SARS-CoV-2 antibody dynamics in blood donors and COVID-19 epidemiology in eight Brazilian state capitals

Carlos Prete Junior
Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica da Universidade de São Paulo
Lewis Buss
Instituto de Medicina Tropical da Faculdade de Medicina da Universidade de São Paulo
Tassila Salomon
Faculdade Ciências Médicas de Minas Gerais https://orcid.org/0000-0002-8123-3118
Marcio Oikawa
Universidade Federal do ABC https://orcid.org/0000-0003-2416-1864
Rafael Pereira
Institute for Applied Economic Research – Ipea
Isabel Moura
Faculdade Ciências Médicas de Minas Gerais, Brazil
Lucas Delerino
Fundação Oswaldo Cruz, Brazil
Manoel Barral-Netto
Instituto Gonçalo Moniz - Fundação Oswaldo Cruz (Fiocruz) https://orcid.org/0000-0002-5823-7903
Natalia Tavares
Fundação Oswaldo Cruz, Brazil
Rafael Franca
Fundação Oswaldo Cruz, Brazil
Viviane Boaventura
Fundação Oswaldo Cruz, Brazil
Fabio Miyajima
Fundação Oswaldo Cruz, Brazil; Universidade Federal do Ceará, Brazil
Alfredo Medrone Junior
Fundação Pró-Sangue - Hemocentro de São Paulo
Cesar de Almeida Neto
Fundação Pró-Sangue Hemocentro de São Paulo https://orcid.org/0000-0002-8490-4634
Nanci Salles
Fundação Pró-Sangue - Hemocentro de São Paulo
Suzete Ferreira
Fundação Pró-Sangue - Hemocentro de São Paulo
Karine Fladzinski
Centro de Hematologia e Hemoterapia do Paraná (HEMEPAR), Brazil
Luana de Souza
Centro de Hematologia e Hemoterapia do Paraná (HEMEPAR), Brazil
Luciane Schier
Centro de Hematologia e Hemoterapia do Paraná (HEMEPAR), Brazil
Patricia Inoue
Centro de Hematologia e Hemoterapia do Paraná (HEMEPAR), Brazil

Nelson Fraiji
Fundação Hospitalar de Hematologia e Hemoterapia do Amazonas
Lilyane Xabregas
Fundação Hospitalar de Hematologia e Hemoterapia do Amazonas (HEMOAM), Brazil
Myuki Crispim
Fundação Hospitalar de Hematologia e Hemoterapia do Amazonas
Fernando Araujo
Fundação de Hematologia e Hemoterapia da Bahia (HEMOBA), Brazil
Luciana Carlos
Fundação de Hematologia e Hemoterapia da Bahia (HEMOBA), Brazil
Veridiana Pessoa
Fundação de Hematologia e Hemoterapia da Bahia (HEMOBA), Brazil
Maisa Ribeiro
Fundação HEMOMINAS, Brazil
Rosenvaldo de Souza
Fundação HEMOMINAS, Brazil
Anna Cavalcante
16. Fundação de Hematologia e Hemoterapia de Pernambuco (HEMOPE), Brazil
Maria Valença
Fundação de Hematologia e Hemoterapia de Pernambuco (HEMOPE), Brazil
Maria da Silva
Fundação de Hematologia e Hemoterapia de Pernambuco (HEMOPE), Brazil
Esther Lopes
Instituto Estadual de Hematologia Arthur de Siqueira Cavalcanti (HEMORIO), Brazil
Luiz Amorim
Hemorio
Sheila Mateos
Instituto Estadual de Hematologia Arthur de Siqueira Cavalcanti (HEMORIO), Brazil
Gabrielle Nunes
Instituto Estadual de Hematologia Arthur de Siqueira Cavalcanti (HEMORIO), Brazil; Fundação Oswaldo Cruz, Brazil
Samara Santos
Universidade Federal da Bahia, Brazil
Alexander Silva-Junior
Universidade Federal do Amazonas, Brazil; Centro Universitário do Norte, Brazil; Fundação Oswaldo Cruz, Brazil
https://orcid.org/0000-0002-5645-252X
Michael Busch
Vitalant Research Institute https://orcid.org/0000-0002-1446-125X
Marcia Castro
Harvard T.H. Chan School of Public Health https://orcid.org/0000-0003-4606-2795
Christopher Dye
Department of Zoology, University of Oxford, UK
Oliver Ratmann
Imperial College London, UK https://orcid.org/0000-0001-8667-4118
Nuno Faria
Department of Zoology, University of Oxford
Charles Whittaker
Imperial College London https://orcid.org/0000-0002-5003-2575

Vitor Nascimento
Universidade de Sao Paulo https://orcid.org/0000-0002-0543-4735
Ester Sabino (sabinoec@usp.br)
Departamento de Molestias Infecciosas e Parasitarias & Instituto de Medicina Tropical da Faculdade de Medicina da Universidade de São Paulo

---

Article

Keywords:

Posted Date: February 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1363260/v1

Abstract

There are large differences in the shape and size of regional SARS-CoV-2 epidemics in Brazil. Here we tested monthly blood donation samples for IgG antibodies from March 2020 to March 2021 in eight of Brazil's most populous cities. There was large variation in the inferred attack rate adjusted for seroreversion across cities, and seroprevalence was consistently smaller in women and donors older than 55 years. The age-specific infection fatality rate differed between cities and consistently increased with age. The infection hospitalisation rate (IHR) increased significantly during the gamma-dominated second wave in Manaus, suggesting increased morbidity of the Gamma VOC compared to previous variants circulating in Manaus. The higher disease penetrance associated with the health system's collapse increased the overall IFR by a minimum factor of 2.91 (95% CrI 2.43–3.53). These results demonstrate large heterogeneity in epidemic spread and highlight the utility of blood donor serosurveillance to monitor SARS-CoV-2 epidemics.

Introduction

Brazil has experienced one of the world's most significant COVID-19 epidemics, with over 22 million cases and 621,000 deaths reported as of January 14, 2022. However, this national picture masks important sub-national heterogeneity, with extensive variation in SARS-CoV-2 spread between population groups[1] and locations[2,3] as well as regional differences in the stringency of non-pharmaceutical interventions[4].

Understanding the drivers of these differences is crucial, both retrospectively as a means of evaluating past attempts at controlling spread, but also as a guide to the potential impact of future transmission. Indeed, a significant fraction of the COVID-19 burden in Brazil was driven by the emergence of the Gamma (P.1) Variant of Concern (VOC) in November 2020, which drove extensive resurgence of transmission following its apparent emergence in the Amazonas State capital city of Manaus. Despite the evidence of high levels of population-level immunity that should have hindered further transmission[5], a phenomenon attributed to the Gamma VOC's likely increased transmissibility and ability to partially evade immune responses[6]. Subsequent spread to the rest of Brazil led to similar resurgence, extensive transmission, and disease burden leading to substantial pressure on health systems [7–9]. As with the first epidemic wave, the degree and extent to which different locations were affected varied markedly. Understanding the drivers of this variation is crucial to shed light on how and why SARS-CoV-2 spreads across different populations, and how past epidemics shape subsequent transmission of the virus. More generally, because previous natural infection may enhance vaccine response[10–12], understanding the extent of previous exposure in the country may have important implications for the development of epidemic waves driven by new variants in the context of the ongoing large-scale, nationwide vaccination campaign.

Here, we analyse the divergent epidemic SARS-Cov-2 dynamics in eight of the biggest Brazilian cities (Belo Horizonte, Curitiba, Fortaleza, Manaus, Recife, Rio de Janeiro, Salvador and São Paulo). We estimate the seroprevalence over time for these cities disaggregated by age and sex using repeated cross-sectional convenience samples of routine blood donors collected from March 2020 to March 2021. We also provide estimates for the age-specific Infection Fatality Rates (IFR, defined as the number of deaths per infection) and Infection Hospitalisation Rates (IHR, the number of hospitalisations per infection) for these cities. In Manaus, the Gamma VOC became dominant before March 2021 (see Supplementary Fig. 1), enabling us to provide estimates of Gamma's IFR and IHR. Our results highlight important differences in the drivers of SARS-CoV-2 epidemic spread across Brazil's major population centres, and underscore the utility of blood donors for regular serosurveillance as a tool to track progression of epidemics of emerging infectious diseases.

Results

Serology assay validation and antibody waning

We used the Abbott SARS-CoV-2 anti-nucleocapsid (N) IgG chemiluminescent microparticle immunoassay (CMIA; AdviseDx, Abbott) to determine antibody prevalence in serial monthly samples of residual whole blood donor sera. All SARS-CoV-2 anti-N

IgG assays suffer from signal waning - resulting in positive-negative transition, or “seroreversion” - during convalescence[13]. This amounts to a fall in assay sensitivity through time. The Abbott anti-N IgG CMIA shows particularly rapid signal decay when compared with other assays[14] (Fig. 1A, Supplementary Fig. 2). These antibody dynamics mean that as an epidemic progresses, the crude proportion of individuals with a positive test result will increasingly underestimate the true attack rate[5,15,16]. To attenuate the effect of seroreversion, we defined test positivity using the lower threshold of the signal-to-cutoff value recommended by the manufacturer (0.49).

Antibody kinetics vary with disease severity[5,16,17], and whole blood donors represent predominantly asymptomatic or mild SARS-CoV-2 infections due to donation eligibility criteria[5]. As such, we sought to estimate a time-to-seroreversion distribution that accurately reflected the blood donor convenience sample used in this study. We identified and tested 7675 repeat whole blood donors in Manaus who had made multiple donations throughout for 2020-21 and used these data to estimate the time-to-seroreversion probability distribution (see Methods). The results are shown in Figs. 1B, 1C, and 1D, which compare the half-life, peak signal-to-cutoff values and time-to-seroreversion of repeat whole blood donors to the cohort of symptomatic convalescent plasma donors used to determine sensitivity. Repeat blood donors had a shorter assay signal half-life than plasma conors (median [IQR] 69.3 [53.0–103.8] *versus* 105.9 [62.7–185.1] days) and a lower observed peak signal-to-cutoff ratio (median [IQR] 2.89 [1.49–4.83] *versus* 5.08 [3.22–6.99]), yielding a shorter time between seroconversion and seroreversion (203 [147–294] *versus* 280 [175–441]). This highlights the importance of choosing a time-to-seroreversion distribution that is appropriate for the use case - i.e., the rate of waning seen in PCR-confirmed symptomatic disease would have resulted in underestimation of SARS-CoV-2 attack rates.

