

1 **Remarkable variability in SARS-CoV-2 antibodies across Brazilian**
2 **regions: nationwide serological household survey in 27 states**

3
4 Pedro C Hallal, Fernando P Hartwig, Bernardo L Horta, Gabriel D Victora, Mariângela
5 F Silveira, Claudio J Struchiner, Luís Paulo Vidaleti, Nelson Arns Neumann, Lúcia C
6 Pellanda, Odir A Dellagostin, Marcelo N Burattini, Ana M B Menezes, Fernando C
7 Barros, Aluísio J D Barros, Cesar G Victora

8
9 **Universidade Federal de Pelotas, Brazil** (Dr. Pedro C Hallal, Prof. Aluísio J D Barros,
10 Dr. Bernardo L Horta, Dr. Mariângela F Silveira, Prof. Odir A Dellagostin, Luis Paulo
11 V Ruas, Dr. Fernando P Hartwig, Prof. Ana B Menezes, Prof. Fernando C Barros, Prof.
12 Cesar G Victora), **Universidade Católica de Pelotas** (Prof. Fernando C. Barros),
13 **Laboratory of Lymphocyte Dynamics, Rockefeller University, USA** (Gabriel D
14 Victora), **Fundação Getúlio Vargas , Brazil** (Prof. Claudio J Struchiner), **Pastorate of**
15 **the Child, Brazil** (Nelson A Neumann), **Fundação Universidade Federal de Ciências**
16 **de Saúde de Porto Alegre, Brazil** (Prof. Lúcia C Pellanda), **Universidade Federal de**
17 **São Paulo, Brazil** (Dr. Marcelo N Burattini), **Universidade de São Paulo** (Dr. Marcelo
18 N. Burattini)

19
20

21 **List of e-mail addresses**

22

23 Pedro C Hallal prchallal@gmail.com

24 Fernando P Hartwig fernandophratwig@gmail.com

25 Bernardo L Horta blhorta@gmail.com

26 Gabriel D Victora victora@rockefeller.edu

27 Mariângela F Silveira mariangelafreitassilveira@gmail.com

28 Claudio J Struchiner claustru@gmail.com

29 Luis Paulo V. Ruas lpvidaletti@equidade.org

30 Nelson A Neumann nneumann@pastoraldacrianca.org.br

31 Lúcia C Pellanda pellanda@ufcspa.edu.br

32 Odir A Dellagostin odirad@gmail.com

33 Marcelo N Burattini mnburatt@gmail.com

34 Ana M B Menezes anamene.epi@gmail.com

35 Fernando C Barros fbarros.epi@gmail.com

36 Aluísio J D Barros abarros.epi@gmail.com

37 Cesar G Victora cvictora@gmail.com

38

39

40

41

42

43 **ABSTRACT**

44

45 Population based data on COVID-19 are essential for guiding public policies. We report
46 on the first of a series of planned seroprevalence surveys relying upon on household
47 probabilistic samples of 133 large sentinel cities in Brazil, including 25,025 participants
48 from all 26 states and the Federal District. Seroprevalence of antibodies to SARS-CoV-
49 2, assessed using a lateral flow rapid test, varied markedly across the country's cities
50 and regions, from below 1% in most cities in the South and Center-West regions to up
51 to 25% in the city of Breves in the Amazon (North) region. Eleven of the 15 cities with
52 the highest seroprevalence were located in the North, including the six cities with
53 highest prevalence which were located along a 2,000 km stretch of the Amazon river.
54 Overall seroprevalence for the 90 cities with sample size of 200 or greater was 1.4%
55 (95% CI 1.3-1.6). Extrapolating this figure to the population of these cities, which
56 represent 25% of the country's population, led to an estimate of 760,000 cases, as
57 compared to the 104,782 cases reported in official statistics. Seroprevalence did not
58 vary significantly between infancy and age 79 years, but fell by approximately two-
59 thirds after age 80 years. Prevalence was highest among indigenous people (3.7%) and
60 lowest among whites (0.6%), a difference which was maintained when analyses were
61 restricted to the North region, where most indigenous people live. Our results suggest
62 that pandemic is highly heterogenous, with rapid escalation in Brazil's North and
63 Northeast, and slow progression in the South and Center-West regions.

64

65

66 Although the need for population-based data on COVID-19 is widely recognized,^{1,2}
67 very few national surveys are available.³ Studies that have attempted to recruit
68 population-based samples – as opposed to volunteers or convenience samples – have
69 been recently carried out in a handful of countries. National studies using RT-PCR
70 showed prevalence of 0.6% in Iceland⁴, 0.3% in Austria⁵, and 0.9% in Sweden.⁶ A
71 national serological survey in Spain found a prevalence of 5.0%, ranging from less than
72 2% in some regions to 11% in Madrid.⁷ Studies among volunteers or convenience
73 samples^{4,8-10} usually show higher prevalence than population-based studies.

74
75 The first case of COVID-19 in Brazil was reported on February 27 in the city of São
76 Paulo, and by May 29, there are over 440,000 reported cases and 26,000 deaths. Based
77 on reported events, eight out of the 27 Federation Units (26 states and the Federal
78 District) present cumulative mortality rates above 10 per 100,000 inhabitants: four in
79 the North (Amazonas, Pará, Amapá and Roraima), two in the Northeast (Ceará and
80 Pernambuco) and two in the Southeast Region (Rio de Janeiro and São Paulo).

