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

Supplementary Information

Validation of the obtained attack rates and IFRs

The seroprevalence and IFRs obtained in December 2020 estimated with our seroreversion correction model were validated using a smaller threshold of 0.1 and correcting only for sensitivity and specificity, without explicitly correcting for seroreversion (Supplementary Figure 25, Supplementary Table 2). Even though this approach underestimates the seroprevalence (thus overestimates the IFR) because a fraction of previously seropositive donors had already seroreverted by December 2020 (leading to a significant number of false negative test results), the obtained attack rates and IFRs were similar to the estimates of our model. The inferred seroprevalence for Manaus and Curitiba in December 2020 was respectively 61.0% (95% CrI 56.5% - 65.4%) and 13.4% (95% CrI 10.0% - 17.2%), compatible with the estimates of our seroreversion correction model (Figure 2, Table 1) given that these quantities underestimate the seroprevalence due to waning. The IFR pattern across cities was also similar in this analysis, being higher in Curitiba and smaller in Manaus for almost all age groups.

An alternative approach to estimating the attack rate in December 2020, in the face of waning antibodies and falling assay sensitivity, is to calculate the IFR early in the epidemic, prior to significant waning, and extrapolate the number of future cases from the reported death time series. To further validate the attack rates, we calculated the age-specific IFRs in June 2020 (when less seroreversion is expected) for the age range eligible to donate blood, and extrapolated using only the deaths within this age bracket. As such, the seroprevalence obtained for the other months are based solely on the number of deaths and the IFR inferred for June 2020 (Supplementary Figure 26). This approach led to an estimated seroprevalence of 90.8% (95% CrI 78.1% - 107.7%) and 10.9% (95% CrI 1.2% - 28.0%) in Manaus and Curitiba respectively in December 2020, which are compatible with our estimates if confidence intervals are considered. Nevertheless, this approach has the limitation of assuming a constant IFR through time, and only using a small amount of the total available serologic data.

To validate the inferred cumulative attack rate in November 2020 prior to the Gamma-dominated second wave and also in April 2021, following this second wave, we re-tested 996 samples from November 2020 in Manaus and tested 769 samples from April 2021 using the Abbott anti-S SARS-CoV-2 IgG CIMA assay (Supplementary Figure 27), which showed less waning than the Abbott anti-N assay used in this work[1]. As such, the usage of the anti-S assay reduced the difference between the seroprevalence obtained without explicitly correcting for seroreversion and the true seroprevalence. A sensitivity of 94.0% was obtained by testing convalescent plasma donors with this assay (Supplementary Table 1), and the specificity was assumed as 100%. The crude prevalence of anti-S antibodies was 56.7% (95% CrI 53.6% - 59.8%) in November 2020 and 78.7% (95% CrI 75.6% - 81.4%) in April 2021. After correcting for sensitivity and reweighting by age and sex, the seroprevalence estimate was 60.0% (95% CrI 58.4% - 62.2%) in November 2020 and 83.3% (95% CrI 81.1% - 86.4%) in April 2021, compared to 68.0% (95% CrI 64.2% - 72.7%) and 99.5% (95% CrI 94.0% - 106.6%) estimated using our seroreversion correction model for November 2020 and March 2021. Note that the attack rate estimated by our model considers both infections and reinfections among seronegative individuals, hence the confidence intervals higher than 100%. Of note, we measured the half-life of this assay using the serial repeat blood donors data available in Prete et al[2], obtaining a median half-life of 124.5 (IQR 74.7 - 258.0) days. In November, six months following the first wave in Manaus, some cases of seroreversion are expected to have occurred; as such, this remains an underestimate of the true cumulative attack rate by this point. Assuming no reinfections before November, the smaller seroprevalence measured with the anti-S assay suggests that 8.0% of previously infected donors seroreverted before November 2020.

Estimating the seroreversion probability from convalescent plasma donors

Unlike repeat blood donors, convalescent plasma donors have a known date of symptoms onset. To compute $p^+[n]$ for plasma donors, we estimate the instant of seroreversion for each plasma donor as in **Algorithm 1** and define the date of seroconversion as 8 days after the reported date of symptoms onset. This interval of 8 days is the average lag between seroconversion and seroreversion reported by Orner et al[3] for a threshold of 1.4, but it can be shorter for a threshold of 0.49 employed in this work. The probability mass function of the time to seroreversion $p^-_{day}[n]$ is then the empirical histogram of $\Delta t_i = {t_i}^- - {t_i}^+$, and $p^+[n]$ is obtained from $p^-_{day}[n]$ as in **Algorithm 1**.

Derivation of the expression of the probability of a positive test $\theta[n,a]$

To shorten the next equations, let us denote $T^+(n,a)$ the event of a test applied to an individual from age-sex group a at week n being positive, $I(n,a)$ the event of an individual from age-sex group a being infected at week n, $I(1{:}\,n,a) = \cup_{k=1}^{n} I(n,a)$ the event of an individual from group a having being infected before week n and $C(a)$ the event of an infected individual from group a seroconverting after infection. Also, we denote the negation of an event E as $\bar{E}$, and the probability of an event E as $P\{E\}$. We consider that the initial sensitivity se_0 (i.e., the sensitivity right after seroconversion) and specificity sp of the assay do not depend on the age-sex group or on time.

The probability $\theta[n,a]$ of a test applied to a person from age-sex group a being positive at week n is

$$\theta[n,a] = P\{T^+(n,a)\} = P\{T^+(n,a), I(1{:}\,n,a)\} + P\{T^+(n,a), \bar{\mathrm{I}}(1{:}\,n,a)\}.$$

The first term can be decomposed as

$$\begin{aligned} P\{T^+(n,a), I(1{:}\,n,a)\} &= \sum_{k=1}^{n} P\{T^+(n,a), I(1{:}\,n,a) | I(k,a)\} P\{I(1{:}\,n,a), I(k,a)\} \\ &= \sum_{k=1}^{n} P\{T^+(n,a) | I(k,a)\} P\{I(k,a)\}. \end{aligned}$$

The term $P\{I(k,a)\}$ is the incidence $u[k,a]$, and $P\{T^+(n,a)|I(k,a)\}$ can be further decomposed into

$$\begin{aligned} P\{T^+(n,a)|I(k,a)\} & \\ &= P\{T^+(n,a)|I(k,a), C(a)\} P\{C(a)|I(k,a)\} \\ &+ P\{T^+(n,a)|I(k,a), \bar{C}(a)\} P\{\bar{\mathrm{C}}(a)|I(k,a)\}. \end{aligned}$$

Assuming that an infected individual that did not seroconvert cannot have a positive test, we have $P\{T^+(n,a)|I(k,a), \bar{C}(a)\} = 0$. We approximate $P\{T^+(n,a)|I(k,a)\}$ to $p^+[n-k]$ (that is, the probability of a test being positive $n-k$ weeks after *seroconversion*), neglecting the delay between infection and seroconversion. Since the mean delay between infection and seroreversion is smaller than 8 days as explained above, and since seroprevalence data is discretized using weeks as time unit, this delay has small influence on seroprevalence estimates. Because $P\{C(a)|I(k,a)\}$ is the sensitivity se_0 of the assay and $P\{I(n,a)\}$ is the incidence $u[k,a]$, we have $P\{T^+(n,a), I(1{:}\,n,a)\} = se_0 \sum_{k=1}^{n} p^+[n-k]u[k,a]$.

The second term of $\theta[n,a]$ is

$$P\{T^+(n,a), \underline{I}(1:n,a)\} = P\{T^+(n,a)|\underline{I}(1:n,a)\}P\{\underline{I}(1:n,a)\} = (1-sp)\left(1-\sum_{k=1}^{n} u[k,a]\right),$$

where $sp = P\{T^+(n,a)|\underline{I}(1:n,a)\}$ is the specificity of the assay, which does not change over time.

Therefore, a simpler expression for $\theta[n,a]$ is obtained:

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

Description of the method used to validate the seroprevalence and IFR for 2020

To validate the seroprevalence and IFRs estimated for 2020, we recalculate these quantities by measuring the prevalence in December 2020 with a smaller threshold equal to 0.1 to partially account for seroreversion and correct for sensitivity and specificity, without explicitly incorporating a method to correct for seroreversion (Supplementary Figure 25).

Let TP, FN, FP and TN be respectively the number of true positives, false negatives, false positives and true negatives obtained from plasma donors and the pre-pandemic cohort in Manaus using a threshold 0.1. We use a uniform distribution in the interval $[0,1]$ as prior for the seroprevalence of age-sex group a, and also for the sensitivity se and specificity sp. The posterior distribution of the sensitivity and specificity are respectively $Beta(1+TP, 1+FN)$ and $Beta(1+TN, 1+FP)$. The seroprevalence $\rho[a]$ of age-sex group a is distributed according to a Binomial distribution of size $T[a]$ (the number of tests for this age-sex group) and probability $se \times \rho[a] + (1-sp)(1-\rho[a])$. To draw a posterior sample of $\rho[a]$, we draw a posterior sample of se and sp and from the auxiliary variable $Z[a] \sim Beta(T^+[a]+1, T[a]-T^+[a]+1)$, which represents the raw measured prevalence. Then, we compute the prevalence adjusted by sensitivity and specificity through $\rho[a] = \left(0, \frac{Z[a]+sp-1}{se+sp-1}\right)$. Finally, a sample of the IFR is then drawn from $Beta(1+D[a], 1+\lfloor pop[a] \times \rho[a] \rfloor - D[a])$.

Calculation of the homestay index

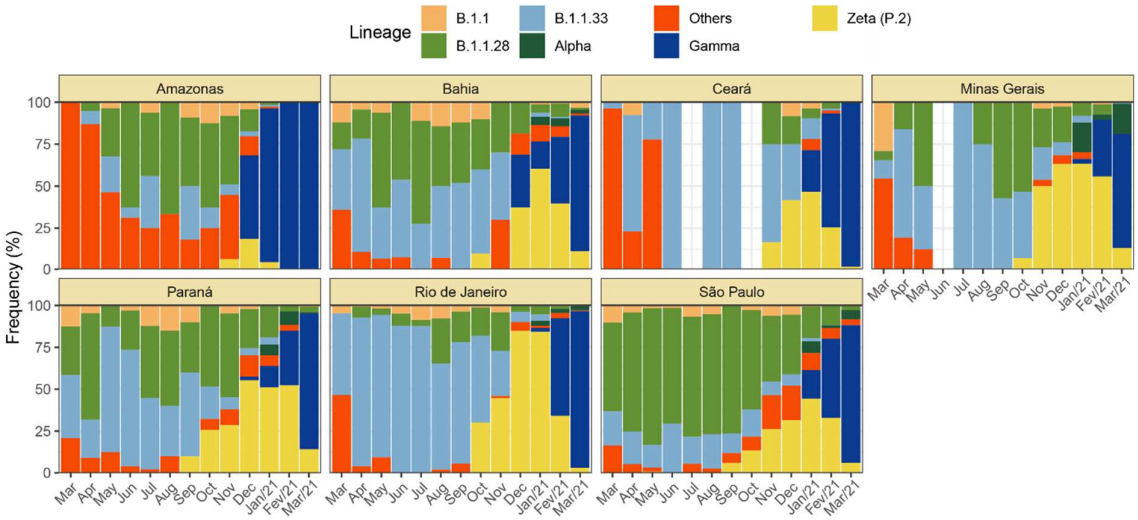
The homestay index for the eight cities shown in Figure 2B was extracted from https://bigdata-covid19.icict.fiocruz.br/. It was calculated using data from Google Mobility reports using the procedure described in[4]. The homestay index is defined as

$$\text{Homestay Index} = X_H - \frac{X_G + X_P + X_T + X_R + X_W}{5},$$

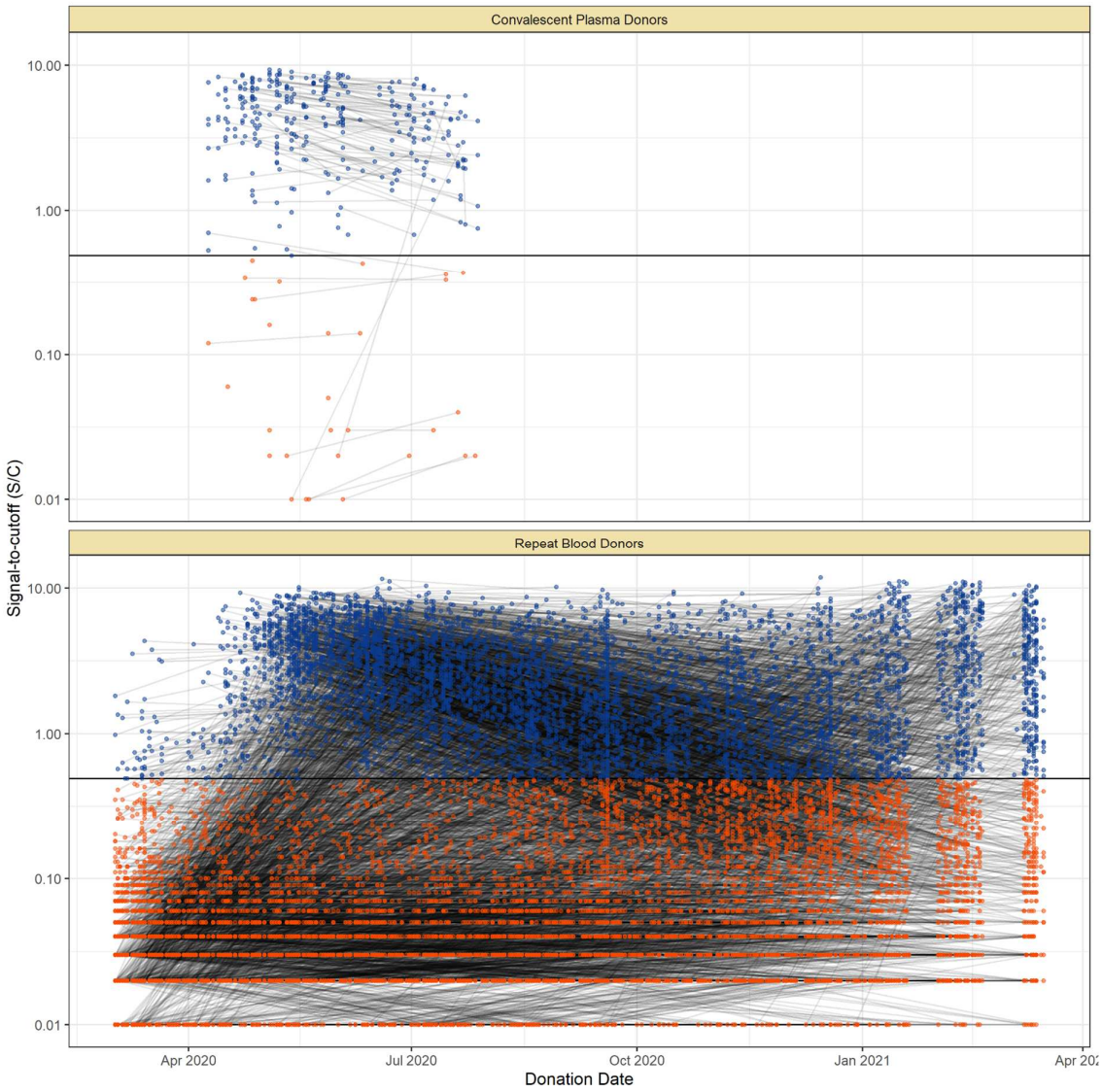
where X_H, X_G, X_P, X_T, X_R and X_W are respectively the variation of mobility (using pre-pandemic mobility levels as baseline) in the following place categories: Residential areas, grocery and pharmacy, parks, transit stations, retail and recreation and workplaces.

