BMJ Open

Seroprevalence of SARS-CoV-2 in urban settings in three sub-Saharan African countries (SeroCoV): a study protocol for a household-based cross-sectional prevalence study using two-stage cluster sampling

Eva Lorenz [1,2,3] Aurélia Souares,[4,5] John Amuasi,[6,7] Wibke Loag,[1]
Christina Deschermeier,[8] Daniela Fusco [1,2] Dominik Benke,[1]
Rivo A Rakotoarivelo,[9] Mandranto Tahinamandranto Rasamoelina,[10]
Raphael Rakotozandrindrainy,[11] Ali Sie,[5,12] Anthony Afum-Adjei Awuah [6,13]
Jürgen May,[1,2] Daniel Eibach,[1,2] Nicole Struck[1,2]

ABSTRACT

Introduction The current COVID-19 pandemic has impacted the entire world with increasing morbidity and mortality and has resulted in serious economic and social consequences. Assessing the burden of COVID-19 is essential for developing efficient pandemic preparedness and response strategies and for determining the impact of implemented control measures. Population-based seroprevalence surveys are critical to estimate infection rates, monitor the progression of the epidemic and to allow for the identification of persons exposed to the infection who may either have been asymptomatic or were never tested. This is especially important for countries where effective testing and tracking systems could not be established and where non-severe cases or under-reported deaths might have blurred the true burden of COVID-19. Most seroprevalence surveys performed in sub-Saharan Africa have targeted specific high risk or more easily accessible populations such as healthcare workers or blood donors, and household-based estimates are rarely available. Here, we present the study protocol for a SARS-CoV-2 seroprevalence estimation in the general population of Burkina Faso, Ghana and Madagascar in 2021.

Methods and analysis The SeroCoV study is a household-based cross-sectional prevalence investigation in persons aged 10 years and older living in urban areas in six cities using a two-stage geographical cluster sampling method stratified by age and sex. The presence of anti-SARS-CoV-2 IgG antibodies will be determined using a sensitive and specific SARS-CoV-2 IgG ELISA. In addition, questionnaires will cover sociodemographic information, episodes of diseases and history of testing and treatment for COVID-like symptoms, travel history and safety measures. We will estimate the seroprevalence of SARS-CoV-2, taking into account test performance and adjusting for the age and sex of the respective populations.

Ethics and dissemination Ethical approval was received for all participating countries. Results will be disseminated through reports and presentations at the country level as well as peer-reviewed publications and international scientific conferences presentations.

Strengths and limitations of this study

► The study assesses the population-based sero-prevalence of SARS-CoV-2 in six urban settings in Burkina Faso, Ghana and Madagascar.
► A two-stage cluster sampling strategy and stratification by age and sex are used to reduce selection bias.
► A highly specific anti-SARS-CoV-2 IgG ELISA reduces the probability of false positive test results.
► Behavioural and biological indicators at both participant and household level are investigated.
► Seroprevalence is estimated within a time frame in 2021 but not repeatedly over the course of the pandemic. The study covers a span of time relatively late in the course of the pandemic; due to antibody waning, seropositive individuals with infections dating back more than 10 months may be missed.

INTRODUCTION

The COVID-19 outbreak was declared a pandemic on 11 March 2020 by WHO. Over the course of 18 months, SARS-CoV-2 caused over 211 million cases and claimed more than 4 million lives worldwide.[1] In Africa, as of 24 August 2021, approximately 5.4 million cases and 130 407 deaths had been reported, representing 3% of the reported global cumulated cases and 3% of the reported global deaths.[1] Poverty, malnutrition, fragile healthcare systems and the burden of other infectious

diseases (eg, HIV) served as the basis for devastating morbidity and mortality predictions for the continent.[2] A younger age distribution, potentially protective immune history, warmer climate, lower overall population density, lower level of urbanisation and connectivity between geographical regions and lack of long-term care facilities have all been discussed as possible factors contributing to the relatively lower morbidity and mortality associated with COVID-19 in Africa.[3 4] Additionally, limited testing capacities, undetected mild cases or asymptomatic infections, and under-reported deaths might have obscured the true burden of COVID-19 in Africa.[5 6]