81
82 Three population-based studies are available from the Southern half of Brazil. Data
83 were collected on April 11-13, April 25-27 and May 9-11 in nine large cities in Rio
84 Grande do Sul state. Prevalence of antibodies using the Wondfo rapid test were 0.05%,
85 0.13% and 0.22% in the three waves, respectively.¹¹ In the city of Ribeirão Preto (São
86 Paulo state) prevalence was equal to 1.4% based on the same test¹² and in six high-risk
87 districts in São Paulo city prevalence was found to be 5.4% using a chemiluminescence
88 immunoassay.¹³

89

90 Governmental response to the pandemic has been marked by controversy, with the
91 country's president opposing social distancing measures and minimizing the importance
92 of COVID-19. Two consecutive Ministers of Health were either fired or resigned in less
93 than one month due to opposition to the president's stance. In contrast, most state
94 governors and city mayors enforced closure of schools, shops and non-essential
95 services, and more recently imposed lockdowns and compulsory use of masks in public
96 spaces. Hospital services, particularly in the eight states mentioned above, are at the
97 brink of collapse due to the high numbers of affected patients, particularly for intensive
98 care.

99

100 From mid-April onwards, a number of mayors and governors have relaxed social
101 distancing policies. While schools remain closed and public gatherings prohibited,
102 industrial, commercial and services sectors are allowed to open daily for limited
103 periods, while use of face masks continues to be enforced.

104

105 In spite of the staggering official statistics, a vast majority of cases are not being
106 reported as testing is limited to severe illnesses, and there is also evidence that COVID-
107 19 deaths are undercounted. In light of the present crisis, there is an urgent need for
108 population-based data on the pandemic.

109

110 We conducted a nationwide seroprevalence survey in 133 sentinel cities in 26 Brazilian
111 states and the Federal District (Figure 1). Cities were selected because they constitute
112 the seats of the country's intermediate regions and represent commercial and services
113 hubs for surrounding urban and rural areas. Over an 8-day period (from to May 14 to
114 21), our field team visited a systematic sample of households in randomly selected

115 census tracts. In each household, we administered a rapid SARS-CoV-2 antibody test to
116 a randomly selected household member, followed by a brief questionnaire to collect
117 sociodemographic data (see Online Methods for details).

118

119 The planned sample included 250 subjects in each of the 133 cities. It was not possible
120 to conduct any interviews in one city; 250 individuals were tested in 46 cities, 200-249
121 in 44 cities, 100-199 in 14 cities, and 1-99 in 28. The total sample size was 25,025. The
122 sample fell short of the planned number due to lockdown measures imposed in several
123 cities with restrictions to mobility of the interviewers, and to lack of coordination
124 between the Ministry of Health (which commissioned the study) and the city and state
125 governments. These difficulties were compounded by the rapid spread of disinformation
126 through social media characterizing the interviewers as swindlers, or of even being part
127 of a plot to spread the virus. In 27 cities, interviewers were arrested and in eight cities
128 the tests were destroyed by local police forces.

129

130 To reach a total of 25,025 interviews in the 133 selected cities, 46,011 attempts were
131 made. Overall, 23% of the contacted households refused to participate and in another
132 23% residents were not available at the time of the visit, resulting in a response rate of
133 55%. Among the 90 cities where 200 or more interviews were completed, 18 had
134 response rates of 80% or higher, 27 between 60% and 79%, and 45 below 60%. In terms
135 of Brazil's five regions, the average numbers of tests per city were 230 in the North, 156
136 in the Northeast, 177 in the Southeast, 242 in the South and 165 in the Center-West
137 region. The national mean sample size by city was 188, or 75.2% of the target of 250
138 per city.

139

140 There are 5,565 cities in Brazil. We compared population sizes, reported COVID-19
141 cases and deaths by April 13, and the Human Development Index¹⁴ in three groups of
142 cities: the 90 where it was possible to conduct 200 or more tests during the survey, the
143 43 cities included in the original sample where fewer than 200 subjects were tested, and
144 the remaining 5,432 cities in the country (Supplementary Table 1). Cities with 200 or
145 more tests tended to have larger populations and higher rates of reported cases and
146 deaths than those with fewer than 200 tests, or the remaining cities in the country. The
147 Human Development Index of the first two groups tended to be higher than in the third
148 group of cities.

149

150 Table 1 shows the characteristics of individuals who provided blood samples. Compared
151 to the whole Brazilian population, our sample of large sentinel cities included more
152 individuals from the North and South regions, and fewer from the Southeast. Men and
153 young people were underrepresented, as were those over the age of 50 years. The
154 distribution in terms of skin color showed fewer individuals who reported being white,
155 and larger percentages of black, Asians (yellow skin color) and indigenous people.

156

157 In total, there were 347 positive results in 24,995 individuals with valid test results
158 (1.39%). Prevalence results were corrected estimating a sensitivity of 84.8% based on
159 four independent validation studies,¹⁵ and two different estimates of specificity,
160 respectively 99.95% based upon our own early study in Southern Brazil,¹¹ and 99.0%
161 specificity as in validation studies using frozen sera.¹⁵ Prevalence results using the latter
162 sets of estimates are presented in Supplementary Table 5; the largest difference in city
163 prevalence levels using the two methods was 1.1 percent point.

164

165 Fifteen cities had prevalence above 2.0% (Supplementary Table 2): 11 in the North, two
166 in the Northeast (Fortaleza with 8.9% and Recife with 3.4%) and two in the Southeast
167 region (São Paulo with 3.3% and Rio de Janeiro with 2.4%). Four cities from the North
168 showed prevalence above 10%: Breves (25.0%), Tefé (19.8%), Castanhal (15.5%),
169 Belém (15.2%) and Manaus (12.7%). Except for São Paulo, where 212 individuals were
170 tested, all other high-prevalence cities had samples of at least 240 subjects. The six
171 cities with highest prevalence were all located along a 2,000 km stretch of the Amazon
172 river, from Tefé in the central Amazon basin to Macapá, Belém, and Castanhal at the
173 mouth of the river (Figure 2).

174

175 Fifty-four (60%) of the 90 cities with at least 200 subjects had no positive tests, and 13
176 (14.4%) had only one positive test. Cities with zero cases represented 100% of those in
177 the Center-West region (9 out of 9), 80% of those in the South (16 out of 20), 55% of
178 those in the Southeast (11 out of 20), 55% of those in the Northeast (12 out of 22), and
179 only 32% of those in the North (6 out of 19). When the index subject was positive, other
180 household members were also tested; 21.6% of the families had at least another positive
181 case. For families with at least two other members, this proportion was equal to 26.3%.