Calculation of age-standardised estimates

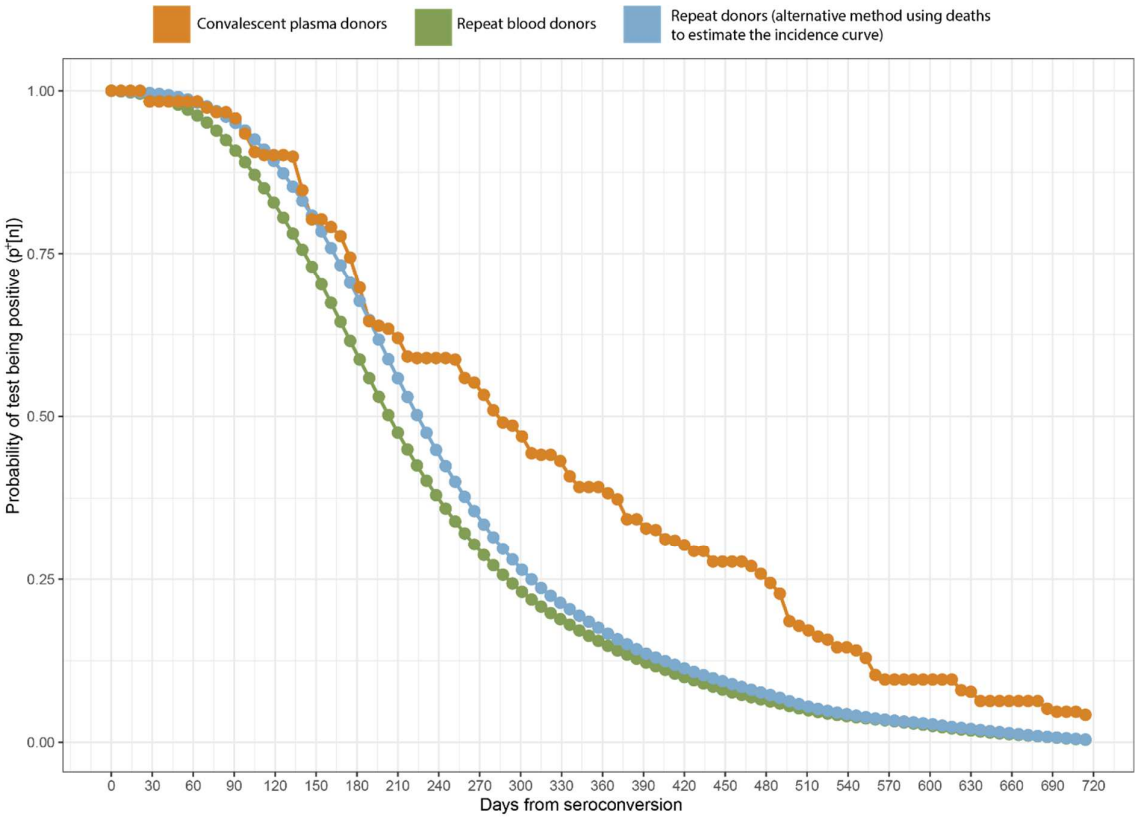
In this work, we calculated the age-standardised mortality, the age-standardised overall IFR and the age-standardised overall IHR. The procedure used to perform age-standardisation was the same for all these quantities. We define an age-standardised variable as the estimate that would be obtained if all cities had the same age structure. Denoting as $\eta[a]$ an age-specific IFR or IHR for a given city and $w[a]$ as the proportion of the combined population of all eight cities belonging to age group a, then the age-standardised overall IFR or IHR is $\sum_{a=1}^{M} w[a]\eta[a]$. Similarly, denoting as $\mu(t,a)$ the mortality for age-group a and day t for a given city, the age-standardised mortality is $\sum_{a=1}^{M} w[a]\mu(t,a)$.



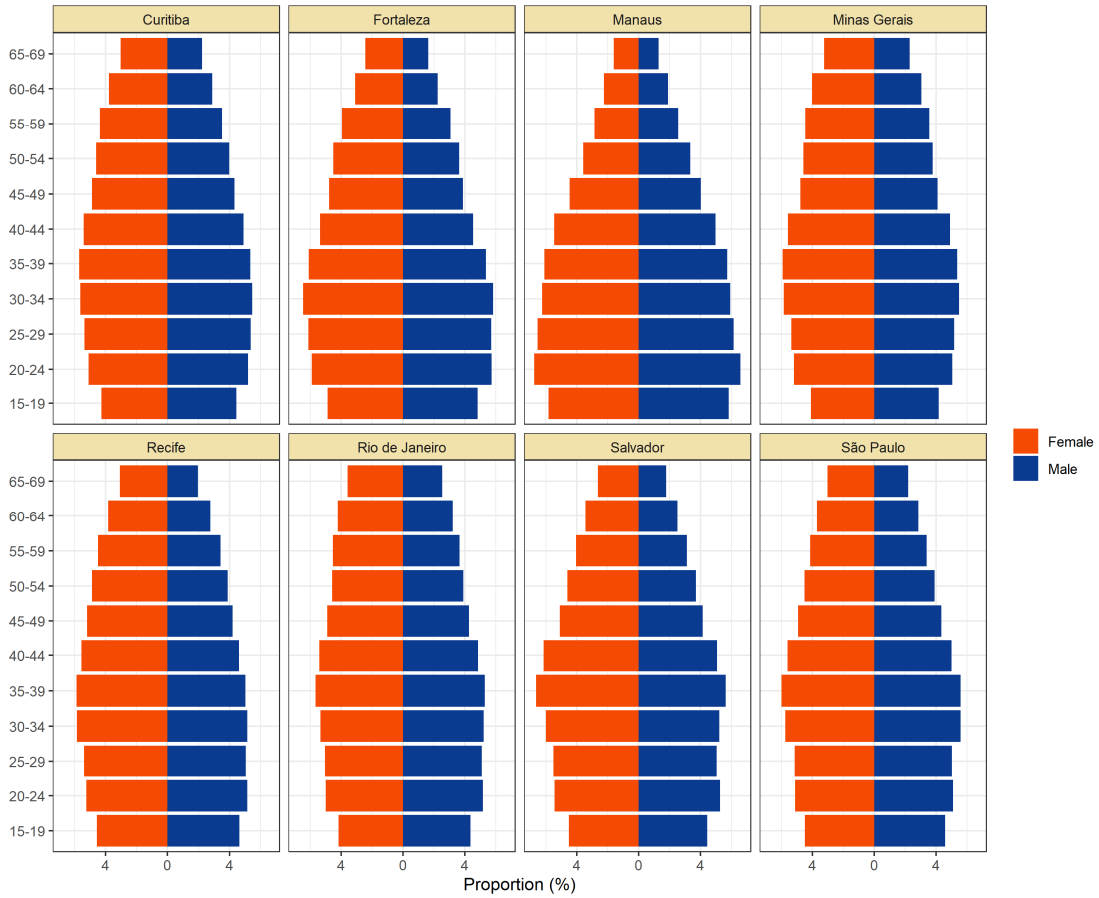
Supplementary Figure 1 - Monthly frequency of each lineage among cases confirmed by PCR for each state from March 2020 to March 2021. Data was extracted from http://www.genomahcov.fiocruz.br/. Lineage data was not available for the state of Pernambuco.



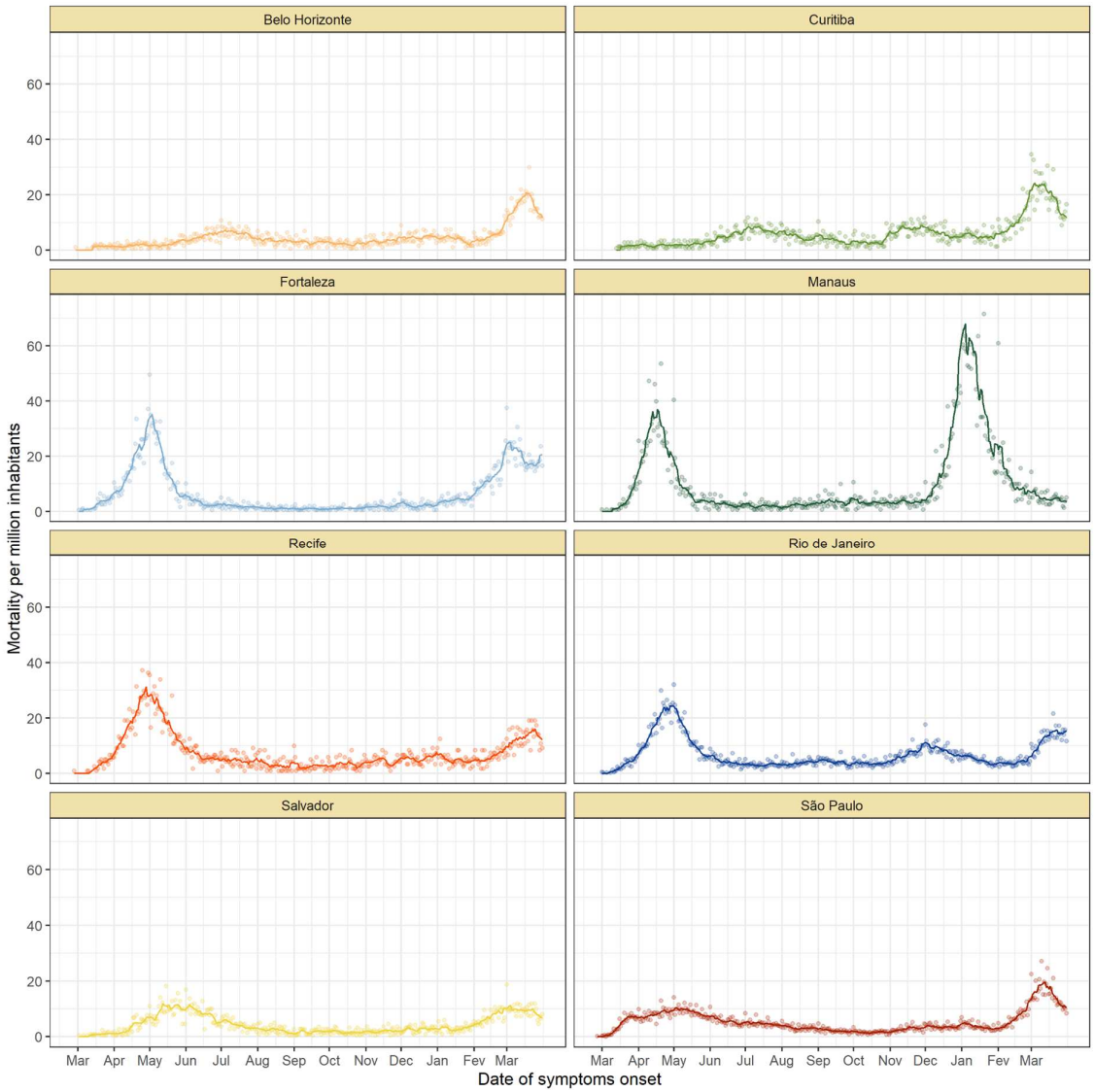
Supplementary Figure 2 - Serial donations of 218 convalescent plasma donors (symptomatic and known date of onset, PCR-positive) and 7,675 repeat whole blood donors (unknown if symptomatic or date of onset, unknown PCR status) included in this study. These cohorts were used to determine the rate of antibody waning and time-to-seroreversion distributions for the anti-N CIMA.



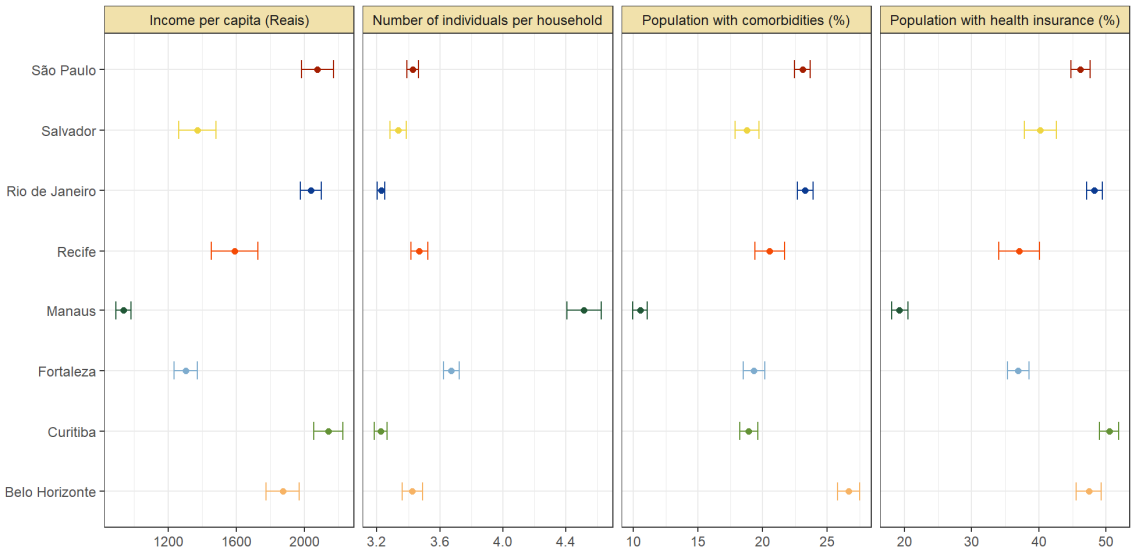
Supplementary Figure 3 - Probability of a test being positive a given number of weeks after seroconversion given that seroconversion occurred ($p^+[n]$) in terms of the method used to estimate it (see Methods). In this work, we use the estimate of $p^+[n]$ obtained from repeat blood donors (green curve).