Population-based serological surveys that measure SARS-CoV-2 antibodies are essential in identifying the proportion of a population that has been exposed to the virus, even without symptoms or testing.[7] This information is crucial in determining the level of immunity within the population, thereby facilitating the implementation of a strategic preparedness and response plan to slow transmission, weaken recurring waves of the outbreak, and minimise the impact of the epidemic on health systems.[5 6 8] Furthermore, seroprevalence data can be used to guide vaccination strategies and deployment plans and to identify and prioritise vulnerable populations. As of 21 July 2021, more than 2031 serosurveys had been conducted worldwide to estimate SARS-CoV-2 seroprevalences.[9–14] However, only four studies were representative population-based studies using ELISAs: in Nigeria,[15] in Zambia,[16] South Sudan[13] and Kenya.[17] So far, no household-based serosurveys have been conducted in Burkina Faso, Madagascar or Ghana. Selection of a sensitive and specific SARS-CoV-2 ELISA is crucial, as some commercially available SARS-CoV-2 ELISAs have been shown to have a lower specificity for Ghanaian and Nigerian serum panels.[18]

Here, we publish a study protocol for a household-based cross-sectional study intended to estimate the prevalence of anti-SARS-CoV-2 IgG antibodies and identify associated risk factors for seropositivity in six urban settings across Burkina Faso, Ghana and Madagascar in 2021.

METHODS AND ANALYSIS
Study design
The SeroCoV study is a population-based one-time household cross-sectional prevalence investigation involving persons aged 10 years and older living in urban areas in selected African cities using a two-stage cluster sampling approach. The protocol was aligned with the WHO seroepidemiological standardisation initiative.[7]

Study setting
The SeroCoV study is conducted in the urban residential neighbourhoods (ie, study sites) of Ouagadougou and Bobo-Dioulasso in Burkina Faso, Accra, Kumasi and Tamale in Ghana, and Fianarantsoa in Madagascar. These urban sites were selected because incidence of SARS-CoV-2 infections has been reported to be higher in urban than in rural areas and because the respective communities were assumed, in the absence of reliable data due to limited access to testing capacities and undetected mild courses or asymptomatic infections, to already have been exposed to the virus.

Primary outcome
The primary outcome measure is the proportion of the population that has been exposed to the virus (referred to as seroprevalence) measured as SARS-CoV-2 IgG seropositivity based on ELISAs for blood plasma samples.

Secondary outcomes
Secondary outcome measures include the proportion of individuals with asymptomatic or subclinical infections, as well as possible indicators for assessing transmission of SARS-CoV-2. The seropositivity within an entire household will be determined for a subset of households.

Sample size
The study sample size was calculated in relation to the primary outcome. We assumed the seroprevalence to have reached at least 20% in the underlying study populations during the first year of the pandemic between March 2020 and until the beginning of the recruitment period in February 2021.[11] Several studies have suggested that antibody levels wane to below the limit of detection around 4–8 months after acute infection.[19 20] Therefore, it is possible that we missed individuals who were initially seropositive but had seroreverted by the time of sample collection. We based the effective sample size (ESS) calculation on a conservative prevalence estimate p of 50%, to enable determining prevalence ranging from 20% to 90% with sufficient statistical power, calculated as follows: ESS = $(z(1-\alpha/2)^2 p(1 p))/d^2$.[21 22] We assumed a significance level α of 5% and a relative precision d of ±5%. To account for the clustered design, as there is greater similarity of participant's SARS-CoV-2 test results within a cluster, we assumed a design effect of 1.45 (intra-cluster correlation coefficient: 0.05, mean cluster size of 10) yielding an actual sample size of 557. To account for 15% non-response of eligible households, we adjusted the number of participants needed to be recruited to 557/0.85=655.