182

183 Figure 3 shows prevalence estimates as they relate to the officially reported COVID-19
184 cases and deaths in the country as of May 13, in the 90 cities with 200 or more tests.
185 Due to the nature of antibody tests, very recent infections will not result in positive
186 tests.¹⁶ This is partially offset by lags in reporting cases due to the time required for
187 confirmation through RT-PCR testing, which according to local health authorities is of
188 approximately two weeks. In addition, deaths are likely undercounted, and there may be
189 a substantial time lag between infection and death, particularly in the case of prolonged

190 hospitalization. In spite of these caveats, there were strong correlations between
191 prevalence estimated by the survey and reported cases and deaths per population (Figure
192 3), with correlation coefficients of 0.66 and 0.85, respectively, both with $P < 0.0001$. Our
193 results regarding time trends were also consistent with officially reported cases and
194 deaths (Figure 4).

195

196 We estimated the completeness of recording and infection-fatality rates for the 90 cities
197 with samples of 200 or more tests. Taken together, these cities reported 104,782 cases
198 by May 13, compared to our estimate of 760,000 based on the survey findings, a seven-
199 fold difference – or, equivalently, 13.8% of underreporting. The ratio of deaths over
200 estimated cases – the estimated infection-fatality rate – was 1.0% (7,640). Our estimates
201 of underreporting at city level, for the 36 cities with non-zero prevalence, ranged from
202 0.4% to 57.3% of cases being officially reported, with estimated infection-fatality rates
203 ranging from 0.0% to 2.4% (Supplementary Table 3).

204

205 Table 1 shows a breakdown of prevalence findings, resulting from two sets of
206 individual-level analyses: unadjusted estimates, and estimates corrected for the test
207 validity parameters and for the clustered nature of the sample. We focus the presentation
208 of results on the second set. Individuals living in the North region had 6.3% prevalence,
209 with the second highest prevalence observed in the Northeast (0.8%). No cases were
210 observed in the Center-West region, and very few in the South.

211

212 Prevalence was similar among men and women. There was no significant difference
213 according to age ($P=0.323$) although individuals aged 80 or more years seem to be less

214 frequently affected. Children aged 0-4 and 5-9 years were just as likely to display
215 antibodies to SARS-CoV-2 as adults.

216

217 There were marked differences in prevalence according to ethnic group, ranging from
218 0.7% among whites to 4.3% in indigenous subjects. Because indigenous individuals in
219 the sample were concentrated in the North region (35.5% of all indigenous in the
220 national sample), and because most positive cases are also in the North, we carried out
221 further analyses to verify whether the higher prevalence in indigenous individuals were
222 due to confounding by region (Supplementary Table 4). In the national sample, the odds
223 ratio for positive serology in indigenous individuals, compared to whites, was 5.89
224 (95%CI 2.99-10.66). Adjustment for region led to substantial reduction of the odds ratio
225 to 2.26 (95%CI 1.13-4.17). Restriction of the analyses to the North region resulted in an
226 odds ratio of 2.57 (95%CI 1.22-5.02), and further restriction to the cities in the North
227 that are not state capitals produced an odds ratio of 2.45 (95%CI 1.00-5.40). P-values
228 for the comparison of the five ethnic groups in the four sets of analyses described above
229 were <0.001, 0.009, 0.003 and 0.030, respectively.

230

231 In terms of ethnicity, the second highest prevalence (2.4%) was found in the “brown”
232 category. This is the most heterogeneous ethnic category in the country. A large
233 nationwide genomic ancestry study¹⁷ showed that in the Northern city of Belém self-
234 classified brown individuals had, on average, 69% European ancestry, followed by 21%
235 Amerindian ancestry and 11% African ancestry, while in the South they had on average
236 44% European, 11% Amerindian and 45% African ancestries. Antibody prevalence in
237 the brown group was 1.6-fold that of whites in the North, 1.3-fold in the Northeast and
238 0.9-fold in the Southeast.

239 To our knowledge this is the largest population-based study of prevalence of antibodies
240 to SARS-CoV-2 in geographical scope, and the second largest – after the national
241 survey in Spain⁷ – in terms of sample size. Possibly the most remarkable finding from
242 our analyses was the cluster of high prevalence in six cities along the Amazon River. In
243 the city of Breves, the prevalence of 25% appears to be the highest ever reported
244 anywhere so far.¹⁸ This finding of high prevalence in a tropical region contradicts
245 common wisdom that continents such as Africa may be protected against COVID-19
246 due to high ambient temperature.¹⁹

247

248 A possible explanation for this cluster is that long river trips (for example 8 hours from
249 Belém to Breves, or 36 hours from Manaus to Tefé) offer the possibility of intense
250 contagion in overcrowded boats where most passengers use hammocks for sleeping or
251 resting in the decks (Figure 5). Except for road travel between Belém and Castanhal (67
252 km), all other transportation among the six cities is by river boat or – for a minority who
253 can afford it – by plane. Unpublished analyses by Manaus city government showed an
254 inverse association (Pearson's $r = 0.78$; $P < 0.0001$) between the daily number of boats
255 leaving the capital to a given city and the number of days elapsed between the first
256 reported case in Manaus and in the city of interest.²⁰ The Ministry of Health started
257 reporting report daily numbers of cases on March 28, by when there were 105 cases in
258 Manaus, 9 in Belém and 4 in Macapa – all state capitals - and one case in Castanhal
259 which is close to Belém. The first case in Tefé was reported on April 10, and the first in
260 Breves on April 20, or only three weeks before the reference date for the survey. The
261 high prevalence in the last two cities suggests that the epidemic was already well
262 underway when the first case was reported.