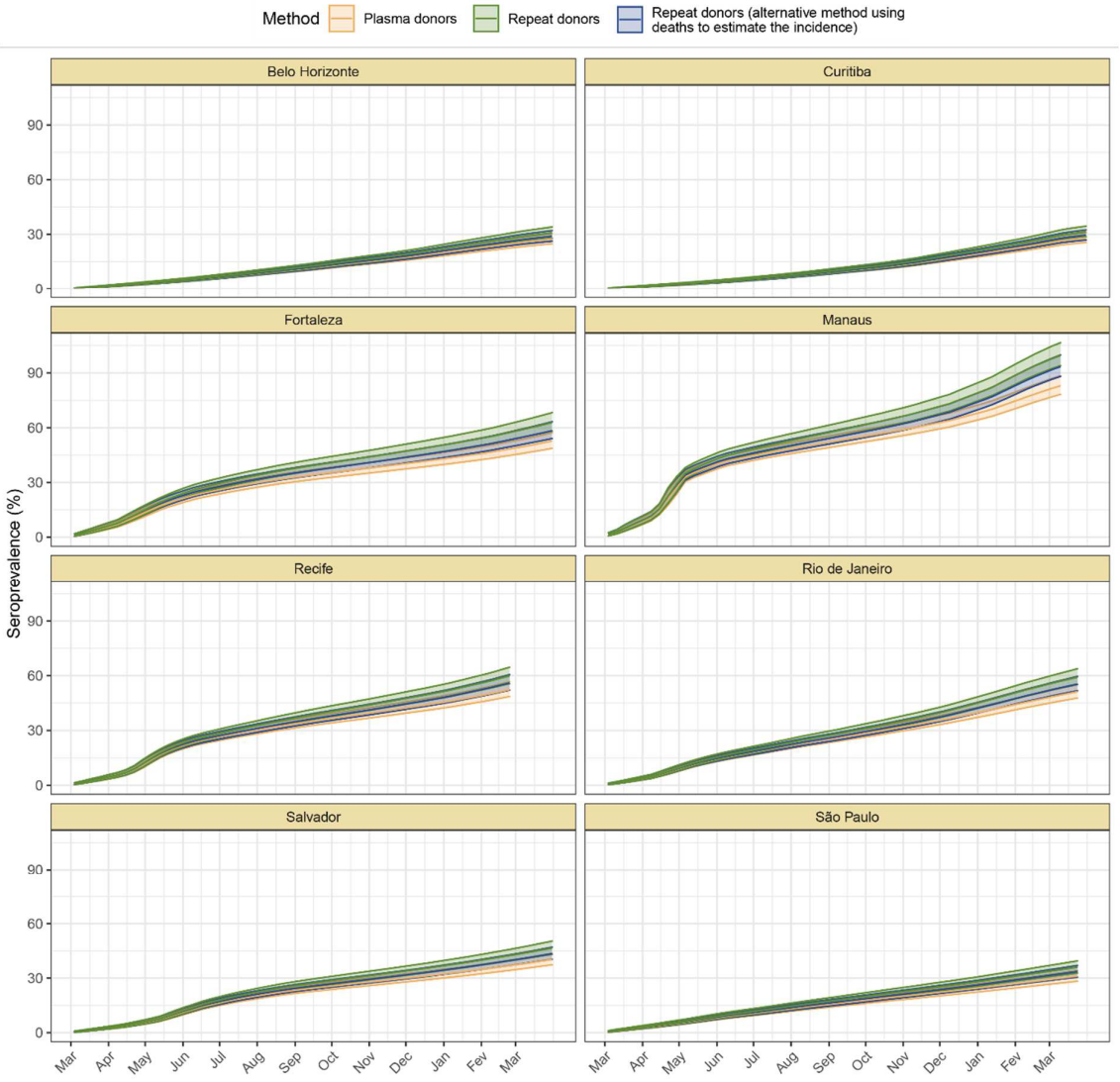
Supplementary Figure 4 - Population pyramids for the eight cities obtained from the projected population estimates for 2020 (extracted from https://demografiaufrn.net/laboratorios/lepp/). The proportions are relative to the population between 15 and 69.



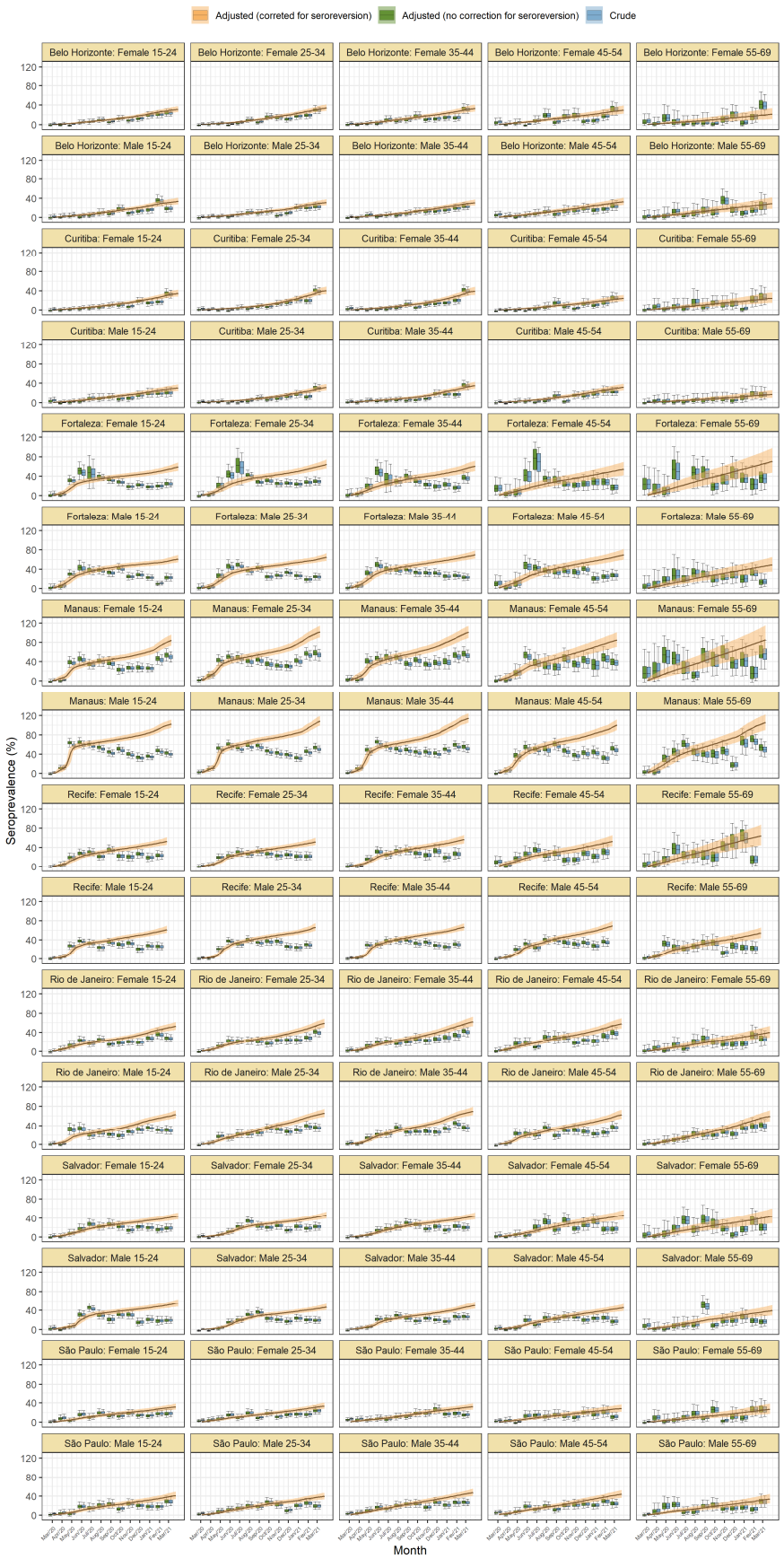
Supplementary Figure 5 - Mortality by SARI per million inhabitants in each of the eight cities without any age standardisation.



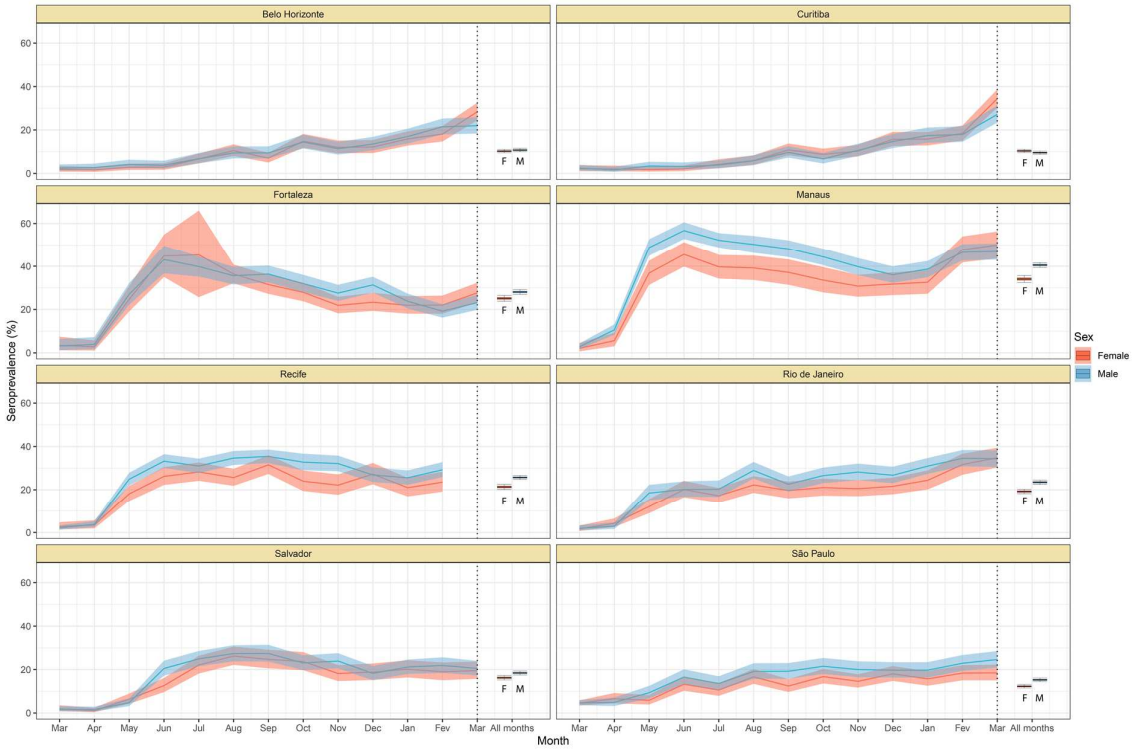
Supplementary Figure 6 - Proportion of population with health insurance, any comorbidity, number of individuals per household and income per capita for each municipality. The Information on health insurance, incidence of comorbidities, number of individuals per household as well as household income per capita was retrieved from the National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios (PNAD) COVID-19), a national telephone survey conducted by IBGE with over 1 888 560 interviews between May and September 2020. The data on comorbidities included chronic obstructive pulmonary disease, diabetes, hypertension or cardiovascular disease such as myocardial infarction, angina or heart failure)**.**



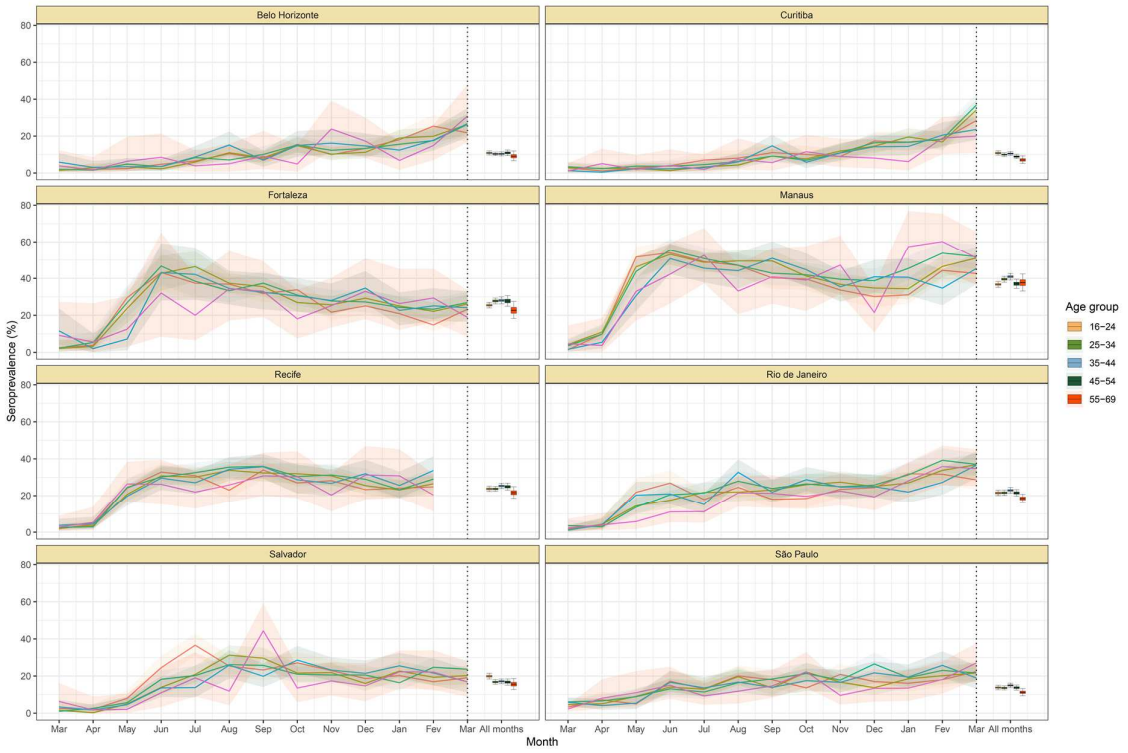
Supplementary Figure 7 - Cumulative seroprevalence estimated with 95% confidence intervals (ribbons) using three different methods to calculate the time-to-seroreversion distribution (see Methods). For the analyses in this paper, we consider the seroprevalence curve in green, obtained using repeat donors to estimate the time to seroreversion.


Adjusted (correted for seroreversion)
Adjusted (no correction for seroreversion)
Crude
Belo Horizonte: Female 15-24
Belo Horizonte: Female 25-34
Belo Horizonte: Female 35-44
Belo Horizonte: Female 45-54
Belo Horizonte: Female 55-69
Belo Horizonte: Male 15-24
Belo Horizonte: Male 25-34
Belo Horizonte: Male 35-44
Belo Horizonte: Male 45-54
Belo Horizonte: Male 55-69
Curitiba: Female 15-24
Curitiba: Female 25-34
Curitiba: Female 35-44
Curitiba: Female 45-54
Curitiba: Female 55-69
Curitiba: Male 15-24
Curitiba: Male 25-34
Curitiba: Male 35-44
Curitiba: Male 45-54
Curitiba: Male 55-69
Fortaleza: Female 15-24
Fortaleza: Female 25-34
Fortaleza: Female 35-44
Fortaleza: Female 45-54
Fortaleza: Female 55-69
Fortaleza: Male 15-24
Fortaleza: Male 25-34
Fortaleza: Male 35-44
Fortaleza: Male 45-54
Fortaleza: Male 55-69
Manaus: Female 15-24
Manaus: Female 25-34
Manaus: Female 35-44
Manaus: Female 45-54
Manaus: Female 55-69
Manaus: Male 15-24
Manaus: Male 25-34
Manaus: Male 35-44
Manaus: Male 45-54
Manaus: Male 55-69
Recife: Female 15-24
Recife: Female 25-34
Recife: Female 35-44
Recife: Female 45-54
Recife: Female 55-69
Recife: Male 15-24
Recife: Male 25-34
Recife: Male 35-44
Recife: Male 45-54
Recife: Male 55-69
Rio de Janeiro: Female 15-24
Rio de Janeiro: Female 25-34
Rio de Janeiro: Female 35-44
Rio de Janeiro: Female 45-54
Rio de Janeiro: Female 55-69
Rio de Janeiro: Male 15-24
Rio de Janeiro: Male 25-34
Rio de Janeiro: Male 35-44
Rio de Janeiro: Male 45-54
Rio de Janeiro: Male 55-69
Salvador: Female 15-24
Salvador: Female 25-34
Salvador: Female 35-44
Salvador: Female 45-54
Salvador: Female 55-69
Salvador: Male 15-24
Salvador: Male 25-34
Salvador: Male 35-44
Salvador: Male 45-54
Salvador: Male 55-69
São Paulo: Female 15-24
São Paulo: Female 25-34
São Paulo: Female 35-44
São Paulo: Female 45-54
São Paulo: Female 55-69
São Paulo: Male 15-24
São Paulo: Male 25-34
São Paulo: Male 35-44
São Paulo: Male 45-54
São Paulo: Male 55-69
Seroprevalence (%)
0
40
80
120
Month

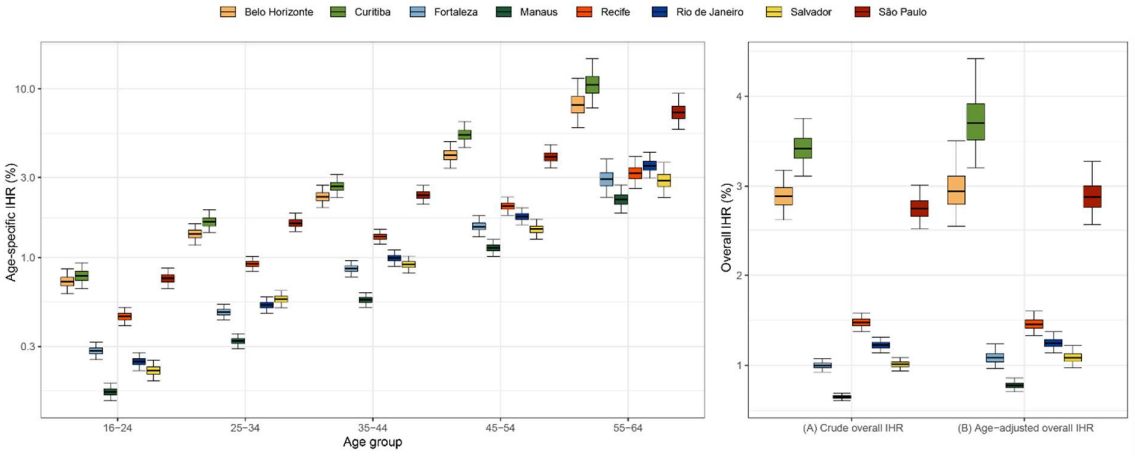
Supplementary Figure 8 – Seroprevalence estimates disaggregated by age and sex for the eight cities. As in Figure 2D, three seroprevalence estimates are shown: (i) Crude seroprevalence; (ii) Seroprevalence adjusted for sensitivity, specificity and reweighted by age and sex, but not corrected for seroreversion; (iii) Adjusted seroprevalence estimated by our seroreversion corrected model (continuous curves), which accounts for seroreversion in addition to sensitivity, specificity and age-sex distribution. This seroprevalence estimate consists in the sum of infections and reinfections among seronegative donors, hence it can be larger than 100%.



