

1 **Longitudinal SARS-CoV-2 seroprevalence in a rural and urban community**
2 **household cohort in South Africa, during the first and second waves July**
3 **2020-March 2021**

4
5 Jackie Kleynhans^{1,2*} (MPH), Stefano Tempia^{1,2,3} (PhD), Nicole Wolter^{1,4} (PhD), Anne von Gottberg^{1,4}
6 (PhD), Jinal N. Bhiman^{1,4} (PhD), Amelia Buys¹ (MTech), Jocelyn Moyes^{1,2} (MD), Meredith L.
7 McMorrow^{3,5} (MD), Kathleen Kahn⁶ (PhD), F. Xavier Gómez-Olivé⁶ (PhD), Stephen Tollman⁶ (PhD),
8 Neil A. Martinson^{7,8} (MD), Floidy Wafawanaka⁶ (MSc), Limakatso Lebina⁷ (MD), Jacques du Toit⁶
9 (MD), Waasila Jassat⁹ (MD), Mzimasi Neti¹ (MPH), Marieke Brauer¹⁰ (MMed Path), Cheryl Cohen^{1,2}
10 (PhD), for the PHIRST-C Group.

11 ¹Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the
12 National Health Laboratory Service, Johannesburg, South Africa. ²School of Public Health, Faculty of
13 Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. ³Influenza Division,
14 Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. ⁴School of
15 Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
16 ⁵Influenza Program, Centers for Disease Control and Prevention, Pretoria, South Africa. ⁶MRC/Wits
17 Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty
18 of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. ⁷Perinatal HIV
19 Research Unit (PHRU), University of the Witwatersrand, Johannesburg, South Africa. ⁸Johns Hopkins
20 University Center for TB Research, Baltimore, Maryland, United States of America. ⁹Division of Public
21 Health Surveillance and Response, National Institute for Communicable Diseases of the National
22 Health Laboratory Service, Johannesburg, South Africa. ¹⁰Immunology Department, National
23 Reference Laboratory, Ampath Pathology, Pretoria, South Africa.

24 *Corresponding author: jackiel@nicd.ac.za

25 **Research in context**

26 *Evidence before this study*

27 Seroprevalence studies provide better estimates of SARS-CoV-2 burden than laboratory-confirmed
28 cases because many infections may be missed due to restricted access to care and testing, or
29 differences in disease severity and health-care seeking behaviour. This underestimation may be
30 amplified in African countries, where testing access may be limited. Seroprevalence data from sub-
31 Saharan Africa are limited, and comparing seroprevalence estimates between countries can be
32 challenging because populations studied and timing of the study relative to country-specific
33 epidemics differs. During the first wave of infections in each country, seroprevalence was estimated
34 at 4% in Kenya and 11% in Zambia. Seroprevalence estimates in South African blood donors is
35 estimated to range between 32% to 63%. South Africa has experienced two waves of infection, with
36 the emergence of the B.1.351/501Y.V2 variant of concern after the first wave. Reported SARS-CoV-2
37 cases may not be a true reflection of SARS-CoV-2 burden and specifically the differential impact of
38 the first and second waves of infection.

39 *Added value of this study*

40 We collected longitudinal blood samples from prospectively followed rural and urban communities,
41 randomly selected, household cohorts in South Africa between July 2020 and March 2021. From 668
42 and 598 individuals included from the rural and urban communities, respectively, seroprevalence
43 was found to be 7% (95%CrI 5-9%) and 27% (95%CrI 23-31%), after the first wave of infection, and
44 26% (95%CrI 22-29%) and 41% (95%CrI 37-45%) after the second wave, in rural and urban study
45 districts, respectively. After standardising for age, we estimated that only 5% of SARS-CoV-2
46 infections were laboratory-confirmed and reported. Infection-hospitalisation ratios in the urban
47 community were higher in the first (2.01%, 95%CI 1.57-2.57%) and second (2.29%, 95%CI 1.63-
48 3.94%) wave than the rural community where there was a 0.75% (95%CI 0.49-1.41%) and 0.66%
49 (95%CI 0.50-0.98%) infection-hospitalisation ratio in the first and second wave, respectively.

50 When comparing the infection fatality ratios for the first and second SARS-CoV-2 waves, at the urban
51 site, the ratios for both in-hospital and excess deaths to cases were significantly higher in the second
52 wave (0.36%, 95%CI 0.28-0.57% in-hospital and 0.51%, 95%CI 0.34-0.93% excess deaths), compared
53 to the first wave in-hospital (0.17%, 95%CI 0.15-0.20%) and excess (0.13%, 95%CI 0.10-0.17%)
54 fatality ratios, $p < 0.001$ and $p < 0.001$, respectively). In the rural community, the point estimates for
55 infection-fatality ratios also increased in the second wave compared to the first wave for in-hospital
56 deaths, 0.13% (95%CI 0.10-0.23%) first wave vs 0.20% (95%CI 0.13%-0.28%) second wave, and
57 excess deaths (0.51%, 95%CI 0.30-1.06% vs 0.70%, 95%CI 0.49-1.12%), although neither change was
58 statistically significant.

59 *Implications of all the available evidence*

60 In South Africa, the overall prevalence of SARS-CoV-2 infections is substantially underestimated,
61 resulting in many cases being undiagnosed and without the necessary public health action to isolate
62 and trace contacts to prevent further transmission. There were more infections during the first wave
63 in the urban community, and the second wave in the rural community. Although there were less
64 infections during the second wave in the urban community, the infection-fatality ratios were
65 significantly higher compared to the first wave. The lower infection-hospitalisation ratio and higher
66 excess infection-fatality ratio in the rural community likely reflect differences in access to care or
67 prevalence of risk factors for progression to severe disease in these two communities. In-hospital
68 infection-fatality ratios for both communities during the first wave were comparable with what was
69 experienced during the first wave in India (0.15%) for SARS-CoV-2 confirmed deaths. To our
70 knowledge, these are the first longitudinal seroprevalence data from a sub-Saharan Africa cohort,
71 and provide a more accurate understanding of the pandemic, allowing for serial comparisons of
72 antibody responses in relation to reported laboratory-confirmed SARS-CoV-2 infections within
73 diverse communities.

74

75 **ABSTRACT**

76 Background

77 SARS-CoV-2 infections may be underestimated due to limited testing access, particularly in sub-
78 Saharan Africa. South Africa experienced two SARS-CoV-2 waves, the second associated with
79 emergence of variant 501Y.V2. In this study, we report longitudinal SARS-CoV-2 seroprevalence in
80 cohorts in two communities in South Africa.

81 Methods

82 We measured SARS-CoV-2 seroprevalence two monthly in randomly selected household cohorts in a
83 rural and an urban community (July 2020-March 2021). We compared seroprevalence to laboratory-
84 confirmed infections, hospitalisations and deaths reported in the districts to calculate infection-case
85 (ICR), infection-hospitalisation (IHR) and infection-fatality ratio (IFR) in the two waves of infection.

86 Findings

87 Seroprevalence after the second wave ranged from 18% (95%CrI 10-26%) and 28% (95%CrI 17-41%)
88 in children <5 years to 37% (95%CrI 28-47%) in adults aged 19-34 years and 59% (95%CrI 49-68%) in
89 adults aged 35-59 years in the rural and urban community respectively. Individuals infected in the
90 second wave were more likely to be from the rural site (aOR 4.7, 95%CI 2.9-7.6), and 5-12 years (aOR
91 2.1, 95%CI 1.1-4.2) or ≥60 years (aOR 2.8, 95%CI 1.1-7.0), compared to 35-59 years. The in-hospital
92 IFR in the urban site was significantly increased in the second wave 0.36% (95%CI 0.28-0.57%)
93 compared to the first wave 0.17% (95%CI 0.15-0.20%). ICR ranged from 3.69% (95%CI 2.59-6.40%) in
94 second wave at urban community, to 5.55% (95%CI 3.40-11.23%) in first wave in rural community.

95 Interpretation

96 The second wave was associated with a shift in age distribution of cases from individuals aged to 35-
97 59 to individuals at the extremes of age, higher attack rates in the rural community and a higher IFR
98 in the urban community. Approximately 95% of SARS-CoV-2 infections in these two communities

99 were not reported to the national surveillance system, which has implications for contact tracing and
100 infection containment.