263

264 Another possibility for high prevalence in the Amazon is that indigenous people,
265 including those who self-classify as having “brown” skin color in the North
266 region, may have higher susceptibility to SARS-CoV-2 infection due to genetic or
267 sociocultural factors. According to the Articulation of Indigenous Peoples in Brazil,²¹
268 COVID-19 cases have already been detected in 71 different indigenous peoples around
269 the country. Comorbidity with metabolic and cardiovascular diseases, which is
270 increasing rapidly among native Brazilians²² would place them at increased risk of
271 death due to COVID-19. Collection of biological samples in future studies may throw
272 light on possible mechanisms.

273

274 We were only able to identify one population-based study of antibodies to SARS-CoV-2
275 in children. In Spain, prevalence was 1.1% among infants, 2.2% among children aged 1-
276 4 years, and 3.0% at ages 5-9 years.⁷ All of these figures were lower than the overall
277 study prevalence of 5%. In our study, young children displayed similar prevalence
278 (1.3%) to that observed at older ages – for example, 1.4% both in 10-19 and 20-29-year-
279 olds.

280

281 The limitations of our analyses include the restriction of the sample to sentinel cities that
282 constitute regional hubs, which are larger, more developed and better equipped with
283 health services than the country as a whole. Our survey response rate of 54.4% is similar
284 to that in the Spanish survey (59.5%) and higher than achieved in national surveys in
285 Iceland and Austria, both of which had response rates of about one third of the intended
286 sample.^{4,5} In addition to the above-mentioned difficulties with local authorities and
287 disinformation, it is known that many families have been moving away from large cities
288 during the social isolation period, to stay with friends or relatives in small cities or rural

289 areas. Our sample had fewer children than expected, which was probably due to their
290 reluctance to undergo a finger prick when randomly selected within the household; in
291 these cases, a second person was randomly selected and if that person also refused the
292 household was replaced.

293

294 While concerns have been raised regarding the use of less than perfect serological tests
295 for clinical decision-making and for defining individuals as immune to reinfection, use
296 of such tests for population-based seroprevalence estimates is much less controversial,
297 provided that sensitivity and specificity are sufficiently high and appropriately corrected
298 for.^{23,24} The rapid lateral flow test used in our analysis (Wondfo SARS-CoV-2 antibody
299 test) underwent four different validation studies, including one carried out by our own
300 team. These studies placed the test's sensitivity and specificity at 84.8% and 99.0%,
301 respectively. This was the second best-performing lateral flow test out of 10 assessed by
302 Whitman and colleagues.²⁵ Given our concern that the test's specificity in these four
303 validation studies that relied on frozen sera is underestimated, we used our own
304 specificity estimate from a statewide survey carried out in early April in nine cities with
305 very few reported cases for COVID-19¹⁵ in the main analyses, while in the
306 Supplementary Materials we also report results corrected for a specificity of 99.0% as
307 calculated from the pooled validation studies (see Online Methods for more detail). The
308 largest difference between the two sets of corrected estimates was 1.1 percent point.

309

310 Our serological results reflect the epidemic curve at a point in time that precedes sample
311 collection by several days¹⁶. Reported cases and deaths are also affected by a variable
312 time lag. Yet, our serological findings are consistent with reported deaths and cases
313 (Figures 3 and 4). Our findings show on average one in seven cases were reported in

314 these large cities, with above-average wealth and health services. However, this value
315 was subject to large variation, with the highest-prevalence city of Breves reporting only
316 1.1% of infection. Our infection-fatality estimate based on reported deaths over survey-
317 estimated cases was only 1.0%, but this statistic may be affected by underreporting of
318 COVID-19 deaths.

319

320 We documented that the COVID-19 pandemic has affected the five regions of Brazil
321 with widely different intensities. The highest prevalence was observed in a tropical area
322 along the Amazon river. Individuals with indigenous ancestry were at particularly high
323 risk. Young children were as likely to present antibodies as adults. This is the first wave
324 of a national survey to be repeated every three weeks in order to monitor progress of the
325 pandemic.

326

327 **Acknowledgments**

328 We acknowledge the support from Instituto Serrapilheira, Pastoral da Criança, the
329 Brazilian Collective Health Association (ABRASCO) and JBS's initiative 'Fazer o Bem
330 Faz Bem'.

331

332 Pedro C Hallal, Fernando P Hartwig, Bernardo L Horta, Gabriel D Victora, Mariângela

333 F Silveira, Claudio J Struchiner, Luis Paulo V Ruas, Lúcia C Pellanda, Odir A

334 Dellagostin, Marcelo Burattini, Ana M B Menezes, Fernando C Barros, Aluísio J D

335 Barros, and Cesar G Victora contributed to the conception and design of the work, to

336 the acquisition, analysis, and interpretation of data and the draft of the manuscript.

337 Nelson A Neumann contributed to the acquisition of data. All authors have approved the

338 submitted version and have agreed to be personally accountable for the author's own

339 contributions and to ensure that questions related to the accuracy or integrity of any part

340 of the work, even ones in which the author was not personally involved, are

341 appropriately investigated, resolved, and the resolution documented in the literature.

