Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors

Short Title: SARS-CoV-2 Sero-surveillance of blood donors

One Sentence Summary: 1 in 25 Kenyan blood donors aged 15-64 had SARS-CoV-2 antibodies in May 2020, when fewer than 100 COVID-19 deaths had been reported nationally.

Authors: Sophie Uyoga[1*§], Ifedayo M.O. Adetifa[1,2§], Henry K. Karanja[1§], James Nyagwange[1§], James Tuju[1], Perpetual Wanjiku[1], Rashid Aman[3], Mercy Mwangangi[3], Patrick Amoth[3], Kadondi Kasera[3], Wangari Ng'ang'a[4], Charles Rombo[5], Christine Yegon[5], Khamisi Kithi[5], Elizabeth Odhiambo[5], Thomas Rotich[5], Irene Orgut[5], Sammy Kihara[5], Mark Otiende[1], Christian Bottomley[2], Zonia N. Mupe[1], Eunice W. Kagucia[1], Katherine E. Gallagher[1,2], Anthony Etyang[1], Shirine Voller[1,2], John N. Gitonga[1], Daisy Mugo[1], Charles N. Agoti[1], Edward Otieno[1], Leonard Ndwiga[1], Teresa Lambe[6], Daniel Wright[6], Edwine Barasa[1], Benjamin Tsofa[1], Philip Bejon[1,6], Lynette I. Ochola-Oyier[1], Ambrose Agweyu[1†], J. Anthony G. Scott[1,2†], George M. Warimwe[1,6†]

Affiliations:

[1] KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya.

[2] Department of Infectious Diseases Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London, United Kingdom.

[3] Ministry of Health, Government of Kenya, Nairobi, Kenya.

[4] Presidential Policy & Strategy Unit, The Presidency, Government of Kenya.

[5] Kenya National Blood Transfusion Services, Ministry of Health, Nairobi, Kenya.

[6] Nuffield Department of Medicine, Oxford University, United Kingdom.

*Correspondence to: SUyoga@kemri-wellcome.org

Abstract:

The spread of SARS-CoV-2 in Africa is poorly described. The first case of SARS-CoV-2 in Kenya was reported on March 12, 2020 and an overwhelming number of cases and deaths were expected but by July 31, 2020 there were only 20,636 cases and 341 deaths. However, the extent of SARS-CoV-2 exposure in the community remains unknown. We determined the prevalence of anti-SARS-CoV-2 IgG among blood donors in Kenya in April-June 2020. Crude seroprevalence was 5.6% (174/3098). Population-weighted, test-performance-adjusted national seroprevalence was 4.3% (95% CI 2.9–5.8%) and was highest in urban counties, Mombasa (8.0%), Nairobi (7.3%) and Kisumu (5.5%). SARS-CoV-2 exposure is more extensive than indicated by case-based surveillance and these results will help guide the pandemic response in Kenya, and across Africa.

Main Text:

Africa accounts for 17% of the global population (*1*) but by late July 2020 accounted for only 5% of the global COVID-19 cases and 3% of global COVID-19 deaths reported (*2*). This disparity has been attributed to limited capacity for diagnosis, timely implementation of stringent containment measures, a younger population structure and a predominance of asymptomatic and mild infections (*3, 4*). The first case of COVID-19 in Kenya was detected on March 12, 2020. Within one week the government instituted containment measures to limit the spread of the virus (*5*). By July 31, national surveillance recorded 20,636 cases and 341 deaths (*6*). This increase in cases is notably slower than the epidemic in Wuhan, Europe or the USA. Recently, it has been suggested that "the virus is spreading… …with an attenuated outcome in Africa" but there are few data available to confirm or refute this assertion (*7*).

In countries affected early in the pandemic, serological surveillance was used to define cumulative incidence. For example, at the release of lockdown in Wuhan, 9.6% of staff resuming work were found to have anti-SARS-CoV-2 antibodies (*8*). At the end of the epidemic wave in Spain, seropositivity was 5.0% in a random population sample of 60,897 (*9*). As the epidemic curve declined in Geneva, seroprevalence rose over three weeks from 4.8% to 10.9% (*10*). Currently, there are few estimates of SARS-CoV-2 seroprevalence in Africa in the literature (*11*).

Movement restrictions, in response to COVID-19, have limited the conduct of fieldwork for population-based serosurveys. Several countries have monitored seroprevalence in blood transfusion donors (*12, 13*) or expectant mothers attending ante-natal clinics (*14*). Here we report

the results of a pragmatic national serosurvey using residual blood samples from transfusion donors across Kenya and a highly sensitive and specific assay for anti-SARS-CoV-2 spike immunoglobulin G (IgG).

We validated a widely-used enzyme linked-immunosorbent assay for SARS-CoV-2 IgG (*15*) with 910 serum samples from the pre-pandemic period and 174 sera from polymerase chain reaction (PCR) defined SARS-CoV-2 cases, and a well-characterized 5 sera panel from the National Institute of Biological Standards and Control (NIBSC) in the UK. For either receptor-binding domain (RBD) or whole spike, specificity was higher when using a ratio of the sample optical density (OD)/negative control OD than when using the raw sample OD plus 3 standard deviations to define seropositivity (Table S1). By using OD ratios, both RBD and spike ELISAs correctly classified 901 of 910 pre-pandemic samples as seronegative (Table S1). However, the spike ELISA detected more seropositives (166 of 179 vs compared to 145 of 179 for RBD ELISA) among sera from SARS-CoV-2 PCR-positive individuals (Figure S2, panels A & B). Based on these data, we defined anti-SARS-CoV-2 IgG seropositivity as an OD ratio >2 and selected the spike ELISA for this study. The sensitivity and specificity, at this threshold, were 92.7% (95% CI 87.9-96.1%) and 99.0% (95% CI 98.1-99.5%), respectively (Table S1, Figure S3 panels A & B, Figure S5 and S6). As previously noted (*15*), the RBD and whole spike ELISA responses were highly correlated (Figure S3, panel C), with very little inter-assay variation (Figure S4).

A total of 3,174 blood transfusion samples were collected from four Kenya National Blood Transfusion Service (KNBTS) regional blood transfusion centers that are supported by several satellites and hospitals between April 30 and June 16, 2020, from individuals aged 15-66 years.

Approximately half of the samples were drawn in Mombasa; the remainder were evenly distributed between Nairobi, Kisumu and Eldoret (Figure 1, Table S2). We excluded 18 duplicate samples, 56 records missing data on age or collection date and two records from individuals aged ≥65 years. Policy in Kenya is to avoid blood donation from individuals >65 years, and we excluded these other data points as potentially unreliable. These exclusions left 3,098 samples for further analysis (Figure 1).

Of the 3,098 samples, 174 were positive for anti-SARS-CoV-2 Spike IgG giving a crude seroprevalence of 5.6% (95% CI 4.8–6.5%). Crude seroprevalence varied by age (p=0.046), ranging between 3.4-7.0% among adults 15-54 years; all 71 donors aged 55-64 years were seronegative (Table 1). Crude seroprevalence did not vary by sex (p=0.50) but did vary geographically, from 1.9% in the Rift Valley region to 10.0% in the Western region (p=0.002, Table 1).

Compared to the 2019 Kenya Population and Housing Census, our participants were more commonly male (82.0% in our study vs 49.3% in the census), had more persons aged 25-34 years (40.1% vs 27.3%) and more residents of coastal Counties (49.2% versus 9.1%, Table 2). We therefore adjusted the prevalence estimate for the demographics of the sample using post-stratification, and for the sensitivity and specificity of the test.

The Bayesian population-weighted and test-adjusted seroprevalence for Kenya was 4.3% (95% CI 2.9-5.8%, Table 1) and the posterior sensitivity and specificity estimates were 92.4% (95% CI 88.0-95.6%) and 98.9 (95% CI 98.2-99.5%), respectively. Seroprevalence was higher (4.2-5.2%) in the younger age groups (15-44 years) and declined in the older age groups (45-64 years) but

was similar for both sexes. Seroprevalence was highest for those living in Mombasa, Nairobi and the Western region, although the number of observations for the Western region was small. The directly standardized seroprevalence estimates are presented in Table S3. Seroprevalence was also calculated for Counties that had at least 120 donors sampled. The three largest urban Counties of Mombasa, Nairobi, and Kisumu had SARS-CoV-2 seroprevalence of 8.0% (95% CI 5.5-11.1%), 7.3% (95% CI 4.2-11.4%) and 5.5% (95% CI 2.8-9.6%), respectively (Table S4).

The frequency of blood donor sampling and crude seroprevalence estimates increased with time over the 7-week study period (Figure 2). The median sample date was May 30, 2020 while the mid-point of the study was May 24, 2020. We did not adjust for sample date because the period of sampling varied for residents of different counties (Figure 2C); instead we show the variation in crude prevalence over time (Figure 2A).

Voluntary non-remunerated donors (VNRDs), who donate blood at community-based 'blood drives' comprised only 7.6% (236/3098) of our sample of donors; the remainder were family replacement donors (FRDs) who provide a unit of blood in compensation for a transfusion received by a sick relative. The two groups did not differ significantly by age (p=0.15) or sex (p=0.51, Table S5). Crude seroprevalence was 8.5% (20/236) for VNRDs and 5.4% (154/2862) for FRDs. The median sample date for VNRDs (June 14, 2020) was two weeks later than that for FRDs (May 29, 2020).