101 Funding

102 US Centers for Disease Control and Prevention

103

104

105

106

107

108

109

110

111

112

113 **Introduction**

114 The first laboratory-confirmed case of coronavirus disease 2019 (COVID-19) in South Africa was
115 announced on March 5, 2020, and the country has since experienced two waves of COVID-19.¹ A
116 nationwide lockdown from 27 March – 30 April 2020 confined all persons to their homes (excluding
117 essential services), which was followed by a gradual easing of restrictions.² The second wave of
118 infections began in November 2020,¹ and the country instituted a less restrictive lockdown from 28
119 December 2020 – 1 March 2021, which prohibited all gatherings with the exception of funerals and
120 mandated a night-time curfew.² Across Africa, the second wave was more severe than the first,³ and
121 specifically in South Africa higher weekly incidence, hospitalisations and deaths were reported for
122 the second wave, compared to the first.⁴⁻⁶ The second wave in South Africa was coupled with the
123 emergence of a new variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2),
124 501Y.V2, also known as B.1.351.⁷

125 South Africa reported more than 1.6 million laboratory-confirmed cases (by reverse transcription
126 polymerase chain reaction (RT-PCR) or rapid antigen tests) by mid-May 2021,⁴ but many cases may
127 go undiagnosed due to mild or absent symptoms or lack of, or reluctance to access care or testing.
128 Data on the proportion of people with serologic evidence of prior SARS-CoV-2 infection are critical to
129 assess infection rates, calculate infection-hospitalisation and infection-fatality ratios, compare
130 infection burden between waves of infection and to guide public health responses.⁸ Previous studies
131 have shown that SARS-CoV-2 seroprevalence is higher in close contacts of cases and high-risk
132 healthcare-workers, and lower in individuals younger than 20 years, or 65 years and older, with no
133 differences between males and females.⁹ It is still unclear whether HIV-infection increases the risk
134 for SARS-CoV-2 infection, and results from studies thus far have varied.^{10,11}

135 We describe the seroprevalence of SARS-CoV-2 by age and HIV-infection status in two household
136 cohorts in a rural and an urban community at five time points from July 2020, during the first wave,
137 to March 2021, after the second epidemic wave. We also compare disease burden between the first

138 and second wave by comparing the seroprevalence by wave to reported laboratory-confirmed
139 infections, hospitalisations and deaths within the respective districts in which these two
140 communities are located.

141 **Methods**

142 *Study population*

143 We conducted a prospective study on a randomly selected household cohort in a rural (Agincourt,
144 Ehlanzeni District, Mpumalanga Province) and an urban (Jouberton, Dr Kenneth Kaunda District,
145 North West Province) community as part of the PHIRST-C study (a **P**rospective **H**ousehold study of
146 **S**ARS-CoV-2, **I**nfluenza, and **R**espiratory **S**yncytial virus community burden, **T**ransmission dynamics
147 and viral interaction in South Africa). Methods for the cohort study are detailed in the appendix.

148 Recruitment to this study began in July 2020 and follow-up will continue through July 2021.

149 Households which previously participated in the PHIRST study (Prospective Household observational
150 cohort study of Influenza, Respiratory Syncytial virus and other respiratory pathogens community
151 burden and Transmission dynamics in South Africa) during 2016-2018,^{12,13} and additional randomly
152 selected households were eligible. Households with three or more household members of any age
153 were enrolled if $\geq 80\%$ of members consented.

154 The study was approved by the University of the Witwatersrand Human Research Ethics Committee
155 (Reference 150808) and the US Centers for Disease Control and Prevention relied on local clearance
156 (IRB #6840).

157 *Seroprevalence*

158 We collected baseline data, including demographics and HIV status, and blood (blood draw 1) at
159 enrolment (20 July – 17 September 2020), and every two months thereafter (blood draw 2 during 21
160 September – 10 October; blood draw 3 during 23 November – 12 December 2020; blood draw 4
161 during 25 January – 20 February 2021 and blood draw 5 during 22 March – 11 April 2021). HIV status

162 was confirmed from medical records (if HIV-infected), and by rapid test for participants with
163 unknown, or self-reported negative status. Prior SARS-CoV-2 infection was determined using the
164 Roche Elecsys® Anti-SARS-CoV-2 assay (Roche Diagnostics, Rotkreuz, Switzerland), using
165 recombinant nucleocapsid (N) protein. The assay was performed on the Cobas e601 instrument, and
166 a cut-off index (COI) of ≥ 1.0 was considered an indication of prior infection (seropositivity). Signal to
167 cut-off ratio was not considered. Data analysis was performed in Stata 14 (StataCorp, College
168 Station, Texas, USA), using six age groups: pre-school (<5 years), primary school (5-12 years),
169 secondary school (13-18 years), young adults (19-34 years), adults (35-59 years) and older adults
170 (≥ 60 years) who are prioritised for SARS-CoV-2 vaccination.¹⁴ Seroprevalence estimates were
171 adjusted for sensitivity and specificity as previous described,¹⁵ based on the manufacturers' reported
172 99.5% sensitivity and 99.8% specificity.¹⁶ Seroprevalence 95% credible intervals (95%CrI) were
173 obtained using Bayesian inference with 10,000 posterior draws.¹⁵ Pearson's chi-squared test was
174 used to assess the statistical significance of differences in SARS-CoV-2 seropositivity across blood
175 collection times, waves of infection and between the two communities.

176 *Calculation of infection-case ratio (ICR), infection-hospitalisation ratio (IHR) and infection-fatality*
177 *ratio (IFR) by wave of infection*

178 To assess the burden of SARS-CoV-2, and compare the severity of illness between the first and
179 second waves, we performed an ecological study comparing estimated number of infections based
180 on seroprevalence in our cohort study, to reported number of cases, hospitalisations and in-hospital
181 and excess deaths in the same district for each wave. We calculated the age-adjusted total number
182 of infections, laboratory-confirmed cases, hospitalisations, deaths, ICR (number of infections
183 compared to laboratory-confirmed cases), IHR and in-hospital and excess death IFR for each wave of
184 infection as described in below equations. The first wave was defined as 1 March 2020 (week 11) to
185 21 November 2020 (week 47) coinciding with the first case of SARS-CoV-2 reported in South Africa,
186 and ending the week before blood draw 3 started, and the second wave as 22 November 2020

187 (week 48) to 27 March 2021 (week 11), starting directly after the defined wave 1 period, and ending
 188 the week before blood draw 5 started (Figure 1). Data sources used in these calculations are
 189 described in the appendix. In brief, all laboratory-confirmed SARS-CoV-2 infections in South Africa
 190 (by RT-PCR and antigen tests) are reported to the Notifiable Medical Conditions Surveillance System
 191 (NMCSS) through automated feeding of data from public and private laboratory information systems
 192 to a data warehouse.⁴ The number of reported laboratory-confirmed SARS-CoV-2 cases from the two
 193 districts where the rural and urban communities were located were obtained from the NMCSS,⁴ and
 194 numbers of hospitalisations and in-hospital deaths were obtained from the COVID-19 National
 195 Hospital Surveillance (DATCOV).⁵ Provincial excess deaths per 100,000 population (based on death
 196 trends for 2014-2019) were obtained from the South African Medical Research Council report on
 197 weekly deaths.⁶ District-level and South African population denominators were obtained from the
 198 StatsSA mid-year population estimates for 2020.¹⁷ Age standardised estimates for the selected
 199 endpoints for each wave were obtained as follows:

$$200 \quad \text{Infections} = \frac{\sum_i (s_i \times SA p_i)}{\sum_i (SA p_i)} \times 100,000$$

201 Where s_i is the seroprevalence in the cohort in the respective community and wave for age group i
 202 and $p_{i(SA)}$ is the South African population for age group i . Calculated for wave 1 as seroprevalence at
 203 blood draw 3, and for wave 2 as seroprevalence at blood draw 5, excluding those who seroconverted
 204 at blood draw 3. Estimates only included participants with a blood draw 3 and 5 pair, and adjusted
 205 for sensitivity and specificity of test.¹⁵

$$206 \quad \text{Laboratory – confirmed cases} = \frac{\sum_i ((c_i \div p_{i(d)}) \times p_{i(SA)})}{\sum_i (p_{i(SA)})} \times 100,000$$

207 Where c_i is the number of laboratory-confirmed cases (RT-PCR and antigen-based tests) from the
 208 respective district reported to the NMCSS (wave 1: 3 March – 21 November 2020, wave 2: 22
 209 November 2020 – 27 March 2021) in age group i , $p_{i(d)}$ is the district population for age group i and
 210 $p_{i(SA)}$ is the South African population for age group i .