Two-stage cluster geo-point sampling
In the first stage of the sampling approach, administrative boundaries were used to allocate clusters. Shapefiles with administrative boundaries and respective population statistics were obtained from in-country or online sources.[23 24] Clusters were selected using the probability proportional to population size method. In the second stage, the sampling frame was based on geographical coordinates (geo-point sampling – any point method) within the polygon(s) where the target population lived. Within each sampled unit, geographical coordinates were randomly selected as households to be recruited into the study.

Participant recruitment

After successful navigation to a set of global position system (GPS) coordinates using OsmAnd[25] software, eligible households will be identified. In case GPS coordinates do not fall onto a household, the closest household within a radius of 100 m will be identified. In case GPS coordinates fall on a building with multiple households, respective standard procedures to identify a household will be followed. In a scenario where several households appear to be similarly close, a 45° clockwise turn after facing north should be performed and the first household on the right selected. A team of healthcare workers and interviewers will approach a household member and introduce the study team and the study rationale. If participants are open to learning more, the study will be explained in greater detail. Each participant or their legal guardian must sign an informed consent form to participate in the study. Depending on the social structure of a study country, the household head (Ghana) or household representative (Burkina Faso and Madagascar) will be asked for permission to approach other members of the household. Only after a 'household consent' form has been signed will other household members be approached and individual informed consent forms obtained. Households are defined as a group of 2 or more individuals that sleep under the same roof and share their meals together (excluding residential institutions such as boarding schools, dormitories, hostels, prisons, other communities hosting groups of people). Household members are eligible for inclusion if they or their guardian provide written consent to participate, are at least 10 years of age, and do not have any known pathology or health problems contraindicated for blood sample collection. Study participants will be selected in line with their country's age and gender distribution. One person per household will be included in the study. In a random subset of 100 households across all six study sites (Bobo-Dioulasso, Ouagadougou, Kumasi, Greater Accra, Tamale, Fianarantsoa), the entire household will be sampled and interviewed in order to get a more accurate seroprevalence estimate and gain insights into infection levels within households. Children under the age of 10 years will be excluded. Qualified health professionals will then take a blood sample from each study participant and perform the interview using the Research Electronic Data Capture (REDCap)[26] Mobile App offline on any mobile device. If a selected household refuses to participate, new GPS coordinates will be selected.

Serological testing

SARS-CoV-2 IgG seropositivity will be determined based on an ELISA for blood plasma samples. Specifically, the IgG antibodies will be detected using a patented platform technology (EP2492689) using a recombinantly expressed, truncated SARS-CoV-2 nucleocapsid protein as an antigen. Assay specificity was determined during assay validation and found to be very high using pre-COVID-19 stock samples from Africa, in contrast to several commercially available tests.[18] Diagnostic and analytical sensitivity was determined using samples from PCR-confirmed SARS-CoV-2 patients.

Data analysis plan

To estimate the seroprevalence, we will first determine unadjusted frequencies of positive tests as a proportion of the final sample size along with exact 95% CIs. Second, we will estimate seroprevalence adjusted for test performance characteristics based on a Bayesian hierarchical model. Third, we will estimate seroprevalence adjusted for both test performance characteristics and population characteristics (ie, age and sex distribution and neighbourhoods) using a Bayesian multilevel regression with poststratification to adjust for differences between the sample and target population.[27] This approach will enable us to adjust estimates for test performance. We will furthermore exploratorily estimate relative risks of seropositivity among exposed and non-exposed individuals for various risk factors (eg, socioeconomic background, travel history, number of contacts per day, perception of COVID-19 and psychological consequences using a WHO scale) using log-binomial regression. Missing data on outcome measurements will be investigated with respect to baseline covariates to assess the potential for selection bias. In addition, location, distance and clustering of actual vs initially sampled GPS coordinates of households will be investigated to assess the potential for selection bias. Furthermore, the choice of priors for Bayes estimation when considering test performance will be considered in sensitivity analyses.

Patients and public involvement

Participants were not involved in the study design or in the setting of research objectives or outcomes of this study protocol.

ETHICS
Ethical approval

Ethical clearance was obtained by the national Ethical Board Committees of each participating country (National Ethical Committee, Ouagadougou, Burkina Faso, Ministry of Public Health in Antananarivo, Madagascar, Committee on Human Research, Publication and Ethics in Ghana, as well as by the Ethical Commission of the Ärztekammer Hamburg). The SeroCoV study protocol, questionnaire, information leaflet, and informed consent forms were reviewed and approved: Reference German Ethics: 2020-10035-BO and 2020-10035-1-BO, Reference Madagascar Ethics: CERBMIORG0000851, No 175-MSANP/SG/AGMED/CNPV/CERBM, Reference Burkina Faso Ethics: No 2020–137/MS/SG/INSP/CRSN, Reference Ghana Ethics: CHRPE/AP/218/20.

Safety considerations and dissemination plan

The risks and side effects from blood drawing will be explained to the participants. The data collected in the study will help improve and guide efforts to understand the extent of COVID-19 transmission in Burkina Faso, Ghana and Madagascar and may help prevent further transmission of the virus. Where applicable and

appropriate, information about the study including the protocol and questionnaires will be shared with key stakeholders including partner and participating institutions, regulators, local government and health authorities. Some countries also disseminated information about the study and its progress on social media. IgG results were shared with participants who donated blood in some participating countries. In addition, country-level data will be shared with respective national COVID-19 management teams and health service authorities to support local efforts in the management of COVID-19, including specific committee to share research results, for example, with Ghana Health Service in Ghana, Centre Régional de Commandement Opérationnel COVID-19. The results of the study will furthermore be disseminated to research and implementation partners, published in peer-reviewed journals, and presented at national and international conferences and events.

Project status

Recruitment started at all sites in February 2021 (3 February BF, 26 February MG, 18 February GH). As of August 2021, participant enrolment, household and member questionnaire assessment and biological sampling has been completed in Ouagadougou and Bobo-Dioulasso, Fianarantsoa and Kumasi, and is ongoing in the Greater Accra Region and Tamale.

Author affiliations

[1]Infectious Disease Epidemiology, Bernhard Nocht Institute of Tropical Medicine, Hamburg, Germany
[2]German Center for Infection Research Hamburg-Lübeck-Borstel-Riems Site, Hamburg, Germany
[3]Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Centre of the Johannes Gutenberg University Mainz, Mainz, Germany
[4]Universitatsklinikum Heidelberg Institut fur Global Health, Heidelberg, Germany
[5]German Center for Infection Research Heidelberg Site, Heidelberg, Baden-Württemberg, Germany
[6]Kumasi Centre for Collaborative Research in Tropical Medicine, Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana
[7]Department of Global and International Health, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[8]Infectious Disease Diagnostics, Bernhard Nocht Institute of Tropical Medicine, Hamburg, Germany
[9]Faculty of Medicine, University of Fianarantsoa, Fianarantsoa, Madagascar
[10]Centre d'Infectiologie Charles Méreiux, Antananarivo, Madagascar
[11]Universite d'Antananarivo, Antananarivo, Madagascar
[12]Centre de Recherche en Sante de Nouna, Nouna, Boucle du Mouhoun, Burkina Faso
[13]Department of Molecular Medicine, School of Medicine and Dentistry, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Disclaimer The content of this publication is solely the responsibility of the authors. Funding organisations did not have a role in the study design, hiring or training of staff, intervention delivery, or data collection.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Eva Lorenz http://orcid.org/0000-0002-6057-0078
Daniela Fusco http://orcid.org/0000-0001-8833-239X
Anthony Afum-Adjei Awuah http://orcid.org/0000-0002-8912-9673