We developed an analytic method to correct raw seroprevalence data for seroreversion, improving on the method used in Buss et al[5]. We first estimate the time-to-seroreversion distribution (Supplementary Fig. 3) using serial donations from repeat blood donors, which determines how sensitivity for a given individual decreases with the time after seroconversion. We then corrected the raw seroprevalence estimates for the changing sensitivity within a Bayesian framework (see online Methods). We first calculated attack rates for each age and gender group in each city, and summed these using the proportion of each group in the Brazilian reference population to obtain standardised estimates. It is worth noting that the model estimates the incidence of primary infections and reinfections among seronegatives and computes the seroprevalence as the cumulative sum of the incidence. Hence, the obtained seroprevalence can surpass 100% due to reinfections among seronegative individuals, and reinfections among seropositive donors are not accounted for in the seroprevalence estimate.

COVID-19 mortality across Brazilian capitals

The location of the eight Brazilian state capitals that contributed serology data are shown in Fig. 2A. They collectively represent approximately 14% of the total Brazilian population. The age distributions of the eight cities differ widely (Supplementary Fig. 4), as such COVID-19 mortality is presented as age-standardised rates (see Supplementary Fig. 5 for the crude mortality curves). Between March 1st 2020 and March 31st 2021, the age-standardised mortality rate varied from 1.7 deaths per 1,000 inhabitants in Belo Horizonte, to 5.3 deaths/1,000 in Manaus, which had twice the mortality of Fortaleza, the city with the next highest mortality (Fig. 2C).

Figure 2B shows the homestay index for the eight cities (see Supplementary Information for the definition). Manaus, the city with the youngest population, returned to pre-pandemic levels of mobility by July 2020, having consistently lower homestay index (i.e., higher mobility) than other cities after June 2020, whereas the other seven cities showed a relatively homogenous mobility pattern. The shape of the mortality curves also varied markedly (Fig. 2D). Manaus was also an outlier in having the lowest income *per capita*, health insurance coverage, and lowest proportion of the population with comorbidities, along with the highest number of residents per household (Supplementary Fig. 6).

Blood donor serosurveillance

Using an average of 951 monthly samples of routine whole blood donations (Mar 2020 to Mar 2021, a total of 97,950 samples) in each of the eight cities, we measured the crude seroprevalence of anti-N IgG antibodies detectable by the Abbott

CIMA (Table 1). However, these raw estimates of seroprevalence are affected by seroreversion dynamics and provide a poor guide for assessing past levels of population exposure.

Table 1

– Attack rate estimates for December 16, 2020 (before the Gamma-dominated wave in Manaus) and February 24, 2021. This table contains the attack rate estimated by our seroreversion correction model along with the crude seroprevalence and the seroprevalence corrected only by sensitivity, specificity and reweighted by age and sex, without any correction for seroreversion. Seroprevalence estimates can surpass 100% due to reinfections. Seroprevalence estimates are only available for all cities simultaneously until February 24, 2021.

	December 2020			February 2021		
	Crude Seroprevalence (%)	Seroprevalence corrected by sensitivity, specificity and reweighted by age and sex (%)	Estimated Seroprevalence (%)	Crude Seroprevalence (%)	Seroprevalence corrected by sensitivity, specificity and reweighted by age and sex (%)	Estimated Seroprevalence (%)
Belo Horizonte	12.8 (10.8–15.1)	13.1 (10–16.5)	20.6 (18.6–22.7)	25.2 (22.5–28.1)	27.8 (23.7–32.3)	27.8 (25.2–30.6)
Curitiba	15.2 (13–17.5)	14.4 (11.5–17.7)	19.3 (17.5–21.2)	30.8 (28–33.7)	31.1 (27.4–35.1)	27.6 (25.2–30.3)
Fortaleza	28.1 (25.4–31)	29.8 (25.6–34.2)	48.8 (45.4–52.7)	24.9 (22.2–27.7)	26.6 (22.6–31)	57.4 (53.3–62.2)
Manaus	34.6 (31.6–37.7)	36.1 (32–40.6)	75.0 (70.8–80.3)	47.7 (44.4–51)	52.6 (47.8–57.8)	95.8 (90.6–102.5)
Recife	27 (24.3–29.8)	30.7 (25.9–35.8)	49.4 (46.1–53.1)	27.4 (24.7–30.2)	27.1 (23.2–31.5)	59.9 (55.6–64.6)
Rio de Janeiro	24.4 (21.7–27.2)	24.8 (21.4–28.4)	42.2 (39.4–45.4)	34.5 (31.6–37.5)	36.8 (32.9–41)	54.7 (51.1–58.9)
Salvador	18.4 (16–21)	18.4 (15.3–22)	35.3 (32.8–38.1)	20.1 (17.6–22.7)	20.1 (16.8–23.8)	42.4 (39.5–45.8)
São Paulo	18.8 (16.5–21.4)	19.0 (15.9–22.4)	26.6 (24.3–29.1)	21.7 (19.1–24.4)	22.3 (19–25.9)	33.3 (30.3–36.5)

Using our Bayesian seroreversion correction model, we present in Fig. 3 the age-standardised SARS-CoV-2 attack rates (i.e., the cumulative rate of the population that was infected or reinfected) as of March 2021 after accounting for test sensitivity, test specificity and IgG seroreversion (coloured lines) along with the directly measured seroprevalence (light grey boxplots) and the estimated seroprevalence adjusting for test sensitivity and specificity (dark grey boxplots). Our results further underscore the significantly different scales of SARS-CoV-2 epidemic impact experienced across the eight cities, with the implied attack rates ranging from only 19.3% in Curitiba, to as high as 75.0% in Manaus by December 2020 (see Table 1). Alternative cumulative incidence estimates produced using different time-to-seroreversion distributions are similar to those in Fig. 3 and shown in Supplementary Fig. 7. We note that even though the seroprevalence estimated by our model includes reinfection in seronegative individuals, the model does not capture reinfection in already positive individuals, and so is likely to underestimate SARS-CoV-2 attack rates in scenarios where reinfection is not rare.

The slope of the seroprevalence curves (Fig. 2D) also differed significantly across cities, showing different dynamics of antibody acquisition at the population level according to the shape and dynamics of the epidemic experienced. Cities with only minimal epidemic peaks as Belo Horizonte and Curitiba showed near constant rates of increase in seropositivity after adjustment for antibody waning. By contrast, cities with substantial epidemic peaks as Fortaleza and Manaus demonstrated

significant variation in the rate at which estimated seropositivity increased in the population, with these rates highest during the epidemic peaks. These findings highlight the capacity of blood donor-based serological data to recapitulate important temporal trends in the intensity and dynamics of the epidemics across these eight cities.

The estimated seroprevalence in June and July in Fortaleza was significantly smaller than the measured seroprevalence without correction for seroreversion, even though the seroprevalence estimate disaggregated by age and sex (Supplementary Fig. 8) lies within the confidence intervals of the measured seroprevalence. This effect happened especially in women, which had a crude seroprevalence that was significantly larger than in men in June and July 2020, but became similar in the following months. It is possible that the seroreversion rate observed in Fortaleza had been faster than the rate estimated from repeat blood donors, in which case we undercorrected for seroreversion, underestimating the attack rate. However, a more likely explanation is that samples between March and July 2020 for Fortaleza are less representative of the population, since only 39.4% from 4,970 selected samples could have been retrieved and tested, compared to 97.0% for the other cities and months. As such, seropositive individuals from Fortaleza may have been more likely to donate in these months, leading to an overestimated crude seroprevalence.

Age-sex patterns in blood donor seroprevalence

We next examine the patterns and dynamics of attack rates across different groups by disaggregating the seroprevalence data by age and sex. The seroprevalence estimates disaggregated by age and sex are shown in Fig. 3 (see Supplementary Figs. 9–10 for seroprevalence disaggregated by only age or sex). Across the eight cities, our results consistently show differences between sexes - on average, men tended to have higher attack rates than women, although the degree and extent of this difference varied between cities. In São Paulo, the seroprevalence in December 2020 for men was 30.6% compared to the 23.0% estimated in women (i.e. 33.5% (95% CrI 17.7–51.9) higher, Fig. 3B). By contrast, seroprevalence in Curitiba in December 2020 was similar for women and men, being only 4.65% (95% CrI − 11.5–18.5) higher in women.

We also observed an extensive variation in the dynamics of population-level seroprevalence between age-groups, with seroprevalence in December 2020 typically highest in younger age-groups. The seroprevalence of individuals below the age of 55 increased in all cities except for Recife when compared to donors aged between 55 and 69, increasing by 34.1% (95% CrI − 2.23–91.2) in Curitiba and decreasing by a small factor of 0.5% (95% CrI − 24.8–19.1) in Recife. Furthermore, in cities with a large increase in seroprevalence during the first epidemic wave (i.e., Manaus, Recife, Fortaleza, and Salvador), this was primarily driven by younger men. In these locations, the differences between age-sex groups slowly narrowed during the long period of less intense transmission (Fig. 3A). This highlights important differences between age-groups in the extent to which they were exposed to the virus and/or contributed to transmission at different points during the regional epidemics - differences that are not evident, or certain, from case or death counts alone.

In addition to the differences in attack rates by age and sex, seroprevalence did not increase homogeneously among different age and sex groups. In Manaus, seroprevalence was significantly larger in men and younger individuals aged 16–44 in July 2020, but between July and December seroprevalence increased faster in women and donors older than 45 years, leading to smaller differences in attack rate by age and sex in December 2020 (Fig. 3B). Similar patterns are also observed in Salvador, Recife and Fortaleza, although with smaller age and sex inequalities.

