitalizations for cancer, labor and delivery, AMI, stroke, fracture or crush, open wound, appendicitis, or accidents (vehicle or other). See Appendix B for definitions. For the general population, the test rate is the fraction of people tested at least once in that week. For the hospitalizations, the test rate is the fraction of hospitalizations admitted in that week with a test between two days prior to admission and four days after. To age-standardize we reweighs the hospitalization samples to match the population age distribution.

Figure 3: Weekly bounds on prevalence under test monotonicity and hospital representativeness assumptions



Notes: The scale in the top figure differs from the others. Figure plots age-standardized bounds on prevalence under test monotonicity and hospital monotonicity or independence, for the indicated hospitalized population, in dark gray. For comparison the figure plots (in light gray) the bound using only population-wide data. See notes to Figure 2 for sample definitions. The ICLI-hospitalization figure plots bounds for hospitalized patient under test monotonicity, but does not plot population bounds,because the representativness assumptions are rejected. The dashed line depict 95% confidence intervals for the bound. To age-standardize we reweight the hospitalization samples to match the population age distribution.

Table 1: Person-level summary statistics

Sample	Full State (1)	Ever Tested (2)	Hospitalized				
			Ever (3)	Has Diagnosis (4)	Not ICLI (5)	Clear cause (6)	ICLI (7)
Age as of 1/1							
Born after 1/1		0.004	0.118	0.129	0.143	0.003	0.005
0-17	0.237	0.142	0.036	0.032	0.030	0.029	0.034
18-29	0.166	0.208	0.121	0.123	0.133	0.181	0.039
30-50	0.250	0.279	0.184	0.180	0.183	0.182	0.138
50-64	0.197	0.202	0.204	0.197	0.189	0.191	0.264
65-74	0.087	0.095	0.162	0.158	0.150	0.171	0.231
75+	0.063	0.070	0.175	0.182	0.172	0.243	0.289
Age unknown		0.001	0.001	0.000	0.000	0.000	0.000
Gender							
Male	0.493	0.442	0.417	0.428	0.420	0.393	0.496
Female	0.507	0.558	0.583	0.572	0.580	0.607	0.504
Unknown		0.012	0.001	0.000	0.000	0.000	0.000
Race/ethnicity							
White	0.848	0.877	0.862	0.834	0.835	0.853	0.828
Black	0.099	0.101	0.121	0.147	0.146	0.128	0.158
Race unknown		0.000	0.000	0.000	0.000	0.000	0.000
Test variables							
Ever tested	0.343	1.000	0.489	0.526	0.498	0.550	0.794
Confirmed positive	0.058	0.169	0.085	0.092	0.061	0.069	0.308
People	6,637,426	2,278,910	539,903	325,410	291,650	66,887	51,870
Counties	92	92	92	92	92	92	92

Notes: Column reports characteristics on the state population, Column 2 reports characteristics for the set of people appearing in the test data, and columns 3-7 for people appearing the hospital data, ever (column 3), with at least one diagnosis (column 4), at least one non-ICLI hospitalization for ICLI (column 5), at least one clear cause hospitalization (column 6, see text for details), or at least one ICLI hospitalization with a diagnosis and not for ICLI (column 7).

Table 2: Bounds on prevalence by cause of admission, pooling all time periods

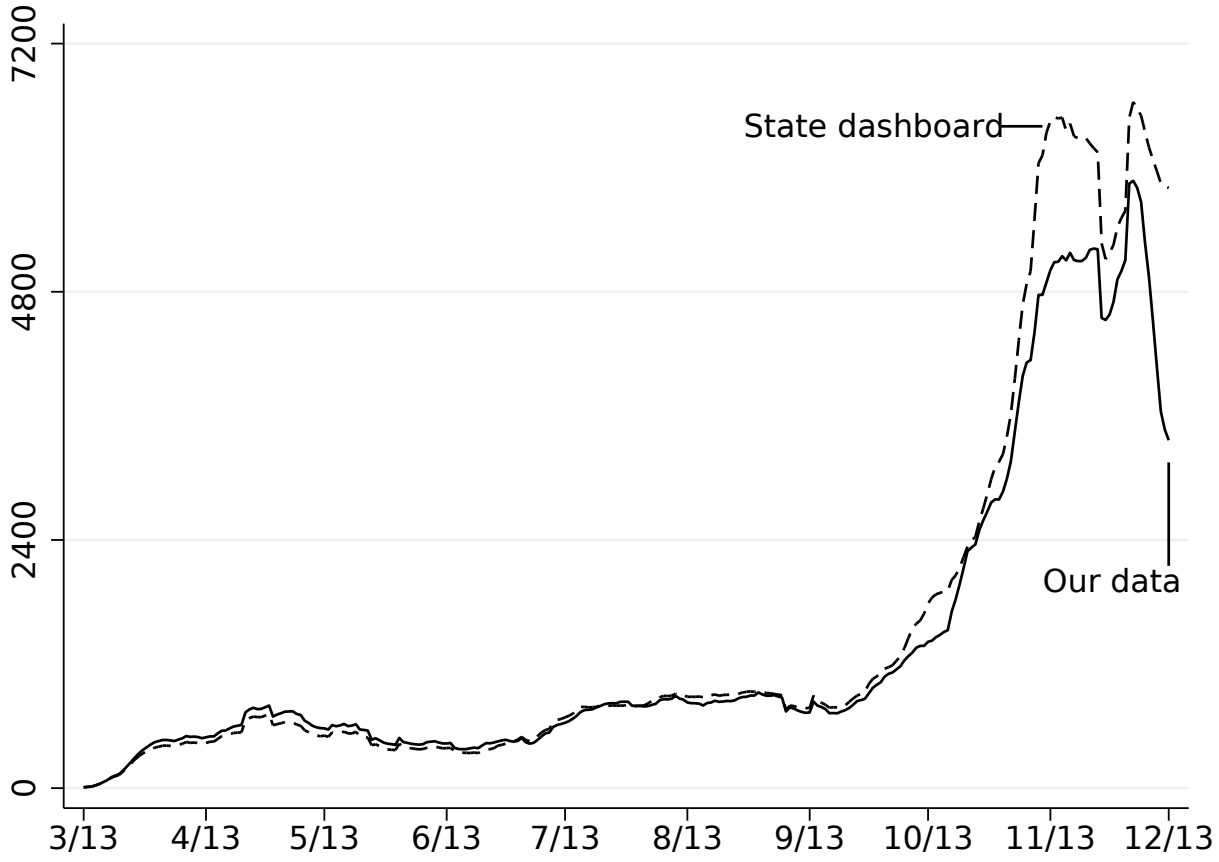
Cause of admission	# Admissions	Test rate	Bound (H-M)	Bound (H-I)
Ami	8,624	0.382	[0.002, 0.085] (0.002, 0.095)	[0.033, 0.085] (0.029, 0.095)
Appendicitis	1,961	0.384	[0.002, 0.045] (0.002, 0.059)	[0.017, 0.045] (0.012, 0.059)
Cancer	9,585	0.337	[0.002, 0.021] (0.002, 0.027)	[0.007, 0.021] (0.006, 0.027)
Fracture	13,718	0.363	[0.002, 0.041] (0.002, 0.046)	[0.015, 0.041] (0.013, 0.046)
Labor Delivery	13,304	0.197	[0.002, 0.041] (0.002, 0.048)	[0.008, 0.041] (0.007, 0.048)
Other Accident	9,782	0.313	[0.002, 0.088] (0.002, 0.098)	[0.027, 0.088] (0.024, 0.098)
Stroke	8,297	0.250	[0.002, 0.066] (0.002, 0.077)	[0.017, 0.066] (0.014, 0.077)
Vehicle Accident	1,944	0.271	[0.002, 0.025] (0.002, 0.038)	[0.007, 0.025] (0.003, 0.038)
Wound	3,642	0.310	[0.002, 0.077] (0.002, 0.093)	[0.024, 0.077] (0.019, 0.093)

Notes: Table reports the number of admissions, in-hospital test rate, and bounds on COVID prevalence, by cause of admission. H-M bounds are valid under hospital monotonicity, and H-I are valid under hospital independence (both require test monotonicity). The H-I and H-M upper bound is identical, and the lower bound for H-M is the same across causes because it is determined by the population confirmed positive rate. The sample consists of all admissions with the indicated cause between March 13 and December 18, 2020. See Appendix B for a precise definition of each cause. We report the point estimate for the bounds in brackets, and 95% confidence interval in parentheses.

For online publication

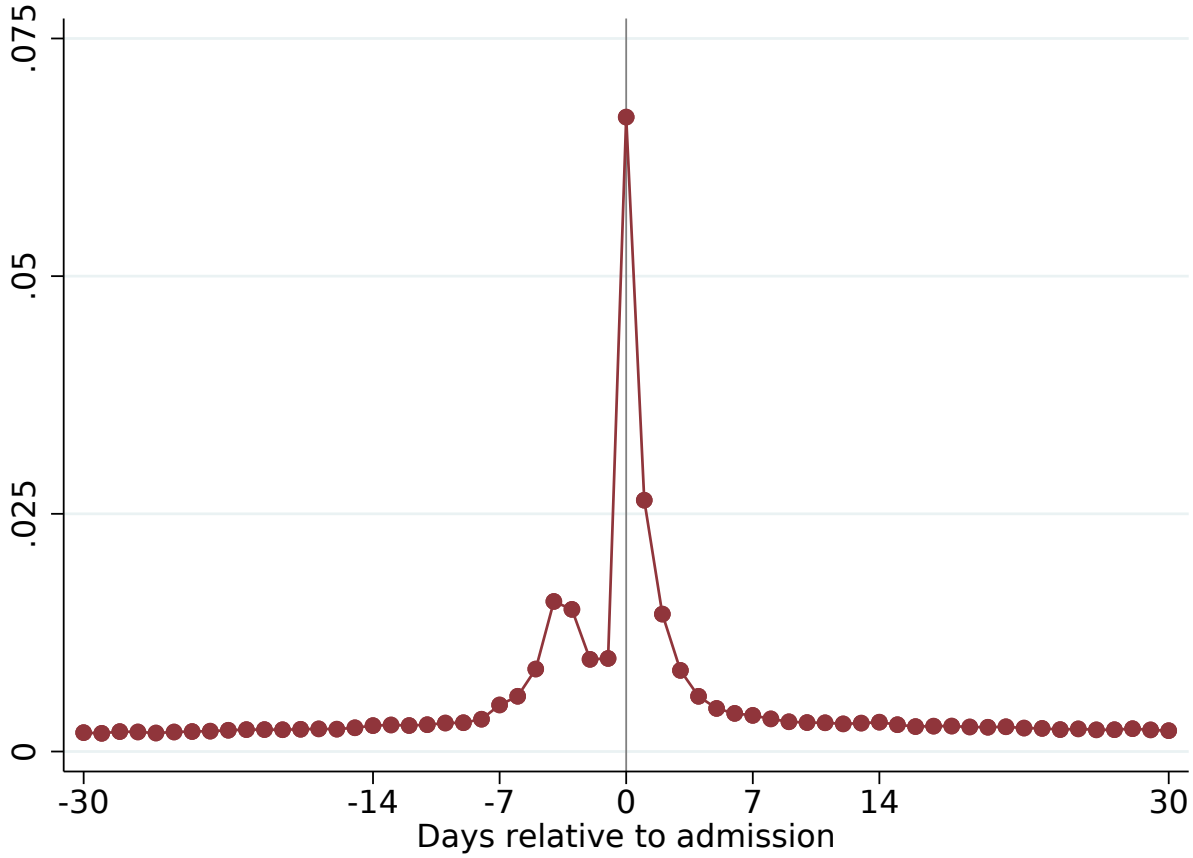
A Appendix figures and tables

Figure A.1: Positive cases in our data and on state dashboard



Notes: Figure plots 7-day moving average of the number of positive cases reported on Indiana's COVID-19 dashboard (Indiana State Department of Health, 2020), as well as the number of positive cases observed in our data.