Population exposure across Kenya, with a population-weighted test-adjusted seroprevalence of 4.3%, is considerably higher than was previously thought, based on the cases and deaths reported to date. Seroprevalence was particularly high in the three urban counties; Mombasa (8.0%),

Nairobi (7.3%) and Kisumu (5.5%). Consistent with other studies, seroprevalence did not vary significantly by sex; (*9, 10, 16*) however, it peaked in 35-44-year-olds and was lowest for those ≥45 years, which is also consistent with existing reports where seroprevalence was found to be lower in older adults (*9, 10*).

SARS-CoV-2 seroprevalence in our study is comparable to estimates from large population-based serosurveys in China, Switzerland, Spain and the USA after the initial epidemic peak and following many tens of thousands of deaths (*9, 10, 17, 18*). Our results are also comparable to other surveys of blood donors in Brazil (*13*), Italy (*12*), and many parts of England (*19*). Kenya has an estimated population of 53 million in 2020 and 57% of the population is aged 15-64 years. If the transfusion donor seroprevalence of 4.3% was applied to all 15-64-year-olds it would suggest approximately 1.3 million infections. However, by the median sample date, May 30, 2020, only 2093 cases had been detected (of which approximately 90% were asymptomatic) and 71 deaths among all ages (*6*). Although it is difficult to extrapolate our data directly to the whole population, they do strongly suggest that the infection is more widespread in Kenya than the current PCR test results suggest and indicate a need for more systematic testing. The current PCR testing strategy targets symptomatic individuals, health care workers, contacts of confirmed cases, international travelers, cross border truck drivers and residents of areas identified as hot spots."

What are the potential explanations for the divergence in the ratio of observed cases or deaths to serologically defined infections inferred from transfusion donors in Kenya, compared to many high-income countries? (i) The seroprevalence could be over-estimated because of bias in the selection or behavior of blood transfusion donors. (ii) Cases could be under-ascertained by

national public health surveillance though it seems unlikely that reporting of deaths and severe cases could be reduced by several orders of magnitude, and hospitals in Kenya were not overwhelmed by admissions with respiratory illness. (iii) The steep demographic age-pyramid results in a smaller vulnerable age group. In Kenya, only 3.9% of the population is aged 65 years or greater which is substantially less than, for example, 23.3% found in Italy; again, this would only explain a several-fold reduction in severe cases or deaths (*4*). (iv) There may be alternative mechanisms of immunity to SARS-CoV-2 including cell-mediated immunity (*20, 21*) perhaps as a result of HCoV-elicited immunity (*22, 23*). Despite our prior work showing HCoVs circulate in Kenya (*24*), we did not identify evidence of cross-reactive antibodies to endemic coronaviruses in our validation study.

Although blood donors are not representative of the Kenyan population as a whole, we adjusted for demographic bias in the sample structure by standardization against the age, sex, and regional distribution of the Kenyan population. A substantial proportion (43%) of the population of Kenya is outside the age-range (15-64 years) sampled in this study and the seroprevalence in children <15 years and adults >65 years is often lower (*9, 10*); our estimate for blood donors may be higher than the estimate for the population as a whole. Blood donors also differ from the general population in their risk of exposure to SARS-CoV-2. For instance, potential donors are excluded from giving blood if they have been ill during the last six months so the sample may underestimate the population prevalence of SARS-CoV-2 antibodies; on the other hand, people who are shielding at home are unlikely to be captured in our sample leading to an overestimate of seroprevalence. Our exploration of the two distinct populations of blood donors, FRDs and VNRDs, suggests variation in the seroprevalence by donor group but, of note, 92% (2862/3098) of our sample came from the group with lower seroprevalence and exclusion of VNRDs reduced

the crude seroprevalence in our study little, from 5.6% to 5.4%. Against these considerations, other countries have relied on blood transfusion donors for an early estimate of seroprevalence but later estimates from random population samples have not been substantially different (*25, 26*).

A key strength of this study is the rigorous validation that included testing positive and negative control samples from the target population, as well as reference plasma from the UK NIBSC as part of a WHO-coordinated effort on SARS-CoV-2 seroepidemiology. In addition, we adopted a conservative seropositivity threshold to optimize assay specificity and sensitivity for our setting.

The pandemic response in countries with limited health care capacity has been driven by the aggressive implementation of control measures to limit transmission. Unfortunately, this strategy has been accompanied by enormous collateral costs, particularly in Africa. Modelled estimates of the disruptions of essential medical services, such as immunization and antenatal care, suggest an additional ~253,500 child deaths and 12,200 maternal deaths over six months in low and middle-income countries (*27*). In the absence of social protection, the economic effects of lockdown are debilitating so it is important to obtain an early measure of the trajectory of the epidemic.

Our study provides a national and regional estimate of population exposure to SARS-CoV-2 in an African country. The 4.3% prevalence in blood transfusion donors is in sharp contrast with the reported COVID-19 cases and deaths and supports the impression that disease may be attenuated in Africa (*7*).

References and Notes:

1. United Nations Department of Economic and Social Affairs Population Division, "World Urbanization Prospects: The 2018 Revision, custom data acquired via https://population.un.org/wup/DataQuery/," (2018).

2. Africa Centres for Diseases Control and Prevention, "Coronavirus Disease 2019 (COVID-19). https://africacdc.org/covid-19/. Accessed 21 July 2020," (2020).

3. B. Z. Diop, M. Ngom, C. Pougue Biyong, J. N. Pougue Biyong, The relatively young and rural population may limit the spread and severity of COVID-19 in Africa: a modelling study. *BMJ Glob Health* **5**, 10.1136/bmjgh-2020-002699 (2020).

4. J. B. Dowd *et al.*, Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci U S A* **117**, 9696-9698, 10.1073/pnas.2004911117 (2020).

5. Ministry of Health Kenya, "Press Statement on the update of the coronavirus in the country and response measures. https://www.health.go.ke/wp-content/uploads/2020/03/Coronavirus-Press-Statement-March-17-2020.pdf," (2020).

6. Ministry of Health Kenya, "COVID-19 Situation Reports (SITREP). https://www.health.go.ke/#1591180376422-52af4c1e-256b " (2020).

7. M. Mbow *et al.*, COVID-19 in Africa: Dampening the storm? *Science* **369**, 624-626, 10.1126/science.abd3902 (2020).

8. X. Wu, B. Fu, L. Chen, Y. Feng, Serological tests facilitate identification of asymptomatic SARS-CoV-2 infection in Wuhan, China. *J Med Virol*, 10.1002/jmv.25904 (2020).

9. M. Pollan *et al.*, Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* **396**, 535-544, 0.1016/S0140-6736(20)31483-5 (2020).

10. S. Stringhini *et al.*, Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *Lancet* **396**, 313-319, 10.1016/S0140-6736(20)31304-0 (2020).

11. M. G. Chibwana *et al.*, High SARS-CoV-2 seroprevalence in Health Care Workers but relatively low numbers of deaths in urban Malawi [version 1; peer review: awaiting peer review]. *Wellcome Open Res* **5**, 199, 10.12688/wellcomeopenres.16188.1 (2020).

12. L. Valenti *et al.*, SARS-CoV-2 seroprevalence trends in healthy blood donors during the COVID-19 Milan outbreak. *medRxiv*, 2020.2005.2011.20098442 (2020).

13. L. Amorim Filho *et al.*, Seroprevalence of anti-SARS-CoV-2 among blood donors in Rio de Janeiro, Brazil. *Rev Saude Publica* **54**, 69, 10.11606/s1518-8787.2020054002643 (2020).

14. D. D. Flannery *et al.*, SARS-CoV-2 seroprevalence among parturient women in Philadelphia. *Science Immunology* **5**, 10.1126/sciimmunol.abd5709 (2020).

15. F. Amanat *et al.*, A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med* **26**, 1033-1036, 10.1038/s41591-020-0913-5 (2020).

16. A. T. Huang *et al.*, A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. *Nat Commun* **11**, 4704, 10.1038/s41467-020-18450-4 (2020).

17. X. Xu *et al.*, Seroprevalence of immunoglobulin M and G antibodies against SARS-CoV-2 in China. *Nat Med* **26**, 1193-1195, 10.1038/s41591-020-0949-6 (2020).

18. F. P. Havers *et al.*, Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. *JAMA Intern Med*, 10.1001/jamainternmed.2020.4130 (2020).

19. Public Health England, "Sero-prevalence epidemiology, England. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/899301/Weekly_COVID19_Surveillance_Report_week_28.pdf in Weekly Coronavirus Disease 2019 (COVID-19) Surveillance Report 2020 Week 28," (2020).

20. A. Grifoni *et al.*, Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489-1501, 10.1016/j.cell.2020.05.015 (2020).

21. N. Le Bert *et al.*, SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature*, 10.1038/s41586-020-2550-z (2020).

22. A. Sette, S. Crotty, Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. *Nat Rev Immunol* **20**, 457-458, 10.1038/s41577-020-0389-z (2020).

23. K. W. Ng *et al.*, Pre-existing and de novo humoral immunity to SARS-CoV-2 in humans. *bioRxiv*, 2020.2005.2014.095414 (2020).

24. G. P. Otieno *et al.*, Surveillance of endemic human coronaviruses (HCoV-NL63, OC43 and 229E) associated with childhood pneumonia in Kilifi, Kenya. [version 2; peer review: 2 approved]. *Wellcome Open Res* **5**, 150, 10.12688/wellcomeopenres.16037.2 (2020).