211
$$\text{Hospitalisations} = \frac{\sum_i((h_i \div p_{i(d)}) \times p_{i(SA)})}{\sum_i(p_{i(SA)})} \times 100,000$$

212 Where h_i is the number of hospitalisations from the respective district reported to COVID-19
 213 Sentinel Hospital Surveillance (DATCOV, wave 1: 5 March – 21 November 2020, wave 2: 22
 214 November 2020 – 27 March 2021)⁵ in age group i , $p_{i(d)}$ is the district population for age group i and
 215 $p_{i(SA)}$ is the South African population for age group i .

216
$$\text{In – hospital deaths} = \frac{\sum_i((d_i \div p_{i(d)}) \times p_{i(SA)})}{\sum_i(p_{i(SA)})} \times 100,000$$

217 Where d_i is the number of in-hospital deaths from the respective districts reported to DATCOV
 218 (wave 1: 5 March – 21 November 2020, wave 2: 22 November 2020 – 27 March 2021)⁵ in age group
 219 i , $p_{i(d)}$ is the district population for age group i and $p_{i(SA)}$ is the South African population for age
 220 group i .

221
$$\text{Excess deaths} = \frac{(ExD \times 0.85) \times p_{(SA)}}{p_{(SA)}} \times 100,000$$

222 Where ExD is the rate of provincial excess deaths adjusted to the South African population reported
 223 by South African Medical Research Council (rural wave 1: 28 June – 21 November 2020, urban wave
 224 1: 21 June – 21 November 2020, both communities wave 2: 22 November 2021 – 26 March 2021),⁶
 225 and $p_{(SA)}$ is the total South African population. According to estimates only 85% of excess deaths are
 226 attributable to COVID-19.⁶

227 For infections, hospitalisations, in-hospital (minimum) and excess (maximum) deaths, 95%
 228 confidence intervals (95% CI) were calculated using the Clopper-Pearson method with a Poisson
 229 distribution.

230
$$\text{Infection – case ratio (ICR)} = \frac{\text{Laboratory – confirmed cases}}{\text{Infections}} \times 100$$

231
$$\text{Infection – hospitalization ratio (IHR)} = \frac{\text{Hospitalisations}}{\text{Infections}} \times 100$$

232
$$\text{Infection – fatality ratio (IFR)} = \frac{\text{Deaths}}{\text{Infections}} \times 100$$

233 Confidence intervals for infection ratios were calculated as ratios from the 95% confidence intervals
234 of infection, hospitalisation and death rates. For international comparisons, we repeated the
235 calculations to standardise to the WHO standard world population,¹⁸ using Sprague multipliers¹⁹ to
236 expand age groups to one-year bands, and aggregate to the age groups used in this study.

237 *Comparison of cases between first and second wave of infection*

238 We further compared the characteristics of participants who seroconverted during the first wave of
239 infections, to those who seroconverted during the second wave of infections using unconditional
240 logistic regression. We compared site, age, sex, HIV status, CD4 count, viral load, other underlying
241 medial conditions, body mass index, employment status, smoking and alcohol use in participants
242 that seroconverted in wave 1 (blood draw 3) to those who seroconverted in wave 2 (blood draw 3,
243 excluding draw 3 seroconversions). For this analysis, we only included participants with a blood 3
244 and 5 paired serum sample. For the multivariable model we assessed all variables that were
245 significant at $p < 0.2$ on univariate analysis, and dropped non-significant factors ($p \geq 0.05$) with
246 manual backward elimination.

247 *Persistence of SARS-CoV-2 antibodies*

248 For all participants for which 5 sera were collected, and who seroconverted during draw 2 to 5, we
249 plotted COI values with the blood draw at which seroconversion took place as point 0. For those who
250 were seropositive at baseline we plotted the COI results from each blood draw. We calculated the
251 mean COI and exact 95% confidence interval at each point using the Clopper-Pearson method. We
252 assessed the percentage of participants with a COI ≥ 1.0 at each subsequent blood draw as the
253 number of participants with COI ≥ 1 divided by the seroconverted participants with a serum sample
254 at the time point.

255 **Results**

256 *Study population*

257 In the rural community, we approached 185 households, 118 (64%) were enrolled and 641/692
258 (92%) of household members consented and/or assented to participate. In the urban community,
259 352 households were approached, 114 (32%) enrolled and 570/607 (93%) of household members
260 consented and/or assented. In both communities, the percentage of children, females, and
261 unemployed individuals included in the cohort were higher than in district census data (Appendix
262 Table 1). Median age was 13 (IQR 7-29) and 21 (IQR 10-43) years, and HIV prevalence was 14%
263 (95%CI 11-17%) and 18% (95%CI 14-21%) in the rural and urban communities, respectively.

264 *Seroprevalence*

265 During blood draw 1-5, blood draw coverage was 67% - 88% and 84-88% in the rural and urban
266 communities, respectively (Appendix Table 2). The majority, 83% (n=553) of participants who lived in
267 the rural and 83% (n=499) in the urban community had both blood draw 3 and blood draw 5 blood
268 collected, with 56% (n=377) and 72% (n=431) of participants at the rural and urban site having all 5
269 bloods collected, respectively.

270 Seroprevalence, adjusted for assay sensitivity and specificity, in the rural community was lower at
271 blood draw 1 than in the urban community (1%, 95%CI 0-2% vs 15%, 95%CI 12-18%, $p<0.001$),
272 increasing after the first wave of infections (at blood draw 3) to 7% (95%CI 5-9) in the rural
273 community and 27% (23-31%) in the urban community ($p<0.001$, Figure 1, Appendix Table 3). After
274 the second wave (blood draw 5), seroprevalence increased by 19% to reach 26% (95%CI 22-29%,
275 $p<0.001$) in the rural community, and by 14% to reach 41% (95%CI 37-45%, $p<0.001$) in the urban
276 community (Figure 1, Appendix Table 3).

277 At blood draw 5, seroprevalence was highest in the 19-34 years age group (37%, 95% CrI 28-47%) in
278 the rural community and the 35-59 years age group (59%, 95%CrI 49-68%) in the urban community

279 (Figure 1, Appendix Table 3). The seroprevalence was lowest in children <5 years, 18% (95% CrI 10-
280 26%) and 28% (95% CrI 17-41%) in the rural and urban communities, respectively.

281 At blood draw 5, SARS-CoV-2 seroprevalence was similar between HIV-infected and HIV-uninfected
282 participants in all age groups and in both communities (Appendix Table 4). For adults aged 19-34
283 years, the seroprevalence in HIV-uninfected individuals in the rural and urban community was 34%
284 (21/66) and 36% (26/73), respectively, compared to HIV-infected individuals at 43% (10/23) and 37%
285 (7/19) in rural and urban community, respectively ($p=0.41$ and $p=0.92$, Appendix Table 4).

286 *Infection-case ratio (ICR), infection-hospitalisation ratio (IHR) and infection-fatality ratio (IFR) by*
287 *district and wave of infection*

288 Weekly incidence of reported laboratory-confirmed SARS-CoV-2 infections in the rural site district
289 peaked in the first wave at 68 cases per 100,000 population in week 31 of 2020, (starting 26 July),
290 and again at 165 cases per 100,000 in week 2 of 2021 (starting 10 January). In the urban site district,
291 the first wave peaked at 106 cases per 100,000 in week 30 of 2020 (starting 19 July), the second
292 wave at 79 cases per 100,000 in week 2 of 2021 (starting 10 January, Figure 1).

293 During the first wave of infections (blood draw 3) in the rural community, 40/553 participants had
294 seroconverted resulting in an age-adjusted seroprevalence of 10% (95%CrI 5-17%) or incidence of
295 10,041 (95%CrI 4,759-17,088) per 100,000 population. Within the rural district, standardised to the
296 South African population, 557 laboratory-confirmed SARS-CoV-2 infections per 100,000 population,⁴
297 75 COVID-19-related hospitalisations per 100,000 population, and 14 in-hospital deaths per 100,000
298 population⁵ were reported by end of week 47 of 2020 (starting 15 November). Excess deaths of 51
299 per 100,000 population were reported for Mpumalanga Province.⁶ Considering the 10%
300 seroprevalence at blood draw 3 in the rural community as a proxy for the district, the ICR was only
301 5.55% (95%CI 3.40-11.23%). There was a 0.75% (95%CI 0.49-1.41%) IHR and an in-hospital IFR of
302 0.13% (95%CI 0.10-0.23%) and an excess deaths IFR of 0.51% (95%CI 0.30-1.06%, Figure 2-3).