342

343 **Competing interests**

344 None declared

345

346

347 References

- 348 1. Pearce, N., Vandenbroucke, J.P., VanderWeele, T.J. & Greenland, S. Accurate
349 Statistics on COVID-19 Are Essential for Policy Guidance and Decisions.
350 *American journal of public health*, e1-e3 (2020).
- 351 2. Barreto, M.L., *et al.* [What is urgent and necessary to inform policies to deal with
352 the COVID-19 pandemic in Brazil?]. *Rev Bras Epidemiol* **23**, e200032 (2020).
- 353 3. Wise, J. Covid-19: Surveys indicate low infection level in community. *Bmj* **369**,
354 m1992 (2020).
- 355 4. Gudbjartsson, D.F., *et al.* Spread of SARS-CoV-2 in the Icelandic Population. *The*
356 *New England journal of medicine*. DOI: 10.1056/NEJMoa2006100 (2020).
- 357 5. Institute for Social Research and Consulting Ogris & Hofinger GmbH (SORA).
358 COVID-19 Prevalence. (2020).
- 359 6. Public Health Agency of Sweden. Nya resultat från undersökning av förekomsten
360 av covid-19 i Sverige. (Stockholm, 2020).
- 361 7. Ministerio de Ciencia e Innovacion and Ministerio de Salud. Spain. Estudio ENE-
362 COVID19: Primera ronda Estudio Nacional de Sero-Epidemiología de la Infección
363 por SARS-CoV-2 en España. Informe preliminar 13 mayo de 2020. (Madrid,
364 Spain, 2020).
- 365 8. Ministry of Health and Welfare (South Korea). Coronavirus disease 19, Republic of
366 South Korea. (2020).
- 367 9. Bendavid, E., *et al.* COVID-19 Antibody Seroprevalence in Santa Clara County,
368 California. Preprint at
369 <https://www.medrxiv.org/content/10.1101/2020.04.14.20062463v2> (2020).
- 370 10. County of Los Angeles Public Health. USC-LA County Study: Early Results of
371 Antibody Testing Suggest Number of COVID-19 Infections Far Exceeds Number
372 of Confirmed Cases in Los Angeles County. (Los Angeles, 2020).
- 373 11. Silveira, M., *et al.* Repeated population-based surveys of antibodies against SARS-
374 CoV-2 in Southern Brazil. Preprint at
375 <https://www.medrxiv.org/content/10.1101/2020.05.01.20087205v2> (2020).
- 376 12. Comitê Técnico do Inquérito Epidemiológico SARS-CoV-2. Avaliação da
377 prevalência de marcadores virológicos e sorológicos do SARS-CoV-2 na população
378 de Ribeirão Preto: um inquérito epidemiológico. Resultados preliminares.
379 (Universidade de São Paulo, Ribeirão Preto, 2020).
- 380 13. Grupo Mapeamento SARS-CoV-2. Inquérito domiciliar para estimar a
381 soroprevalência da infecção por SARS-CoV-2 no município de São Paulo. Estudo
382 piloto. (São Paulo, SP, 2020).
- 383 14. PNUD. *Atlas do Desenvolvimento Humano no Brasil.*, (Programa das Nações
384 Unidas para o Desenvolvimento, Brasília, 2010).
- 385 15. Pellanda, L.C., *et al.* Sensitivity and specificity of a rapid test for assessment of
386 exposure to SARS-CoV-2 in a community-based setting in Brazil. Preprint at
387 <https://www.medrxiv.org/content/10.1101/2020.05.06.20093476v1> (2020).
- 388 16. Long, Q.X., *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19.
389 *Nature medicine* (2020).
- 390 17. Pena, S.D., *et al.* The genomic ancestry of individuals from different geographical
391 regions of Brazil is more uniform than expected. *PloS one* **6**, e17063 (2011).
- 392 18. Popovich, N. & Sanger-Katz, M. The world is still far from herd immunity for
393 coronavirus. in *New York Times* (New York, 2020).
- 394 19. The, L. COVID-19 in Africa: no room for complacency. *The Lancet* **395**, 1669
395 (2020).

- 396 20. Boletim ODS Atlas Amazonas Número 7. Especial COVID-19. Vol. Number 7,
397 May (Universidade Federal do Amazonas Press, Manaus, AM, Brazil, 2020).
- 398 21. Articulacao de Povos Indigenas do Brasil. Vidas indígenas e o COVID-19. (2020).
- 399 22. Schmidt, M.I., *et al.* Chronic non-communicable diseases in Brazil: burden and
400 current challenges. *Lancet* **377**, 1949-1961 (2011).
- 401 23. Theel, E.S., *et al.* The Role of Antibody Testing for SARS-CoV-2: Is There One?
402 *Journal of clinical microbiology* (2020).
- 403 24. Bryant, J.E., *et al.* Serology for SARS-CoV-2: Apprehensions, opportunities, and
404 the path forward. *Science immunology* **5**(2020).
- 405 25. Whitman, J.D., Hiatt, J., Mowery, C.T. & et al. Test performance evaluation of
406 SARS-CoV-2 serological assays. (*unpublished*) (2020).
- 407 26. Petrucelli JL. Saboia AL (editors). Características étnico-raciais da população:
408 classificação e identidades. (Instituto Brasileiro de Geografia e Estatística (IBGE)
409 Rio de Janeiro, Brazil, 2013).