25. H. Ward *et al.*, Antibody prevalence for SARS-CoV-2 following the peak of the pandemic in England: REACT2 study in 100,000 adults. *medRxiv*, 2020.2008.2012.20173690v20173692 (2020).

26. P. H. England, National COVID-19 surveillance report week 40 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/923668/Weekly_COVID19_Surveillance_Report_week_40.pdf (2020).

27. T. Roberton *et al.*, Early estimates of the indirect effects of the COVID-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study. *Lancet Glob Health* **8**, e901-e908, 10.1016/S2214-109X(20)30229-1 (2020).

28. K. Ministry of Health, "Policy Guidelines on Blood Transfusion in Kenya. The National Blood Transfusion Service of Kenya. https://nbtskenya.or.ke/wp-content/uploads/2019/02/Policy-Guidelines-on-Blood-Transfusion-in-Kenya.pdf. Accessed 20 July 2020," (2001).

29. A. Gelman, J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*. A. M. Alvarez, N. L. Beck, L. L. Wu, Eds., Analytical Methods for Social Research (Cambridge University Press, New York, United States, 2007).

30. A. Gelman, C. Carpenter, Bayesian analysis of tests with unknown specificity and sensitivity. *Journal of the Royal Statistical Society C, Applied Statistics. 2020. In press. Applied Statistics.* *http://www.stat.columbia.edu/~gelman/research/unpublished/specificity.pdf* (2020).

31. A. Gelman, Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* **1**, 515-533 (2006).

Acknowledgments: We thank Prof Florian Krammer for providing the plasmids used to generate the RBD, spike protein, and CR3022 mAb used in this work. Development of SARS-CoV-2 reagents was partially supported by the NIAID Centres of Excellence for Influenza Research and Surveillance (CEIRS) contract HHSN272201400008C. The COVID-19 convalescent plasma panel (NIBSC 20/118) and research reagent for SARS-CoV-2 Ab (NIBSC 20/130) were obtained from the National Institute for Biological Standards and Control, UK. We thank the blood donors and KNBTS staff who supported this work. We also thank the WHO SOLIDARITY II network for sharing of protocols and for facilitating the development and distribution of control reagents.

This paper has been published with the permission of the Director, Kenya Medical Research Institute.

This study was approved by the Scientific and Ethics Review Unit (SERU) of the Kenya Medical Research Institute (Protocol SSC 3426). Before the blood draw, donors gave individual consent for the use of their samples for research. Ethical approval was obtained for collection, storage and further use for the sample sets used in the validation assays (SERU numbers: 1433, 3149, 3426).

Funding: This project was funded by the Wellcome Trust (grant numbers 220991/Z/20/Z; 203077/Z/16/Z). Sophie Uyoga is funded by DELTAS Africa Initiative [DEL-15-003], Isabella Ochola-Oyier is funded by a Wellcome Trust Intermediate Fellowship (107568/Z/15/Z), Ambrose Agweyu is funded by a DFID/MRC/NIHR/Wellcome Trust Joint Global Health Trials Award (MR/R006083/1), J. Anthony G. Scott is funded by a Wellcome Trust Senior Research Fellowship (214320) and the NIHR Health Protection Research Unit in Immunisation, Ifedayo

Adetifa is funded by an United Kingdom's Medical Research Council and Department For International Development through a African Research Leader Fellowship (MR/S005293/1) and by the NIHR-MPRU at UCL (grant 2268427 LSHTM). GMW is supported by a fellowship from the Oak Foundation. Charles N. Agoti is funded by the DELTAS Africa Initiative [DEL-15-003], and the Department for International Development and Wellcome (220985/Z/20/Z).

Author contributions:

Conceptualization and methodology: SU, IMOA, AA, JAGS, EW, KG, AE, SV RA, MM, PA, KK WN and GMW.

Investigation: SU, HKK, JN, JT, PW, CR, CY, KKI, EO, TR, IO, SK, ZNM, JNG, DM, CN, EO, LN, LIO and GMW.

Formal analysis: JAGS, MO and CB.

Validation: TL, DW, HKK, JN, JT, LIO and GMW.

Resources and funding acquisition: SU, TL, PB and GMW.

Supervision: EB, BT and PB.

Writing – Original draft preparation: SU, IMOA, AA, JAGS and GMW.

Writing – Review and editing: all authors.

Competing interests: RA, MM, KK and PA are from the Ministry of Health, Government of Kenya. All other authors declare no competing interests.

Data and materials availability: De-identified data has been published on the Havard dataverse server https://doi.org/10.7910/DVN/RENVC9

Supplementary Materials:

Materials and Methods

Figures S1-S6

Tables S1-S5

References ((*28-31*))

Figures

Figure 1. Participant flow diagram for SARS-CoV-2 seroprevalence study of blood donors in Kenya. Exclusion criteria for the selection of samples with complete data.

Figure 2. Timeline of sampling for SARS-CoV-2 seroprevalence in blood donors in Kenya. Against the timeline of the sampling period, panel A shows the weekly crude seroprevalence and 95% confidence interval, panel B shows the daily frequency of samples collected and panel C shows the temporal distribution of samples by region. Proportion, counts and regional distribution of donors during the study period.

Table 1. Crude, population-weighted, and test performance-adjusted SARS-CoV-2 anti-spike protein IgG seroprevalence by participant characteristics and regions. Prevalence estimates calculated using multilevel regression and post-stratification (MLRP) to account for differences in the sample population and the national population, subsequently adjusted for assay sensitivity and specificity.

	All samples	Seropositive samples	Crude seroprevalence		Kenya population (2019 Census)	Bayesian population-weighted seroprevalence*		Bayesian population-weighted, test-adjusted seroprevalence*	
			%	(95% CI)		%	(95% CI)	%	(95% CI)
Age									
15 – 24 years	808	49	6.1	4.5 – 7.9	9,733,174	5.1	3.7 – 6.9	4.4	2.7 – 6.4
25 – 34years	1242	66	5.3	4.1 – 6.7	7,424,967	4.9	3.6 – 6.4	4.2	2.8 – 6.0
35 – 44 years	714	50	7.0	5.2 – 9.1	4,909,191	5.9	4.3 – 8.1	5.2	3.3 – 7.7
45 – 54 years	263	9	3.4	1.6 – 6.4	3,094,771	3.8	1.9 – 6.0	3.0	1.1 – 5.4
55 – 64 years	71	0	0		1,988,062	3.4	0.7 – 6.2	2.9	0.7 – 5.7
Sex									
Male	2540	146	5.7	4.9 – 6.7	13,388,243	4.4	2.9 – 6.2	3.6	1.9 – 5.8
Female	558	28	5.0	3.4 – 7.2	13,761,922	5.5	4.4 – 6.8	4.8	3.5 – 6.4
Regions									
Central	105	7	6.7	2.7 – 13.2	3,452,213	5.6	2.9 – 10.0	4.9	1.9 – 9.7
Mombasa	550	51	9.3	7.0 – 12.0	792,072	8.3	6.1 – 10.9	7.8	5.4 – 10.8
Other Coast	973	39	4.0	2.9 – 5.4	1,671,097	3.7	2.6 – 5.1	2.9	1.6 – 4.6
Eastern / N. Eastern	242	11	4.5	2.3 – 8.0	5,176,080	4.3	2.5 – 7.0	3.5	1.4 – 6.6
Nairobi	235	21	8.9	5.6 – 13.3	3,002,314	7.6	4.9 – 11.2	7.1	4.2 – 11.2
Nyanza	442	30	6.8	4.6 – 9.5	3,363,813	6.0	4.2 – 8.4	5.2	3.1 – 7.9
Rift Valley	481	8	1.7	0.7 – 3.3	7,035,581	2.1	1.1 – 3.6	1.5	0.4 – 3.1
Western	70	7	10.0	4.1 – 19.5	2,656,995	7.0	3.5 – 13.1	6.3	2.5 – 13.1
Total	3,098	174	5.6	4.8 – 6.5	27,150,165	4.9	3.9 – 6.2	4.3	2.9 – 5.8

*Re-weighted prevalence estimates based on demographic data from the 2019 Kenya Population and Housing Census

Table 2. General characteristics of the study population compared to the national population of Kenya. N is the number of individuals in each stratum.