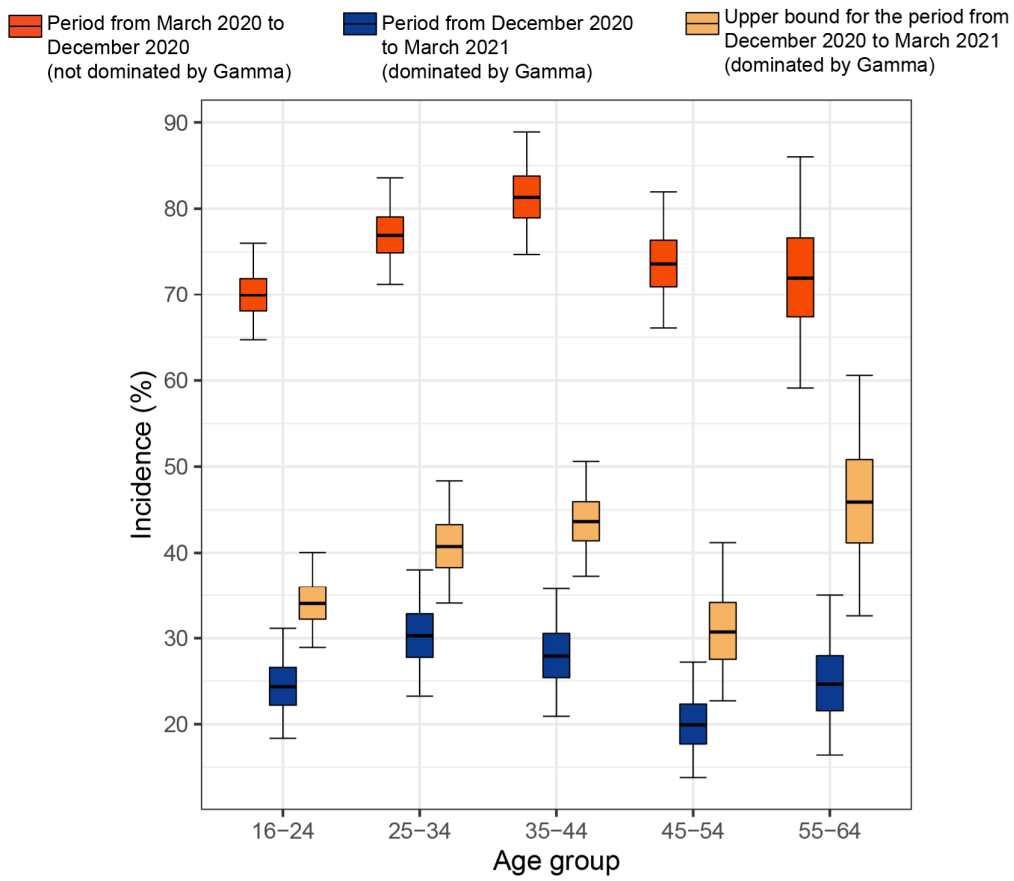
Supplementary Figure 9 – Crude monthly seroprevalence disaggregated by sex. Boxplots represent the crude seroprevalence obtained by aggregating all months, defined as the proportion of positive tests considering all months.



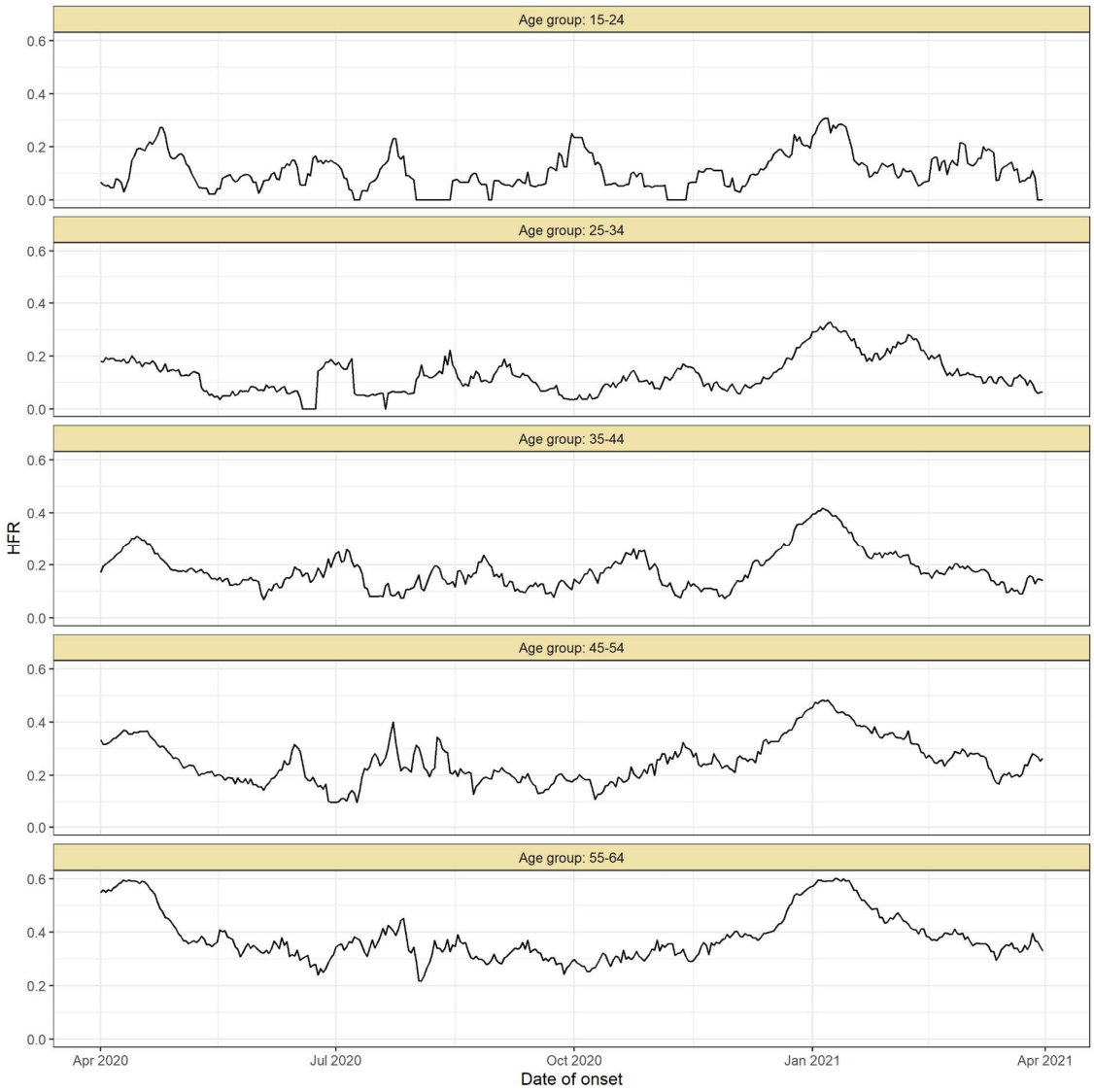
Supplementary Figure 10 - Crude monthly seroprevalence disaggregated by age. Boxplots represent the crude seroprevalence obtained by aggregating all months, defined as the proportion of positive tests considering all months.



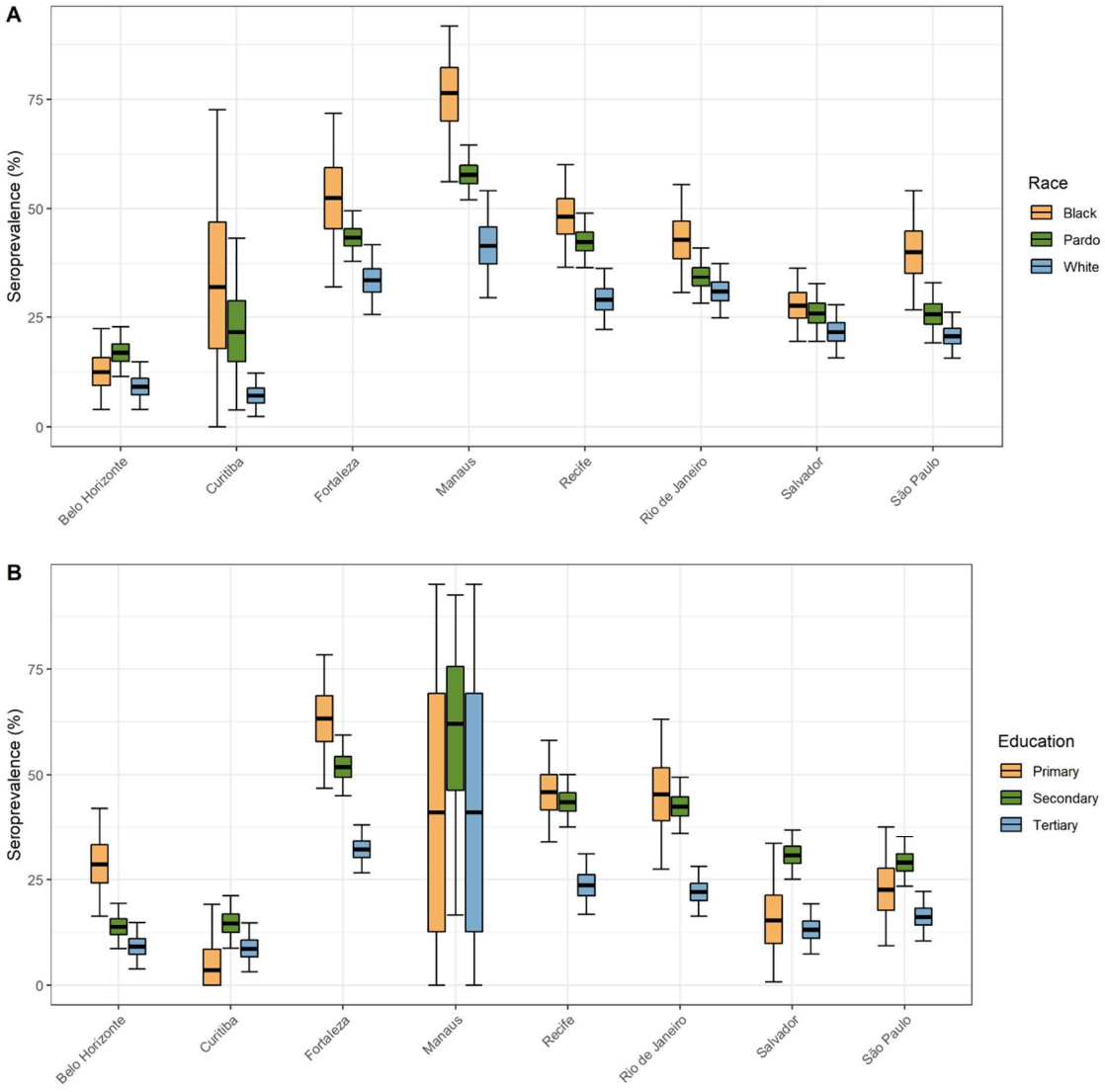
Supplementary Figure 11 - Estimated age-specific and overall Infection Hospitalisation Rates (IHRs) for the eight cities in the period between March and December 2020. Note that IHR depends not only on disease severity, but also on access to healthcare and availability of hospital resources, which can vary across cities.



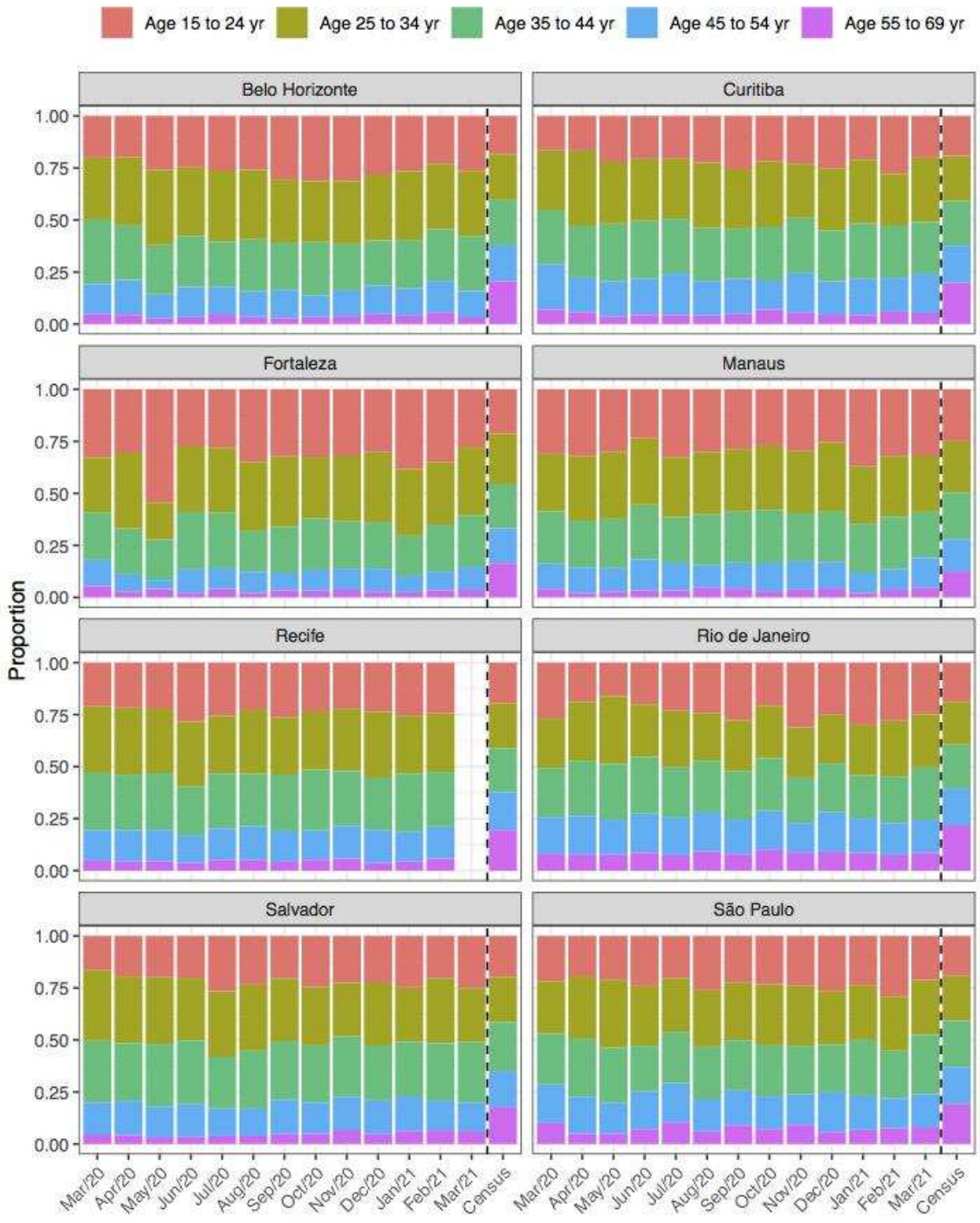
Extended Data Figure 12 - Incidence in Manaus estimated by our seroreversion correction model for the periods March 2020 – December 2020, December 2020 – March 2021 and the estimated upper bound for the incidence (see Methods).