Variation in the SARS-CoV-2 Infection Fatality Rate (IFR) Across Age-Groups, Sexes and Locations

Using estimates of the cumulative number of individuals infected alongside records of COVID-19 deaths available from Brazil's SIVEP-GRIPE platform, we next calculated the infection fatality rate (IFR) for each city and age-group. Figure 5A presents the estimated age-specific IFRs for each municipality as of December 2020, before the Gamma VOC epidemic in Brazil. Our results show the IFR significantly increases with age, ranging from 0.03% in individuals aged 16–24 year, to 1.31% in individuals aged 55–64 year. This is in-keeping with previous work highlighting a strong age-dependency in COVID-19

mortality[5,18,19]. Cities presented different age-standardised overall IFRs, being smaller in Manaus (0.24%) and higher in Curitiba (0.54%).

There was a strong correlation (Pearson's correlation = 0.92) between the age-standardised mortality rate in each city and the attack rate inferred from blood donor serosurveillance data (Fig. 5B). Both the overall IFR and the overall IFR adjusted for the age structure of the city differed significantly between cities (Fig. 5C), showing that the IFR differences cannot be explained only by the different age structures. Despite the differences between cities, the obtained age-specific IFRs were similar to the estimates from Brazeau et al[18] but higher than the estimates from O'Driscoll et al[19] (Fig. 5D). The age-specific and overall Infection-Hospitalisation Rate (IHR) were also estimated (Supplementary Fig. 11) and showed similar patterns, being larger in Belo Horizonte, Curitiba and São Paulo.

The obtained IFRs and attack rates for December 2020 were validated using alternative approaches that do not correct directly for seroreversion, not depending on the proposed seroreversion correction model (see Supplementary Information).

The Dynamics and Epidemiological Impacts of the Gamma VOC in Manaus

In Manaus, the majority of infections after mid-December are attributable to the Gamma variant, which prompted us to estimate attack rates for the period from December 16 2020 to March 31 2021, and compare these to the period March 1st, 2020 to December 15th, 2020 during which almost all infections are attributable to non-Gamma variants[6,20]. Since reinfections were common in the Gamma-dominated wave in Manaus[21], reinfections must be considered to estimate the attack rate. However, we are unable to identify which of the seropositive blood donors are primary infections and which are re-infections, and so we are unable to provide a point estimate of cumulated attack rates in the Gamma-dominated period. Instead, we calculated upper bounds assuming maximum proportions of reinfections (see online Methods).

The inferred upper bound of the age-specific attack rate in the Gamma-dominated period in Manaus ranged from 30.6% (95% Bayesian Credible Interval (CrI) 22.8–41.1) to 46.0% (95% CrI 32.8–60.6) in individuals aged 45–54 and 55–69 (Supplementary Fig. 12), showing small variation of the attack rate among age groups. The estimated upper bound for the age-standardised cumulative attack rate in the second period dominated by the Gamma variant was 37.5% (95% CrI 35.3–42.6), significantly smaller than the cumulative attack rate of 75.0% (Fig. 2) estimated for the first period dominated by non-Gamma variants.

Comparing to the COVID-19 attributable hospitalisations and deaths reported to the SIVEP database, we next used the estimated upper bounds of the age-specific attack rates in the Gamma period in Manaus to calculate lower bounds of the age-specific Infection-Hospitalisation Rate (IHR) and Infection Fatality Rate (IFR) for the Gamma period. We then compared with the IFRs and IHRs obtained with the attack rate estimated for the period during which non-Gamma variants dominated (March 1st, 2020 to December 15th, 2020). The resulting age-specific IFRs and IHRs are shown in Fig. 5A and 6B respectively, and the relative risks obtained using the IFR or IHR in December 2020 as baseline in Figs. 6C and 6D. The lower bound for the IHR increased in all age groups, from 34.4% (95% CrI 6.5–70.0) in individuals aged 16–24 to 163.4% (95% CrI 90.9–264.3) in individuals aged 45–54 when compared to the IHR estimated for the non-Gamma period. The increased hospitalisation risk combined with an increased in-hospital fatality rate (HFR, defined as the number of deaths per hospitalisation) during the second wave (Supplementary Fig. 13) resulted in an increased age-specific IFR, with a lower bound increasing 93.8% (95% CrI 36.4–186.4) in individuals aged 55–64 to 273.5% (95% CrI 167.8% – 423.4%) in individuals aged 45–54 when compared to the first wave (Fig. 5C). As such, even though the IFR and IHR increased for all age groups during the Gamma-dominated period, this difference was more significant in younger age groups. The obtained lower bound for the overall IFR was 0.527% (95% CrI 0.447–0.630), 2.91 (95% CrI 2.43–3.53) times higher than the estimated IFR for the first wave in Manaus.

Discussion

Our results highlight the divergent epidemic dynamics across eight of Brazil's biggest cities as reflected by mortality rates, and show that these differences are recapitulated in blood donor based serial cross-sectional serosurveillance. Despites the large IFR differences observed across cities, seroprevalence was strongly correlated with cumulative age-standardised mortality (Fig. 4B). These results reinforce the validity of blood donors as a convenient population for serosurveillance. A previous study[22] has highlighted the need for a reliable, cost-effective method of immunological surveillance to provide evidence of past infection and to understand the dynamics of emerging disease. Even though serology is less precise for identifying infections on an individual level, it is an effective tool for monitoring epidemics at a populational level. As blood donation programs are an existing component of medical infrastructure globally and in which blood samples are readily available in many locations, this approach can be rapidly implemented and carried out in large populations.

We estimated larger attack rates in individuals aged 16–54 years. This is consistent with previous work examining age patterns of transmission from mobility data in the United States[23], but we have measured infection directly rather than making inferences indirectly on the basis of COVID-19 deaths and movement. Possible reasons for the higher attack rate in people aged 16–54 years include, but are not limited to, different risk perception and shielding practises, and disease biology with more frequent asymptomatic infections in younger people, which increases infection risk in this age group due to greater mixing among working age adults. We also found overall higher levels of seroprevalence in men compared to women, and these patterns changed over time. For instance, in Manaus, a very high seroprevalence was reached rapidly among young men by July 2020, after which relatively little increase in overall seroprevalence occurred in men. By contrast, among older women, who reached less than half the attack rate seen in men by June, the seroprevalence continued to increase. This heterogeneity in transmission in a location with high overall antibody prevalence meant that some groups remained relatively susceptible and perpetuated transmission at a lower level[24,25]. Other works suggest that socioeconomic condition also contributed to heterogeneity of SARS-CoV-2 spread in Manaus[25], which is confirmed by the high seroprevalence observed in Black and less-educated donors (Supplementary Fig. 14).

We also confirm a strong age dependency of COVID-19 IFRs[18,19]. Although age-specific IFRs were roughly similar across the cities (Fig. 4A), and similar to estimates in the literature (Fig. 4D), there were some noticeable differences. For example, the more affluent south and southeastern cities of Belo Horizonte, Curitiba and São Paulo tended to have higher age-specific IFRs, whereas in the northern and northeastern cities of Manaus, Salvador, and Fortaleza, the age-specific IFRs tended to be lower. This may be due to underreporting of deaths, but might also reflect lower prevalence of comorbidities in the latter populations (Supplementary Fig. 6). Cities with larger IFRs also had larger IHRs, suggesting that the differences in IFR reflect the different risks of developing a severe disease. The different lineages circulating in the eight cities may have also contributed to the observed IFR and IHR difference (Supplementary Fig. 1). While most of the cases in the first wave in Amazonas and Ceará were caused by earlier lineages, the lineages B.1.1.28, B.1.1.33 and later P.2 (Zeta) were more prevalent in other states. It is worth noting the IHR also depends on the probability of an individual with severe disease being hospitalised. This probability depends on access to health facilities and availability of healthcare resources, and therefore may vary across cities even if disease severity remains constant.

Our results also clearly demonstrate a higher IHR during the Gamma-dominated observation period compared to the non-Gamma observation period in Manaus for all age groups (Fig. 5B, 6D). This supports observations[26] that the Gamma VOC tends to cause more severe disease than the ancestral non-Gamma variants circulating locally, even among young adults in Manaus. The larger increase in IHR for younger adults aged 25–54 years is compatible with the younger profile of hospitalisations of the Gamma-dominated wave in Brazil[27,] observed before vaccination coverage reached significant levels in older age groups[8]. In Manaus, the increased levels of hospitalisation caused parts of the healthcare system to collapse during the second wave causing an increase in hospital fatality rates (HFR) as previously described[9], further increasing the IFR. The higher IFR associated with Gamma VOC infection during the second wave is therefore due to a combination of two factors – increased disease severity resulting in a greater proportion of infections requiring hospital-based care (the IHR, arising primarily from intrinsic viral properties and pathogenicity), and the impacts of this increased healthcare pressure on mortality within-hospitals (the HFR, arising primarily from healthcare pressure).

Furthermore, there are some relevant limitations to our results that need to be pointed. Firstly, blood donors are a convenience sample and extrapolation to the entire population should be done with caution. Due to eligibility criteria in Brazil, blood donors are limited to those aged 16 to 69 years, with a strong skew towards younger adults even within this eligibility range (Supplementary Fig. 15) in most Brazilian regions. However, our results do suggest that blood donor serosurveillance agrees with other metrics of epidemic size as mortality, both cumulatively (Fig. 5B) and through time (Fig. 3). Moreover, both sensitivity and seroreversion could be an age-dependent process as a proxy for disease severity i.e., older individuals are more likely to be symptomatic, seroconvert and have longer time to seroreversion. Indeed, we see this pattern of longer half-lives and larger peak signal-to-cutoffs in convalescent plasma donors who had recovered from more severe disease (Supplementary Fig. 16). Therefore, correction of crude seroprevalence for antibody waning could possibly be confounded by demographic differences between the eight cities. However, since individuals that had a severe disease are unlikely to donate blood, seropositive whole blood donors are likely fairly homogenous in having had milder or asymptomatic disease, and as such, the rate of waning may not vary significantly between locations. An additional important point to note is that the longer an epidemic last, the more frequent reinfections become due to the natural waning of immunity in the time-period following infection. Our data span over a year of transmission in areas with multiple waves with high SARS-CoV-2 burden and consequently non-negligible reinfection rates, as such it is difficult to reliably infer the attack rates from seroprevalence data towards the end of the time series. For this reason, our model produces upper bounds for cumulative prevalence of > 100% in Manaus by early 2021.

Despite these limitations, blood donors represent an accessible population to detect trends of the epidemic that otherwise could only be obtained through expensive population-based studies, which are difficult to establish in Brazil during the course of a rapidly progressing epidemic. Studies to understand the main differences between blood donors and the general population would help the development of better sampling protocols to mitigate bias and should be part of preparedness for future epidemics of infectious diseases.