410

411

412 **Table 1. Seroprevalence according to sociodemographic characteristics.**

	Sample distribution		Brazil 2019	Unadjusted		Adjusted for sample design and test validity	
	Number	%		Estimate	95% CI	Estimate	95% CI
Region				P<0.001		P<0.001	
Northeast	6552	26.2%	27.2%	0.7%	0.5% - 0.9%	0.8%	0.5% - 1.1%
North	5064	20.3%	8.8%	5.4%	4.8% - 6.0%	6.3%	5.4% - 7.2%
Central-West	2477	9.9%	7.8%	0.0%	0.0% - 0.1%	0.0%	NA% - NA%
Southeast	5833	23.3%	42.1%	0.4%	0.2% - 0.6%	0.4%	0.2% - 0.7%
South	5069	20.3%	14.3%	0.1%	0.1% - 0.3%	0.1%	0.0% - 0.6%
Sex				P=0.168		P=0.225	
Female	14452	57.8%	51.7%	1.3%	1.1% - 1.5%	1.5%	1.2% - 1.8%
Male	10543	42.2%	48.3%	1.5%	1.3% - 1.8%	1.7%	1.4% - 2.0%
Age (years)				P=0.128		P=0.323	
0-4	430	1.7%	7.2%	1.4%	0.5% - 3.0%	1.6%	0.6% - 3.5%
5-9	682	2.7%	7.0%	1.2%	0.5% - 2.3%	1.3%	0.5% - 2.9%
10-19	2287	9.1%	15.1%	1.4%	0.9% - 1.9%	1.5%	1.0% - 2.2%
20-29	3866	15.5%	16.5%	1.4%	1.1% - 1.8%	1.6%	1.2% - 2.1%
30-39	3834	15.3%	16.3%	1.5%	1.2% - 2.0%	1.7%	1.3% - 2.3%
40-49	3975	15.9%	13.5%	1.6%	1.2% - 2.0%	1.8%	1.4% - 2.3%
50-59	4015	16.2%	11.0%	1.7%	1.3% - 2.1%	1.9%	1.4% - 2.5%
60-69	3381	13.5%	7.5%	1.0%	0.7% - 1.4%	1.1%	0.7% - 1.6%
70-79	1797	7.2%	4.0%	1.2%	0.8% - 1.8%	1.4%	0.8% - 2.1%
80+	728	2.9%	2.0%	0.5%	0.1% - 1.4%	0.6%	0.1% - 1.6%
Color/ethnicity				P<0.001		P<0.001	
White	9493	38.7%	45.2%	0.6%	0.5% - 0.8%	0.7%	0.5% - 0.9%
Brown	11042	45.1%	45.1%	2.1%	1.8% - 2.4%	2.4%	2.0% - 2.8%
Black	2961	12.1%	8.9%	1.1%	0.7% - 1.5%	1.2%	0.8% - 1.8%
Asian	685	2.8%	0.5%	1.2%	0.5% - 2.3%	1.3%	0.5% - 2.8%
Indigenous	327	1.3%	0.4%	3.7%	1.9% - 6.3%	4.3%	2.1% - 7.6%

413

419 **Figure 2.** Six cities with the highest prevalence in the study. The small maps show areas of 50 x
420 50 km. line 518: Satellite images are from Google Earth.

421

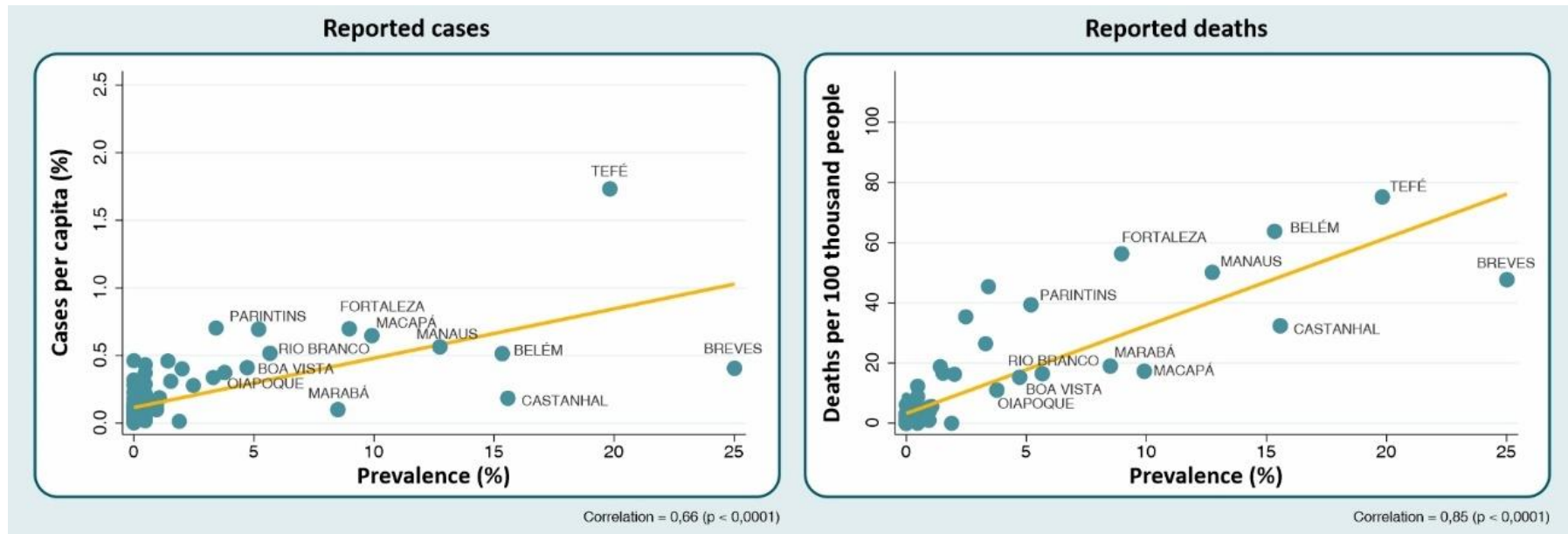


422

423

424 **Figure 3.** Scatter diagram for survey-based seroprevalence versus reported cases and deaths per population.