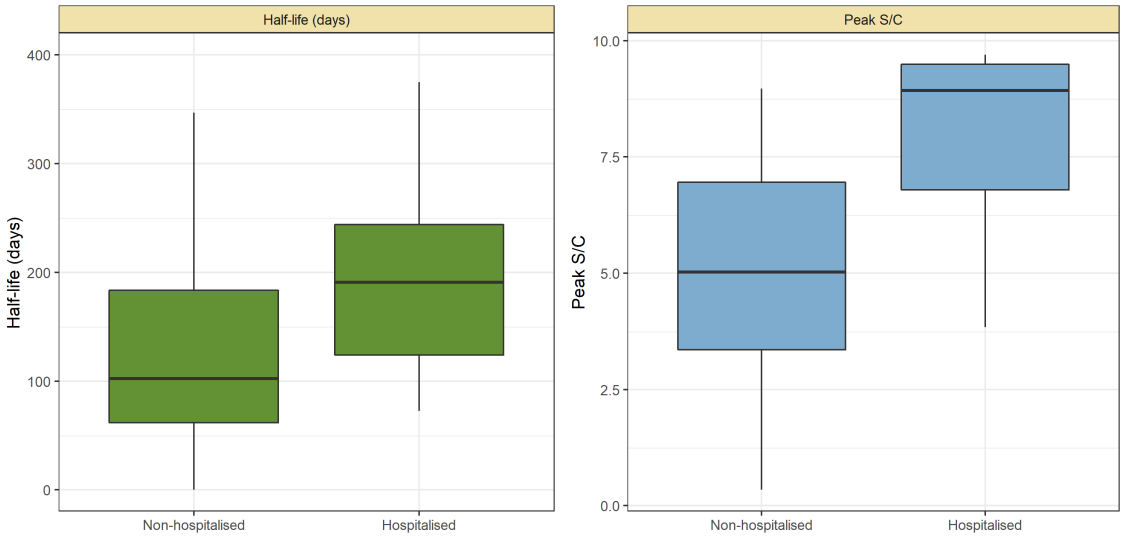
Supplementary Figure 13 - Estimated in-hospital fatality rate (HFR) over time in Manaus. The HFR was calculated by dividing the number of deaths by the number of hospitalisations using a moving window of 14 days. The HFR had a peak in the Gamma-dominated wave in January 2021 in all age groups.



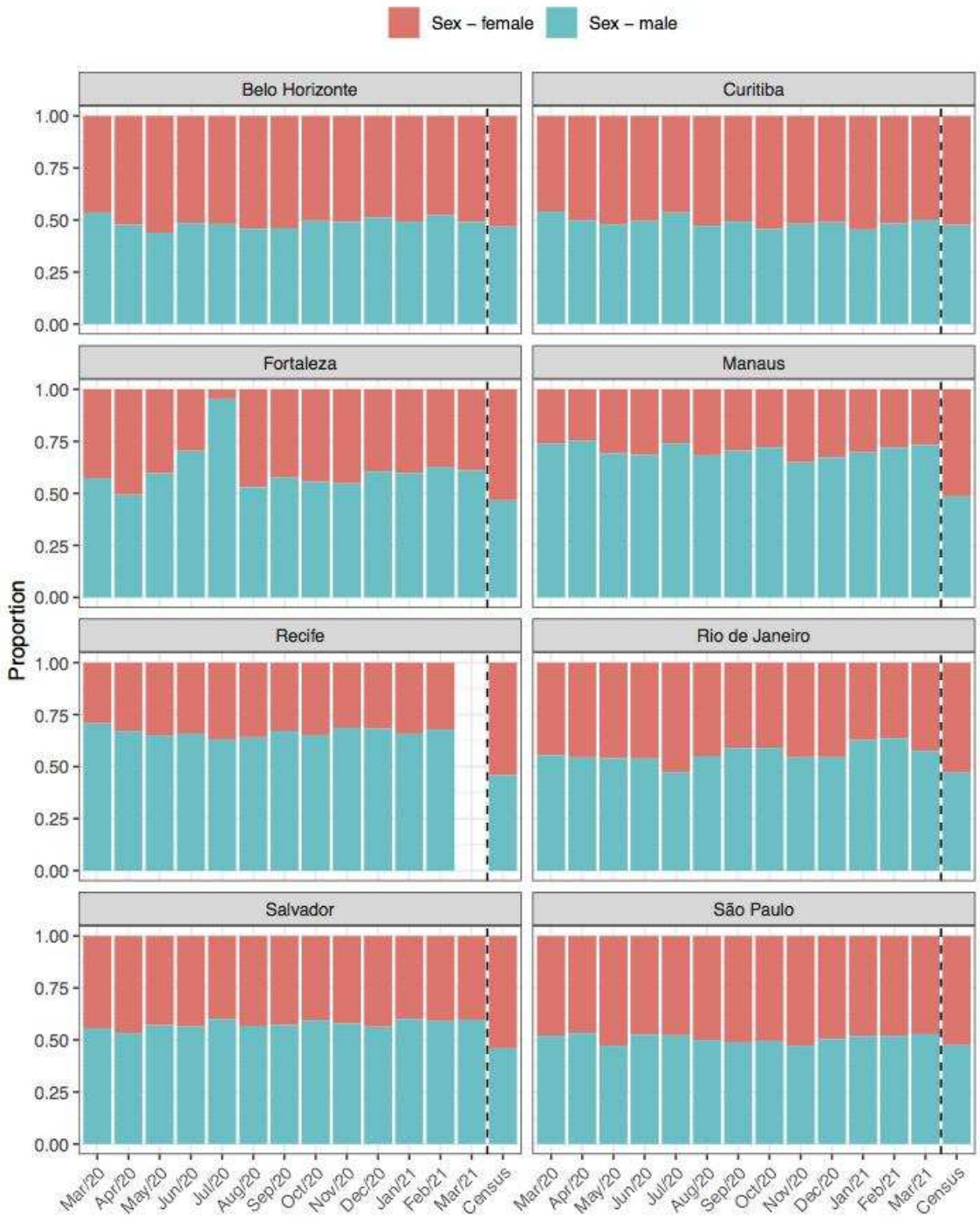
Supplementary Figure 14 - Seroprevalence in December 2020 disaggregated by race and education level. The seroprevalence was estimated using a threshold of 0.1 signal-to-cutoff and correcting for sensitivity and specificity, without any explicit correction for seroreversion. To aid visualization, we merged the White and East Asian races and discarded Indigeneous individuals. Information on education level is available for very few donors in Manaus, hence the large confidence intervals.



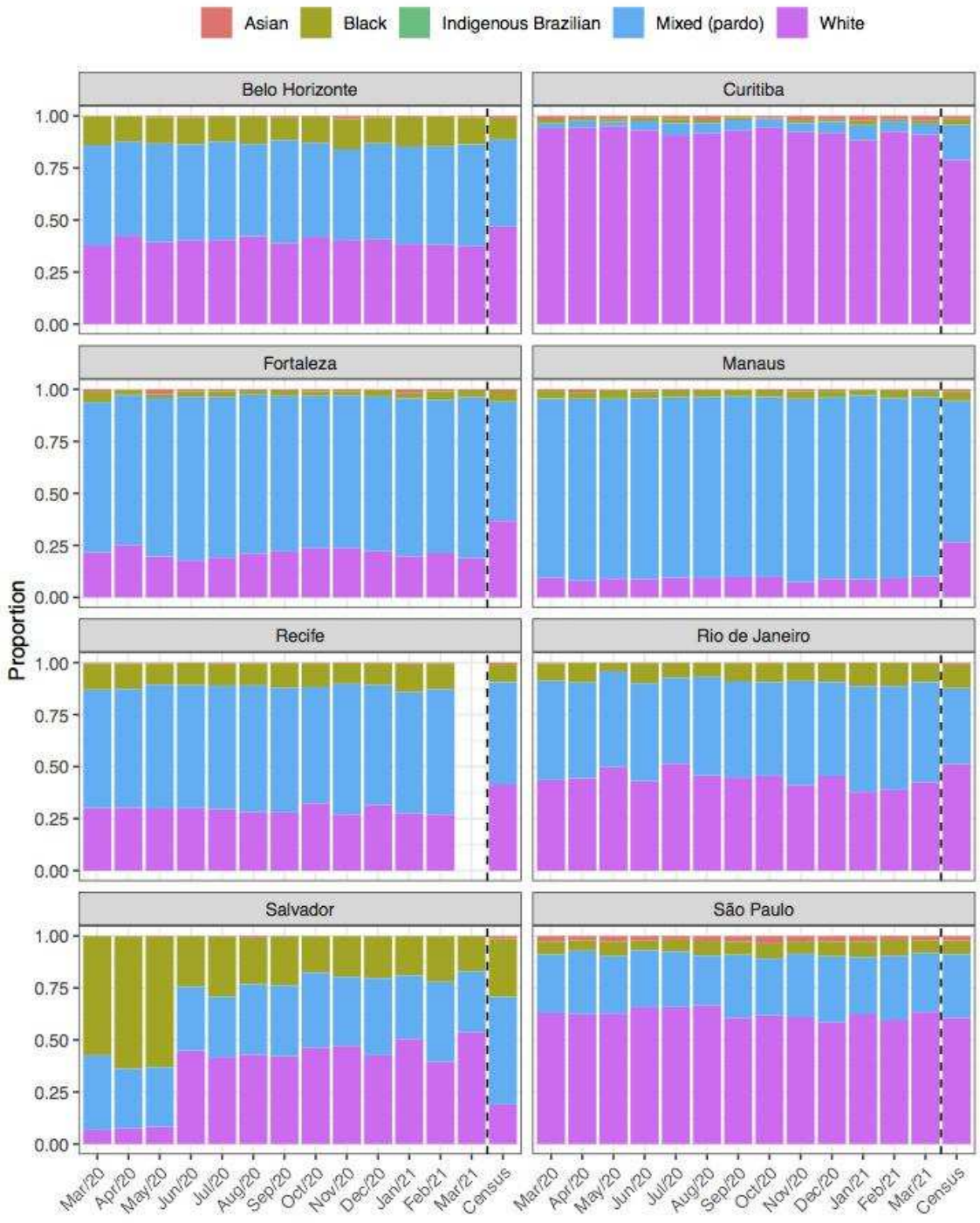
Supplementary Figure 15 - Comparison of age structure in blood donors tested for SARS-CoV-2 IgG antibodies population size projections for 2020 based on the last available 2010 census. y-axis shows. The census distribution for age is out of residents between 15-70yr.



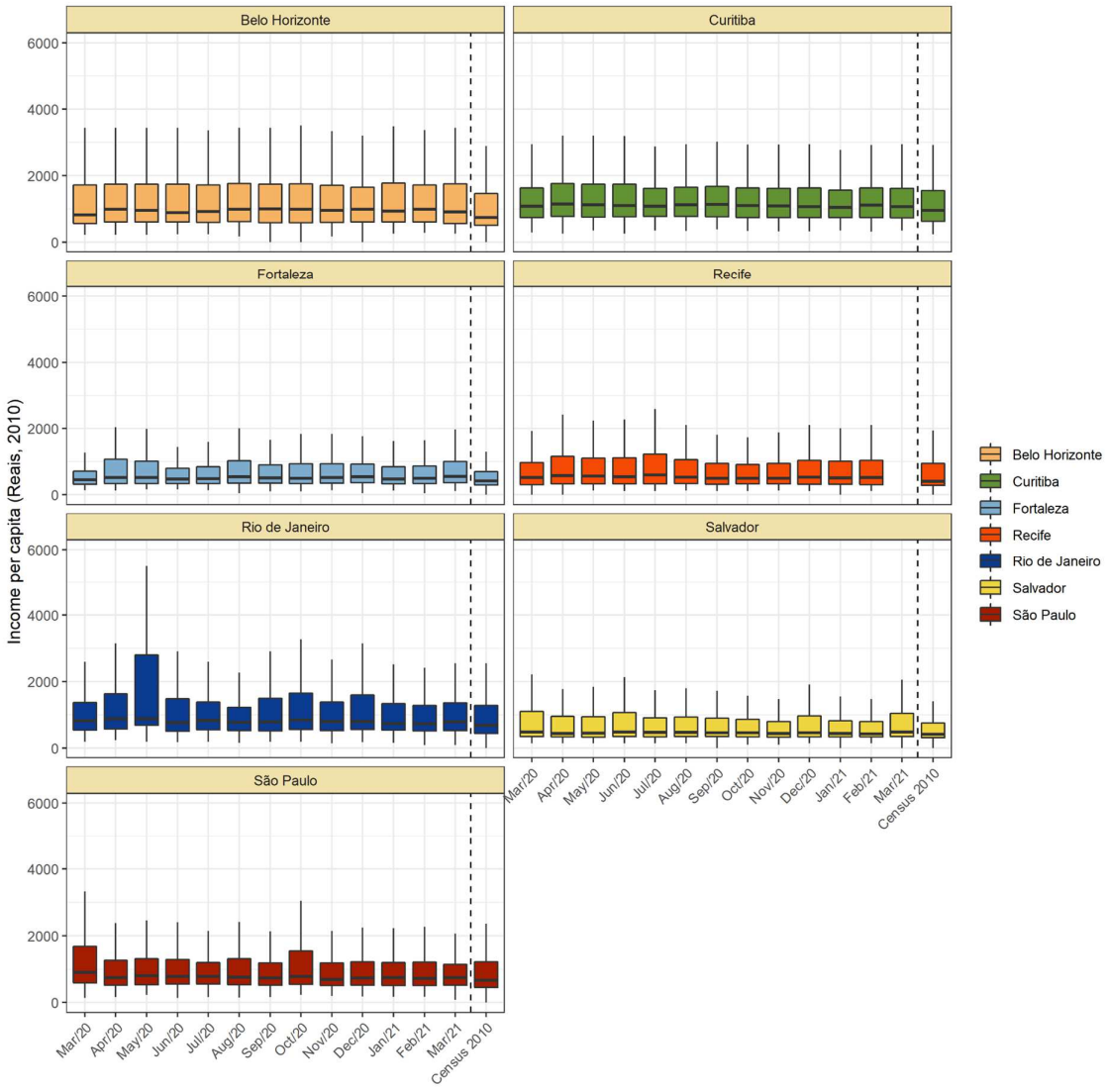
Supplementary Figure 16 - Abbott SARS-CoV-2 anti-N half-life and peak signal-to-cutoff (S/C) estimated in convalescent plasma by hospitalisation status. Instead of estimating one half-life for each plasma donor, a different half-life was calculated for each pair of consecutive donations with decaying S/C reading. Box plots show the median (central lines), interquartile range (hinges), and range extending to 1.5 times the interquartile range from each hinge (whiskers).



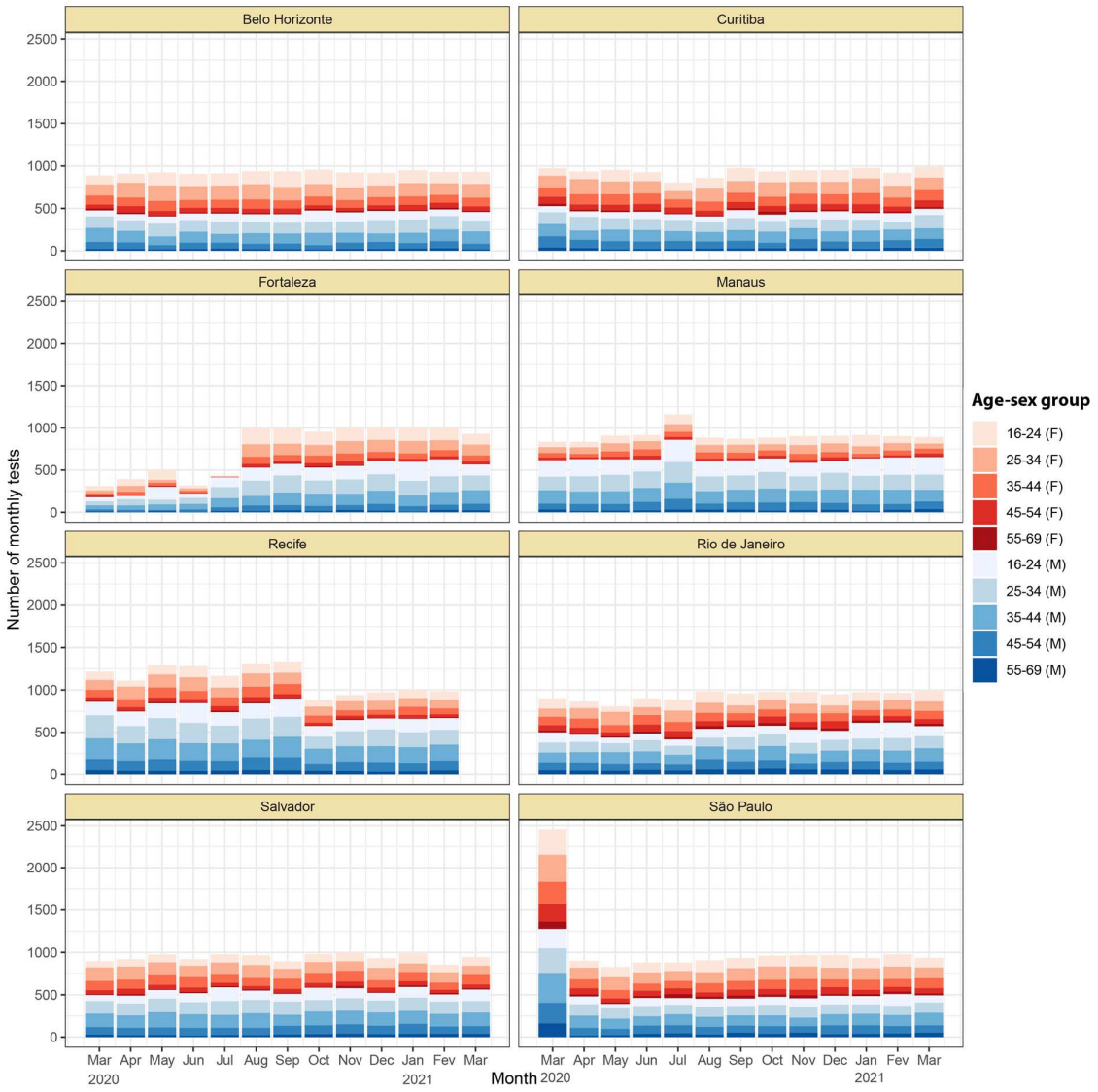
Supplementary Figure 17 - Comparison of sex structure in blood donors tested for SARS-CoV-2 IgG antibodies and the sex distribution at the last Brazilian census (2010).



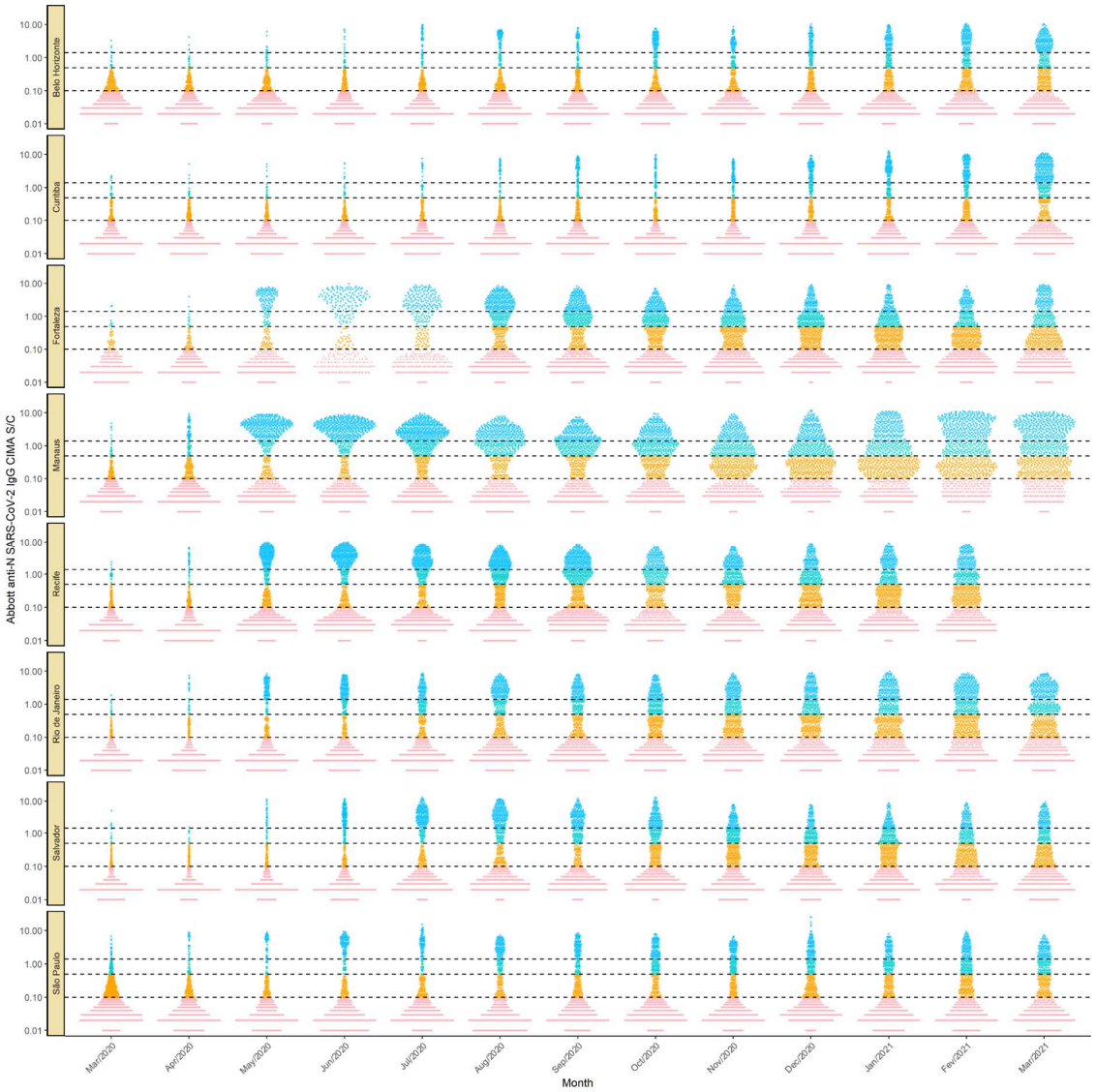
Supplementary Figure 18 - Comparison of self-declared skin colour among blood donors tested for anti-SARS-CoV-2 IgG antibody and the distribution of skin colour at the last available Brazilian census (2010).



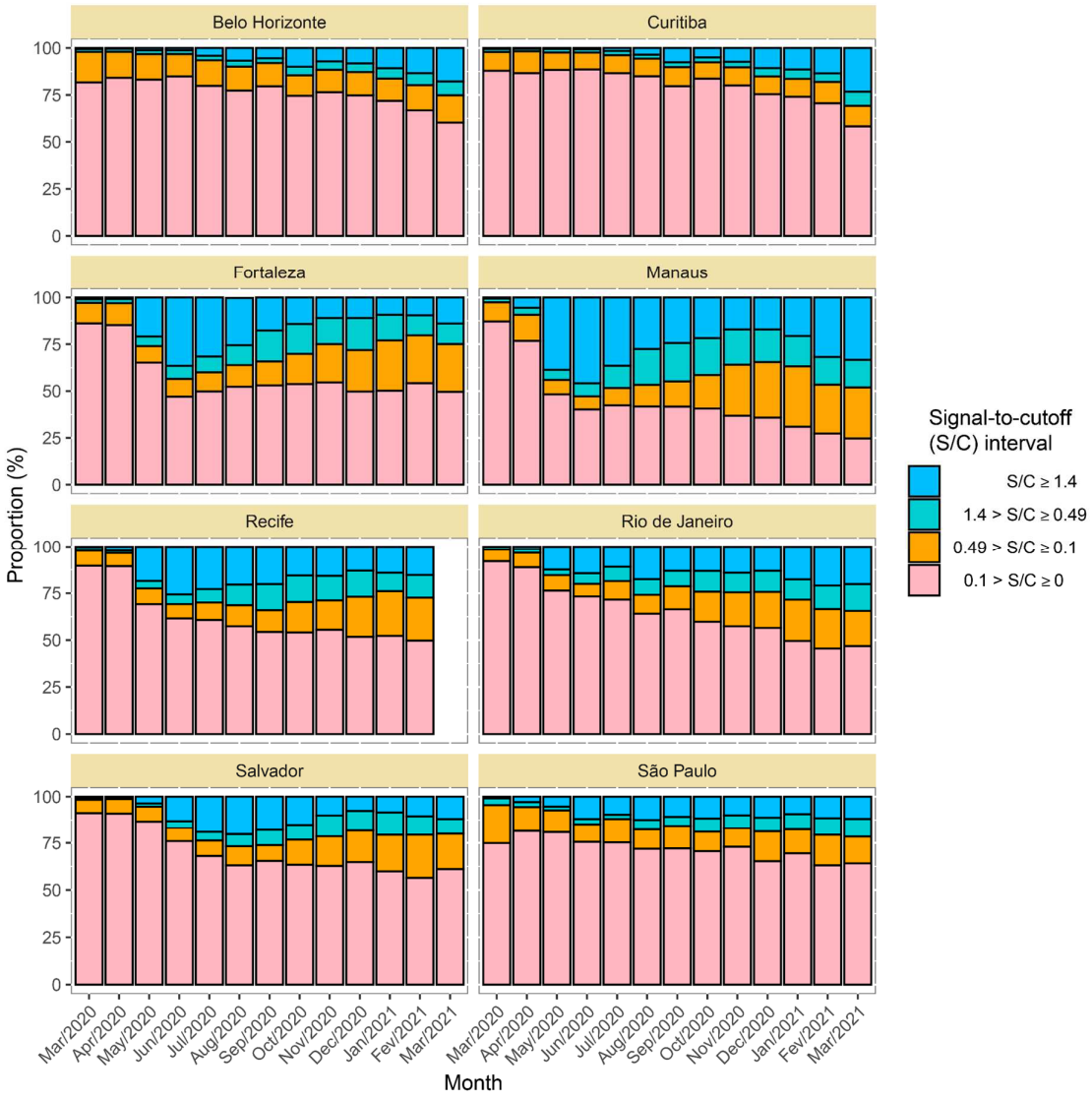
Supplementary Figure 19 - Comparison of income per capita of the census tract of selected blood donors and the income per capita distribution at the last available Brazilian census (2010) for each municipality. Information on census tracts was not available for Manaus.



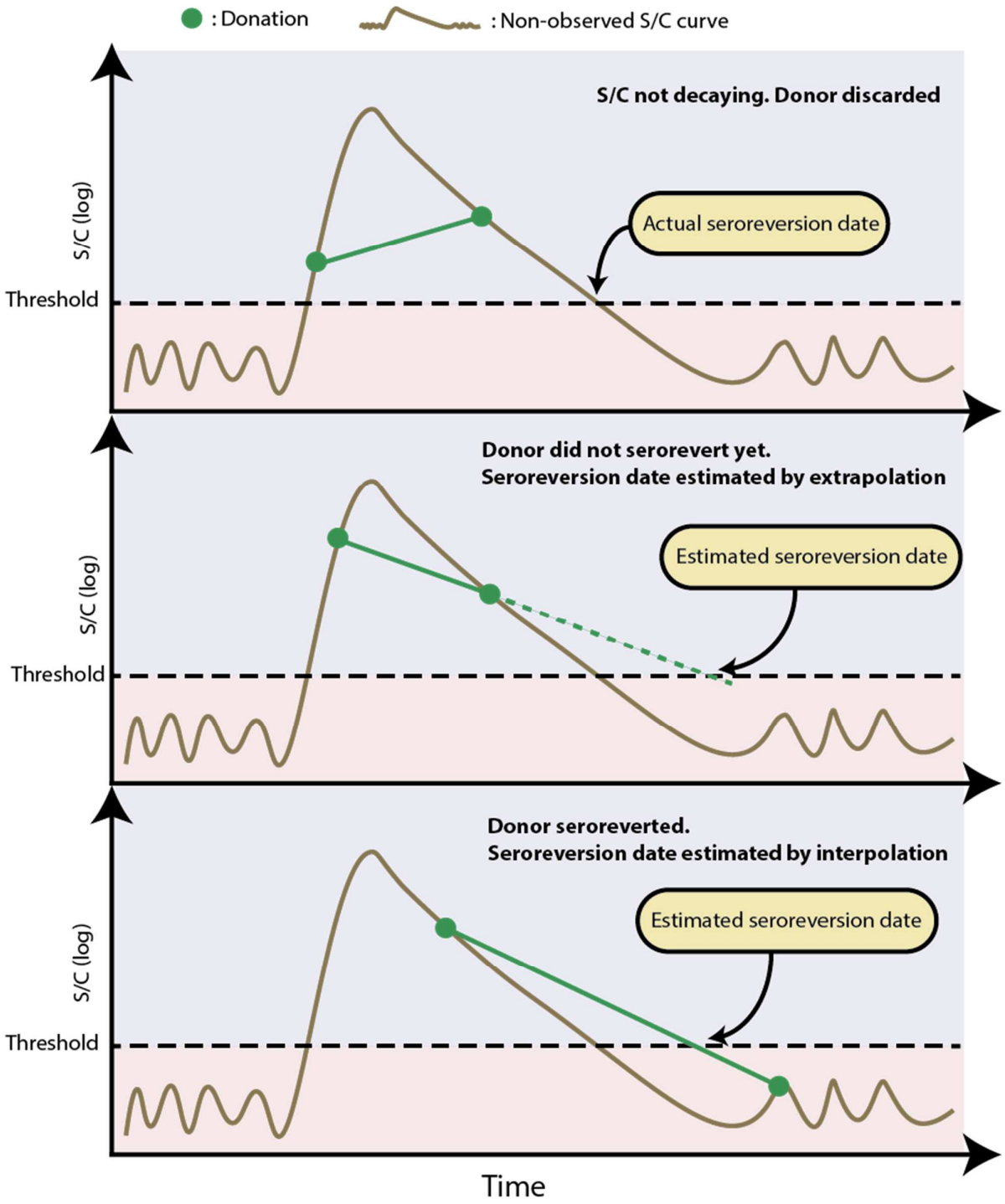
Supplementary Figure 20 - Monthly number of tests disaggregated by age group and sex. Around 1,000 monthly tests were applied between March 2020 and March 2021, except for Recife where tests were applied until February 2021.