		Blood transfusion samples		Kenya National Census 2019	
		N	%	N	%
Age	15-24 years	808	26.1	9,733,174	35.8
	25-34 years	1,242	40.1	7,424,967	27.3
	35-44 years	714	23.0	4,909,191	18.1
	45-54 years	263	8.5	3,094,771	11.4
	55-64 years	71	2.3	1,988,062	7.3
Sex	Male	2540	82.0	13,388,243	49.3
	Female	558	18.0	13,761,922	50.7
Regions	Central	105	3.4	3,452,213	12.7
	Mombasa	550	17.8	792,072	2.9
	Other Coast	973	31.4	1,671,097	6.2
	Eastern / N. Eastern	242	7.8	5,176,080	19.1
	Nairobi	235	7.6	3,002,314	11.1
	Nyanza	442	14.3	3,363,813	12.4
	Rift Valley	481	15.5	7,035,581	25.9
	Western	70	2.3	2,656,995	9.8
Total	Kenya 15-64 years	3098		27,150,165	

5

Science
AAAS

Supplementary Materials for

Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors: a population-based study

Authors: Sophie Uyoga PhD[1*§], Ifedayo M.O. Adetifa FWACP PhD[1,2§], Henry K. Karanja MSc[1§], James Nyagwange PhD[1§], James Tuju PhD[1], Perpetual Wanjiku BSc[1], Rashid Aman[3] PhD, Mercy Mwangangi MSc[3], Patrick Amoth MMed[3], Kadondi Kasera MSc[3], Wangari Ng'ang'a MSc[4], Charles Rombo BSc[5], Christine Yegon BSc[5], Khamisi Kithi BSc[5], Elizabeth Odhiambo BSc[5], Thomas Rotich BSc[5], Irene Orgut BSc[5], Sammy Kihara Dip.[5], Mark Otiende MSc[1], Christian Bottomley PhD[2], Zonia N. Mupe[1] BSc, Eunice W. Kagucia PhD[1], Katherine E Gallagher PhD[1,2], Anthony Etyang PhD[1], Shirine Voller MSc[1,2], John N. Gitonga Dip.[1], Daisy Mugo BSc[1], Charles N. Agoti PhD[1], Edward Otieno BSc[1], Leonard Ndwiga MSc[1], Teresa Lambe PhD[6], Daniel Wright MSc[6], Edwine Barasa PhD[1], Benjamin Tsofa PhD[1], Philip Bejon PhD[1,6], Lynette I. Ochola-Oyier PhD[1], Ambrose Agweyu PhD[1†], J. Anthony G. Scott FMedSci[1,2†], George M. Warimwe PhD[1,6†]

Correspondence to: SUyoga@kemri-wellcome.org

This PDF file includes:

Materials and Methods

Study sample

This study was carried out at the KEMRI-Wellcome Trust Research Programme (KWTRP) in Kilifi, Kenya, in collaboration with the Kenya National Blood Transfusion Service (KNBTS). The KWTRP is the government-designated laboratory for SARS-CoV-2 testing in Coastal Kenya.

Anonymized residual donor serum samples, used for screening of transfusion transmissible infections, were collected at 4 KNBTS regional centers (Mombasa, Nairobi, Eldoret and Kisumu). The KNBTS guidelines (*28*) define eligible blood donors as individuals aged 16-65 years, weighing ≥50kg, with hemoglobin of 12.5g/dl, a normal blood pressure (systolic 120–129 mmHg and diastolic BP of 80–89 mmHg), a pulse rate of 60-100 beats per minute and without any history of illness in the past 6 months. KNBTS normally relies on voluntary non-remunerated donors (VNRD) recruited at public blood drives typically located in high schools, colleges and universities. Since September 2019, because of reduced funding, KNBTS has depended increasingly on family replacement donors (FRD) who provide units of blood in compensation for those received by sick relatives. Family replacement donors are close family members or friends within immediate reach willing to assist the patient.

Although we sampled from only 4 of 6 regional centers, the sample provides a much wider national representation because each regional center serves between 5-10 administrative counties. Supplementary Table S2 shows the distribution of the donors from the four regional blood transfusion centers across hospitals and satellite blood banks in the various administrative counties. Patients at health facilities served by the regional centers typically

would have travelled some distance to obtain secondary and tertiary care, including transfusion, and relatives, acting as FRDs, may be called in from farther afield. Movement restrictions in/out of Nairobi and Mombasa Counties to control the transmission of COVID-19 reduced, but did not completely exclude, this wider representation.

Assuming a COVID-19 seroprevalence ranging from 3-10% during the study duration, a sample of 300-500 blood donors per operations center per month will give a 4-7% margin of error for seroprevalence estimates.

Antibody testing using SARS-CoV-2 Spike Protein ELISA

We adapted the Krammer Enzyme linked Immunosorbent assay as follows (*15*): Nunc MaxiSorp™ flat-bottom 96-well plates (ThermoFisherScientific) were coated with 2µg/ml of whole trimeric spike protein or spike receptor binding domain (RBD) (Figure S1) at 37°C for 1h, washed 3 times in wash buffer (0.1% Tween 20 in 1X phosphate buffered saline) and blocked with Blocker™ Casein (ThermoFisherScientific) for 1h at room temperature. Heat-inactivated serum or plasma samples were diluted 1:800 in Blocker™ Casein, added to the spike- or RBD-coated plates and incubated for 2h at room temperature. After 3 further washes, 100µl horseradish peroxidase-conjugated goat anti-human IgG antibody (Catalogue number 074-1002, KPL-SeraCare), diluted 1:10,000 in wash buffer, was added to the plates. They were incubated for 1h at room temperature, washed 3 times and developed with o-phenylenediamine dihydrochloride (OPD) substrate (Sigma) for 10 min. Plates were read on an Infinite® 200 PRO microplate reader (TECAN) at 492 nm. The CR3022 monoclonal antibody (mAb) was used as positive control. A pool of sera from 50 adults sampled pre-COVID-19 pandemic was used as negative control.

ELISA performance, validation and definition of a threshold for seropositivity

We validated the RBD and whole trimeric spike ELISAs against panels of pre-pandemic and pandemic serum/plasma (Table S1). The gold-standard negative 'pre-pandemic' serum/plasma panels comprised: sera from annual cross-sectional surveys for malaria surveillance in coastal Kenya in April-May 2018 (200 adults and 200 children); (2) sera from 500 adult blood donors collected in 2018 as part of research into the quality of transfused blood in coastal Kenya; (3) convalescent plasma from children (n=9) admitted to Kilifi County Hospital 2011-2013 with PCR-confirmed infection with endemic human coronavirus NL63 (n=4) or OC43 (n=5) (*24*).

The gold-standard positive 'pandemic' plasma panel comprised plasma from 174 COVID-19 patients sampled ≥7 days after their PCR-positive diagnosis (Figure S2), and a well-characterized panel of 5 plasma samples from the National Institute of Biological Standards and Control (NIBSC) in the UK. These included the research reagent for anti-SARS-CoV-2 Ab (NIBSC code 20/130), and the convalescent plasma panel NIBSC code 20/118 that includes a panel of 4 convalescent plasma from recovered COVID-19 patients in the UK and a pre-pandemic plasma pool from healthy UK adults.

Assay performance was assessed by testing these serum/plasma panels against both the RBD and whole trimeric spike protein of SARS-CoV-2 by IgG ELISA. Both assays had the same performance characteristics in our hands and, as the specificity of either was already excellent, in the interests of laboratory efficiency, we chose to pursue our studies with one assay alone. We chose the assay with the highest sensitivity. CR3022 mAb was used as positive control and a pool of pre-pandemic sera from 50 Kenyan adults was used

as a negative control. These were randomly selected from the malaria cross-sectional surveillance samples described earlier.

Ethical Considerations

This study was approved by the Scientific and Ethics Review Unit (SERU) of the Kenya Medical Research Institute (Protocol SSC 3426). Before the blood draw, donors gave individual consent the use of their samples for research. Ethical approval was obtained for collection, storage and further use for the sample sets used in the validation assays (SERU numbers: 1433, 3149, 3426).

Statistical Analysis

We tabulated seropositive results by age, sex and residence (county or region). Because the survey was a non-random sample of the Kenyan population, we also calculated standardized prevalence estimates using population data from the Kenya 2019 census. We used 2 methods of standardization; 1) direct standardization on the observed prevalence and population weights in 80 region-age-sex strata, and 2) multilevel regression and post-stratification (MLRP) adjusted for the sensitivity and specificity of the assay (*29*). To predict stratum prevalence for MLRP we fitted a Bayesian logistic regression model that included sex as a fixed effect, and age and region as random effects (*30*). The Bayesian analysis was done using the rjags package in R version 3.6.1.(*31*). We used vague or weakly informative priors for all parameters (*30*).

We plotted sample distribution and seroprevalence over time to illustrate potential time-trends and tabulated seroprevalence across different donor groups (FRDs vs VNRDs) to explore potential biases.

Multilevel regression and post-stratification (MRP) was used to reweight the regional, age-specific and sex-specific prevalence estimates. The method was implemented in 2 stages:

1) A hierarchical regression model was used to predict the prevalence in 80 age-sex-region strata (5 categories of age, 8 regions and 2 categories of sex)

2) Regional, age-specific and sex-specific prevalence estimates were obtained by appropriately weighting stratum-specific prevalence estimates using data from the 2019 Kenyan census. For example, to estimate the prevalence in Mombasa, the age and sex specific prevalence estimates for the Mombasa were weighted by age and sex specific population totals.

In the hierarchical model, the number of seropositive individuals in age group ($i = 1, \dots 5$), region ($j = 1, \dots, 8$) and sex ($k = 1, 2$) follows a binomial distribution:

$$y_{ijk} \sim \text{Bin}(n_{ijk}, p^*_{ijk}).$$

where p^*_{ijk} is the observed prevalence and n_{ijk} is the number tested.

The observed prevalence depends on the true prevalence as well as the sensitivity and specificity of the assay, i.e.

$$p^*_{ijk} = p_{ijk} \times se + (1 - p_{ijk}) \times (1 - sp).$$

And the true prevalence, p_{ijk}, depends on age, region and sex according to the following logistic model:

$$\text{logit}(p_{ijk}) = b_0 + b_1(k - 1) + (v_j - \bar{v}) + (u_i - \bar{u})$$

$$v_j \sim \mathrm{N}\left(\mu_{\mathrm{region}},\ \sigma^2_{\mathrm{region}}\right)$$

$$u_i \sim \mathrm{N}\left(\mu_{\mathrm{age}},\ \sigma^2_{\mathrm{age}}\right).$$

Note that v_j and u_i are centred around their respective means. Although we could have fitted an equivalent uncentered model with $v_j \sim \mathrm{N}\left(0,\ \sigma^2_{\mathrm{region}}\right)$ and $u_i \sim \mathrm{N}\left(0,\ \sigma^2_{\mathrm{age}}\right)$, we chose instead to centre these variables to reduce autocorrelation in the McMC-generated sample of the posterior distribution [Chapter 19 in Gelman and Hill] (*29*).

