medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Assessing the Age Specificity of Infection Fatality Rates for COVID-19: Meta-Analysis & Public Policy Implications

Andrew T. Levin, Kensington B. Cochran, and Seamus P. Walsh

23 July 2020

Abstract: This paper assesses the age specificity of the infection fatality rate (IFR) for COVID-19. Our benchmark meta-regression synthesizes the age-specific IFRs from four recent large-scale seroprevalence studies conducted in Belgium, Geneva, Spain, and Sweden. The estimated IFR is close to zero for children and younger adults but rises exponentially with age, reaching about 0.3 percent for ages 50-59, 1 percent for ages 60-69, 4 percent for ages 70-79, and 24 percent for ages 80 and above. We compare those predictions to the age-specific IFRs implied by recent seroprevalence studies of six U.S. geographical areas, three small scale studies, and three countries (Iceland, New Zealand, and Republic of Korea) that have engaged in comprehensive tracking and tracing of COVID-19 infections. We also review more than 30 other seroprevalence studies whose design was not well-suited for estimating age-specific IFRs. Our findings indicate that COVID-19 is not just dangerous for the elderly and infirm but also for healthy middle-aged adults, for whom the fatality rate is roughly 50 times greater than the risk of dying in an automobile accident. Consequently, the overall IFR for a given location is intrinsically linked to the age-specific pattern of infections. In a scenario where the U.S. infection rate reaches nearly 30 percent, our analysis indicates that protecting vulnerable age groups could prevent over 200,000 deaths.

Affiliations: Levin is a professor of economics at Dartmouth College, research associate of the NBER, and international research fellow of the Centre for Economic Policy Research. Cochran and Walsh are recent graduates of Dartmouth College.

Declaration: The authors have no financial interests nor any other conflicts of interest related to this study. No funding was received for conducting this study. The views expressed here are solely those of the authors and do not represent the views of any other person or institution.

Corresponding Author: Andrew T. Levin *(andrew.t.levin@dartmouth.edu)*

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Assessing the Age Specificity of Infection Fatality Rates for COVID-19: Meta-Analysis & Public Policy Implications

Andrew T. Levin, Kensington B. Cochran, and Seamus P. Walsh

23 July 2020

Objective: Determine age-specific infection fatality rates for COVID-19 to inform public health policies and communications that help protect vulnerable age groups.

Methods: Studies of COVID-19 prevalence were collected by conducting an online search of published articles, preprints, and government reports identified by online searches. A total of 48 studies were identified covering a wide array of locations in advanced economies. Studies were screened using four specific criteria: (i) transparency of seroprevalence test characteristics and procedures; (ii) representative sample of the general population; (iii) effective pandemic containment by the time of the study; and (iv) reporting of age-specific prevalence and confidence intervals. Age-specific IFRs were computed using reported fatalities four weeks after the midpoint date of each study, reflecting the typical pattern of lags in fatalities and reporting. Four studies were identified as benchmarks and used in meta-regression of the infection fatality rate (IFR) as a function of age, using the STATA *metareg* procedure. The meta-regression results were compared with age-specific IFRs for 12 other locations -- an "out-of-sample" exercise that statisticians commonly use in assessing the validity of forecasting models.

Results: Our analysis finds a highly significant exponential relationship between age and IFR for COVID-19. The estimated age-specific IFRs are close to zero for children and younger adults but rise to about 0.3 percent for ages 50-59, 1 percent for ages 60-69, 4 percent for ages 70-79, and 24 percent for ages 80 and above. Nearly all of the age-specific IFRs from all 16 studies fall within the 95% prediction interval of the meta-regression.

Discussion: Our analysis indicates that the link between IFR and age is far stronger than the link with comorbidities such as chronic pulmonary disease. Other recent studies indicate that other characteristics—including race, ethnicity, and socioeconomic status—are not tightly linked to IFR but are strongly tied to infection rates and hence to disparities in mortality outcomes.

Conclusions: Our results clearly indicate that COVID-19 is not just dangerous for the elderly and infirm but also for healthy middle-aged adults, for whom the fatality rate is roughly 50 times greater than the risk of dying in an automobile accident. Furthermore, the overall IFR for COVID-19 should not be viewed as an exogenously fixed parameter but as intrinsically linked to the age-specific pattern of infections. Consequently, individual and collective efforts that minimize infections in older adults could substantially decrease total deaths. In a scenario where the infection rate of the U.S. population reaches nearly 30%, our analysis indicates that protecting vulnerable age groups could prevent over 200,000 deaths.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1: COVID-19 Cases in New York City

	Total as of July 15, 2020	Share of Infections
NYC Residents	8 Million	*NA*
Estimated Infections	1.6 Million	100%
Symptomatic Cases	1.1 Million	65%
Reported Cases	220,000	12%
Hospitalization	55,000	3%
Fatalities	23,000	1%

Source: New York City Health Department (2020).

1. Introduction

As the COVID-19 pandemic has spread across the globe, some fundamental issues have remained unclear: How dangerous is COVID-19? And to whom? The answers to these questions have crucial implications in determining appropriate public health policies as well as informing prudent decision-making by individuals, families, and communities.

The standard epidemiological approach to gauging the severity of an infectious disease is to determine its *infection fatality rate* (IFR), that is, the ratio of deaths to the total number of infected individuals. The IFR is readily observable for certain viruses, such as Ebola, where nearly every case is associated with severe symptoms and the incidence of fatalities is extremely high; for such diseases, the IFR is practically identical to the *case fatality rate* (CFR), that is, the ratio of deaths to reported cases. By contrast, most people who are infected with SARS-Cov-2—the virus that causes COVID-19—are asymptomatic or experience only mild symptoms such as headache or loss of taste and may be unlikely to receive a viral test or be included in official case reports. Consequently, reported cases tend to comprise a small fraction of the total number of infections, and hence the CFR is *not* an adequate metric for the true severity of the disease.

As shown in Table 1, assessing the IFR for COVID-19 is analogous to finding a needle in a haystack, especially in a dense urban area such as New York City (NYC). The New York State Department of Health recently conducted a large-scale seroprevalence study and estimated the NYC infection rate at about 22 percent, that is, 1.6 million out of 8 million NYC residents.[1] As of mid-July, NYC had about 220,000 reported COVID-19 cases, almost exactly one-tenth of the total number of infections. About one-fourth of those reported cases were severe enough to require hospitalization, many of whom unfortunately succumbed to the disease. All told, fatalities represented about one-tenth of reported cases but only one-hundredth of all infections.

While the NYC data indicate an IFR of about 1 percent, analysis of other locations has produced a puzzlingly wide array of IFR estimates, ranging from around 0.5 percent in Geneva and Zurich to rates above 2 percent in Spain and in the Republic of Korea (henceforth "Korea"). Indeed, a

[1] See New York Department of Health (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

recent meta-analysis noted the high degree of heterogeneity across aggregate estimates of IFR and concluded that research on age-stratified IFR is "urgently needed to inform policymaking."[2]

In this paper, we consider the hypothesis that the observed variation in IFR across locations may primarily reflect the age specificity of COVID-19 infections and fatalities. In particular, the overall IFR for a given location can be expressed as follows:

$$IFR = \sum_{a=1}^{N} POPSHARE_a * INFRATE_a * IFR_a$$

using data for N distinct age groups, where $POPSHARE_a$ denotes the share of age group a in the total population, $INFRATE_a$ denotes that age group's COVID-19 infection rate, and IFR_a denotes that age group's infection fatality rate. Demographic information about the age structure for a given location is readily available from census data. Consequently, a crucial task is to use seroprevalence data to assess age-specific infection rates and IFRs.

Rather than focusing on any single location, we proceed by conducting *meta-analysis* using data from a wide array of distinct locations, drawn from a total of 48 recent studies of COVID-19 prevalence. We begin by highlighting key characteristics of studies that are essential for assessing age-specific IFRs, including the use of a broadly representative sample of the general population, seroprevalence methods with high positive predictive power, and tabulation of fatalities that are appropriately linked to the dates of the seroprevalence testing. Based on those criteria, we exclude 32 other seroprevalence studies that are not suitable for estimating age-specific IFRs.

Using these criteria, we identify four seroprevalence studies that serve as benchmarks: Belgium, Geneva, Spain, and Sweden. Applying meta-regression methods to this set of benchmark studies, we estimate a log-linear relationship between IFR and age and obtain precise coefficients that are not significantly influenced by outliers. In particular, the estimated IFR is close to zero for children and younger adults but increases exponentially with age, reaching about 0.5 percent for ages 55-64, 1.6 percent for ages 65-74, and exceeding 6 percent for ages 70+.

Next, we compare these meta-regression predictions to the age-specific IFRs implied by recent seroprevalence studies of six U.S. geographical areas (Connecticut, South Florida, Missouri, New York, Utah, and Puget Sound) as well as three countries (Iceland, Korea, and New Zealand) that have engaged in comprehensive tracking and tracing of COVID-19 infections. We also compare these results with three small-scale seroprevalence studies (Castiglione d'Adda, Gangelt, and the Diamond Princess cruise ship) and with the pathbreaking study of Ferguson et al. (2020) that was conducted at a very early stage of the pandemic.

[2] See Meyerowitz-Katz and Merone (2020), p.3.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Our analysis has two key conclusions: (1) COVID-19 is not just dangerous for the elderly and infirm but also for healthy middle-aged adults, for whom the fatality rate is roughly 50 times greater than the risk of dying in an automobile accident; and (2) age-specific policy choices and communications can dramatically decrease COVID-19 deaths. In particular, the overall IFR should *not* be viewed as an exogenously fixed parameter but as intrinsically linked to the age composition of the population and the age-specific pattern of infections.[3] Consequently, individual and collective efforts that minimize infections in older adults could substantially decrease total deaths. In a scenario where the infection rate of the U.S. population reaches nearly 30%, our analysis indicates that protecting vulnerable age groups could prevent over 200,000 deaths.

The remainder of this paper is structured as follows: Section 2 describes our methodology. Section 3 presents our meta-analysis results. Section 4 considers these findings in the context of other demographic characteristics (including race and ethnicity) and co-morbidities. Section 5 discusses the public policy implications of our analysis, including comparison to other types of fatality risks and scenario analysis of the age-specific pattern of U.S. infections and deaths.

2. Methodology

2.1 Overview

To perform the present meta-analysis, we collected published papers and preprints that have studied the seroprevalence and/or infection fatality rate of COVID-19. To identify these studies, we performed online searches in MedRxiv and Medline using the criterion ((“infection fatality rate” or “IFR” or “seroprevalence”) and (“COVID-19” or “SARS-Cov-2”)).[4] We identified other studies listed in reports by government agencies such as the U.S. Center for Disease Control & Prevention and the U.K. Parliament Office.[5] Finally, we confirmed the comprehensiveness of our literature search by referring to two recent meta-analysis studies that have assessed overall IFR for COVID-19 and a recent meta-analysis study comparing seroprevalence with reported cases.[6]

Before proceeding further, we restricted our meta-analysis to studies of advanced economies, based on current membership in the Organization for Economic Cooperation and Development (OECD).[7] It should be emphasized that we applaud recent efforts to assess seroprevalence in a

---

[3] Acemoglu, Chernozhukov, Werning, and Whinston (2020) and Chen et al. (2020) analyze optimal targeted lockdowns and reopenings in analytical frameworks with distinct age groups (e.g, young, middle-aged, and retired) using age-specific IFRs calibrated to the findings of Ferguson et al. (2020) and Verity et al. (2020), respectively. By contrast, Hall, Klenow, and Jones (2020) analyze these issues using a more stylized analytical framework in which the aggregate IFR is an exogenously fixed parameter.

[4] These searches were conducted on July 1 and updated on July 10 and July 19.

[5] For example, see U.K. Parliament Office (2020).

[6] See Ioannidis (2020) and Meyerowitz-Katz and Merone (2020) for meta-analysis of the overall IFR for COVID-19 and Byambasuren et al. (2020) for a systematic comparison of seroprevalence with reported cases.

[7] OECD countries include: Australia, Austria, Belgium, Canada, Chile, Colombia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States (*https://www.oecd.org*).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

number of developing countries (including Brazil, Croatia, Ethiopia, and Iran), but we have excluded those studies in light of the distinct challenges associated with health care provision and reporting of fatalities in those locations.[8] We also excluded studies focused exclusively on measuring seroprevalence in a narrow segment of the population such as health care workers or pregnant women.[9] Appendix A lists all of the studies identified in our literature search.

2.2 Prevalence Measures

Our meta-analysis encompasses two distinct approaches for assessing COVID-19's prevalence: (1) extensive tracking and contact-tracing using live-virus testing and (2) seroprevalence studies that test for antibodies produced in response to the virus. Testing for the live virus is done by either a quantitative reverse-transcription polymerase chain reaction (qRT-PCR) molecular test for the viral nucleic acid sequence, or an antigen test for proteins specific to the virus.[10] These tests detect the virus within a few days of disease onset. While using live antigen testing is the optimal approach for determining prevalence, it requires extensive continuous testing of a population, and was only thoroughly implemented in select countries with relatively small populations, notably South Korea, Iceland, and New Zealand.

Most studies of COVID-19 prevalence have proceeded using serological analysis to determine what fraction of the population has developed either IgG or IgM antibodies to the virus. IgM antibodies develop earlier, but decrease over time, while IgG antibodies develop later and remain in high concentrations for several months. Antibodies are tested for using several methods. Enzyme-linked immunosorbent assays (ELISA) proceed by tagging antibody-antigen interactions with a reporter protein. Chemiluminescent immunoassays (CLA) work similarly by tagging the antigen-antibody interaction with a fluorescent protein. Lateral Flow Assays (LFA), also known as rapid diagnostic tests (RDT), produce a colored band upon antigen-antibody interaction.

Recognizing that SARS-Cov-2 is both novel and hazardous, public regulatory agencies have issued "emergency use authorizations" (EUA) to facilitate the rapid deployment of live virus and antibody tests based on the test characteristics reported by each manufacturer.[11] Subsequent studies by independent laboratories have reassessed the characteristics of these test kits, in many cases finding markedly different results than those of the manufacturer. Such differences reflect (a) the extent to which test results may be affected by seemingly trivial differences in its implementation, and (b) the extent to which seriological properties may vary across different segments of the population. For example, a significant challenge in producing accurate tests is to distinguish COVID-19 antibodies from those associated with other coronaviruses (including the

[8] See Silveira et al. (2020), Jerkovic et al. (2020), Kempen et al. (2020), and Shakiba et al. (2020) for seroprevalence analysis of locations in Brazil, Croatia, Ethiopia, and Iran, respectively. Fassihit and Gladstone (2020) highlight the shortcomings of official tabulations of COVID-19 fatalities in Iran during the early stages of the pandemic.
[9] For example, Flannery et al. (2020) assess seroprevalence in parturient women.
[10] Carter et al., 2020
[11] For example, see U.S. Food & Drug Administration (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

common cold). Consequently, the assessment of test characteristics may vary with seemingly innocuous factors such as the season of the year in which the blood samples were collected.