303 If excluding participants who had seroconverted after wave 1, there were 100/553 participants from
304 the rural community that had seroconverted at blood draw 5, an age-adjusted seroprevalence of
305 21% (95%CrI 13-30%) or incidence of 20,756 (95%CrI 13,007-30,011) per 100,000 population for the
306 second wave. The ICR was 4.00% (95%CI 2.86-6.17%), IHR 0.66% (0.50-0.98%), in-hospital IFR 0.20%
307 (95%CI 0.16-0.28%) and excess deaths IFR 0.70% (95%CI 0.49-1.12%, Figure 2-3).

308 In the urban community the age-adjusted seroprevalence at blood draw 3 was 29% (95%CrI 21-40%)
309 or incidence of 28,419 (95%CrI 20,902-38,840) per 100,000 population. There was a 3.68% (95%CI
310 2.82-4.79%) ICR and 2.01% (95%CI 1.57-2.57%) IHR. The in-hospital IFR was 0.17% (95%CI 0.15-
311 0.20%) and excess deaths IFR 0.13% (95%CI 0.10-0.17%, Figure 2-3). During the second wave, the
312 age-adjusted seroprevalence in the urban community was 15% (95%CrI 8%-23%) or incidence 15,132
313 (95%CrI 8,181-22,977) per 100,000 population, resulting in a ICR estimate of 3.69% (95%CI 2.59-
314 6.40%), IHR of 2.29% (95%CI 1.63-3.94%), in-hospital IFR of 0.36% (95%CI 0.28-0.57%) and an excess
315 deaths IHR of 0.51% (95%CI 0.34-0.93%, Figure 2-3). These estimates standardised to the WHO
316 world population are presented in Appendix Figure 2.

317 *Comparison of cases between first and second wave of infection*

318 Compared to the urban community, individuals in the rural community who seroconverted were 4.7
319 (95%CI 2.9-7.6) times more likely to seroconvert during the second wave. Compared to those aged
320 35-59 years, individuals aged 5-12 years and ≥ 60 years were 2.1 (95%CI 1.1-4.2) and 2.8 (95%CI 1.1-
321 7.0) times more likely to seroconvert in the second wave (Table 1).

322 *Persistence of SARS-CoV-2 antibodies*

323 Of the 72 participants seropositive at the baseline blood collection, and with blood draw 1-5 samples
324 collected, 99% (71/72) still had a COI ≥ 1 by blood draw 5. The mean COI at baseline for seropositive
325 participants was 64, which increased to 125 at blood draw 2 and reduced to 59 at blood draw 5
326 (Figure 4a). The participant who no longer had detectable SARS-CoV-2 antibodies at the fifth draw

327 had a starting COI of 9, which subsequently declined to 0.7 six months after the first draw. Of the
328 210 participants with blood draw 1-5 samples, 99% (167/169), 99% (70/71%) and 93% (41/44) still
329 had a COI ≥ 1 in the first, second and third blood draw after initial seroconversion, respectively
330 (Figure 4b). The participants who seroreverted had starting COIs ranging from 2 – 6. The mean COI at
331 the point of seroconversion was 48, which increased to 86 at the first blood post seroconversion and
332 reduced to 61 at the third draw post seroconversion.

333 **Discussion**

334 We assessed SARS-CoV-2 seroprevalence in 1,211 individuals living in two diverse communities in
335 South Africa and show that laboratory-confirmed cases reported from study districts greatly
336 underestimate the true burden of SARS-CoV-2 infections. At baseline, seroprevalence was 1% and
337 15%, increasing to 7% and 27% after the first wave, and by March 2021. Following the second
338 epidemic wave, seroprevalence was 26% and 41% in the rural and urban communities, respectively.
339 The highest seroprevalence was 59% in adults aged 35-59 years in the urban community, and the
340 lowest was 18% in rural community children <5 years. During the second wave, compared to the first
341 wave, the rural site was more affected, and infections in the second wave more likely affected
342 children aged 5-12 years and adults ≥ 60 years. We observed no differences in seroprevalence by HIV
343 status. In the urban community, the IFR was higher in the second wave (0.36-0.51%), compared to
344 the first (0.13-0.17%), even though numbers of infections were lower, suggesting possible increased
345 severity associated with the emergence of novel variant 501Y.V2. Most individuals who
346 seroconverted maintained detectable SARS-CoV-2 antibodies in subsequent serum samples.

347 Low seropositivity was observed at the rural site at baseline, and the seroprevalence remained low
348 until blood draw 3, only reaching seroprevalence of 7% after the first wave of infections, which was
349 considerably lower than the seroprevalence of 27% at the urban site at the same time. This could
350 possibly be related to the relatively isolated location and lower population density in the rural
351 community compared to more densely populated urban community. The seroprevalence in the rural

352 site increased to 26% at blood draw 5, which was after the second wave of infections within the
353 district. This could have been due to possible increased transmissibility of the 501Y.V2 lineage that
354 was circulating in the second wave,²⁰ as well as additional transmission networks in the community
355 during the December holiday period when large scale urban-to-rural migration takes place as people
356 return home for the yearend holidays. The urban site had fewer seroconversions in the second wave
357 compared to the first, which may be due to existing immunity among individuals in the community
358 after the first wave. As seen in previous studies,⁹ adults had the highest seroprevalence levels,
359 although there was still a relatively high seroprevalence of 18% and 28% in children <5 years at the
360 rural and urban community respectively.

361 There are few data available on seroprevalence in Africa, most studies have focused on specialised
362 groups. A study conducted among blood donors in South Africa during the second wave of the
363 pandemic found a seroprevalence of 32% to 63%, in five provinces of South Africa that have both
364 rural and urban communities.²¹ In our study, enrolling a random sample of community members, we
365 observed a seroprevalence in adults ranging between 25-37% in rural households, and 35-59% in the
366 urban households, suggesting that seroprevalence is heterogeneous between communities.

367 Seroprevalence data from sub-Saharan Africa are also limited. In Kenya the seroprevalence in blood
368 donors during the country's first wave of infections was 4%, and was also higher in urban
369 communities.²² In a population-level household sero-survey in Zambia during their first wave of
370 infections, 11% of individuals had evidence of SARS-CoV-2 infection.²³ After the first wave, we
371 estimated a higher seroprevalence in the rural (7%) and in the urban communities (27%) compared
372 to Kenya²² and a higher seroprevalence in the urban community than in Zambia.^{22,23}

373 Based on our estimates, only 4-6% of cases were laboratory confirmed. Our study suggests
374 substantially higher prevalence of infection ascertained through serology and that the differences
375 may have been greater in the urban community than the rural community; however, more extensive
376 studies are needed to assess whether this is consistent in other areas. Compared to the urban

377 community, the rural community had a higher proportion of SARS-CoV-2 infections reported to the
378 NMCSS (4-6% vs 4%) and less than half the rate of hospitalisation (0.8% vs 2-3%). These may be due
379 to differences in referral and testing policies, health-seeking behaviour, access to care and
380 differences in circulating lineages within these districts.

381 A study comparing the severity of the first and second waves of infections in South Africa in
382 hospitalised patients found a higher mortality rate in the second wave, compared to the first, even
383 controlling for increased pressure on health services which was also associated with increased
384 mortality.²⁴ At the urban site, the IFR was significantly higher in the second wave (0.43-0.50%)
385 compared to the first (0.13-0.20%), although no differences were observed in IHR between the two
386 waves. The lower overall number of infections in the second wave in this site means that our finding
387 of increased mortality is unlikely to be related to pressure on health services. The increased severity
388 of the second wave may be related to increased severity of the 501Y.V2/B.1.135 variant, but further
389 studies are needed to confirm this. The excess death IFR during the first wave in the urban site was
390 smaller than the in-hospital IFR. This may be due to uncertainty on the process for excess death
391 estimation, or that the 85% contribution of COVID-19 to excess deaths was an underestimation
392 within the province. However, the in-hospital IFR followed the same trend of increase between wave
393 1 and 2 (0.13% to 0.51%).