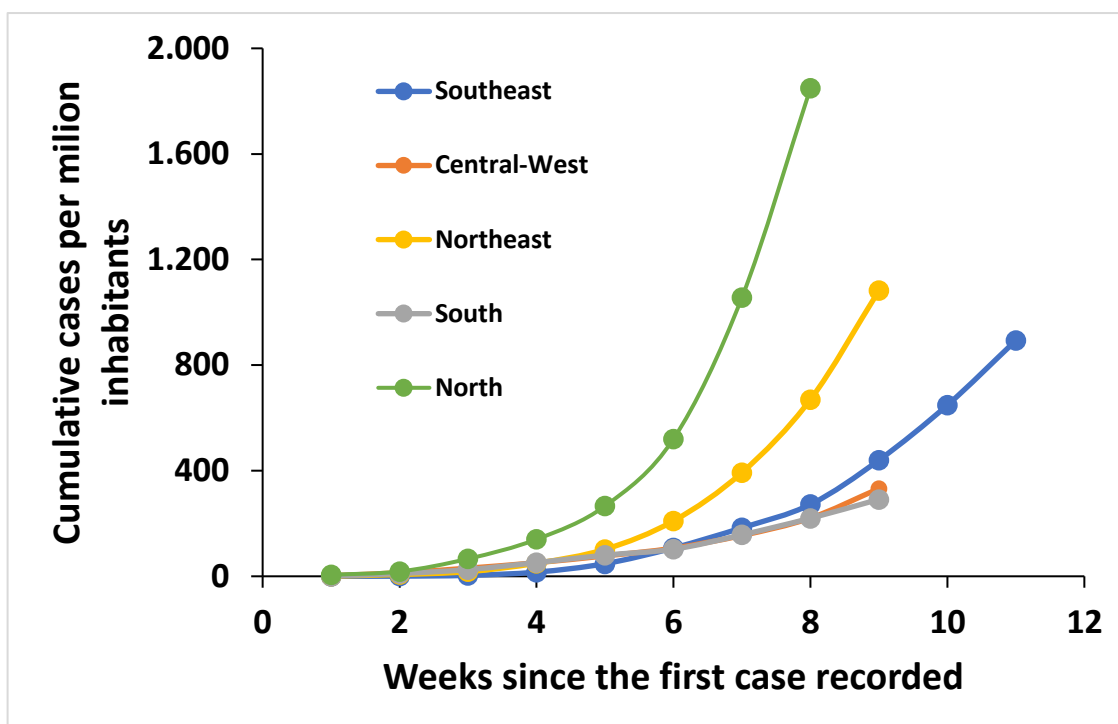
425



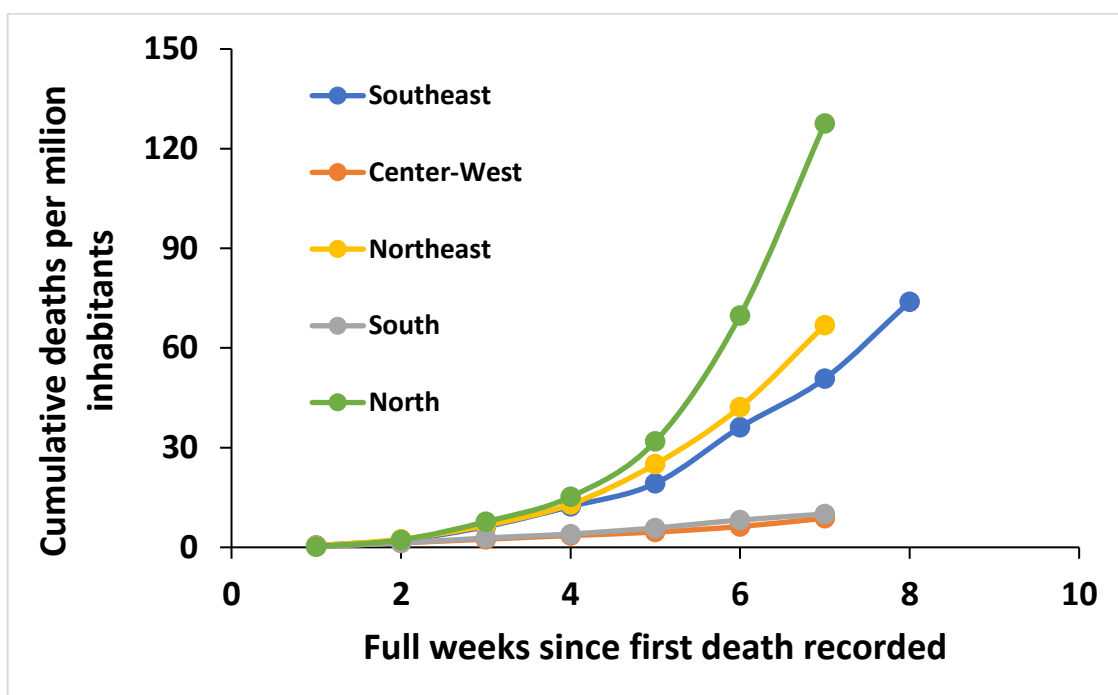
426

427

428 **Figure 4.** Time trends in reported cases and deaths between the start of the epidemic in each
429 region and May 13, 2020 (source: <https://covid.saude.gov.br>)



430



431

432

433 **Figure 5 (illustrative).** A crowded boat on the Amazon river.

434



435

436

437

438 **ONLINE METHODS**

439 Brazil's 27 federative units (26 states and the Brasilia Federal District) are divided by
440 the National Institute of Geography and Statistics in 133 intermediary regions. The main
441 city in each region was selected for the study (Figure 1). Further information on these
442 sentinel cities is provided in the Supplementary Materials.

443

444 Sampling

445 Using multistage sampling, we selected 25 census tracts with probability proportionate
446 to size in each sentinel city, and 10 households at random in each tract. Using the data
447 collection app, one individual was randomly selected from a listing of all household
448 members completed at the beginning of the visit. Data collection took place from May
449 15-22, 2020. With 250 individuals per city, the margins of error for estimating
450 prevalence figures of 2%, 5% and 10% are respectively 1.77, 2.70, and 3.79 percent
451 points. At national level, the total desired sample of 33,250, the corresponding margins
452 of error are 0.15, 0.24 and 0.33. In case the selected individual refused to provide a
453 blood sample, a second household member was randomly selected. If this person also
454 refused, the interviewers moved on to the next household to the right of the one that had
455 been originally selected.

456

457 Laboratory methods

458 Prevalence of antibodies was assessed with a rapid point-of-care test, the WONDFO
459 SARS-CoV-2 Antibody Test (Wondfo Biotech Co., Guangzhou, China), using finger
460 prick blood samples. This test detects immunoglobulins of both IgG and IgM isotypes
461 specific to SARS-CoV-2 antigens in a lateral flow assay. Two drops of blood from a
462 pinprick are sufficient to detect the presence of antibody. The assay reagent consists of

463 colloidal gold particles coated with recombinant SARS-CoV-2 antigens. Following the
464 introduction of the blood sample, reactive antibody:antigen:colloidal gold complexes, if
465 present, are captured by antibodies against human IgM and IgG present on the on the
466 “test” (T) line in the kit's window, leading to the appearance of a dark-colored line.
467 Samples without SARS-CoV-2-reactive antibodies will not lead to appearance of this
468 line. Valid tests are identified by a positive control line (C) in the same window. If this
469 control line is not visible, the test is deemed inconclusive, which is uncommon.