REFERENCES

1 WHO. Weekly epidemiological update on COVID-19, 2021. Available: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-13-july-2021
2 Massinga Loembé M, Tshangela A, Salyer SJ, *et al*. COVID-19 in Africa: the spread and response. *Nat Med* 2020;26:999–1003.
3 Sajadi MM, Habibzadeh P, Vintzileos A, *et al*. Temperature, humidity, and latitude analysis to estimate potential spread and seasonality of coronavirus disease 2019 (COVID-19). *JAMA Netw Open* 2020;3:e2011834–e34.
4 Adams J, MacKenzie MJ, Amegah AK, *et al*. The conundrum of low COVID-19 mortality burden in sub-Saharan Africa: myth or reality? *Glob Health Sci Pract* 2021;9:433–43.
5 Kandel N, Chungong S, Omaar A, *et al*. Health security capacities in the context of COVID-19 outbreak: an analysis of international health regulations annual report data from 182 countries. *Lancet* 2020;395:1047–53.
6 Organisation WH. *Thematic paper on the status of country preparedness capacities*, 2019.
7 WHO. *Population-Based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (COVID-19) infection, 26 may 2020. version 2.0*. Geneva: World Health Organization, 2020.
8 WHO. *Critical preparedness, readiness and response actions for COVID-19*, 2021.
9 Arora RK, Joseph A, Van Wyk J, *et al*. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. *Lancet Infect Dis* 2021;21:e75–6.
10 Bobrovitz N, Arora RK, Yan T. Lessons from a rapid systematic review of early SARS-CoV-2 serosurveys. *medRxiv* 2020.
11 Chen X, Chen Z, Azman AS, *et al*. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. *Lancet Glob Health* 2021;9:e598–609.
12 Rostami A, Sepidarkish M, Leeflang MMG, *et al*. SARS-CoV-2 seroprevalence worldwide: a systematic review and meta-analysis. *Clin Microbiol Infect* 2021;27:331–40.
13 Wiens KE, Mawien PN, Rumunu J, *et al*. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Juba, South Sudan: a population-based study. *medRxiv* 2021. doi:10.1101/2021.03.08.21253009. [Epub ahead of print: 12 Mar 2021].
14 Franceschi VB, Santos AS, Glaeser AB. Population-Based prevalence surveys during the COVID-19 pandemic: a systematic review. *medRxiv* 2020.
15 Majiya H, Aliyu-Paiko M, Balogu VT. Seroprevalence of COVID-19 in niger state. *medRxiv* 2020.
16 Mulenga LB, Hines JZ, Fwoloshi S, *et al*. Prevalence of SARS-CoV-2 in six districts in Zambia in July, 2020: a cross-sectional cluster sample survey. *Lancet Glob Health* 2021;9:e773–81.
17 Ngere IA, Dawa J, Hunsperger E, *et al*. High seroprevalence of SARS-CoV-2 eight months after introduction in Nairobi, Kenya. *SSRN Journal* 2021.

18 Emmerich P, Murawski C, Ehmen C, *et al*. Limited specificity of commercially available SARS-CoV-2 IgG ELISAs in serum samples of African origin. *Trop Med Int Health* 2021;26:621–31.

19 Ripperger TJ, Uhrlaub JL, Watanabe M, *et al*. Orthogonal SARS-CoV-2 serological assays enable surveillance of low-prevalence communities and reveal durable humoral immunity. *Immunity* 2020;53:925–33.

20 Seow J, Graham C, Merrick B, *et al*. Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat Microbiol* 2020;5:1598–607.

21 Arya R, Antonisamy B, Kumar S. Sample size estimation in prevalence studies. *Indian J Pediatr* 2012;79:1482–8.

22 Daniel WW. *Biostatistics: a foundation for analysis in the health sciences*. 7 edn. Hoboken: John Wiley & Sons, Inc, 1999.

23 HDX THDE. *Madagascar - Subnational Administrative Boundaries - Humanitarian Data Exchange*, 2021.

24 HDX THDE. *Madagascar - Subnational Population Statistics - Humanitarian Data Exchange*, 2021.

25 Foundation FS. OsmAnd - Offline Mobile Maps and Navigation, 2021. Available: https://osmand.net/

26 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.

27 Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. *J R Stat Soc C* 2020;69:1269–83.