Online Methods

Selection of blood donors for estimation of seroprevalence

Each of the eight cities had a monthly quota of 1,000 kits for testing selected donation samples in this study. In order to select more representative samples, we selected blood samples so that the spatial distribution of residential location of selected donors match the spatial distribution of population density in each municipality. More specifically, each city was divided into sub-municipal administrative zones, and the original quota (1,000 kits) was divided into sub-quotas following the populational distribution of the city administrative zones. Starting from the second week of each month, we selected consecutive blood donors based on the geolocation of their residential postcode to fill the sub-quotas. In this way, donations with missing or wrong postal code were considered ineligible for selection.

In Manaus, however, donor postcodes were not reliably collected, so that the number of missing and wrong values make this strategy unfeasible. So, samples were selected consecutively with no postal code restrictions. We also developed a study management system to operationalize this sampling strategy, whereby blood donor postcodes and epidemiological data were automatically extracted and selected. After that, the selected donation sample IDs were released for the research assistant to be separated for testing.

In Brazil, blood donation samples are usually saved for six months, so when serological test kits were made available in July 2020, we could retrospectively select and test frozen samples from February to July. After this period samples were selected and tested in real time. Antibody tests results were not made available to the blood donors themselves.

Blood donors are a convenience sample, and thus may not be representative of the wider population in terms of their risk of SARS-CoV-2 exposure. Supplementary Figs. 17–19 shows a comparison between recorded blood donor demographics and the last available Brazilian census conducted in 2010. Donors differ systematically in age, sex and self-reported skin colour

compared to the population, but the income per capita is similar. To account for the differences in the age-sex structure of blood donors, we divide donors in age-sex groups and estimate the prevalence of each age-sex group separately. Then, we calculate the seroprevalence of the population as a weighted sum of the seroprevalences of each age-sex group.

SARS-CoV-2 serology assays

We applied chemiluminescent microparticle immunoassays (CIMA, AdviseDx, Abbott) that detect IgG antibodies against the SARS-CoV-2 nucleocapsid (N) because it was the only automated commercially available kit in Brazil when the study started (July 2020) and we used this kit throughout the study until March in all eight cities except Recife, where we used the kit until February 2021. Supplementary Figs. 20–22 contain the number of tests disaggregated by month, age, sex and the monthly signal-to-cutoff distribution. We also decided to validate the results observed in Manaus, as this represents a unique sentinel population, by retesting all samples in November 2020 using the chemiluminescent microparticle immunoassay (CIMA, AdviseDx, Abbott) that detect IgG antibodies against the SARS-CoV-2 spike (S) protein (see Supplementary Information for the validation analyses).

To determine the test sensitivity, we considered a cohort of 208 non-hospitalised symptomatic SARS-CoV-2 PCR-positive convalescent plasma donors tested within 60 days after symptom onset (Supplementary Table 1). These donors had symptomatic COVID-19 with PCR-confirmed SARS-CoV-2 infection and were recruited to provide convalescent plasma. We found a sensitivity of 90.6% for the anti-N assay using a threshold of 0.49 signal-to-cutoff and 94.0% for the anti-S assay. Specificity for the anti-N assay was 97.5%, with 801 negative results in 821 pre-pandemic blood donation samples[5]. Sensitivity and specificity for other assay thresholds are shown in Supplementary Table 1. The anti-S assay has a specificity of > 99% [13,14], and we assume 100% in this study. Although the sensitivity of both assays declines through time due to waning of the detected antibodies below the positivity threshold, the anti-S IgG antibodies wane more slowly[13,14]. As previously highlighted, the convalescent plasma donors cohort overestimates sensitivity for our use case due to spectrum bias. This is because convalescent donors had moderate-to-severe SARS-CoV-2 infection, and thus differ from the whole blood donor population (used to estimate seroprevalence), who are more likely to have had asymptomatic or mild disease.

We subsequently estimated the distribution of time-to-seroreversion, and thus the sensitivity decreasing through time, for the anti-N assay. We first calculated this in the convalescent donors, in whom the date of symptom onset is known, and whose blood samples were collected longitudinally during convalescence. As such the time-to-seroreversion distribution was computed after accounting for right censoring. However, again due to spectrum bias, the extrapolation of antibody waning from convalescent donors to whole blood donors is unlikely to be valid. As such, we obtained a second cohort of repeat blood donors in Manaus that provided multiple donations during the 2020–2021 period. These donors are expected to have the same antibody dynamics as the seroprevalence cohort, as they are drawn from the same population and have predominantly mild or asymptomatic infections. However, in this group the time of infection is unknown, as infection is inferred by serostatus alone. The procedure to manage this problem is described below.

Methods used to estimate the time-dependent sensitivity

Let se_0 be the sensitivity measured shortly after symptomatic infection (that is, the probability of an infected individual seroconverting to an S/C above the threshold), and $p^+[m]$ the probability of a donor remaining positive m weeks after seroconversion (given that the donor seroconverted). Then, the sensitivity of the test m weeks after seroconversion is then $se_0 \times p^+[m]$ for a given donor. In this section we describe the procedure used to determine $p^+[m]$ from repeat blood donor data, for which time of infection and time of seroreversion are unknown. The seroreversion correction model described in the next section uses the estimate of $p^+[m]$ to calculate the seroprevalence accounting for seroreversion.

The criteria to select repeat blood donors were: 1. at least one positive test, indicating SARS-CoV-2 infection, 2. at least one subsequent blood sample, in order to interpolate the date of seroreversion, and 3. falling S/C between these two samples, because one of the samples used to define the interpolation curve may have occurred before the peak S/C, hence the half-life

and the date of seroreversion cannot be estimated. Therefore, all selected donors had at least one positive sample and at least one subsequent sample (positive or negative) with smaller S/C.

To calculate $p^+[m]$, we first estimate the date of seroreversion for each repeat blood donor using an exponential interpolation (a linear interpolation in the log scale). We choose an exponential interpolation because an exponential decay is frequently used to model antibody dynamics[16]. When seroreversion is interval-censored, i.e., a donor that has a positive test subsequently becomes negative, we interpolate an exponential curve that passes through the last positive sample and the first negative sample. Otherwise, when seroreversion is not interval-censored, then it is right-censored (a donor remains positive on their last sample), in which case we extrapolate an exponential line through the last two positive samples and project this forward. As such, the estimated instant of seroreversion for blood donor i (denoted as t_i^-) is the point where the interpolation curve crosses the threshold for a positive test. The interpolation procedure is illustrated in Supplementary Fig. 23. The proposed method may overestimate t_i^- if an S/C used to define the interpolation curve was sampled shortly after seroconversion before the peak S/C was reached, since in this case the S/C curve does not behave as an exponential, leading to an overestimated half-life. To partially overcome this problem, we discard donors that do not serorevert within 106 weeks (2 years) after seroconversion.

After estimating t_i^-, for each blood donor i, we compute the probability distribution of the date of seroconversion for that donor, p_i. For this, we identify the earliest and latest possible date of seroconversion t_{min} (the date of the last negative result before seroconversion or March 1st 2020 if the donor has no positive results before seroconversion) and t_{max} (the date of the first positive result). The relative probability of seroconversion within this window depends on the incidence of seroconversions due to SARS-CoV-2 infection for the cohort of repeat donors, denoted $u_{repeat}[n]$. To estimate this quantity, we calculate the histogram of the date of first positive donation for repeat blood donors and then apply a 30-day moving average. As a sensitivity analysis, we also calculate $u_{repeat}[n]$ by computing the histogram of the date of onset of SARI deaths observed in Manaus, and applying to it a 7-day window moving average, yielding similar results (Supplementary Figs. 3 and 7).

The distribution of the date of seroconversion is obtained by truncating the incidence curve of repeat blood donors $u_{repeat}[n]$ in the interval $[t_{min}, t_{max}]$ and renormalizing the distribution. We then generate 1,000 samples of the instant of seroconversion $t_i^+ \sim p_i[n]$ and compute the 1000 sample delays between seroconversion and seroreversion $\Delta t_i = t_i^- - t_i^+$.

The probability of the delay between seroconversion and seroreversion being $n \geq 1$ days (denoted as $p^-_{day}[n]$) is calculated with the empirical histogram of the $1000 \times N_{donors}$ samples of Δt_i, $i = 1, \cdots, N_{donors}$. The distribution $p^-_{day}[n]$ is then binned into weeks by taking the average of $(p^-_{day}[7n], p^-_{day}[7n+1], \cdots, p^-_{day}[7n+6])$ for $n \geq 1$. The resulting distribution, denoted as $p^-_{week}[n]$, represents the probability of seroreversion exactly n weeks after seroconversion.

Finally, the probability of a donor remaining positive n weeks after seroconversion

$p^+[n]$ (i.e., the probability of a donor seroconverting after week n) obtained through

$$p^+[n] = 1 - \sum_{k=1}^{n} p^-_{week}[k].$$

The presented method is summarised in **Algorithm 1** in Supplementary Information.

Our proposed seroreversion correction model

Here, we present a Bayesian model that draws posterior samples from the incidence over time corrected by sensitivity, specificity and seroreversion using as input the estimated curve for $p^+[n]$, the number of weekly positive tests and total number of tests. Even though the main output of the model is the incidence at week n for age-sex group a (denoted as $u[n, a]$

), the seroprevalence at week n for group a can be calculated from $u[n, a]$ as $\rho[n, a] = \sum_{k=1}^{n} u[k, a]$. For simplicity, the proposed model ignores the delay between infection and seroconversion, as it should have small impact on the estimate of $u[n, a]$. To define the age-sex groups, age was discretized in the intervals 16–24, 25–34, 35–44, 45–54 and 55–70.

Assuming that the sensitivity se_0 and specificity sp of the assay are independent of the age-sex group, the probability of a random person from age-sex group a being tested positive at week n, denoted as $\theta[n, a]$, is

$$\theta[n, a] = se_0 \sum_{k=1}^{n} p^+[n-k]u[k, a] + (1 - sp)\left(1 - \sum_{k=1}^{n} u[k, a]\right).$$

The derivation of the expression above is presented in Supplementary Information. The left term $se_0 \sum_{k=1}^{n} p^+[n-k]u[k, a]$ represents true positives (previously infected donors that are still seropositive), while the right term $(1 - sp)\left(1 - \sum_{k=1}^{n} u[k, a]\right)$ represents false positives (uninfected donors that test positive).