To estimate se and sp , we used data from 179 true positives (positive on PCR for SARS-CoV-2) and 910 true negatives (samples collected pre COVID-19). Among the true positives x = 166 tested positive on the assay, and among the true negatives $z = 901$ tested negative. These data were included in the likelihood by assuming a binomial model: $x \sim \mathrm{Bin}(se, 179)$ and $z \sim \mathrm{Bin}(sp, 910)$.

The full model was fitted using the following priors:

$$b_0 \sim \mathrm{N}(0, 10^4)$$

$$b_1 \sim \mathrm{N}(0, 10^4)$$

$$\mathrm{logit}(se) \sim \mathrm{N}(1.5, 0.25)$$

$$\mathrm{logit}(sp) \sim \mathrm{N}(5.3, 0.25)$$

$$\mu_{\mathrm{age}} \sim \mathrm{N}(0, 10^4)$$

$$\mu_{\mathrm{region}} \sim \mathrm{N}(0, 10^4)$$

$$\sigma_{\mathrm{age}} \sim \mathrm{N}(0, 0.25)^+$$

$$\sigma_{\text{region}} \sim \text{N}(0, 0.25)^{+}.$$

The half-normal priors for σ_{age} and σ_{region} were chosen to be weakly informative (*31*). To interpret these priors, note that if the prevalence is 5% then a variance of 0.25 on the logit scale corresponds to 95% of estimates being between 1.9% and 12.5%. Non-informative priors were used for all other parameters.

Rjags code:

```
library(rjags)
library(coda)

### Data
# female = 1, male = 2;
# age categories: 1 = 15-24, 2 = 25-34, 3 = 35-44, 4 = 45-54, 5 = 55-64;
# regions: 1 = central, 2 = coast_mombasa, 3 = coast_other, 4 = eastern/north eastern,
# 5 = nairobi, 6 = nyanza, 7 = rift valley, 8 = western

dat <- read.csv("stratum_level_data_prov7mom.csv")
nr <- 8
na <- 5
ns <- 2
y <- array(dat$spikepos, dim = c(na, nr, ns))
n <- array(dat$n, dim = c(na, nr, ns))
pw <- array(dat$pw, dim = c(na, nr, ns))

### Model

model_string <- "model{

# Likelihood

   for(i in 1:na){
     for(j in 1:nr){
       for(k in 1:ns){
                                y[i, j, k] ~ dbinom(se * p[i, j, k]  +
                                                    (1 - sp) * (1 - p[i, j, k]) ,
                                                    n[i, j, k])

                                logit(p[i, j, k]) <- (b0 + b1 * (k - 1)
                                                      + u[i] - mean(u[]) + v[j] - mean(v[]))

       }
     }
   }

# Sensitivity and specificity

   x ~ dbinom(se, 179)
   z ~ dbinom(sp, 910)

# Age effect

  tau_a <- 1/pow(sd_a, 2)
  for(i in 1:na){
    u[i] ~ dnorm(mu_a, tau_a)
            }

# Region effect

  tau_r <- 1/pow(sd_r, 2)
```

```
  for(j in 1:nr){
            v[j] ~ dnorm(mu_r, tau_r)
            }

# Priors

  b0 ~ dnorm(0, 1e-04)
  b1 ~ dnorm(0, 1e-04)
  se ~ dunif(0,1)
  sp ~ dunif(0,1)

# Hyperpriors

  sd_a ~ dnorm(0, 4) T(0,)
  sd_r ~ dnorm(0, 4) T(0,)
  mu_a ~ dnorm(0, 1e-04)
  mu_r ~ dnorm(0, 1e-04)

# Predicted prevalence by age, region and sex

  for(i in 1:na){
    age[i] <- inprod(p[i,1:nr,1:ns], pw[i,1:nr,1:ns])/sum(pw[i,1:nr,1:ns])
  }

  for(j in 1:nr){
    region[j] <- inprod(p[1:na,j,1:ns], pw[1:na,j,1:ns])/sum(pw[1:na,j,1:ns])
  }

  for(k in 1:ns){
    sex[k] <- inprod(p[1:na,1:nr,k], pw[1:na,1:nr,k])/sum(pw[1:na,1:nr,k])
  }

  national <- inprod(p[1:na,1:nr,1:ns], pw[1:na,1:nr,1:ns])

          }"

### Compile and update

model <- jags.model(textConnection(model_string),
                data = list(y = y,
                            n = n,
                            x = 166,
                            z = 901,
                            pw = pw,
                            na = na,
                            nr = nr,
                            ns = ns),
                n.chains = 1,
                inits = list(.RNG.name = "base::Wichmann-Hill",
                             .RNG.seed = 999))

update(model, 1000, progress.bar = "none") # Burn-in period = 1000 samples

samp <- coda.samples(model,
                 variable.names = c("age", "region", "sex", "national",
```

```
                        "se", "sp"),
              n.iter = 10000,
              progress.bar = "none")
summary(samp)
```

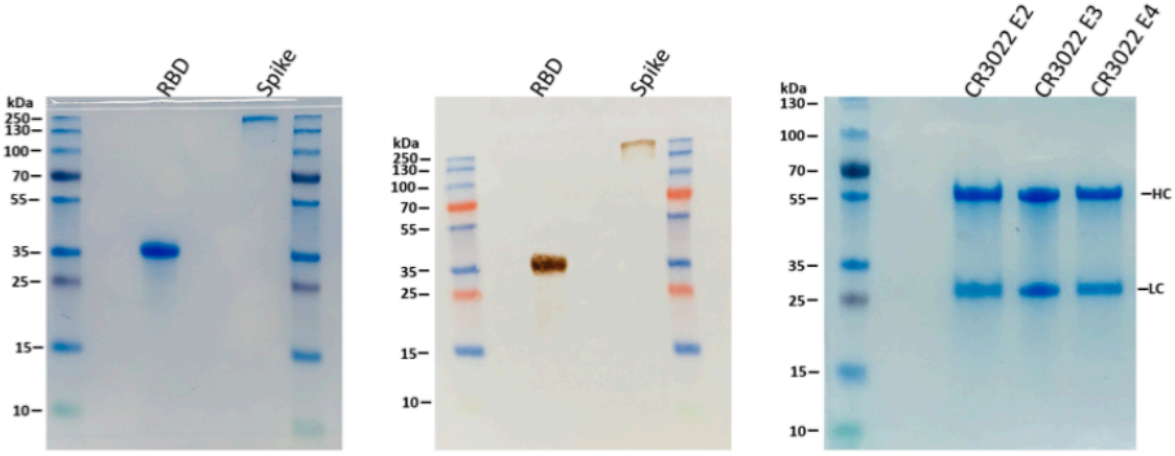
Data: "stratum_level_data_prov7mom"

sex	region	agecat	pop	n	spikepos	pw
1	1	1	519959	12	2	0.01915123
1	1	2	460999	15	0	0.0169796
1	1	3	362528	5	0	0.0133527
1	1	4	250703	1	0	0.00923394
1	1	5	154353	1	0	0.00568516
1	2	1	129100	29	3	0.00475504
1	2	2	133665	40	2	0.00492317
1	2	3	73755	29	3	0.00271656
1	2	4	35204	13	1	0.00129664
1	2	5	17586	3	0	0.00064773
1	3	1	318515	23	0	0.0117316
1	3	2	235979	52	1	0.00869162
1	3	3	143652	25	0	0.00529102
1	3	4	89621	5	0	0.00330094
1	3	5	70401	1	0	0.00259302
1	4	1	917709	12	0	0.03380123
1	4	2	658336	21	2	0.02424796
1	4	3	467552	14	3	0.01722096
1	4	4	296313	4	0	0.01091386
1	4	5	204280	1	0	0.00752408
1	5	1	506240	12	0	0.01864593
1	5	2	546839	19	1	0.02014128
1	5	3	277879	7	0	0.01023489
1	5	4	121945	4	0	0.0044915
1	5	5	50945	2	0	0.00187642
1	6	1	667485	50	7	0.02458493
1	6	2	500100	32	1	0.01841978
1	6	3	275937	21	0	0.01016336
1	6	4	184347	1	0	0.0067899
1	6	5	159604	2	0	0.00587856
1	7	1	1338584	20	0	0.04930298
1	7	2	1004252	45	0	0.0369888
1	7	3	579433	19	0	0.02134179
1	7	4	374981	3	0	0.01381137
1	7	5	233322	0	0	0.00859376
1	8	1	536628	5	2	0.01976518