470

471 The rapid test underwent independent validation studies. According to the manufacturer,
472 it has a sensitivity of 86.4% and specificity of 99.6%

473 ([https://en.wondfo.com.cn/product/wondfo-sars-cov-2-antibody-test-lateral-flow-](https://en.wondfo.com.cn/product/wondfo-sars-cov-2-antibody-test-lateral-flow-method-2/)

474 [method-2/](https://en.wondfo.com.cn/product/wondfo-sars-cov-2-antibody-test-lateral-flow-method-2/)). The tests were acquired by Brazilian Ministry of Health for population

475 surveys and surveillance programs. A validation study carried out by the National

476 Institute for Quality Control in Health (INCQS, Oswaldo Cruz Foundation, RJ, Brazil)

477 showed a sensitivity of 100% and specificity of 98.7%. In an evaluation of 10 different

478 lateral flow assays, Whitman and colleagues²⁵ found that the Wondfo test was one of

479 the two with the best performance, with sensitivity of 81.5% and specificity of 99.1%.

480 Our own evaluation in Brazil found a sensitivity of 77.1% and specificity of 98.0%.¹¹

481 By pooling the results from the four validation studies, weighted by sample sizes,

482 sensitivity is estimated at 84.8% (95% CI 81.4%;87.8%) and specificity at 99.0% (95%

483 CI 97.8%;99.7%).¹¹

484

485 In early April 2020, our team conducted a household probability survey in nine cities in

486 the state of Rio Grande do Sul (Nature Medicine, in press), when the pandemic was at a

487 very early stage in the state. Of a total sample of 4,188 subjects there were only two

488 positive results. We believe that this survey provides a better estimate of the test’s false-
489 positive rate in the field, given that the other four studies relied on frozen samples for
490 specificity estimation. Assuming that all cases in that survey were false-positives leads
491 to a specificity rate of 99.95%. Whitman and colleagues, in their analyses of 10 lateral
492 flow tests, observed “*moderate-to-strong positive bands in several pre-COVID-19 blood*
493 *donor specimens, some of them positive by multiple assays, suggesting the possibility*
494 *of non-specific binding of plasma proteins, non-specific antibodies, or cross-reactivity*
495 *with other viruses.*”²⁵ Our findings suggest the possibility that studies using frozen
496 serum samples may have yielded higher false-positive rates than those associated with
497 testing fingerprick blood. We therefore used as correction parameters in the main
498 analyses a sensitivity of 84.8% and the 99.95% specificity derived from our previous
499 population-based survey (Nature Medicine, in press). Analyses using the same
500 sensitivity level and a specificity of 99.0% which is the weighted mean value of the first
501 four validation studies are presented in the supplementary data.

502

503 Data collection

504 Participants answered short questionnaires including sociodemographic information
505 (sex, age, schooling, skin color and household assets), COVID-19-related symptoms,
506 use of health services, compliance with social distancing measures and use of masks.
507 Due to the presence of widespread miscegenation, the official Brazilian classification of
508 ethnicity recognizes five groups, based on the question: “What is your race or color?”
509 The five response options are “white”, “brown” (“pardo” in Portuguese), “black”,
510 “yellow” and “indigenous”. Interviewers are instructed to check the “yellow” option
511 when the respondent mentions being of Asian descent, and “indigenous” when any of
512 the multiple first nations are mentioned. This system is endorsed by the Afro-

513 descendants movement, which advocates for disaggregation of all national statistics to
514 raise their visibility.²⁶

515

516 Field workers used tablets or smartphones to record the full interviews, register all
517 answers, and photograph the test results. All positive or inconclusive tests were read by
518 a second observer, as well as 20% of the negative tests. If the index subject in a
519 household had a positive result, all other family members were invited to be tested.

520

521 Ethical approval and data availability

522 Interviewers were tested and found to be negative for the virus, and were provided with
523 individual protection equipment that was discarded after visiting each home. Ethical
524 approval was obtained from the Brazilian's National Ethics Committee (process number
525 CAAE 30721520.7.1001.5313), with written informed consent from all participants.
526 Positive cases were reported to the municipal COVID-19 surveillance systems. Data
527 will become publicly available upon request from the corresponding author 30 days
528 after publication.

529

530 Data analyses

531 The survey data was analyzed using two strategies. The first consists of treating the
532 survey as if it were a simple random sample, using the exact binomial method to
533 calculate confidence intervals and the likelihood ratio test (implemented as logistic
534 regression) to compare the prevalence among sociodemographic groups. In the second
535 strategy, we accounted for both the sampling design of the survey and corrected for the
536 test validity, as described in detail in the Supplement. Hypothesis testing was performed
537 using Cochran's Q heterogeneity test implemented as fixed effects meta-regression. All

538 analyses were performed using R version 3.6.1.¹ The “survey” package^{1,2} was used to
539 account for the sampling design. Meta-regression was implemented using the “metafor”
540 package.⁴ Further information on the analytical approach is available in the
541 Supplementary Materials.

542

543 **References**

- 544 1. R Core Team. R: A Language and Environment for Statistical Computing. Vienna,
545 Austria: R Foundation for Statistical Computing, 2018 (<https://www.R-project.org/>).
- 546 2. Lumley, T. Analysis of complex survey samples. *J Stat Soft* **9**, 1-19 (2004).
- 547 3. Lumley, T. survey: analysis of complex survey samples. R package version 3.35-1,
548 2019 (<https://CRAN.R-project.org/package=survey>).
- 549 4. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J Stat*
550 *Soft* **36**, 1-48 (2010).

551