394 Considering the high seroprevalence observed, the age standardised IHR (wave 1: 0.8% rural, 2%
395 urban, wave 2: 0.8% rural, 2.58% urban) and IFR (wave 1: 0.2-0.5% rural, 0.1-0.2% urban, wave 2:
396 0.2-0.4% rural, 0.4-0.5% urban) were lower compared to the non-age standardised estimates from
397 the USA, where the IHR and IFR was estimated as 2% and 1% respectively,²⁵ and from Italy where IFR
398 was estimated as 0.9%.²⁶ Our first wave in-hospital IFR estimates (0.13% rural, 0.17% urban) were
399 more similar to the age-adjusted 0.15% reported from India for the first wave SARS-CoV-2-confirmed
400 deaths,²⁷ and our first wave excess death IFR was higher in the rural (0.51%) and lower in the urban

401 (0.13%) community compared to the age-adjusted 0.28% IFR excess deaths reported from Brazil
402 during their first wave of infections.²⁸

403 We did not observe a difference in SARS-CoV-2 seroprevalence in HIV-infected and –uninfected
404 individuals at either site. Due to HIV causing immune suppression, there is a concern that HIV-
405 infected individuals may be more susceptible to SARS-CoV-2 infection.^{10,11} A meta-analysis including
406 data from North America, Africa, Europe, and Asia found a 24% higher risk of symptomatic SARS-
407 CoV-2 infection in HIV-infected individuals, albeit with high heterogeneity between countries,¹¹
408 whereas a case-control study from the United States of America found a lower SARS-CoV-2
409 seroprevalence in HIV-infected individuals. Although HIV infection may not increase susceptibility to
410 infection, it has been demonstrated to be a risk factor for developing severe COVID-19 and death
411 following infection.^{10,29}

412 Our study is limited by a relatively small sample size, reducing the power for accurate
413 seroprevalence estimates in small age strata; and including only two geographic sites, and therefore
414 may not be representative of other districts and provinces in South Africa. The seroprevalence
415 reported here may be an under-estimate of population-level infection rates as not everyone infected
416 with SARS-CoV-2 develops antibodies. ICR, IHR and IFR formed part of an ecological analysis which is
417 inherently prone to biases. Excess deaths in the first wave may be underestimated since the
418 reporting period only started in June, and were reported at provincial-level which may be different to
419 within the district. Similarly, transmission dynamics within our cohort may not be similar to the
420 district which formed the comparison point for our case, hospitalisation and in-hospital deaths. The
421 ELISA utilised to detect SARS-CoV-2 antibodies was qualitative, and not suited for quantitative
422 analysis of antibody levels. Ongoing follow-up of this cohort will track future infections and monitor
423 antibody waning, and compare these data to laboratory confirmed infections and symptoms from
424 twice weekly follow-up. A strength of our study is the collection of samples from prospectively
425 followed up individuals from randomly selected households within the study communities and

426 inclusion of individuals of all ages. To our knowledge, this is the first seroprevalence data from a
427 cohort in South Africa, which provides the advantage of serial comparisons of antibody responses in
428 relation to reported laboratory-confirmed SARS-CoV-2 infections within the community through two
429 successive SARS-CoV-2 waves.

430 We estimate that approximately 95% of SARS-CoV-2 infections in these two communities were not
431 laboratory-confirmed and reported to the national surveillance system, which has major implications
432 for contact tracing and isolation and other measures to contain infection. We observed
433 heterogeneity between seroprevalence estimates based on pandemic wave, community and age
434 group indicating the need for ongoing studies with inclusion of diverse settings.

435

436 **References**

- 437 1. National Institute for Communicable Diseases. COVID-19 Testing Summary. 12 February 2021
438 2021. [https://www.nicd.ac.za/wp-content/uploads/2021/03/COVID-19-Testing-Summary-Week-8-](https://www.nicd.ac.za/wp-content/uploads/2021/03/COVID-19-Testing-Summary-Week-8-2021.pdf)
439 [2021.pdf](https://www.nicd.ac.za/wp-content/uploads/2021/03/COVID-19-Testing-Summary-Week-8-2021.pdf)16 February 2021).
- 440 2. South African Government. COVID-19 Risk Adjusted Strategy. 2020.
441 <https://sacoronavirus.co.za/covid-19-risk-adjusted-strategy/> (accessed 16 February 2021 16
442 February 2021).
- 443 3. Salyer SJ, Maeda J, Sembuche S, et al. The first and second waves of the COVID-19 pandemic
444 in Africa: a cross-sectional study. *The Lancet* 2021; **397**(10281): 1265-75.
- 445 4. National Institute for Communicable Diseases. COVID-19 Weekly Epidemiology Brief Week
446 19. 22 May 2021 2021. [https://www.nicd.ac.za/wp-content/uploads/2021/05/COVID-19-Weekly-](https://www.nicd.ac.za/wp-content/uploads/2021/05/COVID-19-Weekly-Epidemiology-Brief-week-19-2021.pdf)
447 [Epidemiology-Brief-week-19-2021.pdf](https://www.nicd.ac.za/wp-content/uploads/2021/05/COVID-19-Weekly-Epidemiology-Brief-week-19-2021.pdf)22 May 2021).
- 448 5. National Institute for Communicable Diseases. COVID-19 Sentinel Hospital Surveillance
449 Update Week 18. 14 May 2021 2021. [https://www.nicd.ac.za/wp-content/uploads/2021/05/NICD-](https://www.nicd.ac.za/wp-content/uploads/2021/05/NICD-COVID-19-Weekly-Sentinel-Hospital-Surveillance-update-Week-18-2021.pdf)
450 [COVID-19-Weekly-Sentinel-Hospital-Surveillance-update-Week-18-2021.pdf](https://www.nicd.ac.za/wp-content/uploads/2021/05/NICD-COVID-19-Weekly-Sentinel-Hospital-Surveillance-update-Week-18-2021.pdf)22 May 2021).

- 451 6. South African Medical Research Council. Report on Weekly Deaths in South Africa Week 18.
452 10 May 2021 2021. <https://www.samrc.ac.za/reports/report-weekly-deaths-south-africa> (accessed
453 22 May 20021).
- 454 7. Wibmer CK, Ayres F, Hermanus T, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South
455 African COVID-19 donor plasma. *bioRxiv : the preprint server for biology* 2021: 2021.01.18.427166.
- 456 8. Murhekar MV, Clapham H. COVID-19 serosurveys for public health decision making. *The*
457 *Lancet Global Health* 2021; **9**(5): e559-e60.
- 458 9. Chen X, Chen Z, Azman AS, et al. Serological evidence of human infection with SARS-CoV-2: a
459 systematic review and meta-analysis. *The Lancet Global Health* 2021; **9**(5): e598-e609.
- 460 10. Spinelli MA, Lynch KL, Yun C, et al. SARS-CoV-2 seroprevalence, and IgG concentration and
461 pseudovirus neutralising antibody titres after infection, compared by HIV status: a matched case-
462 control observational study. *Lancet HIV* 2021: S2352-3018(21)00072-2.
- 463 11. Ssentongo P, Heilbrunn ES, Ssentongo AE, et al. Epidemiology and outcomes of COVID-19 in
464 HIV-infected individuals: a systematic review and meta-analysis. *Sci Rep* 2021; **11**(1): 6283-.
- 465 12. Cohen C, McMorrow ML, Martinson NA, et al. Cohort Profile: a Prospective Household
466 cohort study of Influenza, Respiratory Syncytial virus, and other respiratory pathogens community
467 burden and Transmission dynamics in South Africa (PHIRST), 2016-2018. *medRxiv* 2021:
468 2021.01.06.21249313.
- 469 13. Cohen C, Kleynhans J, Moyes J, et al. Asymptomatic transmission and high community
470 burden of seasonal influenza in an urban and a rural community in South Africa, 2017–18
471 (PHIRST): a population cohort study. *The Lancet Global Health* 2021; **9**(6): e863-e74.
- 472 14. South African Government. COVID-19: Vaccine Strategy. 2021.
473 <https://sacoronavirus.co.za/2021/01/03/covid-19-vaccine-strategy/> (accessed 16 February 2021 16
474 February 2021).
- 475 15. Larremore DB, Fosdick BK, Bubar KM, et al. Estimating SARS-CoV-2 seroprevalence and
476 epidemiological parameters with uncertainty from serological surveys. *Elife* 2021; **10**: e64206.