1	8	2	346576	4	0	0.01276515
1	8	3	223106	5	0	0.00821748
1	8	4	162342	0	0	0.00597941
1	8	5	131167	1	0	0.00483117
2	1	1	502372	21	1	0.01850346
2	1	2	435831	30	2	0.01605261
2	1	3	361207	14	2	0.01330404
2	1	4	253810	6	0	0.00934838
2	1	5	150451	0	0	0.00554144
2	2	1	111366	97	11	0.00410185
2	2	2	131180	174	21	0.00483165
2	2	3	87176	106	9	0.00321088
2	2	4	50520	47	1	0.00186076
2	2	5	22520	12	0	0.00082946
2	3	1	310596	174	6	0.01143993
2	3	2	203297	369	13	0.00748787
2	3	3	140459	209	15	0.00517341
2	3	4	96887	87	4	0.00356856
2	3	5	61690	28	0	0.00227218
2	4	1	995785	46	0	0.03667694
2	4	2	639043	75	2	0.02353735
2	4	3	481607	45	3	0.01773864
2	4	4	312435	22	1	0.01150767
2	4	5	203020	2	0	0.00747767
2	5	1	408632	36	3	0.01505081
2	5	2	532179	82	7	0.01960132
2	5	3	321079	50	9	0.01182604
2	5	4	166936	19	1	0.00614862
2	5	5	69640	4	0	0.00256499
2	6	1	613177	143	9	0.02258465
2	6	2	391827	114	8	0.01443185
2	6	3	282754	54	4	0.01041445
2	6	4	162762	17	1	0.00599488
2	6	5	125820	8	0	0.00463423
2	7	1	1338739	113	4	0.04930869
2	7	2	928186	151	3	0.03418712
2	7	3	621283	96	1	0.02288321
2	7	4	392358	29	0	0.0144514
2	7	5	224443	5	0	0.00826673
2	8	1	518287	15	1	0.01908965
2	8	2	276678	19	3	0.01019066
2	8	3	209784	15	1	0.0077268
2	8	4	143607	5	0	0.00528936

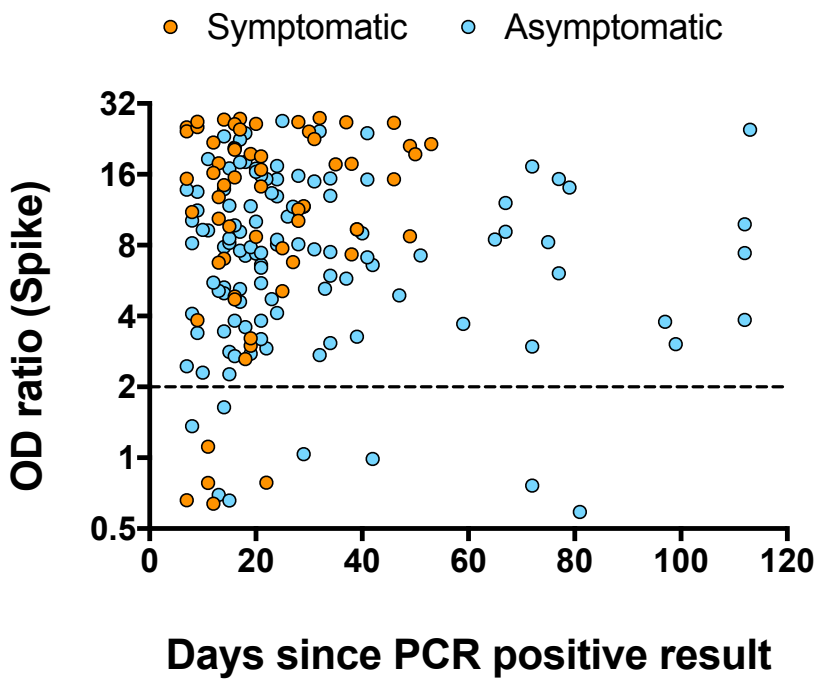
2	8	5	108820	1	0	0.00400808

Fig. S1. SARS-CoV-2 RBD and whole spike protein and CR3022 mAb production



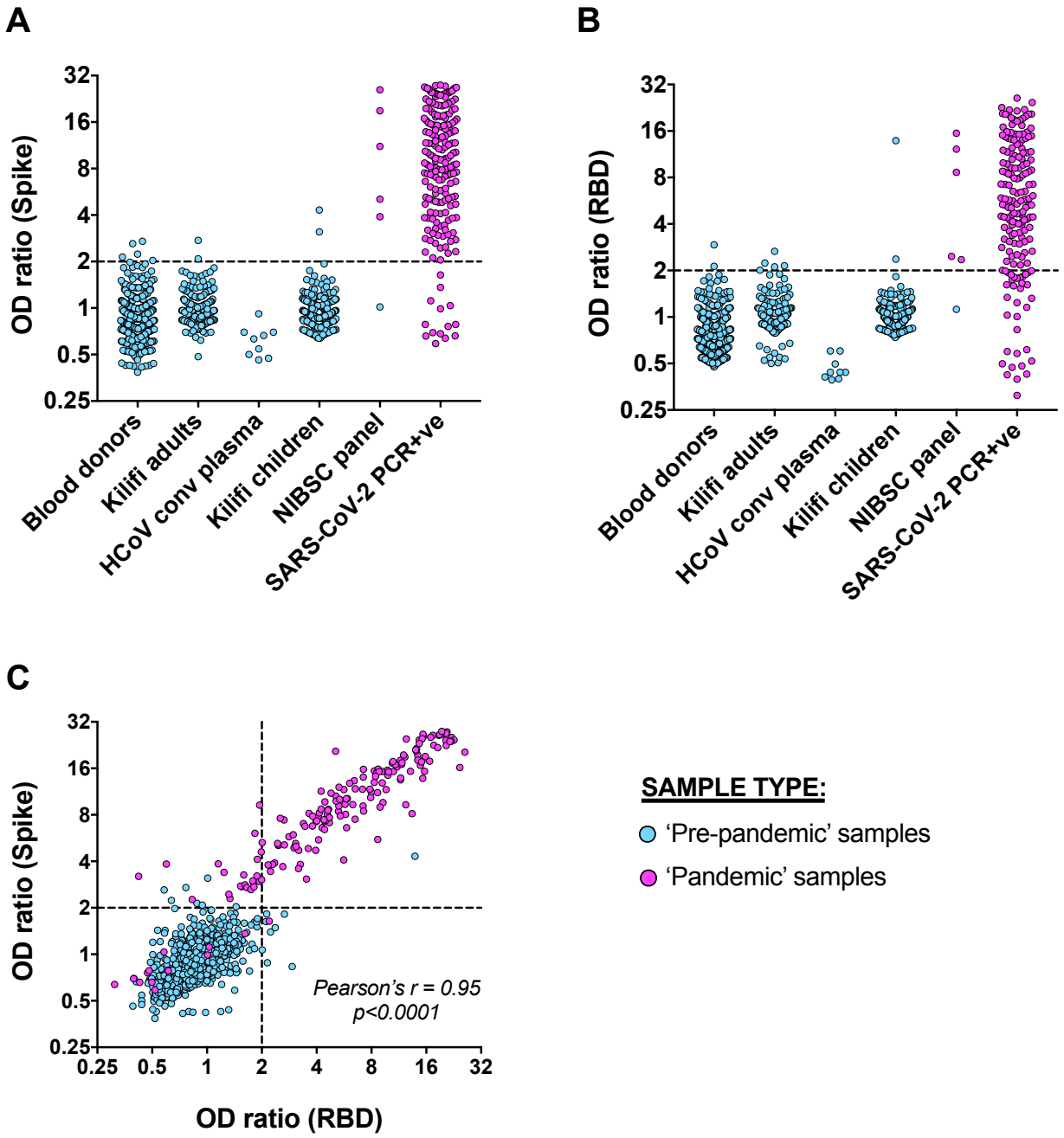
SDS-PAGE and Western blot images of SARS-CoV-2 whole spike and RBD protein, and CR3022 elutions are shown. Concentrated proteins were mixed with 5x protein loading buffer (National Diagnostics) at a ratio of 4:1, heated for 10 min at 100°C and protein integrity analyzed on 10% reducing SDS-PAGE gels made with ProtoGel Quick-cast 12% (National Diagnostics). Gels were run at 100 V for 2 hours and stained with InstantBlue (Expedeon) and destained in distilled water overnight. To check the specificity of expressed and purified proteins, gels were run for 100 V for 2 h and the unstained gels blotted on PVDF Western blotting membrane (Sigma) by semi-dry transfer using Pierce Power Blot Cassette (ThermoFisherScientific). After transfer, the blot was blocked for 1 h in 5% skimmed milk (blocker) and incubated for 1 hour with anti-poly-Histidine-peroxidase antibody (Sigma) in blocker at a ratio of 1:1000. The blot was washed 4x with 0.1% Tween 20 (Sigma) in 1x Phosphate Buffered Saline (PBS) (Thermo Fisher Scientific) and developed with DAB substrate (Sigma) containing 30% H2O2 (Sigma).

Fig. S2. Relationship between antibody response and time since PCR diagnosis



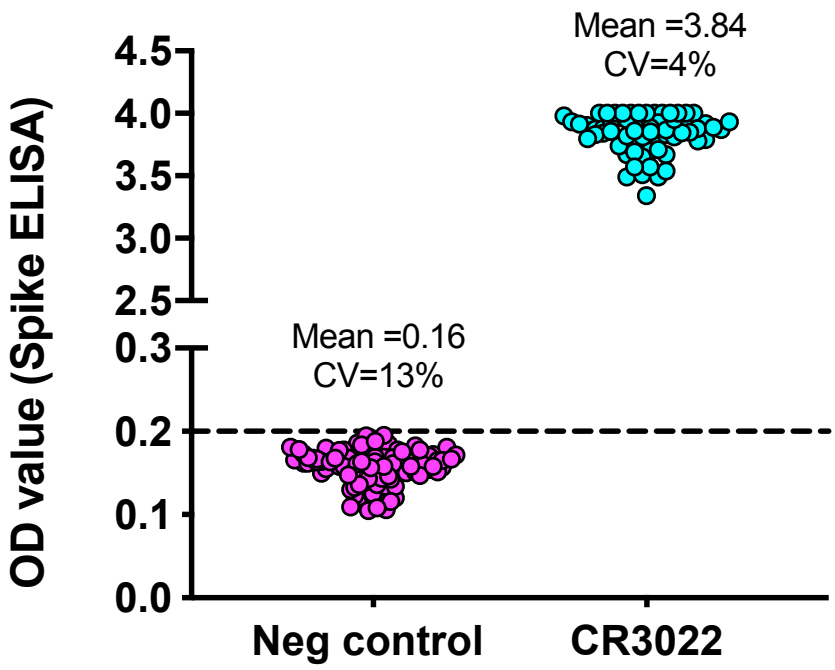
Each point represents one of 174 SARS-CoV-2 positive individuals from Kenya included in the 'pandemic' panel, stratified by whether they had asymptomatic or symptomatic infection. The spike ELISA OD ratios are shown with respect to the duration between the PCR positive result and the sampling date of the serum used in the ELISA.