Let us denote as $T^+[n, a]$ and $T[n, a]$ respectively the number of positive tests and the total number of tests for week n and age-sex group a. Given $\theta[n, a]$, the probability distribution of $T^+[n, a]$ is

$$T^+[n, a] \mid \theta[n, a] \sim Binomial(T[n, a], \theta[n, a]).$$

We use a Bayesian framework to draw posterior samples from u assuming a non-informative prior, but limiting the final seroprevalence in the interval $[0, b]$, where b is a fixed input of the algorithm that can be 1 or 2 depending on whether reinfections are allowed, and we use $b = 2$ in this work. Instead of defining a prior distribution for $u[n, a]$ directly, we decompose it into $u[n, a] = \rho_{max}[a]\, u_{norm}[n, a]$, where $\rho_{max}[a] \sim Uniform(0, b)$ sets the upper bound of the final prevalence to b and $u_{norm}[:, a] \sim Dirichlet(1, 1, \ldots, 1)$ is the normalised incidence which sums to 1. This decomposition is equivalent to assuming a uniformly distributed prior for $u[:, a]$ in the simplex $0 \leq \sum_{n=1}^{N} u[n, a] \leq b$ with $u[n, a] \geq 0 \forall n$.

After drawing posterior samples from $u[n, a]$, we calculate the seroprevalence at week n for age-sex group $a$as $\rho[n, a] = \sum_{k=1}^{n} u[k, a]$and then compute the age-sex weighted seroprevalence $\rho[n]$, given by

$$\rho[n] = \sum_{a=1}^{M} w[a]\rho[n, a],\ w[a] = \frac{\mathrm{pop}[a]}{\sum_{k=1}^{M} \mathrm{pop}[k]},$$

where $\mathrm{pop}[a]$ is the population for the age-sex group a in the corresponding city.

The presented Bayesian model is summarised in **Algorithm 2** in Supplementary Information. The posterior samples are drawn using a Monte-Carlo Markov Chain (MCMC) algorithm with 100,000 iterations.

The incidence returned by the model was validated through posterior predictive checks by randomly selecting 1,000 samples from $u[n, a]$and drawing samples from $T^+[n, a] \mid \theta[n, a] \sim Binomial(T[n, a], \theta[n, a])$. The resulting crude seroprevalence is then compared with the measured crude seroprevalence (Supplementary Fig. 24).

The proposed Bayesian seroreversion correction model can be seen as an improvement on that presented in Buss et al[5]. The model in Buss et al. assumed a parametric form for time to seroreversion and derived the parameters by assuming an increasing cumulative incidence in the repeated cross-sectional samples of blood donors in Manaus. Here we derived the distribution directly from repeat blood donors without assuming any parametric form. Also, Buss et al. applied the seroreversion correction method to the measured seroprevalence corrected for sensitivity, specificity and reweighted by age and sex, while here we estimate the seroprevalence in each age group separately, allowing the identification of non-homogeneous incidence in different age groups.

Despite these differences, the results presented here (Table 1, Fig. 2D) are compatible with the seroprevalence estimates of 28.8% and 76.0% respectively for São Paulo and Manaus in Buss et al. The proposed seroreversion method also differs from other methods in the literature[16,28] in that we use the incidence curve to estimate the time-dependent sensitivity instead of the deaths or confirmed cases curve, producing a seroprevalence that does not depend on case reporting and that can be reliably inferred in epidemics where the IFR changes with time, as was the case in Manaus.

Estimating the infection fatality rate for December 2020

We estimate the infection fatality rate using total deaths due to severe acute respiratory infection (SARI) which include PCR- and clinically-confirmed SARS-CoV-2 infection as well as SARI deaths without a final diagnosis, excluding SARI deaths confirmedly caused by other aetiologies. This approach reduces under-reporting, particularly in 2020 when testing was not widely available. To estimate the IFR in 2020, we use the seroprevalence estimated by our model for December 16 2020, and select only SARI deaths with symptoms between March 1st and December 15, 2020. For the first wave of COVID-19 that occurred in the eight cities, we estimate the number of cases as the age-specific population size (https://demografiaufrn.net/laboratorios/lepp/) multiplied by the estimated seroprevalence in the corresponding age group. We propagate the uncertainty in the prevalence estimate through the calculation of IFR.

Let $\rho[a]$ and $pop[a]$ be the prevalence and the population estimated for age group a. We assume a uniform distribution in the interval $[0,1]$ as a non-informative prior for $IFR[a]$, and the number of deaths $D[a]$ observed for each age group a is Binomial-distributed with size $\lfloor \rho[a] \times pop[a] \rfloor$ (the number of infections) and probability $IFR[a]$. For each sample of $\rho[a]$, we draw a sample of the posterior distribution of $IFR[a]$, given by

$$Beta(1 + D[a], 1 + \lfloor \rho[a] \times pop[a] \rfloor - D[a])$$

and compute the median, interquartile ranges and 95% confidence intervals of the IFR by retrieving the quantiles of the posterior distribution. The same analysis was repeated to compute the Infection-Hospitalisation Rate (IHR), by using the number of hospitalisations with SARI instead of the number of deaths.

To validate the obtained IFRs, we also estimate the IFRs using the measured prevalence corrected only by the sensitivity and specificity of the assay, without explicitly accounting for seroreversion. In this validation analysis, we use a small threshold of 0.1 signal-to-cutoff to avoid underestimating the prevalence due to seroreversion (see Supplementary Information).

Estimating the infection fatality rate for the Gamma VOC

We estimate the IFR and the attack rate separately for the second, Gamma-dominant, SARS-CoV-2 wave that occurred in Manaus. The Gamma variant was first detected in Manaus in November 2020, and its prevalence among PCR-positive patients grew rapidly to 87.0% on January 4 2021[6]. For this reason, it is reasonable to assume that all infections in Manaus that occurred after December 15 are due to the Gamma VOC. The Gamma-dominated wave was characterised by a non-negligible proportion of reinfections[6,21,29]. It is estimated that 13.6–39.3% of the infections in the second wave of COVID-19 epidemic in Manaus were reinfections[21], which is explained by the higher in-vitro reinfection potential of Gamma[30] and partial immunity waning eight months after the first surge. Thus, to calculate the attack rate and IFR of the Gamma-dominated wave, reinfections must be considered.

However, estimating the incidence of reinfections among positive donors is not straightforward - as a positive result may be either primary infection or reinfection, and these cannot be distinguished using a single test result. For this reason, it was not possible to obtain a point estimate for the number of infections that happened in the second wave in Manaus. To overcome this problem, we calculate upper bounds for the attack rate of the Gamma-dominated wave in Manaus (that is, the incidence between December 2020 and March 2021), and conversely lower bounds for the IFR of the Gamma VOC.

We additionally computed the IFR obtained using the seroprevalence estimated by the model. It is worth noting that our seroreversion correction model only estimates the incidence among seronegative individuals, thus an S/C boosting due to

reinfection is not detected by our method. As such, our model estimates the seroprevalence assuming there are no reinfections among positive individuals, underestimating the size of the second wave in Manaus.

To limit the maximum number of reinfections allowed among positive donors, we estimate the incidence using a threshold of 1.4 signal-to-cutoff (the upper threshold recommended by the manufacturer) instead of 0.49 signal-to-cutoff (the lower threshold recommended by the manufacturer), and correct for sensitivity based on 163 true positives and 30 false negatives in the plasma donors cohort. Since the specificity of the test using a threshold of 1.4 is 99.9%, and since it is not straightforward to take the specificity into account when reinfections are allowed, we do not correct for specificity in this analysis.

We compute the monthly number of positive tests $T^+[n]$ from December 2020 to March 2021 for a given age-sex group, as well as number of true positive (TP) and false negatives (FN) from convalescent plasma donors and the number of False Positives (FP) and true negatives (TN) from the pre-pandemic blood donors cohort in Manaus (Supplementary Table 1). Then, we draw posterior samples from the raw estimated monthly incidence and the seroprevalence in December and correct for the sensitivity and specificity of the assay. The lower bound for the attack rate of the Gamma VOC is the estimated incidence between December 2020 and March 2021, and the upper bound is the incidence in this period added by the seroprevalence measured in December. Hence, the upper bound is obtained assuming that all individuals that were seropositive in December get reinfected or that all previously seropositive individuals seroreverted in March 2021. The lower bound for the IFR is then calculated using the upper bound of the attack rate and the number of deaths with symptoms onset between December 16 and March 15. This procedure is repeated for each age-sex group independently, and is summarised in **Algorithm 3** in **Supplementary Information**.

The IHR for the Gamma VOC was estimated using the same procedure, but using the number of hospitalisations by SARI instead the number of deaths.

Data availability

All serological data required to reproduce the analyses and the codes used for the main analyses are available at https://github.com/CADDE-CENTRE/seroprevalence_eight_cities.

Declarations

This project was approved by the Brazilian national research ethics committee, CONEP CAAE - 30178220.3.1001.0068. The Brazilian national research committee (CONEP) waived for informed consent. All methods were performed in accordance with relevant guidelines and regulations.

Acknowledgments

This work was supported by the Itaú Unibanco “Todos pela Saude” program and by CADDE/FAPESP (MR/S0195/1 and FAPESP 18/14389-0) (http://caddecentre.org/); Wellcome Trust and Royal Society Sir Henry Dale Fellowship 204311/Z/16/Z (N.R.F.); the Gates Foundation (INV- 034540 and INV-034652) the National Heart, Lung, and Blood Institute Recipient Epidemiology and Donor Evaluation Study (REDS, now in its fourth phase, REDS-IV-P) for providing the blood donor demographic and zip code data for analysis (grant HHSN268201100007I); and the UK Medical Research Council under a concordat with the UK Department for International Development and Community Jameel and the NIHR Health Protection Research Unit in Modelling Methodology. CAPJ was supported by FAPESP (2019/21858-0). CAPJ, VHN were supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. VHN was supported by CNPq (304714/2018-6). SCF is supported by FAPESP. FM is supported by PROGRAMA INOVA FIOCRUZ-CE/Funcap, Edital 01/2020 Number: FIO-0167-00065.01.00/20 SPU Nº 06531047/2020. MBN is supported by CPNq. RFOF is supported by JBS - Fazer o bem faz bem. OR is supported by Medical Research Council MR/V038109/1.