- 477 16. Roche. Elecsys® Anti-SARS-CoV-2. 2020.
478 <https://diagnostics.roche.com/global/en/products/params/elecsys-anti-sars-cov-2.html> (accessed 21
479 March 2021 21 March 2021).
- 480 17. Statistics South Africa. Statistical release P0302 - Mid-year population estimates 2020. 2020.
481 http://www.statssa.gov.za/?page_id=1854&PPN=P0302&SCH=7263414 May 2021).
- 482 18. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of
483 rates: a new WHO standard. [Geneva]: World Health Organization; 2001.
- 484 19. Shryock HS, Larmon EA, Siegel JS, U.S. Department of Commerce BotC. The methods and
485 materials of demography. Washington: U.S. Gov. Print. Off.; 1980.
- 486 20. Tegally H, Wilkinson E, Giovanetti M, et al. Emergence and rapid spread of a new severe
487 acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike
488 mutations in South Africa. *medRxiv* 2020: 2020.12.21.20248640.
- 489 21. Wendy S, Laurette M, Ronel S, et al. Prevalence of anti-SARS-CoV-2 antibodies among blood
490 donors in Northern Cape, KwaZulu-Natal, Eastern Cape, and Free State provinces of South Africa in
491 January 2021. *Research Square* 2021.
- 492 22. Uyoga S, Adetifa IMO, Karanja HK, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in
493 Kenyan blood donors. *Science (New York, NY)* 2021; **371**(6524): 79-82.
- 494 23. Mulenga LB, Hines JZ, Fwoloshi S, et al. Prevalence of SARS-CoV-2 in six districts in Zambia in
495 July, 2020: a cross-sectional cluster sample survey. *The Lancet Global Health* 2021.
- 496 24. Jassat W, Mudara C, Ozougwu L, et al. Increased mortality among individuals hospitalised
497 with COVID-19 during the second wave in South Africa. *medRxiv* 2021: 2021.03.09.21253184.
- 498 25. Angulo FJ, Finelli L, Swerdlow DL. Estimation of US SARS-CoV-2 Infections, Symptomatic
499 Infections, Hospitalizations, and Deaths Using Seroprevalence Surveys. *JAMA network open* 2021;
500 **4**(1): e2033706-e.

- 501 26. Pota M, Pota A, Sirico ML, Esposito M. SARS-CoV-2 Infections and COVID-19 Fatality:
502 Estimation of Infection Fatality Ratio and Current Prevalence. *Int J Environ Res Public Health* 2020;
503 **17**(24): 9290.
- 504 27. Murhekar MV, Bhatnagar T, Selvaraju S, et al. Prevalence of SARS-CoV-2 infection in India:
505 Findings from the national serosurvey, May-June 2020. *The Indian journal of medical research* 2020;
506 **152**(1 & 2): 48-60.
- 507 28. Silva AAMd, Lima-Neto LG, Azevedo CdMPESd, et al. Population-based seroprevalence of
508 SARS-CoV-2 and the herd immunity threshold in Maranhão. *Rev Saude Publica* 2020; **54**: 131-.
- 509 29. Western Cape Department of Health in collaboration with the National Institute for
510 Communicable Diseases SA. Risk Factors for Coronavirus Disease 2019 (COVID-19) Death in a
511 Population Cohort Study from the Western Cape Province, South Africa. *Clin Infect Dis* 2020.

512 **Acknowledgements**

513 All individuals participating in the study, field teams for their hard work and dedication to the study,
514 especially Dr Tumelo Moloantoa, Dr Kgaugelo Kgasago, Nompumelelo Yende and Sizzy Ngobeni for
515 coordination of site activities, Charlotte Yaldwyn and Shingirai Chipungu for processing of collected
516 bloods, the CRDM laboratory team, Doreen Janse van Rensburg for managing logistics surrounding
517 testing of collected bloods and result extraction, Nevashan Govender, Genevieve Ntshoe, Andronica
518 Moipone Shonhiwa, Darren Muganhiri, Itumeleng Matiea, Eva Mathatha, Fhatuwani Gavhi, Teresa
519 Mashudu Lamola, Matimba Makhubele, Mmaborwa Matjokotja, Simbulele Mdleleni, Masingita
520 Makhubela from the national SARS-CoV-2 NICD surveillance team, Lincoln Darwin, Fazil Mckenna, Dr
521 Trevor Graham Bell, Ndivhuwo Munava, Muzammil Raza Bano and Jimmy Khosa from NICD IT, Prof
522 Rob Dorrington from the Centre for Actuarial Research for providing wave-specific excess death
523 estimates, Ben Cowling, Kanta Subbarao, Juliet Pulliam, Melissa Rolfes and Mosa Moshabela from
524 the PHIRST-C scientific committee and Andrew Whitelaw, June Fabian, Jennifer Verani, Lindiwe
525 Qwabe, Banele Faku and Saheen Methar from the PHIRST-C safety committee.

526 **Funding**

527 This work was supported by the National Institute for Communicable Diseases of the National Health
528 Laboratory Service and the US Centers for Disease Control and Prevention (cooperative agreement
529 number 6U01IP001048-04-02).

530 **Data availability**

531 The investigators welcome enquiries about possible collaborations and requests for access to the
532 dataset. Data will be shared after approval of a proposal and with a signed data access agreement.
533 Investigators interested in more details about this study, or in accessing these resources, should
534 contact the corresponding author.

535 **Disclaimer**

536 The findings and conclusions in this report are those of the authors and do not necessarily represent
537 the official position of the CDC.

538 **Potential conflicts of interest**

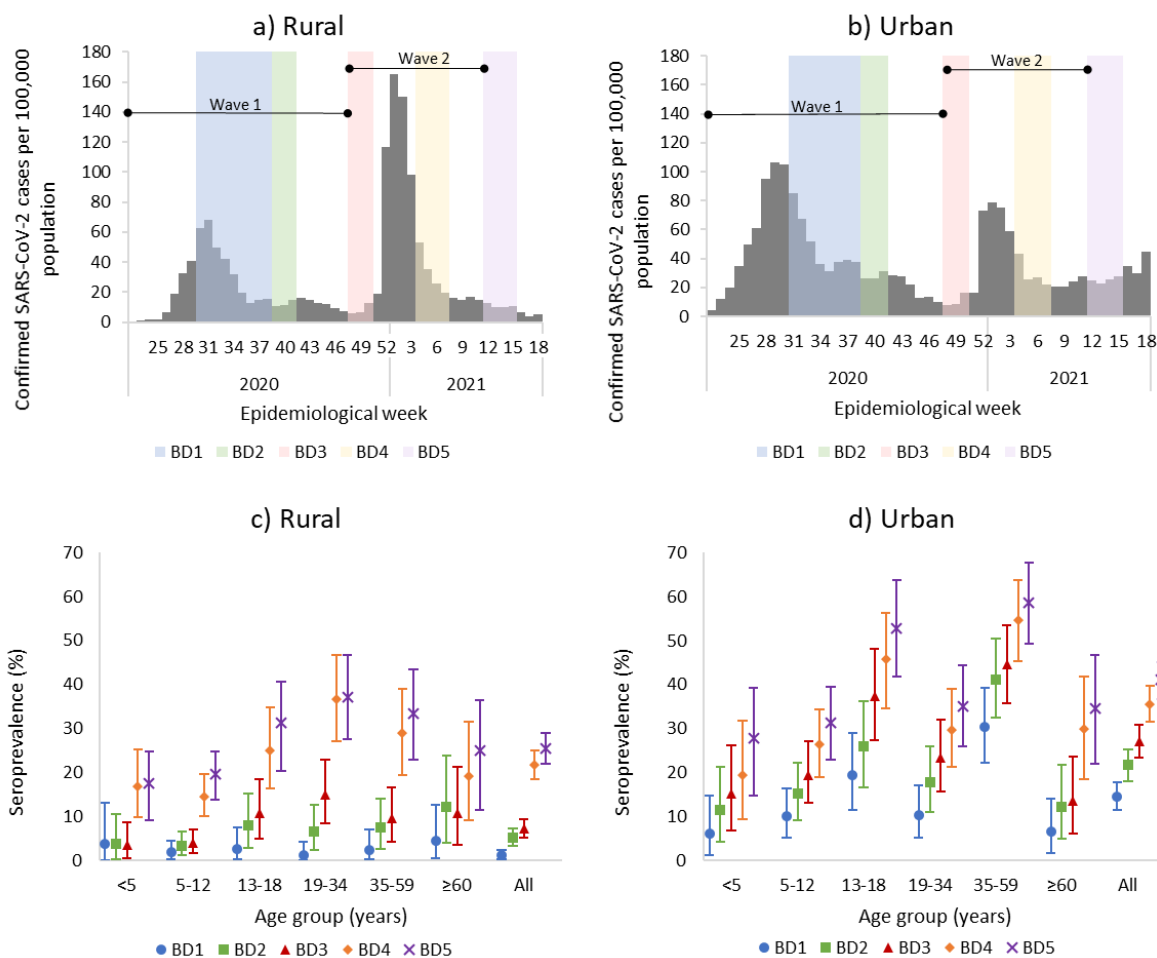
539 Cheryl Cohen reports receiving grant funds from US-Centers for Disease Control and Prevention,
540 Wellcome Trust and South African Medical Research Council.