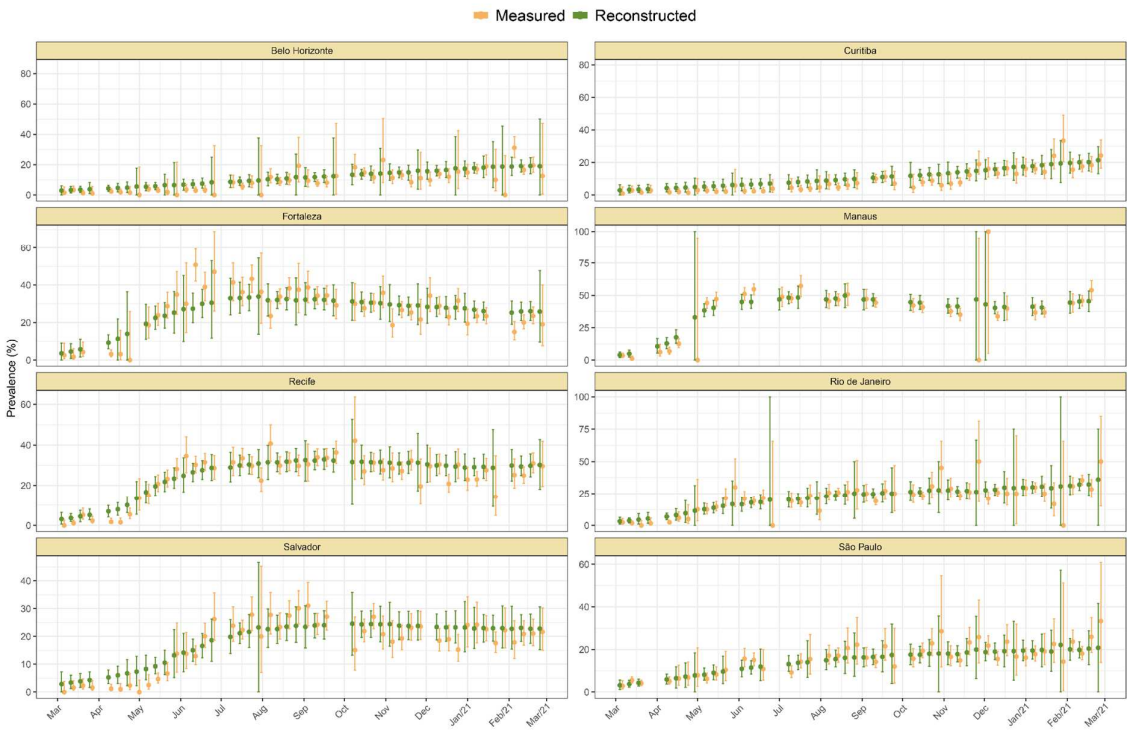
Supplementary Figure 21 - Monthly antibody signal-to-cutoff (S/C) reading in each of the eight cities. Each point represents the test result of a blood donor. The dashed lines represent the thresholds 1.4 (the threshold recommended by the manufacturer), 0.49 (the lower threshold recommended by the manufacturer and used in the main analyses of this paper) and 0.1, which we use in validation analyses.



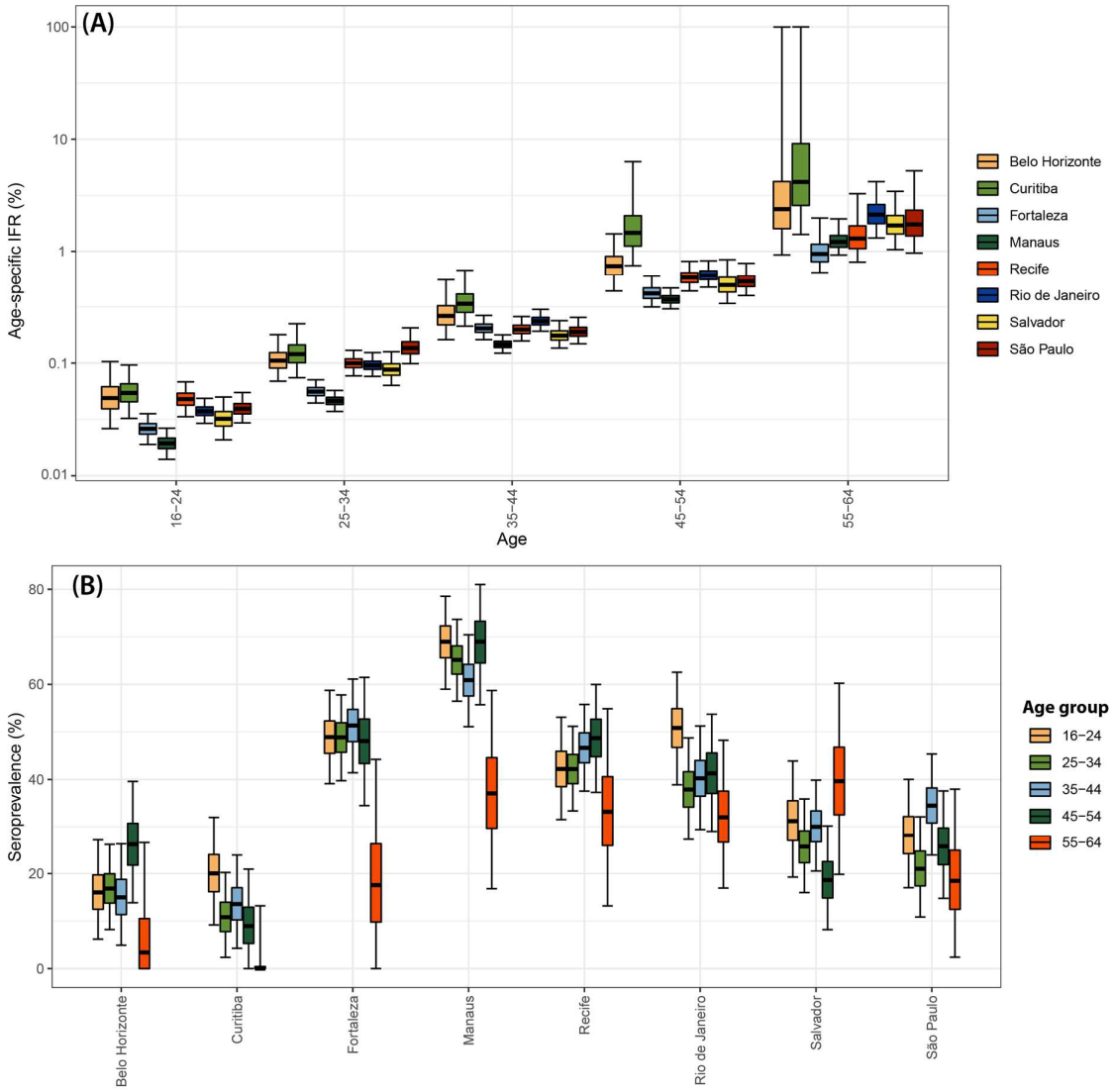
Supplementary Figure 22 - Time series of the proportional distributions of raw signal-to-cutoff readings in monthly blood donor samples. The thresholds were chosen as follows: 1.4 is the manufacturer's recommended upper threshold for assay positivity which maximises specificity; 0.49 is the manufacturer's lower recommended threshold, which improves sensitivity following antibody waning; and 0.1 S/C is an even lower threshold used for this analysis, which still provides specificity of 86% (112 false positives in 821 pre-pandemic blood donation samples) but further improves sensitivity in the face of antibody waning.



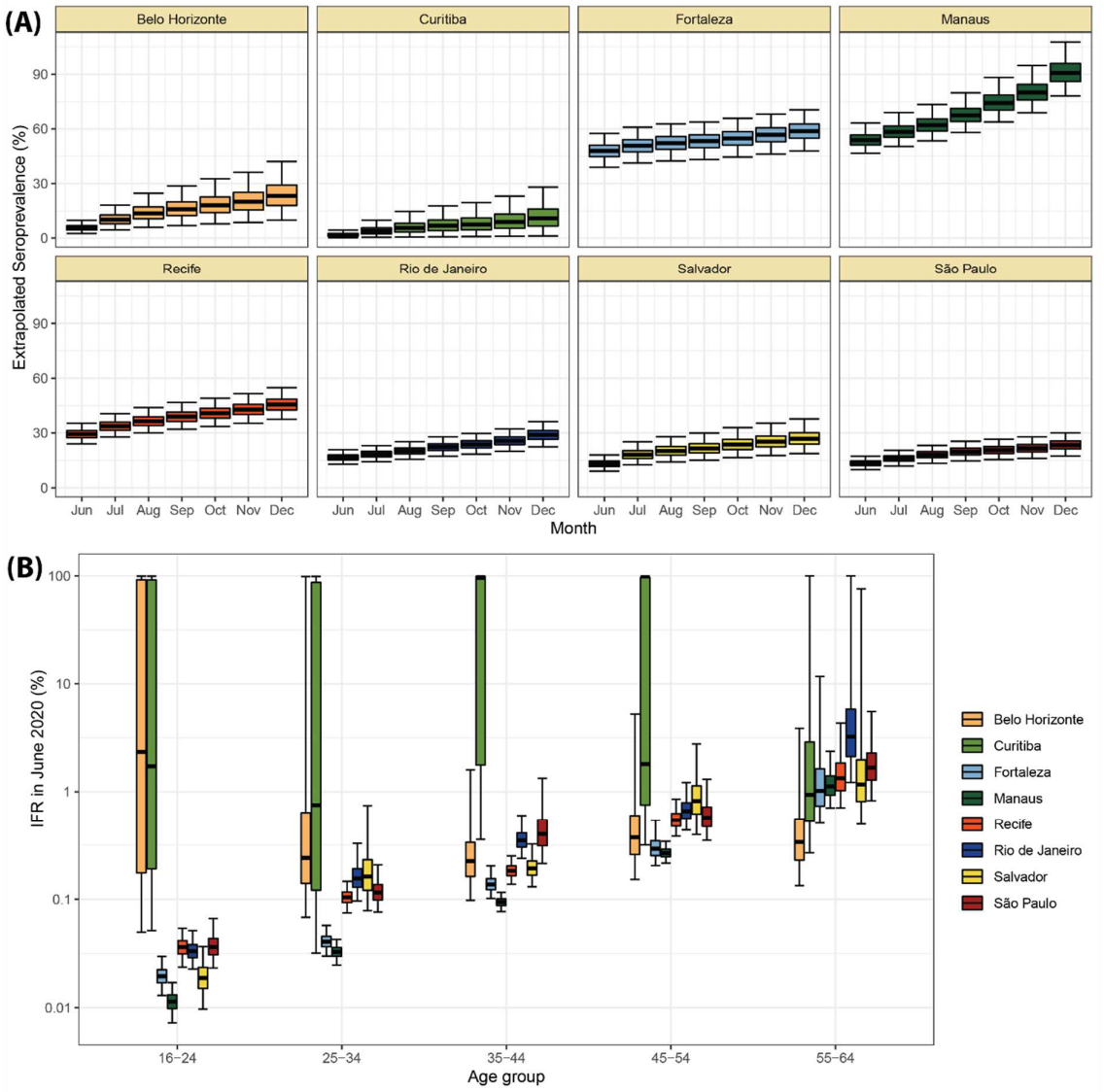
Supplementary Figure 23 - Illustration of the procedure used to estimate the time of seroreversion for each repeat blood donor. This figure shows three donors with the same idealized S/C curve but different donation dates. The first donor was discarded because the observed S/C was rising. The second donor did not become negative yet after seroconversion, thus the seroreversion date is estimated by extrapolating an exponential curve that contains the last two positive results. The third donor became negative after seroconversion, hence the seroreversion date was estimated by applying an exponential interpolation that contains the last positive result and the first negative result after seroconversion.



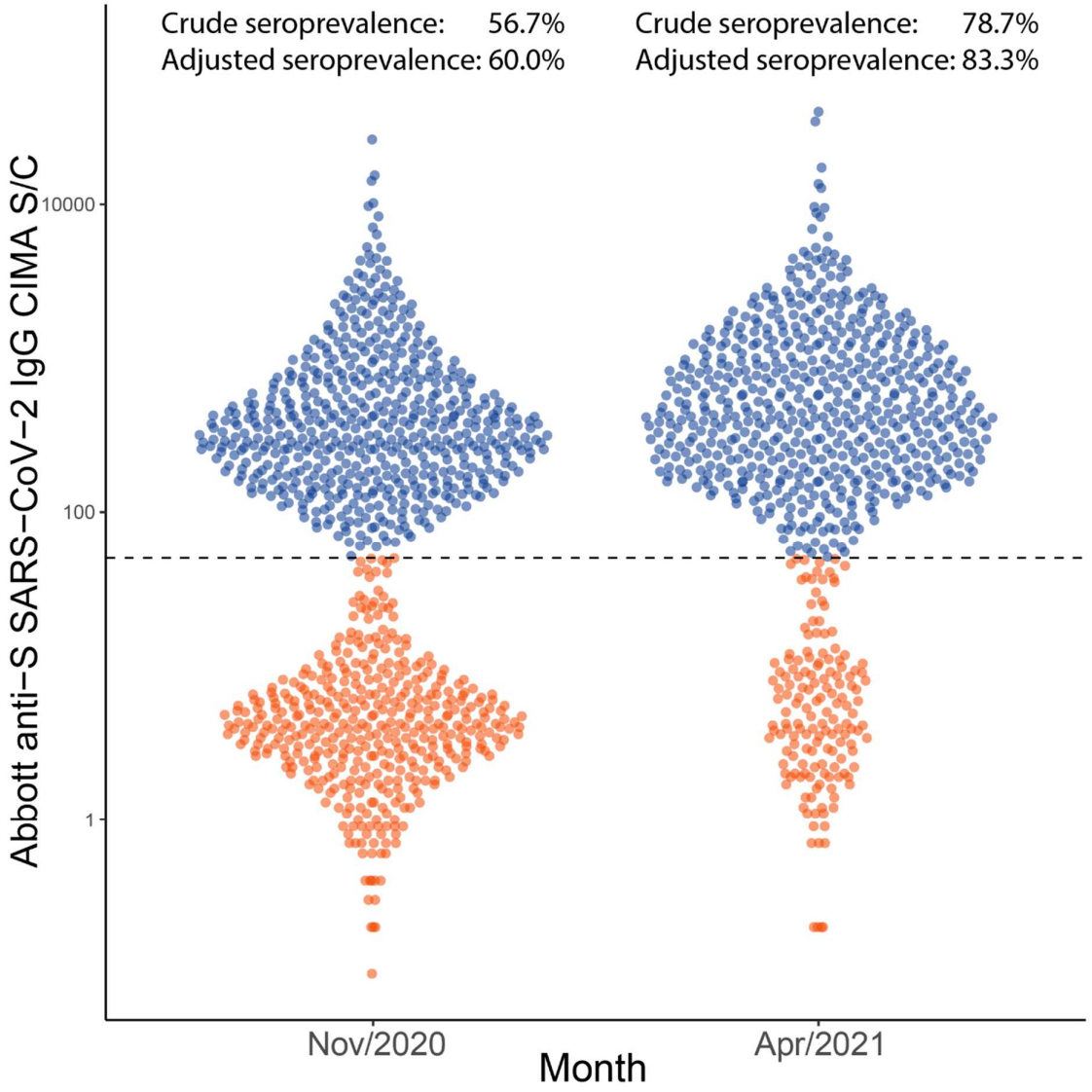
Supplementary Figure 24 - Measured weekly crude seroprevalence compared to the crude seroprevalence reconstructed by our proposed seroreversion correction model. The high similarity between both quantities shows that the seroprevalence estimated by the model is compatible with the observations.