Fig. S3. Anti-SARS-CoV-2 Spike IgG antibody ELISA performance characteristics



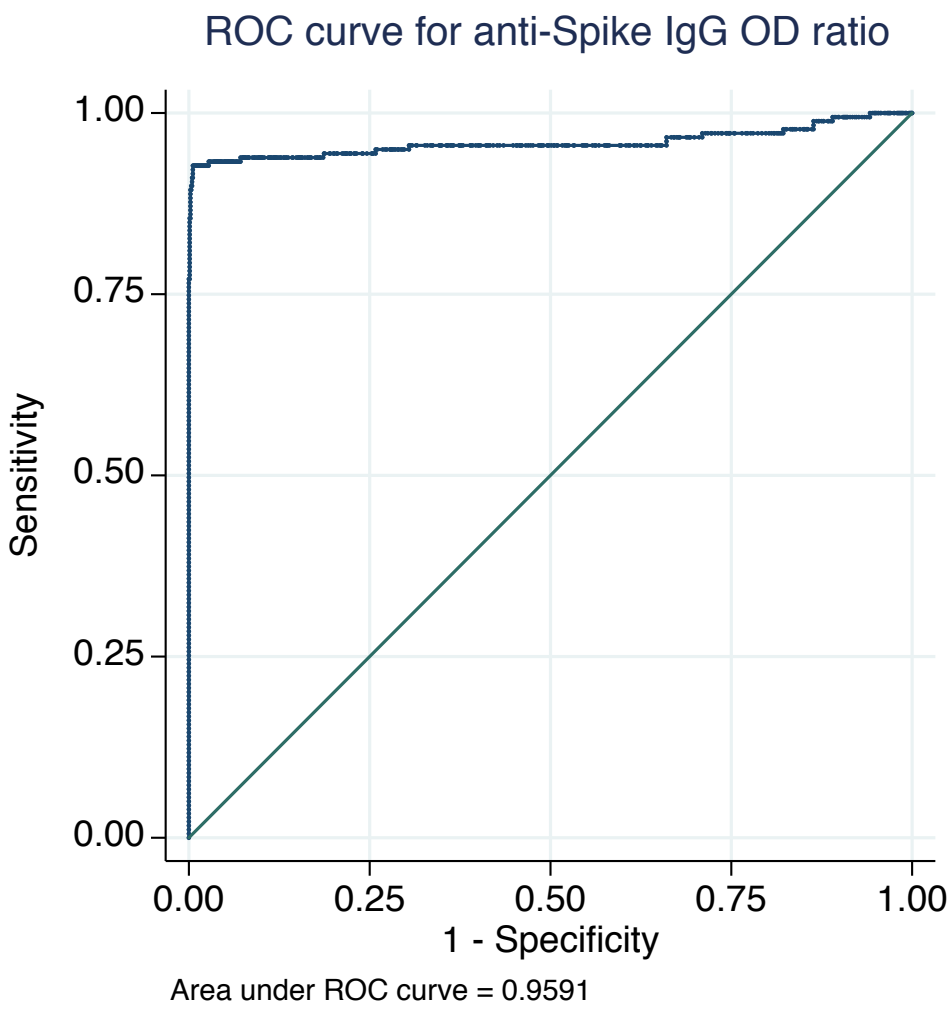
The OD ratios from spike and RBD ELISAs are shown. Each point represents an individual, colour-coded by the sampling period i.e. pre-pandemic or pandemic. The number of samples in each group are summarized in Table S1.

Fig. S4. Assay reproducibility



We assessed reproducibility of the spike ELISA by examining the ODs and coefficient of variation for the negative and positive (CR3022) controls for all the test runs done during the screening of the blood donor samples. All negative control ODs were expected to be <0.2 and >3 for positive controls. Performance was as expected with little inter-assay variation.

Fig. S5. Receiver operating characteristic curve for anti-Spike ELISA



We assessed the discrimination of the OD ratio of anti-Spike IgG in a ROC curve constructed using 910 true negative controls and 179 true positive controls (see Table S1). The assay discriminates well between the two populations and the selected threshold of 2.00 has a sensitivity of 92.7% and a specificity of 99.0% at the upper left-hand corner of the ROC curve.

Fig. S6. Impact of threshold selection on population discrimination

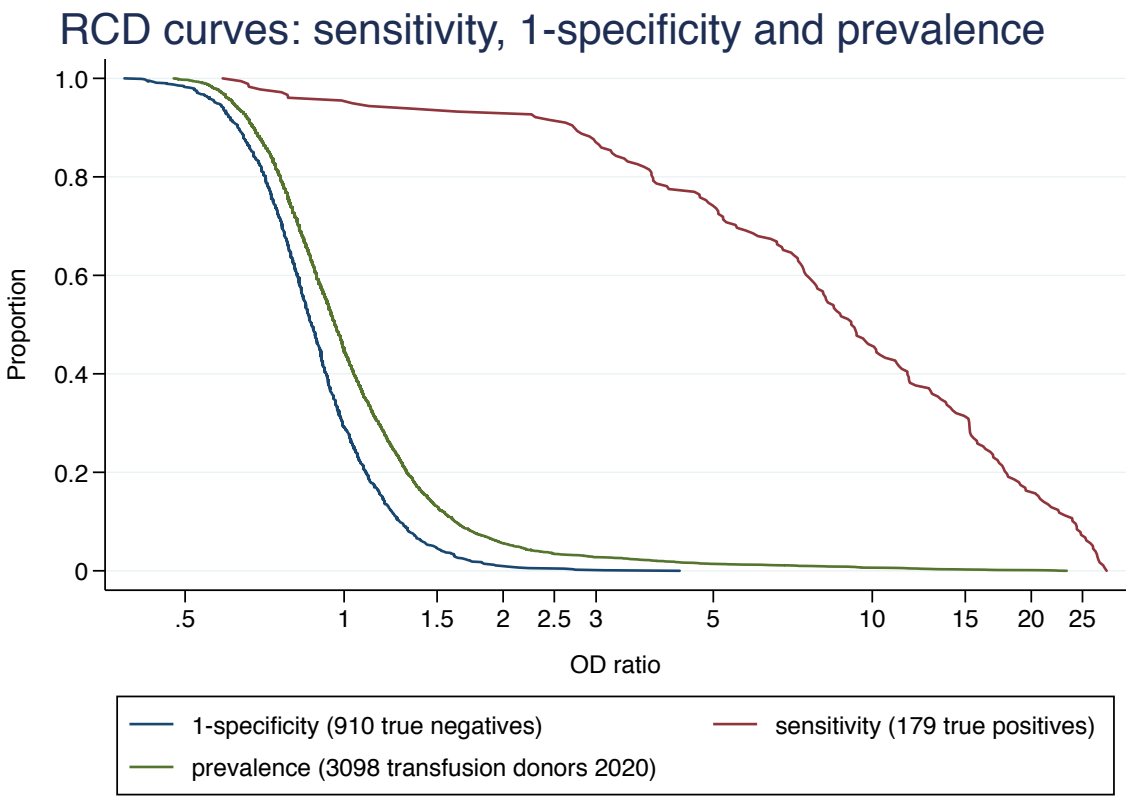


Figure shows reverse cumulative distribution curves to illustrate the dynamic range, distribution and discrimination of the anti-Spike IgG ELISA results (expressed as a ratio of test OD to negative control OD) among the true negative controls, true positive samples and the 2020 blood transfusion donors.

Table S1. Assay validation, selection and definition of seropositivity for SARS-CoV-2 ELISA

		Sample OD>Negative control OD + 3SD			Sample OD-to-Negative control OD ratio >2		
	N	**RBD +ve**	**Spike +ve**	**Both +ve**	**RBD +ve**	**Spike +ve**	**Both +ve**
Pre-pandemic panel (gold standard negatives)							
Adults, Kilifi cross-sectional survey, 2018	**200**	39	22	10	5	2	0
Adults, coastal Kenya blood donors, 2018	**500**	26	56	10	2	5	0
Children, Kilifi cross-sectional survey, 2018	**200**	40	12	7	2	2	1
Children, HCoV (OC43, NL63) convalescent plasma, 2011-13	**9**	0	0	0	0	0	0
NIBSC UK adults pooled plasma	**1**	1	0	0	0	0	0
Pandemic panel (gold standard positives)							
NIBSC UK COVID-19 convalescent plasma	**5**	5	5	5	5	5	5
SARS-CoV-2 PCR positive cohort	**174**	159	164	159	140	161	139

OD- Optical Density
SD- Standard Deviation
NIBSC UK- National Institute of Biological Standards and Control, United Kingdom
HCoV- Human coronaviruses; convalescent plasma from 4 children admitted with HCoV NL63 infection, and 5 with HCoV OC43 infection.