541

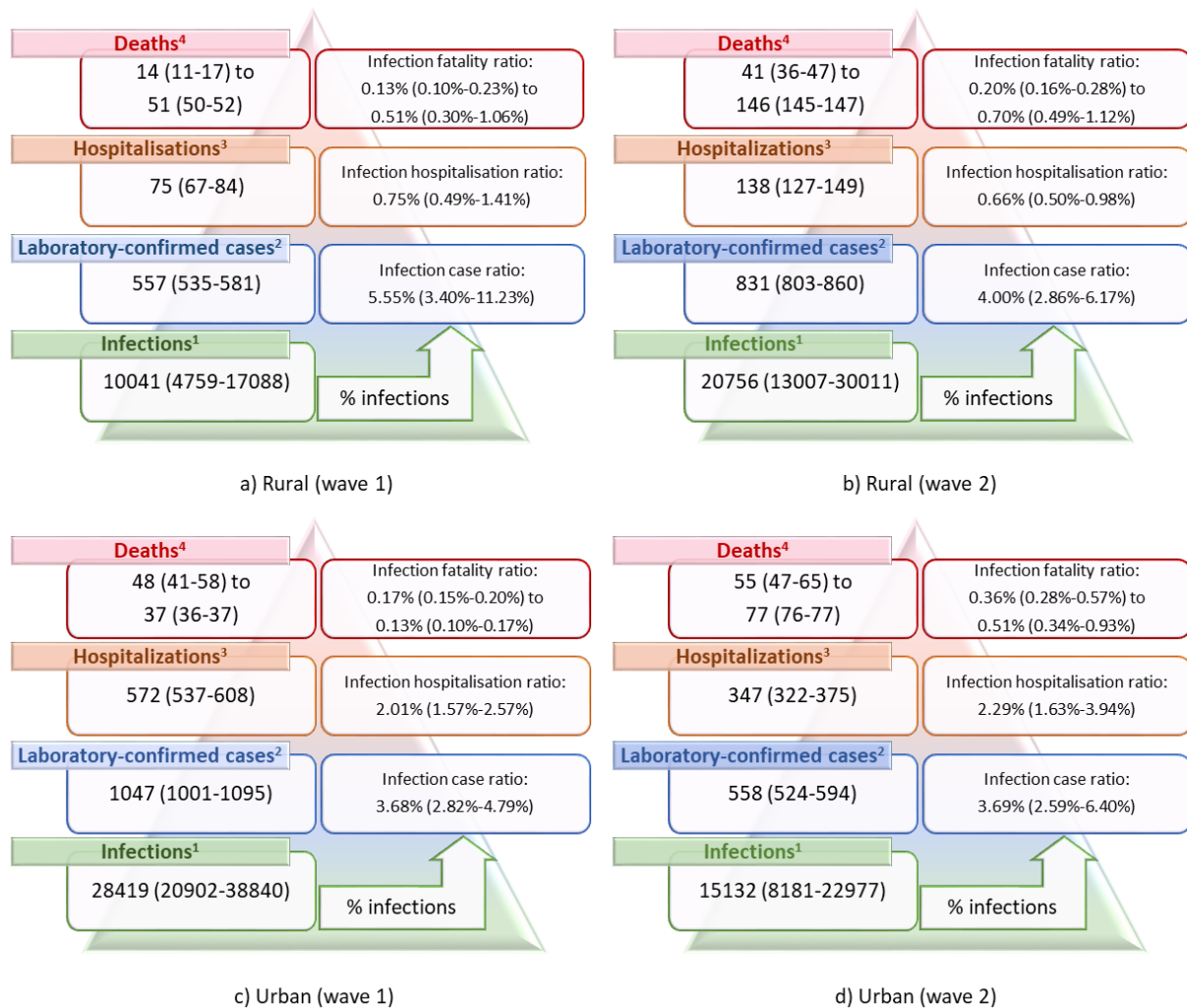
542 **Table**543 **Table 1.** Comparison of participants with detectable SARS-CoV-2 antibodies after the first wave
544 (blood draw 3) and second wave (blood draw 5), July 2020 – April 2021, South Africa.

	Infected in wave 2 (n/N %)	Univariate OR	95% CI	Multivariable aOR	95% CI
Site					
Rural	100/140 (71)	4.9	3.1-7.8	4.7	2.9-7.6
Urban	71/210 (34)	Ref		Ref	
Sex					
Male	58/123 (47)	Ref			
Female	113/227 (50)	1.1	0.7-1.7		
Age group (years)					
<5 years	16/25 (64)	3.2	1.3-8.0	2.7	1.0-7.2
5-12 years	44/74 (59)	2.6	1.4-4.9	2.1	1.1-4.2
13-18 years	28/64 (44)	1.4	0.7-2.7	1.3	0.6-2.6
19-34 years	33/67 (49)	1.7	0.9-3.3	1.3	0.7-2.6
35-59 years	33/92 (36)	Ref		Ref	
≥60 years	17/28 (61)	2.8	1.2-6.6	2.8	1.1-7.0
HIV status					
Negative	132/271 (49)	1.0	0.6-1.7		
Positive	33/68 (49)	Ref			
CD4 count					
≥200	26/54 (48)	1.9	0.2-21.7		
<200	1/3 (33)	Ref			
Viral load					
<1000	23/51 (45)	Ref			
≥1000	5/7 (71)	3.0	0.5-17.2		
Other underlying illness*					
No	155/316 (49)	1.1	0.5-2.2		
Yes	16/34 (47)	Ref			
Body-mass index					
Underweight	12/22 (55)	1.5	0.6-3.9		
Normal weight	76/141 (54)	1.5	0.9-2.5		
Overweight	38/84 (45)	1.1	0.6-1.9		
Obese	45/103 (44)	Ref			
Currently smoking†					
No	81/190 (43)	Ref			
Yes	16/36 (44)	1.1	0.5-2.2		
Alcohol use‡					
No	79/167 (47)	2.0	1.1-3.8		
Yes	18/59 (31)	Ref			
Employment status‡					
Unemployed	58/128 (45)	1.9	0.5-6.4		
Student	4/13 (31)	Ref			
Employed	16/41 (39)	1.4	0.4-5.5		

545 Includes all participants with blood draw 3 and 5 sera pairs, and who had seroconverted at either draw. * Self-reported history of asthma,
546 lung disease, heart disease, stroke, spinal cord injury, epilepsy, organ transplant, immunosuppressive therapy, organ transplantation,
547 cancer, liver disease, renal disease, or diabetes. † Individuals ≥15 years. ‡ Individuals ≥18 years.



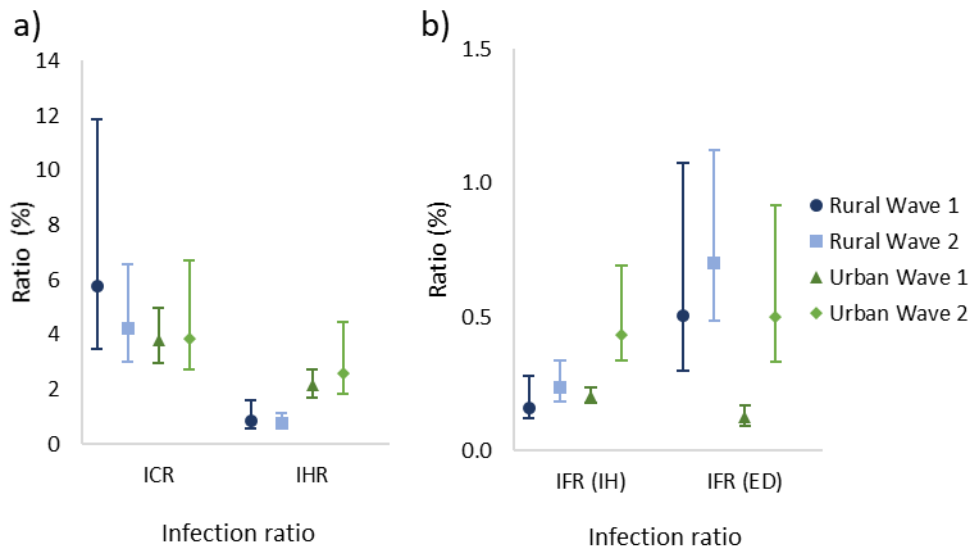
549
 550 **Figure 1.** Timing of blood collection and district SARS-CoV-2 weekly incidence¹ in a) rural community
 551 and b) urban community, and seroprevalence at each blood collection by age group in the c) rural
 552 and d) urban community, March 2020 – March 2021, South Africa.
 553 Blood draw (BD) 1: 20 July – 17 September 2020, blood draw 2: 21 September – 10 October 2020,
 554 blood draw 3: 23 November – 12 December 2020, blood draw 4 25 January – 20 February 2021,
 555 blood draw 5: 22 March – 11 March 2021. Vertical lines represent 95% credible interval.
 556 Seroprevalence estimates adjusted for sensitivity and specificity of assay. Wave 1 (1 March – 21
 557 November 2020) and 2 (22 November 2020 – 27 March 2021) lines indicate period used for analysis.
 558 ¹Laboratory-confirmed SARS-CoV-2 infections (reverse transcription polymerase chain reaction and
 559 antigen tests) reported to the Notifiable Medical Conditions Surveillance System (NMCCS).



560

561 **Figure 2.** South African age-standardised SARS-CoV-2 infection, diagnosis, hospitalisation and deaths
 562 per 100,000 population in a rural community during a) wave 1 and b) wave 2, and urban community
 563 during c) wave 1 and d) wave 2 of infections, March 2020 - March 2021, South Africa.

564 ¹Based on seroprevalence at third/fifth blood draw. ²Laboratory confirmed SARS-CoV-2 cases
 565 reported from districts from national SARS-CoV-2 surveillance (NMCSS, wave 1: 3 March – 21
 566 November 2020, wave 2: 22 November 2020 – 27 March 2021). ³Hospitalisations in the district
 567 based on COVID-19 Sentinel Hospital Surveillance (DATCOV, wave 1: 5 March – 21 November 2020,
 568 wave 2: 22 November 2020 – 27 March 2021). ⁴Minimum estimate: in-hospitalisation deaths in
 569 districts based on COVID-19 Sentinel Hospital Surveillance report (DATCOV, wave 1: 5 March – 21
 570 November 2020, wave 2: 22 November 2020 – 27 March 2021), maximum estimate: provincial
 571 excess deaths reported by South African Medical Research Council (rural wave 1: 28 June – 21
 572 November 2020, urban wave 1: 21 June – 21 November 2020, wave 2: 22 November 2021 – 26
 573 March 2021). ⁵Standardised to South Africa mid-year population estimate for 2020. Wave 1: 1 March
 574 – 21 November 2020), wave 2 22 November 2020 – 27 March 2021. Values in bracket refer to 95%
 575 credible interval for infections and 95% confidence interval for all other estimates.



576

577

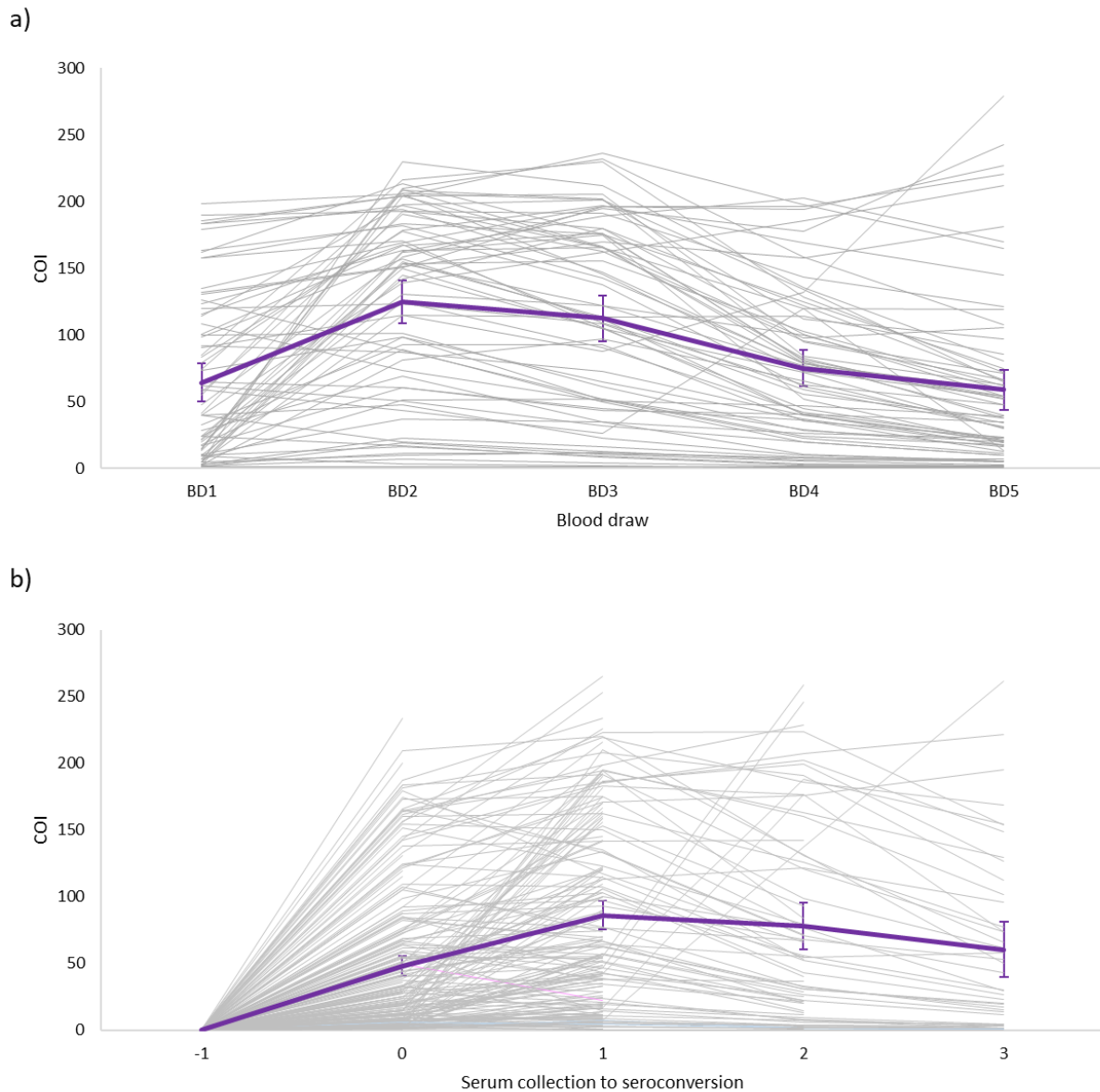
578

579

580

581

Figure 3. SARS-CoV-2 a) infection-case (ICR), infection-hospitalisation (IHR) and b) in-hospital (IH) infection-fatality (IFR) and excess (ED) infection-fatality (IFR) ratios in a rural and urban community during the first and second wave of infections, March 2020 - March 2021, South Africa. Vertical lines represent 95% confidence interval. Wave 1: 1 March – 21 November 2020, wave 2: 22 November 2020 – 27 March 2021.



582
 583
 584
 585
 586
 587

Figure 4. Cut-off index (COI) on Roche Elecsys® Anti-SARS-CoV-2 assay for individuals with blood draw 1 to 5 (BD1 to BD5) samples who a) were seropositive at baseline b) seroconverted during blood draw 2-5, July 2020 – April 2021, South Africa. Mean COI with 95% confidence interval in purple line. COI values in b aligned to first draw prior to seroconversion.