Figure A.2: Timing of tests relative to hospitalization



Notes: Figure plots the fraction of hospitalized patients who had a SARS-CoV-2 test on the indicated day relative to their admission, for non-ICLI hospitalizations, defind as hospitalizations with no diagnosis for influenza-like or COVID-like illness. Patients who are never tested are in the denominator, and a patient can be tested on multiple days.

Table A.1: Weekly test rates, by sample

Sample	Population	Non-ICLI		Clear Cause		ICLI	
Week	% Tested (1)	N (2)	% Tested (3)	N (4)	% Tested (5)	N (6)	% Tested (7)
13mar	0.001	6,094	0.034	1,363	0.021	1,140	0.294
20mar	0.002	5,173	0.115	1,278	0.065	1,226	0.734
27mar	0.003	4,697	0.155	1,189	0.124	1,342	0.776
03apr	0.003	4,677	0.174	1,273	0.108	1,204	0.754
10apr	0.004	4,841	0.177	1,248	0.114	1,078	0.741
17apr	0.005	5,057	0.194	1,348	0.185	1,191	0.702
24apr	0.006	5,303	0.196	1,431	0.150	1,108	0.621
01may	0.007	5,832	0.302	1,442	0.303	1,195	0.755
08may	0.008	6,187	0.325	1,555	0.273	1,106	0.756
15may	0.009	6,774	0.321	1,576	0.268	1,120	0.776
22may	0.007	6,793	0.315	1,581	0.272	1,059	0.729
29may	0.008	6,949	0.318	1,562	0.291	975	0.715
05jun	0.010	7,359	0.314	1,715	0.291	996	0.753
12jun	0.012	7,554	0.323	1,707	0.251	974	0.728
19jun	0.012	7,437	0.320	1,644	0.279	946	0.759
26jun	0.011	7,434	0.325	1,655	0.291	898	0.739
03jul	0.010	7,441	0.325	1,660	0.289	937	0.719
10jul	0.013	7,591	0.346	1,702	0.340	1,050	0.744
17jul	0.015	7,650	0.340	1,658	0.289	992	0.708
24jul	0.014	7,671	0.311	1,675	0.279	1,028	0.750
31jul	0.015	7,516	0.323	1,634	0.278	1,126	0.690
07aug	0.016	7,722	0.308	1,638	0.278	1,106	0.636
14aug	0.017	7,709	0.321	1,751	0.274	1,048	0.725
21aug	0.016	7,745	0.300	1,770	0.262	1,114	0.679
28aug	0.016	7,733	0.290	1,715	0.247	1,006	0.671
04sep	0.015	7,448	0.299	1,637	0.268	1,015	0.657
11sep	0.016	7,871	0.321	1,717	0.286	1,030	0.637
18sep	0.016	7,857	0.301	1,719	0.246	1,074	0.622
25sep	0.016	7,791	0.310	1,678	0.271	1,146	0.641
02oct	0.018	7,655	0.311	1,713	0.265	1,306	0.647
09oct	0.016	7,476	0.303	1,622	0.256	1,414	0.635
16oct	0.022	7,333	0.343	1,597	0.274	1,457	0.650
23oct	0.023	7,356	0.364	1,603	0.292	1,521	0.689
30oct	0.026	7,379	0.346	1,568	0.311	1,629	0.583
06nov	0.029	7,508	0.360	1,705	0.280	2,126	0.612
13nov	0.032	7,013	0.439	1,579	0.418	2,280	0.624
20nov	0.026	6,268	0.431	1,462	0.405	2,152	0.633
27nov	0.029	6,433	0.475	1,466	0.440	2,151	0.643
04dec	0.022	5,668	0.474	1,217	0.456	1,716	0.668
11dec	0.005	2,284	0.480	426	0.532	689	0.667

Notes: Table reports the weekly test rate for the population, and the number of hospitalizations and test rate, by type of hospitalizations, weighted to match the population age distribution.. (The population size is 6.64 million in all weeks.) ICLI hospitalizations have at least one diagnosis for influenza-like or COVID-like illness. Clear cause hospitalizations are hospitalizations for cancer, labor and delivery, AMI, stroke, fracture or crush, open wound, appendicitis, or accidents (vehicle or other). See Appendix B for definitions.

Table A.2: Weekly test rates, by sample, not age weighted

Sample	Population	Non-ICLI		Clear Cause		ICLI	
	% Tested	N	% Tested	N	% Tested	N	% Tested
Week	(1)	(2)	(3)	(4)	(5)	(6)	(7)
13mar	0.001	6,919	0.034	1,368	0.030	1,143	0.294
20mar	0.002	6,011	0.102	1,281	0.085	1,231	0.726
27mar	0.003	5,546	0.158	1,193	0.142	1,344	0.833
03apr	0.003	5,561	0.173	1,281	0.158	1,207	0.795
10apr	0.004	5,703	0.190	1,250	0.174	1,079	0.772
17apr	0.005	5,914	0.196	1,353	0.217	1,195	0.720
24apr	0.006	6,139	0.196	1,434	0.178	1,110	0.642
01may	0.007	6,711	0.287	1,444	0.328	1,200	0.756
08may	0.008	7,101	0.310	1,556	0.330	1,110	0.746
15may	0.009	7,704	0.322	1,580	0.330	1,128	0.762
22may	0.007	7,717	0.305	1,586	0.308	1,061	0.734
29may	0.009	7,877	0.316	1,569	0.327	982	0.737
05jun	0.010	8,270	0.321	1,721	0.317	1,001	0.744
12jun	0.012	8,432	0.318	1,711	0.310	983	0.702
19jun	0.012	8,313	0.319	1,651	0.319	948	0.694
26jun	0.011	8,381	0.318	1,661	0.332	903	0.708
03jul	0.010	8,367	0.311	1,668	0.321	945	0.705
10jul	0.013	8,547	0.325	1,708	0.341	1,052	0.691
17jul	0.015	8,655	0.327	1,661	0.335	1,006	0.706
24jul	0.015	8,601	0.307	1,680	0.305	1,033	0.712
31jul	0.015	8,505	0.320	1,641	0.315	1,135	0.674
07aug	0.016	8,666	0.304	1,642	0.322	1,108	0.663
14aug	0.017	8,619	0.301	1,755	0.292	1,052	0.649
21aug	0.016	8,707	0.287	1,776	0.293	1,121	0.647
28aug	0.016	8,688	0.284	1,715	0.283	1,010	0.617
04sep	0.015	8,381	0.287	1,647	0.293	1,020	0.635
11sep	0.016	8,860	0.305	1,719	0.308	1,040	0.651
18sep	0.016	8,797	0.297	1,724	0.286	1,075	0.635
25sep	0.016	8,717	0.305	1,687	0.305	1,157	0.624
02oct	0.018	8,585	0.306	1,721	0.310	1,310	0.656
09oct	0.016	8,369	0.302	1,629	0.304	1,422	0.655
16oct	0.022	8,191	0.335	1,599	0.326	1,463	0.638
23oct	0.024	8,190	0.348	1,608	0.341	1,530	0.661
30oct	0.026	8,252	0.332	1,574	0.322	1,639	0.610
06nov	0.029	8,326	0.365	1,710	0.341	2,138	0.640
13nov	0.032	7,900	0.429	1,586	0.465	2,287	0.673
20nov	0.027	7,140	0.418	1,467	0.460	2,159	0.661
27nov	0.030	7,253	0.453	1,473	0.491	2,163	0.671
04dec	0.022	6,384	0.452	1,225	0.491	1,725	0.669
11dec	0.005	2,522	0.406	428	0.467	689	0.597

Notes: Table reports the weekly test rate for the population, and the number of hospitalizations and test rate, by type of hospitalizations. (The population size is 6.64 million in all weeks.) ICLI hospitalizations have at least one diagnosis for influenza-like or COVID-like illness. Clear cause hospitalizations are hospitalizations for cancer, labor and delivery, AMI, stroke, fracture or crush, open wound, appendicitis, or accidents (vehicle or other). See Appendix B for definitions.

Table A.3: Weekly bounds on prevalence under test monotonicity, by sample, March-July

Sample	Pop	Non-ICLI		Clear cause	
Representatives assumption		Hosp-M	Hosp-I	Hosp-M	Hosp-I
Week	(1)	(2)	(3)	(4)	(5)
13mar	[0.0001, 0.097]	[0.0001, 0.043]	[0.0019, 0.043]	[0.0001, 0.044]	[0.0017, 0.044]
	(0.0001, 0.105)	(0.0001, 0.066)	(0.0010, 0.066)	(0.0001, 0.095)	(0.0001, 0.095)
20mar	[0.0003, 0.154]	[0.0003, 0.112]	[0.0139, 0.112]	[0.0003, 0.110]	[0.0107, 0.110]
	(0.0003, 0.162)	(0.0003, 0.140)	(0.0108, 0.140)	(0.0003, 0.157)	(0.0063, 0.157)
27mar	[0.0006, 0.166]	[0.0006, 0.159]	[0.0286, 0.159]	[0.0006, 0.113]	[0.0148, 0.113]
	(0.0005, 0.173)	(0.0005, 0.170)	(0.0234, 0.170)	(0.0005, 0.166)	(0.0089, 0.166)
03apr	[0.0006, 0.154]	[0.0006, 0.131]	[0.0238, 0.131]	[0.0006, 0.083]	[0.0135, 0.083]
	(0.0006, 0.161)	(0.0006, 0.157)	(0.0190, 0.157)	(0.0006, 0.151)	(0.0083, 0.151)
10apr	[0.0006, 0.149]	[0.0006, 0.138]	[0.0236, 0.138]	[0.0006, 0.149]	[0.0132, 0.149]
	(0.0006, 0.155)	(0.0006, 0.154)	(0.0182, 0.154)	(0.0006, 0.154)	(0.0074, 0.154)
17apr	[0.0008, 0.169]	[0.0008, 0.051]	[0.0115, 0.051]	[0.0008, 0.026]	[0.0054, 0.026]
	(0.0008, 0.175)	(0.0008, 0.065)	(0.0091, 0.065)	(0.0008, 0.048)	(0.0027, 0.048)
24apr	[0.0008, 0.137]	[0.0008, 0.086]	[0.0182, 0.086]	[0.0008, 0.057]	[0.0108, 0.057]
	(0.0008, 0.142)	(0.0008, 0.119)	(0.0142, 0.119)	(0.0008, 0.088)	(0.0058, 0.088)
01may	[0.0009, 0.122]	[0.0009, 0.049]	[0.0153, 0.049]	[0.0009, 0.064]	[0.0192, 0.064]
	(0.0009, 0.127)	(0.0009, 0.070)	(0.0113, 0.070)	(0.0009, 0.116)	(0.0094, 0.116)
08may	[0.0008, 0.096]	[0.0008, 0.035]	[0.0119, 0.035]	[0.0008, 0.093]	[0.0215, 0.093]
	(0.0007, 0.100)	(0.0007, 0.050)	(0.0078, 0.050)	(0.0007, 0.099)	(0.0091, 0.099)
15may	[0.0008, 0.099]	[0.0008, 0.035]	[0.0120, 0.035]	[0.0008, 0.043]	[0.0121, 0.043]
	(0.0008, 0.102)	(0.0008, 0.043)	(0.0098, 0.043)	(0.0008, 0.065)	(0.0074, 0.065)
22may	[0.0006, 0.087]	[0.0006, 0.051]	[0.0158, 0.051]	[0.0006, 0.040]	[0.0110, 0.040]
	(0.0006, 0.090)	(0.0006, 0.069)	(0.0113, 0.069)	(0.0006, 0.063)	(0.0067, 0.063)
29may	[0.0006, 0.073]	[0.0006, 0.025]	[0.0083, 0.025]	[0.0006, 0.022]	[0.0070, 0.022]
	(0.0006, 0.076)	(0.0006, 0.038)	(0.0058, 0.038)	(0.0006, 0.038)	(0.0032, 0.038)
05jun	[0.0006, 0.059]	[0.0006, 0.022]	[0.0073, 0.022]	[0.0006, 0.021]	[0.0064, 0.021]
	(0.0006, 0.061)	(0.0006, 0.028)	(0.0055, 0.028)	(0.0006, 0.035)	(0.0023, 0.035)
12jun	[0.0005, 0.045]	[0.0005, 0.022]	[0.0067, 0.022]	[0.0005, 0.018]	[0.0053, 0.018]
	(0.0005, 0.047)	(0.0005, 0.036)	(0.0040, 0.036)	(0.0005, 0.031)	(0.0022, 0.031)
19jun	[0.0006, 0.051]	[0.0006, 0.014]	[0.0047, 0.014]	[0.0006, 0.016]	[0.0046, 0.016]
	(0.0005, 0.053)	(0.0005, 0.018)	(0.0032, 0.018)	(0.0005, 0.029)	(0.0013, 0.029)
26jun	[0.0006, 0.060]	[0.0006, 0.022]	[0.0070, 0.022]	[0.0006, 0.021]	[0.0071, 0.021]
	(0.0006, 0.062)	(0.0006, 0.037)	(0.0040, 0.037)	(0.0006, 0.033)	(0.0037, 0.033)
03jul	[0.0007, 0.069]	[0.0007, 0.015]	[0.0049, 0.015]	[0.0007, 0.033]	[0.0088, 0.033]
	(0.0007, 0.071)	(0.0007, 0.026)	(0.0030, 0.026)	(0.0007, 0.067)	(0.0027, 0.067)
10jul	[0.0009, 0.067]	[0.0009, 0.026]	[0.0086, 0.026]	[0.0009, 0.020]	[0.0067, 0.020]
	(0.0009, 0.069)	(0.0009, 0.038)	(0.0054, 0.038)	(0.0009, 0.032)	(0.0033, 0.032)
17jul	[0.0010, 0.069]	[0.0010, 0.011]	[0.0041, 0.011]	[0.0010, 0.012]	[0.0042, 0.012]
	(0.0010, 0.071)	(0.0010, 0.015)	(0.0029, 0.015)	(0.0010, 0.021)	(0.0016, 0.021)
24jul	[0.0010, 0.072]	[0.0010, 0.019]	[0.0064, 0.019]	[0.0010, 0.021]	[0.0065, 0.021]
	(0.0010, 0.074)	(0.0010, 0.024)	(0.0047, 0.024)	(0.0010, 0.033)	(0.0032, 0.033)
31jul	[0.0011, 0.072]	[0.0011, 0.034]	[0.0104, 0.034]	[0.0011, 0.052]	[0.0110, 0.052]
	(0.0010, 0.074)	(0.0010, 0.052)	(0.0067, 0.052)	(0.0010, 0.073)	(0.0041, 0.073)

Notes: Table reports weekly bounds on COVID prevalence in the indicated sample under test monotonicity, as well as the indicated representativeness assumption. Hosp-M means hosptilalization monotonicity, and Hosp-I means hospitalization independence. Hospitalized samples weighted to match the population age distribution. See Table A.1 for sample definitions. Bounds are in brackets, 95% confidence intervals in parentheses.

Table A.4: Weekly bounds on prevalence under test monotonicity, by sample, August-December

Sample	Pop	Non-ICLI		Clear cause	
Representatives assumption		Hosp-M	Hosp-I	Hosp-M	Hosp-I
Week	(1)	(2)	(3)	(4)	(5)
07aug	[0.0010, 0.067]	[0.0010, 0.032]	[0.0087, 0.032]	[0.0010, 0.020]	[0.0065, 0.020]
	(0.0010, 0.068)	(0.0010, 0.048)	(0.0054, 0.048)	(0.0010, 0.033)	(0.0032, 0.033)
14aug	[0.0010, 0.059]	[0.0010, 0.022]	[0.0070, 0.022]	[0.0010, 0.059]	[0.0173, 0.059]
	(0.0010, 0.060)	(0.0010, 0.033)	(0.0043, 0.033)	(0.0010, 0.060)	(0.0079, 0.060)
21aug	[0.0010, 0.060]	[0.0010, 0.022]	[0.0066, 0.022]	[0.0010, 0.041]	[0.0098, 0.041]
	(0.0010, 0.062)	(0.0010, 0.033)	(0.0042, 0.033)	(0.0010, 0.061)	(0.0038, 0.061)
28aug	[0.0010, 0.062]	[0.0010, 0.016]	[0.0052, 0.016]	[0.0010, 0.012]	[0.0035, 0.012]
	(0.0010, 0.063)	(0.0010, 0.021)	(0.0038, 0.021)	(0.0010, 0.022)	(0.0013, 0.022)
04sep	[0.0009, 0.057]	[0.0009, 0.016]	[0.0049, 0.016]	[0.0009, 0.015]	[0.0042, 0.015]
	(0.0008, 0.058)	(0.0008, 0.021)	(0.0036, 0.021)	(0.0008, 0.025)	(0.0020, 0.025)
11sep	[0.0009, 0.053]	[0.0009, 0.016]	[0.0053, 0.016]	[0.0009, 0.014]	[0.0043, 0.014]
	(0.0009, 0.055)	(0.0009, 0.025)	(0.0032, 0.025)	(0.0009, 0.022)	(0.0021, 0.022)
18sep	[0.0009, 0.058]	[0.0009, 0.016]	[0.0051, 0.016]	[0.0009, 0.022]	[0.0064, 0.022]
	(0.0009, 0.059)	(0.0009, 0.021)	(0.0036, 0.021)	(0.0009, 0.036)	(0.0030, 0.036)
25sep	[0.0012, 0.070]	[0.0012, 0.019]	[0.0064, 0.019]	[0.0012, 0.018]	[0.0061, 0.018]
	(0.0011, 0.072)	(0.0011, 0.024)	(0.0046, 0.024)	(0.0011, 0.029)	(0.0026, 0.029)
02oct	[0.0014, 0.081]	[0.0014, 0.018]	[0.0058, 0.018]	[0.0014, 0.024]	[0.0065, 0.024]
	(0.0014, 0.083)	(0.0014, 0.023)	(0.0044, 0.023)	(0.0014, 0.038)	(0.0029, 0.038)
09oct	[0.0017, 0.101]	[0.0017, 0.027]	[0.0086, 0.027]	[0.0017, 0.021]	[0.0072, 0.021]
	(0.0016, 0.102)	(0.0016, 0.040)	(0.0064, 0.040)	(0.0016, 0.031)	(0.0038, 0.031)
16oct	[0.0024, 0.110]	[0.0024, 0.029]	[0.0103, 0.029]	[0.0024, 0.022]	[0.0073, 0.022]
	(0.0024, 0.111)	(0.0024, 0.038)	(0.0078, 0.038)	(0.0024, 0.033)	(0.0039, 0.033)
23oct	[0.0031, 0.132]	[0.0031, 0.030]	[0.0115, 0.030]	[0.0031, 0.047]	[0.0154, 0.047]
	(0.0031, 0.134)	(0.0031, 0.037)	(0.0095, 0.037)	(0.0031, 0.070)	(0.0102, 0.070)
30oct	[0.0042, 0.162]	[0.0042, 0.038]	[0.0136, 0.038]	[0.0042, 0.057]	[0.0190, 0.057]
	(0.0041, 0.163)	(0.0041, 0.048)	(0.0104, 0.048)	(0.0041, 0.093)	(0.0098, 0.093)
06nov	[0.0055, 0.189]	[0.0055, 0.056]	[0.0213, 0.056]	[0.0055, 0.069]	[0.0230, 0.069]
	(0.0054, 0.191)	(0.0054, 0.067)	(0.0180, 0.067)	(0.0054, 0.090)	(0.0170, 0.090)
13nov	[0.0059, 0.181]	[0.0059, 0.060]	[0.0264, 0.060]	[0.0059, 0.115]	[0.0450, 0.115]
	(0.0058, 0.183)	(0.0058, 0.076)	(0.0216, 0.076)	(0.0058, 0.178)	(0.0312, 0.178)
20nov	[0.0052, 0.196]	[0.0052, 0.080]	[0.0352, 0.080]	[0.0052, 0.160]	[0.0519, 0.160]
	(0.0052, 0.198)	(0.0052, 0.099)	(0.0287, 0.099)	(0.0052, 0.197)	(0.0322, 0.197)
27nov	[0.0066, 0.225]	[0.0066, 0.062]	[0.0306, 0.062]	[0.0066, 0.100]	[0.0457, 0.100]
	(0.0066, 0.227)	(0.0066, 0.075)	(0.0253, 0.075)	(0.0066, 0.149)	(0.0332, 0.149)
04dec	[0.0047, 0.213]	[0.0047, 0.048]	[0.0243, 0.048]	[0.0047, 0.060]	[0.0299, 0.060]
	(0.0046, 0.215)	(0.0046, 0.056)	(0.0206, 0.056)	(0.0046, 0.078)	(0.0214, 0.078)
11dec	[0.0010, 0.195]	[0.0010, 0.056]	[0.0277, 0.056]	[0.0010, 0.086]	[0.0480, 0.086]
	(0.0010, 0.199)	(0.0010, 0.084)	(0.0143, 0.084)	(0.0010, 0.177)	(0.0117, 0.177)

Notes: Table reports weekly bounds on COVID prevalence in the indicated sample under test monotonicity, as well as the indicated representativeness assumption. Hospitalized samples weighted to match the population age distribution. See Table A.1 for sample definitions. Bounds are in brackets, 95% confidence intervals in parentheses.

Table A.5: Weekly bounds on prevalence under test monotonicity, by sample, not age-weighted, March-July

Sample	Pop	Non-ICLI		Clear cause	
Representatives assumption		Hosp-M	Hosp-I	Hosp-M	Hosp-I
Week	(1)	(2)	(3)	(4)	(5)
13mar	[0.0001, 0.111]	[0.0001, 0.090]	[0.0034, 0.090]	[0.0001, 0.073]	[0.0022, 0.073]
	(0.0001, 0.120)	(0.0001, 0.113)	(0.0020, 0.113)	(0.0001, 0.114)	(0.0001, 0.114)
20mar	[0.0003, 0.181]	[0.0003, 0.159]	[0.0189, 0.159]	[0.0003, 0.174]	[0.0149, 0.174]
	(0.0003, 0.187)	(0.0003, 0.184)	(0.0155, 0.184)	(0.0003, 0.186)	(0.0089, 0.186)
27mar	[0.0006, 0.197]	[0.0006, 0.197]	[0.0390, 0.197]	[0.0006, 0.147]	[0.0210, 0.147]
	(0.0005, 0.202)	(0.0005, 0.201)	(0.0338, 0.201)	(0.0005, 0.196)	(0.0128, 0.196)
03apr	[0.0006, 0.178]	[0.0006, 0.162]	[0.0331, 0.162]	[0.0006, 0.163]	[0.0259, 0.163]
	(0.0006, 0.183)	(0.0006, 0.180)	(0.0285, 0.180)	(0.0006, 0.182)	(0.0170, 0.182)
10apr	[0.0006, 0.168]	[0.0006, 0.134]	[0.0300, 0.134]	[0.0006, 0.106]	[0.0184, 0.106]
	(0.0006, 0.173)	(0.0006, 0.156)	(0.0257, 0.156)	(0.0006, 0.149)	(0.0109, 0.149)
17apr	[0.0009, 0.188]	[0.0009, 0.085]	[0.0194, 0.085]	[0.0009, 0.058]	[0.0126, 0.058]
	(0.0008, 0.191)	(0.0008, 0.101)	(0.0154, 0.101)	(0.0008, 0.088)	(0.0070, 0.088)
24apr	[0.0008, 0.147]	[0.0008, 0.110]	[0.0249, 0.110]	[0.0008, 0.086]	[0.0154, 0.086]
	(0.0008, 0.150)	(0.0008, 0.129)	(0.0208, 0.129)	(0.0008, 0.123)	(0.0098, 0.123)
01may	[0.0009, 0.126]	[0.0009, 0.058]	[0.0190, 0.058]	[0.0009, 0.057]	[0.0187, 0.057]
	(0.0009, 0.129)	(0.0009, 0.069)	(0.0152, 0.069)	(0.0009, 0.078)	(0.0116, 0.078)
08may	[0.0008, 0.098]	[0.0008, 0.039]	[0.0137, 0.039]	[0.0008, 0.047]	[0.0154, 0.047]
	(0.0008, 0.100)	(0.0008, 0.048)	(0.0109, 0.048)	(0.0008, 0.067)	(0.0093, 0.067)
15may	[0.0008, 0.094]	[0.0008, 0.045]	[0.0164, 0.045]	[0.0008, 0.057]	[0.0190, 0.057]
	(0.0008, 0.097)	(0.0008, 0.053)	(0.0136, 0.053)	(0.0008, 0.078)	(0.0129, 0.078)
22may	[0.0006, 0.084]	[0.0006, 0.047]	[0.0162, 0.047]	[0.0006, 0.053]	[0.0164, 0.053]
	(0.0006, 0.086)	(0.0006, 0.055)	(0.0132, 0.055)	(0.0006, 0.074)	(0.0107, 0.074)
29may	[0.0006, 0.069]	[0.0006, 0.029]	[0.0102, 0.029]	[0.0006, 0.027]	[0.0090, 0.027]
	(0.0006, 0.071)	(0.0006, 0.036)	(0.0079, 0.036)	(0.0006, 0.042)	(0.0050, 0.042)
05jun	[0.0006, 0.056]	[0.0006, 0.026]	[0.0092, 0.026]	[0.0006, 0.028]	[0.0087, 0.028]
	(0.0006, 0.058)	(0.0006, 0.032)	(0.0072, 0.032)	(0.0006, 0.042)	(0.0047, 0.042)
12jun	[0.0005, 0.041]	[0.0005, 0.018]	[0.0064, 0.018]	[0.0005, 0.023]	[0.0070, 0.023]
	(0.0005, 0.042)	(0.0005, 0.024)	(0.0046, 0.024)	(0.0005, 0.036)	(0.0035, 0.036)
19jun	[0.0006, 0.049]	[0.0006, 0.017]	[0.0062, 0.017]	[0.0006, 0.015]	[0.0049, 0.015]
	(0.0006, 0.050)	(0.0006, 0.022)	(0.0044, 0.022)	(0.0006, 0.027)	(0.0019, 0.027)
26jun	[0.0006, 0.058]	[0.0006, 0.020]	[0.0070, 0.020]	[0.0006, 0.031]	[0.0103, 0.031]
	(0.0006, 0.060)	(0.0006, 0.025)	(0.0051, 0.025)	(0.0006, 0.046)	(0.0058, 0.046)
03jul	[0.0007, 0.067]	[0.0007, 0.015]	[0.0052, 0.015]	[0.0007, 0.026]	[0.0084, 0.026]
	(0.0007, 0.069)	(0.0007, 0.020)	(0.0037, 0.020)	(0.0007, 0.041)	(0.0037, 0.041)
10jul	[0.0009, 0.067]	[0.0009, 0.020]	[0.0074, 0.020]	[0.0009, 0.024]	[0.0082, 0.024]
	(0.0009, 0.069)	(0.0009, 0.026)	(0.0056, 0.026)	(0.0009, 0.037)	(0.0041, 0.037)
17jul	[0.0010, 0.069]	[0.0010, 0.017]	[0.0061, 0.017]	[0.0010, 0.020]	[0.0066, 0.020]
	(0.0010, 0.071)	(0.0010, 0.022)	(0.0044, 0.022)	(0.0010, 0.032)	(0.0030, 0.032)
24jul	[0.0010, 0.072]	[0.0010, 0.027]	[0.0093, 0.027]	[0.0010, 0.039]	[0.0119, 0.039]
	(0.0010, 0.073)	(0.0010, 0.034)	(0.0073, 0.034)	(0.0010, 0.057)	(0.0072, 0.057)
31jul	[0.0011, 0.071]	[0.0011, 0.028]	[0.0100, 0.028]	[0.0011, 0.035]	[0.0110, 0.035]
	(0.0010, 0.073)	(0.0010, 0.034)	(0.0078, 0.034)	(0.0010, 0.049)	(0.0063, 0.049)

Notes: Table reports weekly bounds on COVID prevalence in the indicated sample under test monotonicity, as well as the indicated representativeness assumption. See Table A.1 for sample definitions. Bounds are in brackets, 95% confidence intervals in parentheses.

Table A.6: Weekly bounds on prevalence under test monotonicity, by sample, not age-weighted, August-December

Sample	Pop	Non-ICLI		Clear cause	
Representatives assumption		Hosp-M	Hosp-I	Hosp-M	Hosp-I
Week	(1)	(2)	(3)	(4)	(5)
07aug	[0.0011, 0.067]	[0.0011, 0.023]	[0.0079, 0.023]	[0.0011, 0.030]	[0.0098, 0.030]
	(0.0010, 0.069)	(0.0010, 0.029)	(0.0059, 0.029)	(0.0010, 0.046)	(0.0054, 0.046)
14aug	[0.0010, 0.061]	[0.0010, 0.023]	[0.0077, 0.023]	[0.0010, 0.047]	[0.0137, 0.047]
	(0.0010, 0.062)	(0.0010, 0.029)	(0.0058, 0.029)	(0.0010, 0.061)	(0.0087, 0.061)
21aug	[0.0010, 0.064]	[0.0010, 0.022]	[0.0071, 0.022]	[0.0010, 0.031]	[0.0090, 0.031]
	(0.0010, 0.065)	(0.0010, 0.027)	(0.0054, 0.027)	(0.0010, 0.046)	(0.0054, 0.046)
28aug	[0.0010, 0.066]	[0.0010, 0.022]	[0.0069, 0.022]	[0.0010, 0.021]	[0.0058, 0.021]
	(0.0010, 0.068)	(0.0010, 0.028)	(0.0051, 0.028)	(0.0010, 0.033)	(0.0024, 0.033)
04sep	[0.0009, 0.060]	[0.0009, 0.023]	[0.0075, 0.023]	[0.0009, 0.025]	[0.0073, 0.025]
	(0.0009, 0.061)	(0.0009, 0.030)	(0.0058, 0.030)	(0.0009, 0.040)	(0.0035, 0.040)
11sep	[0.0009, 0.055]	[0.0009, 0.019]	[0.0064, 0.019]	[0.0009, 0.030]	[0.0093, 0.030]
	(0.0009, 0.056)	(0.0009, 0.023)	(0.0048, 0.023)	(0.0009, 0.046)	(0.0052, 0.046)
18sep	[0.0009, 0.059]	[0.0009, 0.021]	[0.0071, 0.021]	[0.0009, 0.032]	[0.0093, 0.032]
	(0.0009, 0.061)	(0.0009, 0.027)	(0.0052, 0.027)	(0.0009, 0.049)	(0.0047, 0.049)
25sep	[0.0012, 0.072]	[0.0012, 0.024]	[0.0082, 0.024]	[0.0012, 0.031]	[0.0095, 0.031]
	(0.0011, 0.073)	(0.0011, 0.031)	(0.0062, 0.031)	(0.0011, 0.048)	(0.0052, 0.048)
02oct	[0.0015, 0.082]	[0.0015, 0.026]	[0.0090, 0.026]	[0.0015, 0.032]	[0.0099, 0.032]
	(0.0014, 0.084)	(0.0014, 0.033)	(0.0071, 0.033)	(0.0014, 0.047)	(0.0053, 0.047)
09oct	[0.0017, 0.102]	[0.0017, 0.034]	[0.0115, 0.034]	[0.0017, 0.051]	[0.0154, 0.051]
	(0.0016, 0.104)	(0.0016, 0.042)	(0.0094, 0.042)	(0.0016, 0.070)	(0.0095, 0.070)
16oct	[0.0024, 0.111]	[0.0024, 0.038]	[0.0140, 0.038]	[0.0024, 0.046]	[0.0150, 0.046]
	(0.0024, 0.113)	(0.0024, 0.046)	(0.0115, 0.046)	(0.0024, 0.064)	(0.0097, 0.064)
23oct	[0.0032, 0.134]	[0.0032, 0.047]	[0.0181, 0.047]	[0.0032, 0.069]	[0.0237, 0.069]
	(0.0031, 0.135)	(0.0031, 0.055)	(0.0148, 0.055)	(0.0031, 0.091)	(0.0165, 0.091)
30oct	[0.0042, 0.164]	[0.0042, 0.043]	[0.0161, 0.043]	[0.0042, 0.067]	[0.0217, 0.067]
	(0.0042, 0.165)	(0.0042, 0.052)	(0.0132, 0.052)	(0.0042, 0.088)	(0.0145, 0.088)
06nov	[0.0055, 0.191]	[0.0055, 0.071]	[0.0288, 0.071]	[0.0055, 0.105]	[0.0358, 0.105]
	(0.0055, 0.193)	(0.0055, 0.080)	(0.0253, 0.080)	(0.0055, 0.132)	(0.0272, 0.132)
13nov	[0.0059, 0.184]	[0.0059, 0.066]	[0.0317, 0.066]	[0.0059, 0.114]	[0.0532, 0.114]
	(0.0058, 0.186)	(0.0058, 0.074)	(0.0278, 0.074)	(0.0058, 0.139)	(0.0436, 0.139)
20nov	[0.0053, 0.198]	[0.0053, 0.093]	[0.0442, 0.093]	[0.0053, 0.116]	[0.0534, 0.116]
	(0.0052, 0.200)	(0.0052, 0.103)	(0.0387, 0.103)	(0.0052, 0.139)	(0.0420, 0.139)
27nov	[0.0067, 0.227]	[0.0067, 0.084]	[0.0426, 0.084]	[0.0067, 0.142]	[0.0703, 0.142]
	(0.0066, 0.229)	(0.0066, 0.093)	(0.0375, 0.093)	(0.0066, 0.169)	(0.0577, 0.169)
04dec	[0.0047, 0.214]	[0.0047, 0.068]	[0.0344, 0.068]	[0.0047, 0.108]	[0.0534, 0.108]
	(0.0047, 0.216)	(0.0047, 0.077)	(0.0298, 0.077)	(0.0047, 0.133)	(0.0412, 0.133)
11dec	[0.0010, 0.196]	[0.0010, 0.051]	[0.0228, 0.051]	[0.0010, 0.085]	[0.0399, 0.085]
	(0.0010, 0.201)	(0.0010, 0.064)	(0.0177, 0.064)	(0.0010, 0.120)	(0.0219, 0.120)

Notes: Table reports weekly bounds on COVID prevalence in the indicated sample under test monotonicity, as well as the indicated representativeness assumption. See Table A.1 for sample definitions. Bounds are in brackets, 95% confidence intervals in parentheses.

Table A.7: Demographics and test rates among hospitalized patients, by group

Group	Number of Admissions	Female	Age						
			Newborn	0-17	18-29	30-49	50-64	65-74	>74
All	781,587	0.555	0.080	0.028	0.105	0.181	0.228	0.182	0.196
Has diagnosis	355,425	0.557	0.100	0.026	0.112	0.178	0.214	0.173	0.198
ICLI	49,904	0.493	0.005	0.023	0.037	0.139	0.269	0.239	0.287
Non-ICLI	305,521	0.568	0.115	0.027	0.124	0.184	0.205	0.162	0.183
Clear cause	61,682	0.592	0.003	0.033	0.165	0.175	0.195	0.179	0.249
Cancer	9,585	0.465	0.001	0.053	0.026	0.122	0.322	0.284	0.192
Labor/delivery	13,304	0.995	0.009	0.023	0.611	0.357	0.000	0.000	0.000
AMI	8,624	0.405	0.000	0.000	0.007	0.112	0.315	0.265	0.301
Stroke	8,297	0.487	0.001	0.004	0.011	0.092	0.269	0.256	0.368
Fracture	13,718	0.546	0.003	0.034	0.063	0.128	0.178	0.187	0.408
Open wound	3,642	0.420	0.002	0.047	0.097	0.197	0.224	0.167	0.266
Appendicitis	1,961	0.465	0.000	0.224	0.199	0.274	0.181	0.084	0.038
Vehicle accident	1,944	0.356	0.001	0.090	0.216	0.297	0.195	0.119	0.082
Other accident	9,782	0.541	0.003	0.033	0.034	0.082	0.154	0.208	0.486

Notes: Table reports the number and age distribution of admissions, for different categories of admissions, over the time period March 13, 2020 through June 18, 2020. See Appendix B for definitions of the different causes of admissions (Cancer-other accident).

B Defining causes of admissions

This section provides more details on our definition of ICLI, non-ICIL, and "clear cause" hospitalization, listing the ICD-10 codes used to define each.

Following Armed Forces Health Surveillance Center (2015), the codes for influenza-like illness are B97.89, H66.9, H66.90, H66.91 H66.92, H66.93, J00, J01.9, J01.90, J06.9, J09, J09.X, J09.X1, J09.2, J09.X3, J09.X9, J10, J10.0, J10.00, J10.01, J10.08, J10.1, J10.2, J10.8, J10.81, J10.82, J10.83, J10.89, J11, J11.0, J11.00, J11.08, J11.1, J11.2, J11.8, J11.81, J11.82, J11.83, J11.89, J12.89, J12.9, J18, J18.1, J18.8, J18.9, J20.9, J40, R05, and R50.9. We say a hospitalizaiton is for an influenza-like illness if it has any of these diagnosis codes in any position. We say a hospitalization is for a COVID-like illness if it has any ICD-10 code among those that is among the CDC's lists of diagnosis codes for COVID-19 Center for Disease Control and Prevention (2020). These codes are J12.89, J20.8, J22, J40, J80, J98.8, O95.5, R05, R06.02, R50.9, U07.1, Z03.818, Z11.58, and Z20.828.

We define ICLI-related hospitalizations as ones with at least one ILI or CLI diagnosis code. We define non-ICLI related hospitalizations as hospitalized with diagnosis codes, but no ILI or CLI code.

We also define "clear cause" hospitalizations. These are hospitalizations for labor and delivery, AMI, stroke, fractures and crushes, wounds, vehicle accidents, other accidents, appendicitis, or cancer. With the exception of cancer, we define a hospitalization as belonging to one of these groups if it has any diagnosis codes for that group, listed below. Cancer is treated differently because it can be a comorbidity. We say a hospitalization is for cancer if a cancer diagnosis (listed below) is an admitting diagnosis, the primary final

diagnosis, or if chemotherapy diagnosis is present. We use the following ICD-10 codes.

- **AMI** I21, I22.
- **Appendicitis** K35-K38.
- **Cancer** C00-C97 (in primary or admitting diagnosis), or Z51.0-Z51.2 (in any position).
- **Fracture/Crush** S02, S12, S22, S32, S42, S52, S62, S72, S82, S92, T02, S07, S17, S37, S47, S57,S67, S77, S87, S97, T07.
- **Labor and delivery** O60-O75, O80-O84.
- **Other accidents** W00-W99, X00-X59.
- **Stroke** I61-I64.
- **Vehicle accident** V01-V99.
- **Wound** S01, S11, S21, S31, S41, S51, S61, S71, S81, S91, T01.

C Calculating negative predictive values with test-retest data

Setup and identification Here we show how to use data on multiple tests to simultaneously identify prevalence and test error rates, and how to use this information to obtain the negative predictive value (NPV) of at test under a narrow set of assumptions. Assume in particular that people have been tested exactly twice, with $R1_i$ the outcome of the first test and $R2_i$ the outcome of the second test for person i. Let C_i be person i's true infection status, which we assume is fixed between the tests. Let $p = Pr(C_i = 1)$ be the prevalence of active SARS-CoV-2 infections in this twice-tested population.

Test outcomes may differ from true infection status because of test errors. In general, therefore, there are four possible sequences of test outcomes: $(0,0), (0,1), (1,0), (1,1)$. We let $P_{ab} = Pr(R1_i = a, R2_i = b)$ for $(a,b) \in \{0,1\}^2$.

We make three strong assumptions to simplify the analysis.

Assumption 4. *The specificity of the test is 1. That is,* $\beta = Pr(Rj_i = 0|C_i = 0) = 1$.

Assumption 5. *The sensitivity of the test,* $\alpha = Pr(Rj_i = 1|C_i = 1)$*, does not depend on the initial test result.*

Assumption 6. *Retesting is random, i.e. independent of* $R1_i$ *and* C_i.

Assumption 4 is the weakest of these assumptions. It implies that there are no false positives, which is consistent with typical practice (UCSF Health Hospital Epidemiology and Infection Prevention, 2020). The remaining assumptions are stronger. Assumption 5 says that the test errors are independent of the initial test result. It would be violated,

for example, if false negatives are more common for patients with high levels of mucus, and mucus levels are correlated across test results. Assumption 6 says that retesting rates do not depend on possible testing errors. We would expect this condition to fail if highly symptomatic people with negative tests are especially likely to test negative. We view this assumption as the most suspect.

Under these assumptions, the test outcome probabilities P_{ab} simplify considerably. Since the probabilities sum to one, and the assumptions imply that $P_{10} = P_{01}$, the only non-redundant probabilities are:

$$P_{00} = (1-p) + p(1-\alpha)^2$$

$$P_{11} = p\alpha^2.$$

We can observe P_{00} and P_{11}. Solving for the unknowns p and α, we have

$$p = \frac{(P_{00} - P_{11} - 1)^2}{4P_{11}}$$

$$\alpha = \frac{2P_{11}}{1 - P_{00} + P_{11}}$$

This shows how to get p and α from two tests, and the assumption that specificity (β) equals 1. Our goal is to find the negative predictive value (NPV), which can be computed given knowledge of α, β and p. In general, for a single test $NPV = Pr(C_i = 0 | R_i = 0)$. Applying Bayes rule shows that:

$$NPV = \frac{1-p}{p(1-\alpha)+(1-p)}$$

Results To implement this approach, we construct a sample of all people who are tested on a given day, not tested the previous day, and then tested again in the next day. There are 835,195 such test pairs. We find $P_{00} = 0.884$ and $P_{11} = 0.113$. Nearly all the mass is on the diagonals; test results switch less than 1% of the time. This fact, together with the assumption that specificity is equal to 1, implies very low false negative rates. Plugging these values into our formula, we have $p = 0.116$ and $\alpha = 0.987$, which implies $NPV = 0.998$. Using instead, all people who are retested once within a three day period, we find similar results: $p = 0.118, \alpha = 0.972, NPV = 0.996$.

We emphasize that these estimates are valid for the twice-tested population and under assumptions 4-6, in particular, random retesting. The prevalence estimate is the prevalence among people tested twice, not the population prevalence. And it is only a valid estimate under assumptions 1-3. In reality, it is likely that retests are most common among suspected false negatives (i.e. when a highly symptomatic patient tests negative). We see some evidence for this: $P_{01} = 0.0013$ and $P_{10} = 0.0016$, a slight but significant difference implying that negative-then-positive is slightly more common than positive-than-negative, inconsistent with the random retesting assumption. We therefore do not view our estimates of prevalence and sensitivity as definitive; rather we think of the sensitivity estimate as a lower bound on sensitivity, because we have selected a retest sample which has a disproportionate number of false negatives. As NPV is increasing in sensitivity, α, our implied estimate of $1 - NPV$ is likely an upper bound on $1 - NPV$.

D Measurement Error In Testing

Virological tests for the presence of SARS-CoV-2 may not be perfectly accurate, and so far there are no detailed studies of the performance of the PCR tests that Indiana is using to test people for SARS-CoV-2. To clarify how error-ridden tests complicate our prevalence estimates, we augment the notation to distinguish between test results and virological status. We continue to use C_{it} and D_it to represent a person's true infection and testing status at date t. But now we introduce R_{it}, which is a binary measure set to 1 if the person tests positive and 0 if the person tests negative. Using this notation, $Pr(C_{it} = 1|D_{it} = 1, R_{it} = 1)$ is called the Positive Predictive Value (PPV) of the test among people who are tested and who test positive. $Pr(C_{it} = 0|D_{it} = 1, R_{it} = 0)$ is called the Negative Predictive Value (NPV) among people who are tested and who test negative. $1 - NPV = Pr(C_{it} = 1|D_{it} = 1, R_{it} = 0)$ is the fraction of people who test negative who are actually infected with SARS-CoV-2.

Our initial worst case bounds assumed no test errors. Relaxing that assumption yields a different set of upper and lower bounds on prevalence. Following Manski and Molinari (2020), we assume that (i) $PPV = 1$ so that none of the positive tests are false, but (ii) $Pr(C_{it} = 1|D_{it} = 1, R_{it} = 0) \in [\lambda_l, \lambda_u]$. The second condition imposes a bound on $1-NPV$, which is the fraction of people who test negative who are actually infected. Under these two restrictions, the new worst case bounds work out to:

$$L_{w,\lambda} = L_w + \lambda_l Pr(R_{it} = 0|D_{it} = 1)Pr(D_{it} = 1)$$

$$U_{w,\lambda} = U_w + \lambda_u Pr(R_{it} = 0|D_{it} = 1)Pr(D_{it} = 1)$$

Allowing for test errors increases the worst case lower bound by the best-case fraction of missing positives, and increases the worst case upper bound by the worst-case fraction of missing positives. Similar expressions hold for prevalence bounds under test monotonicity and other independence assumptions.

The upshot is that knowledge of test accuracy is important for efforts to learn about prevalence. In their study of the cumulative prevalence of SARS-CoV-2 infections, Manski and Molinari (2020) computed upper and lower bounds on prevalence under the assumption that $\lambda_l = .1$ and $\lambda_u = .4$, citing Peci et al. (2014). Manski and Molinari (2020) view this choice of $.1 \leq 1 - NPV \leq .4$ as an expression of scientific uncertainty about test errors, and they refer to the resulting prevalence bounds as "illustrative." However, the structure of the test error bounds makes it clear that assumptions about the numerical magnitude of test errors have inferential consequences. For example, setting $\lambda_u = .4$ implies that, regardless of the outcome of the test, at least 40 percent of the people who are tested for SARS-CoV-2 are infected.

Although there is little published evidence on the properties of the SARS-CoV-2 PCR test, previous research suggests that PCR test errors are uncommon in other settings. For example, Peci et al. (2014) study the performance of rapid influenza tests using PCR-based tests as a *gold standard*. PCR tests are used as a gold standard because they are expected to have very high PPV and NPV.

To shed more light on test errors, we constructed a sample of people who are tested and retested in a short interval, specifically people who were (i) tested on day t, (ii) not tested on day $t-1$, and (iii) were tested again on day $t+1$. We show in Appendix C how these data can be used to estimate error rates, under assumptions of random retesting

and no false positives. Our data include 835,000 test-retest events. Using $R1_i$ and $R2_i$ to represent the results of a person's first and second test, we found that $Pr(R1_i = 1, R2_i = 1) = .11$ and $Pr(R1_i = 0, R2_i = 0) = .88$ among the people in the twice-tested sample. The two tests were discordant for less than 1 percent of the twice-tested sample. These results imply a negative predictive value of 99.8 percent.

This estimate of NPV depends on our assumptions of random retesting and no false positives. While the no false positive assumption appears plausible, random retesting is not necessarily satisfied. In particular, a patient with a suspected COVID case who initially tests negative may be retested; this selective retesting would bias us towards finding false negatives. Another reason for retesting is delays in processing results. If a patient was tested prior to a planned hospitalization, and the result is not available at the time of the hospitalization, the attending physician may order an in-hospital test, which would be available within hours. This type of retesting is less likely to lead to bias. As we explain in Appendix C, we can test for selection into retesting by looking for symmetry in test results. Under random retesting (and no false positives), the sequences "positive-then-negative" and "negative-then-positive" should be equally likely. In practice we find that "negative-then-positive" is slightly more common, meaning that our test-retest sample likely disproportionately selects people with initial false negatives.

Overall, we think that a plausible value for λ_l is nearly zero, and a plausible value for λ_u is 0.005. Accounting for test errors in this range would have almost no effect on the upper and lower bounds reported in the paper. Test-retest data are potentially informative about test errors, but a limitation of is that retested people are not necessarily representative of the population.

E Small bias from excluding ICLI-hospitalizations

Our main sample uses non-ICLI hospitalizations to bound COVID prevalence in the general population. This approach therefore yields bounds on the prevalence of non-severe COVID-19, where "non-severe" means "not severe enough to induce a COVID-related hospitalization." These bounds are of course biased for bounds on overall COVID-19 prevalence. However this bias is quite small, small enough that it is unlikely to be decision relevant. We show this in two separate arguments.

To begin we abuse notation slightly and let H in this section be an indicator for an ICLI-related hospitalization, rather than any hospitalization. Both arguments start from the observation that COVID prevalence is equal to COVID prevalence among the hospitalized population plus its prevalence among the unhospitalized population:

$$Pr(C = 1) = Pr(C = 1, H = 1) + Pr(C = 1, H = 0).$$

Since our main sample is limited to non-ICLI hospitalizations, our bounds can be interpreted as bounds on $Pr(C = 1, H = 0)$, and the bias is (at most) the bias from omitting $Pr(C = 1, H = 1)$.

E.1 Argument from rarity of ICLI-related hospitalizations

Our first argument that this bias is small is to observe that $Pr(C = 1, H = 1) \leq Pr(H = 1)$. That is, the overall rate of ICLI-related hospitalizations in the population is an upper bound on the fraction of people in the population who are COVID-19 positive

and have an ICLI-related hospitalization. Fortunately, $Pr(H = 1)$ is nearly observable in our data.

In particular, we don't quite observe $Pr(H = 1)$ because not all hospitals report diagnosis information. We can therefore bound $Pr(H = 1)$ by assuming that when diagnosis information is not reported, $H = 1$. Taking this approach, Figure E.1 shows $Pr(H = 1)$ in our data. This is the weekly count of ICLI-related hospitalizations, scaled by the population of Indiana. An alternative approach, also shown in Figure E.1 is to measure ICLI-related hospitalizations as the total number of hospitalizations, scaled by the share of ICLI-related hospitalizations among hospitalizations with diagnoses. We see that $Pr(H = 1)$ is always less than 0.3%, typically less than 0.2%, using the more conservative bound. Thus the population prevalence of COVID-19 exceeds our upper bound by at most 0.3%. A more precise estimate of the bias uses the estimated $Pr(H = 1)$ from observed diagnoses, about 0.05 percent, and uses Figure 3 to infer that $Pr(C = 1|H = 1)$ is typically less than 50 percent, and so the bias from excluding ICLI related hospitalizations is likely less than 0.025 percent, that is, about 1700 cases out of a population of 6.8 million. Reassuringly, this number is similar to the average reported COVID-19 hospitalizations in the state of Indiana in 2020 (Indiana State Department of Health, 2020).[13]

E.2 Argument from low infection hospitalization rate

A second argument shows, similarly, that there is little bias from conditioning on ICLI-unrelated hospitalizations. This argument is based on the fact that the infection hospital-

[13] Because hospitalizations last a few days, our weekly admission count is comparable to the state's daily count of then umber of people in the hospital.

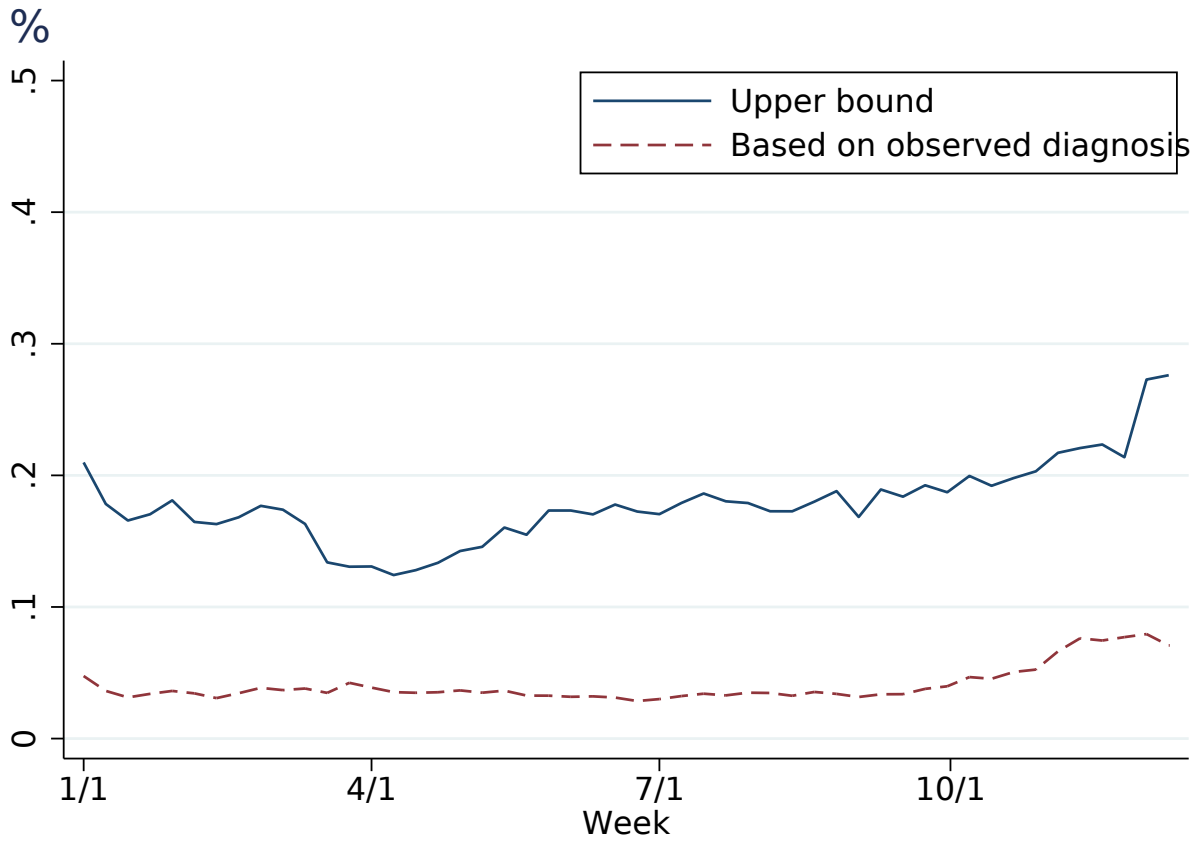
ization rate, $Pr(H = 1|C = 1)$, is known to be low.

After substituting $Pr(H = 1|C = 1)Pr(C = 1)$ for $Pr(C = 1, H = 1)$ in the equation above, and a bit of algebra, we have

$$Pr(C = 1)\frac{Pr(C = 1, H = 0)}{1 - Pr(H = 1|C = 1)}.$$

By focusing on ICLI-unrelated hospitalizations, we bound the numerator. The expression above shows that our bound is off by a factor of at most $(1 - Pr(H = 1|C = 1))^{-1}$. If $Pr(H = 1|C = 1)$ were known to be low, then the bias in our bound would be low as well. The available evidence indicates that the infection hospitalization rate – $Pr(H = 1|C = 1)$ – is small, not more than 10 percent in unvaccinated populations, and likely smaller. Menachemi et al. (2021) estimate 2.1 percent in Indiana (excluding nursing homes) and Mahajan et al. (2021) estimate 7 percent in Connecticut, both using random sample testing to establish population prevalence and treating the number of hospitalizations as known. Salje et al. (2020) estimate 2.9 percent, using a model-driven approach. All estimates imply that our upper bound is too low by, at most, 7.5 (1/.93) percent (we emphasize: percent, not percentage point). As our upper bound is usually less than 5 percent, we are left with a bias of, at most, .4 percentage points.

Figure E.1: Estimate ICLI-related hospitalizations as a share of the population, by week



Notes: Figure plots, for each week, two estimates of the share of the population of Indiana admitted for an ICLI-related hospitalization. Not all hospitals report diagnosis information, so the upper bound assumes hospitalizations are ICLI-related if the diagnosis information is unreported. The "based on observed diagnosis" line assumes that the share of ICLI-related hospitalizations among the hospitals with missing diagnosis information is equal to their share among the hospitals with reported information.

F Inference and age adjustment details

F.1 Inference for Intersection Bounds

The sample analogue estimators we use to construct the test monotonicity, hospital monotonicity, and hospital independence bounds reported in the paper are all asymptotically consistent. However, the hospital monotonicity and hospital independence bounds are examples of "intersection bounds". The sample analogue estimators are asymptotically consistent but their sampling distribution is somewhat complicated and the point estimates may include finite sample bias because the minimum and maximum operators are non-linear.

To understand the finite sample bias of the intersection bounds, consider the upper bound on population prevalence under under test monotonicity and hospital monotonicity:

$$\begin{aligned} U_{mH} &= \min\left\{Pr(C=1|D=1), Pr(C=1|D=1,H=1)\right\} \\ &= \min\left\{\text{Population test positivity}, \text{Hospitalized test positivity}\right\}. \end{aligned}$$

We estimate this bound by using the sample analogs of $Pr(C=1|D=1)$ and $Pr(C=1|D=1,H=1)$, say $\hat{P}(C=1|D=1)$ and $\hat{P}(C=1|D=1,H=1)$. Because the minimum operator is not linear, $E[U_{mH}]$ is not equal to the minimum of the two expectations. Suppose for illustration that, in the population, $Pr(C=1|D=1,H=1) < Pr(C=1|D=1)$, so that the hospital test positivity binds. In that case, finite sample bias may arise because in any given random sample, there is a positive probability that $\hat{P}(C=1|D=1) < \hat{P}(C=$

$1|D = 1, H = 1)$. But if this probability is small, then so is the bias.

We using the bootstrap method described in Manski and Pepper (2009); Kreider and Pepper (2007) to estimate confidence intervals for the test monotonicity, hospital monotonicity, and hospital independence bounds and to assess concerns about finite sample bias in the hospital monotonicity and hospital independence bounds.

We use 500 bootstrap simulations. In each bootstrap replication we formed each set of bounds. We used percentiles of the bootstrap distribution of the upper and lower bounds to form a 95 percent confidence interval around the identified set. The lower bound of the 95 percent confidence interval is the 2.5th percentile of the bootstrapped lower bounds, and the upper bound of the 95 percent confidence interval is the 97.5th percentile of the upper bounds.

We also used the bootstrap to estimate the degree of finite sample bias associated with the hospital monotonicity and hospital independence bounds. We estimate the finite sample bias as the difference between the average estimate in the bootstrap sample and the actual point estimate in the full sample. Table F.1 shows bootstrap estimates of the bias in the upper bound under test monotonicity and hospitalization monotonicity applied to the non-ICLI hospitalized population. The bootstrap results suggest that the finite sample bias is negligible in our application. The estimated bias is less than 1 percent (not percentage point) in most weeks. It makes sense that the bias is small because the sample size in our analysis is very large, and also because there is such a large gap between population test positivity and hospitalized test positivity. The results is that there is a low probability across bootstraps that $\hat{P}(C = 1|D = 1) < \hat{P}(C = 1|D = 1, H = 1)$. We show this probability week-by-week in Figure F.1. In the first month of the sample there is a non

trivial probability – 10-30% – that the inequality does not hold in a given random sample. After mid-April, however, this probability becomes essentially zero in every week. Accordingly, finite sample bias is not an important worry in our application and we do not attempt to correct our point estimates or confidence intervals for finite sample bias.

F.2 Age Adjustment

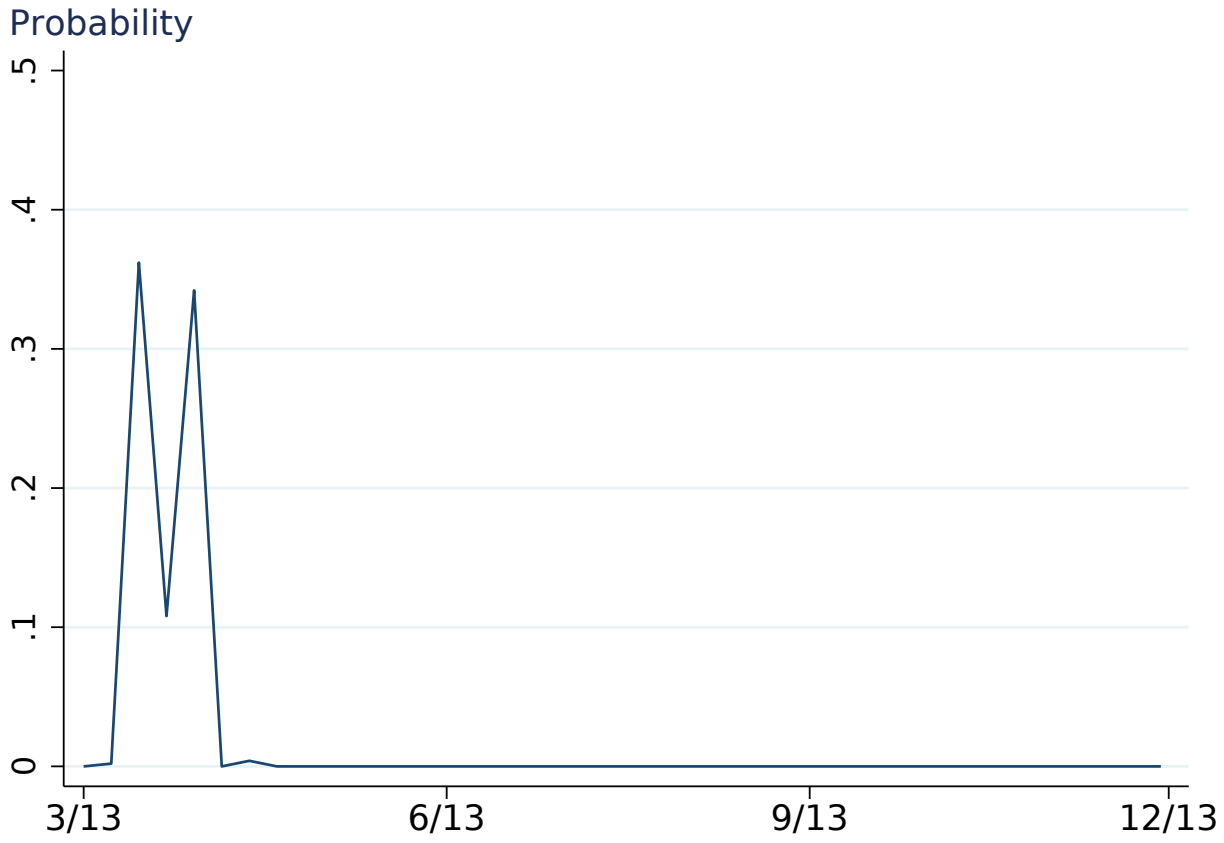
Because the tested and hospitalized samples are not age representative of the general population, throughout the paper, we report both unadjusted results and age-standardized upper and lower bounds. This simply means that we stratify the data six age groups (0-17, 18-30, 30-50, 50-64, 65-74, and 75 and older) and then compute the upper and lower bounds within each age-strata. Afterwards, we average the age group specific bounds by weighting each age-specific bound by that age group's share of the Indiana population. We construct confidence intervals for the age-adjusted bounds using the bootstrap; in each bootstrap iteration we calculate the age-adjusted bound or intersection bound (as appropriate), and our confidence intervals for the bound are the 2.5th percentile of the lower bound confidence interval and the 97.5th percentile of the upper bound confidence interval.

Table F.1: Small bias in intersection bounds

	Upper bound	Bias	Bias/bound
mean	0.044	-0.000	-0.003
min	0.011	-0.005	-0.037
p25	0.020	-0.000	-0.009
p50	0.031	-0.000	-0.002
p75	0.053	0.000	0.004
max	0.159	0.001	0.016

Notes: Table reports statistics on the estimated upper bound (under test monotonicity and hospitalization monotonicity applied to the non-ICLI hospitalization), the bias in the upper bound, and the ratio of the bias to the bound. These statistics vary across weeks in the sample. All bounds are age-adjusted. We estimate the bias as the difference between the average estimate in the bootstrap samples and the actual estimate.

Figure F.1: Estimated probability that $\hat{P}(C = 1|D = 1) < \hat{P}(C = 1|D = 1, H = 1)$, by week



Notes: Figure shows the estimated probability, for each week, that $\hat{P}(C = 1|D = 1) < \hat{P}(C = 1|D = 1, H = 1)$, estimated using a bootstrap. All bounds are age-adjusted.

G Validity Tests

Our main results show that the test monotonicity bounds on prevalence are much tighter for the non-ICLI hospitalized population than for the population as a whole. These tighter bounds are informative for general population prevalence only under additional assumptions about hospital representativeness, either a monotonicity assumption or an equal prevalence assumption. How valid are these assumptions? Assessing them directly is of course impossible because we lack data on prevalence in the population as a whole or in the hospital sample.

Our main analysis provides one type of indirect evidence in support of our hospital representativeness assumptions. The non-ICLI and clear-cause samples generate similar bounds, and, within the clear-cause sample, there are not large differences in bounds across different causes of admission. This suggests that prevalence does not vary with the exact set of hospitalizations studied, although of course this does not prove hospitalization monotonicity or hospitalization independence are credible assumptions.

In this section, we provide two additional pieces of evidence on the hospital IV assumptions. First we show that the hospital bounds are consistent with the estimates of population prevalence from the Indiana COVID-19 Random Sample Study (Menachemi et al., 2020; Richard M. Fairbanks School of Public Health, 2020).[14] Second, we compare the hospital sample to the general population in terms of their likelihood of prior testing (prior to the hospital data) and the test rate of their home counties. We take these to be proxies for their concern about COVID, although other interpretations are possible.

[14] Our data do not contain the test results from the Random Sample Study, so we compare our bounds to the published results.

G.1 Comparison to random sample testing

A valuable benchmark for the hospital-based prevalence bounds comes from a large-scale study of SARS-CoV-2 prevalence in Indiana. The study invited a representative sample of Indiana residents (aged 12 and older) to obtain a SARS-CoV-2 test. The first wave of the study took place April 25-29, and the second wave took place June 3-7. The preliminary results are reported in Menachemi et al. (2020) and Richard M. Fairbanks School of Public Health (2020). The response rate was roughly 25 percent, and no attempt was made to correct for non-random response. Nonetheless this survey appears to be the best benchmark available. We report the point estimates for prevalence (assuming random nonresponse) and their confidence intervals in the top panel of Table G.1. The first wave estimates 1.7 percent prevalence and the second 0.5 percent.[15]

We compare our prevalence bound during the weeks containing the random sample survey, in the bottom panel of the table. Using population testing we obtain very wide bounds that contain the random sample study estimates. This fact provides some support for the test monotonicity assumption. Under our hospital representativeness assumptions, the bounds are tighter, especially in June. Our bounds under hospital monotonicity always contain the random sample study point etimates. Under hospital independence, the point estimate lies slightly below the lower bound. However the 95% confidence interval always overlap. Thus for both dates the prevalence point estimates are consistent with the bounds we obtain under our hospital representativeness assumptions.

[15] The estimates in Table G.1 are slightly different from those reported by Richard M. Fairbanks School of Public Health (2020). We report updated calculations, based on correspondence with the authors.

Table G.1: Do our bounds contain estimates of prevalence from random-sample testing?

Time period	April 25-29	June 3 -7
Random Sample Study		
Prevalence estimates	0.0170	0.005
95% confidence interval	(0.011, 0.025)	(0.002, 0.013)
Bounds from...		
Population testing	[0.0008, 0.137]	[0.0006, 0.059]
	(0.0008, 0.142)	(0.0006, 0.061)
(0.0008, 0.001)	(0.0006, 0.001)	
Non-ICLI hospitalizations (H-M)	[0.0008, 0.086]	[0.0006, 0.022]
	(0.0008, 0.119)	(0.0006, 0.028)
(0.0008, 0.118)	(0.0006, 0.028)	
Non-ICLI hospitalizations (H-I)	[0.0182, 0.086]	[0.0073, 0.022]
	(0.0142, 0.119)	(0.0055, 0.028)
(-0.0551, 0.118)	(-0.0087, 0.028)	
Clear cause hospitalizations (H-M)	[0.0008, 0.057]	[0.0006, 0.021]
	(0.0008, 0.088)	(0.0006, 0.035)
(0.0008, 0.088)	(0.0006, 0.035)	
Cause hospitalizations (H-I)	[0.0108, 0.057]	[0.0064, 0.021]
	(0.0058, 0.088)	(0.0023, 0.035)
(-0.0407, 0.088)	(-0.0117, 0.035)	

Notes: The first two rows of the table report the estimated population prevalence and 95% confidence interval from the Indiana COVID-19 Random Sample Study, conducted over the indicated dates, which assumes random nonresponse (Menachemi et al., 2020; Richard M. Fairbanks School of Public Health, 2020). The remaining rows report the (age-adjusted) bounds on prevalence, in brackets, with 95-percent confidence intervals, in parentheses, from our different samples, under test monontonicity, as well as hospital monotonicity (H-M) or hospital independnence (H-I) as indicated, for the week containing the random sample study period.

G.2 Comparison of prior testing and community testing

A standard way of measuring representativeness is to compare the distribution of covariates in a study population to their distribution in the target population. In our case, this approach is most convincing if we have well-measured covariates that proxy for SARS-CoV-2 infection risk. Two candidate covariates are the community SARS-CoV-2 testing rate and the prior testing rate. The idea behind these proxies is that people who come from areas with high test rates, or who have been tested in the past, may themselves have a higher current likelihood of being infected with the virus.

To operationalize these measures, we define the community testing rate for person i as the fraction of people in i's county who have ever been tested, as of the end of our sample period. We define the prior test rate of person i as of date t as the probability that i was tested at least once during the week-long period $[t-15, t-9]$. We focus on this window because it is the second week prior to our hospital testing window (which runs from $t-2$ to $t+4$ for a patient admitted at t). We allow for a week of time to elapse between the hospitalization and the "prior" testing because it is possible that some pre-hospital testing would occur in the window $[t-8, t-3]$. When studying prior tests, we limit the sample to each person's first hospitalization after March 1, 2020, to avoid picking up the higher testing that mechanically results from the fact that people hospitalized once are more likely than the general population to have been previously hospitalized. As with our bounds, we weight the data to match the population age distribution.

Table G.2 shows the community testing rate. The average county in Indiana has a testing rate of 25%, with an interquartile range of 22% to 28%. The average person lives

Table G.2: Hospitalized patients are not drawn from counties with high test rates

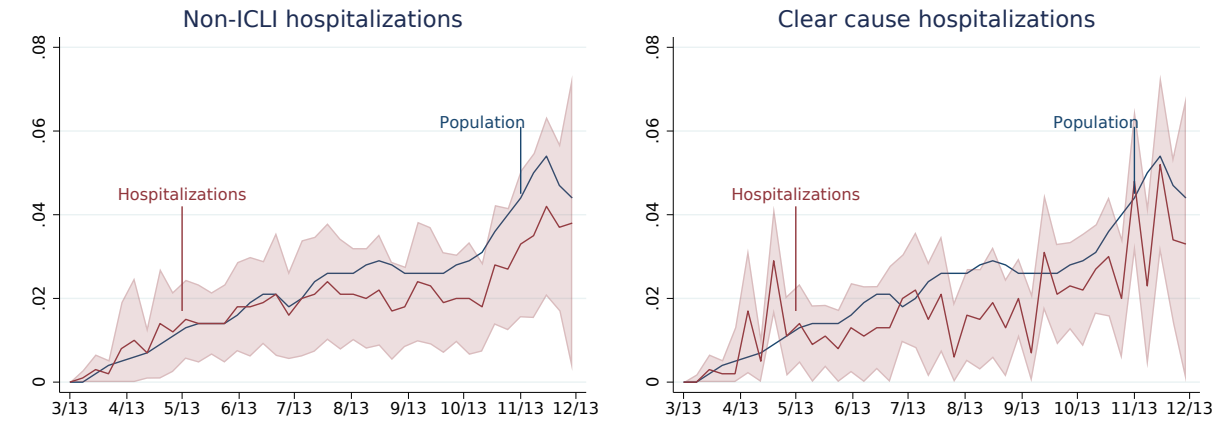
	County test rate
Average	0.252
25th percentile	0.219
75th percentile	0.280
Population	
Average person	.267
Hospitalizations	
Non-ICLI	.266
	[20.8]
Clear cause	.265
	[16.9]
ICLI	.267
	[2.5]

Notes: The county test rate is the share of the county population tested at least once in our test data. Table reports county-level statistics, as well as the average county test rates for the general population, the non-ICLI hospitalizations, clear cause hospitalizations, and ICLI hospitalizations, as well as t-statistic (in brackets) for the null hypothesis that the average person and the average hospitalization have the same county test rate.

in a county with a test rate of 26.7%. The average non-ICLI hospitalized patient comes from a county with a test rate of 26.6%, and the average clear-cause hospitalization patient comes from a county with a test rate of 26.5%. Among ICLI hospitalizations it is 26.7%. Our sample size is large enough that these differences are all statistically significant. Practically, however, the differences are very small. Hospitalized patients appear to come from counties that are roughly representative in terms of their testing rates. These rates are all significantly different from the population average.

Figure G.1 shows the prior testing rate as a function of admission date for the non-ICLI hospitalization sample, the clear-cause hospitalization sample, and the general population (for which the prior test rate on day t is defined as the fraction tested between $t - 15$

Figure G.1: Prior test rates, population and hospitalization samples



Notes: The prior test rate is the fraction of the group at date t that was tested between $t-15$ and $t-9$. Figure plots the average prior test rate for the population, for non-ICLI hospitalizations (in the left panel) and for clear cause hospitalizations (right panel). The shaded area is the 95% confidence interval for each week and hospitalization sample.

and $t-9$). The rates in the hospitalization samples are initially close to the population rate (when testing is low in general), but the lines diverge. By the last week of the sample, the prior testing rate is 1-2 percentage points lower in the hospitalization samples, than in the population. Although the differences in weekly testing rates are not statistically significant, the lower prior testing rate in the hospital sample could indicate that the hospital sample is negatively selected on SARS-CoV-2 infection risk.