Table S2. Distribution of blood donors across the various administrative regions in the country covered by the four regional blood centers.

County	**Mombasa**	**Nairobi**	**Kisumu**	**Eldoret**	Total
Kilifi	485	0	0	0	485
Kwale	236	0	0	0	236
Lamu	82	0	0	0	82
Mombasa	550	0	0	0	550
Taita Taveta	133	0	0	0	133
Tana River	37	0	0	0	37
Garissa	0	3	0	0	3
Kiambu	0	25	0	0	25
Kitui	0	5	0	0	5
Machakos	0	145	0	0	145
Makueni	0	41	0	0	41
Marsabit	0	11	0	0	11
Nairobi	0	235	0	0	235
Nyeri	0	80	0	0	80
Turkana	0	15	0	26	41
Wajir	0	37	0	0	37
Bungoma	0	0	1	19	20
Busia	0	0	23	0	23
Homabay	0	0	56	0	56
Kericho	0	0	1	0	1
Kisii	0	0	63	0	63
Kisumu	0	0	197	0	197
Migori	0	0	26	0	26
Narok	0	0	1	0	1
Nyamira	0	0	1	0	1
Siaya	0	0	99	0	99
Vihiga	0	0	27	0	27
Trans Nzoia	0	0	0	45	45
Uasin Gishu	0	0	0	376	376
West Pokot	0	0	0	17	17
	1523	597	495	457	3098

Table S3. Crude, directly-standardized, population-weighted, and test performance-adjusted SARS-CoV-2 IgG antibody seroprevalence estimates by participant characteristics and geography in 8 regions

	All samples	Seropositive samples	Crude seroprevalence % (95% CI)		Kenya population (2019 Census)	Directly standardized seroprevalence* % (95% CI)		Bayesian population weighted seroprevalence** % (95% CI)		Bayesian population weighted, test-adjusted seroprevalence** % (95% CI)	
Age											
15 - 24 years	808	49	6.1	4.5 - 7.9	9,733,174	6.3	4.0 - 8.6	5.1	3.7 - 6.9	4.4	2.7 – 6.4
25 - 34years	1242	66	5.3	4.1 - 6.7	7,424,967	4.3	2.5 - 6.2	4.9	3.6 - 6.4	4.2	2.8 – 6.0
35 - 44 years	714	50	7.0	5.2 - 9.1	4,909,191	6.3	3.3 - 9.3	5.9	4.3 - 8.1	5.2	3.3 – 7.7
45 - 54 years	263	9	3.4	1.6 - 6.4	3,094,771	1.3	0.0 - 3.4	3.8	1.9 - 6.0	3.0	1.1 – 5.4
55 - 64 years	71	0	0		1,988,062	0	0	3.4	0.7 - 6.2	2.9	0.7 – 5.7
Sex											
Male	2540	146	5.7	4.9 - 6.7	13,388,243	4.8	3.9 - 5.7	4.4	2.9 – 6.2	3.6	1.9 – 5.8
Female	558	28	5.0	3.4 - 7.2	13,761,922	4.6	2.7 - 6.5	5.5	4.4 – 6.8	4.8	3.5 – 6.4
Regions											
Central	105	7	6.7	2.7 - 13.2	3,452,213	5.5	1.2 - 9.9	5.6	2.9 – 10.0	4.9	1.9 – 9.7
Mombasa	550	51	9.3	7.0 - 12.0	792,072	8.5	6.1 - 10.9	8.3	6.1 – 10.9	7.8	5.4 – 10.8
Other Coast	973	39	4.0	2.9 - 5.4	1,671,097	2.2	1.0 - 3.4	3.7	2.6 – 5.1	2.9	1.6 – 4.6
Eastern / N. Eastern	242	11	4.5	2.3 - 8.0	5,176,080	4.4	2.4 - 6.3	4.3	2.5 – 7.0	3.5	1.4 – 6.6
Nairobi	235	21	8.9	5.6 - 13.3	3,002,314	5.8	2.3 - 9.3	7.6	4.9 – 11.2	7.1	4.2 – 11.2
Nyanza	442	30	6.8	4.6 - 9.5	3,363,813	6.1	3.8 - 8.4	6.0	4.2 – 8.4	5.2	3.1 – 7.9
Rift Valley	481	8	1.7	0.7 - 3.3	7,035,581	1.0	0.0 - 2.3	2.1	1.1 – 3.6	1.5	0.4 – 3.1
Western	70	7	10.0	4.1 - 19.5	2,656,995	11.5	4.4 - 18.7	7.0	3.5 - 13.1	6.3	2.5 – 13.1
Total	3,098	174	5.6	4.8 - 6.5	27,150,165	5.4	4.2 – 6.6	4.9	3.9 –6.2	4.3	2.9 – 5.8

*Prevalence estimates for regions were directly standardized on age and sex; those for age were directly standardized on sex and region; and those for sex were directly standardized on age and region. The total prevalence estimate was directly standardized on age, sex and region.
**Re-weighted prevalence estimates based on demographic data from the 2019 Kenyan census.

Table S4. Crude, directly-standardized, population-weighted, and test-performance adjusted SARS-CoV-2 IgG antibody seroprevalence estimates by geography in 13 counties/regions

	County	All samples	Seropositive samples	Crude seroprevalence %	(95% CI)	Kenya population (2019 Census)	Directly standardized seroprevalence* %	(95% CI)	Bayesian population weighted seroprevalence** %	(95% CI)	Bayesian population weighted, test-adjusted seroprevalence** %	(95% CI)
Region												
Central	All	105	7	6.7	2.7 - 13.2	3,452,213	5.5	1.2 - 9.9	5.5	2.7 – 10.0	4.9	1.9 – 9.6
Coast	Kilifi	485	24	4.9	3.2 - 7.3	784,069	2.2	0.3 - 4.2	4.5	2.9 - 6.5	3.9	2.2 – 6.3
	Kwale	236	12	5.1	2.7 - 8.7	446,434	4.0	0.8 - 7.1	4.5	2.5 - 7.4	3.8	1.6 – 7.0
	Mombasa	550	51	9.3	7.0 - 12.0	792,072	8.5	6.1 - 10.9	8.3	6.1 - 11.0	8.0	5.5 – 11.1
	Other	252	3	1.2	0.2 - 3.4	440,594	0.7	0.0 - 2.0	1.9	0.8 - 3.8	1.4	0.4 – 3.4
Eastern /	Machakos	145	9	6.2	2.9 - 11.4	878,729	7.1	3.6 - 10.5	5.3	2.8 - 9.1	4.6	1.8 – 8.7
N Eastern	Other	97	2	2.1	0.3 - 7.3	4,297,351	0.8	0.0 - 2.4	2.9	1.0 - 6.3	2.4	0.6 – 5.7
Nairobi	All	235	21	8.9	5.6 - 13.3	3,002,314	5.8	2.3 - 9.3	7.6	4.9 - 11.4	7.3	4.2 – 11.4
Nyanza	Kisumu	197	14	7.1	3.9 - 11.6	657,677	6.3	2.9 - 9.7	6.1	3.6 - 9.7	5.5	2.8 – 9.6
	Other	245	16	6.5	3.8 - 10.3	2,706,136	5.9	2.4 - 9.4	5.6	3.4 - 8.7	4.9	2.5 – 8.4
Rift Valley	Uasin Gishu	376	3	0.8	0.1 - 2.3	700,908	0.6	0.0 - 1.7	1.5	0.6 - 2.9	1.1	0.3 – 2.6
	Other	105	5	4.8	1.6 - 10.7	6,334,673	2.7	0.0 - 6.9	4.3	2.0 - 8.3	3.7	1.3 – 8.0
Western	All	70	7	10.0	4.1 - 19.5	2,656,995	11.5	4.4 - 18.7	7.0	3.4 - 13.2	6.4	2.5 – 13.5
Total		3,098	174	5.6	4.8 – 6.5	27,150,165	5.1	3.6 - 6.6	5.2	4.0 – 6.9	4.7	3.2 – 6.5

*Prevalence estimates for counties were directly standardized on age and sex. The total prevalence estimate was directly standardized on age, sex and region.
**Re-weighted prevalence estimates based on demographic data from the 2019 Kenyan census.

Table S5. A comparison of the general characteristics and SARS-CoV-2 seroprevalence in the blood donor populations in Kenya

	Family Replacement Donors				Voluntary Non-Remunerated Donors			
	Number sampled	**% of sample**	**Antibody positive**	**Seroprevalence (%)**	**Number sampled**	**% of sample**	**Antibody positive**	**Seroprevalence (%)**
All donors	2,862	100.0%	154	5.4%	236	100.0%	20	8.5%
Male	2,346	82.0%	132	5.6%	194	82.2%	14	7.2%
Female	516	18.0%	22	4.3%	42	17.8%	6	14.3%
15-24 years	755	26.4%	43	5.7%	53	22.5%	6	11.3%
25-34 years	1,146	40.0%	57	5.0%	96	40.7%	9	9.4%
35-44 years	663	23.2%	47	7.1%	51	21.6%	3	5.9%
45-54 years	236	8.2%	7	3.0%	27	11.4%	2	7.4%
55-64 years	62	2.2%	0	0.0%	9	3.8%	0	0.0%