Supplementary Figure 25 - IFR and seroprevalence measured in December 2020 using 0.1 as threshold and correcting for sensitivity, specificity and reweighting by age and sex. No seroreversion correction was performed to estimate the seroprevalence. That the wide confidence intervals for the IFR of Curitiba and Belo Horizonte are due to the small number of infections and deaths in these cities.



Supplementary Figure 26 - (A) Seroprevalence obtained by extrapolating the IFR measured in 2020 to the following months. The seroprevalence was obtained by dividing the cumulative number of deaths in each month by the IFR estimated for June 2020. **(B)** Age-specific IFR estimated in June 2020. Cities with small attack rate in June 2020 as Belo Horizonte and Curitiba present a large uncertainty in the estimated IFR.



Supplementary Figure 27 - Distribution of signal-to-cutoff in Manaus obtained in November 2020 and April 2021 with the anti-S assay. Each points represents the test result for a blood donor. Using a threshold of 50 (dashed line), the crude seroprevalence was 56.7% in November and 78.7% in April. Adjusting for sensitivity, these estimates increase to 78.7% and 83.3%.

	Convalescent plasma donors		Pre-pandemic blood donors cohort			
Method (threshold)	**Number of positive tests (TP)**	**Number of negative tests (FN)**	**Number of positive tests (FP)**	**Number of negatives tests (TN)**	**Sensitivity (%) and 95% CrI**	**Specificity and 95% CrI**
Anti-N assay; (1.4 S/C)	174	34	1	820	83.8 (78.0 - 88.4)	99.8 (99.3, 100.0)
Anti-N assay; (0.49 S/C)	189	19	20	801	90.6 (86.2, 94.0)	97.5 (96.3 - 98.4)
Anti-N assay; (0.1 SC)	197	11	112	709	94.4 (90.8 - 97.0)	86.3 (83.8 - 88.5)
Anti-S assay; (50 S/C)	231	14	-	-	94.0 (90.6 - 96.5)	Assumed 100%

Supplementary Table 1 - Number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) used to estimate the sensitivity of the anti-N assay for the thresholds 1.4, 0.49 and 0.1, and for the anti-S assay for the threshold 50.0. Sensitivity and specificity were calculated by computing the quantiles of Beta(1+TP, 1+FN) and Beta(1+TN, 1+FP), respectively.

	IFR obtained using the seroprevalence estimated with the proposed model (%)					IFR obtained using a threshold of 0.1 signal-to-cutoff to estimate the seroprevalence (%)				
Age group	**16-24**	**25-34**	**35-44**	**45-54**	**55-64**	**16-24**	**25-34**	**35-44**	**45-54**	**55-64**
Belo Horizonte	0.03 (0.02 - 0.05)	0.09 (0.07 - 0.11)	0.17 (0.14 - 0.21)	0.46 (0.37 - 0.57)	1.36 (0.99 - 1.96)	0.05 (0.03 - 0.1)	0.11 (0.07 - 0.18)	0.26 (0.16 - 0.55)	0.74 (0.44 - 1.43)	2.38 (0.93 - 99.84)
Curitiba	0.05 (0.03 - 0.07)	0.08 (0.06 - 0.11)	0.22 (0.17 - 0.27)	0.74 (0.61 - 0.91)	2.2 (1.6 - 3.16)	0.05 (0.03 - 0.1)	0.12 (0.07 - 0.22)	0.34 (0.21 - 0.68)	1.46 (0.74 - 6.31)	4.16 (1.42 - 99.96)
Fortaleza	0.03 (0.02 - 0.03)	0.05 (0.04 - 0.06)	0.16 (0.14 - 0.19)	0.38 (0.33 - 0.45)	0.9 (0.7 - 1.2)	0.03 (0.02 - 0.04)	0.06 (0.04 - 0.07)	0.20 (0.16 - 0.27)	0.42 (0.32 - 0.6)	0.95 (0.65 - 1.99)
Manaus	0.02 (0.01 - 0.02)	0.04 (0.03 - 0.05)	0.11 (0.1 - 0.13)	0.31 (0.27 - 0.36)	0.97 (0.8 - 1.19)	0.02 (0.01 - 0.03)	0.05 (0.04 - 0.06)	0.15 (0.12 - 0.18)	0.37 (0.3 - 0.47)	1.21 (0.93 - 1.95)
Recife	0.04 (0.03 - 0.05)	0.09 (0.07 - 0.1)	0.17 (0.15 - 0.2)	0.48 (0.41 - 0.55)	1.05 (0.85 - 1.33)	0.05 (0.03 - 0.07)	0.10 (0.08 - 0.13)	0.2 (0.16 - 0.26)	0.58 (0.44 - 0.81)	1.3 (0.8 - 3.27)
Rio de Janeiro	0.04 (0.03 - 0.04)	0.08 (0.07 - 0.1)	0.20 (0.18 - 0.23)	0.46 (0.41 - 0.52)	1.37 (1.16 - 1.65)	0.04 (0.03 - 0.05)	0.10 (0.08 - 0.12)	0.24 (0.19 - 0.3)	0.61 (0.47 - 0.82)	2.13 (1.32 - 4.19)
Salvador	0.02 (0.02 - 0.03)	0.06 (0.05 - 0.08)	0.16 (0.14 - 0.19)	0.34 (0.29 - 0.41)	0.97 (0.77 - 1.25)	0.03 (0.02 - 0.05)	0.09 (0.06 - 0.13)	0.18 (0.14 - 0.24)	0.49 (0.34 - 0.84)	1.71 (1.04 - 3.43)
São Paulo	0.04 (0.03 - 0.05)	0.11 (0.09 - 0.12)	0.20 (0.18 - 0.23)	0.53 (0.46 - 0.63)	1.65 (1.31 - 2.13)	0.04 (0.03 - 0.05)	0.14 (0.1 - 0.21)	0.19 (0.15 - 0.25)	0.53 (0.4 - 0.78)	1.74 (0.97 - 5.23)

Supplementary Table 2 - Age-specific IFRs and 95% credible intervals obtained with the seroprevalence estimated with our seroreversion correction model and with a threshold of 0.1 signal-to-cutoff corrected for sensitivity and specificity.

Algorithm: Method proposed to estimate $p^+[n]$

Input : Set of serial donations from N_{donors} repeat blood donors who have at least one positive result and a second result with decaying S/C; Daily incidence over time for repeat blood donors $u_{\text{repeat}}[n]$; Number of samples N_{samples} used to estimate the distribution of the time to seroreversion.

Output: Probability of an individual remaining positive n weeks after seroconversion $p^+[n]$

foreach $i = 1, \cdots, N$ **do**

Calculate the date of seroreversion for donor i by computing the instant where the exponential curve that passes through the last positive donation and first negative donation after seroconversion (if seroreversion occurred) or the two last positive donations crosses the threshold. Denote it by t_i^- ;

Denote $t_{\min}$ as the last negative result of the donor before seroconversion, or set $t_{\min}$ as March 1st 2020 if the donor has no positive results before seroconversion. Denote $t_{\max}$ as the date of the first positive result. The unobserved date of seroconversion belongs to the interval $[t_{\min}, t_{\max}]$ and its probability mass function is given by

$$p_i[n] = \begin{cases} \frac{u_{\text{repeat}}[n]}{\sum_{k=t_{\min}}^{t_{\max}} u_{\text{repeat}}[k]}, & t_{\min} \leq n \leq t_{\max} \\ 0, & \text{otherwise} \end{cases}.$$

Generate N_{samples} samples $\Delta t_i^{(1)}, \Delta t_i^{(1)}, \cdots, \Delta t_i^{(N_{\text{samples}})}$ from $\Delta t_i = t_i^- - t_i^+$, where $t_i^+ \sim p_i[n]$.

end

Calculate the empirical probability mass function of Δt by computing the empirical histogram of the generated samples $\Delta t_i^{(1)}, \Delta t_i^{(1)}, \cdots, \Delta t_i^{(N_{\text{samples}})}$ and denote it as $p_{\text{day}}^-[n]$.

Convert $p_{\text{day}}^-[n]$ from days to weeks as

$$p_{\text{week}}^-[n] = \frac{1}{7} \sum_{i=1}^{7} \sum_{j=7n+1}^{7(n+1)} p_{\text{day}}^-[j-i].$$

Calculate the probability of an individual remaining positive n weeks after seroconversion $p^+[n]$ as

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

Algorithm 1 - Method proposed to estimate the probability of an individual remaining positive n weeks after seroconversion ($p^+[n]$), the time-dependent component of the sensitivity.

Algorithm: Bayesian model to estimate the seroprevalence

Input : Probability of positiveness n weeks after seroconversion $p^+[n]$; Weekly number of tests $T[n,a]$ and number of positive tests $T^+[n,a]$ for each age-sex group a; Number of true positives TP, true negatives TN, false positives FP and false negatives FN used to determine the sensitivity and specificity of the test; Maximum seroprevalence allowed b.

Output: Weekly incidence $u[n,a]$ for age-sex group $a = 1, 2, \cdots, M$ and week $n = 1, 2, \cdots, N$

Priors:

$$\mathrm{se}_0 \sim \mathrm{Beta}(1 + \mathrm{TP}, 1 + \mathrm{FN})$$

$$\mathrm{sp} \sim \mathrm{Beta}(1 + \mathrm{TN}, 1 + \mathrm{FP})$$

$$u_{\mathrm{norm}}[1:N, a] \sim \mathrm{Dirichlet}(1, \cdots, 1) \qquad \forall a$$

$$\rho_{\max}[a] \sim \mathrm{Uniform}(0, b) \qquad \forall a$$

Auxiliary variables:

$$u[n,a] = u_{\mathrm{norm}}[n,a] \times \rho_{\max}[a] \qquad \forall n, a$$

$$\theta[n,a] = \mathrm{se}_0 \sum_{k=1}^{n} p^+[n-k]u[k,a] + (1 - \mathrm{sp}) \left(1 - \sum_{k=1}^{n} u[k,a]\right) \qquad \forall n, a$$

Likelihood:

$$T^+[n,a] \sim \mathrm{Binomial}(T[n,a], \theta[n,a]) \qquad \forall n, a$$

Algorithm 2 - The proposed Bayesian model to estimate the seroprevalence corrected by sensitivity, specificity and seroreversion. The seroprevalence is the cumulative sum of the obtained incidence, and can therefore be higher than 1 due to reinfections. In this work we use $b = 2$ to partially account for reinfections. Since this model assumes all infections occur in seronegative donors, $u[n,a]$ can be interpreted as the incidence in seronegative donors, and reinfections among seropositive individuals are not detected.

Algorithm: Method used to estimate the lower and upper bounds for the attack rate of the Gamma VOC in Manaus and the IFR.

Input : Number of deaths D; Monthly number of positives $T^{+}[n]$ and number of tests $T[n]$, for $n \in \{12, 13, 14, 15\}$ (December 2020 to March 2021); Population pop; Number of true positives TP and false negatives FN from the convalescent plasma donors cohort; Number of true negatives TN and false positives FP from the pre-pandemic cohort in Manaus; Population size pop

Output: Posterior samples of the upper bound for the attack rate of the second wave (AR_{max}) and the lower bound for the IFR (IFR_{min}).

Bayesian model:

$$\rho_{\text{December}} \sim \text{Uniform}(0, 1)$$

$$(u_{\text{norm}}[12], u_{\text{norm}}[13], u_{\text{norm}}[14]) \sim \text{Dirichlet}(1, 1, 1)$$

$$u_{\text{max}} \sim \text{Uniform}(0, 1)$$

$$u[n] = u_{\text{norm}}[n] \times u_{\text{max}}$$

$$T^{+}[n] \sim \text{Binomial}\left(T[n], \rho_{\text{December}} + \sum_{k=12}^{n} u[k]\right)$$

foreach *Posterior sample of* $\rho_{December}$, $u[12]$, $u[13]$, $u[14]$ **do**

Draw a sample from $\text{se} \sim \text{Beta}(1 + \text{TP}, 1 + \text{FN})$ and a sample from $\text{sp} \sim \text{Beta}(1 + \text{TN}, 1 + \text{FP})$;

Compute the incidence and seroprevalence in December corrected by sensitivity and specificity:

$$\rho^{\text{corr}}_{\text{December}} = \frac{\rho_{\text{December}} + \text{sp} - 1}{\text{se} + \text{sp} - 1}$$

$$u^{\text{corr}}[n] = \frac{u[n] + \text{sp} - 1}{\text{se} + \text{sp} - 1}$$

Compute $\text{AR}_{\text{max}} = \sum_{n=12}^{14} u^{\text{corr}}[n]$;

Draw samples from the lower bound for the IFR as $\text{IFR}_{\text{min}} \sim \text{Beta}(1 + D, 1 + \lfloor \text{AR}_{\text{max}} \times \text{pop} \rfloor - D)$.

end

Algorithm 3 - Method used to estimate the lower and upper bound for the attack rate of the Gamma VOC in Manaus and the IFR. This algorithm is executed for each age-sex group independently.

References

1. 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.
2. Prete, C. A., Jr *et al.* Reinfection by the SARS-CoV-2 Gamma variant in blood donors in Manaus, Brazil. *bioRxiv* (2021) doi:10.1101/2021.05.10.21256644.
3. Orner, E. P. *et al.* Comparison of SARS-CoV-2 IgM and IgG seroconversion profiles among hospitalized patients in two US cities. *Diagn. Microbiol. Infect. Dis.* **99**, 115300 (2021).
4. Barreto, I. C. de H. C. *et al.* Health collapse in Manaus: the burden of not adhering to non-pharmacological measures to reduce the transmission of Covid-19. *Saúde em Debate* **45**, 1126–1139 (2021).