The Blood Center SARS-CoV-2 Prevalence group is also composed by Cláudia M. M. Abrahim, Martirene A. Silva, Fabíola S. A. Hanna, Adriana S. N. Ramos, Juqueline R. Cristal and Samara Alves. We also thank Robert Verity for his critical review of the paper and suggestions.

Contributions

CAPJ, LFB, ECS, NRF, TS conceived the study. CAPJ, LFB, CW, ECS, VHN, NRF wrote the first draft. TS, MKO, ALSJ, AMJ, AFC, CAN, EL, FM, FLVA, GTN, KAF, LAX, LMS, LD, LMBC, LKS, LAF, MAR, MBN, MIBV, MVS, MAEC, NAS, NMT, NF, PMI, RFOF, RE, SAS, SOGM, SCF, VP, VSB collected and pre-processed the data. CAPJ, RHMP, ICGM collected and analysed socio-economic data. CAPJ, LFB, NRF, VHN, ECS did the main analysis. ECS, VHN, CW, NRF, OR, CD, MCC, MPB supervised the work. All authors reviewed and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Materials & Correspondence

Correspondence and requests for materials should be addressed to Carlos Prete (carlos.prete@usp.br) and Ester Sabino (sabinoec@usp.br).

References

1. Li, S. L. *et al.* Higher risk of death from COVID-19 in low-income and non-White populations of São Paulo, Brazil. BMJ Glob Health **6**, (2021).
2. Hallal, P. C. *et al.* SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. Lancet Glob Health **8**, e1390–e1398 (2020).
3. Castro, M. C. *et al.* Spatiotemporal pattern of COVID-19 spread in Brazil. Science **372**, 821–826 (2021).
4. de Souza Santos, A. A. *et al.* Dataset on SARS-CoV-2 non-pharmaceutical interventions in Brazilian municipalities. Sci Data **8**, 73 (2021).
5. Buss, L. F. *et al.* Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. Science **371**, 288–292 (2021).
6. Faria, N. R. *et al.* Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science (2021) doi:10.1126/science.abh2644.
7. Martins, A. F. *et al.* Detection of SARS-CoV-2 lineage P.1 in patients from a region with exponentially increasing hospitalisation rate, February 2021, Rio Grande do Sul, Southern Brazil. Euro Surveill. **26**, (2021).
8. de Oliveira, M. H. S., Lippi, G. & Henry, B. M. Sudden rise in COVID-19 case fatality among young and middle-aged adults in the south of Brazil after identification of the novel B.1.1.28.1 (P.1) SARS-CoV-2 strain: analysis of data from the state of Parana. *bioRxiv* (2021) doi:10.1101/2021.03.24.21254046.
9. Brizzi, A. *et al.* Report 46: Factors driving extensive spatial and temporal fluctuations in COVID-19 fatality rates in Brazilian hospitals. *medRxiv* (2021) doi:10.1101/2021.11.01.21265731.
10. Crotty, S. Hybrid immunity. Science **372**, 1392–1393 (2021).
11. Stamatatos, L. *et al.* mRNA vaccination boosts cross-variant neutralizing antibodies elicited by SARS-CoV-2 infection. Science (2021) doi:10.1126/science.abg9175.
12. Reynolds, C. J. *et al.* Prior SARS-CoV-2 infection rescues B and T cell responses to variants after first vaccine dose. Science (2021) doi:10.1126/science.abh1282.
13. Stone, M. *et al.* Evaluation of commercially available high-throughput SARS-CoV-2 serological assays for serosurveillance and related applications. *bioRxiv* (2021) doi:10.1101/2021.09.04.21262414.

14. Di Germanio, C. *et al.* SARS-CoV-2 antibody persistence in COVID-19 convalescent plasma donors: Dependency on assay format and applicability to serosurveillance. Transfusion (2021) doi:10.1111/trf.16555.
15. Takahashi, S., Greenhouse, B. & Rodríguez-Barraquer, I. Are Seroprevalence Estimates for Severe Acute Respiratory Syndrome Coronavirus 2 Biased? J. Infect. Dis. **222**, 1772–1775 (2020).
16. Takahashi, S. *et al.* SARS-CoV-2 serology across scales: a framework for unbiased seroprevalence estimation incorporating antibody kinetics and epidemic recency. *medRxiv* (2021) doi:10.1101/2021.09.09.21263139.
17. Lumley, S. F. *et al.* The duration, dynamics and determinants of SARS-CoV-2 antibody responses in individual healthcare workers. Clin. Infect. Dis. (2021) doi:10.1093/cid/ciab004.
18. Brazeau, N., Verity, R., Jenks, S., Fu, H. & Whittaker, C. Report 34: COVID-19 infection fatality ratio: estimates from seroprevalence. (2020).
19. O'Driscoll, M. *et al.* Age-specific mortality and immunity patterns of SARS-CoV-2. Nature **590**, 140–145 (2021).
20. Naveca, F. G. *et al.* COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. Nat. Med. **27**, 1230–1238 (2021).
21. Prete, C. A., Jr *et al.* Reinfection by the SARS-CoV-2 Gamma variant in blood donors in Manaus, Brazil. BMC Infect. Dis. **22**, 127 (2022).
22. Mina, M. J. *et al.* A Global Immunological Observatory to meet a time of pandemics. Elife **9**, (2020).
23. Monod, M. *et al.* Age groups that sustain resurging COVID-19 epidemics in the United States. Science **371**, (2021).
24. Buss, L. F. & Sabino, E. C. Intense SARS-CoV-2 transmission among affluent Manaus residents preceded the second wave of the epidemic in Brazil. The Lancet. Global health vol. 9 e1475–e1476 (2021).
25. Lalwani, P. *et al.* High anti-SARS-CoV-2 antibody seroconversion rates before the second wave in Manaus, Brazil, and the protective effect of social behaviour measures: results from the prospective DETECTCoV-19 cohort. Lancet Glob Health **9**, e1508–e1516 (2021).
26. Banho, C. A. *et al.* Effects of SARS-CoV-2 P.1 introduction and the impact of COVID-19 vaccination on the epidemiological landscape of São José Do Rio Preto, Brazil. *bioRxiv* (2021) doi:10.1101/2021.07.28.21261228.
27. de Souza, F. S. H., Hojo-Souza, N. S., da Silva, C. M. & Guidoni, D. L. Second wave of COVID-19 in Brazil: younger at higher risk. Eur. J. Epidemiol. **36**, 441–443 (2021).
28. Shioda, K. *et al.* Estimating the cumulative incidence of SARS-CoV-2 infection and the infection fatality ratio in light of waning antibodies. Epidemiology **32**, 518–524 (2021).
29. Coutinho, R. M. *et al.* Model-based estimation of transmissibility and reinfection of SARS-CoV-2 P.1 variant. Commun Med **1**, (2021).
30. Lucas, C. *et al.* Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced immunity. Nature **600**, 523–529 (2021).

