BMJ Global Health

Assessing the burden of COVID-19 in developing countries: systematic review, meta-analysis and public policy implications

Andrew T Levin ,[1,2] Nana Owusu-Boaitey,[3] Sierra Pugh,[4] Bailey K Fosdick,[5] Anthony B Zwi,[6] Anup Malani ,[7] Satej Soman ,[8] Lonni Besançon,[9] Ilya Kashnitsky,[10] Sachin Ganesh,[11] Aloysius McLaughlin,[11] Gayeong Song,[11] Rine Uhm,[11] Daniel Herrera-Esposito,[12] Gustavo de los Campos,[13] Ana Carolina Peçanha Antonio ,[14] Enyew Birru Tadese,[15] Gideon Meyerowitz-Katz [16,17]

To cite: Levin AT, Owusu-Boaitey N, Pugh S, *et al*. Assessing the burden of COVID-19 in developing countries: systematic review, meta-analysis and public policy implications. *BMJ Global Health* 2022;**7**:e008477. doi:10.1136/bmjgh-2022-008477

Handling editor Seye Abimbola

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjgh-2022-008477).

Received 9 January 2022
Accepted 5 May 2022

For numbered affiliations see end of article.

Correspondence to
Gideon Meyerowitz-Katz; gideon.meyerowitzkatz@health.nsw.gov.au

ABSTRACT

Introduction The infection fatality rate (IFR) of COVID-19 has been carefully measured and analysed in high-income countries, whereas there has been no systematic analysis of age-specific seroprevalence or IFR for developing countries.

Methods We systematically reviewed the literature to identify all COVID-19 serology studies in developing countries that were conducted using representative samples collected by February 2021. For each of the antibody assays used in these serology studies, we identified data on assay characteristics, including the extent of seroreversion over time. We analysed the serology data using a Bayesian model that incorporates conventional sampling uncertainty as well as uncertainties about assay sensitivity and specificity. We then calculated IFRs using individual case reports or aggregated public health updates, including age-specific estimates whenever feasible.

Results In most locations in developing countries, seroprevalence among older adults was similar to that of younger age cohorts, underscoring the limited capacity that these nations have to protect older age groups. Age-specific IFRs were roughly 2 times higher than in high-income countries. The median value of the population IFR was about 0.5%, similar to that of high-income countries, because disparities in healthcare access were roughly offset by differences in population age structure.

Conclusion The burden of COVID-19 is far higher in developing countries than in high-income countries, reflecting a combination of elevated transmission to middle-aged and older adults as well as limited access to adequate healthcare. These results underscore the critical need to ensure medical equity to populations in developing countries through provision of vaccine doses and effective medications.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Prior meta-analyses of data from high-income countries have shown that the COVID-19 infection fatality rate (IFR) increases exponentially with age while seroprevalence (as measured by antibodies against SARS-CoV-2) has been markedly lower for older adults relative to younger adults.

WHAT THIS STUDY ADDS

⇒ We analyse serology and mortality data from 62 studies of 25 developing countries, and we find that age-stratified IFRs are about two times higher than the benchmark metaregression for high-income countries.

⇒ Indeed, population IFR in developing countries is similar to that of high-income countries, because differences in population age structure are roughly offset by disparities in healthcare access and elevated infection rates among older age cohorts.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings underscore the urgency of disseminating vaccines and effective medications throughout the developing world.

INTRODUCTION

An important unknown during the COVID-19 pandemic has been the relative severity of the disease in developing countries compared with higher-income nations. The incidence of fatalities in many developing countries appeared to be low in the early stages of the pandemic, suggesting that the relatively younger age structure of these countries might have protected them against the harms of the disease. More recently, however, it has become clear that the perceived differences in mortality may have been illusory, reflecting poor vital statistics systems leading to under-reporting of COVID-19 deaths.[1 2] Moreover, relatively low mortality outcomes in developing countries would be starkly different from the typical pattern observed for many

Table 1 Confirmed COVID-19 deaths as of 20 March 2022

Country	Cumulative deaths	Mortality rate per million
USA	971 162	2917.1
Brazil	657 495	3072.5
India	516 510	370.7
Russia	357 234	2448.3
Mexico	322 072	2472.5
Peru	211 865	6351.0
UK	163 658	2399.4
Italy	157 785	2613.7
Indonesia	153 738	556.3
France	141 002	2091.3
Iran	139 610	1641.9
Colombia	139 452	2720.2

Source: Our World In Data.[56]

other communicable diseases, reflecting the generally lower access to good-quality healthcare in these locations.[3 4]

As shown in table 1, mortality attributable to COVID-19 in many developing locations exceeds 2000 deaths per million. Of the 12 nations with the highest number of deaths attributed to COVID-19, eight are developing countries. Furthermore, these statistics may understate the true death toll in a number of lower-income and middle-income countries. Numerous studies of excess mortality have underscored the limitations of vital registration and death reporting, particularly in developing countries.[1 2 5–9] For example, recent studies of India have found that actual deaths from COVID-19 were about 10 times higher than those in official reports.[2 5] Similarly, a study in Zambia found that only 1 in 10 of those who died with COVID-19 symptoms and whose postmortem COVID-19 test was positive were recorded as COVID-19 deaths in the national registry.[10] Strikingly, the continuation of that study has demonstrated the catastrophic impact of COVID-19 in Zambia, raising the overall mortality by as much as 5–10 times relative to a normal year.[11]

There has, however, been a relative dearth of systematic research concerning the early experience of COVID-19 and the associated infection fatality rate (IFR) in developing countries. Previous evaluations have largely focused on assessing these patterns in high-income countries, where high-quality data on seroprevalence and fatalities has been readily available throughout the pandemic.[12 13] In particular, seroprevalence studies conducted in high-income countries in 2020 found low overall prevalence of antibodies to COVID-19 (generally less than 10%),[14] with much lower prevalence among older adults compared with younger cohorts. Analysis of these data has clearly underscored the extent to which the IFR of COVID-19 increases exponentially with age; that is, the disease is far more dangerous for middle-aged and older adults compared with children and young people.[12 13 15] Two prior meta-analytical studies have considered variations in IFR by age but did not consider the possibility that IFR in developing locations might differ systematically from high-income countries due to healthcare quality, access and other socioeconomic factors.[12 16]

Objectives

1. Determine overall prevalence of COVID-19 infection in locations in developing countries.
2. Assess age-specific patterns of seroprevalence in these locations.
3. Estimate age-specific IFRs and compare to benchmark values for high-income countries.
4. Investigate possible reasons for differences in population IFR between locations.

METHODS

To perform this meta-analysis, we collected published papers, preprints and government reports of COVID-19 serology studies for which all specimens were collected before 1 March 2021 and that were publicly disseminated by 17 December 2021. The full search methodology is given in online supplemental appendix 1A. The study was registered on the Open Science Foundation: https://osf.io/edpwv/

We restricted the scope of our analysis to locations in developing countries using the classification system of the International Monetary Fund (IMF); that is, we excluded locations that the IMF classifies as 'high-income countries'.[17] In some contexts developing countries are also described as low-income to middle-income countries or as emerging and developing economies.

Inclusion/exclusion criteria

Our analysis only included studies that had a random selection of participants from a sample frame representative of the general population.[18 19] Consequently, studies of convenience samples—such as blood donors or residual sera from commercial laboratories—were excluded. Such samples are subject to intrinsic selection biases that may vary across different settings and hence would detract from systematic analysis of the data. Indeed, there is abundant evidence from the pandemic that convenience samples provide inaccurate estimates of seroprevalence, with assessments indicating that they are likely to overestimate the true proportion infected.[20 21]

A crucial part of our analysis entailed adjusting raw seroprevalence to reflect the sensitivity and specificity of the particular assay used in each serology study, and to construct credible intervals that reflect uncertainty about assay characteristics as well as conventional sampling uncertainty. Where a reported study did not include that information, we requested it from study authors. Other data needed and extracted for the analysis included start and end dates of specimen collection, the specific assay used and age-specific serology data.

See online supplemental appendix 1b for further details on inclusion and exclusion criteria.

Deaths

For locations with publicly available databases of all individual cases, we tabulated the fatality data to match the age brackets of that serology study, using cumulative fatalities as of 14 days after the midpoint date of specimen collection to reflect the time lags between infection, seropositivity and fatal outcomes. In the absence of individual case data, we searched for contemporaneous public health reports and tabulated cumulative deaths as of 28 days after the midpoint date of specimen collection to incorporate the additional time lags associated with real-time reporting of COVID-19 fatalities (see online supplemental appendix 1d for discussion of death lags).

Matching prevalence estimates with subsequent fatalities is not feasible if a serology study was conducted in the midst of an accelerating outbreak. Therefore, as in previous work,[15] we estimated seroprevalence but did not analyse IFRs for locations where the cumulative death toll increased by three times or more over the 4-week period following the midpoint date of specimen collection. For details, see the online supplemental appendix 1d. In instances where we were not able to match deaths to serology data, or there were accelerating outbreaks, we used this information to look at serology only.

Additionally, we extracted data on excess deaths for all countries that were included in our IFR analysis. We used two primary sources of estimates on excess mortality: the Institute for Health Metrics and Evaluation (IHME)[22] and the World Mortality Dataset (WMD).[1] The IHME produces national or regional estimates of excess mortality for every location included in this review, while the WMD has estimates for a subset of those locations. We then computed the ratio of excess mortality to reported fatalities for each location in order to assess the impact of potential death under-reporting, and calculated adjusted IFRs using excess mortality as the numerator, as well as the ratio between IFRs calculated using reported and excess deaths. We used excess mortality as it is likely to better represent the true burden of COVID-19 accurately in developing nations.[1]

Adjustment for seroreversion

For those assays used in serology studies included in our analysis, we classified each assay's risk of seroreversion (high, medium or low) based on longitudinal serology studies and serological analysis of prior RT-PCR positive cases. For each location for which the assay used in serology was classified as having high risk of seroreversion, we made adjustments to the data on assay sensitivity. See online supplemental appendix 2a for further details.

Statistical analysis

We use a Bayesian modelling framework to simultaneously estimate age-specific prevalence and IFRs for each location in our study. First, we model age-specific prevalence for each location at the resolution of the serology data reported. Then, we model the number of people that test positive in a given study location and age group as coming from a binomial distribution with a test positivity probability that is a function of the true prevalence, sensitivity and specificity, accounting for seroconversion and seroreversion (see the online supplemental appendix 2B).

As in Carpenter and Gelman (2020),[23] we consider sensitivity and specificity to be unknown and directly model the lab validation data (eg, true positives, true negatives, false positives, false negatives) for each test. Independent weakly informative priors are placed on the seroprevalence parameters, and independent, informative priors akin to those in Carpenter and Gelman[23] are placed on the sensitivity and specificity parameters. To avoid assumptions about the variability of prevalence across age within a serology age bin, we aggregate deaths for each location to match their respective serology age bins. Independent mildly informative priors are assumed on the age-group-specific IFR parameters.

Prevalence for a given age group and location is estimated by the posterior mean and equal-tailed 95% credible interval. Uniform prevalence across age is deemed plausible for locations where the 95% credible intervals for the ratio of seroprevalence for age 60 years and older over the seroprevalence estimate for ages 20 years to 60 years contains 1.

IFR calculation and comparison

We model the number of individuals at a given location and age group that are reported as dying of COVID-19 as Poisson distributed with rate equal to the product of the age group IFR, age group population and age group prevalence. For locations where deaths were reported separately for different age bins this model provides IFR estimates for specific age groups and for broader population cohorts, including adults aged 18–65 years. For locations where death data were not disaggregated by age the model provides a population IFR. The model was implemented in the programming language R, with posterior sampling computation implemented with the Stan software package.[24]

To perform a meta-analysis of age-specific IFRs across locations, we conduct a metaregression with random effects. In the metaregression, the dependent variable is the estimated IFR for a specific age group in a specific geographical location, the explanatory variable is the median age of that particular age group, and the SD of each idiosyncratic error is taken from the Bayesian analysis described above. We used a random-effects procedure to allow for residual heterogeneity between studies and across age groups by assuming that these divergences are drawn from a Gaussian distribution. We also allowed for fixed effects by location, to account for locations that deviate from the norm. Since the metaregression used IFR estimates based on reported deaths, we compared the location-specific fixed effects to two estimates of the ratio

of excess mortality to COVID-19 deaths in each location. We also compared these metaregression results to a prior metaregression of age-specific IFR for high-income countries;[15] further details are given in online supplemental appendix 2d and table A5. This was performed using the meta regression procedure in Stata V.17. Finally, we computed population IFRs adjusted for COVID-19 death undercounting and compared these estimates to the proportion of well-certified deaths.

Covariates

We selected covariates that were judged likely to have an impact either on the IFR of COVID-19 itself or on the accuracy of official data on COVID-related mortality based on prior research and expertise. Such covariates included GDP per capita and measures of healthcare capacity; the complete list is provided in online supplemental appendix 1f. Where possible, we extracted these covariates at a state or regional level within a country; otherwise, they were identified at a national level. In instances where a covariate was only available at the national level, we aggregated location-specific seroprevalence and IFRs by weighting each location using the square root of the number of serology specimens collected in that location.

RESULTS

We identified a total of 2384 study records, with 2281 records identified from online databases and a further 124 from Twitter, Google Scholar and a prior publication.[25] After excluding 2062 records, we assessed 343 records and determined that 97 studies satisfied the criteria for inclusion in the final analyses, of which 62 studies (representing a total of 25 developing countries) could be used to produce IFR estimates; see online supplemental appendix 1c for details. The geographical distribution of these studies is shown in figure 1, while table 2 provides a list of the studies used in producing IFR estimates, including the specimen collection dates and the assay used in each study. Further details are provided in our GitHub repository https://covid-ifr.github.io/.

Seroprevalence

As shown in figure 2A and 2B, seroprevalence reached relatively high levels in numerous locations in developing countries during the time frame covered by our analysis. The upper panel shows estimates from studies where specimens were collected between April and September 2020, while the lower panel shows corresponding estimates for the period from October 2020 to February 2021.

In most developing country locations, seroprevalence was roughly uniform across age strata. Figure 3 shows the heatmap of age-specific seroprevalence across all age cohorts. As shown in figure 4, the ratio of seroprevalence for older adults (ages 60+ years) compared with middle-aged adults (ages 40–59 years) is indistinguishable from unity in most of these locations. While many locations had a ratio below 1, the majority of the areas were very substantially above the ratio for higher-income areas (green shaded region), and the point estimates were not markedly below 1, indicating minimal difference in infection rates between older and younger adults in developing nations.

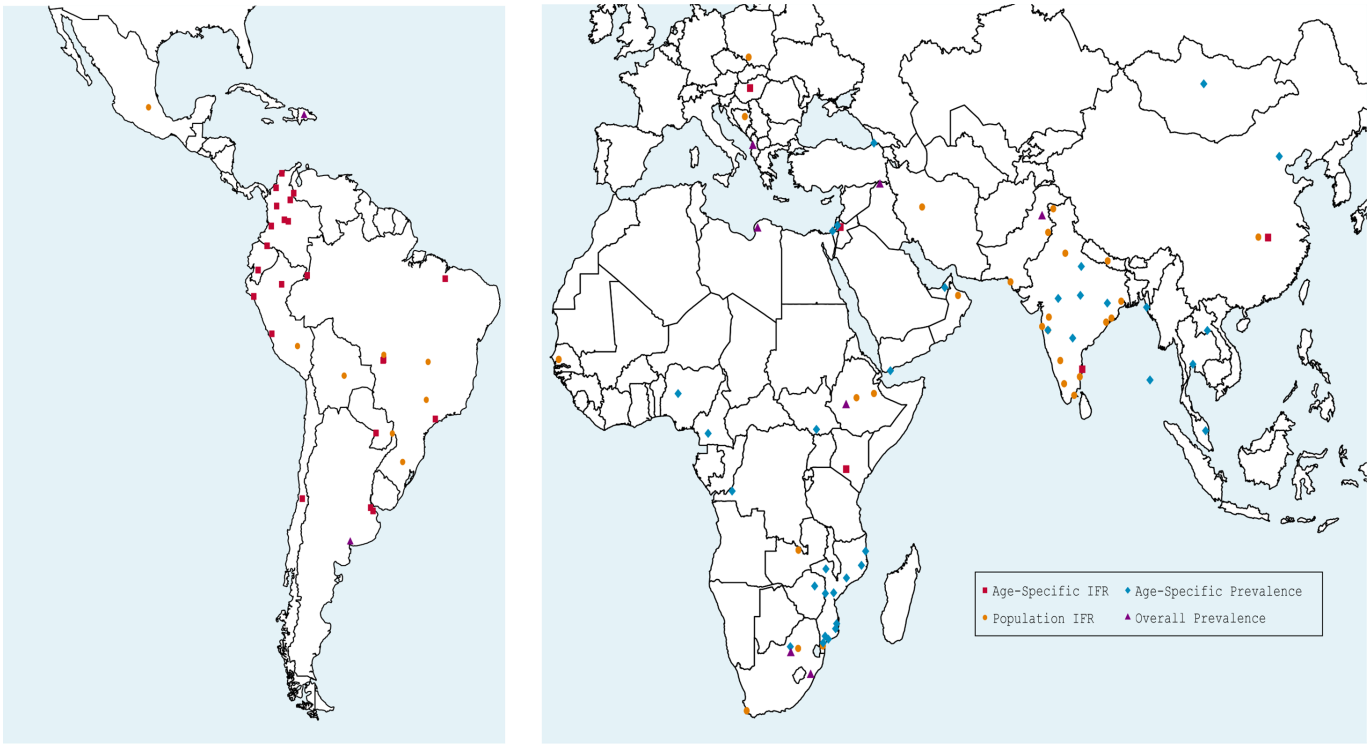


Figure 1 Map of study locations. IFR, infection fatality rate.

Table 2 Included studies for infection fatality rate (IFR)

	Location	Date range	Assay used	Citation
Latin America				
Argentina	Buenos Aires City	10 September to 18 October 2020	COVIDAR IgG	68
	Municipality of Hurlingham	26 November to 10 December 2020	COVIDAR IgG	69
Bolivia	Santa Cruz	22 August to 13 September 2020	Standard Q IgG/IgM	70
Brazil	Cuiabá	16 September to 15 October 2020	DiaSorin Liaison IgG	71
	Distrito Federal	2–17 December 2020	CTK Biotech Onsite IgG/IgM	72
	Foz do Iguaçu	14 May to 9 June 2020	Proprietary	73
	Maranhao	27 July to 8 August 2020	Roche Elecsys IgG/IgM	74
	Mato Grosso	16 September to 15 October 2020	DiaSorin Liaison IgG	71
	Pitangueiras	24 August to 29 September 2020	ECO IgG/IgM	75
	Rio Grande do Sul	5–7 February 2021	University of Rio de Janeiro	76
	Sao Paulo City	1–10 October 2020	Roche Elecsys IgG/IgM	77
	Várzea Grande	16 September to 15 October 2020	DiaSorin Liaison IgG	71
Chile	Santiago/ Coquimbo/Talca	26 September to 25 November 2020	Roche Elecsys IgG/IgM	78
Colombia	Barranquilla	20–30 September 2020	Siemens Advia IgG/IgM	79
	Bogotá	10 October to 5 November 2020	Siemens Advia IgG/IgM	79
	Bucaramanga	27 September to 9 October 2020	Siemens Advia IgG/IgM	79
	Cali	18–28 November 2020	Siemens Advia IgG/IgM	79
	Córdoba (eight cities)	1 July to 29 October 2020	INgezim DR IgG/IgM/IgA	80
	Cucuta	5–15 October 2020	Siemens Advia IgG/IgM	79
	Ipiales	3–11 December 2020	Siemens Advia IgG/IgM	79
	Leticia	15–25 September 2020	Siemens Advia IgG/IgM	79
	Medellín	5 October to 20 December 2020	Siemens Advia IgG/IgM	79
	Villavicencio	20–30 October 2020	Siemens Advia IgG/IgM	79
Ecuador	Cuenca	11 August to 1 November 2020	Standard Q IgG/IgM	81
Mexico	Nationwide	18 August to 13 November 2020	Roche Elecsys IgG/IgM	82
Paraguay	Asunción & Central Dept.	23 December to 16 February 2021	Beijing Kewei IgG/IgM	83
Peru	Cusco Province	12–27 September 2020	Roche Elecsys IgG/IgM	84
	Iquitos, Loreto	13–18 July 2020	Orient Gene Biotech IgG/IgM	85
	Lambayeque	24 June to 10 July 2020	Coretest IgG/IgM	86
	Lima and Callao	28 June to 9 July 2020	Standard Q IgG/IgM	40
Africa				
Ethiopia	Addis Ababa	22 July to 10 August 2020	Core Technology IgG	87
	Dire Dawa	15 June to 30 July 2020	Abbott Architect IgG	88
Kenya	Nairobi County	2–23 November 2020	Wantai IgG/IgM	89
Mozambique	Maputo city	3–21 August 2020	Abbott PanBio IgG/IgM	90
Senegal	Nationwide	24 October to 26 November 2020	Wantai IgG/IgM	91
South Africa	Gauteng	4 November to 22 January 2021	Luminex S IgG	92
	Mitchells Plain	8 December to 31 January 2021	Wantai IgG/IgM	93
Zambia	Lusaka & Ndola	4–27 July 2020	Euroimmun IgG	94
Middle East				
Iran	Nationwide	3 August to 31 October 2020	Pishtaz Teb IgG/IgM	95
Jordan	Nationwide	27 December to 6 January 2021	Wantai IgG/IgM	96
Oman	Nationwide	12–19 July 2020	DiaSorin Liaison IgG	97
Europe				

Continued

Table 2 Continued

	Location	Date range	Assay used	Citation
Bosnia & Herzegovina	Republika Srpska	4 November to 16 December 2020	Wantai IgG/IgM	25
Hungary	National Study	1–16 May 2020	Abbott Architect IgG	98
Poland	Katowice region	1 October to 30 November 2020	Euroimmun IgG	99
Russia	St. Petersburg	25 May to 28 June 2020	Genetico CoronaPass Total	21
South Asia				
India	Berhampur	6–6 August 2020	Roche Elecsys IgG/IgM	100
	Bhubaneswar	10–10 July 2020	Roche Elecsys IgG/IgM	100
	Chennai	17–28 July 2020	Abbott Architect IgG	101
	Delhi	1–7 August 2020	Zydus Kavach IgG	102
	Karnataka	15 June to 29 August 2020	THSTI IgG	103
	Malegaon	25 July to 20 August 2020	Karwa Kavach IgG	104
	Mumbai	29 June to 19 July 2020	Abbott Architect IgG	42
	Paschim Medinipur	27 July to 7 August 2020	ErbaLisa IgG	105
	Pimpri-Chinchwad	7–17 October 2020	Abbott Architect IgG	106
	Puducherry	10–16 September 2020	Roche Elecsys IgG/IgM	107
	Srinagar	17–20 October 2020	Abbott Architect IgG	108
	Tamil Nadu	19 October to 30 November 2020	iFlash IgG & Vitros IgG	109
Nepal	Nationwide	9–22 October 2020	Wantai IgG/IgM	110
Pakistan	Karachi	15–31 July 2020	Roche Elecsys IgG/IgM	111
	Lahore	15–31 July 2020	Roche Elecsys IgG/IgM	111
East Asia				
China	Hubei (excluding Wuhan)	10–18 April 2020	Bioscience IgG/IgM	112
China	Wuhan	10–18 April 2020	Bioscience IgG/IgM	112

Infection fatality rates

Our statistical analysis produced age-specific IFRs and CIs for 28 locations, and population IFRs for those locations as well as an additional 27 places. The full results of this analysis are shown in the online supplemental appendix 3. We obtain the following metaregression results:

$$log_{10}(IFR) = -2.75 + 0.0478 * age$$
$$(0.10)\,(0.0023)$$

where IFR is expressed in percentage points, and the SE for each estimated coefficient is given in parentheses. These estimates are highly significant with t-statistics of −28.7 and 21.0, respectively, and p values below 0·0001. The residual heterogeneity is τ^2=0.039 (p<0.0001) and I^2=92.5, confirming that the random effects are essential for capturing unexplained variations across studies and age groups. The adjusted R^2 is 91.1%. Location-specific fixed effects are only distinguishable from zero for three locations: Maranhão, Brazil (−0.50); Chennai, India (−0.68); and Karnataka, India (−1.29).

The metaregression results can be seen in figure 5. Nearly all of the observations fall within the 95% prediction interval. The importance of the location-specific effects is readily apparent. Indeed, these effects imply that the age-specific IFRs for Maranhão are about 1/3 of the metaregression prediction, while those for Chennai and Karnataka are 1/5 and 1/20, respectively.

This metaregression analysis uses age-specific IFRs based on reported COVID-19 deaths in each location. As a crosscheck, table 3 reports the ratio of excess mortality to reported deaths for each of these locations. For nearly all of these locations, the ratio is indistinguishable from unity; that is, reported COVID-19 deaths are broadly consistent with the evidence from excess mortality assessments. There were three exceptions (Chennai, Karnataka and Nairobi, Kenya), two of which had significant location-specific effects in the metaregression.

The precision of IFR estimates varied by age. At lower age groups, the number of deaths becomes very small, and thus the uncertainty is large regarding the IFR. Conversely, at older ages the number of infections and deaths can be very small in countries with extremely small populations of those aged over 65 years, and thus these estimates are also uncertain. The detailed analysis of age-specific IFR for each location is provided in online supplemental appendix figure A6.

Figure 6 shows that these age-specific IFRs are systematically higher than those of a prior metaregression estimated using studies of high-income countries.[15] That

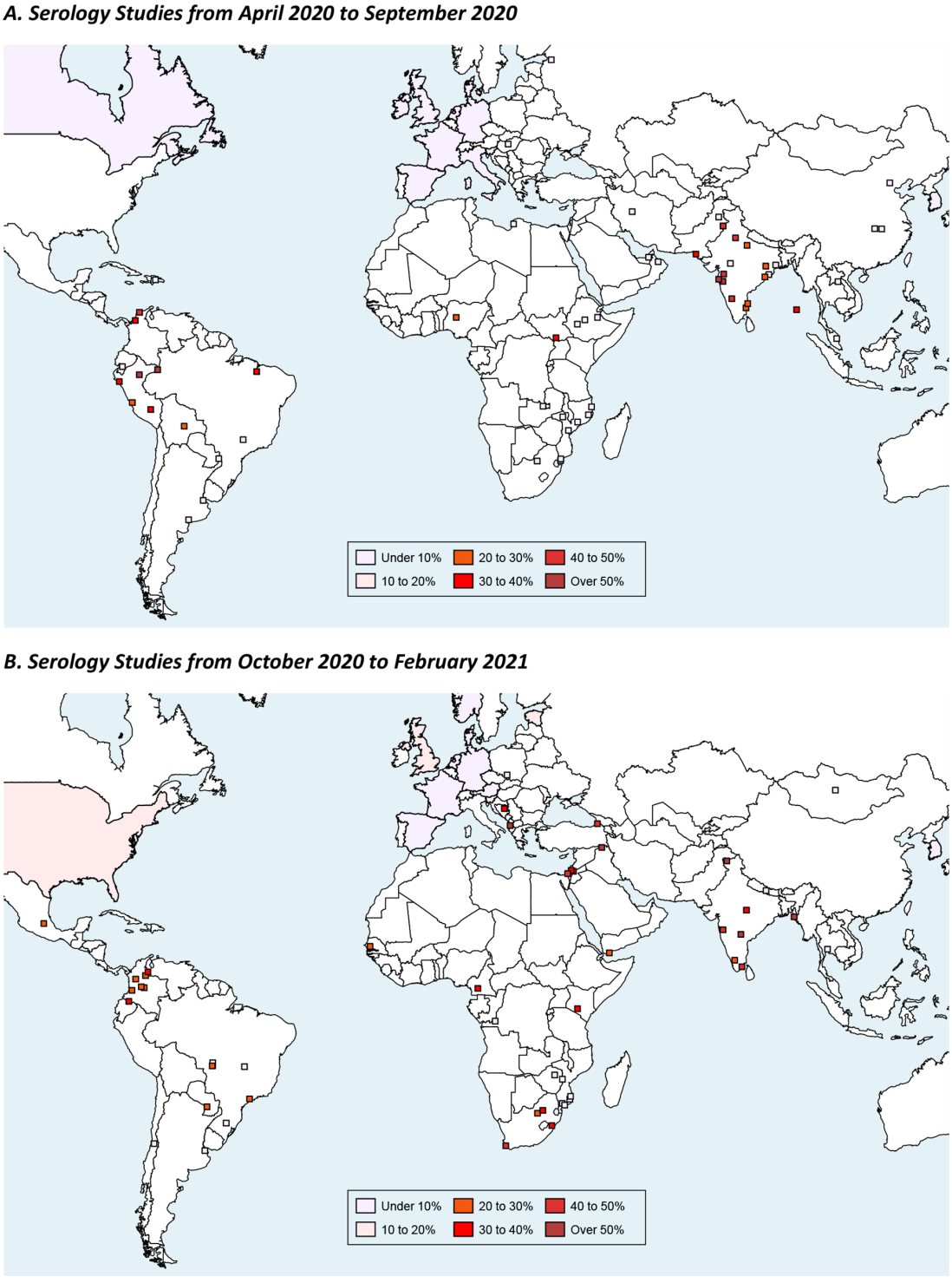


Figure 2 Estimates of seroprevalence.

benchmark metaregression has a slope of 0.0524 (95% CI 0.0499 to 0.0549), and a Welch test strongly rejects the hypothesis of equality in the slope parameters for developing countries versus high-income countries with a value of $p<0.0001$. This figure also shows a variant of our metaregression, estimated using studies of developing country locations conducted over the same time frame as in the benchmark metaregression (April to September 2020) and excluding the three outlier locations (Maranhão, Chennai and Karnataka); the estimated intercept and slope coefficient of this variant (−2.68 and 0.0480, respectively) are statistically indistinguishable from the baseline values shown above.

Figure 7 shows estimates of population IFR at ages 18–65 years, adjusted for excess mortality using the ratios shown in table 3. To facilitate comparability across locations, these estimates use a standardised age structure to aggregate the age-specific prevalence and fatalities in

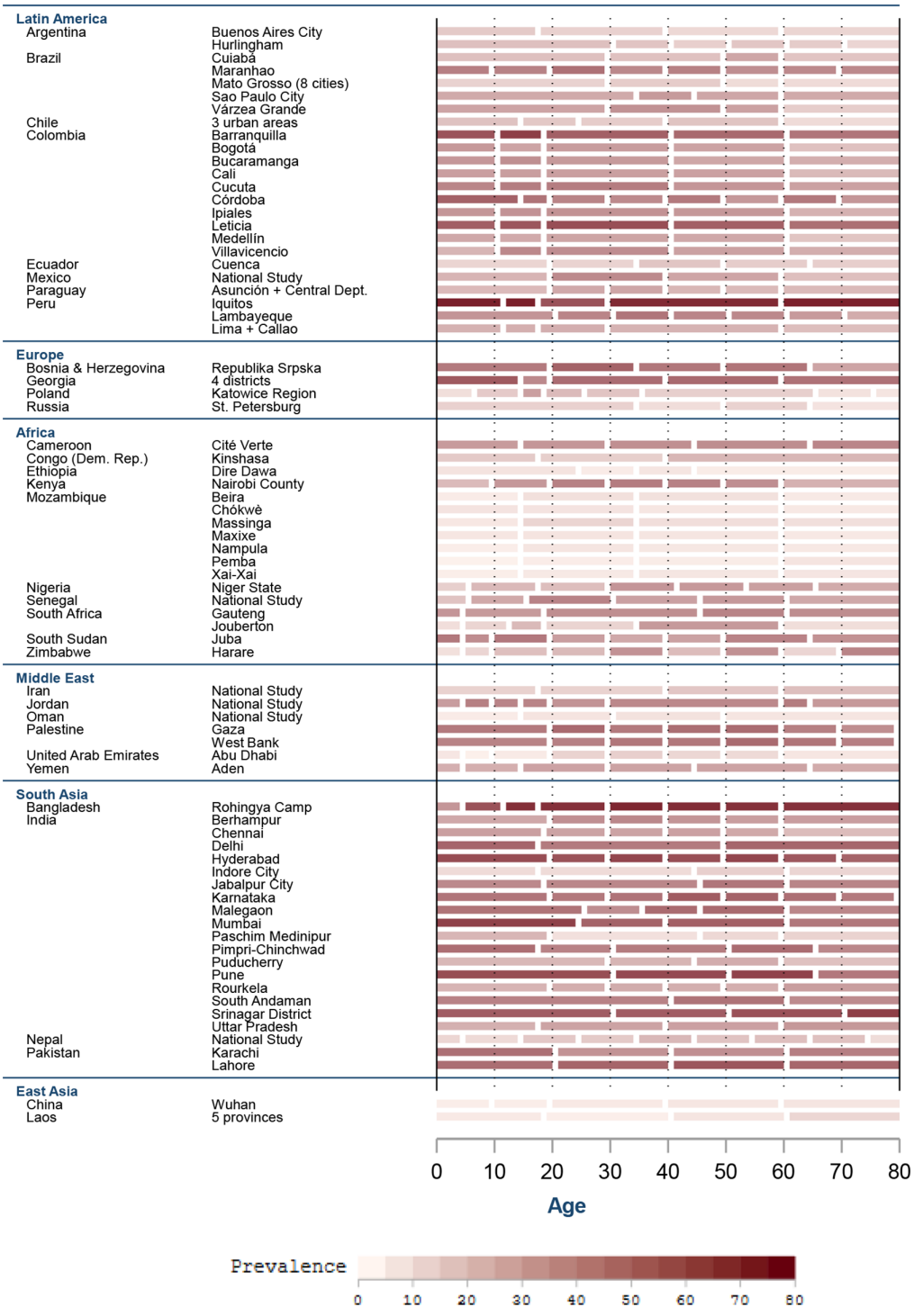


Figure 3 Age-specific seroprevalence by location.

each location. Corresponding estimates, using the actual population age structure of each location, are shown in online supplemental appendix figure A12.

Assessment of death reporting

For the full set of locations for which population IFR can be assessed, we found that the adequacy of death certification was highly significant in explaining cross-country variations. As shown in figure 8, the median value of population IFR was about 0.5% in countries where a majority of deaths were well certified (using Sustainable Development Goal (SDG) assessments[26] conducted prior to the pandemic), compared with only 0.05% in countries with lower proportions of well-certified deaths. In the latter set of countries, adjustments for excess mortality shift the population IFR upwards by an order of magnitude, to a median of 0.6%. Indeed, the population IFR for Zambia increases from 0.23% to 1.96%—the highest value for any country in our sample. In contrast, the excess mortality

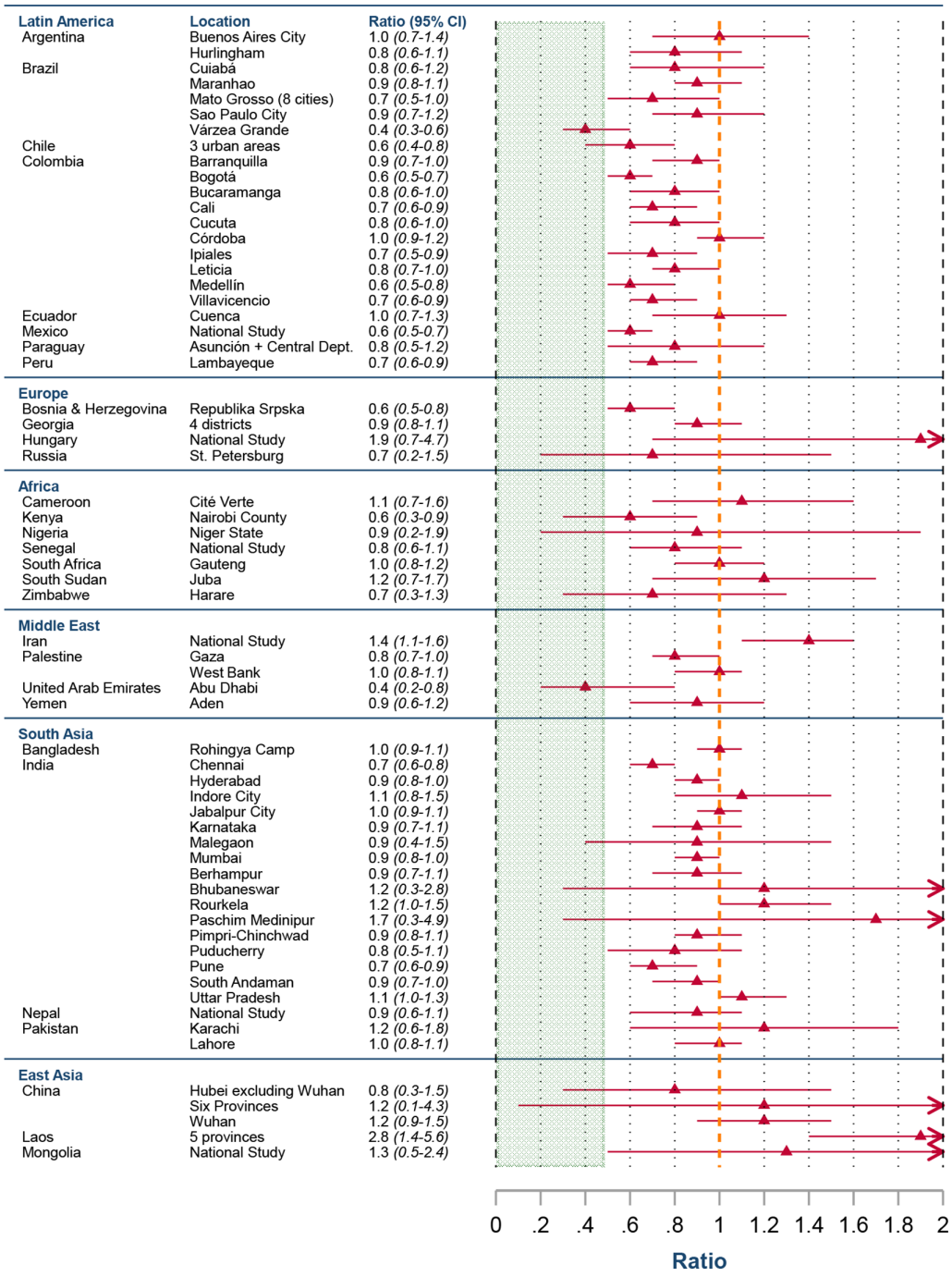


Figure 4 Ratio of seroprevalence for older adults (60+ years) compared with adults (40–59 years).

adjustments make relatively little difference for countries with a majority of well-certified deaths.

Finally, we considered the extent to which the adjusted measures of population IFR were robust to alternative estimates of the ratio of excess mortality to reported deaths. As shown in figure 9 the estimates from IHME and WMD were generally well aligned, with just a small number of exceptions.

The adjusted population IFRs had a median value of 0.49% using the IHME estimates and 0.58% using the WMD estimates.

DISCUSSION

COVID-19 has had a severe burden on developing countries. Prevalence in developing countries is roughly uniform across age groups, in contrast to the typical pattern in high-income countries where seroprevalence is markedly lower among middle-aged and older adults who are most vulnerable to this disease. Moreover, the IFR is substantially higher in developing countries compared with high-income countries.

At 20 years of age, the mean IFR in developing countries is 2.7 times higher than that in high-income countries

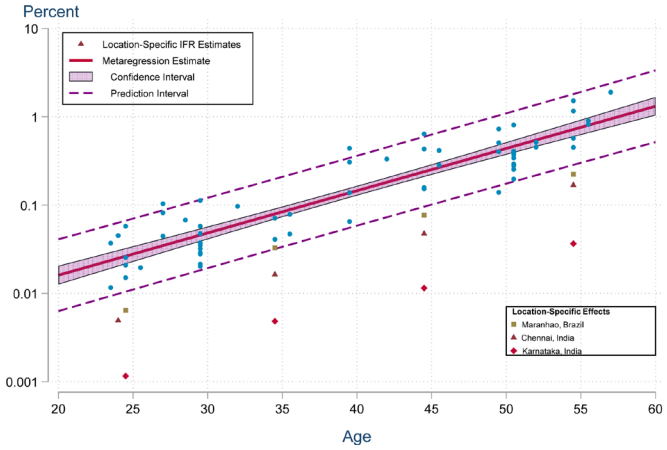


Figure 5 metaregression results. IFR, infection fatality rate.

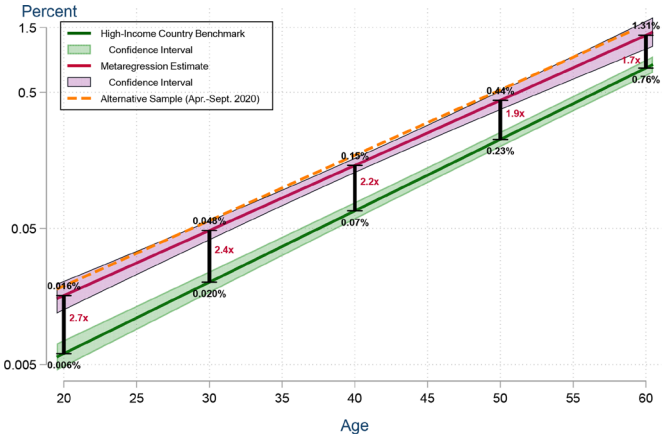


Figure 6 IFR in developing countries compared to high-income countries. IFR, infection fatality rate.

Table 3 Ratio of excess mortality to reported COVID-19 deaths

Country	Location	Ratio (95% CI)
Argentina	Buenos Aires City	1.07 (1.0 to 1.5)
Argentina	Municipality of Hurlingham	1.07 (1.0 to 1.5)
Brazil	Maranhao	1.41 (1.0 to 2.4)
Brazil	Sao Paulo City	1.02 (1.0 to 1.3)
Brazil	Cuiaba, Mato Grosso	1.00 (1.0 to 1.0)
Brazil	Varzea Grande, Mato Grosso	1.00 (1.0 to 1.0)
Chile	Coquimbo-La Serena, Greater Santiago, Talca	1.00 (1.0 to 1.0)
China	Wuhan	1.00 (1.0 to 1.0)
Colombia	Leticia (Amazonas)	1.09 (1.0 to 1.6)
Colombia	Barranquilla (Atlantico)	1.09 (1.0 to 1.6)
Colombia	Medellin (Antioquia)	1.09 (1.0 to 1.6)
Colombia	Bucaramanga (Santander)	1.09 (1.0 to 1.6)
Colombia	Cucuta (Norte Santander)	1.09 (1.0 to 1.6)
Colombia	Villavicencio (Meta)	1.09 (1.0 to 1.6)
Colombia	Bogota	1.09 (1.0 to 1.6)
Colombia	Cali (Valle del Cauca)	1.09 (1.0 to 1.6)
Colombia	Ipiales (Narino)	1.09 (1.0 to 1.6)
Colombia	Cordoba: 8 cities	1.09 (1.0 to 1.6)
Ecuador	Cuenca (Azuay)	1.01 (1.0 to 1.1)
Hungary	National Study	1.04 (1.0 to 1.4)
India	Karnataka	4.89 (2.6 to 8.2)
India	Chennai	4.80 (2.7 to 7.9)
Jordan	National Study	1.57 (1.0 to 3.0)
Kenya	Nairobi County	13.29 (7.1 to 23.1)
Paraguay	Asuncion+Central Department	1.10 (1.0 to 1.6)
Peru	Lambayeque	1.09 (1.0 to 1.6)
Peru	Lima (Metropolitana)+Callao	1.09 (1.0 to 1.6)
Peru	Iquitos, Loreto	1.09 (1.0 to 1.6)

Note: This table shows Institute for Health Metrics and Evaluation (IHME) estimates of the ratio of excess mortality to reported COVID-19 deaths (constrained to be 1.0 or greater).[22] The 95% CIs, enclosed in parentheses, are also taken directly from IHME, with a one-tailed interval for each location where the estimated undercount ratio is constrained by the lower bound of unity.

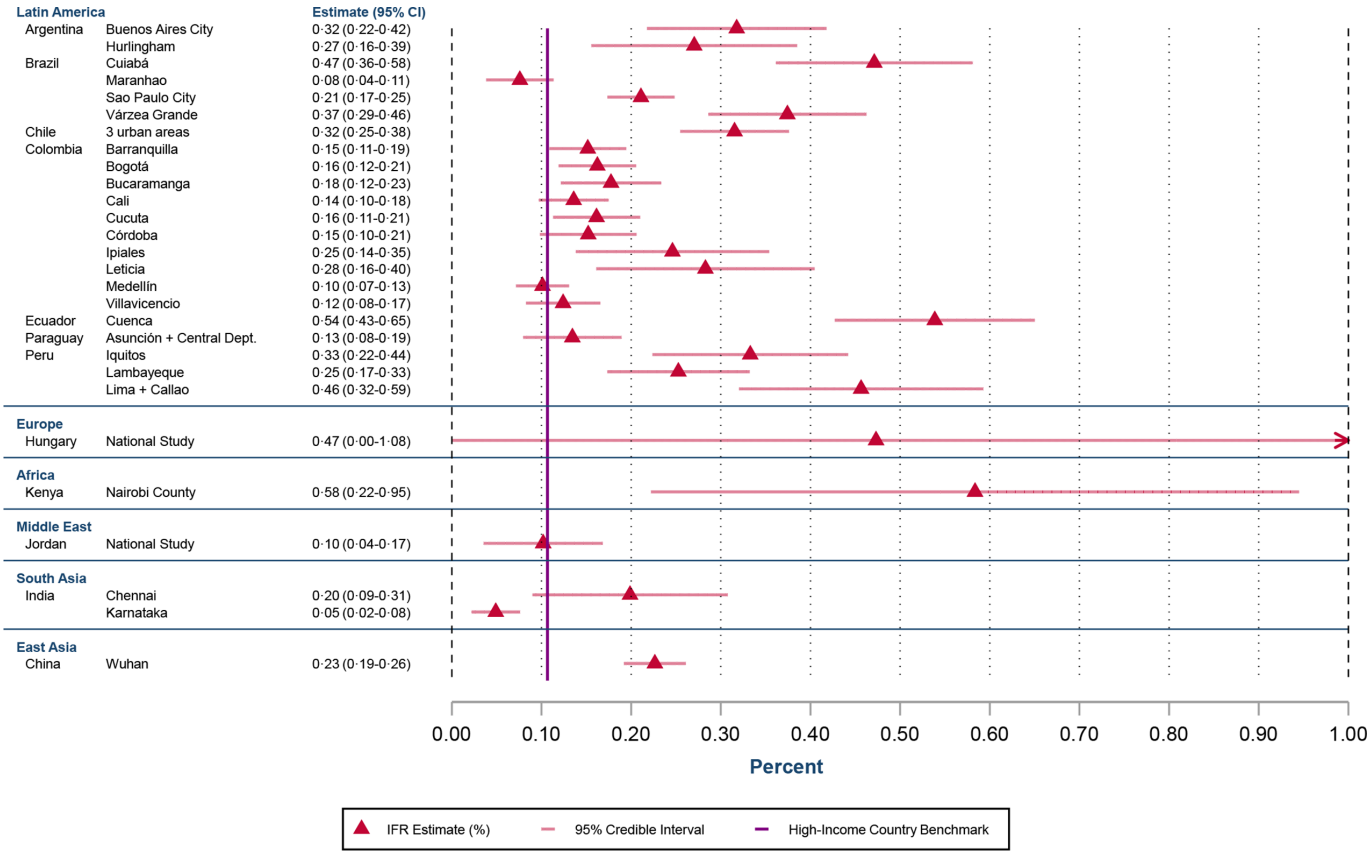


Figure 7 Population IFR for ages 18–65 years. IFR, infection fatality rate.

and at age 60 years the risk is doubled. At the oldest ages, this discrepancy is reduced, with only a modestly increased risk at age 80 years. These relationships have also been found with socioeconomic status within specific places such as Santiago, Chile.[7] This warrants further research to understand why access to healthcare and other socioeconomic issues appear to have a larger impact on survival at younger ages. This finding does not

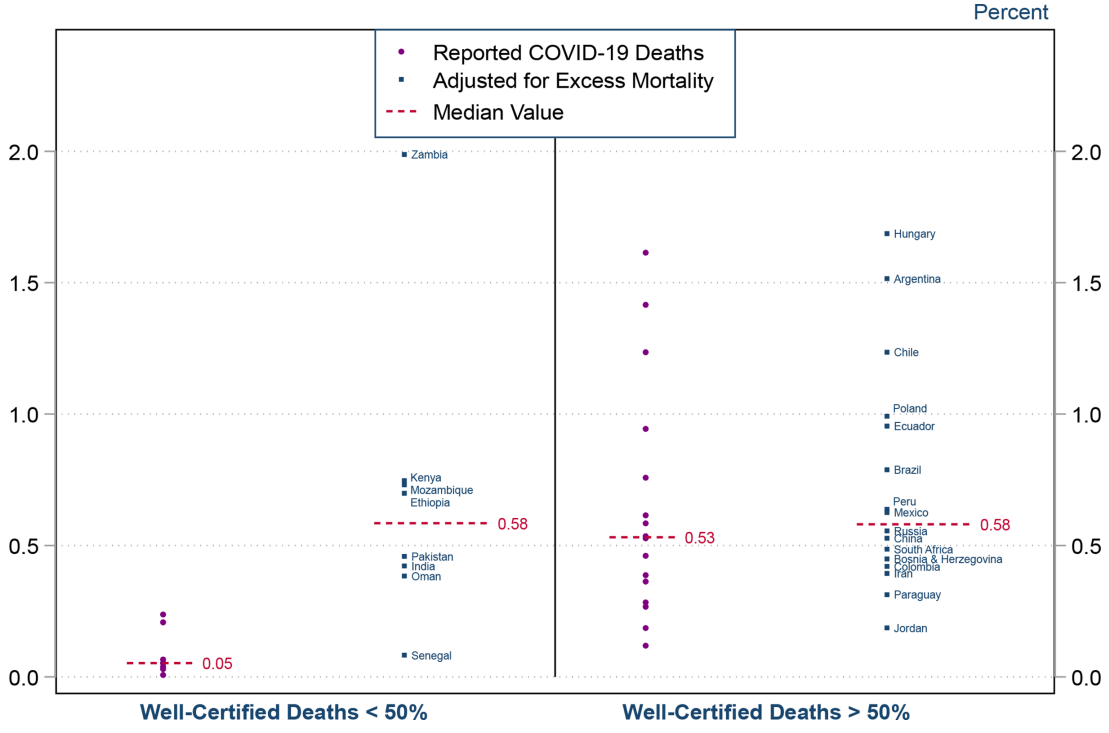


Figure 8 Population IFR and well-certified death registrations. IFR, infection fatality rate.

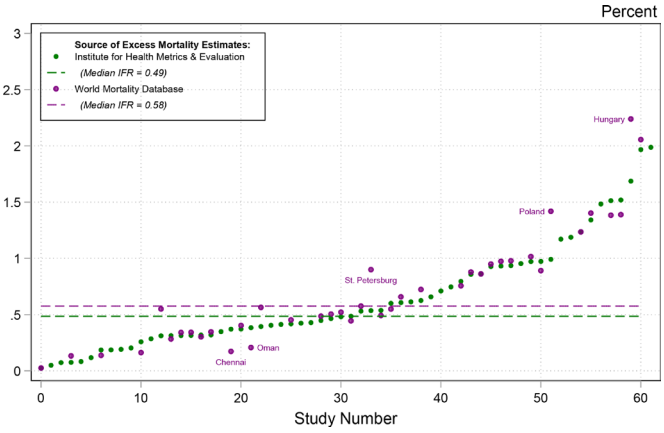


Figure 9 Excess mortality adjusted population IFRs. IFR, infection fatality rate.

rely on any specific modelling assumptions such as log-linearity, which is shown by the readily apparent disparity in figure 7, showing the age-standardised IFR for each individual location compared with the benchmark of high-income countries.

The elevated IFR in developing nations only becomes apparent when stratifying by age and adjusting for death under-reporting. Indeed, the quality of the vital statistics system tends to be linked to the overall level of economic development, and hence some previous studies of unadjusted data have incorrectly inferred that population IFRs are lower in developing countries than in high-income countries.[27–30]

Our results are important for addressing questions that have arisen about whether COVID-19 was less dangerous for populations in sub-Saharan Africa compared with locations elsewhere.[31–33] As shown in figure 7, the age-standardised population IFR of Nairobi County, Kenya is about five times higher than the high-income country benchmark. Likewise, figure 8 shows that the population IFR (adjusted for under-reporting of COVID-19 fatalities) exceeds 0.5% for locations in Ethiopia, Mozambique and South Africa; the sole exception is Senegal, perhaps due to even more severe death undercounting than captured by the estimated ratio. These results underscore the importance of drawing inferences from representative samples rather than from convenience samples.[34–36]

These results are consistent with the pattern observed for most other communicable diseases.[3 4] In locations with little ability to work from home, where quarantine is difficult or impossible, where opportunities for physical distancing and access to sanitation are poor, with lower healthcare resources, and where even basic resources such as supplemental oxygen are in short supply, people have fared substantially worse during the pandemic than in high-income settings. Indeed, in low-income settings where fewer hospital beds and healthcare workers are available, COVID-19 has caused great devastation and an enormous death toll. With a much higher IFR, particularly in younger people, the ultimate burden for developing nations from COVID-19 is likely to be very high.

Another important facet of our results is that seroprevalence was both higher and consistent across age groups in developing countries—a striking contrast to the typical pattern in high-income countries, where prevalence among older adults was markedly lower than among younger adults.[15 37 38] Evidently, it is very difficult to insulate elderly people from the virus in a slum or a rural village. This is likely also impacted by the higher proportion of multigenerational families in developing countries,[39] a known risk-factor for COVID-19 infection and death.[40 41] For example, seroprevalence in slum neighbourhoods of Mumbai was about four times higher than in non-slum neighbourhoods.[42] Our analysis indicates that the relatively uniform prevalence of COVID-19 in developing countries has dramatically increased the number of fatalities in these locations.

Our findings reinforce the conclusions of previous studies that have assessed the IFR of COVID-19.[13 43] In particular, COVID-19 is dangerous for middle-aged adults, not just the elderly and infirm.[15] Our metaregression results are well aligned with IFR estimates produced for specific locations in developing countries (see supplementary online supplemental appendix table A8).

Our analysis underscores that incomplete death reporting is a crucial source of apparent differences in COVID-19 death rates. In particular, this is related to the proportion of deaths that are assigned to so-called 'garbage codes'.[26 44 45] These deaths are, by definition, not included in national tallies of the population that has died from COVID-19. As shown in figure 8, the IFR is on average 10 times higher in locations with reasonably adequate vital statistics compared with other locations where a majority of deaths are not well certified.