The reliability of seroprevalence testing depends on three key factors: (1) the seroprevalence test's sensitivity (odds the test detects the virus in an infected person); (2) the seroprevalence test's specificity (odds the test returns a negative result for a uninfected person); and (3) the true disease prevalence in the sample. In a population where the actual prevalence is relatively low, the frequency of false-positive tests is crucial for determining the reliability of the test results. Consequently, a key metric of test reliability is positive predictive value (PPV), that is, the likelihood that a positive test result is a true positive. The PPV can be evaluated as follows:

$$\text{PPV} = \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})}$$

Evidently, lower prevalence can markedly diminish the reliability of seroprevalence testing. For example, in a seroprevalence study of Dutch blood donors using the Wantai Total Antibody ELISA, the crude prevalence rate was found to be 2.7%.[12] However, that antibody test has a PPV of 42.4%, and hence the adjusted prevalence is only 0.6 %, with a 95 percent confidence interval of 0% to 5.2%. In effect, *practically all of the positive tests obtained in this study might be false positives*. By contrast, a seroprevalence study of New York City found a much higher crude prevalence of 20.0% using a Wadsworth Pan-Ig test with a PPV of 94.8%.[13] Consequently, the adjusted prevalence for this study is *higher* than the crude prevalence, namely, 21.7% with a 95 percent confidence interval of 19.2% to 24.4%.[14]

Test sensitivity and specificity also have a high impact on PPV. For example, in a serological study of Santa Clara County, researchers used a Premier Biotech LFA test and estimated prevalence at 1.5% based on a test specificity of 99.5%.[15] However, a subsequent study found the specificity of that test to be only 97.2%.[16] That revision to the test specificity reduces its PPV in the Santa Clara study from 71.6% to 31.1%, and the adjusted prevalence for Santa Clara residents declines to 0%; that is, the prevalence in that population was so low that it could not be distinguished from zero.[17]

These examples underscore why the sensitivity and specificity of COVID-19 antibody tests should *not* be treated as fixed parameters that are known with a high degree of certainty, as would generally be the case for medical tests of other diseases that have been authorized via standard regulatory procedures. Thus, the 95% confidence interval for each seroprevalence

[12] See Slot et al. (2020)
[13] See Rosenberg et al. (2020).
[14] See Appendix Table B2 for further details.
[15] See Bendavid et al. (2020).
[16] See Whitman et al. (2020).
[17] See Appendix Table B3 for further details.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

estimate should reflect the degree of uncertainty about its sensitivity and specificity as well as the conventional uncertainty that reflects the size of the sample used in producing that estimate.[18]

In light of these testing reliability concerns, our meta-analysis excludes seroprevalence studies that do not disclose the test method or report estimates and confidence intervals that reflect the characteristics of the test. We also exclude studies that relied on test kits that were subsequently withdrawn by the manufacturer due to concerns about inadequate reliability.

2.2 Constructing a Representative Seroprevalence Sample

In order to accurately use antibody tests to estimate population prevalence, the study sample must accurately reflect the sampled population. We exclude four types of seroprevalence studies from our paper that do not provide accurate estimates of population-wide age-specific infection rates: (1) studies from clinics including COVID-19 patients; (2) studies which employed active recruitment; (3) studies whose samples are heavily skewed by age; and (4) studies of blood donors. Many of these studies were extremely useful in their originally intended contexts, but are not useful in the current context, for reasons outlined below.

Studies from serum samples at health care facilities produced inflated estimates of prevalence when those samples included patients who were obtaining treatment for symptoms associated with COVID-19. For example, a New York City study from an outpatient clinic in May yielded a seroprevalence estimate double that of two April studies of the same area. As the entire state was in lockdown between the two time points, it is difficult to believe that the prevalence grew so drastically during that time period. Instead, it is likely that the majority of primary care and urgent care patients in the month of May sought medical attention for COVID-like symptoms, and when these patients were not excluded, the estimate of COVID-19 prevalence was inflated

Studies which employ active recruitment also inflate the number of positive patients, as people who think they are positive are more likely to enroll for the free testing. For example, in a study in Luxemburg, of the 35 participants who tested positive, 19 had previously interacted with a person who they knew was positive or had been tested for SARS-CoV-2 previously. Excluding these patients from the sample doubles this particular study's implied overall IFR.

Even random samples may only draw from younger people, making them less useful for estimating age-specific prevalence and IFRs across the full spectrum of the population. For example, in the French town of Oisie, seroprevalence tests were run on schoolchildren, their teachers, and their immediate families; the entire study only included two individuals older than sixty-five.[19] Since our analysis is aimed at gauging the relationship between IFR and age, we restrict our meta-analysis to studies that report age-specific results for a broad spectrum of age groups.

Finally, seroprevalence studies of blood donors also likely are non-representative and inflate the actual infection rate. In the discussion of their Milan blood donor study, Valenti et al. (2020)

[18] See Manski and Molinari (2020) and Larremore et al. (2020).
[19] See Fontanet et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

note that blood donors are generally healthier than the general population and "might have a higher number of social interactions than other groups."[20] For example, a blood donor study in the United Kingdom estimated 7.8% (CI: 7.1 to 8.6%)[21] of the population had COVID-19, while a randomized seroprevalence study of the UK population at a later date estimated the prevalence of COVID-19 to be 5.41% (CI: 4.3 to 6.5%).[22] Though blood donor studies were useful while few other samples were available, randomized seroprevalence studies provide more reliable population prevalence estimates.

Blood specimen from commercial labs may also not be representative of the location's broader population. The CDC conducted seroprevalence testing on blood samples from commercial labs in New York City (NYC), Connecticut, Utah, Missouri, and Puget Sound, Washington.[23] Though these studies mostly control for demographic data such as age, zip code, race, and sex, some other factors may skew the data. Patients who received healthcare during quarantine may have been more cautious of infection due to their underlying medical conditions; alternatively, a decision to enter medical spaces may have corresponded with a less cautious demographic. The CDC compares their estimated seroprevalence in each location to the number of reported cases. Four locations have ratios of seroprevalence to reported cases of around 11:1, but the Connecticut study reports a much lower ratio of 6:1, suggesting the study may have underestimated prevalence there, while the Missouri study has a ratio of 25:1 suggesting that prevalence may have been overestimated. These observations are further discussed in Section 3.

Concerns about potential sample selection issues in studies of health care specimens are underscored by findings from a seroprevalence study that analyzed specimens from two commercial laboratories in New York City.[24] In that study, the estimated seroprevalence differed markedly between the two labs, with non-overlapping 95% confidence intervals of 6.5 to 12.3% for Lab A and 3.7 to 6.1% for Lab B.

2.3 Matching Infections to Deaths

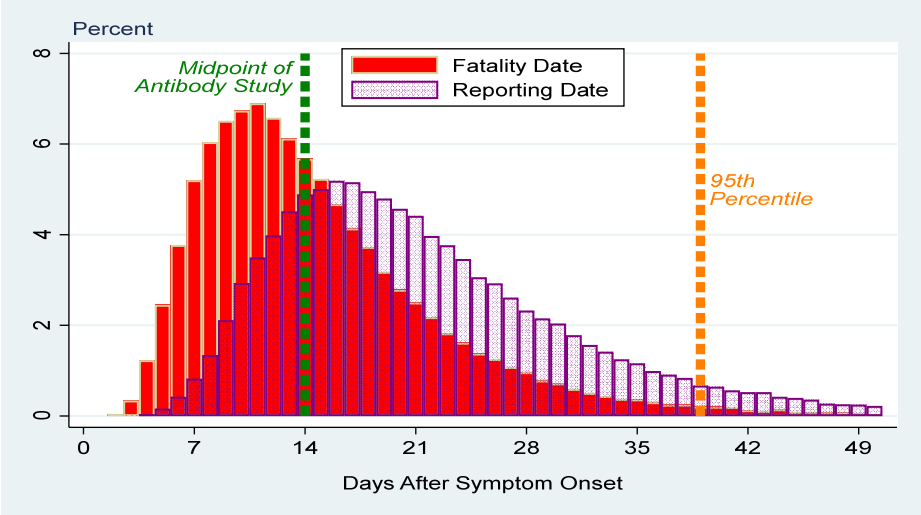
Accurately measuring total deaths, the numerator of the IFR calculation, is perhaps more difficult than assessing prevalence. The time lags from onset of symptoms to death and from death to official reporting are crucial. Symptoms typically develop within 6 days after exposure, but may develop as early as 2 days or as late as 14 days.[25] More than 95% of symptomatic COVID patients have positive antibody (IgG) tests within 17-19 days of symptom onset.[26]

The CDC estimates that the mean time interval from symptom onset to death is 15 days for ages 18-64 (interquartile range of about 9 to 24 days) and 12 days for ages 65+ (IQR of 7 to 19 days). The mean interval from date of death to the reporting of that person's death is about 7 days

[20] See Valenti et al. (2020).
[21] See Public Health England (2020).
[22] See UK Office of National Statistics (2020).
[23] See Havers et al. (2020).
[24] See Havers et al. (2020).
[25] See McAloon et al. (2020).
[26] See Long et al. (2020) and Sethurman et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 1: Time Lags in the Incidence and Reporting of COVID-19 Fatalities



Source: authors' calculations based on CDC estimates; see text.

(interquartile range of about 2 to 19 days). Consequently, the upper bound of the 95% confidence interval between symptom onset and reporting of fatalities is about six weeks (41 days).[27]

Figure 1 illustrates a scenario in which the pandemic ended two weeks prior to the date of a seroprevalence study. This figure shows the results of a stochastic simulation calibrated to reflect the CDC's estimated distribution for the time lags between symptom onset, death, and inclusion in official fatality reports. The histogram shows the frequency of deaths and reported fatalities associated with the infections that occurred on the last day prior to full containment. Consistent with the CDC confidence intervals, about 95% of cumulative fatalities are reported within roughly four weeks of the date of the seroprevalence study.

These considerations underscore the pitfalls of constructing IFRs based on the death toll at the midpoint date of a seroprevalence study, which is the approach that has been taken in most previous studies (including both of the meta-analysis studies of the overall IFR for COVID-19).

In particular, as shown in Table 2, the cumulative fatalities at the time of a seroprevalence study can markedly understate the full death count as of four weeks later. All of these studies were conducted in locations where the pandemic had been contained by the time that seroprevalence was measured, as evident from the fact that the fatality count leveled off over the subsequent month.[28]

Evidently, the precise timing of the count of cumulative fatalities is relatively innocuous in locations (such as Spain and Castiglione d'Adda) where the outbreak had been contained for more than a month prior to the date of the seroprevalence study. But for the other studies shown

[27] See U.S. Center for Disease Control & Prevention (2020e), Table 2.

[28] In each of the locations shown in Table 2, cumulative fatalities stabilized over the month following the seroprevalence study; in each case, the percent change from week 4 to week 5 was less than 10% with the exceptions of Missouri (18%) and South Florida (11%). See Appendix Table D1 for further details.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2: Timing of Reported Fatalities for Selected Seroprevalence Studies

	Cumulative Fatalities		
Location	**Study Midpoint**	**4 Weeks Later**	**% Change**
Belgium	6,262	8,843	41
Geneva, Switzerland	255	287	13
Spain	26,834	27,136	1
Sweden	2,586	3,831	48
Connecticut	2,257	3,637	61
South Florida	513	1,160	126
Missouri	218	562	158
New York State	20,212	28,663	42
Utah	41	96	134
Puget Sound, Washington	536	732	37
Castiglione d'Adda, Italy	62	62	0
Gangelt, Germany	7	9	29
Diamond Princess	8	10	25

Sources: See Appendix A.

in Table 2, the outbreak had only recently been contained, and hence the death count continued rising markedly for several more weeks after the midpoint of the seroprevalence study. For each of those locations, matching seroprevalence to the death count at the midpoint date of the study would significantly underestimate the true level of the IFR. For example, in the case of New York state, computing the IFR using the 4-week fatality count is nearly 1.5 times higher than using the fatality count at the midpoint date of that study (which was conducted in late April).

By contrast, matching seroprevalence estimates with subsequent fatalities is infeasible if the seroprevalence study takes place in the midst of an accelerating outbreak. In particular, if infections and fatalities continue rising exponentially over subsequent weeks, there is *no* precise way of determining what fraction of those deaths resulted from infections before vs. after the date of the seroprevalence study.

Therefore, a crucial criterion for seroprevalence studies to be included in our meta-analysis is that the pandemic is well contained in advance of the study, as indicated by the stabilization of cumulative fatalities within the next several weeks after the midpoint date of the study.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 3: Seroprevalence Studies Conducted during Accelerating COVID-19 Outbreaks

		Cumulative Fatalities		Percent Change
		Midpoint Date	4 Weeks Later	
Los Angeles County	April 10-11	265	1,468	454%
New York City	March 23 – April 1	1,066	14,261	1,238%
Santa Clara County	April 3-4	39	113	190%
Scotland	March 21-23	72	2,273	3,057%

Sources: See Appendix A.7.

As shown in Table 3, four studies are clearly inconsistent with that criterion: Los Angeles County (mid-April), New York City (late March), Santa Clara County (early April), and Scotland (late March).[29] It should be emphasized that these studies provided valuable information about seroprevalence in the midst of an active outbreak, but these seroprevalence results are not well-suited for gauging the IFR of COVID-19.[30]

Finally, it should be noted that reported deaths may not fully capture all fatalities resulting from COVID-19 infections, especially in locations where a substantial fraction of such deaths occur outside of healthcare institutions. In the absence of accurate COVID-19 death counts, an alternative measure, referred to as *excess mortality*, can be computed by comparing the number of deaths for a given time period in 2020 to the average number of deaths over the comparable time period in prior calendar years, e.g., 2015 to 2019. For example, a recent Belgian study used seroprevalence results in conjunction with excess mortality to compute age-specific IFRs, noting that their measure of excess mortality over the period from March to May coincided almost exactly with the tally of reported COVID-19 cases.[31] Similar approaches have been used in other European countries (including Spain, Italy, and the U.K.), while the U.S. Center for Disease Control & Prevention provides regular updates on excess mortality for U.S. geographical locations.[32]

[29] These four studies were conducted by Sood et al. (2020), Havers et al. (2020), Bendavid et al. (2020), and Thompson et al. (2020), respectively.
[30] Nonetheless, all four of these studies have been used in prior meta-analyses aimed at assessing the overall IFR, using reported fatalities as of the midpoint date of each seroprevalence study.
[31] See Molenberghs et al. (2020).
[32] See Rinaldi and Paradisi (2020), Modi et al. (2020), and U.S. Center for Disease Control & Prevention (2020c).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

2.4 Meta-Regression Methodology

The goal of our meta-analysis is to systematically assess previous studies of mortality and infection rates to determine how age and fatality risk are related. To perform this analysis quantitatively, we use random-effects meta-regressions, using the STATA *metareg* procedure.[33] Meta-regressions are a useful tool for comparing study-level summary data. Since the individual observations are not available for any study we use, comparing summary-level data is the only way to compare the studies. We use summary level data from each age group in each study, so effectively one study has multiple "groups" in our meta-regressions.[34]

We treat each age group separately because there are likely random variations in age-specific IFR both across studies and across age groups within a study. Random-effects procedures allow for such random variation between groups (referred to as *residual heterogeneity*) by assuming that these effects are drawn from a Gaussian distribution. The procedure provides reasonable results even if the errors are not strictly normal but may be unsatisfactory if the sample includes large outliers or the distribution of groups is not unimodal. In analytical terms, this framework can be expressed as follows:

$$log\left(IFR_{ij}\right) = \alpha \ + \ \beta * age_{ij} \ + \ \epsilon_{ij} \ + \ u_{ij}$$

$$\text{where } \ u_{ij} \sim N(0, \tau^2) \ \text{ and } \ \epsilon_{ij} \sim N\left(0, \sigma_{ij}^2\right)$$

In this specification, IFR_{ij} is the estimated IFR in study *i* for age group *j*, age_{ij} denotes the median age of that group, ϵ_{ij} denotes the source of idiosyncratic variations for that particular location and age group, and u_{ij} denotes the random effects that characterize any systematic deviations in outcomes across locations and age groups. Under the maintained assumption that each idiosyncratic term ϵ_{ij} has a normal distribution, the idiosyncratic variance is $\sigma_{ij}^2 = ((U_{ij} - L_{ij})/3.96)^2$, where U_{ij} and L_{ij} denote the upper and lower bounds of the 95% confidence interval for that study-age group. The random effects u_{ij} are assumed to be drawn from a homogeneous distribution with zero mean and variance τ^2. The null hypothesis of $\tau^2 = 0$ characterizes the case in which there are no systematic deviations across studies or age groups. If that null hypothesis is rejected, then the estimated value of τ^2 encapsulates the magnitude of those systematic deviations.

Under our baseline specification, the infection fatality rate increases exponentially with age.[35] In particular, this meta-regression is specified in logarithmic terms, with the slope coefficient β encapsulating the impact of higher age on *log(IFR)*. Consequently, the null hypothesis that IFR

[33] See Harbord and Higgins (2008) and Higgins, Thompson, and Spiegelhalter (2009).
[34] We also replicated this analysis using fixed effects for studies and random effects for age groups within studies.
[35] Bonanad et al. (2020) conducted a meta-analysis study of COVID-19 case fatality rates as a function of age using aggregate data from China, Italy, New York, Spain, and the U.K. and found a very strong exponential pattern of mortality: ages 40-49: 1.1%; ages 50-59: 3%; ages 60-69: 9.5%; ages 70-79 22.8%; ages 80+: 29.6%. Similarly, Doherty et al. (2020) investigated a large sample of U.K. hospitalized COVID-19 patients and identified an exponentially increasing mortality hazard rate as a function of patient age.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

is unrelated to age can be evaluated by testing whether the value of β is significantly different from zero. If that null hypothesis is rejected, then the estimated values of α and β characterize the estimated relationship between log(IFR) and age. Consequently, the predicted relationship between IFR and age can be expressed as follows:

$$IFR = e^{\alpha + \beta * age_{ij}}$$

The 95% confidence interval for this prediction can obtained using the delta method. In particular, let IFR_a denote the infection fatality rate for age a, and let σ_c denote the standard error of the meta-regression estimate of log(IFR_a). If IFR_a has a non-zero value, then the delta method indicates that its standard error equals σ_c / IFR_a , and this standard error is used to construct the confidence interval for IFR_a at each age a. Likewise, the prediction interval for log(IFR_a) is computed using a standard error of $\sigma_c + \tau$ that incorporates the systematic variation in the random effects across studies and age groups, and hence the corresponding prediction interval for IFR_a is computed using a standard error of $(\sigma_c + \tau)/IFR_a$.

2.5 Study Selection

In subsections 2.1 to 2.4, we have identified four specific criteria for determining whether a given study should be included in our meta-analysis: (i) transparency about the characteristics and positive predictive power of the prevalence test procedure; (ii) use of a sample data frame that is broadly representative of the general population; (iii) effective containment of the pandemic prior to the initiation of the prevalence survey; and (iv) and reporting of prevalence estimates and confidence intervals for specific age groups as required for the estimation of age-specific IFRs. Based on those four criteria, we have determined that 32 studies are not suitable for assessing the IFR of COVID-19 even though each of those studies has made significant contributions along other lines. Those 32 studies are listed in Appendix A along with the rationale for excluding each of them from our meta-analysis.

Consequently, our meta-analysis focuses on synthesizing IFR data from thirteen locations; see Appendix A for further details. These locations can be classified into four distinct groups:

- ***Benchmark Studies: Belgium, Geneva, Spain, and Sweden***. Each of these locations has been the subject of a large-scale seroprevalence study using a test procedure with high positive predictive power and a sample frame that is broadly representative of the general public and that covers a wide array of age groups.[36] For Belgium and Geneva, estimates of age-specific IFRs have been reported based on each location's seroprevalence results. For Spain, we construct age-specific IFRs using the seroprevalence data in conjunction with excess mortality data published by the Spanish National Institute of Statistics. For Sweden, we construct age-specific IFRs using the seroprevalence data in conjunction

[36] See Molenberghs et al. (2020), Perez-Saez et al. (2020), Pollan et al. (2020), and Sweden PHA (2020c,d).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

with cumulative fatalities four weeks after the midpoint date of the seroprevalence study.[37]

- ***Other Large-Scale Seroprevalence Studies*: *Connecticut, Missouri, New York, Puget Sound, South Florida, Utah.*** Five of these studies were conducted by the U.S. Center for Disease Control & Prevention, using sample specimens from two commercial laboratory companies, and the sixth was conducted by the New York Department of Health using samples collected at supermarkets and grocery stores.[38] In light of potential concerns about sample data frames and limited age-specific seroprevalence data, we include these studies in our meta-analysis but not in the benchmark group. We construct age-specific IFRs using seroprevalence results matched to cumulative fatalities four weeks after the midpoint date of each study.

- ***Comprehensive Tracking and Tracing Countries: Iceland, Korea, and New Zealand.*** These three countries engaged in extensive testing and tracing to halt the spread of infections. Iceland researchers also conducted a large-scale seroprevalence study, and we use the results of that study in computing age-specific IFRs for Iceland.[39] That study also indicates that reported cases in Iceland substantially understated actual prevalence; see Appendix C for details.[40] Thus, we make corresponding adjustments to the reported cases for Korea and New Zealand in constructing age-specific IFRs for each of those locations.

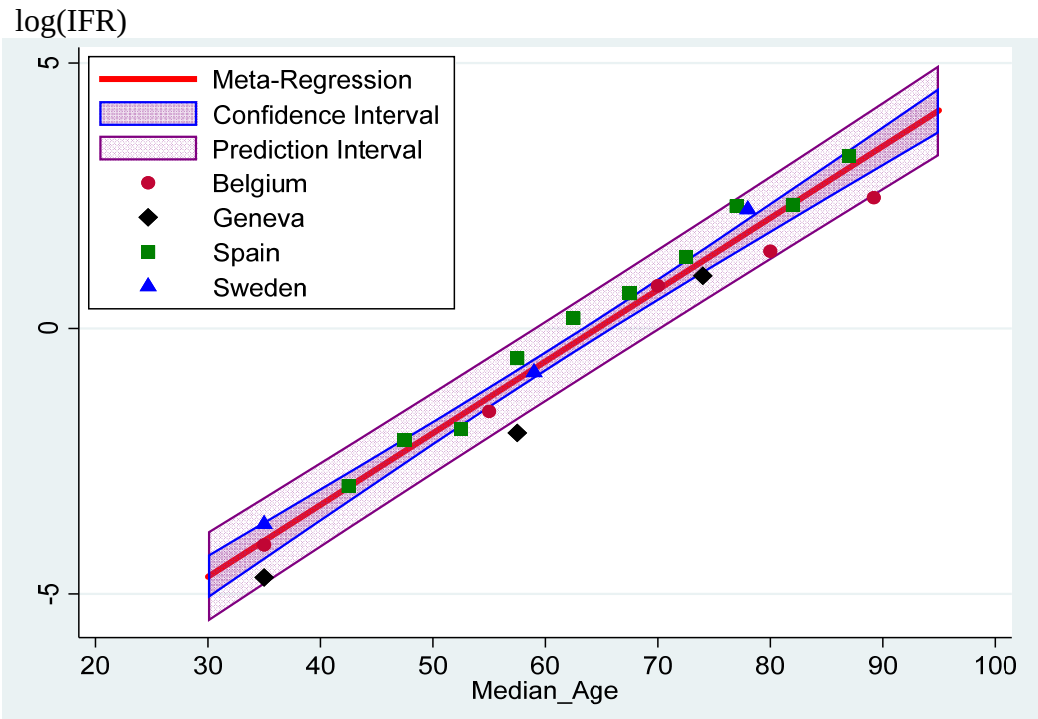
- ***Small-scale studies: Castiglione d'Adda, Gangelt, and Diamond Princess cruise ship***. The first two locations have had seroprevalence studies based on random samples, while data from Diamond Princess has been influential in informing subsequent studies.[41]

---

[37] Sweden Public Health Agency (2020f) recently produced estimates of the infection fatality rate in Stockholm for two age groups (ages 0-69 and 70+) using a novel methodology that links live virus tests, reported cases, and mortality outcomes. Given the markedly different methodology and the breadth of the two age groups, that study is not included in our meta-regression analysis, but it should be noted that their estimated IFR of 4.3% for ages 70+ is well aligned with the results of our meta-regression analysis.
[38] See Havers et al. (2020) and Rosenberg et al. (2020), respectively.
[39] See Gudbjartsson et al. (2020).
[40] See Aspelund, Droste, Stock, and Walker (2020) for statistical analysis of the deviation between reported cases and true prevalence in Iceland.
[41] See Pagani et al. (2020), Streeck et al. (2020), Russell et al. (2020), and Salje et al. (2020a,b).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 2: The Log-Linear Relationship between Age and IFR for COVID-19



3. Results

3.1 Benchmark Analysis

As shown in Figure 2, the results from a random-effects meta-regression of ln(age) on fatality risk reveal a clear, exponential relationship between age and IFR. Note that the relationship between age and IFR is exponential for all ages, not just the elderly. Thus, the predicted IFR for a forty-year-old parent is exponentially higher than for their ten-year-old child, even though both face relatively low absolute IFRs.

As noted in Section 2.4, one key issue for assuring the validity of our meta-regression method is confirming that the distribution of observations is consistent with a normal distribution, without any extreme outliers or clustering of observations. The validity of those assumptions is evident from Figure 2: the observations provide relatively even coverage of the age interval from 30 to 85 years, and nearly all of the observations fall within the 95 percent confidence interval. The only exception is the middle-aged group from the Geneva study, but even that observation is only slightly outside the confidence interval, as one might expect with one out of 21 observations for this meta-regression.[42]

[42] We have also used the output of the Stata *metareg* procedure to confirm that the estimated random effects are consistent with a normal distribution.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 3: The Link between Age and IFR for COVID-19

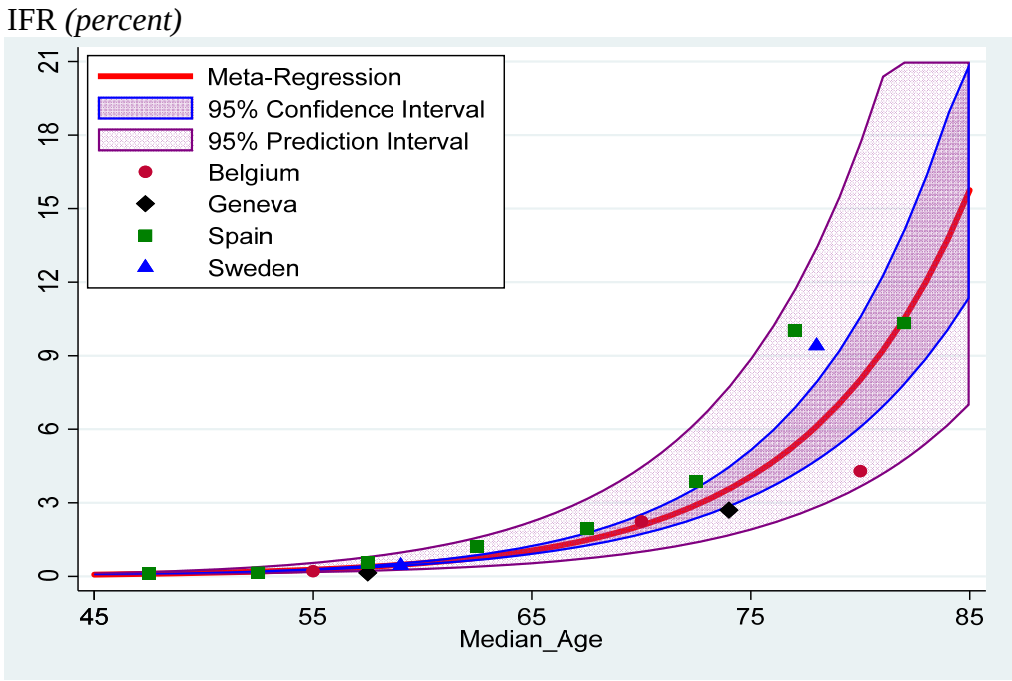


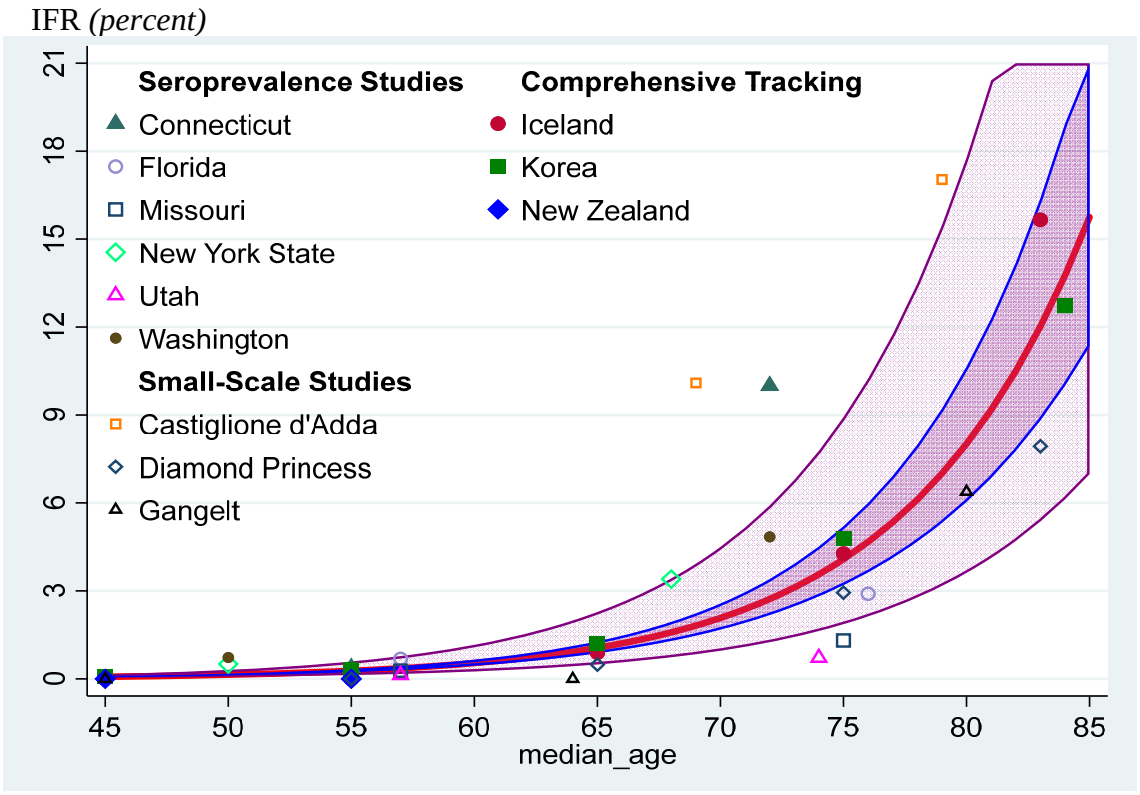
Figure 3 depicts the exponential relationship between IFR and age using the transformed output of the Stata *metareg* procedure. The thick red line is the meta-regression's predicted age-specific IFR curve. The dark purple area is the meta-regression's 95% confidence interval. This means there is a 95% chance the true relationship between IFR and age is within this confidence interval. The wider, light purple area is the meta-regression's 95% prediction interval. Taking into account random variation across studies and locations, there should be a 95% chance a subsequent study of age-specific IFR would land within that prediction interval. The prediction interval is much wider than the confidence interval, consistent with the inclusion of random effects in the meta-regression, i.e., there is substantial random variation across studies as well as age groups.

Figure 3 demonstrates a clear, exponential relationship between age and fatality risk. The disease poses a substantial mortality risk for middle-aged adults and even higher risks for elderly people: A 50 year-old faces a 0.14% chance of dying if infected, and that rate rises to 0.5% for a 60-year-old, 1.6% for a 70-year-old, and over 15% for people ages 85 and above.

As indicated in Figure 3, the second-oldest Spanish age-group (70 – 80 years) appears at the high end of the prediction interval. This may be due to sampling error that underestimated the infection rate (resulting in an inflated IFR estimate). Spanish researchers succeeded in conducting a very large-scale seroprevalence study encompassing about 100,000 people, with associated challenges in maintaining consistency across locations and age groups. Thus, it may be plausible that the infection rate was underestimated among the elderly, thereby diminishing the denominator in our IFR calculations. Finally, it is conceivable that elderly people in Spain

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 4: Assessing Age-Specific IFRs from Other Locations



might have a higher incidence of comorbidities than their peers in other European countries -- a possibility that we revisit in Section 4.

3.2 Comparing Benchmark Results to Other Studies

We now compare the benchmark meta-regression results with the age-specific IFRs for 12 other locations. In effect, this approach is comparable to "out-of-sample" exercises that statisticians commonly use in assessing the validity of a particular model.

As shown in Figure 4, the observations from these twelve studies generally fit within the benchmark prediction intervals, broadly confirming the usefulness of the meta-regression in assessing age-specific IFRs beyond the four locations covered by the set of benchmark studies. Moreover, the outliers in this figure are mostly above the prediction interval rather than below it; that is, the unexplained variations outside the prediction interval tend to be observations with unusually high IFRs.

We now consider the specific results for each group of studies:

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Seroprevalence Studies

Every seroprevalence study in this group except for New York State was conducted using commercial lab blood samples. Connecticut's oldest age-group (the green triangle) appears far above the prediction interval. As discussed in Section 2.2, the Connecticut study may have underestimated prevalence; if the prevalence were revised upwards to align with the typical ratio of reported cases to infections, then this particular observation would likely fall within the prediction interval. Another potentially significant issue is whether Connecticut experienced a relatively high infection rate among residents of assisted care facilities, who would be particularly vulnerable due to elevated age as well as a higher incidence of co-morbidities.

By contrast, the observations for Missouri lie below the prediction interval. In this case, the ratio of estimated seroprevalence to reported cases was extraordinarily high (24:1), more than double the typical 10:1 ratio for other seroprevalence studies.[43] In effect, Missouri's seroprevalence might be significantly overstated, perhaps reflecting non-representative aspects of the sample data frame (i.e., specimens collected by two commercial labs). Moreover, Missouri has a relatively low estimated prevalence of 2.7% compared with most of the other studies in our meta-analysis, which raises concerns about the PPV of the antibody test and the possibility that false-positives may have distorted the estimated prevalence.

Finally, these results suggest that Utah may indeed have a systematically different pattern of age-specific IFR compared to most other locations. As in Missouri, there are potential concerns about PPV for a location with relatively low estimated prevalence of about 2.2%. Unlike Missouri, however, the ratio of estimated infections to reported cases is about 11:1, similar to most other studies. One plausible factor is that a large fraction of Utah residents abstain from use of alcohol, tobacco, and narcotic drugs and hence may have a lower incidence of co-morbidities compared to most other locations in the United States, Europe, and East Asia. The bottom line is that Utah's remarkably low IFR should not simply be dismissed as an outlier; rather, further study of that location may yield significant insights that are applicable elsewhere.

Comprehensive Tracking and Tracing Countries

The age-specific IFRs for all three countries that employed widespread testing and contact-tracing to control the virus's spread all fit well within the benchmark's 95% prediction interval. Of the three, there has only been a seroprevalence study published about one country, Iceland. The study found that Iceland's ratio of actual to reported cases was about 1.4x, enabling us to reasonably reliably compute Iceland's age-specific IFRs.[44]

Both Iceland and New Zealand were able to fairly rapidly control the virus, and thus very few deaths occurred (10 in Iceland and 22 in New Zealand). As three of the deaths in Iceland were in healthcare workers, the death data could be greatly skewed by conflicting factors such as viral

[43] Havers, et al (2020). See Table 3.
[44] See Gudbjartsson et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

load; excluding healthcare workers alters their IFR by 30%. Since each region's death counts were so low, little can be inferred from their age-specific IFRs.

Small-Scale Studies

It is remarkable how well the small-scale studies of Gangelt, Germany, and the Diamond Princess cruise ship fit the benchmark analysis once we account for delays in the timing and reporting of fatalities. As shown in Table 2 above, the number of deaths at the time of the initial infection study was only about half of the final death count a few weeks later.

These two locations also demonstrate why large sample sizes should be weighted much more heavily than small sample sizes. The benchmark's estimated IFR for a 55 year-old is about 0.5%. So, on average, 1 out of 200 infected 55 year-olds should die. However, in Gangelt only about 150 people aged 50-60 were infected, and on the Diamond Princess only 59 people in that age range were infected. It would be quite consistent with the benchmark finding if no people in that age group died at either location. For a virus that only kills a small percentage of patients, observers should be wary of basing their findings on small absolute death counts, since a few deaths can significantly change the implied IFR. For such small samples, the appropriate confidence intervals are too wide to draw many conclusions from. Iceland (10 deaths) and New Zealand (22 deaths) share this limitation.

Castiglione d'Adda's age-specific IFRs for individuals aged 65-85 are outliers above our prediction interval. This outlier status was likely due to the fact that the town was affected relatively severely and early by COVID-19 in the pandemic's first wave, and hospitals became overwhelmed, leading to rationing of medical care. It should not be surprising that the region's IFR for people aged 85 years and older falls within our benchmark estimation; at such high ages the disease is so dangerous that medical care may not influence mortality to any significant degree. The relatively high IFR for the next age cohort (65 to 85 years) could also reflect a higher incidence of co-morbidities compared to similar cohorts in other locations. Higher IFRs might also reflect dense multi-generational urban housing that resulted in increased viral load for elderly people. In sum, the extraordinarily high IFR for individuals ages 65-85 years plays a key role in explaining Castiglione d'Adda's overall IFR of about 5%.

Finally, while not included in our formal meta-analysis, it should be noted that the pathbreaking study of Ferguson et al. (2020) is broadly consistent with our findings. That study was completed at an early stage of the COVID-19 pandemic, drawing on data from expatriation flights to estimate infection rates in Wuhan and then computing age-specific IFRs based on reported fatalities in Wuhan. As in our meta-regression results, the IFR estimates in that study increase exponentially as a function of age, with rates near zero for ages 0-39 and far higher rates for older adults.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 4: Comorbidity Prevalence in Hospitalized Patients vs. General Population

Comorbidities	NYC Hospitalized COVID Patients	NYC Population (Ages 50+)	Difference
Obesity	42%	27%	15%
Diabetes	32%	28%	4%
Cardiovascular Disease	23%	22%	1%
Cancer	6%	7%	-1%
Chronic Respiratory Disease	5%	6%	-1%
Kidney Disease	4%	7%	-3%
HIV	1%	3%	-2%
Liver Disease	0.2%	0.5%	-0.3%

Sources: see Appendix D.

4. Discussion

While age and fatality risk are closely related, it is very unlikely that age alone explains differences in IFR across regions and populations. The remaining variation may be explained by either comorbidities or other population demographics. Section 4.1 discusses the effect of co-morbidities on fatality risk. Section 4.2 discusses how other demographic and socioeconomic factors may explain variations in IFR across locations.

4.1 Comorbidities

Researchers have debated whether certain conditions predispose patients to have more severe cases of COVID-19. A recent study in New York City reported the incidence of several chronic medical conditions in hospitalized COVID-19 patients; the median age of the patients in that sample was 59 years.[45] Table 4 compares those results to the average incidence of comorbidities among New York City residents ages 50 years and above – that is, the relevant age group for comparison with the sample of hospitalized patients.[46] Most comorbidities, including

[45] See Richardson et al. (2020).
[46] New York City does not publish data on the incidence of x and y, and hence Table 4 uses U.S. data to gauge the prevalence of those comorbidities.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 5: Fatality Hazard Ratios for Hospitalized U.K. COVID-19 Patients

Age	Hazard Ratio	Comorbidity	Hazard Ratio
20 to 49	1	Diabetes	1.1
50 to 59	2.7	Malignant Cancer	1.1
60 to 69	5.5	Chronic Cardiac Disease	1.2
70 to 79	9.8	Chronic Pulmonary Disease	1.2
80+	13.5	Chronic Kidney Disease	1.3
		Obesity	1.3
		Liver Disease	1.5

Source: Doherty et al. (2020), Figure 5.

hypertension and diabetes, do not explain which infected individuals are likely to require hospitalization for COVID-19. The only striking positive comorbidity was obesity. However, in NYC, obesity is also much more prevalent among low-income groups who are more likely to live in densely populated neighborhoods and to work in high-exposure jobs. Thus, it is quite possible that NYC's obese population has a higher infection rate, which would explain the higher hospitalization rate for obese individuals. A more detailed version of Table 4 can be found in Appendix D.

A recent U.K. study of hospitalized COVID-19 patients analyzed the impact of comorbidities on the risk of a fatal outcome.[47] Table 5 shows the most relevant portion of their results. Evidently, age influences the estimated hazard ratio far more than any specific comorbidity. In effect, comorbidities appear to be a scaling factor that has a noticeable impact on the hazard ratio compared to a healthy peer with no comorbidities, but that impact is much smaller than the impact of higher age. For example, the fatality risk for an obese 40-year-old hospital patient is moderately higher than for a non-obese individual of the same cohort but far lower (less than one-tenth) of the fatality risk for a non-obese 75-year-old hospital patient.

Both of these studies suggest that differences in comorbidity across geographical locations is likely to be a relatively modest factor in explaining the dispersion in overall IFRs for COVID-19, especially compared to the very strong link between IFR and age.

[47] See Doherty et. al (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

4.2 Demographic and Socioeconomic Factors

Our results indicate that age is a crucial determinant of infection fatality rates for COVID-19. However, our meta-analysis has not directly considered the extent to which IFRs may vary with other demographic factors, including race and ethnicity. Fortunately, some valuable insights can be garnered from other recent studies. In particular, one recent seroprevalence study of residents of two urban locations in Louisiana found no significant difference in IFRs between whites and Blacks.[48]

Nonetheless, the incidence of COVID-19 mortality among people of color is extraordinarily high due to markedly different infection rates that reflect systematic racial and ethnic disparities in housing and employment. For example, a recent infection study of a San Francisco neighborhood found that 80% of positive cases were Latinx – far higher than the proportion of Latinx residents in that neighborhood.[49] That study concluded as follows: *"Risk factors for recent infection were Latinx ethnicity, inability to shelter-in-place and maintain income, frontline service work, unemployment, and household income less than $50,000 per year."* Recent CDC analysis has reached similar conclusions, attributing elevated infection rates among Blacks and Hispanics to dense housing of multi-generational families, increased employment in high-contact service jobs, high incidence of chronic health conditions, and lower quality of health care.[50]

In summary, while the present study has investigated the effects of age on the IFR of COVID-19, further research needs to be done on how infection and fatality rates for this disease are affected by demographic and socioeconomic factors.

[48] See Feehan et al. (2020).
[49] See Chamie et al. (2020).
[50] See Azar et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 6: Age-Specific Fatality Rates for COVID-19 vs. Accidental Injuries

Age Group	COVID-19 Infection Fatality Rate (%)	Automobile Accident Annualized Fatality Rate (%)	Other Accidental Injury Annualized Fatality Rate (%)
0 to 44	0.01	0.01	0.03
45 to 54	0.14	0.01	0.04
55 to 64	0.48	0.01	0.04
65 to 74	1.65	0.01	0.04
75 to 84	5.72	0.02	0.09
85+	22.5	0.02	0.35

Source: National Center for Health Statistics (2020)

5. Conclusion

Age and fatality risk for COVID-19 are exponentially related. In non-technical terms, COVID-19 poses a very low risk for children and younger adults but is hazardous for middle-aged adults and extremely dangerous for elderly people. Table 6 contextualize these risks by comparing the age-specific IFRs from our meta-regression analysis to the annualized risk of a fatal auto accident or other accidental injury. For the youngest age groups, the risk from COVID-19 is broadly comparable to those everyday activities. By contrast, for adults ages 55 to 64, the COVID-19 fatality risk is roughly 50 times greater than the risk of driving a car, and that hazard ratio rises to about 1000:1 for people ages 65 to 74. Moreover, as discussed in Section 4, comorbidities have only modest effects on these risks; that is, being in good health does not necessarily ensure that a middle-aged or older adult will survive a COVID-19 infection.

Our analysis facilitates comparisons between the COVID-19 pandemic and the Spanish Flu pandemic of 1918-20. The U.S. CDC estimates that about 28 percent of the U.S. population was infected by the Spanish Flu and that the death toll was about 675,000. However, that disease was most dangerous for young adults, with an IFR of about 4 percent for people ages 20 to 40 years old, who comprised roughly one-third of the U.S. population at that time. By contrast, COVID-19 is far more dangerous for middle-aged and older adults, whereas the Spanish Flu caused relatively few deaths among those age groups.

Our meta-regression analysis also confirms that COVID-19 is far more deadly than seasonal flu, especially for older adults and elderly people. For example, the U.S. CDC estimates that during winter 2018-19 influenza was associated with about 50 million infections and 34,000 fatalities,

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 7: U.S. Scenario Analysis

Scenario	Infection Rate by Age (*percent*) All	0-49	50-64	65+	Deaths *(thousands)*	IFR *(percent)*
Baseline	*6.4*	*7.6*	*4.5*	***3.7***	***145***	*0.7*
Scenario #1: *current pattern of age-specific prevalence*	28	35	18	14	**465**	0.5
Scenario #2: *uniform prevalence*	28	28	28	28	**895**	1.0
Scenario #3: *protection of vulnerable age groups*	28	39	9	7	**241**	0.3

that is, an overall IFR of about 0.07 percent. By comparison, recent seroprevalence data from U.S. public health laboratories indicates that COVID-19 had infected more than 20 million people by the third week of June, that is, about 6.4% of the U.S. population.[51] Cumulative U.S. fatalities reached nearly 150,000 as of July 21 (four weeks after the date of the seroprevalence data, appropriately reflecting time lags as discussed in Section 2). These figures indicate that the overall IFR of COVID-19 is currently about 0.7%, in line with recent guidance from the CDC.[52] That IFR indicates that COVID-19 is roughly ten times more deadly than the seasonal flu.

Nonetheless, the current level of the overall U.S. IFR should *not* be interpreted as a fixed parameter. Rather, our meta-analysis clearly underscores the rationale for public health measures and communications aimed at reducing the aggregate IFR by mitigating the incidence of new COVID-19 infections among middle-aged and older adults.[53]

To illustrate these considerations, Table 7 outlines three alternative scenarios for the U.S. trajectory of COVID-19 infections and fatalities. All three scenarios assume that the infection rate continues rising to a plateau of 28%, matching the U.S. prevalence of the Spanish Flu. However, the age-specific infection rates vary markedly across the three scenarios:

[51] See U.S. CDC (2020c). Seroprevalence estimates are reported in the U.S. CDC's Weekly COVID Surveillance Summary, based on data collected by 85 state and local public health laboratories spanning the entire country. These reports include age-specific seroprevalence but no details regarding sample selection, test characteristics, or confidence intervals and hence could not be used in our meta-regression analysis.
[52] See U.S. CDC (2020e).
[53] See Davies et al. (2020), De Salazar et al. (2020), Gleser, Gorback, and Redding (2020), Harris (2020),

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

- Scenario #1 assumes that age-specific prevalence will remain similar to the average pattern that has prevailed to date, as indicated by seroprevalence data from U.S. public health laboratories.

- Scenario #2 assumes that the prevalence will eventually become uniform across all age groups, similar to the Spanish Flu pandemic.

- Scenario #3 assumes that public health measures and communications will restrain the incidence of new infections among middle-aged and older adults while prevalence continues rising rapidly among children and younger adults.

To assess the implications of these three alternative assumptions, we use the age-specific IFRs from our meta-regression analysis to determine the death toll for each age group as follows:

$$Deaths_{age} = Population_{age} \times Infection\ Rate_{age} \times IFR_{age}$$

Scenario #1 shows that, if the current age-specific infection pattern continues until 28% of the U.S. is infected, deaths will increase by a factor of 3 to around 450,000. The outcome is far worse in Scenario #2, where the virus spreads uniformly across age groups and causes nearly one million deaths. In contrast, Scenario #3 is associated with a far lower proportion of older adults contracting the virus; thus, most of the growing prevalence occurs among children and younger adults, and the death toll is held to about 240,000.

This scenario analysis underscores the possibility that the United States could plausibly end up with an overall infection fatality rate of about 1%, similar to the outcome in New York City (as shown in Table 1). Such an outcome would be associated with a total death toll of nearly 1 million people. Simply maintaining the current pattern of prevalence across age groups would be somewhat less catastrophic, but still result in nearly 500,000 deaths. By comparison, policy measures and communications that protect vulnerable age groups could halve the current overall IFR to around 0.3%, and prevent over 200,000 deaths.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Appendix A: Comprehensive List of Seroprevalence Studies

A.1 Benchmark Studies

Location	Date	Sample Size	Test Method	Fatality Estimation Procedure
Belgium [54]	April 20-26	3,397	IgG ELISA	Excess mortality relative to pre-COVID baseline
Geneva [55]	April 6 – May 9	2,766	IgG ELISA	Bayesian framework
Spain [56]	April 27 – May 11	35,883	IgG/IgM LFA & CMA	Excess mortality relative to pre-COVID baseline
Sweden [57]	April 27 – May 24	4,800	IgG MBA	Reported Fatalities as of June 18

Note: CMA = chemiluminescent microparticle assay; ELISA = enzyme-linked immunosorbent assay; LFA = lateral flow analysis; MBA = multiplex bead array; *MIA = microsphere immunoassay*.

[54] See Molenberghs et al. (2020), Table 6. This study used the seroprevalence findings of Herzog et al. (2020).
[55] See Perez-Saez et al. (2020), Table S2. This study used the seroprevalence findings of Stringhini et al. (2020).
[56] Age-specific IFRs were constructed using the seroprevalence findings of Pollán et al. (2020), Table S7 (both tests positive) and excess mortality data for Week 25 reported by Spain National Institute of Statistics (2020).
[57] See Sweden Public Health Authority (2020a,b,c,d,e) for information about the seroprevalence program design, antibody test standards, results for weeks 18 to 21, and COVID-19 fatalities as of week 24, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A.2 Other Seroprevalence Studies

Location	Date	Sample Size	Test Method	Fatality Estimation Procedure
Connecticut [58]	April 26 – May 3	1,431	IgG ELISA	Reported Fatalities as of May 28 [59]
Florida [48] (South Region)	April 6 - 10	1,742	IgG ELISA	Reported Fatalities as of May 6 [60]
Missouri [48]	April 20 – 26	1,882	IgG ELISA	Reported Fatalities as of May 23 [61]
New York State [62]	April 19-28	15,101	IgG MIA	Reported Fatalities as of May [63]
Utah [48]	April 20 – May 3	1,132	IgG ELISA	Reported Fatalities as of May 25 [64]
Washington [48] (Pugent Sound Region)	March 23 – April 1	3,264	IgG ELISA	Reported Fatalities as of April 26 [65]

[58] See Havers et al. (2020). Population data for each study region by single year of age as of July 1, 2019 was obtained from U.S. Vital Statistics System (2020). Some seroprevalence age brackets were adjusted (+/- 5 years) to match the age structure of each state's COVID-19 fatality report; see the technical appendix for further details.
[59] See Connecticut Department of Health & Human Services (2020).
[60] See Florida Department of Health (2020).
[61] See Utah Department of Health (2020).
[62] See Rosenberg et al. (2020). Population data by single year of age as of July 1, 2019 was obtained from U.S. Vital Statistics System (2020). Some seroprevalence age brackets were adjusted (+/- 5 years) to match the age structure of New York's COVID-19 fatality report; see the technical appendix for further details.
[63] See New York Department of Health (2020) and Leffler and Hogan (2020).
[64] See Missouri Department of Health & Senior Services (2020).
[65] See Washington Department of Health (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A.3 Comprehensive Tracking & Tracing of COVID-19 Infections

Location	Reporting Date	Population	Cases	Fatalities
Iceland[66]	June 14	341,250	1,734	10
New Zealand [67]	July 9	1,910,760	1,417	22
Republic of Korea[68]	July 11	51,269,183	10,086	262

A.4 Small-Scale Studies

Location	Date	Sample Size	Test Method	Fatality Estimation Procedure
Castiglione d'Adda, Italy [69]	May 18 - 25	509	IgG CLA	Excess Mortality [70] (January 1 – May 31)
Gangelt, Germany [71]	March 31 – April 6	919	IgG ELISA	Reported Fatalities as of May 29 [72]
Diamond Princess Cruise Ship [73]	February 1 – March 7	712	PCR	Reported Fatalities as of June 26 [74]

[66] See Iceland Directorate of Health (2020) for finalized data thru June 14, when Iceland had 1,796 recovered cases, 10 fatalities, and 4 individuals in isolation (none hospitalized).
[67] See New Zealand Ministry of Health (2020).
[68] See Korea Center for Disease Control (2020).
[69] See Pagani et al. (2020).
[70] See Italy National Institute of Statistics (2020a,b) for Castiglione d'Adda population by age and excess mortality by age in 2020 compared to the average mortality during the same calendar dates in 2015 to 2019, respectively.
[71] See Streeck et al. (2020).
[72] See Kreis Heinsberg District Administration (2020) and Stat Germania (2020) for Gangelt COVID-19 fatalities and population by age, respectively.
[73] See Mizumoto et al. (2020), Russell et al. (2020), Leffler and Hogan (2020), and Salje et al. (2020a,b).
[74] See Japan National Institute for Infectious Diseases (2020), Mizumoto et al. (2020), and Salje et al. (2020a,b).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A.5 Studies Excluded Due to Transparency Concerns

Location	Transparency Concern
Denmark[75]	The seroprevalence test kit used in this study was subsequently withdrawn from the market by the manufacturer due to reliability concerns.[76]
Indiana[77]	Indiana University issued two press releases regarding initial findings, but as of July 15, no report had been issued regarding the test procedure, sampling data frame, or construction of test-adjusted confidence intervals.
Slovenia[78]	Public officials issued a press release regarding initial findings, but as of July 15, no report had been issued regarding the test procedure, sampling data frame, or construction of test-adjusted confidence intervals.

A.6 Studies Excluded Due to Sample Selection Bias

(a) Studies of Blood Donors

Rationale: Prior research has shown that blood donors tend to be younger and healthier than the general population, with very few blood donors over age 60. Moreover, individuals who donate blood during a pandemic may be more gregarious and less risk-averse than non-donors. Recent U.K. seroprevalence studies indicated that English blood donors had a COVID-19 infection rate of 7.9% compared to a rate of 5.4% for a random sample of the English population.

Excluded Studies: (1) Apulia, Italy.[79] (2) Denmark.[68] (3) England.[80] (4) Milan, Italy.[81] (5) Netherlands.[82] (6) Scotland.[83] (7) San Francisco, CA.[84] (8) Zurich, Switzerland.[85]

[75] See Erikstrup et al. (2020).
[76] https://www.reuters.com/article/us-health-coronavirus-denmark-kits/denmark-to-send-back-inaccurate-antibody-tests-from-chinas-livzon-idUSKBN22W2TC
[77] See Indiana Department of Health (2020).
[78] See Slovenia Government Communication Office (2020).
[79] See Fiorel et al. (2020).
[80] See England Public Health (2020).
[81] See Valenti et al. (2020).
[82] See Slot et al. (2020).
[83] See Thompson et al. (2020).
[84] See Ng et al. (2020).
[85] See Emmenegger et al. (2020); this study covers two distinct set of samples, one of which is from blood donors.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(b) Hospitals or Outpatient Clinics without Screening of COVID-Related Cases

Rationale: A substantial fraction of individuals seeking health care at a hospital or outpatient clinic may have symptoms related to prior exposure to COVID-19 and hence exhibit a higher prevalence of positive test results compared to a random sample of the general population.

Location	Description
Brooklyn, NY[86]	This study used samples from an outpatient clinic and yielded a much higher infection rate than other seroprevalence studies of the New York metropolitan area.
Kobe, Japan[87]	This study tested for IgG antibodies in 1,000 specimens from an outpatient clinic and found 33 positive cases. However, the study did not screen out samples from patients who were seeking treatment for COVID-related symptoms. Moreover, the study reported raw prevalence and confidence interval but did not report statistics adjusted for test characteristics. The manufacturer (ADS Biotec / Kurabo Japan) has indicated that this test has specificity of 100%, based on a sample of 14 pre-COVID specimens, but that specificity has not been evaluated by any independent study. If the true specificity is 98%, then the adjusted prevalence would not be significant. The authors concluded by noting the selection bias and recommended that "further serological studies targeting randomly selected people in Kobe City could clarify this potential limitation."
Tokyo, Japan[88]	The authors of this study specifically cautioned against interpreting their results as representative of the general population. In particular, the sample of 1,071 participants included 175 healthcare workers, 332 individuals who had experienced a fever in the past four months, 45 individuals who had previously taken a PCR test, and 9 people living with a COVID-positive cohabitant. The study obtained a raw infection rate of 3.8%, but the rate is only 0.8% if those subgroups are excluded.
Zurich, Switzerland[89]	This study analyzed two distinct set of samples: (i) blood donors and (ii) hospital patients. Nearly all blood donors were ages 20 to 55, so that sample is not useful for assessing age-specific IFRs for older adults. The sample of hospital patients was not screened to eliminate cases directly related to COVID-19, so that sample may not be representative of the broader population. Moreover, inhabitants of the city of Zurich constituted a relatively large fraction of seropositive results compared to residents from the remainder of the canton of Zurich (which is predominantly rural). The study computes an overall IFR of 0.5%, similar to that of Geneva.

[86] See Reifer et al. (2020).
[87] See Doi et al. (2020).
[88] See Takita et al. (2020a,b).
[89] See Emmenegger et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(c) Active Recruitment of Participants

Rationale: With active recruitment, a substantial fraction of the sample may be comprised of individuals who are aware of or concerned about prior exposure to COVID-19 and hence exhibit a higher prevalence of infections compared to a random sample of the general population.

Location	Description
Luxembourg[90]	Of the 35 participants who tested positive, 19 had previously interacted with a person who was known to be infected or had a prior test for SARS-CoV-2.
Boise, Idaho[91]	This study was promoted during a "Crush the Curve" publicity campaign and required participants to sign up for a test.
Santa Clara, CA [70]	Participants were recruited via social media and needed to drive to the testing site. Stanford Medicine subsequently released a statement indicating that the study was under review due to concerns about potential biases.[92]
Frankfurt, Germany[93]	This study was conducted at a industrial worksite. Among the 5 seropositive participants, 3 had prior positive tests or direct contact with a known positive case.

(d) Other Sample Selection Issues

Location	Description
Oisie, France[94]	This sample of 1,340 participants included elementary school teachers, pupils, and their families. Only two individuals in the sample were ages 65 years and above.
Saxony, Germany[95]	This study analyzed specimen samples from students and teachers at thirteen secondary schools in eastern Saxony and found very low seroprevalence (0.6%).

A.7: Studies Excluded Due to Accelerating Outbreak[96]

Rationale: As discussed in Section 2, if a seroprevalence study takes place in the midst of an accelerating outbreak, then there is no preicse way to determine which of the subsequent fatalities resulted from new infections vs. infections prior to the date of the study.

Excluded Studies (see Section 2.2 and Table 4): (1) Los Angeles County.[97] (2) New York City.[98] (3) Santa Clara County.[99] (4) Scotland.[100]

---

[90] See Snoeck et al. (2020).
[91] See Bryan et al. (2020).
[92] https://www.dailymail.co.uk/health/article-8358003/Stanford-researchers-investigation-tipping-scale-antibody-studies.html
[93] See Fraehling et al. (2020).
[94] See Fontanet et al. (2020).
[95] See Armann et al. (2020).
[96] Studies on expatriate flights and Japanese evacuees from Wuhan also occurred extremely early in the outbreak, and infection and deaths in their respective populations rose exponentially following the initial study.
[97] See Sood et al. (2020).
[98] See Havers et al. (2020).
[99] See Bendavid et al. (2020).
[100] See Thompson et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A.8 Studies Excluded due to Lack of Age-Specific Prevalence

Location	Description
British Columbia, Canada[101]	This study analyzed 885 laboratory specimens from outpatient clinics for the period May 15-27 and found only four positive cases (0.6%). This sample is not well-suited for assessing age-specific prevalence or age-specific IFRs.
Czech Republic[102]	The Czech Ministry of Health conducted a large-scale seroprevalence survey on April 23-May 1, collecting specimens from a random sample of 22,316 residents and testing for IgG antibodies using the Wantai test kit. Only 107 positive cases were identified (raw prevalence = 0.4%), and hence the test-adjusted confidence intervals include the lower bound of zero prevalence. That result is consistent with the very low number of reported cases in the Czech Republic as of early May; for example, Prague had only 1,638 reported cases for a population of 1,3 million.
Expatriate Flights[103]	This study performed PCR tests on 689 individuals expatriated from Wuhan, China on six international flights during January 31-February 2. There were six positive tests (raw prevalence = 0.87%), but assessment of age-specific prevalence or IFRs is not feasible given the sample size, low prevalence, and lack of case outcomes.
Japanese Evacuees[104]	This study performed PCR tests on 565 Japanese citizens expatriated from Wuhan, China. There were eight positive tests, indicating a raw prevalence of 1.4%, but assessment of age-specific prevalence or IFRs is not feasible given the small sample, low prevalence, and lack of data on case outcomes.
Jersey (U.K.)[105]	This study collected samples from 629 households comprising 1,062 individuals and estimated seroprevalence at 4.2% (CI 2.9 to 5.5%), indicating that about 3,300 Jersey residents have been infected. Jersey has had 30 COVID-19 fatalities (as of July 15), and hence the overall IFR is about 1% (similar to that of NYC). However, the seroprevalence sample is too small to facilitate accurate assessments of age-specific IFRs; for ages 55+, there were 258 samples and 12 positive cases,
New Orleans, LA[106]	This study analyzed a random sample of 2,640 participants and obtained a seroprevalence estimate of 6.86% and an IFR of 1.63% (CI 1.53 to 1.74%). The study reported race-specific results but not age-specific seroprevalence or IFRs.
Mount Sinai Hospital, New York City[107]	This study analyzed seroprevalence using specimens from four groups of patients (Cardiology, OB/GYN, Oncology, and Surgery) starting in mid-February. For the final week of the study (April 19), positive results were obtained for 47 of 243 patients; that seroprevalence estimate of 19.3% is well-aligned with the results of the New York Department of Health study. However, the sample size of this cohort is too small for assessing age-specific IFRs.

[101] See Skowronski et al. (2020).
[102] See Czech Ministry of Health (2020).
[103] See Verity et al. (2020).
[104] See Nishiura et al. (2020).
[105] See Jersey (U.K.) Health & Community Services (2020a,b).
[106] See Feehan et al. (2020).
[107] See Stadlbauer et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Neustadt-am-Rennsteig, Germany[108]	This study analyzed seroprevalence of 626 residents (71% of the population of this municipality) and estimated seroprevalence of 8.4% (52 positive cases). However, this sample size is too small for assessing age-specific IFRs.
San Francisco Mission District, CA[109]	This study analyzed active infections and seroprevalence of 3,953 residents in a densely population majority Latinx neighborhood in downtown San Francisco. Positive seroprevalence in older adults was very low (22 out of 3,953) and hence too small for assessing age-specific IFRs.
San Miguel County, CO[110]	The San Miguel County Health Department assessed seroprevalence in March and April using samples from 5,283 participants (66% of county residents). Raw prevalence was very low (0.53%), with only 3 confirmed positive results for adults ages 60 years and above.
Vo, Italy[111]	Vo' is a municipality of 3,300 people, nearly all of whom (87%) participated in an infection survey in late February. However, there were only 54 infections among people ages 50+, so assessing age-specific IFRs is not feasible.

[108] See Weis et al. (2020).
[109] See Chamie et al. (2020).
[110] See San Miguel County Department of Health & Environment (2020).
[111] See Lavezzo et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Appendix B: Positive Predictive Value of Seroprevalence Tests

Table B1: Impact of Crude Prevalence on Positive Predictive Value

Location	**Netherlands**			**New York City**		
Crude Prevalence	2.7%			20%		
Test	*Wantai Total Antibody ELISA*			*Wadsworth Pan-Ig*		
		95% Confidence Bounds			**95 % Confidence Bounds**	
	Mean	**Lower**	**Upper**	**Mean**	**Lower**	**Upper**
Sensitivity	0.621	0.520	0.720	0.880	0.805	0.928
Specificity	0.977	0.950	1.000	0.988	0.973	0.995
PPV	0.424	0.224	1.000	0.948	0.882	0.979
NPV	0.989	0.986	0.992	0.971	0.952	0.982
Adjusted Prevalence	**0.006**	0.000	0.052	**0.217**	0.192	0.244

Table B2: Impact of Specificity on Positive Predictive Value

Location	**Santa Clara County**					
Crude Prevalence	1.5%					
Test	*Premier Biotech LFA*					
Source	**Bendavid et al.**			**Whitman et al.**		
		95% Confidence Bounds			**95 % Confidence Bounds**	
	Mean	**Lower**	**Upper**	**Mean**	**Lower**	**Upper**
Sensitivity	0.828	0.760	0.884	0.828	0.760	0.884
Specificity	0.995	0.992	0.997	0.972	0.921	0.994
PPV	0.716	0.591	0.818	0.311	0.128	0.692
NPV	0.997	0.996	0.998	0.997	0.996	0.998
Adjusted Prevalence	0.012	0.009	0.014	0.000	0.000	0.010

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Appendix C: Comparison of Seroprevalence vs. Reported Cases in Iceland

Age Group	Reported Cases	Estimated Infections	*Confidence Interval*		Ratio of Infections to Reported Cases	*Confidence Interval*	
			Lower	*Upper*		*Lower*	*Upper*
30-39	**289**	**469**	*469*	*703*	**1.6**	*1.6*	*2.4*
40-49	**357**	**644**	*473*	*859*	**1.8**	*1.3*	*2.4*
50-59	**306**	**337**	*211*	*547*	**1.1**	*0.7*	*1.8*
60-69	**213**	**225**	*188*	*375*	**1.1**	*0.9*	*1.8*
70-79	**63**	**70**	*63*	*304*	**1.1**	*1.0*	*4.8*
80+	**25**	**26**	*13*	*319*	**1.0**	*0.5*	*12.8*
All 30+	**1,253**	**1,771**	*1,415*	*3,109*	**1.41**	*1.13*	*2.48*

Sources: cases are reported by Iceland Directorate of Health (2020) as of June 14, when Iceland had 1,796 recovered cases, 10 fatalities, and 4 individuals in isolation (none hospitalized). Estimated infections and 95% confidence intervals are taken from the seroprevalence study of Gujbjartsson et al. (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Appendix D: Comorbidities

Table D1: Comorbidity Prevalence in New York City Hospitalized COVID-19 Patients vs. General Population

Comorbidity	NYC Hospitalized COVID Patients	NYC Population (Ages 50+)	Difference
Cancer	5.6%	6.3%	-0.7%
Cardiovascular Disease			
Hypertension	53.1%	49.2%	3.9%
Coronary artery disease	10.4%	10.5%	-0.1%
Congestive heart failure	6.5%	6.9%	-0.4%
Chronic Respiratory Disease			
Asthma	8.4%	8.6%	-0.2%
Chronic obstructive pulmonary disease	5.0%	7.7%	-2.7%
Obstructive sleep apnea	2.7%	2.8%	-0.1%
Immunosuppression			
HIV	0.8%	2.7%	-2.0%
History of solid organ transplant	1.0%	NA	NA
Kidney Disease			
Chronic	4.7%	13.1%	-8.4%
End-Stage	3.3%	0.6%	2.6%
Liver Disease			
Cirrhosis	0.3%	0.9%	-0.6%
Hepatitis B	0.1%	0.5%	-0.3%
Hepatitis C	0.1%	0.1%	0.0%
Metabolic Disease			
Obesity (BMI>=30)	41.7%	26.9%	**14.8%**
Morbid Obesity (BMI>=35)	19.0%	NA	NA
Diabetes	31.7%	27.6%	4.1%
Ever Smoked	15.6%	43.8%	-28.2%

Note: The following sources were used to gauge the prevalence of comorbidities among New York City residents ages 50 years and above. *Asthma*: U.S. Center for Disease Control & Prevention (2018). *Cancer*: New York State Cancer Registry (2016). *Cardiovascular Diseases*: New York Department of Health (2020). *Diabetes*: New York State Comptroller (2015). *HIV*: New York City Department of Health (2018). *Kidney Disease*: IPRO End-Stage Renal Disease Network of New York (2014). *Liver Disease*: Moon et al. (2019) and Must et al. (1999). *Chronic Pulmonary Disease*: New York Department of Health (2019). *Obesity*: New York City Department of Health (2019).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

Acemoglu, D., Chernozhukov, V., Werning, I., Whinston, M. (2020). Optimal Targeted Lockdowns in a Multi-Group SIR Model. NBER Working Paper 27102. *https://www.nber.org/papers/w27102*

AdminStat Germania (2020). Maps, Analysis and Statistics about the Resident Population, Municipality of Gangelt – population by age. *https://ugeo.urbistat.com/AdminStat/en/de/demografia/eta/gangelt/20158575/4*

Armann, J., Unrath, M., Kirsten, C., Luck, C., Dalpke, A., Berner, R. (2020). Anti-SARS-CoV-2 IgG antibodies in adolescent students and their teachers in Saxony, Germany (SchoolCoviDD19): very low seroprevalence and transmission rates. July 17. *https://doi.org/10.1101/2020.07.16.20155143*

Aron, J., Muellbauer, J. (2020, May). Excess mortality: England is the European outlier in the Covid-19. Retrieved July 12, 2020, from https://voxeu.org/article/excess-mortality-england-european-outlier-covid-19-pandemic

Aspelund, K., Droste, M., Stock, J., Walker, C. (2020). Identification and Estimation of Undetected COVID-19 Cases using Testing Data from Iceland. July 19. NBER Working Paper 27528. *https://www.nber.org/papers/w27528*

Azar, K., Shen, Z., Romanelli, R., Lockhart, S., Smits, K., Robinson, S., Brown, S., Pressman, A. (2020). Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California. *Health Affairs.* *https://doi.org/10.1377/hlthaff.2020.00598*

Banerjee, A., Alsan, M., Breza, E., Chandrasekhar, A., Chowdhury, A., Duflo, E., Goldsmith-Pinkham, P., Olken, B. (2020). Messages on COVID-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients with Similar Effects on Non-Recipient Members of Their Communities. July. NBER Working Paper 27496. *https://www.nber.org/papers/w27496*

Bendavid E, Mulaney B, Sood N, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. *https://www.doi.org/10.1101/2020.04.14.20062463*

Bonanad, C., García-Blas, S., Tarazona-Santabalbina, F., Sanchis, J., Bertomeu-González, V., Fácila, L., Ariza, A., Núñez, J., Cordero, A. (2020). The Effect of Age on Mortality of Patients with COVID-19: A metanalysis with 611,583 subjects. *Journal of the American Medical Directors Association.* May 18. *https://doi.org/10.1016/j.jamda.2020.05.045*

Bryan A, Pepper G, Wener MH, Fink SL, Morishima C, Chaudhary A, et al. (2020). Performance Characteristics of the Abbott Architect SARS-CoV-2 IgG Assay and Seroprevalence Testing in Idaho. *https://doi.org/10.1101/2020.04.27.20082362*

Byambasuren, O., Dobler, C., Bell, K., Rojas, D., Clark, J., McLaws, M., Glasziou, P. (2020). Estimating the Seroprevalence of SARS-CoV-2 Infections: Systematic Review. *https://www.medrxiv.org/content/10.1101/2020.07.13.20153163v1*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Carter, L. J., Garner, L. V., Smoot, J. W., Li, Y., Zhou, Q., Saveson, C. J., Sasso J. M., Gregg A.C., Soares D. J., Beskid T. R., Jervey S.R., Liu, C. (2020). Assay Techniques and Test Development for COVID-19 Diagnosis. *ACS Central Science*, *6*(5), 591-605. *https://pubs.acs.org/doi/10.1021/acscentsci.0c00501*

Chamie, G., Marquez, C., Crawford, E., Peng, J., Petersen, M., Schwab, D., Schwab, J., Martinez, J., Jones, D., Black, D., Gandhi, M., Kerkhoff, A., Jain, V., Sergi, F., Jacobo, J., ROjas, S., Tulier-Laiwa, V., Gallardo-Brown, T., Appa, A., Chiu, C., Rodgers, M., Hackett, J., CLIAhub Consortium, Kistler, A., Hao, S.,, Kamm, J., Dynerman, D., Batson, J., Greenhouse, B., DeRisi, J., Havlir, D. (2020). SARS-Cov-2 Community Transmission during Shelter-in-Place in San Francisco. *https://doi.org/10.1101/2020.06.15.20132233*

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Yang Q., Wang J., Liu Y., Wie Y., Xia J., Yu T., Zhang X., Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *The Lancet*, *395*(10223), 507-513. *https://doi.org/10.1016/s0140-6736(20)30211-7*

Chen, V., Zhou, Y., Fallahi, A., Viswanatha, A., Yang, J., Ghasemi, Y., Ohol, N., Rosenberger, J. (2020). An Optimization Framework to Study the Balance between Expected Fatalities due to COVID-19 and the Reopening of U.S. Communities. July 20. *https://doi.org/10.1101/2020.07.16.20152033*

Chowkwanyun, M., Reed A. L. (2020). Racial Health Disparities and Covid-19—Caution and Context. New England Journal of Medicine. Retrieved June 30, 2020. https://www.nejm.org/doi/full/10.1056/NEJMp2012910

Connecticut Department of Health & Human Services (2020). COVID-19 Cases and Deaths by Age Group. *https://data.ct.gov/Health-and-Human-Services/COVID-19-Cases-and-Deaths-by-Age-Group/ypz6-8qyf*

Czech Ministry of Health (2020). Collective Immunity Study SARS-CoV-2: Czech Prevalence. https://covid-imunita.uzis.cz/index.php?pg=vystupy-a-vysledky

Davies, N., Klepac, P., Liu, Y., Prem, K., Jit, M., CMMID COVID-19 working group, Eggo, R. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*. June 16. *https://www.nature.com/articles/s41591-020-0962-9*

Day, M. (2020). Covid-19: Four-fifths of cases are asymptomatic, China figures indicate. *British Medical Journal*, M1375. *https://www.bmj.com/content/369/bmj.m1375*

De Salazar, P., Gomez-Barroso, D., Pampaka, D., Gil, J. Penaver, B., Fernandez-Escobar, C., Lipsitch, M., Larrauri, A., Goldstein, E., Hernan, M. (2020). Lockdown measures and relative changes in the age-specific incidence of SARS-CoV-2 in Spain. July 2. *https://www.medrxiv.org/content/10.1101/2020.06.30.20143560v1*

Doi, A., Iwata, K., Kuroda, H., Hasuike, T., Nasu, S., Kanda, A., Nagao T., Nishioka H., Tomii K., Morimoto T., Kihara, Y. (2020). Estimation of seroprevalence of novel coronavirus disease (COVID-19) using preserved serum at an outpatient setting in Kobe, Japan: A cross-sectional study. *https://www.medrxiv.org/content/10.1101/2020.04.26.20079822v2*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Emmenegger, M., Cecco, E. D., Lamparter, D., Jacquat, R. P., Ebner, D., Schneider, M. M., Morales I.C., Schneider D., Dogancay B., Guo J., Widmer A., Domange J., Imeri M., Moos R., Zografou C., Trevisan C., Gonzalez-Guerra A., Carella A., Dubach I., Xu C., Meisl G., Kosmoliaptsis V., Malinauskas T., Burgess-Brown N., Owens R., Mongjolsapaya J., Hatch S., Screaton G., Schubert K., Huck J., Liu F., Pojer F., Lau K., Hacker D., Probst-Mueller E., Cervia C., Nilsson J., Boyman O., Saleh L., Spanaus K., Eckardstein A., Schaer D., Ban N., Tsai C., Marino J., Schertler G., Ebert N., Thiel V., Gottschalk J., Frey B., Reimann R., Hornemann S., Ring A., Knowles T., Xenarios I., Stuart D., Aguzzi A. (2020). Early plateau of SARS-CoV-2 seroprevalence identified by tripartite immunoassay in a large population. *https://www.medrxiv.org/content/10.1101/2020.05.31.20118554v2*

Erikstrup, C., Hother, C. E., Pedersen, O. B., Mølbak, K., Skov, R. L., Holm, D. K., Saejmose S., Nilsson A. C., Brooks P. T., Boldsen J.K., Mikkelsen C., Gybel-Brask M., Sorenson E., Dinh K. M., Mikkelsen S., Moller B. K., Haunstrup T., Haritshoj L., Jensen B. A., Hjalgrim H., Lillevang S. T., Ullum, H. (2020). Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors. *https://www.medrxiv.org/content/10.1101/2020.04.24.20075291v1*

Fassihi, F., Gladstone, R. (2020). Confusion and Contradiction Reign in Iran's Coronavirus Reporting. *New York Times*, February 28. *https://www.nytimes.com/2020/02/28/world/middleeast/coronavirus-iran-confusion.html*

Feehan, A., Fort, D., Garcia-Diaz, Price-Haywood, E., Velasco, C., Sapp, E., Pevey, D., Seoane, L. (2020). Point Prevalence of SARS-CoV-2 and Infection Fatality Rate in Orleans and Jefferson Parish, Louisiana, May 9-15, 2020. June 24. *https://doi.org/10.1101/2020.06.23.20138321*

Ferguson, N, Laydon, D, Nedjati Gilani, G, Imai, N., Ainslie, K,. Baguelin, M,. Bhatia, S,. Boonyasiri, A., Cucunuba Perez, Z,. Cuomo-Dannenburg, G,. Dighe, A,. Dorigatti, I,. Fu, H,. Gaythorpe, K,. Green, W,. Hamlet, ,. Hinsley, W,. Okell, L,. Van Elsland, S,. Thompson, H,. Verity, R,. Volz, E,. Wang, H,. Wang, Y,. Walker, P,. Walters, Winskill, P,. Whittaker, C,. Donnelly, C,. Riley, S,. Ghani, A (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. March 16. *https://doi.org/10.25561/77482*

Fiorel, J., Centra, M., De Carlo, A., Granato, T., Rosa, A., Sarno, M., De Feo, L., Di Stefano, M., D'Erricol, M., Lo Caputo, S., De Nittis, R., Arena, F., Corso, G., Margaglione, M., Santantonio, T. (2020). Far Away from Herd Immunity to SARS-Cov-2: results from a survey in healthy blood donors in South Eastern Italy. *https://doi.org/10.1101/2020.06.17.20133678*

Flannery, D., Gouma, S., Dhudasia, M., Mukhopadhyay, S., Pfeifer, M., Woodford, E., Gerber, J., Arevalo, C., Bolton, M., Weirick, M., Goodwin, E., Anderson, E., Greenplate, A., Kim, J., Han, N., Pattekar, A., Dougherty, J., Kuthuru, O., Mathew, D., Baxter, A., Vella, L., Weaver, J., Verma, A., Leite, R., Morris, J., Rader, D., Elovitz, M., Wherry, E., Puopolo, K., Hensley, S. (2020). SARS-CoV-2 Seroprevalence among Parturient Women. *https://www.medrxiv.org/content/10.1101/2020.07.08.20149179v1*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Florida Department of Health (2020). Florida COVID-19 Case Line Data. *https://open-fdoh.hub.arcgis.com/datasets/florida-covid19-case-line-data/data*.

Fontanet, A., Tondeur, L., Madec, Y., Grant, R., Besombes, C., Jolly, N., Pellerin S.F., Ungeheuer M. N., Cailleau I., Kuhmel L., Temmam S., Huon C. Chen K. Y., Crescenzo B., Munier S., Demeret C., Grzelek L., Staropoli I., Bruel T., Gallian P., Cauchemez., Werf S., Schwartz O., Eliot M., Hoen, B. (2020). Cluster of COVID-19 in northern France: A retrospective closed cohort study. *https://www.medrxiv.org/content/10.1101/2020.04.18.20071134v1*

Glaeser, D., Gorback, C., Redding, S. (2020). How much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities. July. NBER Working Paper 27519. *https://www.nber.org/papers/w27519*

Gudbjartsson, D., Helgason, A., Jonsson, H., Magnusson, O., Melsted, P., Norddahl, G., Saemundsdottir, J., Sigurdsson, A., Sulem, P., Agustsdottir, A., Eiriksdottir, B., Fridriksdottir, R., Gardarsdottir, E., Georgsson, G., Gretarsdottir, O., Gudmundsson, K., Gunnarsdottir, T., Gylfason, A., Holm, H., Jensson, B., Jonasdottir, A., Jonsson, F., Josefsdottir, K., Kristjansson, T., Magnusdottir, D., le Roux, L., Sigmundsdottir, G., Sveinbjornsson, G., Sveinsdottir, K., Sveinsdottir, M., Thorarensen, E., Thorbjornsson, B., Love, A., Masson, G., Jonsdottir, I., Moller, A., Gudnason, T., Kristinsson, K., Thorsteinsdottir, U., Stefansson, K. (2020). Spread of SARS-Cov-2 in the Icelandic Population. *New England Journal of Medicine*. June 11. *https://www.nejm.org/doi/full/10.1056/NEJMoa2006100*

Grewelle, R., Leo, G. (2020). Estimating the Global Infection Fatality Rate of COVID-19. *https://www.medrxiv.org/content/10.1101/2020.05.11.20098780v1*

Hall, R., Jones, C., Klenow, P. (2020). Trading Off Consumption and COVID-19 Deaths. June. NBER Working Paper 27340. *https://www.nber.org/papers/w27340*

Harris, J. (2020). Data from the COVID-19 epidemic in Florida suggest that younger cohorts have been transmitting their infections to less socially mobile older adults. July 5. *https://doi.org/10.1101/2020.06.30.20143842*

Havers, F. P., Reed, C., Lim, T. W., Montgomery, J. M., Klena, J. D., Hall, A. J., Fry A. M., Cannon D. L., Chiang C. F., Gibbons A., Krapiunya I., Morales-Betoulle M., Roguski K., Rasheed M., Freeman B., Lester S., Mills L., Carroll D. S., Owen S. M., Johnson J. A., Semenova V. A., Group S. G., Schiffer J., Thornburg, N. P. (2020). Seroprevalence of Antibodies to SARS-CoV-2 in Six Sites in the United States, March 23-May 3, 2020. *https://www.medrxiv.org/content/10.1101/2020.06.25.20140384v1*

Herzog, S., Bie, J. D., Abrams, S., Wouters, I., Ekinci, E., Patteet, L., Coppens, A., Spiegeleer, D., Beutels P., Van Damme P., Hens N., Theeten, H. (2020). Seroprevalence of IgG antibodies against SARS coronavirus 2 in Belgium: A prospective cross-sectional study of residual samples. *https://www.medrxiv.org/content/10.1101/2020.06.08.20125179v1*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Higgins, J., Thompson, S., Spiegelhalter, D. (2009). A Re-Evaluation of Random-Effects Meta-Analysis. *Journal of the Royal Statistical Society A*, 172:137-159. *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2667312/*

Iceland Directorate of Health (2020). COVID-19 in Iceland: Statistics for 28 February to 14 June 2020 - Not Updated after 14 June. *https://www.covid.is/data-old*

Indiana State Department of Health (2020). IUPUI, ISDH release findings from Phase 2 of COVID-19 Testing in Indiana. June 17. *https://news.iu.edu/stories/2020/06/iupui/releases/17-fairbanks-isdh-second-phase-covid-19-testing-indiana-research.html*

Ioannidis, J. (2020). The infection fatality rate of COVID-19 inferred from seroprevalence data. June 8. *https://doi:10.1101/2020.05.13.20101253*

IPRO End-Stage Renal Disease Network of New York (2014). Annual Report. *https://esrd.ipro.org/wp-content/uploads/2016/01/Network-2-2014-Annual-Report.pdf*

Italy National Institute of Statistics (2020a). Resident Population by Age on January 1, 2020. *http://demo.istat.it/pop2020/index.html*

Italy National Institute of Statistics (2020b). Mortality Trend for January 1 to May 31 for the Years 2015-2020, with Preliminary data for 2020 using the ANPR System - L'andamento dei decessi nel periodo 1 gennaio - 31 maggio per gli anni 2015-2020, dati anticipatori per l'anno 2020 sulla base del sistema ANPR. *https://public.tableau.com/views/Mortalit_15858412215300/Mortalit*

Japan National Institute for Infectious Diseases (2020). Field Briefing: Diamond Princess COVID-19 Cases, 20 February Update. *https://www.niid.go.jp/niid/en/2019-ncov-e/9407-covid-dp-fe-01.html*

Jerkovic, I., Ljubic, T., Basic, Z., Kruzic, I., Kunac, N., Bezic, J., Vuko A., Markotic A., Andjelinovic, S. (2020). SARS-CoV-2 antibody seroprevalence in industry workers in Split-Dalmatia and Sibenik-Knin County, Croatia. *https://www.medrxiv.org/content/10.1101/2020.05.11.20095158v1*

Jersey (U.K.) Health & Community Services (2020a). Coronavirus (COVID-19) Tests and Cases in Jersey. *https://www.gov.je/Health/Coronavirus/Pages/CoronavirusCases.aspx*

Jersey (U.K.) Health & Community Services (2020b). Prevalence of Antibodies – Community Survey Round 2. June 9. *https://www.gov.je/SiteCollectionDocuments/Government%20and%20administration/R%20Prevalence%20of%20antibodies%2020200609%20SJ.pdf*

Kempen, J., Abashawl, A., Kinfemichael, H., Difabachew, M., Kempen, C., Debele, M., Menkir A., Assefa, M., Asfaw, E., Habtegabriel, L., Addisie, Y., Nilles E., Longenecker, J. (2020). SARS CoV-2 Serosurvey in Addis Adaba, Ethopia. *https://www.medrxiv.org/content/10.1101/2020.06.23.20137521v1.full.pdf*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Khalili, M., Karamouzian, M., Nasiri, N., Javadi, S., Mirzazadeh, A., & Sharifi, H. (2020). Epidemiological Characteristics of COVID-19: A Systemic Review and Meta-Analysis. *https://www.medrxiv.org/content/10.1101/2020.04.01.20050138v1*

Knittel C. R., Ozaltun B. (2020). What Does and Does Not Correlate with COVID-19 Death Rates. *NBER*, June 2020. *https://www.nber.org/papers/w27391*

Kotchen T. (2010). Obesity-Related Hypertension: Epidemiology, Pathophysiology, and Clinical Management. *American Journal of Hypertension*. *https://doi.org/10.1038/ajh.2010.172*

Kraehling, V., Kern, M., Halwe, S., Mueller, H., Rohde, C., Savini, M., Schmidt M., Wilhelm J., Becker S., Ciesek S., Gottschalk, R. (2020). Epidemiological study to detect active SARS-CoV-2 infections and seropositive persons in a selected cohort of employees in the Frankfurt am Main metropolitan area. *https://www.medrxiv.org/content/10.1101/2020.05.20.20107730v1*

Kreis Heinsberg District Administration (2020). Coronavirus in Kreis Heinsberg Current Information. *https://www.kreis-heinsberg.de/aktuelles/aktuelles/?pid=5149*

Korea Center for Disease Control (2020). Weekly Report on the COVID-19 Situation in the Republic of Korea as of July 11, 2020. *https://www.cdc.go.kr/board/board.es?mid=a30501000000&bid=0031&list_no=367822&act=view*

Krähling, V., Kern, M., Halwe, S., Muller, H., Rohde, C., Savini, M., Schmidt, M., Wilhelm, J., Becker, S., Ciesek, S., Gottschalk, R. (2020). Epidemiological study to detect active SARS-CoV-2 infections and seropositive persons in a selected cohort of employees in the Frankfurt am Main metropolitan area. *https://doi.org/10.1101/2020.05.20.20107730*

Larremore, D., Fosdick, B., Bubar, K., Zhang, S., Kissler, S., Metcalf, C., Buckee, C., Grad, Y. (2020). Estimating SARS-CoV-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys. *https://doi.org/10.1101/2020.04.15.20067066*

Lavezzo, E., Franchin, E., Ciavarella, C., Cuomo-Dannenburg, G., Barzon, L., Vecchio, C. D., Rossi L., Manganelli R., Loregian A., Navarin N., Abate D., Sciro M., Merigliano S., Decanale E., Vanuzzo M. C., Saluzzo F., Onelia F., Pacenti M., Parisi S., Carretta G., Donato D., Flor L., Cocchio S., Masi G., Sperduti A., Cattarino L., Salvador R., Gaythorpe K. A. M, Imperial College London COVID-19 Response Team, Brazzale A. R., Toppo S., Trevisan M., Baldo V., Donnelly C. A., Ferguson N. M., Dorigatti I., Crisanti, A. (2020). Suppression of COVID-19 outbreak in the municipality of Vo, Italy *https://www.medrxiv.org/content/10.1101/2020.04.17.20053157v1*

Leffler, C., Hogan, M. (2020). Age-dependence of mortality from novel coronavirus disease (COVID-19) in highly exposed populations: New York transit workers and residents and Diamond Princess passengers. *https://doi.org/10.1101/2020.05.14.20094847*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Malani, A., Soman, S., Ahser, S., Novosad, P., Imbert, C., Tandel, V., Agarwal, A., Alomar, A., Sarker, A., Shah, D., Shen, D., Gruber, J., Sachdeva, S., Kaiser, D., Bettencourt, L. (2020). Adaptive Control of COVID-19 Outbreaks in India: Local, Gradual, and Trigger-Based Exit Paths from Lockdown. July. NBER Working Paper 27532. *https://www.nber.org/papers/w27532*

Manski, C., Molinari, F. (2020). Estimating the COVID-19 Infection Rate: Anatomy of an inference problem. *Journal of Econometrics*. April 15. *https://www.sciencedirect.com/science/article/pii/S0304407620301676*

McAloon, C., Collins, A., Hunt, K., Barber, A., Byrne, A., Butler, F., Casey, M., Griffin, J., Lane, E., McEvoy, D., Wall, P. (2020). The incubation period of COVID-19: A rapid systematic review and meta-analysisof observational research. *https://www.medrxiv.org/content/10.1101/2020.04.24.20073957v1*

McLaren J. (2020). Racial Disparity in COVID-19 Deaths: Seeking Economic Roots with Census Data. *NBER*, June 2020. *https://www.nber.org/papers/w27407*

Meyerowitz-Katz, G., & Merone, L. (2020). A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates. July 7. *https://www.medrxiv.org/content/10.1101/2020.05.03.20089854v4*

Millett G. A., Jones A. T., Benkeser D., Baral S., Mercer L., Beyrer C., Honermann B., Lankiewicz E., Mena L., Crowley J. S., Sherwood J., Sullivan P. (2020). Assessing Differential Impacts of COVID-19 on Black Communities. *https://doi.org/10.1016/j.annepidem.2020.05.003*

Missouri Department of Health & Senior Services (2020). Missouri COVID-19 Dashboard. *http://mophep.maps.arcgis.com/apps/MapSeries/index.html?appid=8e01a5d8d8bd4b4f85add006f9e14a9d*. Reported fatalities as of May 23 obtained using *https://archive.org*.

Modi, C., Boehm, V., Ferraro, S., Stein, G., Seljak, U. (2020). How deadly is COVID-19? A rigorous analysis of excess mortality and age-dependent fatality rates in Italy. *https://www.medrxiv.org/content/10.1101/2020.04.15.20067074v3*

Moon A., Singal, A., Tapper, E. (2019). Contemporary Epidemiology of Chronic Liver Disease and Cirrhosis. *Journal of Clinical and Gastroenterology and Hepatology*. *https://www.cghjournal.org/article/S1542-3565(19)30849-3/abstract*

Moritsugu, K. (2020). How accurate are China's virus numbers? April 1. *https://www.pbs.org/newshour/world/how-accurate-are-chinas-virus-numbers*

Mizumoto, K., Kagaya, K., Zarebski, A., Chowell, G. (2020). Estimating the Asymptomatic Proportion of Coronavirus Disease 2019 (COVID-19) Cases on Board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveillance*, March 12. *https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.10.2000180*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Molenberghs, G., Faes, C., Aerts, J., Theeten, H., Devleesschauwer, B., Sierra, N., Braeye, T., Renard, F., Herzog, S., Lysyne, P., Van der Heyden, J., Van Oyen, H., Van Damme, P., and Hens, N. (2020). "Belgian Covid-19 Mortality, Excess Deaths, Number of Deaths per Million, and Infection Fatality Rates (8 March - 9 May 2020)." *https://doi.org/10.1101/2020.06.20.20136234*

Moritsugu, K. (2020). How accurate are China's virus numbers? April 1. *https://www.pbs.org/newshour/world/how-accurate-are-chinas-virus-numbers*

Must A., Spadano J., Coakley E., Field A., Golditz, G., Dietz W. (1999). The Disease Burden Associated with Overweight and Obesity. *Journal of the American Medical Association*. *https://jamanetwork.com/journals/jama/fullarticle/192030*

Ng, D., Goldgof, G., Shy, B., Levine, A., Balcerek, J., Bapat, S. P., Prostko J., Rodgers M., Coller K., Pearce S., Franz S., Du L., Stone M., Pillai S., Sotomayor-Gonzalez A., Servellita V., Martin C., Granados A., Glasner D. N., Han L. H., Truong K., Qazi D., Hsu E., Gu W., Santos Y. A., Custer B., Green V., Williamson P., Hills N. K., Lu C. M., Reyes K., Hakim J., Sujishi K., Alazzeh F., Pharm L., Oon C., Miller S., Kurtz T., Hakkett J., Simmons G., Busch M. P., Chiu, C. (2020). SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood from the San Francisco Bay Area. *https://www.medrxiv.org/content/10.1101/2020.05.19.20107482v2*

New York City Department of Health (2018). HIV/AIDS Annual Surveillance Report. *https://www1.nyc.gov/assets/doh/downloads/pdf/dires/hiv-surveillance-annualreport-2018.pdf*

New York City Department of Health (2019). Health of Older Adults in New York City. *https://www1.nyc.gov/site/doh/about/press/pr2019/health-of-older-adults-in-new-york-city*

New York City Department of Health. (2020). COVID-19 Data. Downloaded on July 16 from: *https://www1.nyc.gov/site/doh/covid/covid-19-data.page*.

New York Department of Health (2018). Information for Action Report. *https://www.health.ny.gov/statistics/prevention/injury_prevention/information_for_action/docs/2018-08_ifa_report.pdf*

New York Department of Health (2020a). COVID-19 Tracker – Fatalities by Age Group. *https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-Fatalities*.

New York Department of Health (2020b). BRFSS Cardiovascular Disease Brief. *https://www.health.ny.gov/statistics/brfss/reports/docs/2004_brfss_cardiovascular_disease.pdf*

New York State Cancer Registry (2019). *Cancer Incidence and Mortality in New York State, 1976 – 2016.* *https://www.health.ny.gov/statistics/cancer/registry/*

New York State Comptroller (2015). Diabetes in New York State. *https://www.osc.state.ny.us/sites/default/files/reports/documents/pdf/2018-12/health-diabetes-2015.pdf*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

New Zealand Ministry of Health (2020). COVID-19 Current Cases. Downloaded on July 9. *https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-current-situation/covid-19-current-cases#age*

Nishiura, H., Kobayashi, T., Miyama, T., Suzuki, A., Jung, S., Hayashi, K., Kinoshita R., Yang Y., Yuan B., Akhmetzhanov A. R., Linton, N. M. (2020a). Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases, 94*, 154-155. *https://www.ijidonline.com/article/S1201-9712(20)30139-9/pdf*

Nishiura, H., Kobayashi, T., Yang, Y., Hayashi, K., Miyama, T., Kinoshita, R., Linton N. M., Jung S., Yuan B., Suzuki A., Akhmetzhanov, A. R. (2020b). The Rate of Underascertainment of Novel Coronavirus (2019-nCoV) Infection: Estimation Using Japanese Passengers Data on Evacuation Flights. *Journal of Clinical Medicine, 9*(2), 419. *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7074297/*

Nopsopon, T. (2020). COVID-19 Antibody in Thai Community Hospitals. *https://www.medrxiv.org/content/10.1101/2020.06.24.20139188v3*

Perez-Saez, F. J., Lauer, S. A., Kaiser, L., Regard, S., Delaporte, E., Guessous, I., Stringhini S., Azman, A. (2020). Serology-informed estimates of SARS-COV-2 infection fatality risk in Geneva, Switzerland. *https://www.medrxiv.org/content/10.1101/2020.06.10.20127423v1*

Pollán, M., Perez-Gomez, B., Pastor-Barriso, R., Oteo, J. Hernan, M., Perez-Olmeda, M., Sanmartin, J., Fernandez-Garcia, A., Cruz, I., Fernandez de Larrea, N., Molina, M., Rodriguez, Cabrera, F., Martin, M., Merino-Amador, P., Leon Paniagua, J., Munoz-Montalvo, J., Blanco, F., Yotti, R. (2020). Prevalence of SARS-COV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. *Lancet*. *https://doi.org/10.1016/S0140-6736(20)31483-5*

Reifer, J., Hayum, N., Heszkel, B., Klagsbald, I., Streva, V. A. (2020). SARS-CoV-2 IgG Antibody Responses in New York City. *https://www.medrxiv.org/content/10.1101/2020.05.23.20111427v2*

Rinaldi, G., Paradisi, M. (2020). An empirical estimate of the infection fatality rate of COVID-19 from the first Italian outbreak. *https://www.medrxiv.org/content/10.1101/2020.04.18.20070912v2*

Rosenberg, E., Tesoriero, J., Rosenthal, E., Chung, R., Barranco, M., Styer, L., Parker, M., Leung, S., Morne, J., Greene, D., Holtgrave, D., Hoefer, D., Kumar, J., Udo., T., Hutton, B., Zucker, H. (2020). Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York. *Annals of Epidemiology*. *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7297691/*

Russell, T., Hellewell, J., Jarvis., C., van Zandvoort, K., Abbott, S., Ratnayake, R., CMMID COVID-19 working group, Flasche, S., Eggo, R., Edmunds, W., Kucharski, A. (2020). Estimating the Infection and Case Fatality Ratio for Coronavirus Disease (COVID-19) using Age-Adjusted Data from the Outbreak on the Diamond Princess Cruise Ship, February 2020. *Euro Surveillance*, March 26. *https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.12.2000256*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Salje, H., Kiem, C., Lefrancz, N., Courtejoie, N., Bosetti, P., Paireau, J., Andronico, A., Hoze, N., Richet, J., Dubost, C.-L., Le Strat, Y., Lessler, J., Levy-Bruhl, D., Fontanet, A., Opatowski, L., Boelle, P.-Y., Cauchemez, S. (2020a). Estimating the Burden of SARS-Cov-2 in France. *Science*, June 26. *https://doi.org/10.1126/science.abc3517*

Salje, H., Kiem, C., Lefrancz, N., Courtejoie, N., Bosetti, P., Paireau, J., Andronico, A., Hoze, N., Richet, J., Dubost, C.-L., Le Strat, Y., Lessler, J., Levy-Bruhl, D., Fontanet, A., Opatowski, L., Boelle, P.-Y., Cauchemez, S. (2020b). Erratum for the Report: Estimating the Burden of SARS-Cov-2 in France. *Science*, June 26. *https://doi.org/10.1126/science.abd4246*

Saltzman, J. (2020). Nearly a third of 200 blood samples taken in Chelsea show exposure to coronavirus. *Boston Globe*, April 17. *https://www.bostonglobe.com/2020/04/17/business/nearly-third-200-blood-samples-taken-chelsea-show-exposure-coronavirus/*

San Miguel County Department of Health & Environment (2020). IgG Antibody Tests: Statistics and Demographics. *https://www.sanmiguelcountyco.gov/590/Coronavirus*

Sethuraman, N., Jeremiah, S., Ryo, A. (2020). Interpreting Diagnostic Tests for SARS-CoV-2. *Journal of the American Medical Association*. *https://jamanetwork.com/journals/jama/fullarticle/2765837*

Silveira, M., Barros, A., Horta, B., Pellanda, L., Victora, G., Dellagostin, O., Struchiner C., Burattini M., Valim A., Berlezi E., Mesa J., Ikeda M. K., Mesenburg M., Mantesso M., Dall'Agnol M., Bittencourt R., Hartwig F. P., Menezes A. M., Barros F. C., Hallal P., Victora, C. G. (2020). Repeated population-based surveys of antibodies against SARS-CoV-2 in Southern Brazil. https://www.medrxiv.org/content/10.1101/2020.05.01.20087205v2

Shakiba M, Nazari S, Mehrabian F, et al. Seroprevalence of COVID-19 virus infection in Guilan province, Iran. medRxiv doi: *https://doi.org/10.1101/2020.04.26.20079244*

Skowronski, D., Sekirv, I., Sabaiduc, S., Zou, M., Morshed, M., Lawrence, D., Smolina, K., Ahmed, M., Galanis, E., Fraser, M., Singali, M., Naus, M., Patrick, D., Kaweski, S., Mills, C., Reyes, R., Kelly, M., Levett, P., Petric, M., Henry, B., Krajden, M. (2020). Low SARS-CoV-2 seroprevalence based on anonymized residual sero-survey before and after first-wave measures in British Columbia, Canada, March-May 2020. July 13. *https://www.medrxiv.org/content/10.1101/2020.07.13.20153148v1*

Slot, E., Hogema, B. M., Reusken, C. B., Reimerink, J. H., Molier, M., Karregat, J. H., Ijlst J., Novotny V. M. J., Lier R. A. W., Zaaijer, H. L. (2020). Herd immunity is not a realistic exit strategy during a COVID-19 outbreak. https://www.researchsquare.com/article/rs-25862/v1

Slovenia Government Communication Office. (2020). First study carried out on herd immunity of the population in the whole territory of Slovenia. May 6. Downloaded on July 4 from: *https://www.gov.si/en/news/2020-05-06-first-study-carried-out-on-herd-immunity-of-the-population-in-the-whole-territory-of-slovenia/*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Snoeck, C. J., Vaillant, M., Abdelrahman, T., Satagopam, V. P., Turner, J. D., Beaumont, K., Gomes C. P. C., Fritz J. V., Schroder V. E., Kaysen A., Pavelka L., Stute L., Meyers G. R., Pauly L., Hansen M., Pauly C., Aguayo G. A., Perquin M., Hanff A. M., Ghosh S., Gantenbein M., Huiart L., Ollert M., Krüger, R. (2020). Prevalence of SARS-CoV-2 infection in the Luxembourgish population: The CON-VINCE study. *https://www.medrxiv.org/content/10.1101/2020.05.11.20092916v1*

Sood, N., Simon, P., Ebner, P., Eichner, D., Reynolds, J., Bendavid, E., Bhattacharya, J. (2020). Seroprevalence of SARS-CoV-2–Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020. *JAMA-Journal of the American Medical Association*. *https://jamanetwork.com/journals/jama/fullarticle/2766367*

Spain National Institute of Statistics – Instituto Nacional de Estadistica (2020). Estimación del Número de Defunciones Semanales durante el Brote de Covid-19: diferencia absoluto del acumulado en lo que va de año. Downloaded on July 1 from: *https://www.ine.es/jaxiT3/Tabla.htm?t=35179*.

Stadlbauer, D., Tan, J., Jiang, K., Hernandez, M., Fabre, S., Amanat, F., Teo, C., Arunkumar, G., McMahon, M., Jhang, J., Nowak, M., Simon, V., Sordillo, E., van Bakel, H., Krammer, F. (2020). Seroconversion of a City: Longitudinal monitoring of SARS-CoV-2 seroprevalence in New York City. June 29. *https://doi.org/10.1101/2020.06.28.20142190*

Streeck, H., Schulte, B., Kümmerer, B., Richter, E., Höller, T., Fuhrmann, C., Bartok, E., Dolscheid, R., Berger, M., Wessendorf, L., Eschbach-Bludau, M., Kellings, A., Schwaiger, A., Coenen, M., Hoffmann, P., Stoffel-Wagner, B., Nöthen, M., Eis-Hübinger, A., Exner, M., Schmithausen, R., Schmid, M, Hartmann, G. (2020). Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. June 2. *https://doi.org/10.1101/2020.05.04.20090076*

Stringhini S, Wisniak A, Piumatti G, Azman AS, Lauer SA, Baysson H, et al. Repeated seroprevalence of anti 1 -SARS-CoV-2 IgG antibodies in a population-based sample. *https://doi.org/10.1101/2020.05.02.20088898*

Sweden Public Health Authority – Sveriges Folkhälsomyndighetens (2020a). Study Plan for Assessing the Spread of COVID-19 by Measuring SARS-COV-2 Seroprevalence in a Demographic Cross-Section - Studieupplägg för Folkhälsomyndighetens undersökning av seroprevalens mot SARS-CoV-2 för skattning av genomgången covid-19 i ett tvärsnitts av befolkningen. June 13. *https://www.folkhalsomyndigheten.se/contentassets/6561cbb8fcb8435788fb69f5fd1b4356/studieupplagg-seroprevalens-covid-19.pdf*

Sweden Public Health Authority – Sveriges Folkhälsomyndighetens (2020b). Guidance for Antibody Testing - Vägledning för antikroppspåvisning. Version 3, June 30. *https://www.folkhalsomyndigheten.se/contentassets/2c3d8e40926e4bcc942aa640922bb758/vagledning-antikroppspavisning.pdf*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Sweden Public Health Authority – Sveriges Folkhälsomyndighetens (2020c). Seroprevalence of Antibodies following COVID-19 Infection in Blood Samples from Outpatient Care, Interim report 1, updated on June 18 using data thru week 21 - Påvisning av antikroppar efter genomgången covid-19 i blodprov från öppenvården, delrapport 1 - uppdaterad 2020-06-18 med data för prover insamlade vecka 21. June 18. *https://www.folkhalsomyndigheten.se/publicerat-material/publikationsarkiv/p/pavisning-av-antikroppar-efter-genomgangen-covid-19-hos-blodgivare-delrapport-2/*.

Sweden Public Health Authority – Sveriges Folkhälsomyndighetens (2020d). COVID-19 Report for Week 24 - COVID-19 veckorapport vecka 24. June 18. *https://www.folkhalsomyndigheten.se/globalassets/statistik-uppfoljning/smittsamma-sjukdomar/veckorapporter-covid-19/2020/covid-19-veckorapport-vecka-24_final.pdf* .

Sweden Public Health Authority – Sveriges Folkhälsomyndighetens (2020e). COVID-19 Analytical Statistics. Excel spreadsheet downloaded on June 19 from: *https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/statistik-och-analyser/bekraftade-fall-i-sverige/*.

Sweden Public Health Authority – Sveriges Folkhälsomyndighetens (2020e). The Infection Fatality Rate of COVID-19 in Stockholm – Technical Report. *https://www.folkhalsomyndigheten.se/publicerat-material/publikationsarkiv/t/the-infection-fatality-rate-of-covid-19-in-stockholm-technical-report/*

Takita, M., Matsumura, T., Yamamoto, K., Yamashita, E., Hosoda, K., Hamaki, T., Kusumi, E. (2020a). Challenges of community point-of-care antibody testing for COVID-19 herd-immunity in Japan. *QJM: An International Journal of Medicine*. *https://academic.oup.com/qjmed/article/doi/10.1093/qjmed/hcaa182/5848417*

Takita, M., Matsumura, T., Yamamoto, K., Yamashita, E., Hosoda, K., Hamaki, T., Kusumi, E. (2020b). Regional Difference in Seroprevalence of SARS-CoV-2 in Tokyo: Results from community point-of-care antibody testing. *https://journals.sagepub.com/doi/10.1177/2150132720942695*

Thompson, C., Grayson, N., Paton, R., Lourenço, J., Penman, B., Lee, L. N., Odon V., Mongkolsapaya J., Chinnakannan S., Dejnirattisai W., Edmans M., Fyfe A., Imlach C., Kooblall K., Lim N., Liu C., Lopez-Camacho C., McInally C. A., Ramamurthy N., Ratcliff J., Supasa P., Wang B., Mentzer A. J., Turner M., Semple C., Baillie J. K., ISARIC4C Investigators., Harvala H., Screaton G., Tempterton N., Klenerman P., Jarvis L. Gupta S., Simmonds, P. (2020). Neutralising antibodies to SARS coronavirus 2 in Scottish blood donors - a pilot study of the value of serology to determine population exposure. *https://www.medrxiv.org/content/10.1101/2020.04.13.20060467v1*

U.K. Parliament Office (2020). Antibody Tests for COVID-19. June 16. *https://post.parliament.uk/analysis/antibody-tests-for-covid-19/*

U.S. Food & Drug Administration (2020). EUA Authorized Serology Test Performance. July 1. *https://www.fda.gov/medical-devices/emergency-situations-medical-devices/eua-authorized-serology-test-performance*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

U.S. Center for Disease Control and Prevention (2018). Asthma Data by State or Territory. *https://www.cdc.gov/asthma/most_recent_data_states.htm*

U.S. Center for Disease Control and Prevention (2020a). Interim Guidelines for COVID-19 Antibody Testing. May 23. *https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html*

U.S. Center for Disease Control and Prevention (2020b). COVID-19 in Racial and Ethnic Minority Groups. July 18. *https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html*

U.S. Center for Disease Control and Prevention (2020c). COVID-19 Case Updates. *https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html*

U.S. Center for Disease Control and Prevention (2020d). COVIDView Weekly Report: U.S. Virologic Surveillance - Public Health Laboratories. *https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html*

U.S. Center for Disease Control and Prevention (2020e). COVID-19 Pandemic Planning Scenarios. Updated July 10. *https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html*

U.S. National Vital Statistics System (2020). Bridge-Race Population Estimates – Data Files and Documentation. *https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm*

Utah Department of Health (2020). COVID-19 Surveillance. *https://coronavirus-dashboard.utah.gov/#demographics*.

Valenti, L., Bergna, A., Pelusi, S., Facciotti, F., Lai, A., Tarkowski, M., Berzuini A., Caprioli F., Santoro L., Baselli G., Ventura C. D., Erba E., Bosari S., Galli M., Zehender G., Prati, D. (2020). SARS-CoV-2 seroprevalence trends in healthy blood donors during the COVID-19 Milan outbreak. *https://www.medrxiv.org/content/10.1101/2020.05.11.20098442v2*

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg G., Thompson H., Walker P., Fu H., Dighe A., Griffin J., Cori A., Baguelin M., Bhatia S., Boonyasiri A., Cucuniba Z. M., Fitzjohn R., Gaythorpe K. A. M., Green W., Hamlet A., Hinsley W., Laydon D., Nedjati-Gilani G., Riley S., van-Elsand S., Volz E., Wang H., Wang Y., Xi X., Donnelly C., Ghani A., Ferguson, N. (2020). Estimates of the severity of COVID-19 disease. *https://www.medrxiv.org/content/10.1101/2020.03.09.20033357v1*

Washington Department of Health (2020). 2019 Novel Coronavirus Outbreak (COVID-19) Cases and Deaths by Week of Illness Onset, County, and Age. *https://www.doh.wa.gov/Portals/1/Documents/1600/coronavirus/data-tables/PUBLIC_CDC_Event_Date_SARS.xlsx*

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20160895; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Weis, S., Scherag, A., Baier, M., Kiehntopf, M., Kamradt, T., Kolanos, S., Ankert, J., Glockner, S., Makarewicz, O., Hagel, S., Bahrs, C., Kimmig, A., Proquitte, H., Guerra, J., Loffler, B., Pletz, M., CoNAN study group (2020). Seroprevalence of SARS-CoV-2 antibodies in an entirely PCR-sampled and quarantined community after a COVID-19 outbreak: The CoNAN study. July 17. *https://doi.org/10.1101/2020.07.15.20154112*

Wu, J., Leung, K., Bushman, M., Kishore, N., Niehus, R., Salazar, P. M., Cowling B. J., Lipsitch M., Leung, G. M. (2020). Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, *26*(4), 506-510. *https://www.nature.com/articles/s41591-020-0822-7*

Yancy, C. W. (2020). COVID-19 and African Americans. *Journal of the American Medical Association*, *323*(19), 1891. *https://jamanetwork.com/journals/jama/fullarticle/2764789*

Zeek, A., Briggs, A. (2020). COVID-19 study's preliminary findings show spread of virus in Indiana. May 13. *https://news.iu.edu/stories/2020/05/iupui/releases/13-preliminary-findings-impact-covid-19-indiana-coronavirus.html*