Figures

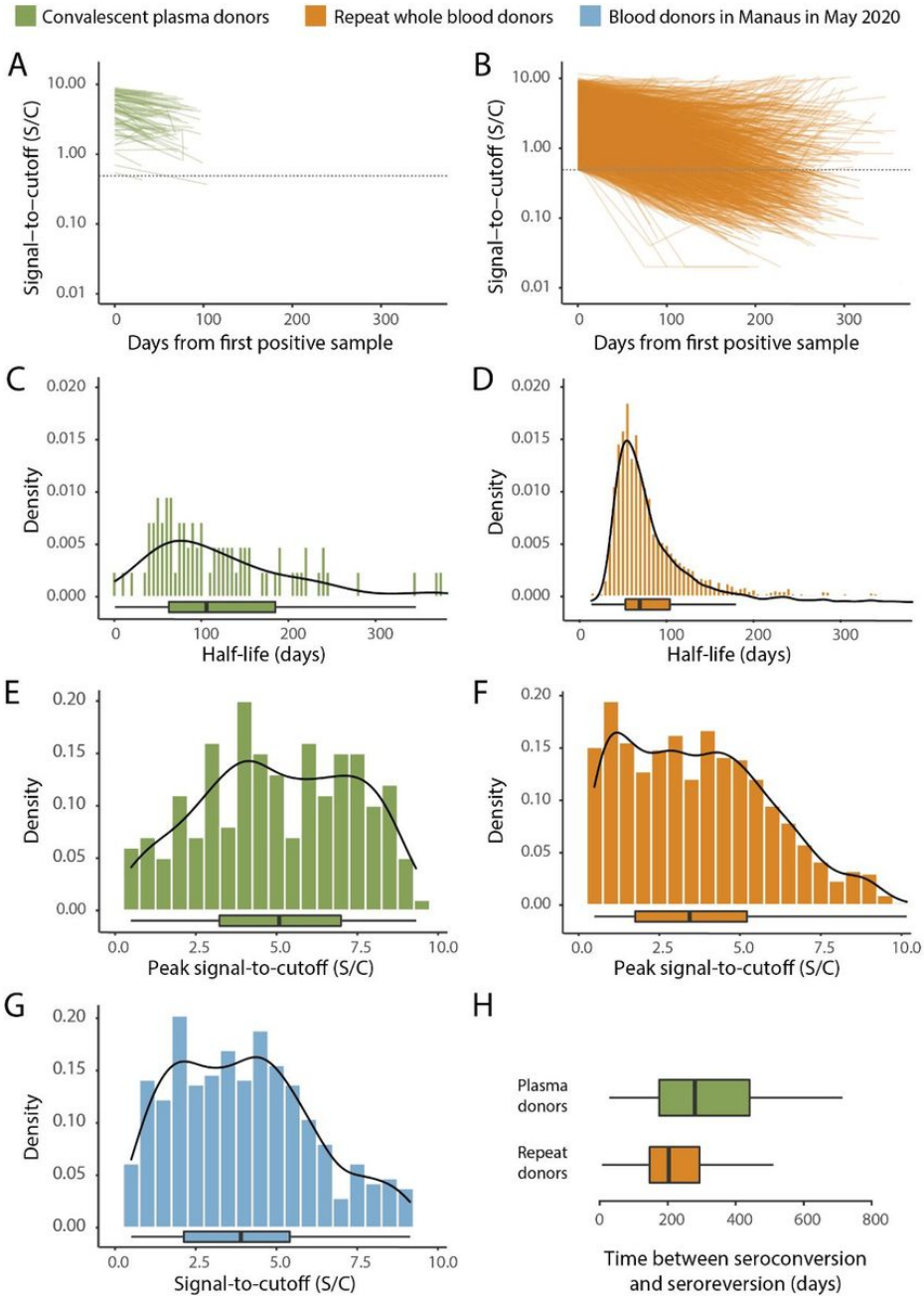

Figure 1

SARS-CoV-2 anti-N IgG dynamics in mild and moderate disease cohorts. (A) and (B): Trajectories of signal-to-cutoff (S/C) values for the Abbott anti-N CIMA in 218 SARS-CoV-2 infected convalescent plasma donors (A) and 7,675 repeat whole blood donors (B). Time is measured from the first positive test. (C) and (D): Probability distribution of the half-lives following infection in SARS-CoV-2 infected convalescent plasma donors (C) and repeat whole blood donors (D). Binned (bars) and smoothed kernel (lines) densities are shown. (E), (F) and (G): Comparison of the probability distribution of the highest signal-to-cutoff measured in plasma donors and positive repeat blood donors that donated before May 31, 2020 and the S/C distribution in Manaus in May 2020. (H): Estimated time between seroconversion and seroreversion (positive-negative conversion) at a threshold of 0.49 signal-to-cutoff for repeat blood donors and convalescent plasma donors. In all figures, box plots show the median (central lines), interquartile range (hinges), and range extending to 1.5 times the interquartile range from each hinge (whiskers).

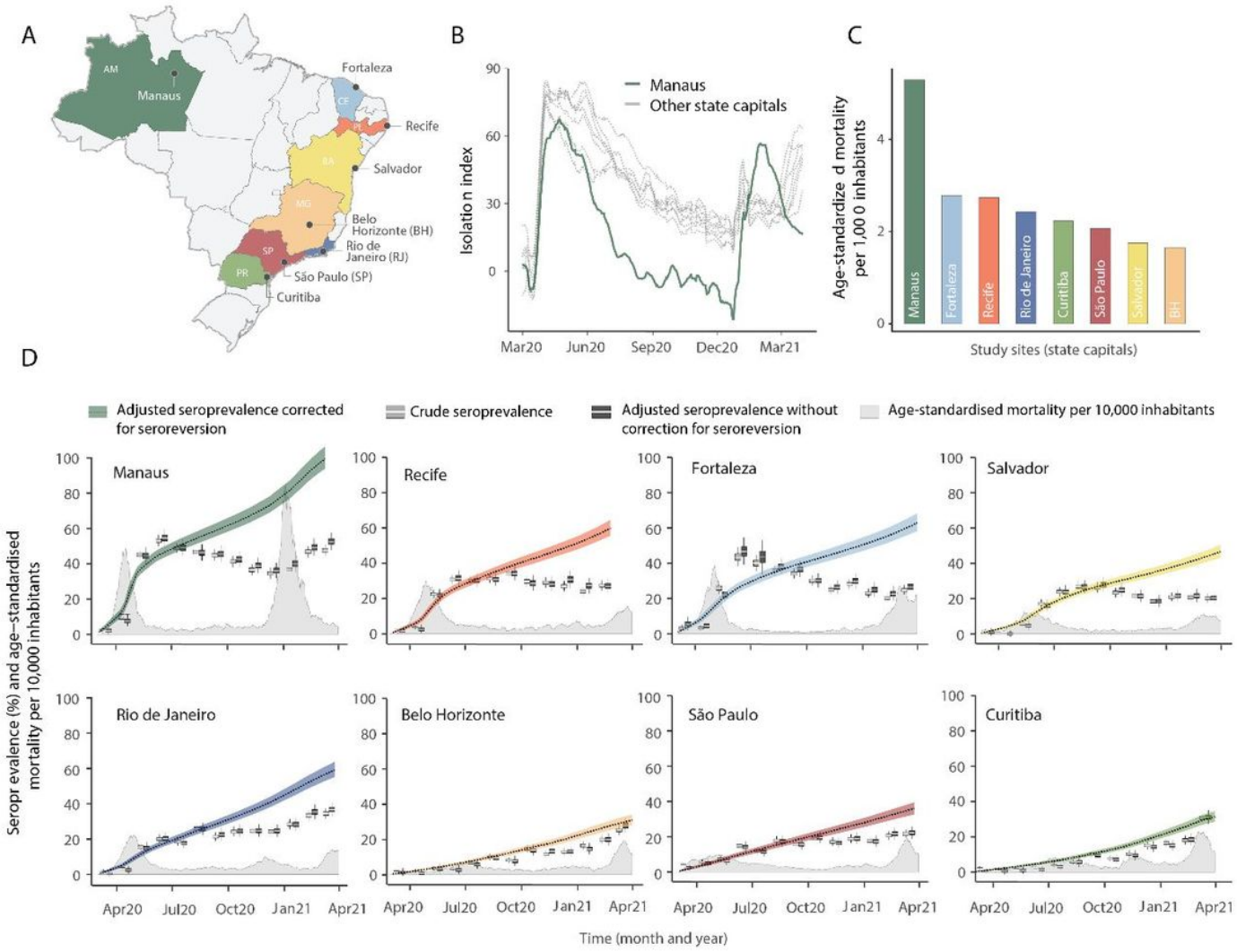

Figure 2

Overview of study site locations, mortality, and mobility data. (A) Map of the Brazilian states with the location of the eight capital cities. (B) Homestay index for the eight cities. Data were obtained from Fiocruz, available at https://bigdata-covid19.icict.fiocruz.br/. (C) Cumulative mortality due to severe acute respiratory syndrome (SIVEP-Gripe system) standardized for age and sex by the direct method using the total Brazilian age-sex structure as reference. Cumulative over the period from March 1st 2020 to March 31st 2021. (D) Weekly SARS-CoV-2 seroprevalence in blood donors across eight Brazilian state capitals. Three seroprevalence estimates are shown: (i) Crude seroprevalence (that is, the proportion of positive tests); (ii) Seroprevalence adjusted for sensitivity, specificity and reweighted by age and sex, but not corrected for seroreversion; (iii) Adjusted seroprevalence estimated by our seroreversion corrected model (continuous curves), which accounts for seroreversion in addition to sensitivity, specificity and age-sex distribution. Both infections and reinfections in seronegative donors are considered to estimate the adjusted seroprevalence, which can surpass 100%. The grey-filled curve shows age-standardised mortality per 10,000 residents. Ribbons and whiskers represent 95% Bayesian Credible Intervals.

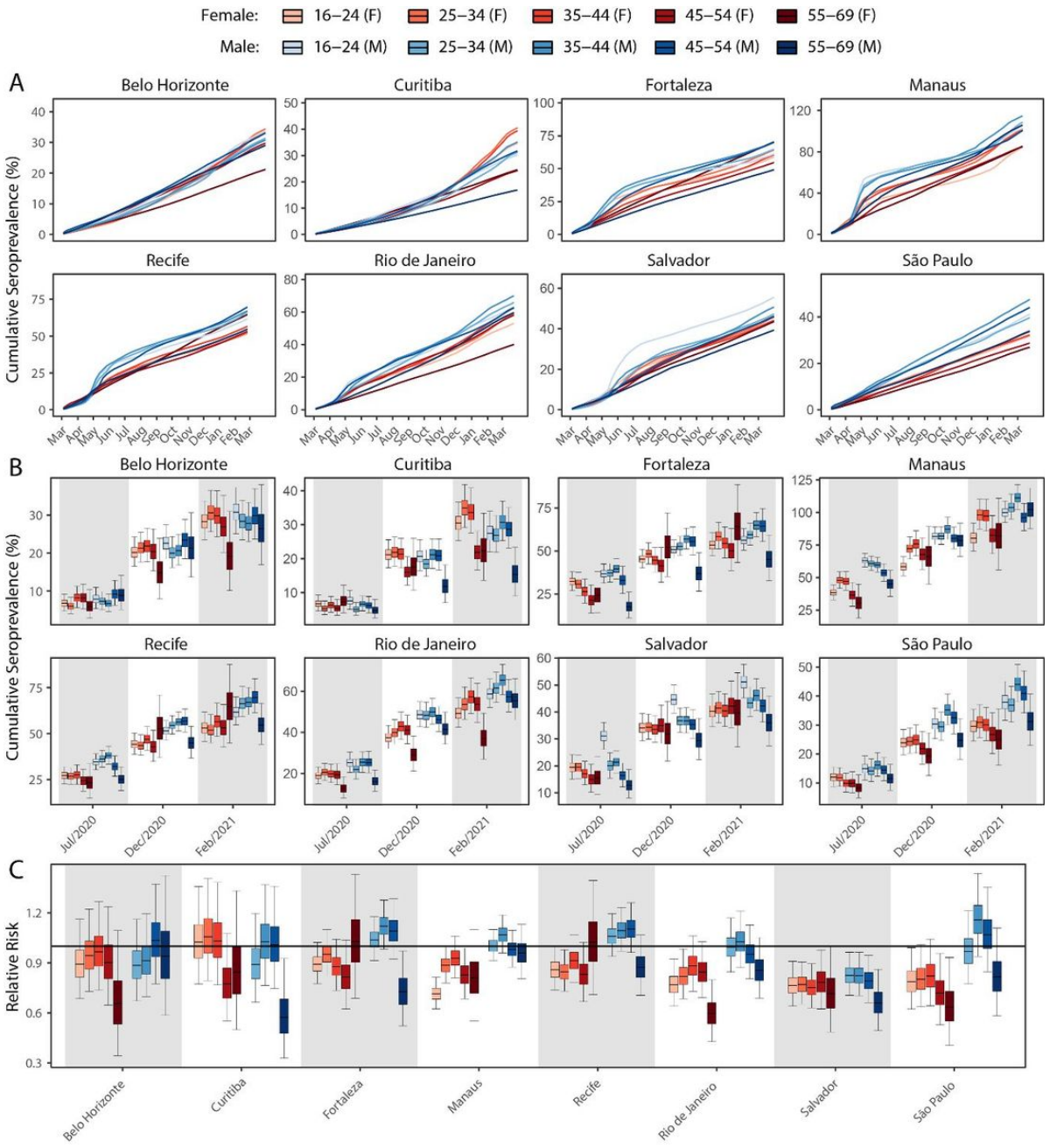

Figure 3

Age-sex patterns in blood donor seroprevalence in eight Brazilian cities. (A) Estimated cumulative seroprevalence by age-sex group. (B) Transversal cuts of figure (A) on July 8th, 2020, December 16th, 2020, and February 24th, 2021 (last week where seroprevalence was estimated for all cities). (C) Relative risk of the cumulative seroprevalence estimated in December 2020 with men aged 16-24 as the reference category in each city. Note that since cities use different values as reference, only relative risks of age-sex groups from the same city can be compared. Box plots show posterior distributions of the relative risks, with the median (central lines), interquartile range (hinges) and 95% confidence intervals (whiskers).