The divergence between population IFRs for locations is similar whether adjusted for death certification or excess mortality. Adjustment for estimates of excess mortality produced location population IFRs that were consistent with IFRs produced in the age-stratified analysis, aside from a few minor outliers. As shown in figure 8, the median of these population IFRs for developing nations, once adjusted for undercounting of COVID-19 deaths, was 0.58%, very similar to the median estimates of IFR for high-income countries.[46]

Excess mortality is a useful metric for adjusting IFR estimates in areas where deaths are well registered but not well certified; that is: captured in national vital statistics but without a specific cause of death.[1] Nonetheless, caution is warranted in applying national estimates of excess mortality to specific regions within a country, recognising that death reporting systems may vary markedly with the degree of urbanisation and other socioeconomic factors. In the case of Ecuador, for example, the national estimate for the ratio of excess mortality to reported COVID-19 deaths in 2020 was 2.6 (1, 22), whereas that ratio was only 1.01 in the province of Azuay.[47]

Moreover, estimates of excess mortality may partly reflect indirect effects of the pandemic on other sources of mortality. On the one hand, non-pharmaceutical

interventions (such as business closures) may reduce mortality from causes such as vehicle accidents.[1 48] Conversely, mortality may be elevated by impaired access to healthcare for non-infectious diseases such as chronic cardiovascular disease or cancer,[49] or by higher burdens of non-COVID infectious diseases such as malaria, tuberculosis or parasitic infections.[50]

Finally, the true burden of COVID-19 may be practically impossible to assess in locations where many deaths are never entered into the national vital statistics system.[51] For example, total mortality in Kenya was lower in 2020 than in 2019, but those statistics should certainly not be interpreted as suggesting that Kenya was unscathed by the pandemic.[22] Indeed, assessments of Kenya's vital statistics found that only two-thirds of actual deaths were recorded in the system.[51] Such considerations may explain other outliers in our analysis, such as Senegal, which remains far below similar locations even when estimates are adjusted for excess mortality.

A useful example in this case is Ethiopia. Despite national statistics not showing a large increase in deaths in Ethiopia during the pandemic, an epidemiological investigation of burial sites has revealed a huge increase in mortality during this period that is not part of the official reporting of COVID-19.[52]

In the absence of better death reporting, it is challenging to assess the extent to which differences in IFR across locations reflect systematic disparities in healthcare access, socioeconomic status and other indicators. Nonetheless, such effects have been clearly demonstrated by studies that have assessed distinct socioeconomic groups within specific regions such as Santiago, Chile.[7] Moreover, these considerations are almost certainly relevant in interpreting our finding that age-stratified IFR is markedly higher in developing countries compared with high-income countries.[53] Indeed, our results underscore the tragedy that a Zambian young adult with COVID-19 would be far more likely to die than a Swiss person of similar age.

Accounting for seroreversion and other assay characteristics is crucial for assessing seroprevalence accurately. Our analysis makes a novel contribution in providing a systematic assessment of the implications of seroreversion; that is, the proportion of people who develop antibodies but whose tests will fall below the limit of detection at a later date. Prior studies have either ignored this issue or have assumed that seroreversion occurs at a fixed geometric rate regardless of the assay used.[12 13] In contrast, we have collated detailed information about the characteristics of all assays used in the serology studies included in our analysis, including data on seroreversion as well as test specificity and sensitivity; that information is fully described in online supplemental appendix 2a and b. Our analysis clearly indicates that the extent of seroreversion differs in magnitude depending on the assay used. Moreover, accounting for seroreversion had substantial implications for a number of locations in our analysis.

Our analysis makes a strong case for swifter action on vaccine and other medication equity. While countries have largely sought to protect their own populations, there is increasing commitment to ensuring that key populations in low-income and middle-income countries receive protection, at a minimum for their frontline health and other personnel. It is widely accepted that failing to control the pandemic across the globe will contribute to the emergence of additional strains of COVID-19, potentially undermining the efficacy of available vaccines.[54] Current medication distribution efforts are grossly inequitable.[55] Recent estimates suggest that fewer than 10% of people in low-income countries have received an immunisation, while the majority of people in high-income countries have had at least one vaccination.[56] Similarly, the availability of effective medications such as Paxlovid is grossly inequitable across the globe.[57]

As with all research, our study is subject to a number of limitations. First, while we made every effort to capture seroprevalence data, including corresponding with dozens of researchers and public health officials worldwide, it is possible that some studies have been missed. However, it is unlikely that any small number of additional studies would make a material difference to our results.

Our analysis did not incorporate time series data on the evolution of COVID-19 deaths. However, some studies of high-income countries have shown how such data can be useful in refining assessment of IFR to incorporate the stochastic timing of COVID-19 deaths.[13 58] Such analysis should be a priority for future research about IFR in developing countries.

While our analysis excluded convenience samples and focused exclusively on representative samples of the population, we recognise that such studies may also be susceptible to selection bias. Research conducted at various stages of the pandemic has found that individual preferences for testing can be associated with substantial bias in estimates of seroprevalence, with corresponding implications for estimates of population IFR.[20 59] Such uncertainty can be incorporated into statistical models of prevalence and IFR.[60 61] However, we did not follow such an approach here, because our statistical model already incorporates a number of other substantial sources of uncertainty.

Our work also did not consider non-mortality harms from COVID-19. Recent work has shown that even at younger ages a substantial fraction of infected individuals will have severe, long-lasting adverse effects from COVID-19.[62] Consequently, the impact on the healthcare system and society may be far greater than would be reflected in mortality rates alone. Focusing only on survival rates obscures the large number of deaths that occur when many people are infected,[63] the relatively high fatality rate of COVID-19 in comparison to other diseases and other causes of death,[64] and non-mortality harms of COVID-19, such as hospitalisation from serious disease.[62]

Future work should address these non-mortality harms, including long COVID-19.

Another potentially serious limitation of our analysis is cross-reactivity in serological tests due to malaria. An investigation in Nigeria found that the commonly used Abbott and Euroimmun serological assays had a false-positive rate of 6.1% against prepandemic samples due to cross-reactivity with malarial antibodies.[65] This would substantially lower specificity of the assay in areas with a high prevalence of past malaria infection, which would have the practical result of producing an upward bias of seroprevalence estimates and downward bias of IFR estimates. Thus, it is plausible that in areas with a large burden of malaria, that the IFR we have calculated represents a substantial underestimate.

Finally, our analysis only includes serology studies where specimen collection was completed by the end of February 2021. Consequently, our results do not reflect any potential changes in IFR that may have resulted from more recent advances in COVID-19 care, most notably, the development of novel antiviral medications and dissemination of vaccines. Of course, the IFR could also shift with the spread of new variants of SARS-CoV-2.[66] However, given that the first major variant of COVID-19 was only identified in late 2020, and most vaccination campaigns in developing nations only began in early 2021, our time frame limits the impact that these factors should have on the results.

CONCLUSION

The prevalence and IFR by age of COVID-19 is far higher in developing countries than in high-income countries, reflecting a combination of elevated transmission to middle-aged and older adults, as well as limited access to adequate healthcare. These results underscore the critical need to accelerate the provision of vaccine boosters and newer effective medications to vulnerable populations in developing countries. Moreover, many developing countries require ongoing support to upgrade the quality of their vital statistics systems to facilitate public health decisions and actions, not only for the COVID-19 pandemic but for future global health concerns.

Code and data

All data and code are available publicly online.[67]

Author affiliations
[1]Economics, Dartmouth College, Hanover, New Hampshire, USA
[2]National Bureau for Economic Research, Cambridge, Massachusetts, USA
[3]School of Medicine, Case Western Reserve University, Cleveland, Ohio, USA
[4]Statistics, Colorado State University, Fort Collins, Colorado, USA
[5]Department of Statistics, Colorado State University, Fort Collins, Colorado, USA
[6]School of Social Sciences, University of New South Wales, Sydney, New South Wales, Australia
[7]Law School, University of Chicago, Chicago, Illinois, USA
[8]Harris School of Public Policy, University of Chicago, Chicago, Illinois, USA
[9]Faculty of Information and Technology, Monash University, Clayton, Victoria, Australia
[10]Interdisciplinary Centre on Population Dynamics, University of Southern Denmark, Odense, Denmark
[11]Department of Economics, Dartmouth College, Hanover, New Hampshire, USA
[12]Laboratorio de Neurociencias, Universidad de la República, Montevideo, Uruguay
[13]Department of Epidemiology & Biostatistics, Michigan State University, East Lansing, Michigan, USA
[14]Adult Intensive Care Unit, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil
[15]Ethiopian Public Health Institute, Addis Ababa, Ethiopia
[16]Western Sydney Diabetes, Western Sydney Local Health District, Blacktown, New South Wales, Australia
[17]School of Health and Society, University of Wollongong, Wollongong, New South Wales, Australia

Twitter Anthony B Zwi @HEARDatUNSW, Anup Malani @anup_malani, Satej Soman @satejsoman, Ana Carolina Peçanha Antonio @AnaCarolPecanha and Enyew Birru Tadese @Enyew54639156

Acknowledgements The authors thank Ariel Karlinsky for assistance with death registration and mortality data.

Contributors ATL and GM-K initiated and provided leadership for the project, and act as guarantors for the project. BKF and SP designated the Bayesian statistical framework. NO-B took primary responsibility for the search procedures, and performed the review of assay characteristics and seroreversion. ATL and NO-B reviewed each of the studies identified in the initial screening, and assessed and applied the exclusion criteria. SS took the lead in designing the data management procedures and setting up the GitHub repository. LB has developed an interactive tool that will be linked to the GitHub repository. SG, AM, GS and RU assisted with data extraction and verification. ABZ, AM and IK reviewed the methodology and contributed to the discussion of key findings. DH-E, GdIC, ACPA and EBT contributed insights that reflected their experience with health issues in developing countries. GM-K drafted the main text; NO-B and SP drafted the supplementary materials. ATL was responsible for conducting the metaregressions and produced all the figures and tables included in the manuscript. ATL, GM-K, NO-B and SP edited the text of the manuscript and the supplementary materials.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study exclusively used publicly available aggregate data sets and published research, and hence no ethics approval was required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. https://covid-ifr.github.io/

ORCID iDs
Andrew T Levin http://orcid.org/0000-0002-1136-6900
Anup Malani http://orcid.org/0000-0002-2594-5778

Satej Soman http://orcid.org/0000-0001-8450-7025
Ana Carolina Peçanha Antonio http://orcid.org/0000-0001-6146-1247
Gideon Meyerowitz-Katz http://orcid.org/0000-0002-1156-9345

REFERENCES

1 Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the world mortality dataset. *Elife* 2021;10:e69336.
2 Ramachandran S, Malani A. All-Cause mortality during SARS-CoV-2 pandemic in India: Nationally-representative estimates independent of official death registry. *medRxiv* 2021;2021.
3 Gilks CF, Crowley S, Ekpini R, *et al*. The who public-health approach to antiretroviral treatment against HIV in resource-limited settings. *Lancet* 2006;368:505–10.
4 Dye C. Global epidemiology of tuberculosis. *The Lancet* 2006;367:938–40.
5 Deshmukh Y, Suraweera W, Tumbe C. Excess mortality in India from June 2020 to June 2021 during the COVID pandemic: death registration, health facility deaths, and survey data. *medRxiv* 2021;2021.
6 *The true death toll of COVID-19: World health Organization*2021
7 Mena GE, Martinez PP, Mahmud AS, *et al*. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* 2021;372:eabg5298.
8 Wetzler HP, Wetzler EA. COVID-19 excess deaths in the United States, New York City, and Michigan during April 2020. *medRxiv* 2020;2020.
9 Modi C, Boehm V, Ferraro S. How deadly is COVID-19? A rigorous analysis of excess mortality and age-dependent fatality rates in Italy. *medRxiv* 2020.
10 Mwananyanda L, Gill CJ, MacLeod W, *et al*. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. *BMJ* 2021;372:n334.
11 Gill CJ. Latest data from Lusaka morgue analysis shows spike in COVID-19 deaths. *The Conversation* 2021.
12 O'Driscoll M, Ribeiro Dos Santos G, Wang L, *et al*. Age-Specific mortality and immunity patterns of SARS-CoV-2. *Nature* 2021;590:140–5.
13 Brazeau N, Verity R, Jenks S. *Report 34: COVID-19 infection fatality ratio: estimates from seroprevalence*. Imperial College London, 2020.
14 Chen X, Chen Z, Azman AS, *et al*. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. *Lancet Glob Health* 2021;9:e598–609.
15 Levin AT, Hanage WP, Owusu-Boaitey N, *et al*. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol* 2020;35:1123–38.
16 Meyerowitz-Katz G, Merone L. A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates. *medRxiv* 2020.
17 World economic and financial surveys world economic outlook Database—WEO groups and aggregates Information 2021.
18 Community assessment for public health emergency response toolkit: CDC 2019.
19 Population-Based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection: World health Organization 2020.
20 Gajda M, Kowalska M, Zejda JE. Impact of two different recruitment procedures (random vs. volunteer selection) on the results of seroepidemiological study (SARS-CoV-2). *Int J Environ Res Public Health* 2021;18:9928.
21 Barchuk A, Shirokov D, Sergeeva M, *et al*. Evaluation of the performance of SARS--CoV--2 antibody assays for a longitudinal population-based study of COVID--19 spread in St. Petersburg, Russia. *J Med Virol* 2021;93:5846–52.
22 COVID-19 estimate downloads. Seattle: Institute for health metrics and Evaluation 2021.
23 Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. *medRxiv* 2020.
24 Team SD. *RStan: the R interface to Stan. R package version 2.19.3*, 2020.
25 Bergeri I, Whelan M, Ware H. Global epidemiology of SARS-CoV-2 infection: a systematic review and meta-analysis of standardized population-based seroprevalence studies, Jan 2020-Oct 2021. *medRxiv* 2021.
26 Vos T, Lim SS, Abbafati C. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet* 2020;396:1204–22.
27 Ayoub HH, Mumtaz GR, Seedat S, *et al*. Estimates of global SARS-CoV-2 infection exposure, infection morbidity, and infection mortality rates in 2020. *Glob Epidemiol* 2021;3:100068.
28 Ioannidis JPA. Infection fatality rate of COVID-19 inferred from seroprevalence data. *Bull World Health Organ* 2021;99:19–33.
29 IHME COVID-19 Forecasting Team. Variation in the COVID-19 infection–fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis. *The Lancet* 2022.
30 Campbell H, Gustafson P. Inferring the COVID-19 infection fatality rate in the community-dwelling population: a simple Bayesian evidence synthesis of seroprevalence study data and imprecise mortality data. *Epidemiol Infect* 2021;149:e243.
31 Maeda JM, Nkengasong JN. The puzzle of the COVID-19 pandemic in Africa. *Science* 2021;371:27–8.
32 Mariam SH. The severe acute respiratory syndrome Coronavirus-2 (SARS-CoV-2) pandemic: are Africa's prevalence and mortality rates relatively low? *Adv Virol* 2022;2022:3387784.
33 Okonji EF, Okonji OC, Mukumbang FC, *et al*. Understanding varying COVID-19 mortality rates reported in Africa compared to Europe, Americas and Asia. *Trop Med Int Health* 2021;26:716–9.
34 Escobar LE, Molina-Cruz A, Barillas-Mury C. Bcg vaccine protection from severe coronavirus disease 2019 (COVID-19). *Proc Natl Acad Sci.* 2020;117:17720–6.
35 Lewis H, Ware H, Whelan M. SARS-CoV-2 infection in Africa: a systematic review and meta-analysis of standardised seroprevalence studies, from January 2020 to December 2021. *medRxiv* 2022.
36 Wolday D, Gebrecherkos T, Arefaine ZG, *et al*. Effect of co-infection with intestinal parasites on COVID-19 severity: a prospective observational cohort study. *E Clinical Medicine* 2021;39
37 Stringhini S, Wisniak A, Piumatti G. Repeated seroprevalence of anti-SARS-CoV-2 IgG antibodies in a population-based sample from Geneva, Switzerland. *medRxiv* 2020.
38 Pollán M, Pérez-Gómez B, Pastor-Barriuso R, *et al*. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* 2020;396:535–44.
39 Living arrangements of older persons around the world: department of economic and social Affairs. 2019.
40 Reyes-Vega MF, Soto-Cabezas MG, Cárdenas F, *et al*. SARS-CoV-2 prevalence associated to low socioeconomic status and overcrowding in an LMIC megacity: a population-based seroepidemiological survey in Lima, Peru. *EClinicalMedicine* 2021;34:100801.
41 Ghosh AK, Venkatraman S, Soroka O, *et al*. Association between overcrowded households, multigenerational households, and COVID-19: a cohort study. *Public Health* 2021;198:273–9.
42 Malani A, Shah D, Kang G, *et al*. Seroprevalence of SARS-CoV-2 in slums versus non-slums in Mumbai, India. *Lancet Glob Health* 2021;9:e110–1.
43 Verity R, Okell LC, Dorigatti I. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis* 2020.
44 GBD 2016 SDG Collaborators. Measuring progress and projecting attainment on the basis of past trends of the health-related sustainable development goals in 188 countries: an analysis from the global burden of disease study 2016. *Lancet* 2017;390:1423–59.
45 Collaborators GS, GBD 2017 SDG Collaborators. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related sustainable development goals for 195 countries and territories: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:2091–138.
46 Meyerowitz-Katz G, Merone L. A systematic review and meta-analysis of published research data on COVID-19 infection fatality rates. *Int J Infect Dis* 2020;101:138–48.
47 Cuéllar L, Torres I, Romero-Severson E, *et al*. Excess deaths reveal the true spatial, temporal, and demographic impact of COVID-19 on mortality in Ecuador. *medRxiv* 2021. doi:10.1101/2021.02.25.21252481. [Epub ahead of print: 01 Mar 2021]. 10.1093/ije/dyab163.
48 Calderon-Anyosa RJC, Kaufman JS. Impact of COVID-19 lockdown policy on homicide, suicide, and motor vehicle deaths in Peru. *Prev Med* 2021;143:106331.
49 Meyerowitz-Katz G, Bhatt S, Ratmann O, *et al*. Is the cure really worse than the disease? the health impacts of lockdowns during COVID-19. *BMJ Glob Health* 2021;6:e006653.
50 Buonfrate D, Bisanzio D, Giorli G, *et al*. The Global Prevalence of *Strongyloides stercoralis* Infection. *Pathogens* 2020;9:468.
51 Karlinsky A. International completeness of death registration 2015-2019. *medRxiv* 2021.

52 Ahmed Y. COVID-19 mortality. Twitter, 2021. Available: https://twitter.com/yakob_son/status/1450448132047790082?s=20 [Accessed 06/01/2021].
53 Demombynes G. COVID-19 age-mortality curves are Flatter in developing countries, 2021. World Bank Group. Public Research Working Paper(9313)
54 COVAX: with a fast-moving pandemic, no one is safe, unless everyone is safe, 2020. World Health organization. Available: https://www.who.int/initiatives/act-accelerator/covax
55 Keith Collins JH. See how rich countries got to the front of the vaccine line, 2021. New York times
56 COVID-19 Data Explorer. Our world in data, 2021. Available: https://ourworldindata.org/explorers/coronavirus-data-explorer [Accessed 27/09/2021].
57 Ledford H, Maxmen A. African clinical trial denied access to key COVID drug Paxlovid. *Nature* 2022;604:412–3.
58 Folkhälsomyndigheten. *The infection fatality rate of COVID-19 in Stockholm – technical report*. Sweden: Public Health Agency of Sweden, 2020.
59 Parrott JC, Maleki AN, Vassor VE, *et al*. Prevalence of SARS-CoV-2 antibodies in New York City adults, June-October 2020: a population-based survey. *J Infect Dis* 2021;224:188–95.
60 Campbell H, de Valpine P, Maxwell L. Bayesian adjustment for preferential testing in estimating infection fatality rates, as motivated by the COVID-19 pandemic. *The Annals of Applied Statistics* 2022;16:436–59.
61 Campbell H, Gustafson P. Inferring the COVID-19 IFR with a simple Bayesian evidence synthesis of seroprevalence study data and imprecise mortality data. *medRxiv.* 2021.
62 Herrera-Esposito D, de los Campos G. Age-Specific rate of severe and critical SARS-CoV-2 infections estimated with multi-country seroprevalence studies. *medRxiv* 2021.
63 Moser W, Fahal MAH, Abualas E. Retrospective mortality and prevalence of SARS-CoV-2 antibodies in greater Omdurman, Sudan: a population–based cross–sectional survey. *medRxiv* 2021.
64 Lapidus N, Paireau J, Levy-Bruhl D, *et al*. Do not neglect SARS-CoV-2 hospitalization and fatality risks in the middle-aged adult population. *Infect Dis Now* 2021;51:380–2.
65 Steinhardt LC, Ige F, Iriemenam NC. Cross-Reactivity of two SARS-CoV-2 serological assays in a setting where malaria is endemic. *J Clin Microbiol.* 2021;59:e00514–21.
66 Classification of omicron (B.1.1.529): SARS-CoV-2 variant of concern. Geneva: World health Organization. 2021.
67 Levin A, Owusu-Boaitey N, Pugh S. Data from: assessing the burden of COVID-19 in developing countries: systematic review, meta-analysis, and public policy implications. *Github Repository* April 12, 2022. [dataset] https://covid-ifr.github.io/
68 Buenos Aires City seroprevalence: city of Buenos Aires2020.
69 COVID-19 seroprevalence study carried out by the University of the Municipality of Hurlingham.
70 The first phase of the Serovigilance epidemiological study concludes: Gobierno municipal de Santa Cruz. 2020.
71 Research indicates that 12.5% of the Mato Grosso population has already been infected by the coronavirus. Governo de Mato Grosso. 2021. http://www.mt.gov.br/-/15990626-pesquisa-aponta-que-12-5-da-populacao-mato-grossense-ja-foi-infectada-pelo-coronavirus
72 Saúde presents partial data from the Covid-19 seroepidemiological survey in DF: federal district health DEPARTMENT. 2020.
73 Serological surveys show a drop in antibody levels against Covid-19. *Federal University of Latin American integration: federal University of Latin American integration*, 2020.
74 AAMd S, Lima-Neto LG, CdMPeSd A. Population-Based seroprevalence of SARS-CoV-2 is more than halfway through the herd immunity threshold in the state of Maranhão, Brazil. *medRxiv* 2020.
75 Crotti Peixoto A, Marques D, de Castro L, *et al*. Avaliação dA prevalência de marcadores virológicos E sorológicos dA infecção pelo SARS-CoV-2 Na população de Pitangueiras. *Sao Paulo: inquérito epidemiológico populacional - Relatório Final* 2020.
76 Hartwig FP, Vidaletti LP, Barros AJD. Combining serological assays and official statistics to describe the trajectory of the COVID-19 pandemic: results from the EPICOVID19-RS study in Rio grande do Sul (southern Brazil). *medRxiv* 2021.
77 Tess BH, Granato CFH, Porto Alves MCG. SARS-CoV-2 seroprevalence in the municipality of São Paulo, Brazil, ten weeks after the first reported case. *medRxiv* 2020.
78 PAaG V, Claudia L, Ramirez-Santana G. Seroprevalence, spatial distribution, and social determinants of SARS-CoV-2 in three urban centers of Chile. *SSRN* 2021.
79 Mercado-Reyes M-R, Jeadran N, Zapata S. Seroprevalence of Anti-Sars-Cov-2 antibodies in Colombia, 2020: a population-based study. *SSRN* 2021.
80 Alvis Guzman N, De la Hoz Restrepo F, Serrano-Coll H, *et al*. Using serological studies to assess COVID-19 infection fatality rate in developing countries: a case study from one Colombian department. *Int J Infect Dis* 2021;110:4–5.
81 Acurio-Páez D, Vega B, Orellana D, *et al*. Seroprevalence of SARS-CoV-2 infection and adherence to preventive measures in Cuenca, Ecuador, October 2020, a cross-sectional study. *Int J Environ Res Public Health* 2021;18:4657.
82 Informe de Resultados de la Encuesta Nacional de Salud y Nutrición - Continua COVID-19: Encuesta Nacional de Salud y Nutrición. 2020.
83 Sequera Guillermo CA, Margarita S, Cynthia V, *et al*. *Infección POR COVID 19: estudio seroepidemiológico de cohorte de base poblacional estratificado POR edad en Asunción Y central: Consejo Nacional de Ciencia Y Tecnologia*, 2020.
84 Huamaní C, Velásquez L, Montes S, *et al*. Population-Based seroprevalence of SARS-CoV-2 antibodies in a high-altitude setting in Peru. *medRxiv* 2021.
85 Álvarez-Antonio C, Meza-Sánchez G, Calampa C, *et al*. Seroprevalence of anti-SARS-CoV-2 antibodies in Iquitos, Peru in July and August, 2020: a population-based study. *Lancet Glob Health* 2021;9:e925–31.
86 Díaz-Vélez C, Failoc-Rojas VE, Valladares-Garrido MJ. SARS-CoV-2 seroprevalence study in Lambayeque, Peru. June–July 2020. *PeerJ*;2021:e11210.
87 Alemu BN, Addissie A, Mamo G. Sero-Prevalence of anti-SARS-CoV-2 antibodies in Addis Ababa, Ethiopia. *bioRxiv* 2020.
88 Shaweno T, Abdulhamid I, Bezabih L, *et al*. Seroprevalence of SARS-CoV-2 antibody among individuals aged above 15 years and residing in congregate settings in dire Dawa City administration, Ethiopia. *Trop Med Health* 2021;49:55.
89 Ngere I, Dawa J, Hunsperger E, *et al*. High seroprevalence of SARS-CoV-2 but low infection fatality ratio eight months after introduction in Nairobi, Kenya. *Int J Infect Dis* 2021;112:25–34.
90 Inquérito Sero-epidemiológico de SARS-CoV-2 na Cidade de Maxixe e Vila de Massinga (InCOVID 2020) - Resultados Preliminares -. República de Moçambique: Ministério da Saúde. 2020.
91 Barry A, Roka J, Talla C. Seroprevalence of Anti-SARS-CoV-2 antibodies in Senegal: a national population-based cross-sectional survey, between October and November 2020, 2021. SSRN. Available: http://dx.doi.org/10.2139/ssrn.3890833
92 Mutevedzi PC, Kawonga M, Kwatra G. Estimated SARS-CoV-2 infection rate and fatality risk in Gauteng Province, South Africa: a population-based seroepidemiological survey. *Int J Epidemiol.* 2021.
93 Sue Aitken JY, Fellows T, Makamadi T, *et al*. *COVID-19 seroprevalence during the second wave of the pandemic in three districts of south africa - preliminary findings: national institute for communicable diseases*, 2021.
94 Mulenga LB, Hines JZ, Fwoloshi S, *et al*. Prevalence of SARS-CoV-2 in six districts in Zambia in July, 2020: a cross-sectional cluster sample survey. *Lancet Glob Health* 2021;9:e773–81.
95 Khalagi K, Gharibzadeh S, Khalili D. Prevalence of COVID-19 in Iran: results of the first survey of the Iranian COVID-19 serological surveillance program. *medRxiv* 2021.
96 Bellizzi S, Alsawalha L, Sheikh Ali S, *et al*. A three-phase population based sero-epidemiological study: assessing the trend in prevalence of SARS-CoV-2 during COVID-19 pandemic in Jordan. *One Health* 2021;13:100292.
97 Al-Abri SS, Al-Wahaibi A, Al-Kindi H, *et al*. Seroprevalence of SARS-CoV-2 antibodies in the general population of Oman: results from four successive nationwide sero-epidemiological surveys. *Int J Infect Dis* 2021;112:269–77.
98 Merkely B, Szabó AJ, Kosztin A, *et al*. Novel coronavirus epidemic in the Hungarian population, a cross-sectional nationwide survey to support the exit policy in Hungary. *Geroscience* 2020;42:1063–74.
99 Zejda JE, Brożek GM, Kowalska M, *et al*. Seroprevalence of Anti-SARS-CoV-2 antibodies in a random sample of inhabitants of the Katowice region, Poland. *Int J Environ Res Public Health* 2021;18:3188.
100 Kshatri JS, Bhattacharya D, Praharaj I, *et al*. Seroprevalence of SARS-CoV-2 in Bhubaneswar, India: findings from three rounds of community surveys. *Epidemiol Infect* 2021;149:e139.
101 Selvaraju SKM, Thangaraj J, Bhatnagar T, *et al*. Population-Based serosurvey for severe acute respiratory syndrome coronavirus 2 transmission, Chennai, India. *Emerg Infect Dis.* 2021;27:586–9.

102 Sharma N, Sharma P, Basu S. The seroprevalence of severe acute respiratory syndrome coronavirus 2 in Delhi, India: a repeated population-based seroepidemiological study. *Trans R Soc Trop Med Hyg Suppl*. 2021.
103 Mohanan M, Malani A, Krishnan K. Prevalence of COVID-19 in rural versus urban areas in a low-income country: findings from a State-Wide study in Karnataka, India. *medRxiv* 2020.
104 Saple P, Gosavi S, Pawar T, *et al*. Seroprevalence of anti-SARS-CoV-2 of IgG antibody by ELISA: community-based, cross-sectional study from urban area of Malegaon, Maharashtra. *J Family Med Prim Care* 2021;10:1453–8.
105 Satpati P, Sarangi S, Gantait K. Sero-surveillance (IgG) of SARS-CoV-2 among asymptomatic general population of Paschim Medinipur, West Bengal, India. *medRxiv* 2020.
106 Banerjee A, Gaikwad B, Desale A, *et al*. Severe acute respiratory syndrome-coronavirus-2 seroprevalence study in Pimpri-Chinchwad, Maharashtra, India coinciding with falling trend - Do the results suggest imminent herd immunity? *Indian J Public Health* 2021;65:256–60.
107 Kar SS, Sarkar S, Murali S, *et al*. Prevalence and time trend of SARS-CoV-2 infection in Puducherry, India, August-October 2020. *Emerg Infect Dis* 2021;27:666–9.
108 Khan SMS, Qurieshi MA, Haq I, *et al*. Seroprevalence of SARS-CoV-2-specific IgG antibodies in Kashmir, India, 7 months after the first reported local COVID-19 case: results of a population-based seroprevalence survey from October to November 2020. *BMJ Open* 2021;11:e053791.
109 Malani A, Ramachandran S, Tandel V. SARS-CoV-2 seroprevalence in Tamil Nadu in October-November 2020. *medRxiv* 2021.
110 Enhanced surveillance on sero-prevalence of SARS-COV-2 in general population government of Nepal. 2020.
111 Haq M, Rehman A, Ahmad J, *et al*. SARS-CoV-2: big seroprevalence data from Pakistan-is herd immunity at hand? *Infection* 2021;49:983–8.
112 Li Z, Guan X, Mao N, *et al*. Antibody seroprevalence in the epicenter Wuhan, Hubei, and six selected provinces after containment of the first epidemic wave of COVID-19 in China. *Lancet Reg Health West Pac* 2021;8:100094.

Supplementary Appendices: Assessing the Burden of COVID-19 in Developing Countries: Systematic Review, Meta-Analysis & Public Policy Implications

Andrew Levin[1,2,3], Nana Owusu-Boaitey[4], Sierra Pugh[5], Bailey K. Fosdick[5], Anthony B. Zwi[6], Anup Malani [2,7], Satej Soman[8], Lonni Besançon[9], Ilya Kashnitsky[10], Sachin Ganesh[1], Aloysius McLaughlin[1], Gayeong Song[1], Rine Uhm[1], Daniel Herrera-Esposito[11], Gustavo de los Campos[12], Ana Carolina Pecanha Antiono[13], Enyew Birru Tadese[14], Gideon Meyerowitz-Katz[15,16]

1. Dartmouth College, Hanover, USA
2. National Bureau for Economic Research, Cambridge, USA
3. Centre for Economic Policy Research, London, United Kingdom
4. Case Western Reserve University School of Medicine, Cleveland, USA
5. Colorado State University, Fort Collins, USA
6. School of Social Sciences, University of New South Wales, Australia
7. Law School, University of Chicago, Chicago, USA
8. Mansueto Institute for Urban Innovation, University of Chicago, Chicago, USA
9. Faculty of Information and Technology, Monash University, Australia
10. Interdisciplinary Centre on Population, University of Southern Denmark, Denmark
11. Laboratorio de Neurociencias, Facultad de Ciencias, Universidad de la República, Uruguay
12. Department of Epidemiology & Biostatistics, Michigan State University, USA
13. Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil
14. Ethiopian Public Health Institute, Ethiopia
15. School of Health and Society, University of Wollongong, Australia
16. Western Sydney Local Health District, Sydney, Australia

Contents

1. Data Collection

a. Systematic Review Methodology

Search Procedure

We searched Serotracker using *"household and community samples"* and *"persons living in slums"* in the *"demographics"* field. We also searched MedRxiv, PubMed, Google Scholar, Biorxiv, SSRN, Twitter, and the Pan American Health Organization database using the pre-specified search term *"COVID-19 seroprevalence".* Then we cross-checked with recently published systematic reviews of worldwide seroprevalence (1-4), while identifying further studies by searching the grey literature and government websites where appropriate. This included searching using the term *"COVID-19 seroprevalence"* on Google in languages on Google Translate such as Portuguese, Spanish, English, French, German, and Italian, with an additional search for *"inquérito sorológico, COVID-19".* Duplicates were reviewed by authors on Google sheets and resolved independently.

We completed searches on October 22nd, 2020, and at least monthly afterwards until July 14, 2021. We also performed searches monthly until September 22, 2021 during initial drafting of our paper. After our completing initial draft, we performed additional searches monthly up to December 17, 2021, which represented the final cut-off date for studies included in our analysis. Only studies with results from an official source were included, such as a published paper, pre-print, presentation by government officials, or the website of the institution that performed the study. If a press report or another unofficial source was found, we performed more detailed searches using information from the unofficial source to find a matching official source. Study authors were contacted by email or Twitter for further information, when needed. We also ran detailed searches after September 22, 2021 on older studies for which preliminary results were found by September 22, but for which updates were posted after September 22. We include links to the studies at each location in the appendix folder of our GitHub repository.

Searches were conducted by one member of the team and then repeated to ensure consistency by another. Data were similarly extracted by one member then cross-checked by another. No data collection was automated. This process was recorded by the team working across regions in Google sheets. Data collection is more fully described below but included extracting seroprevalence information from included studies by age where available, as well as death data specific to COVID-19 from each country/region with valid seroprevalence information. Where serology data was not evident in publicly available reports, we reached out to researchers and public health officials using both email and social media. For death data we largely relied on publicly available official reports.

Studies were reviewed by two authors and screened for inclusion. Disagreements were resolved through discussion between all authors at weekly meetings and via email. Where essential data were missing despite efforts to access them, we excluded the study from our synthesis, as noted in supplementary appendix section 1.b. Our aim was to provide the most robust estimate of age-specific IFR in developing countries, and thus we considered it inappropriate to rely on potentially flawed assumptions regarding these studies in our analysis.

Exclusion of Convenience Samples

Blood donor studies are widely used as blood donors are a convenient population from which to draw a population estimate – donors already have blood taken, can be tested easily, and tend to include people from a relatively wide area (5). However, as has been noted in research prior to the pandemic, donors are a highly selected sample and donor studies often have a large bias in terms of estimates of seroprevalence for other diseases (6, 7). Moreover, in many areas, particularly at initial stages of the pandemic, donating blood was one of few methods available to access a serological test. It is unclear which direction this bias generally runs, especially considering the dynamics of a novel pathogen in the community (8).

Residual sera studies examine clinical blood samples taken initially for other reasons. These samples have an obvious bias in that they are representative of people going to have blood taken for reasons other than SARS-CoV-2 tests, a group that may not be representative of the general population (9). Bayesian procedures can be used to incorporate studies of convenience samples (such as residual sera from blood donors or commercial lab tests) in producing estimates of population infection rates by accounting for uncertainty about the magnitude and direction of bias (10); however, we excluded such studies from our analysis to avoid introducing these additional sources of bias. Convenience sampling of such populations may be sufficient for other purposes, but probabilistic selection from a representative sample frame better facilitates accurate estimation of population-wide infection rates (3, 11).

Risk of Bias

In assessing the risk of bias for each location, we considered three specific factors: (1) the serology study's rate of non-response; (2) the risk of bias due to seroreversion if the study used an assay with high risk of seroreversion but information was not sufficient for adjusting sensitivity accordingly; and (3) the risk of death undercounting was elevated due to low proportion of well-certified deaths. These risk of bias assessments are provided in the appendix folder of our GitHub repository.

Publication Bias

In this context, publication bias in which studies exhibiting certain findings are more likely (or not) to be published, is very unlikely to have an impact, as studies with both high and low seroprevalence estimates are of interest to the scientific literature. Consistent with this, in prior work we found no evidence of publication bias for seroprevalence studies from high-income countries (12). However, to mitigate the risk of publication bias influencing our results, we included lengthy searches of grey literature, following up on media reports of seroprevalence studies to ensure that every age-stratified that we were able to identify was in our metasynthesis.

b. Full Inclusion / Exclusion Criteria

We included only studies that met both of the following conditions:

1. Report seroprevalence from a representative sample in developing countries, meaning: random selection of participants from a sample frame representative of the general population, such as household sampling, or sampling >50% of the general population by census (13-15), conducted in countries classified by the International Monetary Fund as "Emerging and Developing Economies" (16).
2. Available online and accessible in English, or via translatable text if not in English.

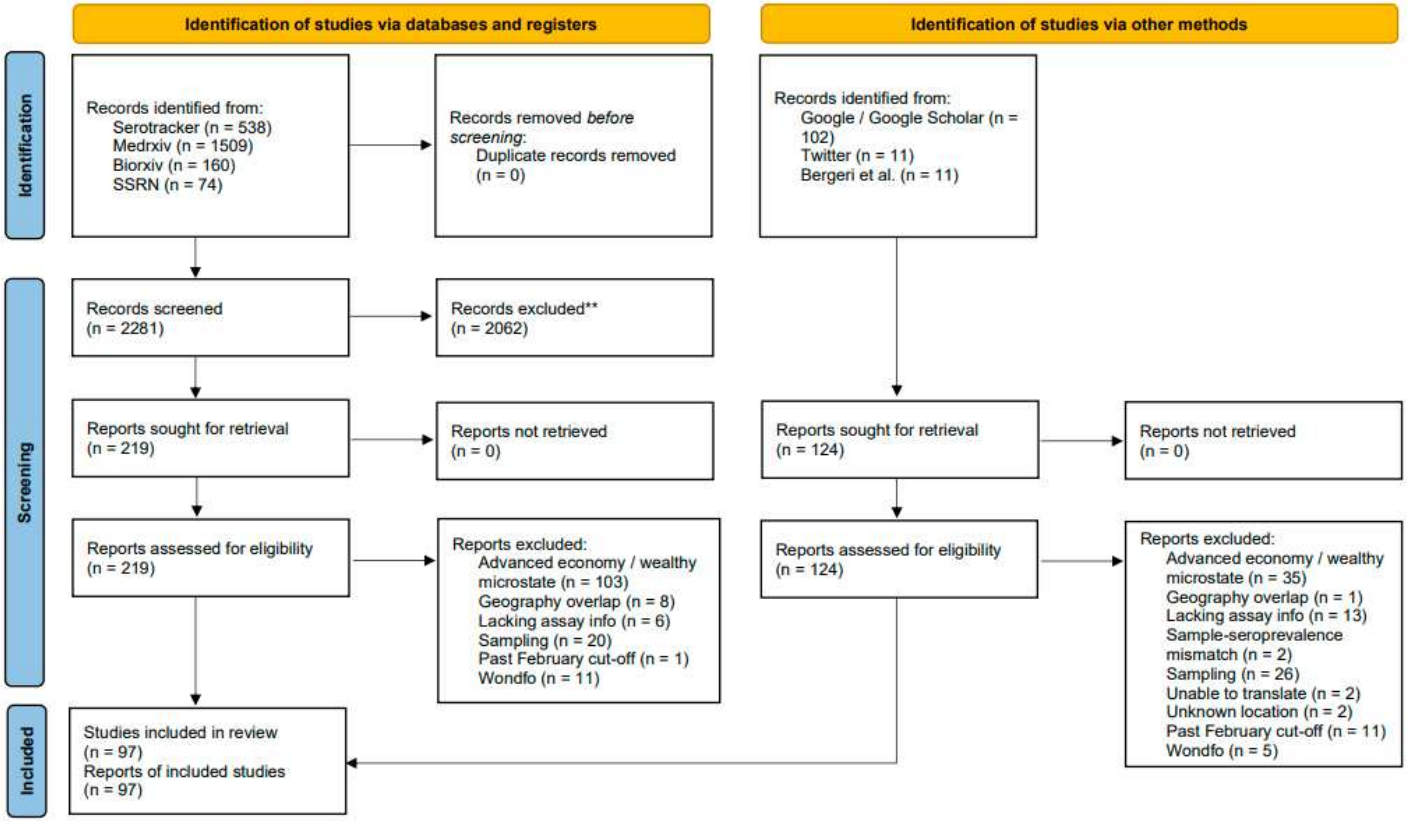
For studies with no reported age-stratified seroprevalence, but sufficient information to otherwise calculate age-stratified IFRs, we calculated these IFRs assuming equal seroprevalence across age-groups instead of excluding the study. When total sample size and age-specific seroprevalence were known, but age-specific sample sizes were not precisely reported, age-specific sample size was imputed based on the age-distribution of the general population. We excluded:

1. Convenience samples (3), including those utilizing residual sera from clinical specimens and blood donors (see the "*Blood Donors and Residual Sera*" section below), dialysis centres, healthcare workers, and actively recruited participants constituting less than 50% of the total population sampled (3, 12, 17).
2. Studies sampling a high-income country, as classified by the International Monetary Fund (16), or a wealthy micronation such as Andorra or Monaco.
3. Studies in which sampling extended after February 2021, to help mitigate the risk of seroreversion on longer timeframes (see supplementary appendix section 2.a).
4. If gender ratios were reported and less than 35% of the sample reported as male or female, in the absence of cited evidence that the study's gender ratio matched the general population.
5. Studies that used the Wondfo serology assay, for the reasons discussed in section 2.b below.
6. Studies that did not report the total number of individuals tested or seroprevalence.
7. Studies using serology assays with insufficient data for estimating sensitivity and specificity from a known number of tested samples.
8. IFR estimate excluded if: A) the sampling start-week or end-week was not known to allow for accurate determination of the corresponding number of COVID-19 deaths, B) test-adjusted population-wide seroprevalence overlapped with 0%, or C) samples were taken during an accelerating outbreak in which reported COVID-19 deaths increased by a factor of three or more from the midpoint date of sampling to 4 weeks later (12).
9. Seroprevalence estimate excluded if both of the following conditions were met: A) IFR estimate was excluded for other reasons listed above, and B) the study overlapped geographically with another included study. This geographical exclusion avoided oversampling the same location (12). IFR estimates that met condition B but not condition A are discussed in the out-of-sample analysis below.

Our "out-of-sample" analysis included studies that met at least one of the following conditions:

1. IFR estimate for a location that geographically overlapped with an included study, and thus inclusion of both estimates risked oversampling the same location. IFR estimates for those locations are provided in supplementary appendix section 3.i.
2. Zero COVID-19 deaths reported for the sampled location, making the calculated IFR non-robust (18). Consequently, our GitHub repository includes seroprevalence estimates for these out-of-sample locations, but IFR estimates were not computed.
3. The total population from which the sample was drawn was less than 30,000, which may not reflect the wider population of the region. Consequently, our GitHub repository includes seroprevalence estimates for these out-of-sample locations, but IFR estimates were not computed.
4. Seroprevalence data from five cities in Pakistan did not become available until after our final cutoff.

c. PRISMA Flow Diagram



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).
**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

d. Death Data

Building on our prior work (12), we assess the length of the lag between the midpoint of serology sampling and the time at which COVID-19 deaths were reported. The time interval between symptom onset and death had an interquartile range (IQR) of:

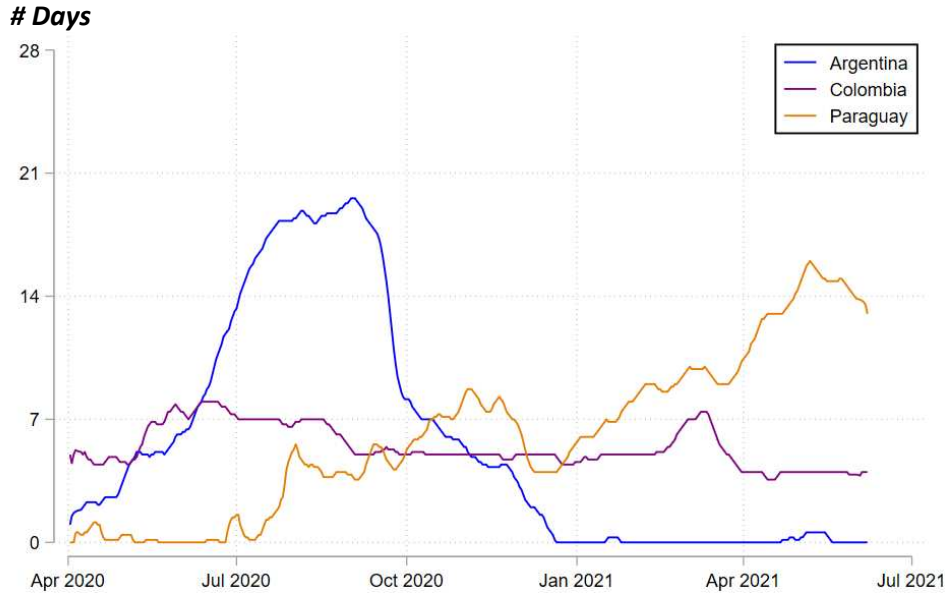
- 7 to 22 days for Argentina (19)
- 9 to 24 days for Colombia (20)
- 10 to 26 days for the Brazilian states of Espírito Santo (21) and Parana (22)

These intervals largely agree with IQRs reported for the USA as:

- 9 to 24 days for ages 18-64
- 7 to 19 days for ages >64 (12)

The IQR for the interval between death and official reporting for the USA was 2 to 19 days (12). This largely matches the interval ranges for Argentina (19), Colombia (20), and Paraguay (23) before March 2021 when included seroprevalence studies stopped collecting samples (see section 1.b), as shown below:

Figure A1 – Death Reporting Lags



Argentina, Colombia, and Paraguay may be outliers with respect to the systematic collection and publication of vital statistics during the pandemic (24); so other developing countries may have substantially longer reporting lags that may not be documented in the absence of detailed case data. The time interval between symptom onset and official death reporting thus appears roughly similar in developing countries as in our previous analysis of high-income countries such as the USA (12).

For some study locations, we were able to extract COVID-19 fatality data from case databases that specified the actual date of death; in those instances, we used the cumulative number of fatalities as of two weeks after the midpoint date of serology specimen collection. In other locations, fatality data was only available from official epidemiological bulletins, in which the official number of cumulative deaths announced at a particular date reflected reporting lags. In those instances, as in

our previous work (12), we extracted death information four weeks after the midpoint of specimen collection. These timing specifications reflect approximate 95th percentiles as follows:

- 2-week interval between symptom onset and seropositivity,
- 4-week interval between symptom onset and death,
- 2-week interval between death and official reporting.

There is also the question of what is the most appropriate death data to use. In most countries there are two sets of COVID-19 deaths: confirmed or suspected. In some countries the government will also present a third tally of deaths, which is modelled using excess mortality statistics or similar. Confirmed COVID-19 deaths may under-estimate the total number of COVID-19 deaths due to insufficient testing (25-28). This may be detected by comparing reported COVID-19 deaths with excess deaths, as reflected in the Peruvian government increasing their tally of reported deaths in a manner that better approximated total excess deaths (24, 29). Nepal's government also later substantially increased their reported tally of COVID-19 deaths (30).

Problems with death reporting are well-illustrated by the case of Mexico, a country whose vital statistics system has notable gaps and which experienced a huge number of COVID-19 deaths. When looking at the raw data on individuals, 90% of those who died did not have a date of death entered into the publicly available data. Previous research also demonstrated that large numbers of people who died from COVID-19 in Mexico failed to access a test and thus are not included in the country's mortality statistics (28). This means that the reported death data available for Mexico is not sufficient to derive a high-quality COVID-19-related IFR. We therefore instead used an alternative official source in Mexico that accounted for this COVID-19 death under-estimation (31).

So for the purposes of the primary analysis, we included the confirmed + suspected death figures where available instead of only confirmed deaths, as confirmed + suspected is the more robust estimate of reported COVID-19 deaths in developing countries. Death data were extracted from national datasets in each country where possible, with alternative sources noted where applicable. Where death data were not immediately available, we contacted the national or local authority through email or social media. We also attempted to confirm death data using the most robust source, and in most cases took the estimate directly from the relevant health authority rather than data aggregation websites. We include our informal assessment of risk of COVID-19 death under-estimation in the appendix folder of our GitHub repository. This is assessment is based on percentage of deaths well-certified in the past decade (32), and on comparison of reported COVID-19 deaths to excess deaths.

e. Assay Characteristics and Seroconversion

For the assays used in the serology studies that were included in our analysis, we catalogued the assay manufacturer's estimates of sensitivity and specificity as well as third-party assessments of its performance characteristics. In addition, we conducted a review to assess serological assays for risk of seroreversion. This review was restricted to studies that tested the same individuals at two different time-points separated by at least two months, or tested individuals at least two months after their first known positive test for SARS-CoV-2. We placed emphasis on commercial assays or assays used in seroprevalence studies.

The search used the terms *"COVID-19 seroreversion"* and *"COVID-19 longitudinal, antibody waning"* in Medrxiv, Biorxiv, Google Scholar, and SSRN. Searches were completed at least monthly from March 2021 to June 2021, with the final search performed on June 30, 2021. We supplemented this with seroreversion studies found during searches up to July 14 for seroprevalence studies with representative sampling, and for which further information was released after July 14 (see "*Systematic Review Methodology*"). Finally, we selected the studies that contained information

about assays that had been used in the serology studies included in our analysis (as described in the preceding subsections of this appendix).

Our systemic review of assay characteristics revealed that the Wondfo assay exhibited extreme variations in test sensitivity across batches, apparently reflecting defects in its manufacturing process (33-35); consequently, any serology study which used this assay was excluded from our analysis.

For seroprevalence to approximate the number of people infected, almost all infected people need to seroconvert by increasing antibody levels after infection. Studies of large populations suggest that >85% or >90% of SARS-CoV-2-infected individuals seroconvert by approximately 2 weeks after infection (36-39). This increases confidence in the accuracy of seroprevalence-based infection estimates that use tests with sufficiently high sensitivity (36, 40, 41). Moreover, >50% of the total population seroconverted in several locations, which would not occur if a substantial proportion of infected individuals failed to seroconvert. Table A1 illustrates this with studies reporting >50% seroprevalence before the onset of widespread SARS-CoV-2 vaccination:

These high seroprevalence estimates may shed light on high vs. low herd immunity thresholds (42-44). For example, Leticia suffered another wave of SARS-CoV-2 infections after reported seroprevalence of 62%, as did Delhi after reported seroprevalence of 56%, the state of Maranhão after reported seroprevalence of 40%, and Jordan after reported seroprevalence of 34% (20, 45, 46). Cross-reactivity also likely does not account for elevated seroprevalence in many of the regions listed in table A1, since cross-reactivity did not significantly reduce test specificity in locations such as Colombia, Ethiopia, and Iran (47-49). These high seroprevalence estimates instead imply that the vast majority of infected individuals seroconverted, increasing the reliability of seroprevalence-based infection estimate (50).

Some serological assays exhibit significantly lower specificity in African populations, possibly due to cross-reactivity with other pathogens (48, 51). This may contribute to divergent seroprevalence estimates between two studies performed in Addis Ababa, Ethiopia (49, 52) (see supplementary appendix section 3.i). However, specificity likely remains high in African populations for many of the assays used in our included studies (49, 53, 54).

Finally, it should be noted that the Gladen-Rogan procedure of adjusting for assay specificity and sensitivity (55)) assumes that those characteristics are precisely known, without accounting for the uncertainty that comes with inferring characteristics from a limited number of tested samples (56). By contrast, our statistical model uses Bayesian methods that incorporate this form of uncertainty.

Table A1 - Locations with Seroprevalence Exceeding 50%

Region	Location	Reported seroprevalence	Number tested for seroprevalence estimate
Latin America	Argentina: Buenos Aires (Barrio Padre Mugica)	53.4% (CI: 52.8 - 54.1%)	873
	Argentina: Metropolitan Area of Buenos Aires (17 de Noviembre)	56.7% (CI: 55.8 - 57.6%)	300
	Colombia, 10 cities: Barranquilla	53% (CI: 41 - 65%)	1426
	Colombia, 10 cities: Guapi	78% (CI: 65 - 91%)	721
	Colombia, 10 cities: Leticia	62% (CI: 51 - 73%)	1417
	Colombia, Córdoba: Montería	55.3% (CI: 52.5 - 57.8%)	1368
	Peru: Iquitos; July, August	70% (CI: 67 - 73%) 66% (CI: 62 - 70%)	716 621
Africa	Ethiopia: Addis Ketema	54.2% (CI: 47.5 - 60.7%)	218
Middle East	Afghanistan: Kabul	53% (CI: < +/-6%)	-
	Iran, 18 cities: Qom, Rasht	58.5% (CI: 37.2 - 83.9%) 72.6% (CI: 53.9 - 92.8%)	108 99
	Iraq: Duhok city	62.6%	743
South Asia	Bangladesh: Dhaka (“slums”)	74%	-
	India: Delhi	56.1% (CI: 55.5 - 56.8%)	28,169
	India: Hyderabad	54.2% (CI: 53.2 - 55.2%)	9363
	India: Karnataka (urban areas)	53.8% (CI: 48.4 - 59.2%)	453
	India: Mumbai, 3 “slums” (in: Chembur West, Dahisar, Matunga)	56.4% 55.1% (CI: 52.4 - 57.8%) 51.1% (CI: 46.4 - 55.8%) 57.0% (CI: 54.7 - 59.2%)	4202 1511 570 2121
	India: Pune, 5 subwards (Lohiyanagar-Kasewadi, Navi Peth-Parvati, Yerwada)	51.3% (CI: 39.9% - 62.4%) 66.4% (CI: 57.8% - 74.1%) 54.1% (CI: 48.3% - 61.7%) 55.5% (CI: 46.6% - 64.1%)	1659 307 331 367

Notes: CI refers to confidence interval. This analysis is restricted to studies with at least 75 people tested. Links to these studies are provided in the Appendix folder of our GitHub repository.

f. Covariates

We extracted data from the most recent year prior to the pandemic, which in most cases was 2019 or earlier. For some estimates such as workforce, we relied on the best available data, some of which was several years old for some countries.

The covariates are:

1. GDP per capita
2. Healthcare spending
3. GNI per capita
4. Hospital beds per capita
5. Life expectancy at birth
6. Healthy life expectancy at age 60
7. Global health security index
8. Skilled healthcare workers per capita
9. Universal health coverage index
10. % of deaths well-certified (32)

Briefly, these covariates were chosen because they either relate to the expected quality of the health system itself (i.e. doctors/nurses per population) or to how likely a country was to be accurately recording the burden of COVID-19 (WHO indicators, human development index). We also included the ratio of life expectancy between age 60 and 20 to account for the potential for survivorship bias – if there was a significant element of survivorship bias in the countries examined, we would expect the ratio to be higher as more elderly people survived longer periods in places with higher mortality in youth.

2. Statistical Methodology

a. Adjustment for Seroreversion

Seroreversion occurs when the specific antibodies a serological assay tests decline to below the assay's level of detection, preventing the assay from identifying infected individuals. As many studies conducted in developing countries were performed long after initial COVID-19 waves passed, the risk of seroreversion could be high. This could lead to underestimation of the proportion of infected people and thus unreliable estimates in our computed IFRs. Moreover, any modelling using assumptions about seroreversion for one serological test would almost certainly lead to errors in other places as different tests can substantially differ in characteristics (57).

For all other assays that were used in the serology studies included in our analysis, we classified each assay's risk of seroreversion (high, medium, or low) based on two sources of data:

- *Longitudinal serology studies*, in which specimens were collected periodically from a given sample of individuals over an extended period of time.
- *Serology analysis of prior RT-PCR positive cases*, i.e., collection of specimens from individuals who had previously tested positive for COVID-19.

Some seroprevalence studies tested a representative sample of the general population, including those with a prior positive SARS-CoV-2 PCR test weeks or months before serological testing, a previous COVID-19 diagnosis weeks or months before serology, etc. If many of these prior-positive individuals later tested seronegative, then that is unlikely to represent failed seroconversion, as previously discussed. It instead likely indicates a high risk of seroreversion during the time following their initial positive test (58). A threshold of <75% sensitivity was selected for this risk of seroreversion because at least 75% of prior-positives tested seropositive using the Roche assay that is at low risk of seroreversion (see Table A2), and the vast majority of sources reported sensitivity of at least 75% before seroreversion, as shown in the input data of our GitHub repository.

Table A2 indicates our assessment of the seroreversion risk of each assay for which sufficient information was available from longitudinal data or analysis of prior confirmed RT-PCR positive cases. For each assay, this table also shows the locations for which we have estimated IFR from serology results obtained using that assay.

Although not shown in the table, three of these assays were also used to estimate seroprevalence in *"Sero-Only"* locations where IFR could not be estimated due to lack of corresponding fatality data: (1) *Roche Elecsys (anti-nucleocapsid)* was used in Duhak, Iraq; Hyderabad and Rourkela, India; Gaza and West Bank, Palestine; and Jourberton, South Africa. (2) *Euroimmun IgG* was used in Tirana, Albania; Pune, India. (3) *Wantai IgG/IgM* was used in Cox's Bazar Rohingya camps, Bangladesh; Georgia *(4 districts)*; Malaysia *(nationwide)*; Mongolia *(nationwide)*, Klerksdorp & Pietermaritzburg, South Africa; and Phuket, Thailand.

For every location for which the assay used in serology was classified as having high risk of seroreversion, we made corresponding adjustments to the data on assay sensitivity as follows:

- *Abbott Architect assay.* We used information from prior studies to assess how the sensitivity of this assay diminishes over time following the onset of infection at each monthly interval from 0 to 6 months. For each of the seven locations where this assay was used, we

computed its weighted sensitivity as of the midpoint date of the serology study, where the weights were determined by the time path of confirmed SARS-CoV-2 cases in that location.

- *Other assays with high risk of seroreversion*. For each of the three locations that used such assays, we extracted information about the seropositivity of specimens from individuals with a prior positive RT-PCR test. This approach automatically accounts for the variation in time intervals since infection, because each of these serology studies used a representative sample of the general population, and is consistent with prior work on how waning of antibodies reduces the proportion of prior-positives who test seropositive (58, 59).

Table A2 – Seroreversion Assessments and Sources

Risk Category	Assay	IFR Locations	Seroreversion Data: *Sequential Tests*	Seroreversion Data: *Prior Positives*	Citations
High	Abbott Architect IgG	Ethiopia: *Dire Dawa*	X		(37, 60-70)
		Hungary *(nationwide)*	X		
		Bosnia & Herzegovina: *Republika Sprska*	X		
		India: *Chennai, Mumbai, Pimpri-Chinchwad, Srinagar*	X		
	ErbaLisa IgG	India: *Paschim Medinipur*		X	(71, 72)
	Zydus Kavach IgG	India: *Delhi*		X	(71)
	Luminex S	South Africa: *Gauteng*		X	(58)
Moderate	Euroimmun IgG	Zambia: *Lusaka & Ndola*	X	X	(67, 73-77)
		Poland: *Katowice*	X		
	Roche Elecsys IgG/IgM *(anti-nucleocapsid)*	Brazil: *Maranhao, Sao Paulo*	X		(37, 38, 64, 69, 70, 73, 78-83)
		Chile: *3 urban areas*	X	X	
		India: *Berhampur, Bhubaneswar, Puducherry*	X	X	
		Mexico *(nationwide)*	X	X	
		Pakistan: *Karachi, Lahore*	X		
Low	COVIDAR IgG	Argentina: *Buenos Aires City, Hurlingham*	X		(84, 85)
	DiaSorin Liaison IgG	Brazil: *Cuiabá, Mato Grosso, Pitangueiras, Várzea Grande*	X		(37, 57, 60, 69, 70, 86)
		Oman *(nationwide)*	X		
	Genetico CoronaPass Total	Russia: *St. Petersburg*	X		(63, 65)
	Ortho Vitros IgG	India: *Tamil Nadu*	X		(57)
	Roche Elecsys IgG/IgM *(anti-spike)*	N/A	X		(64, 87)
	Siemens Advia IgG/IgM	Colombia: *Barranquilla, Bogotá, Bucaramanga, Cali, Cucuta, Ipiales, Leticia, Medellin, Villavicencio*	X		(69, 70)
	University of Rio de Janeiro	Brazil: *Rio Grande do Sul*	X		(33, 88)
	Wantai SARS-CoV-2 Total	Jordan *(nationwide)*	X		(64, 89, 90)
		Kenya: *Nairobi*	X		
		Nepal *(nationwide)*	X		
		Senegal *(nationwide)*	X		
		South Africa: *Mitchells Plain*	X		

Note: *This table shows the seroreversion risk category assigned to each assay for which sufficient information was available.*

- *Assays with medium risk of seroreversion*. For locations that used either of these assays, we extracted information about the seropositivity of specimens from individuals with a prior positive RT-PCR test, and we utilized that data if the seropositivity rate was less than 75% (corresponding to a significant degree of seroreversion in that location.)

Given the seroreversion-adjusted sensitivity for each location, we imputed the corresponding number of seropositive specimens that would be obtained using the actual sample size for that serology study, and then those values serve as inputs to the Bayesian model described below. This approach is conceptually similar to prior studies that have imputed the number of specimens and the number of confirmed cases by inverting seroprevalence confidence intervals (10, 91).

Finally, Table A3 lists the assays for which seroreversion could not be assessed due to insufficient information. For each assay, this table shows the IFR and "Sero-Only" locations for which we relied on the baseline characteristics of that assay.

Table A3 – Assays with Unknown Seroreversion

Assay	IFR Locations	Sero-Only Locations
Abbott PanBio IgG/IgM	Mozambique: *Maputo*	Mozambique: *Beira, Chókwè, Matola, Quelimane, Xai-Xai* Pakistan: *Islamabad*
Beijing Kewei IgG/IgM	Paraguay: *Asuncion & Central Dept.*	
Bioscience IgG/IgM	China: *Wuhan, Hubei ex. Wuhan*	China: *3 provinces*
Core Technology IgG	Ethiopia: *Addis Ababa*	Ethiopia: *3 towns*
Coretest IgG/IgM	Peru: *Lambayeque*	
CTK Biotech Onsite IgG/IgM	Brazil: *Distrito Federal*	
ECO IgG/IgM	Brazil: *Pitangueiras*	
GenBody IgG		Dominican Republic: *10 provinces*
Healgen IgG/IgM		Yemen: *Aden*
INgezim DR IgG/IgM/IgA	Colombia: *Córdoba*	
Karwa Kavach IgG	India: *Malegaon*	India: *Indore, Jabalpur*
Luminex N		Dem. Rep. of Congo: *Kinshasa*
Orient Gene Biotech IgG/IgM	Brazil: *Iquitos, Loreto*	
Pishtaz Teb IgG/IgM	Iran *(nationwide)*	
Proprietary assay #1	Brazil: *Foz do Iguaçu*	
Proprietary assay #2		Laos: *5 provinces*
Proprietary assay #3		Zimbabwe: *Budiriro & Highfield*
Proprietary assay #4		South Sudan: *Juba*
Qingdao Hightop IgG/IgM		Mozambique: *Chimoio, Tete, Massinga, Maxixe, Pemba*
RightSign IgG/IgM		Libya: *Benghazi*
Shenzhen iFlash IgG	India: *Tamil Nadu* Ecuador: *Cuenca*	
Standard Q IgG/IgM	Peru: *Lima & Callao* Bolivia: *Santa Cruz*	Mozambique: *Nampula*
THSTI IgG	India: *Karnataka*	
UNCOV-40 IgG/IgM		Nigeria: *Niger State*

b. Bayesian Model for Estimating Seroprevalence and IFR

Model for COVID-19 infections

Let $R^{\star}_{\mathrm{l},A}$ be the number of individuals who tested seropositive in age group *A* at location l, and $n_{\mathrm{l},A}$ give the number of individuals tested in that age group for this location. We model the number of individuals with a positive serology test in the study as

$$R^{\star}_{\mathrm{l},A} \sim \text{Binomial}\left(n_{\mathrm{l},A}, p_{\mathrm{l},A}\right), \text{ where} \tag{1}$$

$$p_{\mathrm{l},A} = \text{sens}_{t_{\mathrm{l}}}\pi_{\mathrm{l},A} + \left(1 - \text{spec}_{t_{\mathrm{l}}}\right)\left(1 - \pi_{\mathrm{l},A}\right). \tag{2}$$

To account for the error rates of the test, the test positivity probability, $p_{\mathrm{l},A}$, is defined as a function of test sensitivity ($\text{sens}_{t_{\mathrm{l}}}$), test specificity ($\text{spec}_{t_{\mathrm{l}}}$), and the true seroprevalence ($\pi_{\mathrm{l},A}$) for the associated location and age group at the time of the study. For many studies, we did not have seropositivity by age, in which case *A* represented all ages.

To account for uncertainty in the test characteristics, we model the lab validation data directly. Let $n_{\text{sens},t}$ denote the number of positive specimens tested with test *t*, and $x_{\text{sens},t}$ the number of positive specimens that correctly tested positive. Similarly, let $n_{\text{spec},t}$ and $x_{\text{spec},t}$ denote the number of negative specimens tested and the number of negative specimens that correctly tested negative with test *t*, respectively. We model these quantities as follows:

$$x_{\text{sens},t} \sim \text{Binomial}(n_{\text{sens},t}, \text{sens}_t) \tag{3}$$

$$x_{\text{spec},t} \sim \text{Binomial}(n_{\text{spec},t}, \text{spec}_t). \tag{4}$$

Model for COVID-19 deaths

Let $D^{\star}_{\mathrm{l},A}$ give the number of recorded COVID-19 deaths, for age group *A* at location l. Note that if only a single death record is available, then *A* represents the entire age range. We model the recorded COVID-19 deaths as

$$D^{\star}_{\mathrm{l},A} \sim \text{Poisson}\left(N_{\mathrm{l},A} \times \pi_{\mathrm{l},A} \times \text{IFR}_{\mathrm{l},A}\right) \tag{5}$$

where $N_{\mathrm{l},A}$ gives the number of individuals at location l in age group *A*. Then $N_{\mathrm{l},A} \times \pi_{\mathrm{l},A}$ gives the expected number of infected individuals, and $\text{IFR}_{\mathrm{l},A}$ is the infection fatality rate for location l and age group *A*, representing the probability an individual dies from COVID-19, given the individual had COVID-19. Note, the Poisson distribution reflects the relative rarity of a COVID-19 fatality relative to the entire population.

Accounting for data collected in varying age bins

Notice that the models above for deaths and infections in (1) and (5) are functions of prevalence and IFR, respectively, defined on discrete age bins. However, the discrete age bins are not necessarily the same for the death data and the seroprevalence studies. The following adjustments were made to match serology and death age bins:

- **Death bins nested within a serology bin:** We aggregate deaths for each location to match the respective serology age bins to avoid placing assumptions about the variability of prevalence across ages within a single serology age bin.
- **Serology bins nested within a death bin:** The average seroprevalence for the death age bin is calculated as an average of the serology age bins, weighted by the percent of the population in each age bin.

- **Bin endpoints slightly off:** When age bins were within one or two years of matching, serology age bins were adjusted to match the corresponding death age bins.

All modifications to age bins are documented in a spreadsheet in the data folder.

Population Age Distribution

Let $f_l(a)$ denote the number of individuals of age *a* at location l for *a* ∈ (0,1,...,84+). Note, if population age structure is only available in 5 year age bins, then define

$$f_l(a) = \sum_{b\in[0,5,\ldots,80]} \frac{f_l\big([b,b+5)\big)}{5} I_{[b,b+5)}(a)$$

where $f_l\big([b,b+5)\big)$ is the proportion of the population ages $[b,b+5)$.

In cases where the location specific age structure is only available in large bins, but the national age structure is available in 5 year age bins, we leverage the national age structure to inform the location specific age structure as follows. Let *A* denote an interval the location specific age structure is available for (e.g., [0,18)). If *f*(*A*) is the proportion of the population at location l with an age in *A* and $f_n(a)$ is the proportion of the population aged *a* at the national level, then we estimate $f_l(a)$, the proportion at location l that is age *a*, as

$$f_l(a) = f(A)\frac{f_n(a)}{\sum_{b\in A\cap N} f_n(b)}. \quad (6)$$

Essentially, we rescale $f_n(a)$ such that the total mass in *A* matches the observed total mass in *A* at location l, *f*(*A*). Since we model seroprevalence as constant past age 85, we let $f_l(85)$ represent the proportion of the population aged 85 or older, rather than just the proportion aged 85.

Calculating Average Seroprevalence within a Death Age Bin

Define the population age density for age bin *A* as

$$f_{l,A}(a) = \frac{f_l(a)}{\sum_{b\in A\cap N} f_l(b)},\ a \in (0,1,\ldots,84+) \quad (7)$$

in order to truncate $f_l(a)$ to age bin *A*.

The prevalence for age bin $B = \cup_{A\in\mathcal{A}} A$ is then defined

$$\pi_{l,B} = \sum_{A\in\mathcal{A}}\big[\pi_{l,A}\sum_{b\in A\cap N} f_{l,A}(b)\big]. \quad (8)$$

For the locations where we only have serology study information with no corresponding fatality data, the proportion of all study participants that were in a given age bin was assumed representative of the proportion of the population in each age bin since the studies were designed to have representative samples. For the locations with both serology and fatality data, population data were recorded in the Population Distributions tabs with citations.

Priors

Because there are infinitely many combinations of prevalence, sensitivity, and specificity that can result in the same test positivity rate, we used weakly informative priors for the seroprevalence parameters and informative priors for sensitivity and specificity to avoid a multimodal posterior, similar to Gelman and Carpenter (2020). For the seroprevalence parameters, $\pi_{l,A}$, we used independent, weakly informative priors:

$$\pi_{l,A} \sim \text{Beta}(2,6) \quad \text{for all l}, A. \tag{9}$$

These priors assume a mode around 0.15 with a prior probability of 0.8 that $\pi_{l,A}$ falls between 0 and 0.8.

We also used independent priors for the test sensitivities and specificities. For each test assay *t*, the priors on the sensitivity and specificity were

$$\text{sens}_t \sim \text{Beta}(10,1) \tag{10}$$

$$\text{spec}_t \sim \text{Beta}(50,1). \tag{11}$$

To further narrow the seroprevalence, sensitivity, and specificity combinations, we used independent, mildly informative priors for each IFR parameter based on expert knowledge. IFR for COVID-19 is known to increase with age. We also expect IFR to be more extreme (smaller than average or larger than average) when the age bin is small. For example, we would expect an age bin from 20-80 to look similar to the country average, but we would expect an age bin from 70-80 to be much higher than the country average. To formulate a prior that reflects these characteristics, we modeled

$$\text{IFR}_{l,A} \sim \text{Beta}\left(1, \text{IFR}_{l,A}^{\text{prior}}\right) \tag{12}$$

where

$$\text{IFR}_{l,A}^{\text{prior}} = 30 - 20\left[\frac{U_{l,A}-50}{50}\left(1 - \frac{U_{l,A}-L_{l,A}}{100}\right)\right]. \tag{13}$$

The lower and upper bounds of age bin *A* at location l are given by $L_{l,A}$ and $U_{l,A}$, respectively. For open-ended upper ages, we set $U_{l,A} = 100$. As an example, $\text{IFR}_{l,[0,100)}^{\text{prior}} = 30$, while $\text{IFR}_{l,[80,100)}^{\text{prior}} = 14$, allowing for larger IFR estimates when focusing on the older individuals.

Model Implementation

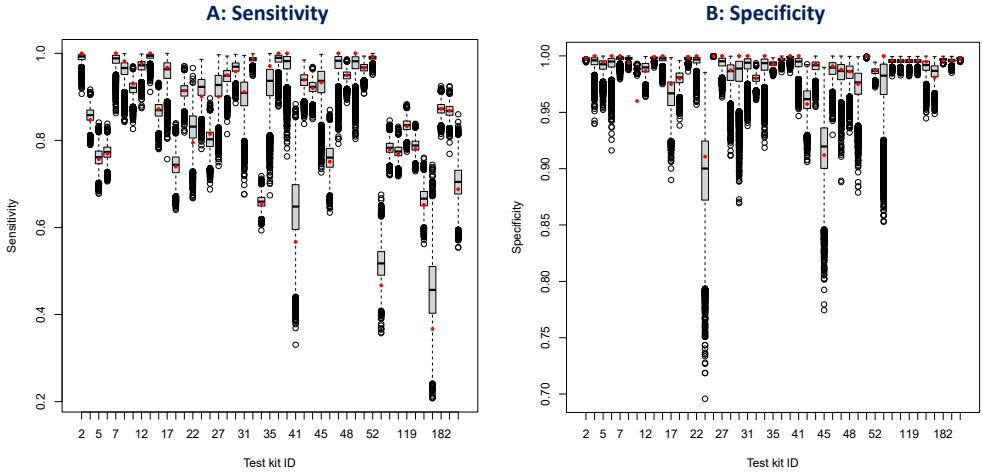
The model was implemented in version 4.0.2 of the programming language R, and posterior samples were obtained via the software package Stan (version 2.21.1). We ran three chains for 10,000 iterations, where the first 5,000 iterations were discarded as warm-up samples. All parameters had an effective sample size greater than 1,200. Additionally, the $\hat{R}$ value was within 0.0016 of 1 for each parameter, suggesting convergence. Examination of traceplots also suggested convergence.

Out-of-sample observation were run as a separate model, so information on test sensitivity and specificity was not pooled between in-sample observations and out-of-sample observations. Traceplots, effective sample size (minimum of 2100), and $\hat{R}$ values (within 0.0029 of 1) suggested convergence of the out-of-sample model as well.

We compared plugin estimates for parameters to the posterior distribution for each parameter to check model fit. In each case there was good agreement, or the Bayesian estimate was superior. For example, in Figure A2, we compare the posterior distribution of the sensitivity and specificity estimates to the raw estimate. In most cases, the middle 50% of the posterior distribution contains

the raw estimate. However, for test kit ID 11 (Qingdao Hightop Biotech IgM/IgG Duo), the raw estimate of specificity is outside the range of the posterior draws. In the case of this test, it was used in locations with extremely low prevalence, such that the Gladen-Rogan (55) adjustment results in a negative estimate of prevalence (meaning the expected number of false positives is greater than the number that tested positive.) Since this is unreasonable, the Bayesian model raises the specificity estimate, lowering the expected number of false positives.

Figure A2 – Sensitivity and Specificity Posterior Distributions



Note: These boxplots show the posterior distribution of each assay's sensitivity (panel A) and specificity (panel B) compared to the raw estimate (red dot) based on lab validation data. Further details (including the name of each assay) are given in the appendix folder of our GitHub repository.

Model Outputs

For each model parameter, we use the posterior mean as the point estimate and produce 95% equal-tail credible intervals to describe uncertainty.

Total seroprevalence

Similar to calculating the average seroprevalence for a death age bin, we estimate total seroprevalence for a location by taking an average of the age bin seroprevalences, weighting by the population distribution at that location:

$$\pi_{l,[0,100+)} = \sum_{A \in \mathcal{A}_\ell} \left[\pi_{l,A} \sum_{b \in A \cap N} f_{l,A}(b) \right] \quad (14)$$

where $\mathcal{A}_\ell$ are the serology age bins associated with location l.

Assessing Uniformity of Seroprevalence Across Age

We calculated total seroprevalence for younger adults and middle aged adults, compared to older adults. The age bins used for each location were selected as follows:

- **Younger adults (approximately 18 to 59):** Any age bins such that 15 ≤ lower age < 60 and 20 < upper age ≤ 65 were included

- **Middle aged adults (approximately 40 to 59):** Any age bins such that 40 < lower age < 60 and 40 < upper age ≤ 65 were included
- **Older adults (approximately 60 and older):** Any age bins such that 60 ≤ lower age were included.

This resulted in sets of bins where the 18-59 bins and the 40-59 bins did not overlap the 60+ bins.

We then calculated total seroprevalence from age *a* to *b* using appropriate age bins, $\mathcal{A}$, similar to equation (8)

$$\pi_{l,a-b} = \sum_{A\in\mathcal{A}} \pi_{l,A} \frac{n_{l,A}}{\sum_{B\in\mathcal{A}} n_{l,B}} \quad (15)$$

where $\frac{n_{l,A}}{\sum_{B\in\mathcal{A}} n_{l,B}}$ estimates the percent of the total population in that age bin, assuming representative age distributions in the serology studies.

For each draw from the posterior distribution, we calculated $\pi_{l,a-b}$ for each of our three age intervals of interest. We then calculated the ratios $\frac{\pi_{l,60+}}{\pi_{l,18-59}}$ and $\frac{\pi_{l,60+}}{\pi_{l,40-59}}$ for each draw.

Total IFR and Comparison to High-income (EJE) Prediction

Total IFR

Suppose location l has death age bins $\mathcal{A}_\ell$. Let $\sum_{b\in A\cap N} f_{l,A_i}(b) = \text{pop}_{l,A}$ for $A \in \mathcal{A}_\ell$. Then

$$\begin{aligned}\text{IFR}_{\text{total}} &= \frac{\text{number of deaths}}{\text{number of infections}} \\ &= \frac{\sum_{A\in\mathcal{A}_\ell} \text{IFR}_{l,A} \times \pi_{l,A} \times \text{pop}_{l,A}}{\sum_{B\in\mathcal{A}_\ell} \pi_{l,B} \times \text{pop}_{l,B}} \\ &= \sum_{A\in\mathcal{A}_\ell} \text{IFR}_{l,A} \times \left(\frac{\pi_{l,A}\times\text{pop}_{l,A}}{\sum_{B\in\mathcal{A}_\ell} \pi_{l,B}\times\text{pop}_{l,B}}\right) \quad (16)\end{aligned}$$

estimates the total IFR for location l. By calculating $\text{IFR}_{\text{total}}$ for each posterior sample, we can then obtain posterior mean and credible intervals for $\text{IFR}_{\text{total}}$.

High-income country benchmark

We compare the IFR estimate to a high income country benchmark based on results from Levin et. al. (4) which found a log-linear relationship between age and IFR. Define

$$\text{HICB}_a = \begin{cases} \int_a^{a+1} \frac{1}{100} 10^{-3.27+0.0524a} da & \text{if } a < 85 \\ \frac{1}{100} 10^{-3.27+0.0524(85)} & \text{if } a \geq 85 \end{cases}. \quad (17)$$

Then HICB_a represents the IFR predicted by the high-income countries line averaged over the interval [*a*,*a*+ 1) for ages less than 85 and assumes the high-income countries line flattens out and becomes uniform for ages 85 and older.

Then if we assume uniform prevalence for a location, the total IFR estimate over age bin *A* for high-income countries is

$$\text{HICB}_A = \frac{\sum_{a\in A} \text{EJE}_a \times f_l(a)}{\sum_{b\in A} f_l(b)}. \quad (18)$$

Subsetting to ages 18-65

To estimate the IFR between ages 18 and 65, we used the same strategy in picking age bins as we did when testing uniform prevalence. That is, we selected $\mathcal{B}_\ell$ to be the death age bins in $\mathcal{A}$ such that the lower age of the bin is greater than or equal to 18 and the upper age of the bin is less than 66. We then applied equation (16), replacing $\mathcal{A}_\ell$ with $\mathcal{B}_\ell$. We were not able to calculate the IFR between 18 and 65 for locations there were no age bins in $\mathcal{B}_\ell$.

Baseline population

In order to compare the impact of the age specific IFR while controlling for population age distribution, we calculated the Total IFR substituting $f_{\mathrm{l}}(a)$ for a baseline population age distribution, $\bar{f}(a)$ in (16). We calculated $\overline{f_A}(a)$ following (7). The baseline population was calculated as a median across locations for each age, then rescaled to sum to one:

$$\overline{f_\star}(a) = \text{median}\{\, f_{\mathrm{l}}(a) \mid \mathrm{l} \text{ is one of the observed locations with fatality data} \,\} \tag{19}$$

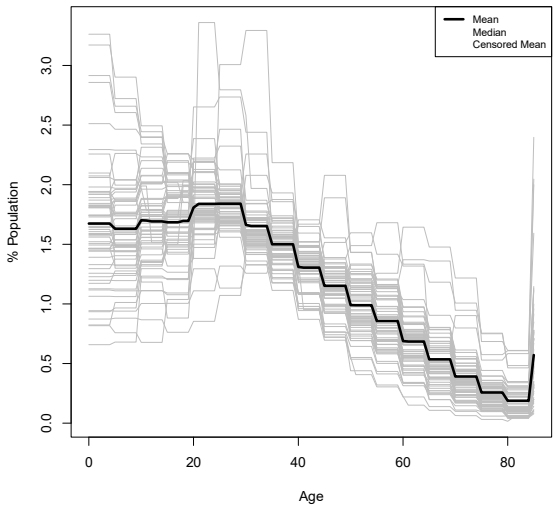
$$\bar{f}(a) = \frac{\overline{f_\star}(a)}{\sum_{a=0}^{85} \overline{f_\star}(a)}. \tag{20}$$

We also considered taking the mean across locations for each age and taking the mean across locations for each age after removing the top five and bottom five locations for that age (a censored mean). All three approaches gave similar values as shown in Figure A3.

Baseline age distribution

Various estimates of baseline age distribution (mean, median, and censored mean) plotted on age distribution for observed locations in grey:

Figure A3 – Estimates of Baseline Age Distribution



Country average

To get an average total IFR estimate for each country, we took a weighted average of the total IFR estimate of the locations within that country. We chose to weight by $1/\sqrt{n_{\mathrm{l}}}$ in order to give more weight to locations with more certain seroprevalence estimates. In locations with multiple age bins, we took the average across $n_{\mathrm{l},A}$ as n_{l}. We weighted by certainty in the seroprevalence estimates rather than the IFR estimates because larger IFR estimates tend to be due to small seroprevalence

estimates and consequently have more uncertainty (i.e., small differences in the denominator, seroprevalence, can result in large changes in the IFR estimate when seroprevalence is small). We did not want to bias the average by down weighting all of the larger IFR estimates.

We followed the same process to estimate the country average IFR between 18 and 65, with the added step of removing any locations in the country where $\mathcal{B}_\ell$ was empty.

c. National Serology Studies of High-Income Countries

As shown in Table A4, many high-income countries succeeded in limiting SARS-CoV-2 transmission during 2020 and early 2021 (92-94). Japan and South Korea (as well as several other East Asian countries) were particularly successful at limiting infection rates (95, 96). Some subnational high-income country locations reported higher seroprevalence, e.g., about 21% in New York City, USA (97), up to 25% in some Swiss cantons (59), and 42% at an Austrian ski area (18). However, the available serology data, based on representative samples of the general population, indicates that no location in any high-income country experienced non-vaccine-induced seroprevalence above 45% prior to March 2021.

Table A4 –National Seroprevalence Estimates for High-Income Countries

Timeframe	Location	Reported seroprevalence	Midpoint date
April 2020 to	Slovenia	0.9% (CI: 0.4-1.4%)	April 25
Sept. 2020	Spain	5.0% (CI: 4.7-5.4%)	May 4
	France	4.5% (CI: 3.9 - 5.0%)	May 17
	Italy	2.5% (CI: 2.3 - 2.6%)	June 19
	United Kingdom	6.0% (CI: 5.8 - 6.1%)	July 1
	Canada	1.9% (CI: 1.4 - 2.0%)	July 15
	South Korea	0.01%	July 19
	Germany	0.7%	July 21
	Denmark	2.0% (CI: 1.7-2.4%)	Sept. 19
	Netherlands	4.7% (CI: 4.0-5.5%)	Sept. 28
Oct. 2020 to	USA	11.9% (CI: 10.5 - 13.5%)	Oct. 30
January 2021	England	5.6% (CI: 5.4-5.7%)	Nov. 3
	Germany	1.1% (CI: 0.9 - 1.3%)	Nov. 6
	Slovenia	4.1% (CI: 3.0 - 5.2%)	Nov. 11
	Austria	4.7% (CI: 3.8 - 5.6%)	Nov. 13
	South Korea	0.1%	Nov. 21
	Spain	9.9% (CI: 9.4 - 10.4%)	Nov. 22
	France	6.2% (CI:5.9-6.6%)	Nov. 26
	Denmark	4.1% (CI: 3.1 - 4.9%)	Dec. 16
	Norway	0.9% (CI: 0.7 - 1.0%)	Jan. 5

Table A5 – Variants of Benchmark Metaregression for High-Income Countries

Description	# Observations	Intercept	Slope Coefficient
Benchmark	104	-3.27 *(0.073)*	0.0524 *(0.0013)*
Exclusion of Convenience Samples	68	-3.32 *(0.089)*	0.0532 *(0.0015)*
Adjustment for Seroreversion	104	-3.22 *(0.070)*	0.0516 *(0.0012)*
Adjustment for Death Undercounting	104	-3.18 *(0.075)*	0.0526 *(0.0013)*

d. Metaregression benchmark for high-income countries

To provide a benchmark for our analysis of IFR in developing countries, we consider the findings from a prior meta-analysis of age-specific IFRs for high-income countries (12). That study conducted a metaregression using 104 observations on age-specific IFRs from 28 locations (using samples collected between April and July 2020) and obtained the following results:

$$\underset{}{log_{10}(IFR)} = \underset{(0.07)}{-3.27} + \underset{(0.0013)}{0.0524 * age}$$

where the standard error for each estimated coefficient is given in parentheses.

To determine whether those results can serve as a suitable benchmark, we must consider several distinct methodological issues. First, the prior study used serology data from convenience samples as well as from representative samples of the general population, whereas our present analysis excludes convenience samples. Second, the prior study computed assay-adjusted seroprevalence using the baseline characteristics of each assay, whereas our present analysis incorporates adjustments for seroreversion over time. Third, the prior study used official reports on confirmed COVID-19 deaths, without incorporating any information about underreporting of COVID-19 fatalities, but subsequent analysis has shown that such underreporting has been substantial in some locations in high-income countries.

To assess the significance of these methodological issues, we have replicated the prior metaregression analysis along with three variants. In the first variant, the metaregression excludes 36 observations from convenience samples. In the second variant, we make seroreversion adjustments for two locations (Italy and Spain) that utilized the Abbott Architect assay, using the same approach as in our present analysis described above. In the third variant, we adjust fatalities using IHME estimates of COVID-19 death undercounts. Table A5 reports the results of this sensitivity analysis. Evidently, the results for each variant are nearly identical to those of the prior benchmark regression, with no statistically significant differences in the estimates of the intercept or slope coefficient. These results underscore the robustness of this metaregression for high-income countries and support its use as a benchmark for assessing IFR in developing countries.

3. Additional Results

a. Seroreversion Estimates

Table A6: Seroreversion Adjustments and Assay Sensitivity

Country	*Location*	*Assay*	*Baseline (%)*	*Adjusted (%)*	*Ratio*
Chile	3 urban areas	*Elecsys*	99.5	68.8	0.69
Ethiopia	Diredawa	*Abbott*	100	86.8	0.87
Hungary	National	*Abbott*	100	87.2	0.87
India	Delhi	*Kawach*	92.1	65.2	0.71
	Pimpri-Chinchwad	*Abbott*	100	77.9	0.78
	Paschim Medinipur	*Erbalisa*	98.3	36.7	0.37
	Chennai	*Abbott*	100	83.2	0.83
	Srinagar	*Abbott*	100	78.1	0.78
	Mumbai	*Abbott*	100	76.9	0.77
South Africa	Gauteng	*Luminex*	100	46.7	0.47

Note: This table shows the characteristics of the assay used in the serology study of each of the specified locations, including the assay sensitivity at baseline, the seroreversion-adjusted sensitivity, and the ratio of adjusted to baseline sensitivity. In denoting these assays, *Elecsys* refers to the Elecsys Anti-SARS-Cov-2 Roche assay, *Abbott* refers to the Abbott Architect IgG assay, Kavach refers to the Kawach IgG assay, *Erbalisa* refers to the ErbaLisa IgG assay, and *Luminex* refers to the Luminex protein trimer assay.

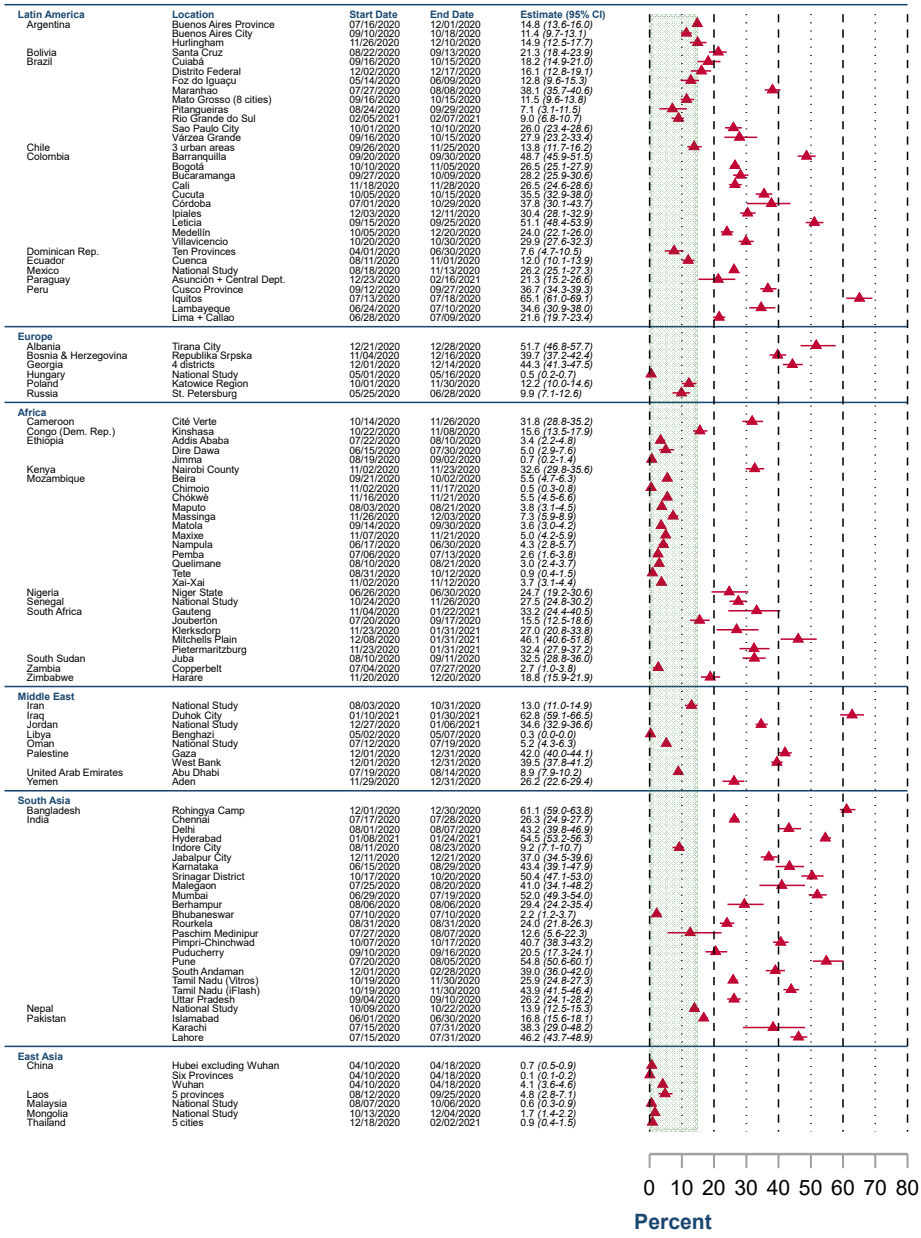
Table A7: Implications for Seroprevalence and IFR

		Seroprevalence (%)			
Country	*Location*	*Baseline Sensitivity*	*Seroreversion-Adjusted Sensitivity*	*Seroprevalence Ratio*	*IFR Ratio*
Chile	3 urban areas	10.1	13.8	1.4	0.73
Ethiopia	Diredawa	4.4	5.0	1.2	0.87
Hungary	National	0.4	0.5	1.4	0.79
India	Delhi	31.2	43.2	1.4	0.72
	Pimpri-Chinchwad	32.9	40.7	1.2	0.83
	Paschim Medinipur	6.8	12.6	1.9	0.55
	Chennai	22.1	26.3	1.2	0.84
	Srinagar	40.2	50.4	1.3	0.80
	Mumbai	40.1	52.0	1.3	0.78
South Africa	Gauteng	18.8	33.2	1.8	0.56

Note: For each location, this table reports the seroprevalence estimate obtained using the baseline sensitivity of the assay used in that serology study as well as the corresponding estimate obtained using the seroreversion-adjusted sensitivity for that assay. The penultimate column shows the ratio of seroreversion-adjusted to baseline-adjusted seroprevalence, while the final column shows the ratio for the corresponding estimates of population IFR for that location.

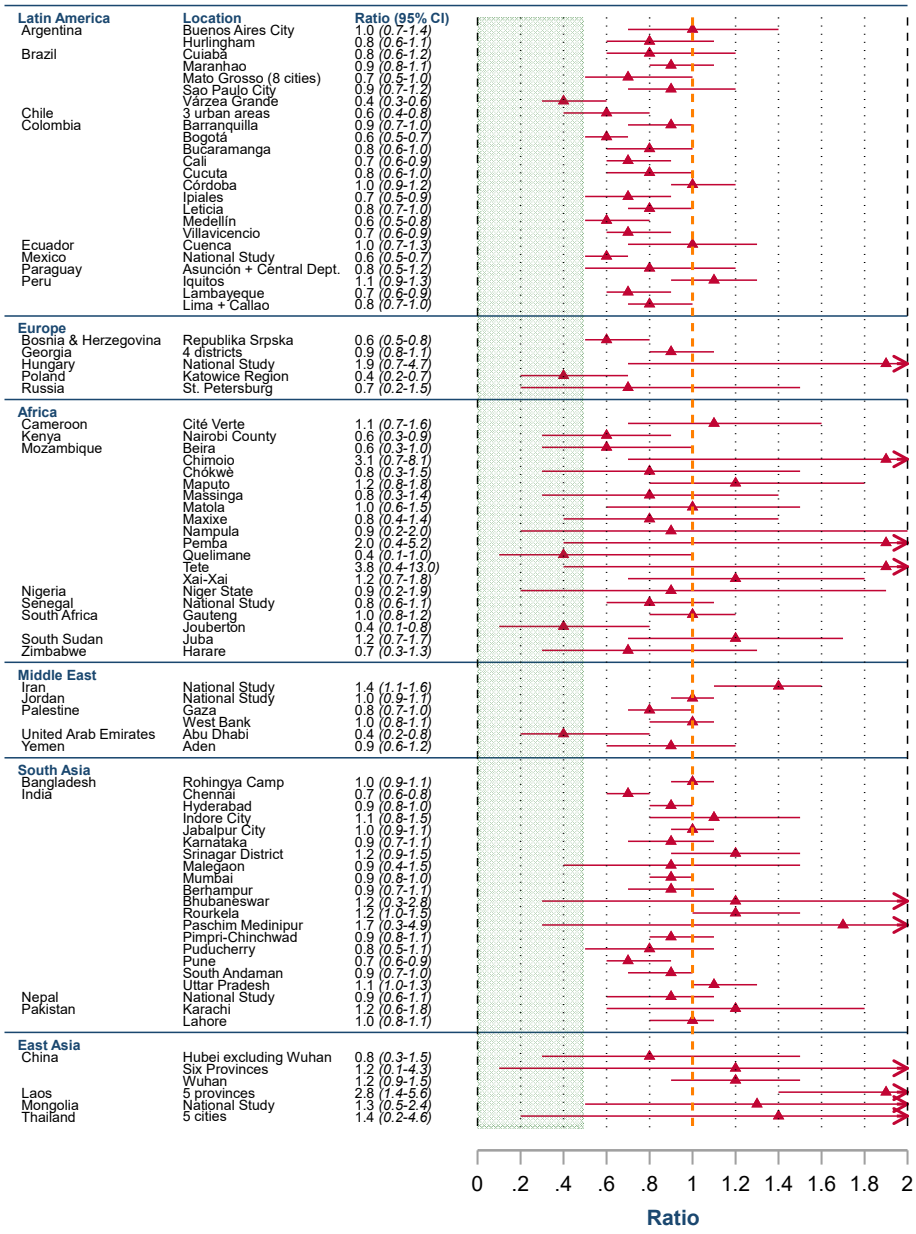
b. Seroprevalence Estimates

Figure A4 – Population-wide Seroprevalence



Notes: The green shading represents the range of national seroprevalence for high-income countries in Table A4. Links to these studies are in the Appendix folder of our GitHub repository.

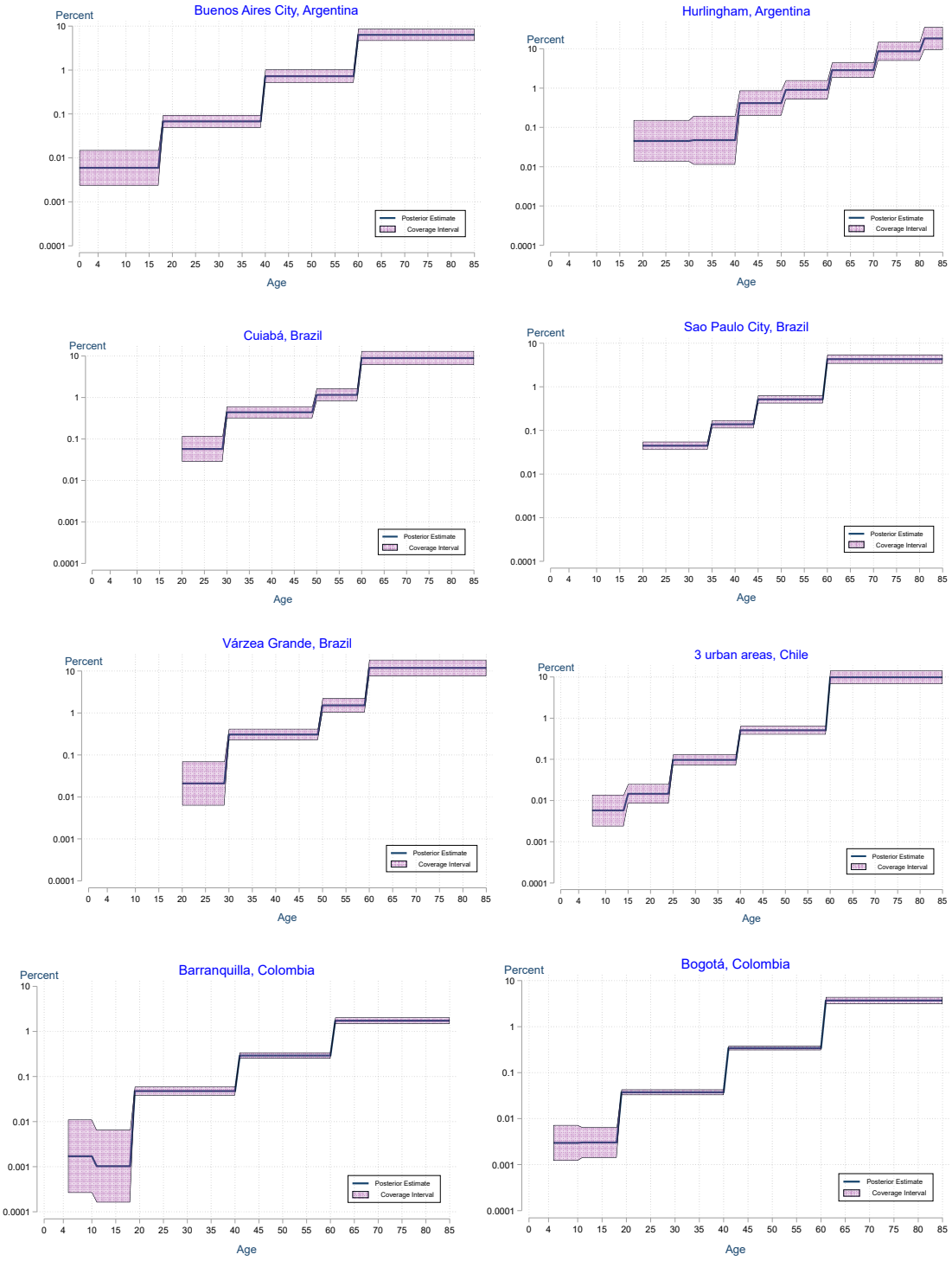
Figure A5 – Ratio of Seroprevalence for Older Adults (60+ years) Compared to Younger Adults (18-59 years)



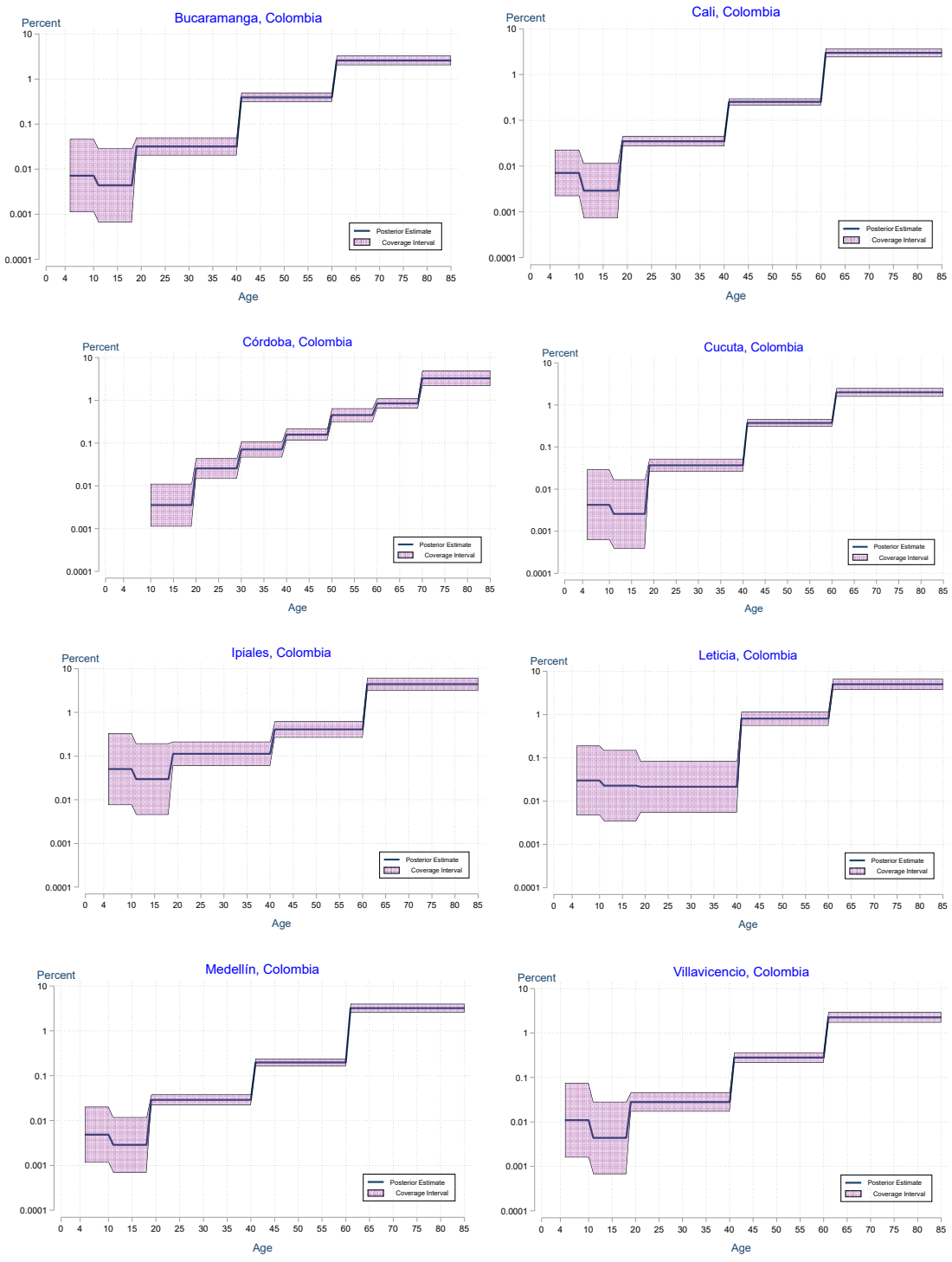
Note: The green shading represents the range of national seroprevalence for high-income countries from our prior work (12). Links to these studies are in the Appendix folder of our GitHub repository.

c. Age-specific IFR curves by location

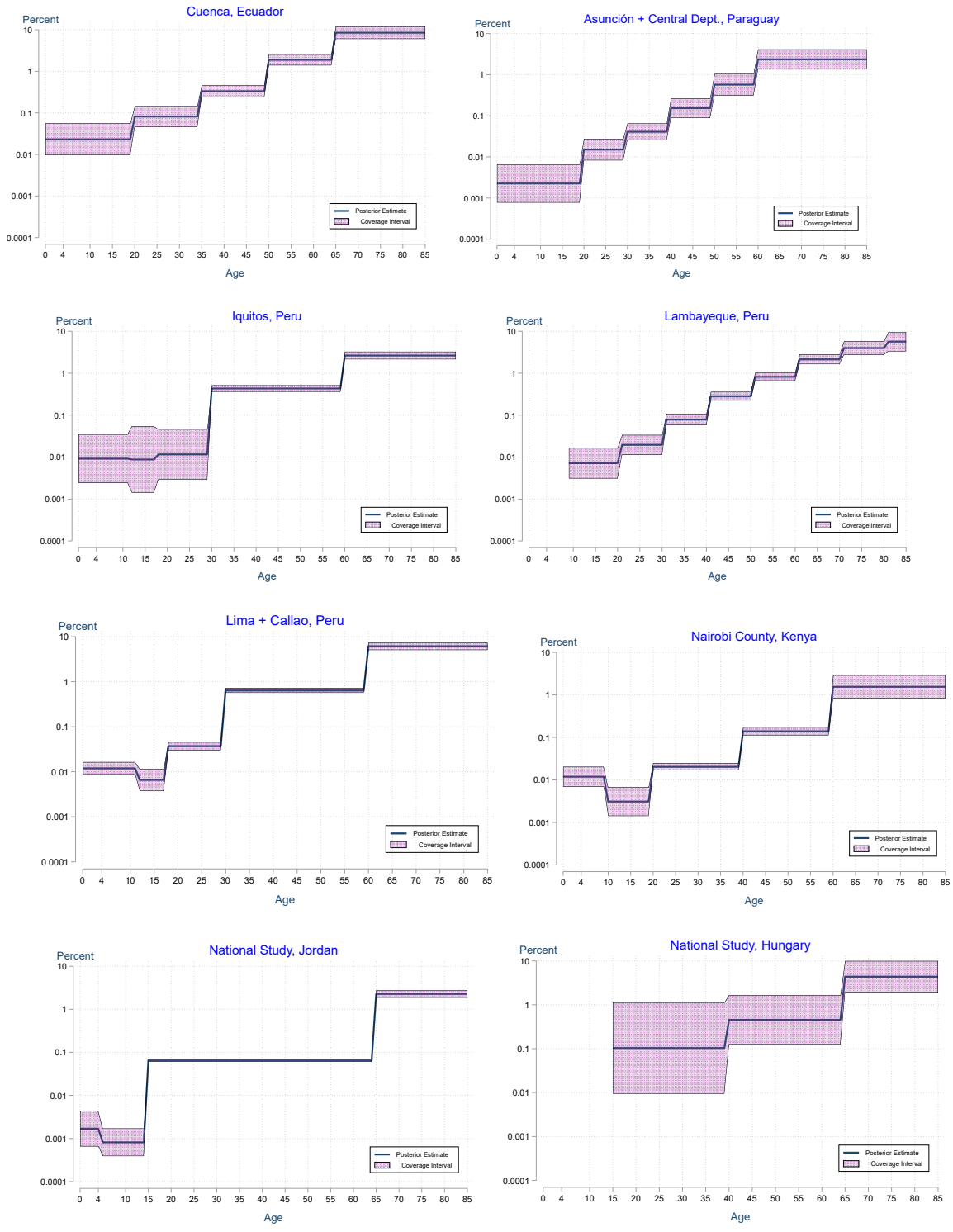
Figure A6 – Age-Specific IFR Curves By Location



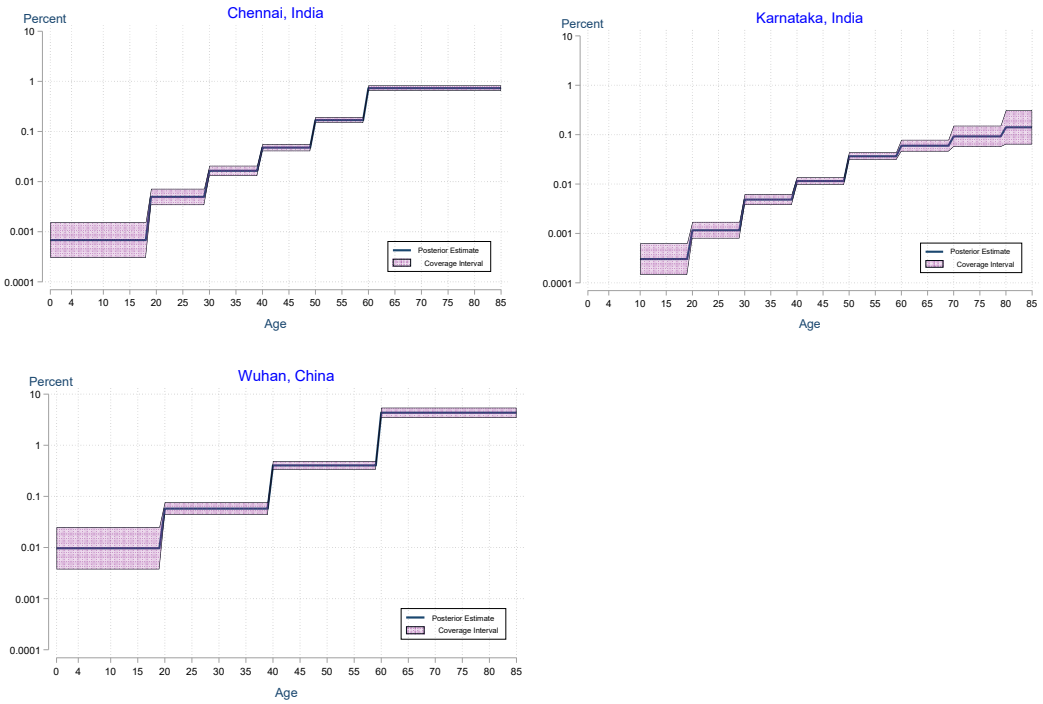
(figure continues on next page)



(figure continues on next page)



(figure continues on next page)



Note: For each location, this figure shows the posterior estimate and 95% credible interval of IFR for each of the age brackets reported in the serology study of that location; each estimate reflects the reported number of COVID-19 fatalities for that age bracket in that location and has not been adjusted for death undercounting.

d. Age-specific IFRs by Age Cohort

Figure A7 – IFR Estimates for Children

Country	Location	Ages	Estimate (95% CI)
India	Karnataka	10–19	0.000 (0.000–0.001)
India	Chennai	0–18	0.001 (0.000–0.001)
Jordan	National Study	5–14	0.001 (0.000–0.002)
Colombia	Barranquilla	11–18	0.001 (0.000–0.004)
Jordan	National Study	0–4	0.002 (0.001–0.004)
Colombia	Barranquilla	5–10	0.002 (0.000–0.006)
Paraguay	Asunción + Central Dept.	0–19	0.002 (0.001–0.005)
Colombia	Cucuta	11–18	0.003 (0.000–0.010)
Colombia	Medellin	11–18	0.003 (0.000–0.008)
Colombia	Cali	11–18	0.003 (0.000–0.008)
Colombia	Bogotá	5–10	0.003 (0.001–0.006)
Colombia	Bogotá	11–18	0.003 (0.001–0.006)
Kenya	Nairobi County	10–19	0.003 (0.001–0.006)
Brazil	Maranhao	10–19	0.003 (0.002–0.005)
Brazil	Maranhao	0–9	0.003 (0.002–0.006)
Colombia	Córdoba	10–19	0.004 (0.001–0.009)
Colombia	Cucuta	5–10	0.004 (0.000–0.016)
Colombia	Villavicencio	11–18	0.004 (0.000–0.016)
Colombia	Bucaramanga	11–18	0.004 (0.000–0.016)
Colombia	Medellin	5–10	0.005 (0.001–0.014)
Chile	3 urban areas	7–14	0.006 (0.002–0.012)
Argentina	Buenos Aires City	0–17	0.006 (0.002–0.013)
Peru	Lima + Callao	12–17	0.007 (0.004–0.011)
Colombia	Cali	5–10	0.007 (0.001–0.018)
Peru	Lambayeque	9–20	0.007 (0.003–0.014)
Colombia	Bucaramanga	5–10	0.007 (0.000–0.027)
Peru	Iquitos	12–17	0.009 (0.000–0.032)
Peru	Iquitos	0–11	0.009 (0.001–0.025)
China	Wuhan	0–19	0.010 (0.003–0.021)
Colombia	Villavicencio	5–10	0.011 (0.000–0.041)
Kenya	Nairobi County	0–9	0.012 (0.007–0.019)
Peru	Lima + Callao	0–11	0.012 (0.009–0.016)
Colombia	Leticia	11–18	0.023 (0.001–0.085)
Ecuador	Cuenca	0–19	0.023 (0.008–0.048)
Colombia	Ipiales	11–18	0.030 (0.001–0.110)
Colombia	Leticia	5–10	0.030 (0.001–0.110)
Colombia	Ipiales	5–10	0.050 (0.001–0.188)

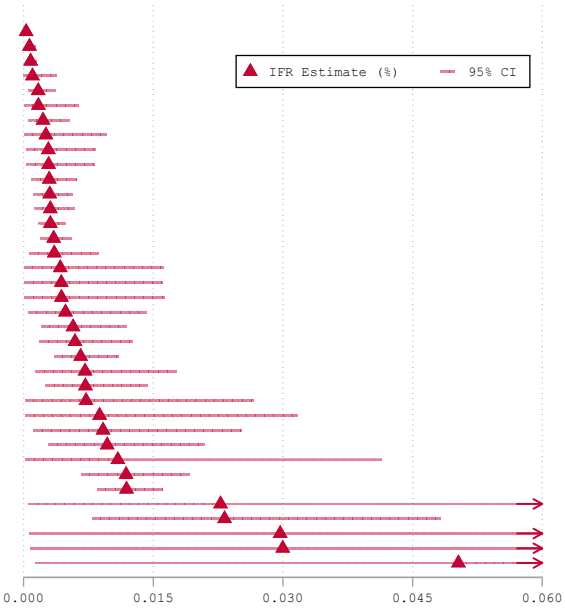


Figure A8 – IFR Estimates for Young Adults

Country	Location	Ages	Estimate (95% CI)
India	Karnataka	20–29	0.001 (0.001–0.002)
India	Karnataka	30–39	0.005 (0.004–0.006)
India	Chennai	19–29	0.005 (0.003–0.007)
Brazil	Maranhao	20–29	0.006 (0.004–0.009)
Peru	Iquitos	18–29	0.012 (0.001–0.033)
Chile	3 urban areas	15–24	0.015 (0.008–0.024)
Paraguay	Asunción + Central Dept.	20–29	0.015 (0.008–0.026)
India	Chennai	30–39	0.016 (0.013–0.020)
Peru	Lambayeque	21–30	0.020 (0.011–0.032)
Kenya	Nairobi County	20–39	0.020 (0.017–0.024)
Brazil	Várzea Grande	20–29	0.021 (0.004–0.053)
Colombia	Leticia	19–40	0.021 (0.003–0.060)
Colombia	Córdoba	20–29	0.026 (0.014–0.042)
Colombia	Villavicencio	19–40	0.028 (0.016–0.043)
Colombia	Medellin	19–40	0.029 (0.022–0.037)
Colombia	Bucaramanga	19–40	0.032 (0.019–0.048)
Brazil	Maranhao	30–39	0.033 (0.026–0.041)
Colombia	Cali	19–40	0.035 (0.027–0.044)
Colombia	Cucuta	19–40	0.037 (0.026–0.050)
Peru	Lima + Callao	18–29	0.037 (0.030–0.045)
Colombia	Bogotá	19–40	0.037 (0.033–0.043)
Paraguay	Asunción + Central Dept.	30–39	0.041 (0.026–0.063)
Brazil	Sao Paulo City	20–34	0.045 (0.036–0.055)
Argentina	Hurlingham	18–30	0.045 (0.009–0.116)
Argentina	Hurlingham	31–40	0.047 (0.005–0.137)
Colombia	Barranquilla	19–40	0.048 (0.038–0.059)
China	Wuhan	20–39	0.058 (0.043–0.075)
Brazil	Cuiabá	20–29	0.058 (0.026–0.106)
Argentina	Buenos Aires City	18–39	0.068 (0.049–0.093)
Colombia	Córdoba	30–39	0.071 (0.047–0.106)
Peru	Lambayeque	31–40	0.079 (0.057–0.105)
Ecuador	Cuenca	20–34	0.082 (0.044–0.136)
Chile	3 urban areas	25–39	0.097 (0.072–0.129)
Hungary	National Study	15–39	0.103 (0.014–0.500)
Colombia	Ipiales	19–40	0.113 (0.054–0.195)

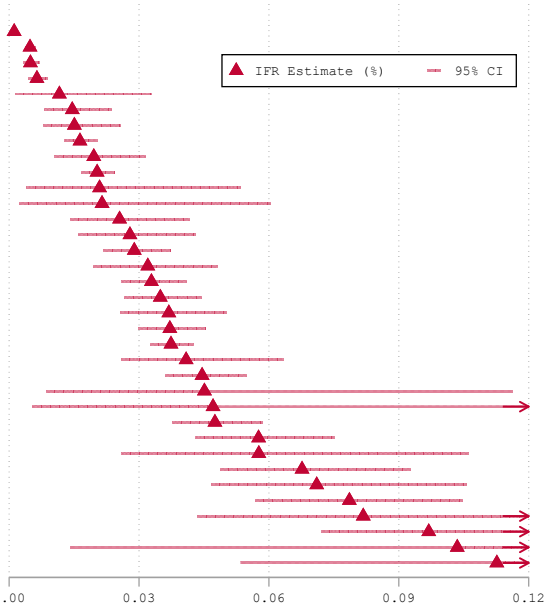


Figure A9 – IFR Estimates for Middle-aged Adults

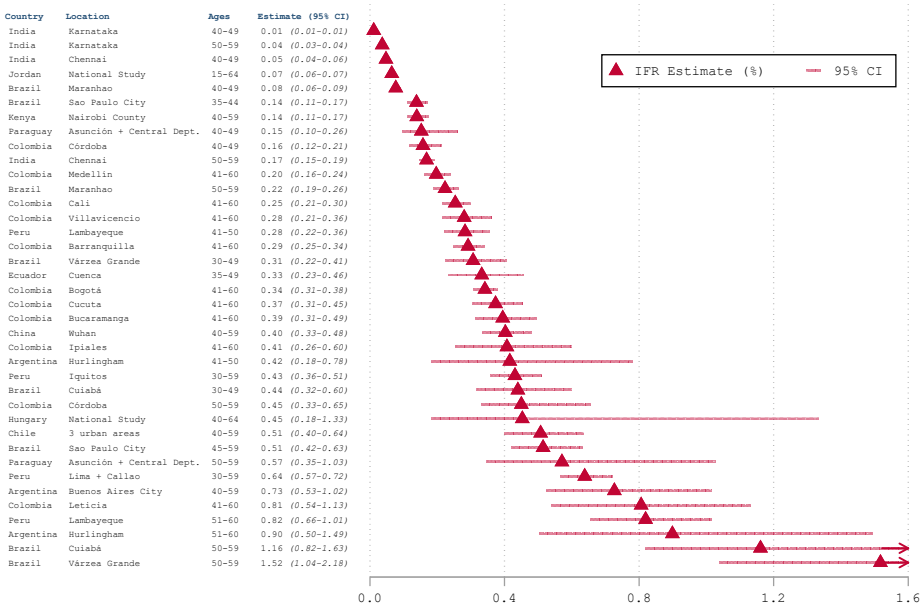
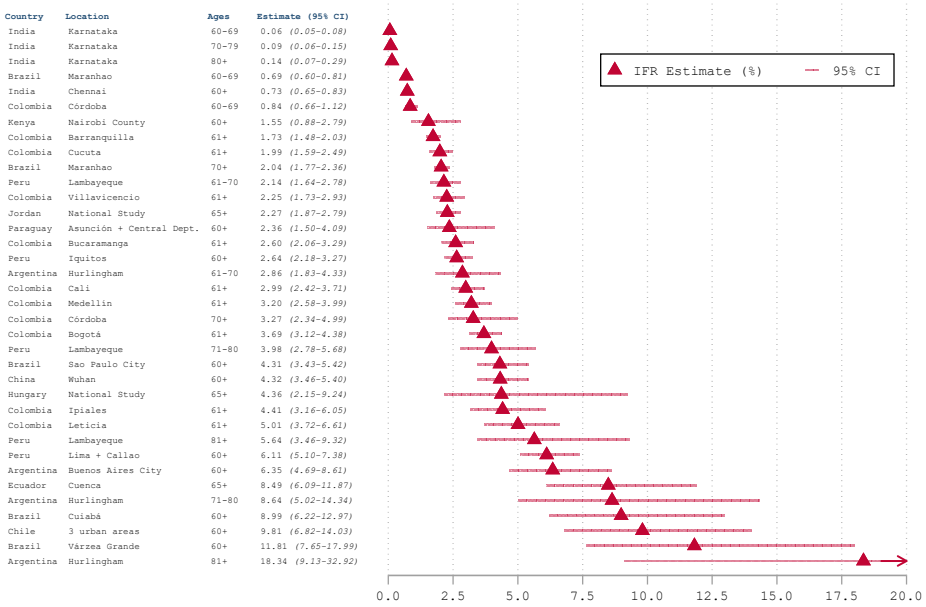


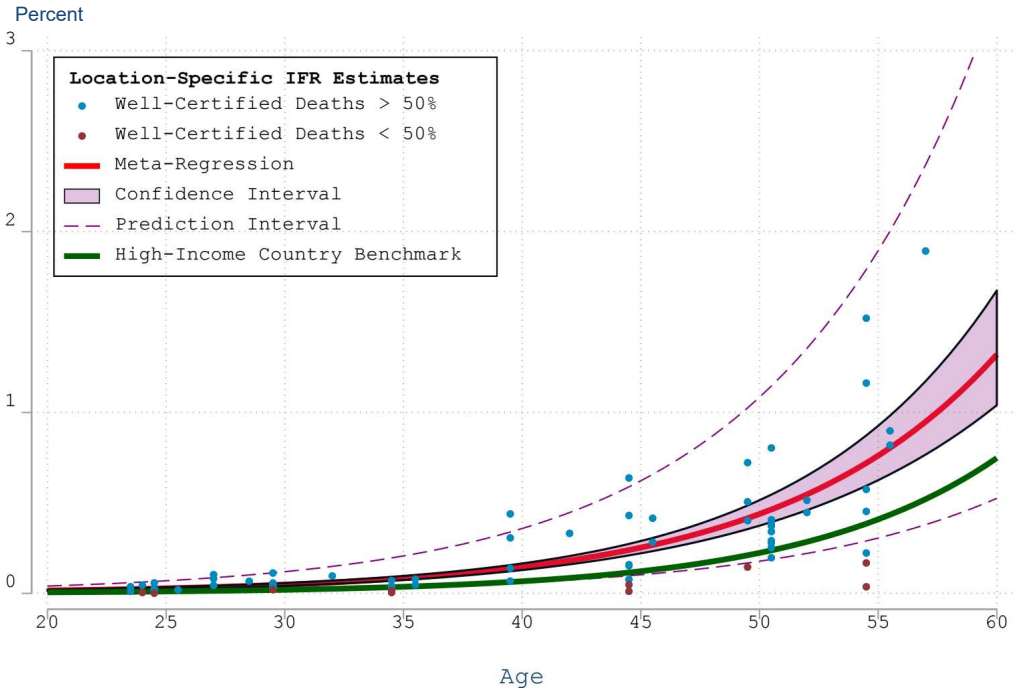
Figure A10 – IFR Estimates for Older Adults



Note: Links to these studies are in the Appendix folder of our GitHub repository.

e. Metaregression Results

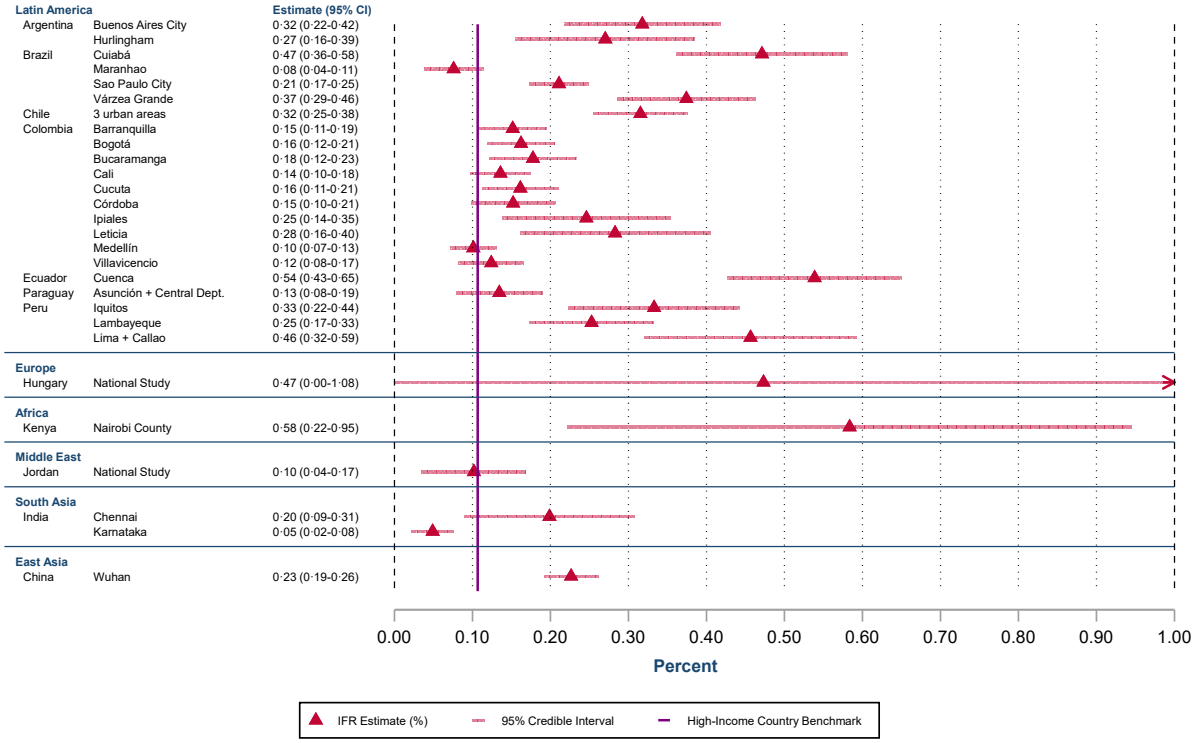
Figure A11 – Age-specific IFR Metaregression in Levels



Note: Links to the studies at each location and categorization by percentage of deaths well-certified are provided in the appendix folder of our GitHub repository.

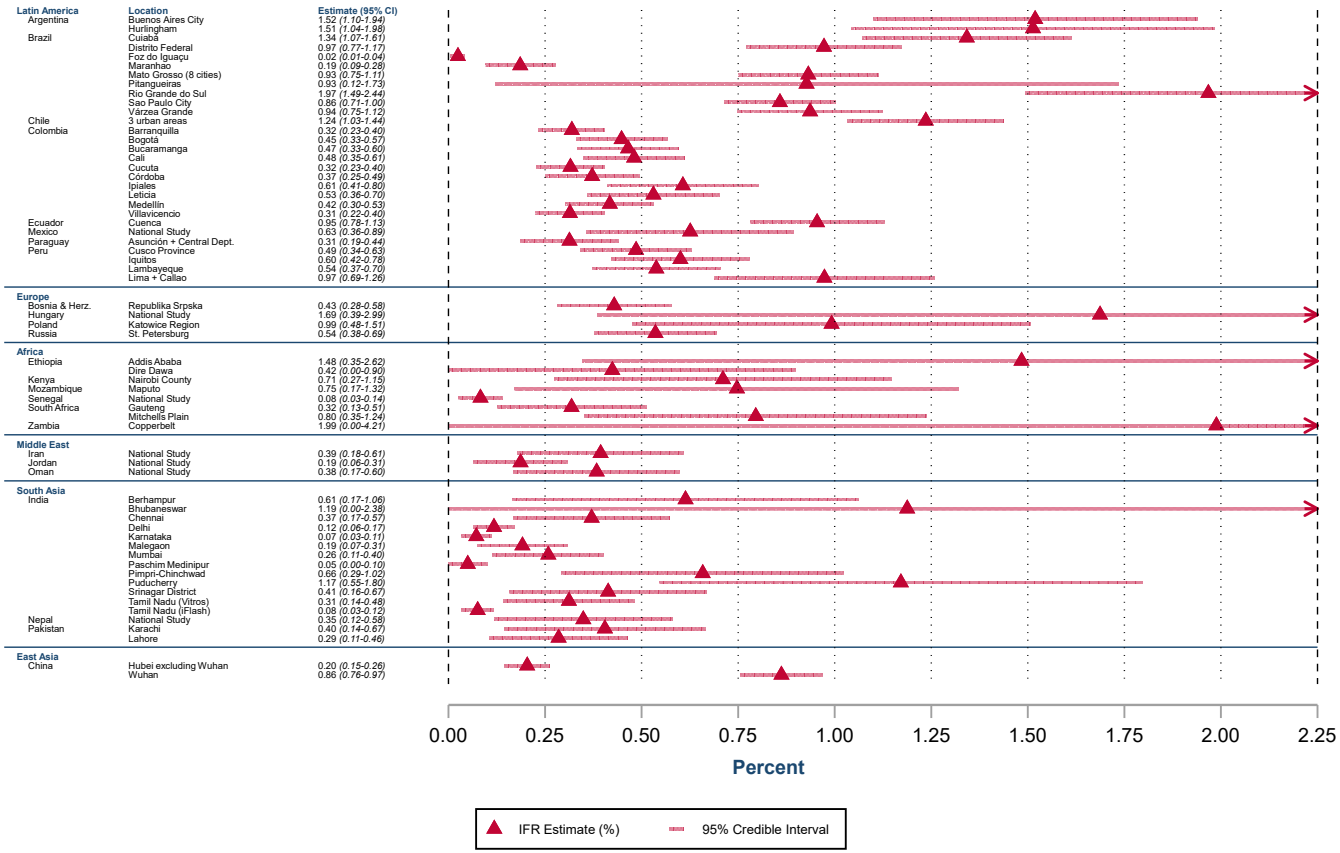
f. Population IFR

Figure A12 – Population IFR (ages 18 to 64)



Note: This figure shows IFR estimates for the population aged 18-64 based on the age structure and age-specific seroprevalence in each location.

Figure A13 – Population IFR (all ages)



g. IFR Estimates from Other Sources

Table A8 compares reported IFRs from studies included in our literature search to our meta-analysis IFR estimates for the corresponding locations, while Table A9 lists reported IFRs from studies identified in our literature search that were excluded from our IFR analysis. It should be noted that IFRs based on reported COVID-19 deaths may be biased due to death undercounting, especially for countries that have a low percentage of well-certified deaths (32).

Table A8 – Comparison of Meta-analysis IFRs to Reported IFRs

Location	IFR stated in the study	Meta-analysis IFR
Brazil: Maranhão*	0.14% (CI: 0.13 - 0.16%) [reported deaths] 0.28% (CI: 0.25 - 0.32%) [excess deaths]	0.13% (CI: 0.12 - 0.14%)
Chile: Coquimbo-La Serena, Greater Santiago, and Talca	1.67% (CI: 1.64 - 1.70%)	1.23% (CI: 1.04 - 1.45%)
Colombia: Córdoba (8 cities)*	0.24% (CI: 0.23 - 0.25%)	0.34% (CI: 0.29 - 0.43%)
Mexico: National	0.47% (CI: 0.44-0.50%)	0.46% (CI: 0.44-0.48%)
Peru: Lambayeque	0.5%	0.49% (CI: 0.43 - 0.56%)
Poland: Katowice region	0.62% (CI: 0.53 - 0.74%)	0.62% (CI: 0.50 - 0.76%)
Russia: St. Petersburg*	0.83% (CI: 0.62 - 1.00%) [excess deaths]	0.54% (CI: 0.41 - 0.73%)
Ethiopia: Addis Ababa #2	≥0.09%	0.20% (CI: 0.09 - 0.63%)
Kenya: Nairobi County*	0.04%	0.06% (CI: 0.05 - 0.07%)
South Africa: Gauteng province	0.28% (CI: 0.27 - 0.30%) [reported deaths] 0.67% (CI: 0.64 - 0.71%) [excess deaths]	0.12% (CI: 0.10 - 0.17%)
South Africa: Mitchells Plain	0.3% (CI: 0.3 - 0.4%) [in-hospital deaths] 0.5% (CI: 0.4 - 0.6%) [excess deaths]	0.31% (CI: 0.27 - 0.37%)
India: national	0.08% (CI: 0.07 - 0.09%) to 0.11% (CI: 0.10 - 0.12%)	0.06% (CI: 0.05 - 0.06%)
India: Chennai*	0.17% (CI: 0.14 - 0.22%)	0.08% (CI: 0.07 - 0.08%)
India: Delhi	0.079% (CI: 0.076 - 0.081%)	0.055% (CI: 0.05 - 0.06%)
India: Kashmir	0.03% (CI: 0.03 - 0.04%)	0.03% (CI: 0.026 - 0.030%)
India: Madurai district*	0.04% (CI: 0.04 - 0.05%)	0.03% (CI: 0.02 - 0.03%)
India: Mumbai (3 wards)*	0.12%	0.08% (CI: 0.08 - 0.09%)
India: Pimpri-Chinchwad	0.17%	0.22% (CI: 0.20 - 0.23%)
India: Puducherry*	0.08%	0.18% (CI: 0.15 - 0.22%)

Note: IFRs are based on reported deaths, not excess deaths, unless otherwise noted. Studies with an asterisk also reported age-specific IFRs. Age-specific IFRs were also reported for Karnataka, India (98).

Table A9 – Reported IFRs for Sero-Only Studies

Location	IFR stated in the study	Type of death	Reason IFR estimate was not generated
Brazil: Rio de Janeiro (multiple regions of the city)	Rio das Pedras: 0.2%; Maré: 0.3%; Rocinha: 0.3%; Cidade de Deus: 0.4%; Realengo: 1.2%; Campo Grande: 1.8%	reported	no death data 4 weeks post-midpoint, insufficient information on assay
South Africa: Jouberton	wave 1: 0.12% (CI: 0.09 – 0.20%) wave 1: 0.16% (CI: 0.13 – 0.23%) wave 2: 0.50% (CI: 0.29 – 1.17%) wave 2: 0.36% (CI: 0.24 – 0.72%)	excess in-hospital excess in-hospital	no death data 4 weeks post-midpoint
South Africa: Klerksdorp, Pietermaritzburg	Klerksdorp: 0.3% (CI: 0.2 – 0.3%) 0.3% (CI: 0.3 – 0.3%) Pietermaritzburg: 0.3% (CI: 0.3 – 0.3%) 0.6% (CI: 0.5 – 0.6%)	in-hospital excess in-hospital excess	no death data 4 weeks post-midpoint
Sudan: Omdurman*	0.64% (CI: 0.62 – 0.75%)	excess	no death data 4 weeks post-midpoint, sampling after February 2021
Iran: Guilan province	0.12%	reported	no death data 4 weeks post-midpoint
Iran: Mazandaran province	0.33%	reported	no death data 4 weeks post-midpoint, no stated start-week and end-week
Palestine	0.11%	reported	no stated start-week and end-week
India: Indore (city, not district)	0.17%	reported	no death data 4 weeks post-midpoint
India: Pune, 5 subwards*	0.21%	reported	no death data 4 weeks post-midpoint
India: Tamil Nadu*	0.05%	reported	Tamil Nadu split into 2 regions based on assay

Note: IFRs are based on reported deaths, not excess deaths, unless otherwise noted. Studies with an asterisk also reported age-specific IFRs.

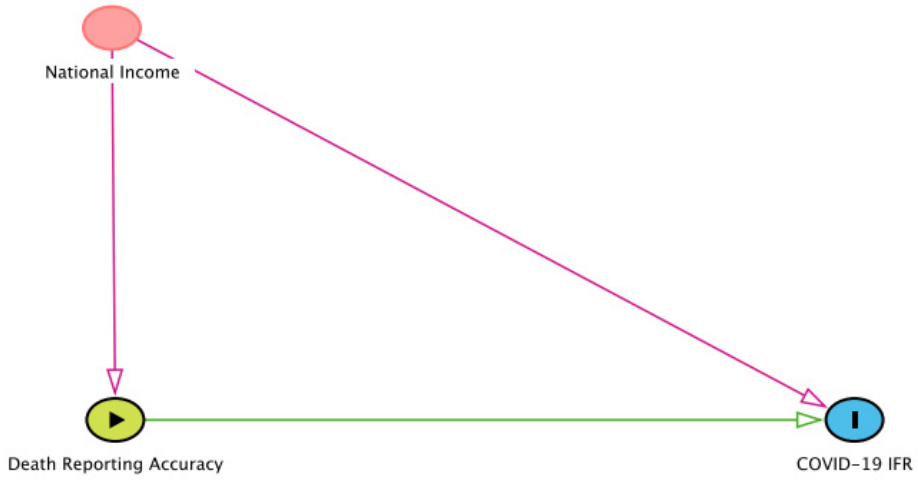
h. Covariates

Table A10 – Correlations Between Covariates, IFR, and Well-Certified Deaths

Covariate	Population IFR	Well-Certified Deaths
Human Development Index	0.63 (0.27-0.83)	0.90 (0.78-0.96)
Log of GDP per capita	0.60 (0.23-0.82)	0.88 (0.73-0.95)
Log(Healthcare Spending)	0.60 (0.22-0.82)	0.93 (0.82-0.97)
Log of GNI per capita	0.60 (0.22-0.82)	0.88 (0.72-0.95)
Hospital Beds Per Capita	0.57 (0.18-0.80)	0.48 (0.06-0.76)
Universal Health Coverage Index	0.55 (0.16-0.79)	0.95 (0.88-0.98)
Skilled Healthcare Workers Per Capita	0.49 (0.08-0.76)	0.69 (0.36-0.86)
Global Health Security Index	0.47 (0.05-0.75)	0.59 (0.21-0.81)
Life Expectancy at Birth	0.43 (0.0-0.73)	0.71 (0.40-0.88)
Healthy Life Expectancy at Age 60	0.40 (-0.04-0.71)	0.83 (0.61-0.93)

This table demonstrates the relationship between various covariates (32), IFR, and the measure of well-certified deaths. This shows that well-certified death is likely to be a primary explanatory variable, which is confounded by relationships with GDP and other national measures when these are used instead. An example of this is shown in the Directed Acyclic Graph below, made using the Daggity online tool: http://www.dagitty.net/dags.html#

Figure A14 – Directed Acyclic Graph of Covariate Relationships



i. Out-of-Sample Analysis

Our analysis excludes seroprevalence estimates that geographically overlap with an already included location, as discussed in supplementary appendix section 1.b. The body of the paper also excludes IFR estimates that overlap with an included IFR, though we still calculated population-wide IFRs for these out-of-sample locations. The table below lists these IFR estimates:

Table A11 – IFRs for Out-of-Sample Locations with Geographical Overlap

Location	Excluded from body of paper	Meta-analysis IFR
Brazil: São Paulo*	No	0.84% (CI: 0.76 - 0.93%)
Brazil: São Paulo #2*	Yes	0.77% (CI: 0.66 - 0.88%)
Ethiopia: Addis Ababa*	No	0.11% (CI: 0.07 - 0.16%)
Ethiopia: Addis Ababa #2	Yes	0.20% (CI: 0.09 - 0.63%)
Ethiopia: Addis Ababa #3*	Yes	0.002% (CI: 0.001 - 0.005%)
India: national*	Yes	0.06% (CI: 0.05 - 0.06%)
India: Kashmir (Srinagar district)*	No	0.06% (CI: 0.06 - 0.07%)
India: Kashmir*	Yes	0.03% (CI: 0.026 - 0.030%)
India: Tamil Nadu (Vitros districts)	No	0.07% (CI: 0.06 - 0.07%)
India: Madurai district (in Tamil Nadu)*	Yes	0.03% (CI: 0.02 - 0.03%)
China: Wuhan*	No	0.86% (CI: 0.76 - 0.97%)
China: Wuhan #2*	Yes	0.71% (CI: 0.59 - 1.06%)

*age-specific seroprevalence also reported in the paper

IFRs from studies of the same location may differ by sampling time due to factors such as improved treatment or new SARS-CoV-2 variants. Despite this, population-wide IFRs were relatively similar for studies that sampled the same location, as illustrated in the table above. These consilient results increase confidence that methodological differences between studies likely do not strongly bias our IFR estimates, in contrast to the order of magnitude difference in IFR between locations stratified by percentage of well-certified deaths, as shown in the body of the paper. Addis Ababa #3 remains the only outlier, possibly due to lower test specificity resulting from cross-reactivity (see supplementary appendix section 2.a), low sample size in comparison to the other two Addis Ababa studies, or sampling in late April 2020 when under-estimation of COVID-19 deaths may have been greater than the July/August 2020 time period during which the other two studies sampled.

Appendix References

1. Bobrovitz N, Arora RK, Yan T, et al. Lessons from a rapid systematic review of early SARS-CoV-2 serosurveys. medRxiv. 2020:2020.05.10.20097451. doi:10.1101/2020.05.10.20097451
2. Chen X, Chen Z, Azman AS, et al. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. The Lancet Global Health. doi:10.1016/S2214-109X(21)00026-7
3. Byambasuren O, Dobler CC, Bell K, et al. Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: Systematic review. PLOS ONE. 2021;16(4):e0248946. doi:10.1371/journal.pone.0248946
4. Bergeri I, Whelan M, Ware H, et al. Global epidemiology of SARS-CoV-2 infection: a systematic review and meta-analysis of standardized population-based seroprevalence studies, Jan 2020-Oct 2021. medRxiv. 2021:2021.12.14.21267791. doi:10.1101/2021.12.14.21267791
5. Laurette M, Marion V, Eduard G, Alex W. Research Square. 2021. doi:10.21203/rs.3.rs-707813/v2
6. Golding J, Northstone K, Miller LL, Davey Smith G, Pembrey M. Differences between blood donors and a population sample: implications for case-control studies. International journal of epidemiology. 2013;42(4):1145-56. doi:10.1093/ije/dyt095
7. Pham D, Nguyen D, Nguyen T, et al. Seroprevalence of HTLV-1/2 Among Voluntary Blood Donors in Vietnam. AIDS Research and Human Retroviruses. 2019;35(4):376-81. doi:10.1089/aid.2018.0240
8. He D, Artzy-Randrup Y, Musa SS, Gräf T, Naveca F, Stone L. The unexpected dynamics of COVID-19 in Manaus, Brazil: Was herd immunity achieved? medRxiv. 2021:2021.02.18.21251809. doi:10.1101/2021.02.18.21251809
9. Boyce RM, Shook-Sa BE, Aiello AE. A Tale of 2 Studies: Study Design and Our Understanding of Severe Acute Respiratory Syndrome Coronavirus 2 Seroprevalence. Clinical Infectious Diseases. 2020. doi:10.1093/cid/ciaa1868
10. Campbell H, de Valpine P, Maxwell L, et al. Bayesian adjustment for preferential testing in estimating infection fatality rates, as motivated by the COVID-19 pandemic. The Annals of Applied Statistics. 2022;16(1):436-59, 24.
11. Shook-Sa BE, Boyce RM, Aiello AE. Estimation Without Representation: Early Severe Acute Respiratory Syndrome Coronavirus 2 Seroprevalence Studies and the Path Forward. The Journal of Infectious Diseases. 2020;222(7):1086-9. doi:10.1093/infdis/jiaa429
12. Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. European Journal of Epidemiology. 2020;35(12):1123-38. doi:10.1007/s10654-020-00698-1
13. Community Assessment for Public Health Emergency Response Toolkit: CDC2019.
14. Bobrovitz N, Noël K, Li Z, et al. SeroTracker-RoB: an approach to automating reproducible risk of bias assessment of seroprevalence studies. medRxiv. 2022:2021.11.17.21266471. doi:10.1101/2021.11.17.21266471
15. World Health Organization. Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection. Geneva: World Health Organization2020 17 March 2020 Contract No.: WHO/2019-nCoV/Seroepidemiology/2020.1.
16. World Economic and Financial Surveys World Economic Outlook Database—WEO Groups and Aggregates Information2021.
17. Gajda M, Kowalska M, Zejda JE. Impact of Two Different Recruitment Procedures (Random vs. Volunteer Selection) on the Results of Seroepidemiological Study (SARS-CoV-2). International Journal of Environmental Research and Public Health. 2021;18(18):9928.

18. Knabl L, Mitra T, Kimpel J, et al. High SARS-CoV-2 seroprevalence in children and adults in the Austrian ski resort of Ischgl. Communications Medicine. 2021;1(1):4. doi:10.1038/s43856-021-00007-1
19. COVID-19 Cases. Ministerio de Salud Argentina. http://datos.salud.gob.ar/dataset/covid-19-casos-registrados-en-la-republica-argentina/archivo/fd657d02-a33a-498b-a91b-2ef1a68b8d16.
20. COVID-19 in Colombia. Instituto Nacional De Salud. 2021. https://www.ins.gov.co/Noticias/paginas/coronavirus.aspx.
21. COVID-19 PANEL - ESPIRITO SANTO STATE. 2021. https://coronavirus.es.gov.br/painel-covid-19-es.
22. Coronavirus - COVID-19. Parana Governo Do Estado. 2021. https://www.saude.pr.gov.br/Pagina/Coronavirus-COVID-19.
23. COVID19 Registro Fallecidos. 2021. https://public.tableau.com/app/profile/mspbs/viz/COVID19PY-Registros/FALLECIDOS.
24. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. eLife. 2021;10:e69336. doi:10.7554/eLife.69336
25. Ramachandran S, Malani A. All-cause mortality during SARS-CoV-2 Pandemic in India: Nationally-representative estimates independent of official death registry. medRxiv. 2021:2021.07.20.21260577. doi:10.1101/2021.07.20.21260577
26. Watson OJ, Alhaffar M, Mehchy Z, et al. Leveraging community mortality indicators to infer COVID-19 mortality and transmission dynamics in Damascus, Syria. Nature Communications. 2021;12(1):2394. doi:10.1038/s41467-021-22474-9
27. Dyer O. Covid-19: Russia admits to understating deaths by more than two thirds. BMJ. 2020;371:m4975. doi:10.1136/bmj.m4975
28. Dyer O. Covid-19: Mexico acknowledges 50 000 more deaths than official figures show. BMJ. 2020;371:m4182. doi:10.1136/bmj.m4182
29. Technical criteria to update the death toll from COVID-19 in Peru: Peruvian State2021.
30. Nepal Dashboard: World Health Organization2021.
31. COVID-19 MÉXICO Comunicado Técnico Diario.
32. Fullman N, Barber RM, Abajobir AA, et al. Measuring progress and projecting attainment on the basis of past trends of the health-related Sustainable Development Goals in 188 countries: an analysis from the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1423-59. doi:10.1016/S0140-6736(17)32336-X
33. Silveira MF, Mesenburg MA, Dellagostin OA, et al. Time-dependent decay of detectable antibodies against SARS-CoV-2: A comparison of ELISA with two batches of a lateral-flow test. Braz J Infect Dis. 2021;25(4):101601-. doi:10.1016/j.bjid.2021.101601
34. Conklin SE, Martin K, Manabe YC, et al. Evaluation of Serological SARS-CoV-2 Lateral Flow Assays for Rapid Point-of-Care Testing. Journal of clinical microbiology. 2021;59(2):e02020-20. doi:10.1128/JCM.02020-20
35. Hartwig FP, Vidaletti LP, Barros AJD, et al. Combining serological assays and official statistics to describe the trajectory of the COVID-19 pandemic: results from the EPICOVID19-RS study in Rio Grande do Sul (Southern Brazil). medRxiv. 2021:2021.05.21.21257634. doi:10.1101/2021.05.21.21257634
36. Lipsitch M, Grad YH, Sette A, Crotty S. Cross-reactive memory T cells and herd immunity to SARS-CoV-2. Nature Reviews Immunology. 2020;20(11):709-13. doi:10.1038/s41577-020-00460-4
37. Dorigatti I, Lavezzo E, Manuto L, et al. SARS-CoV-2 antibody dynamics and transmission from community-wide serological testing in the Italian municipality of Vo'. Nature Communications. 2021;12(1):4383. doi:10.1038/s41467-021-24622-7
38. Kshatri JS, Bhattacharya D, Praharaj I, et al. Seroprevalence of SARS-CoV-2 in Bhubaneswar, India: findings from three rounds of community surveys. Epidemiology and infection. 2021;149:e139. doi:10.1017/s0950268821000972

39. Laxmaiah A, Rao NM, Arlappa N, et al. SARS-CoV-2 seroprevalence in the city of Hyderabad, India in early 2021. medRxiv. 2021:2021.07.18.21260555. doi:10.1101/2021.07.18.21260555
40. Wagner R, Peterhoff D, Beileke S, et al. Estimates and Determinants of SARS-Cov-2 Seroprevalence and Infection Fatality Ratio Using Latent Class Analysis: The Population-Based Tirschenreuth Study in the Hardest-Hit German County in Spring 2020. Viruses. 2021;13(6):1118.
41. Prevention USCfDC. Science Brief: SARS-CoV-2 Infection-induced and Vaccine-induced Immunity. 2021.
42. Sharma N, Sharma P, Basu S, et al. Second wave of the Covid-19 pandemic in Delhi, India: high seroprevalence not a deterrent? medRxiv. 2021:2021.09.09.21263331. doi:10.1101/2021.09.09.21263331
43. Álvarez-Antonio C, Meza-Sánchez G, Calampa C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in Iquitos, Peru in July and August, 2020: a population-based study. The Lancet Global Health. 2021;9(7):e925-e31. doi:10.1016/S2214-109X(21)00173-X
44. Fox SJ, Potu P, Lachmann M, Srinivasan R, Meyers LA. The COVID-19 herd immunity threshold is not low: A re-analysis of European data from spring of 2020. medRxiv. 2020:2020.12.01.20242289. doi:10.1101/2020.12.01.20242289
45. Bellizzi S, Alsawalha L, Sheikh Ali S, et al. A three-phase population based sero-epidemiological study: Assessing the trend in prevalence of SARS-CoV-2 during COVID-19 pandemic in Jordan. One Health. 2021;13:100292. doi:https://doi.org/10.1016/j.onehlt.2021.100292
46. Sharma N, Sharma P, Basu S, et al. Second Wave of the COVID-19 Pandemic in Delhi, India: High Seroprevalence Not a Deterrent? Cureus. 2021;13(10): e19000. doi:doi:10.7759/cureus.19000
47. Poustchi H, Darvishian M, Mohammadi Z, et al. SARS-CoV-2 antibody seroprevalence in the general population and high-risk occupational groups across 18 cities in Iran: a population-based cross-sectional study. The Lancet Infectious Diseases. 2021;21(4):473-81. doi:10.1016/S1473-3099(20)30858-6
48. Emmerich P, Murawski C, Ehmen C, et al. Limited specificity of commercially available SARS-CoV-2 IgG ELISAs in serum samples of African origin. Tropical Medicine & International Health. 2021;26(6):621-31. doi:https://doi.org/10.1111/tmi.13569
49. Abdella S, Riou S, Tessema M, et al. Prevalence of SARS-CoV-2 in urban and rural Ethiopia: Randomized household serosurveys reveal level of spread during the first wave of the pandemic. EClinicalMedicine. 2021;35. doi:10.1016/j.eclinm.2021.100880
50. Brazil Civil Registry. COVID-19 Portal da Transparencia. 2022.
51. Steinhardt LC, Ige F, Iriemenam NC, et al. Cross-Reactivity of Two SARS-CoV-2 Serological Assays in a Setting Where Malaria Is Endemic. Journal of Clinical Microbiology. 2021;59(7):e00514-21. doi:doi:10.1128/JCM.00514-21
52. Alemu BN, Addissie A, Mamo G, et al. Sero-prevalence of anti-SARS-CoV-2 Antibodies in Addis Ababa, Ethiopia. bioRxiv. 2020:2020.10.13.337287. doi:10.1101/2020.10.13.337287
53. Nkuba AN, Makiala SM, Guichet E, et al. High Prevalence of Anti–Severe Acute Respiratory Syndrome Coronavirus 2 (Anti–SARS-CoV-2) Antibodies After the First Wave of Coronavirus Disease 2019 (COVID-19) in Kinshasa, Democratic Republic of the Congo: Results of a Cross-sectional Household-Based Survey. Clinical Infectious Diseases. 2021. doi:10.1093/cid/ciab515
54. Wiens KE, Mawien PN, Rumunu J, et al. Seroprevalence of Severe Acute Respiratory Syndrome Coronavirus 2 IgG in Juba, South Sudan, 2020(1). Emerg Infect Dis. 2021;27(6):1598-606. doi:10.3201/eid2706.210568
55. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. American journal of epidemiology. 1978;107(1):71-6. doi:10.1093/oxfordjournals.aje.a112510
56. Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. medRxiv. 2020:2020.05.22.20108944. doi:10.1101/2020.05.22.20108944

57. Peluso MJ, Takahashi S, Hakim J, et al. SARS-CoV-2 antibody magnitude and detectability are driven by disease severity, timing, and assay. Science Advances. 2021;7(31):eabh3409. doi:doi:10.1126/sciadv.abh3409
58. Mutevedzi PC, Kawonga M, Kwatra G, et al. Estimated SARS-CoV-2 infection rate and fatality risk in Gauteng Province, South Africa: a population-based seroepidemiological survey. International Journal of Epidemiology. 2021. doi:10.1093/ije/dyab217
59. Perez-Saez J, Lauer SA, Kaiser L, et al. Serology-informed estimates of SARS-COV-2 infection fatality risk in Geneva, Switzerland. medRxiv. 2020:2020.06.10.20127423. doi:10.1101/2020.06.10.20127423
60. Stefanelli P, Bella A, Fedele G, et al. Longevity of seropositivity and neutralizing titers among SARS-CoV-2 infected individuals after 4 months from baseline: a population-based study in the province of Trento. medRxiv. 2020:2020.11.11.20229062. doi:10.1101/2020.11.11.20229062
61. Pérez-Olmeda M, Saugar JM, Fernández-García A, et al. Evolution of antibodies against SARS-CoV-2 over seven months: experience of the Nationwide Seroprevalence ENE-COVID Study in Spain. medRxiv. 2021:2021.03.11.21253142. doi:10.1101/2021.03.11.21253142
62. NEW STATEWIDE DATA SHOW EVIDENCE OF FOUR-FOLD INCREASE IN RECENT COVID-19 INFECTIONS: University of Wisconsin-Madison2020.
63. Barchuk A, Shirokov D, Sergeeva M, et al. Evaluation of the performance of SARS--CoV--2 antibody assays for a longitudinal population-based study of COVID--19 spread in St. Petersburg, Russia. Journal of Medical Virology. 2021;93(10):5846-52. doi:https://doi.org/10.1002/jmv.27126
64. Beverland A, Keogan M, Connell J, De Gascun C, Igoe D. Longitudinal Serological Analysis Following a National Seroprevalence Study to Investigate COVID-19 Infection in People Living in Ireland. The Journal of Infectious Diseases. 2021;224(6):1100-1. doi:10.1093/infdis/jiab346
65. European University of St. Petersburg. By mid-August, less than 13% of St. Petersburg residents had contracted the coronavirus. Most recoveries retain antibodies. 2020.
66. Carreño JM, Mendu DR, Simon V, et al. Longitudinal analysis of SARS-CoV-2 seroprevalence using multiple serology platforms. medRxiv. 2021:2021.02.24.21252340. doi:10.1101/2021.02.24.21252340
67. Kahre E, Galow L, Unrath M, et al. Kinetics and seroprevalence of SARS-CoV-2 antibodies – a comparison of 3 different assays. medRxiv. 2021:2021.03.10.21253273. doi:10.1101/2021.03.10.21253273
68. Thiruvengadam R, Chattopadhyay S, Mehdi F, et al. Longitudinal serology in SARS-CoV-2 infected individuals in India – a prospective cohort study. medRxiv. 2021:2021.02.04.21251140. doi:10.1101/2021.02.04.21251140
69. Muecksch F, Wise H, Batchelor B, et al. Longitudinal Serological Analysis and Neutralizing Antibody Levels in Coronavirus Disease 2019 Convalescent Patients. The Journal of Infectious Diseases. 2020;223(3):389-98. doi:10.1093/infdis/jiaa659
70. Sim M, Cockcroft C, Darby D, et al. Paired sensitivity analysis of four SARS-CoV-2 serological immunoassays in a longitudinal cohort of convalescent hospital staff. Annals of Clinical Biochemistry. 2021:00045632211030957. doi:10.1177/00045632211030957
71. Sharma N, Sharma P, Basu S, et al. The seroprevalence of severe acute respiratory syndrome coronavirus 2 in Delhi, India: a repeated population-based seroepidemiological study. Transactions of The Royal Society of Tropical Medicine and Hygiene. 2021. doi:10.1093/trstmh/trab109
72. Satpati P, Sarangi S, Gantait K, et al. Sero-surveillance (IgG) of SARS-CoV-2 among Asymptomatic General population of Paschim Medinipur, West Bengal, India. medRxiv. 2020:2020.09.12.20193219. doi:10.1101/2020.09.12.20193219
73. Perez-Saez J, Zaballa M-E, Yerly S, et al. Persistence and detection of anti-SARS-CoV-2 antibodies: immunoassay heterogeneity and implications for serosurveillance. medRxiv. 2021:2021.03.16.21253710. doi:10.1101/2021.03.16.21253710

74. Ladage D, Rösgen D, Schreiner C, et al. Persisting Antibody Response to SARS-CoV-2 in a Local Austrian Population. Frontiers in Medicine. 2021;8(881). doi:10.3389/fmed.2021.653630
75. Choe PG, Kim K-H, Kang CK, et al. Antibody Responses 8 Months after Asymptomatic or Mild SARS-CoV-2 Infection. Emerging Infectious Disease journal. 2021;27(3):928. doi:10.3201/eid2703.204543
76. Robert Koch Institut. Local Corona Monitoring: Key Data for Berlin-Mitte. 2021. doi:10.25646/8986
77. Robert Koch Institut. Local Corona Monitoring: Key data for Straubing. 2021.
78. Kar S SS, Murali S, Dhodapkar R, Joseph N, Aggarwal R. Prevalence and Time Trend of SARS-CoV-2 Infection in Puducherry, India. Emerg Infectious Diseases. 2021;27(2):666-9. doi:https://doi.org/10.3201/eid2702.204480
79. Vial PAaG, Claudia and Icaza, Gloria and Ramirez-Santana, Muriel and Quezada-Gaete, Ruben and Nuñez-Franz, Loreto and Apablaza, Mauricio and Vial, M. Cecilia and Rubilar, Paola and Correa, Juan and Pérez, Claudia and Florea, Andrei and Guzman, Eugenio and Lavin, Maria-Estela and Concha, Paula and Najera-de Ferrari, Manuel and Najera-de Ferrari, Manuel and Aguilera, Ximena, . Seroprevalence, Spatial Distribution, and Social Determinants of SARS-CoV-2 in Three Urban Centers of Chile. . SSRN. 2021.
80. Radon K, Bakuli A, Pütz P, et al. From first to second wave: follow-up of the prospective Covid-19 cohort (KoCo19) in Munich (Germany). medRxiv. 2021:2021.04.27.21256133. doi:10.1101/2021.04.27.21256133
81. He Z, Ren L, Yang J, et al. Seroprevalence and humoral immune durability of anti-SARS-CoV-2 antibodies in Wuhan, China: a longitudinal, population-level, cross-sectional study. The Lancet. 2021;397(10279):1075-84. doi:10.1016/S0140-6736(21)00238-5
82. Goto A, Go H, Miyakawa K, et al. Sustained Neutralizing Antibodies 6 Months Following Infection in 376 Japanese COVID-19 Survivors. Frontiers in Microbiology. 2021;12(1039). doi:10.3389/fmicb.2021.661187
83. Domènech-Montoliu S, Puig-Barberà J, Pac-Sa MR, et al. Persistence of Anti-SARS-CoV-2 Antibodies Six Months after Infection in an Outbreak with Five Hundred COVID-19 Cases in Borriana (Spain): A Prospective Cohort Study. COVID. 2021;1(1):71-82.
84. Pagotto V, Luna L, Salto J, et al. Long-Term Duration of Antibody Response to SARS CoV-2 in One of the Largest Slums of Buenos Aires. medRxiv. 2021:2021.03.05.21253010. doi:10.1101/2021.03.05.21253010
85. Rodeles LM PL, Benitez R, Benzaquen N, Serravalle P, Long AK, et al. Seroprevalence of anti-SARS-CoV-2 IgG in asymptomatic and pauci-symptomatic people over a 5 month survey in Argentina. Rev Panama Salud Publica. 2021;45(66).
86. Peluso MJ, Takahashi S, Hakim J, et al. SARS-CoV-2 antibody magnitude and detectability are driven by disease severity, timing, and assay. medRxiv. 2021:2021.03.03.21251639. doi:10.1101/2021.03.03.21251639
87. Perez-Saez J, Zaballa M-E, Yerly S, et al. Persistence of anti-SARS-CoV-2 antibodies: immunoassay heterogeneity and implications for serosurveillance. Clinical Microbiology and Infection. 2021;27(11):1695.e7-.e12. doi:10.1016/j.cmi.2021.06.040
88. Alvim RGF, Lima TM, Rodrigues DAS, et al. Development and large-scale validation of a highly accurate SARS-COV-2 serological test using regular test strips for autonomous and affordable finger-prick sample collection, transportation, and storage. medRxiv. 2021:2020.07.13.20152884. doi:10.1101/2020.07.13.20152884
89. Petersen MS, Hansen CB, Fríðheim Kristiansen M, et al. SARS-CoV-2 natural antibody response persists up to 12 months in a nationwide study from the Faroe Islands. medRxiv. 2021:2021.04.19.21255720. doi:10.1101/2021.04.19.21255720
90. Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the Icelandic Population. New England Journal of Medicine. 2020. doi:10.1056/NEJMoa2006100

91. Campbell H, Gustafson P. Inferring the COVID-19 infection fatality rate in the community-dwelling population: a simple Bayesian evidence synthesis of seroprevalence study data and imprecise mortality data. Epidemiology and infection. 2021;149:e243. doi:10.1017/S0950268821002405
92. Lu N, Cheng K-W, Qamar N, Huang K-C, Johnson JA. Weathering COVID-19 storm: Successful control measures of five Asian countries. American Journal of Infection Control. 2020;48(7):851-2. doi:10.1016/j.ajic.2020.04.021
93. Okell LC, Verity R, Watson OJ, et al. Have deaths from COVID-19 in Europe plateaued due to herd immunity? The Lancet. 2020;395(10241):e110-e1. doi:10.1016/S0140-6736(20)31357-X
94. Fuller JA HA, Victory KR, et al. Mitigation Policies and COVID-19–Associated Mortality — 37 European Countries, January 23–June 30, 2020. MMWR Morb Mortal Wkly Rep. 2020;70:58-62.
95. Miyawaki A, Tsugawa Y. Health and Public Health Implications of COVID-19 in Asian Countries. Asian Economic Policy Review.n/a(n/a). doi:https://doi.org/10.1111/aepr.12358
96. Lee K, Jo S, Lee J. Seroprevalence of SARS-CoV-2 antibodies in South Korea. Journal of the Korean Statistical Society. 2021;50(3):891-904. doi:10.1007/s42952-021-00131-7
97. Parrott JC, Maleki AN, Vassor VE, et al. Prevalence of SARS-CoV-2 Antibodies in New York City Adults, June–October 2020: A Population-Based Survey. The Journal of Infectious Diseases. 2021;224(2):188-95. doi:10.1093/infdis/jiab296
98. Cai R, Novosad P, Tandel V, Asher S, Malani A. Representative estimates of COVID-19 infection fatality rates from four locations in India: cross-sectional study. BMJ Open. 2021;11(10):e050920. doi:10.1136/bmjopen-2021-050920