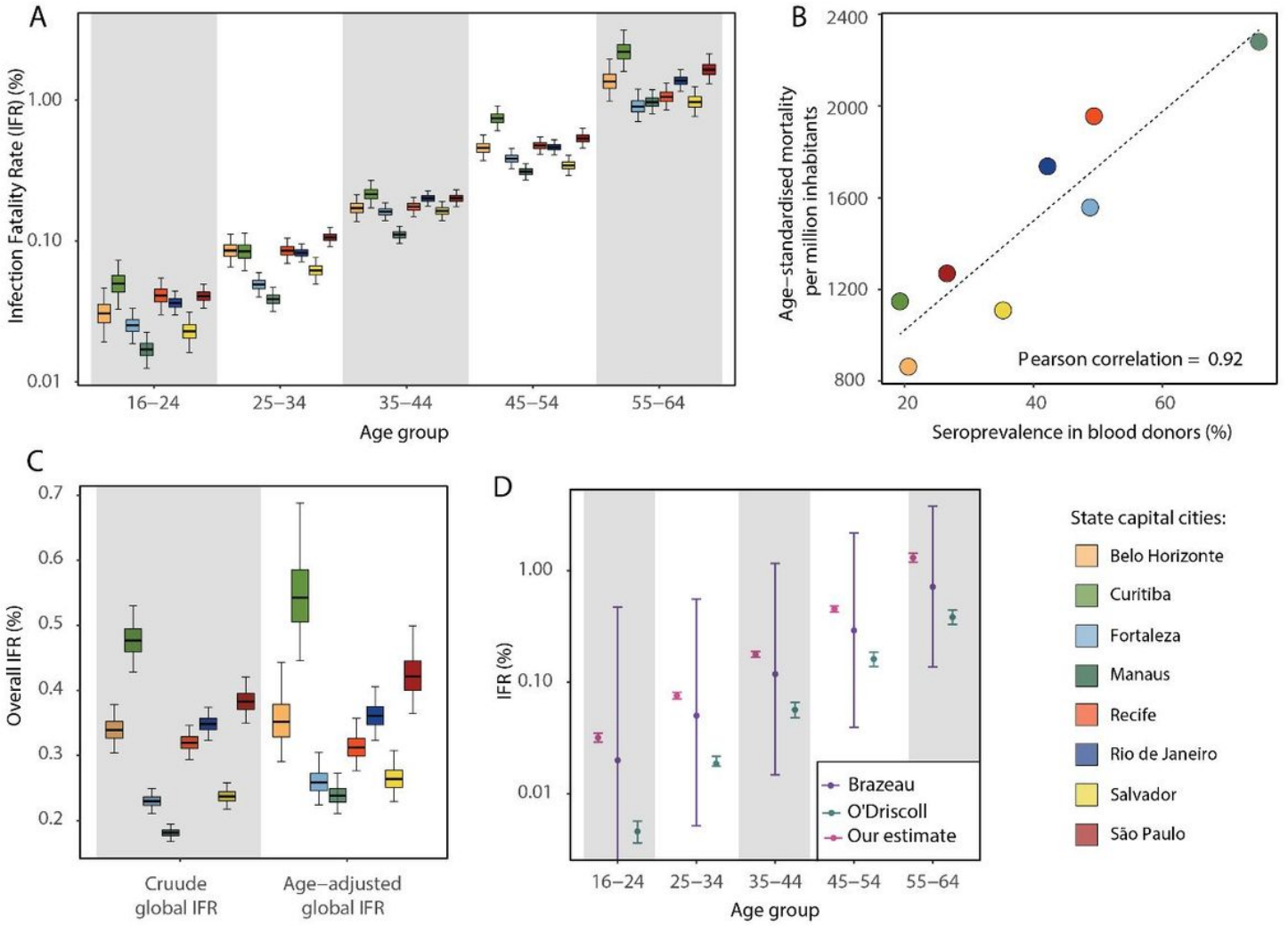

Figure 4

Infection fatality rates (IFR) in eight Brazilian state capitals as of December 15, 2020. The number of deaths was obtained from the SIVEP-Gripe reporting system including all SARI deaths between March 1st, 2020 and December 15th, 2021. (A) Age-specific IFRs (B) Overall IFRs, for the age range of 16-69yr, in each of the eight participating cities. (C) Association between age-standardised mortality rate and cumulative seroprevalence in blood donors for each of the eight cities by December 2020. The black line is a linear regression fit to the coloured points, each representing one of the eight cities. (D) Overall IFR of the eight cities obtained with our age-specific IFR estimates compared with the overall IFR calculated using age-specific IFRs from Brazeau et al[18] and O'Driscoll et al[19].

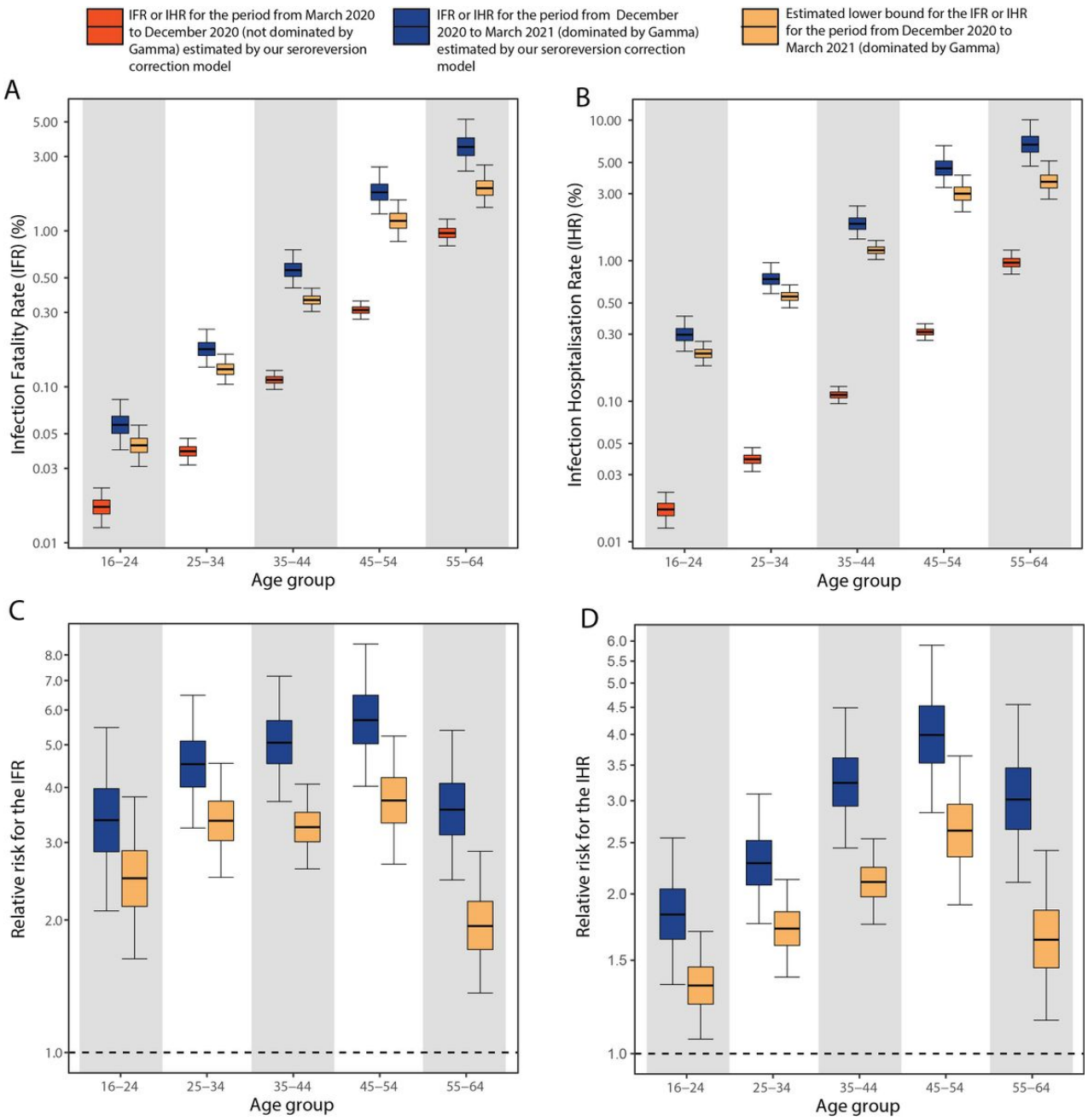

Figure 5

Comparison of infection-to-hospitalisation and infection-to-fatality rate estimates during the non-Gamma and Gamma period in Manaus. (A) Estimated infection-fatality rates and (B) infection-hospitalisation rates for Manaus in the periods from March 1 2020 to December 15 2020 (non-Gamma dominated) and December 16 2020 to March 31 2021 (Gamma dominated). For the Gamma dominated period, estimates shown are lower bounds that were calculated assuming a maximum proportion of reinfections (see Online Methods). (C) Relative Risks of the lower bound estimate of the infection-fatality rates in the Gamma dominated period using the estimated infection-fatality rates in the non-Gamma period as reference, and (D) similarly for infection-hospitalisation rates.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf