

**UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF**

Bernhard-Nocht-Institut für Tropenmedizin

Prof. Dr. med. Jürgen May

**Estimation and Characterization of the SARS-CoV-2  
Seroprevalence in Fianarantsoa, Madagascar**

**Dissertation**

zur Erlangung des Grades eines Doktors der Medizin  
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Leonard Gunga  
aus Berlin

Hamburg 2023

**Angenommen von der Medizinischen Fakultät der Universität Hamburg am:  
12.11.2024**

**Veröffentlicht mit Genehmigung der Medizinischen Fakultät der Universität  
Hamburg.**

**Prüfungsausschuss, der/die Vorsitzende:** Prof. Dr. Michael Ramharter

**Prüfungsausschuss, zweite/r Gutachter/in:** Prof. Dr. Jürgen May

# Table of Content

<b>1. Introduction.....</b>	<b>5</b>
1.1 Background .....	5
1.2 Primary Research Questions .....	6
1.3 Secondary Research Questions.....	7
<b>2. Methods .....</b>	<b>8</b>
2.1 Study design .....	8
2.2 Study setting .....	8
2.3 Sample size.....	10
2.4 Field teams .....	12
2.4.1 Staff .....	12
2.4.2 Equipment.....	13
2.5 Sampling strategy .....	14
2.5.1 Household identification .....	16
2.5.2 Participant recruitment .....	16
2.5.3 Questionnaire.....	17
2.6 Serological testing .....	18
2.6.1 Blood sampling .....	18
2.6.2 Samples transportation and storage .....	19
2.6.3 Samples processing .....	19
2.6.4 ELISA.....	20
2.6.5 Principle of the ELISA.....	20
2.7 Community engagement strategy .....	22
2.8 Compensation for participation .....	22
2.9 Data collection.....	22
2.10 Data analysis.....	22
2.11 Ethical considerations .....	23
2.12 Literature research terms.....	23
<b>3. Results.....</b>	<b>24</b>
<b>4. Discussion .....</b>	<b>44</b>
4.1 Major findings .....	44
4.2 SARS-CoV-2 seroprevalence among subgroups.....	44
4.3 Food and economic insecurities .....	47
4.4 Socioeconomic score .....	48
4.5 Mental health and well-being .....	49
4.6 Sources of information.....	51
4.7 Non-pharmaceutical protection measures .....	52

4.8 Strengths and limitations.....	53
<b>5. Conclusion .....</b>	<b>55</b>
<b>6. Abstract .....</b>	<b>57</b>
<b>7. Zusammenfassung .....</b>	<b>58</b>
<b>8. List of abbreviations.....</b>	<b>59</b>
<b>9. References.....</b>	<b>60</b>
<b>10. Acknowledgment .....</b>	<b>70</b>
<b>11. Curriculum vitae .....</b>	<b>71</b>
<b>12. Appendix(ces) .....</b>	<b>72</b>
12.1 Appendix A: Additional tablets .....	72
12.2 Appendix B: Additional figures.....	78
12.3 Appendix C: Code.....	79
12.4 Appendix D: Literature search terms .....	129
<b>13. Eidesstattliche Versicherung.....</b>	<b>131</b>

# 1. Introduction

## 1.1 Background

In December 2019, the zoonotic transmission of a novel coronavirus occurred at an animal food market in the Wuhan region, China (Guan et al. 2020; Smith Rodgers 2022). This new virus **Severe Acute Respiratory Syndrome Coronavirus type 2** (SARS-CoV-2) belongs to the group of betacoronavirus. It can cause the **Coronavirus Disease 2019** (COVID-19), which is characterized by the clinical appearance of flu-like symptoms, such as fever, dry cough, shortness of breathing or malaise, but also severe forms of pneumonia (Hu et al. 2021; Li et al. 2021). These severe forms are associated with respiratory failure and multi-organ dysfunction (National Institutes of Health 2023). The elderly people and those with underlying medical conditions, such as chronic diseases (e.g., diabetes or dialysis) or immunosuppression, are at increased risk for severe forms of COVID-19 infection (Centers for Disease Control and Prevention 2023). In the 21st century, two similar viruses, SARS-CoV-1 (2002) and MERS-CoV (2012) already caused outbreaks in China and in Middle East, also with severe manifestations of respiratory infection (Cui, Li, and Shi 2019). Similar to these viruses, respiratory person-to-person transmission of SARS-CoV-2 is also possible. In particular, aerosols containing SARS-CoV-2 and poorly ventilated indoor settings have been shown to increase the likelihood of transmission (Meyerowitz et al. 2021). Further, the infectiousness of SARS-CoV-2-infected individuals is the highest at the time of symptom onset (Meyerowitz et al. 2021). Unlike SARS-CoV-1 and MERS-CoV, in which mainly symptomatic patients were responsible for transmission, SARS-CoV-2 infected individual can be infectious even if they are pre-symptomatic, have no or only mild symptoms (Rice et al. 2022). In particular, this feature of SARS-CoV-2 accelerated its rapid global spread, leading to an exponential increase of cases and the declaration of a pandemic by the WHO on March 11, 2020 (World Health Organization 2020b). From then on, worldwide governments had to use public health measures (or non-pharmaceutical interventions (NPI)) to contain the spread of COVID-19. This ‘toolbox’ of measures ranged from simple individual protection through facemasks, regular handwashing, and coughing etiquette, to more drastic ones, such as mandatory quarantine, travel bans and nationwide lockdowns (Hale et al. 2021). As of March 21, 2023, the world had faced numerous waves of infection driven by the wild type coronavirus and its variants, resulting in a global toll of more than 761 Million infections (European Centre for Disease Prevention and Control 2023; World Health Organization 2020a). Despite the tragedy of nearly 7 Million officially confirmed deaths

worldwide, as well as major setbacks to economies and social challenges, the research community was capable to closely survey the course of the pandemic, and accomplished the rapid development of effective vaccines against COVID-19 as early as December 2020 (European Medicines Agency 2023). The research community provided an unprecedented and detailed assessment of the pandemic's impact across different scientific disciplines. However, there are huge regional differences in coverage. This applies especially to low-income countries, of which many are in sub-Saharan Africa (SSA) (World Bank 2021b; Wagner et al. 2022). Early predictions considered the SSA countries particularly vulnerable to the pandemic due to weaker health care systems, and limited testing and tracking capacities (BBC 2020). Instead, the pandemic seemed to have taken a milder course with fewer casualties in SSA, probably due to several factors, including younger median age and less travel activities (Cabore et al. 2022; Adams et al. 2021). Yet, it remains crucial to assess the pandemic's impact on SSA populations, as future outbreaks of infectious diseases or even pandemics are likely to occur in the next decades (World Health Organization 2023; Smitham and Glassman 2021).

In order to unravel the underlying level of exposure of COVID-19 in SSA, a cross-sectional study was conducted in urban populations in Ghana, Burkina Faso, and Madagascar in 2021. The Malagasy study site was in Fianarantsoa, which is the focus of this doctoral thesis (Lorenz et al. 2021). Given the scarce research regarding the impact of COVID-19 in Madagascar and an estimated nationwide reporting rate of less than 1 % of all cases, participants' seroprevalence (i.e., exposure to the virus) was identified, using a SARS-CoV-2 IgG ELISA (Cabore et al. 2022; Deschermeier et al. 2022). In addition, the social and economic implications caused by the pandemic (i.e., lockdowns) were assessed, by administering a questionnaire inquiring about socioeconomic, psychological and health status, and the compliance with non - pharmaceutical interventions.

## 1.2 Primary Research Questions

- (1) What was the SARS-CoV-2 seroprevalence overall and in subgroups of the surveyed-population, based on an IgG Enzyme-linked Immunosorbent Assays (ELISA)?
- (2) Which characteristics of the participants, such as sex, age, socioeconomic, individual consequences of non-pharmaceutical interventions, or comorbidities, were associated with a SARS-CoV-2 seropositivity?

### 1.3 Secondary Research Questions

- (3) Did the household-based sampling strategy yield enrollment of participants from a typical urban Malagasy population in terms of demographic, socioeconomic, health-related aspects (e.g., age and sex distribution, level of education and socioeconomic status)?

This doctoral thesis will contribute to the body of knowledge on the impact of COVID-19 in SSA by determining the SARS-CoV-2 seroprevalence in Fianarantsoa, Madagascar. Further, the investigation will help address the persistent lack of research in this area and provide a country-specific assessment to inform infectious disease prevention and treatment strategies for future outbreaks to policymakers involved.

## 2. Methods

### 2.1 Study design

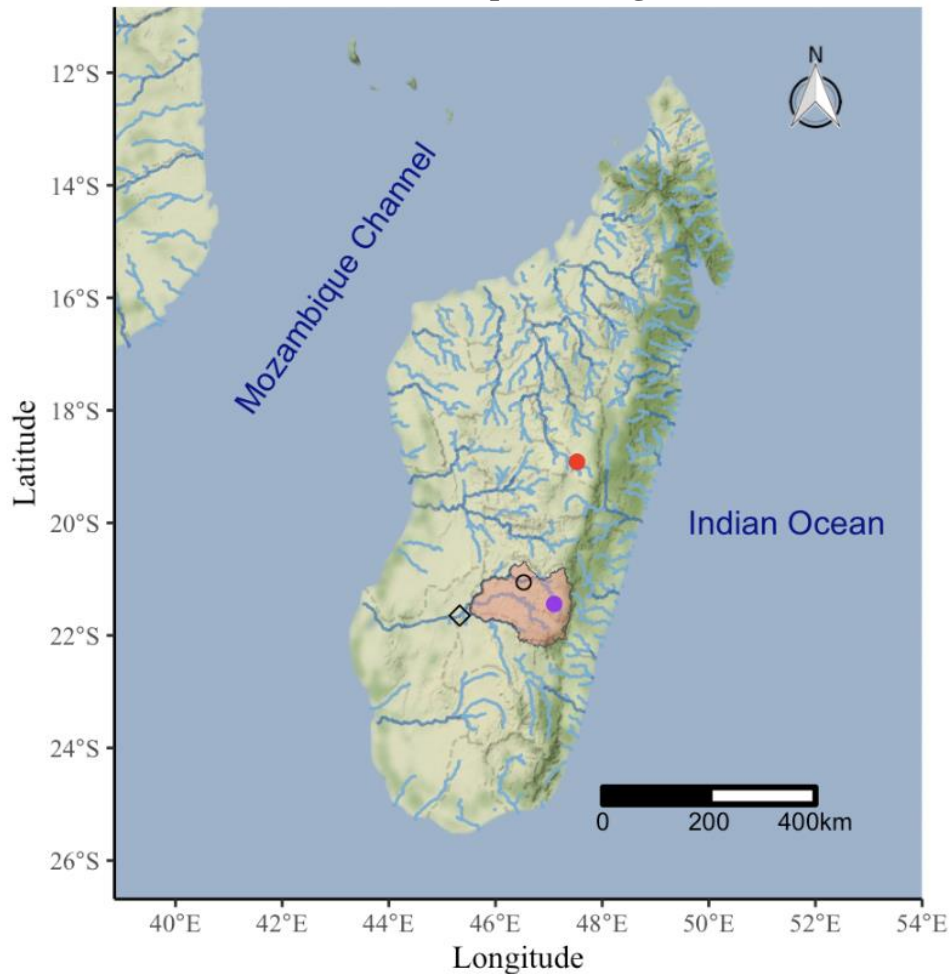
Shortly after the WHO declared the COVID-19 to be a pandemic, a study to unravel the spread and level of exposure to SARS-CoV-2 in sub-Saharan Africa was prepared (World Health Organization 2020b). During that time, many of the characteristics of SARS-CoV-2 that are now common knowledge were still unknown or in the process of coming to light. Hence, some decisions for the study design, sample size calculation or sampling strategy were based on assumptions rather than evidence. However, a cross-sectional study was conducted in urban residential areas of five cities in three African countries. Households were randomly enrolled for a two-stage cluster population-based IgG seroprevalence survey. To assess detailed information on the demography of the study population, the individuals' behavior, economic consequences of the imposed lockdown, and exposure level in urban settings, a survey questionnaire was administered to all participants (Lorenz et al. 2021).

### 2.2 Study setting

This doctoral thesis focuses on the Malagasy city Fianarantsoa in the southern highlands of Madagascar. The Republic of Madagascar is an African island-nation located in the Indian Ocean. The state is characterized by a coastal area and central Highlands (Britannica, The Information Architects of Encyclopaedia 2022). It has a surface of 587,295 km<sup>2</sup> and a population of 29 million inhabitants, with a population density of 49 people per km<sup>2</sup>. Most Malagasy live in rural settings. Yet, the percentage of people living in urban settings is steadily increasing since the 1960s, reaching 39% in 2021 (World Bank 2021c). As a former French colony, it gained independence on June 26<sup>th</sup>, 1960. The age distribution is typical for developing countries in the Global South, with a median age of 19.0 years, and with a Gross National Income (GNI) per capita of \$490 in 2021 (United Nations, Department of Economic and Social Affairs, Population Division 2022; World Bank 2021a). It was classified as a low-income country by the World Bank in 2021 (World Bank 2021b).



## Overview Map of Madagascar



**Figure 1**

The map shows the island nation of Madagascar with its capital Antananarivo (red dot) located in the Indian Ocean. The SeroCoV study site Fianarantsoa (purple dot) is in the regional capital of Haute-Matsiatra (orange area). The *Région* is named after the river Matsiatra (black circle), which flows into the Mangoky (black rectangle). The HydroRivers® data base was used for mapping the rivers (Lehner and Grill 2013). Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

Madagascar is administratively organized into 22 *Régions*, including the *Région* Haute-Matsiatra. This *Région* is named after the river Matsiatra and subdivided into seven districts, including the district of Fianarantsoa, which is equal to the surface of the city and the regional capital of Haute-Matsiatra (Fig. 1). The city is located at an average altitude of 1,200 m above sea level. The climate is subtropical with a warm rainy season (November to March), and a cold dry season (June to August). Its nearly 200,000 inhabitants in 2018, Fianarantsoa represented the fourth largest city in Madagascar. The city is located on highway 7 (*Route nationale*), which connects the Malagasy capital Antananarivo with southern regions (Institut National de la Statistique 2021). As such, it is an important economic hub. Furthermore, the Université de Fianarantsoa is one of the

largest, nation-wide universities with more than 24,000 Students ('Université de Fianarantsoa' 2022). Fianarantsoa translated into English means "Good education", emphasizing its cultural and educational importance. Fianarantsoa is further organized into *Communes*: Tanana Ambony, Tanana Ambany, Andrainjato Avaratra, Andrainjato Sud, Manolafaka, Lalazana, Vatosola. These Communes in turn consist of 50 Fokontanys, which are the smallest administrative unit. The population size for each Fokontany was available through subnational population statistics provided by the UN Office for the Coordination of Humanitarian Affairs (OCHA) and used for the weighted sampling strategy (The Humanitarian Data Exchange 2021b).

### 2.3 Sample size

The effective sample size was calculated using a level of significance  $\alpha$  of 5%, a conservatively assumed seroprevalence  $p$  of 50% after the first wave and a relative precision of  $\pm 5\%$ . Design effect of 1.45 was assumed for the two-staged cluster sampling strategy (Arya, Antonisamy, and Kumar 2012). The survey assumed a 15 % non – responder rate, which results in a targeted sample size of 650 participants.

$$\text{Effective Sample Size} = \frac{\left(z \left(1 - \frac{\alpha^2}{2}\right)\right)^2 * (P(1-P))}{d^2} * 1.45 * (1 - 0.15)^{-1}$$

$d$  (allowable margin of error) =  $\pm 5\%$

$z$  -value = 1.96 (if Confidence interval = 95%)

$\alpha = 1 - \text{CI}$

$P$  = expected prevalence of 50 %

Design effect = 1.45

Non-response of eligible households = 15 %

In Fianarantsoa, an additional 22 households were to be included in the survey, resulting in a total of 672 households. The reason for this was that multiples of 12 households were to be recruited depending on the population size of the Fokontany (Fig. 2). For example, 12 households were to be included in the less populated Fokontany Ambalabe and 48 households in the more populated Fokontany Ankofafalahy Ambany.



## 2.4 Field teams

### 2.4.1 Staff

The data collection was carried out by two local teams. Each team consisted of one physician, two nurses, and a driver. Medical staff were native Malagasy speakers and fluent in French. Before the start of the study, teams were trained by the local principal investigator and a BNITM field coordinator in three sessions. These trainings covered Good Clinical Practice (GCP) in clinical surveys and the implementation of study-specific Standard Operating Procedures (SOP), comprising proper household identification, venous blood sampling, sample storage and data collection with REDCap® (Research Electronic Data Capture) (Harris et al. 2009).

Off-road jeeps (Fig. 3) were used to navigate to households in the different Fokontanys of Fianarantsoa. In most cases, teams had to walk the last meters to the household because of poor roads (Fig. 4).

### SeroCoV Staff Preparing for a Daily Field Visit



**Figure 3**

The picture shows one of two borrowed off-road jeeps for sample transport in Fianarantsoa and to Ambositra and was taken in the courtyard of the infectious disease department at CHU Tambohobe in Fianarantsoa.

Left: Study physician

Right: BNITM staff member.

## Household Identification in Fianarantsoa



**Figure 4**

SeroCoV field team equipped with backpacks navigating to household coordinates in May 2021.

Front row: One of the two study nurses was carrying an official SeroCoV umbrella and the cooling box for the blood samples.

Back row: Study physician

### 2.4.2 Equipment

The navigation software OsmAnd® (OpenStreetMap Automated Navigation Directions) and a list of Fokontany were utilized to correctly identify eligible households ('OsmAnd' 2023). Both teams had backpacks for the study equipment and wore matching SeroCoV vests with badges to identify them as research team members. To inform about the study and recruit participants, the team was equipped with an information leaflet and study flyer explaining the background, purpose, and the procedure of the SeroCoV survey, as well as to introduce all involved investigators.

Two paper-based lists were used to ensure for traceability of enrolled households, and for matching the blood samples to the survey and household demographics. The first list consisted of the date, the name and the address of the participant, the Screening Identifier (SID), the Household Identifier (HID), the Member Identifier (MID) and checkboxes, indicating whether informed consent, blood sampling and questionnaire had been completed. Further, any remarks were notified in an open text section. The second list

was utilized for entire household sampling, comprising the unique HID and the MIDs for each sampled household member.

Informed consent was obtained before conducting the interviews and taking a blood sample. The answers were directly entered into the REDCap® data collection application installed on tablets (Fig. 5). All documents were available in either Malagasy or French. Batteries and power banks were kept in backpacks as a reserve for field-use, as well as one kilo packs of rice and sugar. Sampling kits for blood samples and a cooler were carried by staff. To guarantee adequate sample transportation, tracing forms were both used and signed upon takeover of responsibility, indicating the name of the person responsible and the date.

### Data Collection in the Field



**Figure 5**

The picture shows the living room of an included household. The SeroCoV study nurse filled out the questionnaire, using a tablet and the data collection software REDCap®.

Right: Study nurse

Left: Participant

## 2.5 Sampling strategy

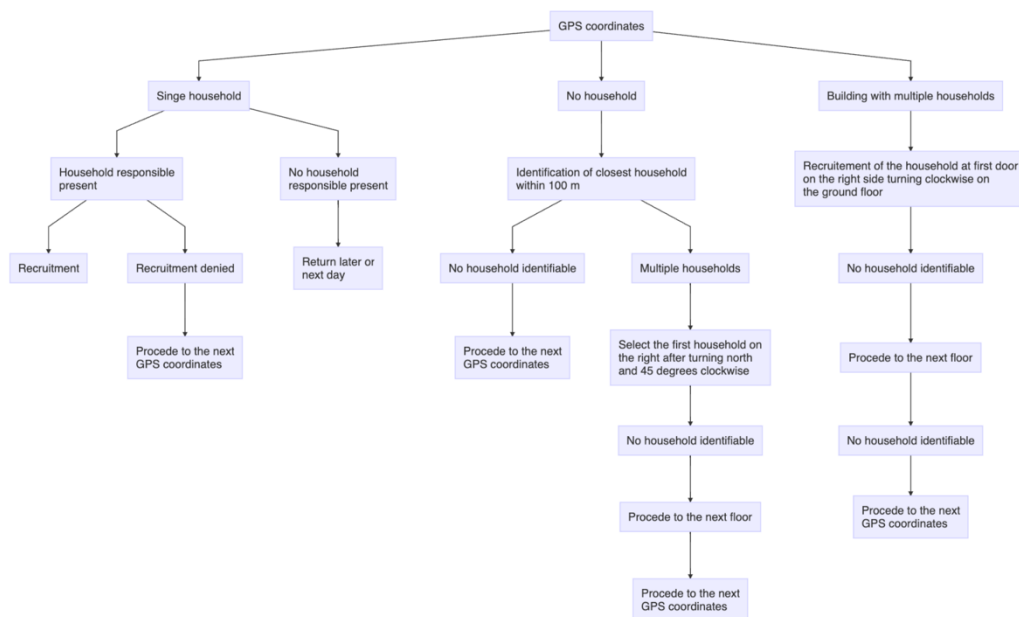
Eligible households were selected applying a two-staged randomized cluster approach. First, each Fokontany represents one or more clusters depending on the size of the population (Fig. 2). Second, randomly distributed GPS coordinates within these Fokontany were used to identify eligible households. Background information on



### 2.5.1 Household identification

An eligible household was defined as a home in which two or more individuals live under the same roof and share their meals. Residential institutions for instance boarding schools, dormitories, hostels, prisons, or other communities hosting groups of people were excluded. To identify eligible households, both field teams followed a SOP by navigating to GPS coordinates with OsmAnd® (Fig. 7).

**Flowchart of Household Identification with GPS Coordinates**



**Figure 7** Flowchart that illustrates the Standard Operating Procedure (SOP) for identifying households using GPS coordinates. The rules for different scenarios are shown, including selection procedure for single households, buildings with multiple households, households with no responsible person present, and no identifiable household.

For each Fokontany, a 30% back-up of GPS coordinates was provided in case no household could be identified.

### 2.5.2 Participant recruitment

To be included in the survey, participants needed to be 10 years or older and live in an eligible household. In addition, participants or their legal representative had to sign the informed consent form. Exclusion criteria were an incapacity to sign the informed consent or medical conditions which would not allow blood sampling. Within a household, participants were recruited depending on the above-mentioned stratification objectives (Tab. 1). Furthermore, in a subset of 100 randomly selected households, all household members were sampled and interviewed.



Before entering a household, physicians outlined the overall design of the study to the household representative and asked whether there is an interest in participating. If so, the physician asked for permission to enter the household and explained the survey in greater detail, including its background, aim, risks, questionnaire content and sampling procedure. In case of willingness to participate, informed consent was obtained by signing the form, or in case of illiteracy, leaving a thumbprint. For minor participants, the parent or legal guardian was required to provide consent. Participants could withdraw their consent at any time without giving a reason. Further, oral consent was obtained prior to taking photographs in participants' homes. After obtaining informed consent, the physician conducted the interview, while the nurses entered the data and took a blood sample. Further, the nurses were responsible for the appropriate (i.e., cooled) transport to the jeep, and then to the fridge in the data entry base, where the Clinical Research Assistant (CRA) took the responsibility for the samples.

### 2.5.3 Questionnaire

The survey consisted of 340 variables, providing information on participants' health status, socioeconomic status, and knowledge of the COVID-19 pandemic. The survey was administered in the more comfortable language of the participant, which was either Malagasy or French. The questionnaire was divided into four sections: Contact form, household, individual and closure. The first section contained general information for the participant's identification, which included SID, GPS coordinates of the enrolled household, date of interview and personal contact details. The subsequent household section comprised items of Demographic and Health Survey (DHS) Wealth Index (Rutstein and Johnson 2004) and previous contact to suspected or confirmed COVID-19 cases within or outside of the household. Further, the applied protection measures and economic consequences for the participant were assessed. Assessing living conditions and household assets, such as sanitary facilities, electricity, or electronic devices, the wealth index was used to classify the study population into five wealth quintiles ranging from lowest to highest, regardless of salary. The third section provided information on demographic, educational, and clinical background, e.g., age, sex, level of education, salary, daily social contacts, pre-existing health conditions, previous tests for acute SARS-CoV-2 infections, hospitalization, and medical treatments. Level of mental exhaustion over the last two weeks was measured with the WHO-5 Well-Being Index (WHO Regional Office for Europe 1998), which is a screening tool for depression. The

WHO-5 Well-Being Index has a sensitivity of 86% and specificity of 81%, ranging from 0 (bad well-being) to 100 (excellent well-being). The cut-off value for a 'screening diagnosis of depression lies at a score lower than 51 points (Topp et al. 2015). Questionnaire completeness was checked in the closure section.

## 2.6 Serological testing

### 2.6.1 Blood sampling

After completion of the questionnaire, venous blood was collected from the participant or the entire household (Fig. 8). Sampling material consisted of disinfectant (70% alcohol), cotton, plasters, tourniquet, venous puncture needle (Vacutainer® 22 G), 9 ml EDTA tube, member ID label and standard tube holder. Personal protection equipment for the team were non-sterile gloves, sharps container, biological waste bag and hydroalcoholic hand gel. Samples were stored in a cooler (styrofoam box) with cooling pads and daily placed into a fridge at the data entry base, located at the largest hospital in Fianarantsoa, Centre Hospitalier Université (CHU) Tambohobe.

### Venous Sampling Preparation



**Figure 8**

The picture shows the living room of an included household. The SeroCoV field team prepared a venous blood sampling of a minor household member after obtaining the caretaker's consent in May 2021.

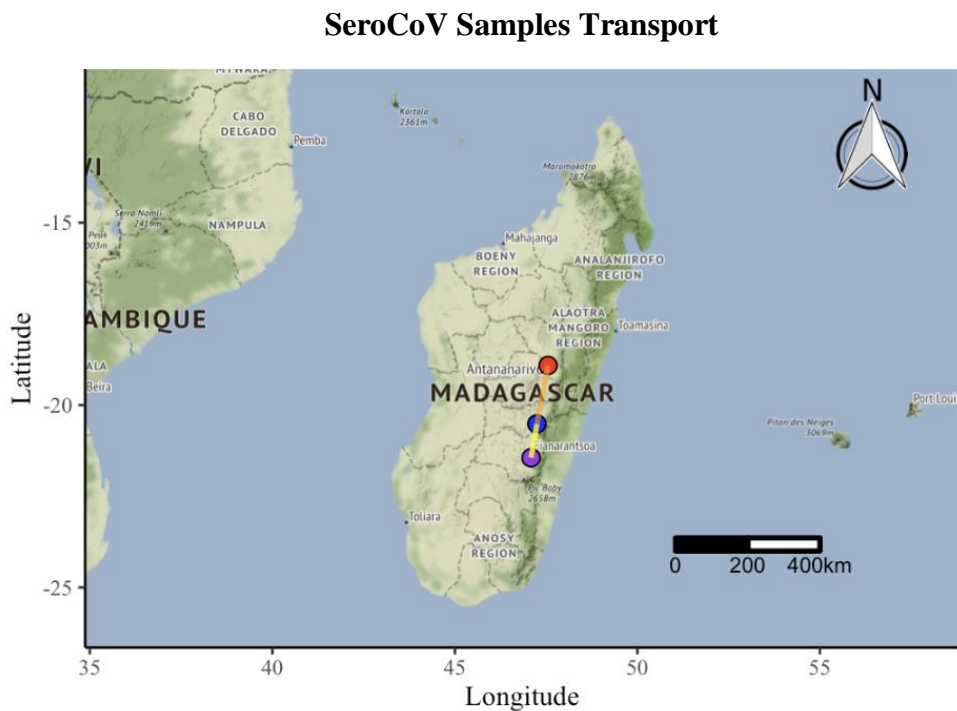
Right: Study physician

Center: Study nurse

Left: Participant

### 2.6.2 Samples transportation and storage

Samples were transported to the Centre d'Infectiologie Charles Mérieux (CICM) in the capital Antananarivo with a stop-over in Ambositra (Fig. 9). First, samples were transported from the CHU Tambohobe to a data entry station at Sokela hotel in Ambositra on Wednesday. There, samples were stored overnight in a fridge at 4 °C. On the next day, they were transported in a cooling box with a regularly commuting car to the CICM laboratory in Antananarivo. Here, they were stored at 4 °C until being processed in the laboratory.



**Figure 9**

The overview map shows the SeroCoV samples transport from the CHU Tambohobe in Fianarantsoa (purple dot) to the Centre d'Infectiologie Charles Mérieux (CICM) in Antananarivo (red dot) with a stopover in the Hôtel Sokela in Ambositra. The first route was carried out on Wednesday and took about 4 hours (yellow line). The second route was performed on Thursday by a regularly commuting car and took about 6 hours (orange line).

Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

### 2.6.3 Samples processing

Participants' blood sample were prepared for ELISA at the CICM laboratory facilities, using a BSL-2 safety cabinet, and following study specific procedures, in which participating staff were trained prior to the application.

First, the laboratory staff verified the correct sample identification by checking that the EDTA blood tube had a Member-ID sticker (S20-MXXXX-MG) on it. They also verified that the two cryotubes and one Eppendorf tube were labelled with the corresponding Sample-ID sticker (S20-MXXXX-MG-BEP-1, S20-MXXXX-MG-BEP-2, S20-

MXXXX-MG-BEP-3). BEP stands for Blood EDTA Plasma. EDTA tubes were constantly kept cool with ice while handling outside the fridge. The EDTA tubes were centrifugated at 2000 x g for 15 min at room temperature. Using a Pasteur pipette, 4 ml of the top layer, the blood plasma, were separated. Then, 1.7 ml of the plasma were transferred to cryotubes for long term storage and 0.6 ml to the Eppendorf tube. The latter (S20-MXXXX-MG-BEP-3) was subsequently used for the ELISA. Finally, the laboratory staff verified the correct sample identification, and used a laboratory sheet to notify any variation regarding the color or particles in the plasma.

To mitigate the risk of infection, potential SARS-CoV-2 cells in plasma were inactivated by adding 6 µl of 1 % Triton X-100, a non-ionic detergent, to the Eppendorf tube. The samples were stored in a 96-tube rack in the fridge, until the rack was full. For positive (1) and negative (2) test controls, 3 slots were left empty.

#### 2.6.4 ELISA

To measure SARS-CoV-2 seroprevalence in Fianarantsoa, an ELISA test kit was used for the detection of IgG antibodies against the SARS-CoV-2 nucleocapsid protein (NCP). The protocol of the applied ELISA has been previously published (Deschermeier et al. 2022). ELISA sensitivity and specificity was determined using samples from PCR-confirmed SARS-CoV-2 patients and pre-pandemic serum samples from German blood donors. Sensitivity was measured at two different concentrations at several time points post infection. Both assay dilutions showed a waning effect, resulting in significantly different sensitivities of 11.8 [2.0 – 35.6] vs. 82.3% [58.2 – 94.6].

Specificity of the ELISA was assessed prior to testing the study population. Pre-pandemic serum samples from Malagasy donors were used. Specificity was high in the Malagasy sub-sample set with 98.8% [95.4 – 100].

#### 2.6.5 Principle of the ELISA

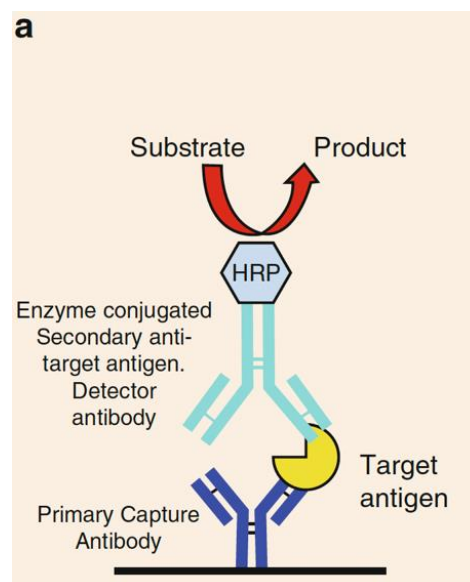
The Fcγ-Receptor-based SARS-CoV-2 ELISA used the concept of a “Sandwich – ELISA” (Fig. 10). The antigen of interest (i.e., IgG antibody against the NCP antigen) is thereby trapped between a primary capture antibody (i.e., the microplate coated with the Fcγ-Receptor) and a second soluble biotinylated detector antibody (Lin 2015). Both antibodies bind to different epitopes.

At the beginning of the assay, serum from the participants was added to a 96-well microplate coated with the Fcγ-Receptor. The Biotin-linked antibody was then added to

the coated microwells. In the next step, Streptavidin, which strongly binds to Biotin, linked to horseradish peroxidase (HRP) was added. Tetramethylbenzidine (TMB) is metabolized by HRP, resulting in a color change at a wavelength of 450 nm. Adding an acidic solution denaturizes the enzyme and stops the reaction. In between each step, the microplate was carefully washed with wash buffer to remove unbounded residuals.

Optical density (OD) measurement was used to determine the SARS-CoV-2 IgG antibody status, evaluating the wavelength absorbance at 450 nm to the reference wavelength of 620 nm. Test results were considered valid if two criteria were met. First, the OD difference (450 nm – 620 nm) of the positive control had to be > 90 % of the value reported in the lot specific certificate of analysis. Second, the OD difference of the negative controls had to be < 10 % of the value reported in the lot specific certificate of analysis. The following formula was used to calculate the optical density of the cut-off (ODCO): Average absorbance value of the negative controls ( $OD_{\text{negative av.}}$ ) + cut-off value reported in the lot specific certificate of analysis. The index values for the samples were calculated, using the formula:  $OD_{450} - OD_{620} (\text{sample}) / ODCO$ . An index value greater than 1.1 was considered positive, whereas an index value smaller than 0.9 was classified as negative. Index values of ambiguous results ranged from 0.9 to 1.1. Results were saved in an Excel® spreadsheet and uploaded into REDCap®.

### The Principle of a 'Sandwich'-ELISA



**Figure 10**

The figure illustrates the principle of a 'Sandwich'-ELISA, using a coated antibody (blue) to trap the antigen (yellow) and a second soluble biotinylated detector antibody (azure). Horse-radish-peroxidase (grey) metabolizes the substrate, resulting in a color change.

Source: Lin, A. V. Direct ELISA. in ELISA (ed. Hnasko, R.) vol. 1318 61–67 (Springer New York, 2015).

## 2.7 Community engagement strategy

In the Malagasy society, community agents (*“Agent communautaire (AC)”*) play an important role in the communication of health-related topics e.g., support of regular vaccination, disease awareness programs or distribution of protection items such as mosquito nets (Razanaboninahitra 2020). Therefore, the local principal investigator (PI) contacted all relevant ACs prior to the study. He informed them about the planned project, study procedures and risks for participants. The ACs assessed the general interest among the population to participate in the study. In addition, the PI announced the study on the local university radio “Radio ROFIA 91.4 MHz”. The test results were communicated to the regional Malagasy coronavirus taskforce Régional de Commandement Opérationnel COVID-19 (CRCO) and the Ministry of Health.

## 2.8 Compensation for participation

At the end of the visit, sampled participants were offered a compensation of one kilo of rice and sugar per household. Participants with pre-existing diabetes received two packages of rice.

## 2.9 Data collection

Data collection took place from 26<sup>th</sup> February to 18<sup>th</sup> June 2021. A clinical research assistant verified data entry from the field on a daily basis and she prepared weekly reports of the stratification targets, which were then shared with the field coordinator and principal investigator. The reports were discussed in online team meetings via Skype® with extended team bi-weekly. After completion of data collection, the database was checked for wrong entries (e.g., typos or double entries) and inconsistencies.

## 2.10 Data analysis

Data analysis was performed using statistical software R®4.2.2 and associated packages (R Core Team 2021). Tidyverse®, here®, jpeg®, magick®, pyramid®, ggcorrplot®, cowplot®, DiagrammeR®, ggtext® and ggrepel® were used for data preparation and plotting, ggmap®, sf® and ggsn® for spatial visualization, psych® for calculating the wealth index quintiles, and gtsummary® and gt® packages to create tables presented here (Wickham et al. 2019; Müller 2020; Urbanek 2022; Ooms 2023; Kamvar 2023; Kassambara 2022; Wilke 2020; Iannone 2022; Wilke and Wiernik 2022; Kamvar

2023; Kahle and Wickham 2013; Pebesma 2018; Baquero 2019; Revelle 2022; Hjelm et al. 2017; Sjoberg et al. 2021; Iannone et al. 2023).

For readability of the main text some of the tables and figures, which contain more detailed information, can be found in the appendices. These are the tables 3, 4, 5.2, 5.3, 5.4, 6.2, 6.3, 6.4, 9.2, 10.2, 10.3, 11 and 16 in appendix A, and the figures 18.1 and 18.2 in the appendix B.

### 2.11 Ethical considerations

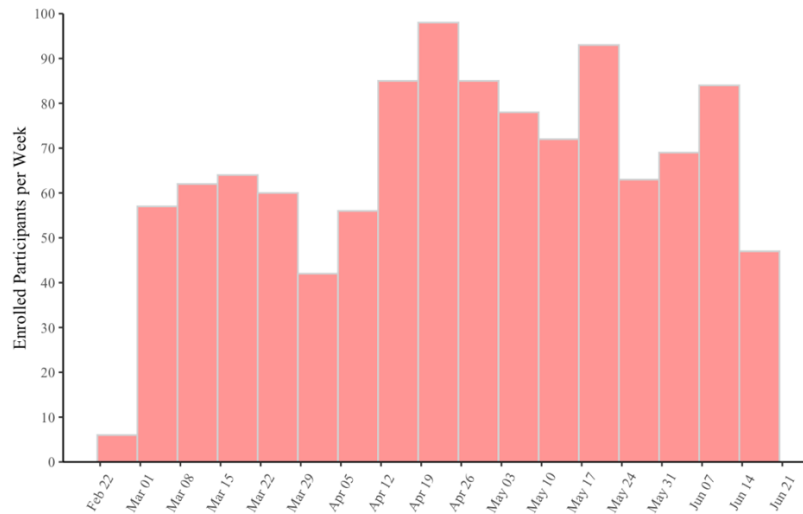
Ethical approval was obtained in all participating countries. National ethics committees reviewed and accepted the SeroCoV study protocol as well as the questionnaire, information leaflet, and informed consent form. The approval reference from Malagasy ethical authority Comité d’Ethique de la Recherche Biomédicale (CERBM) was CERBMIORG0000851, No 175-MSANP/SG/AGMED/CNPV/CERBM. German Ethical Commission of the Ärztekammer Hamburg accepted the submitted document under the references 2020-10035-BO and 2020-10035-1-BO.

### 2.12 Literature research terms

Prior to analyzing the dataset for Madagascar, a systematic literature search was conducted in PubMed® on published SARS-CoV-2 studies. The SARS-CoV-2 pandemic has generated a high number of publications worldwide. The articles included focused on publications regarding the impact of COVID-19 in sub-Saharan African countries in general, as well as local SARS-CoV-2 seroprevalence studies. Further, the literature was screened for similar urban SSA study populations, living in comparable wealth conditions, and consequences for mental Health during the pandemic. The search terms used are summarized in the appendix D.

### 3. Results

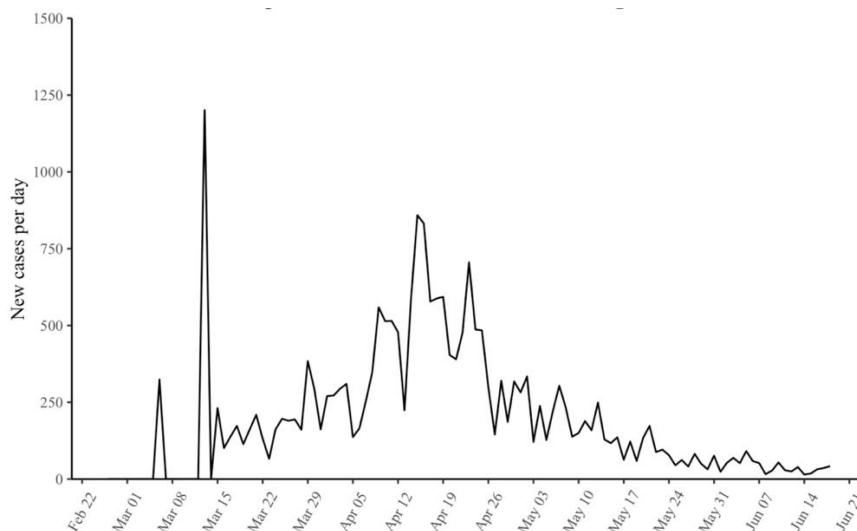
#### Data Collection Over Time



**Figure 11**  
The histogram shows the SeroCoV sampling period and the weekly inclusion rate of participants.

Data collection was performed from the February 22<sup>nd</sup> to June 18<sup>th</sup>, 2021. Figure 11 shows number of included participants per week. Successful recruitment was lowest during the first week of the study due to initial team organization issues and initial low participation within the population. The dip in recruitment around April 2, 2021 is due to the Easter holiday. Highest participation rates were reached from in the weeks of April 19<sup>th</sup> to 25<sup>th</sup>, May 17<sup>th</sup> to 23<sup>rd</sup>, and June 07<sup>th</sup> to 13<sup>th</sup> 2021.

#### Development of COVID-19 Cases in Madagascar



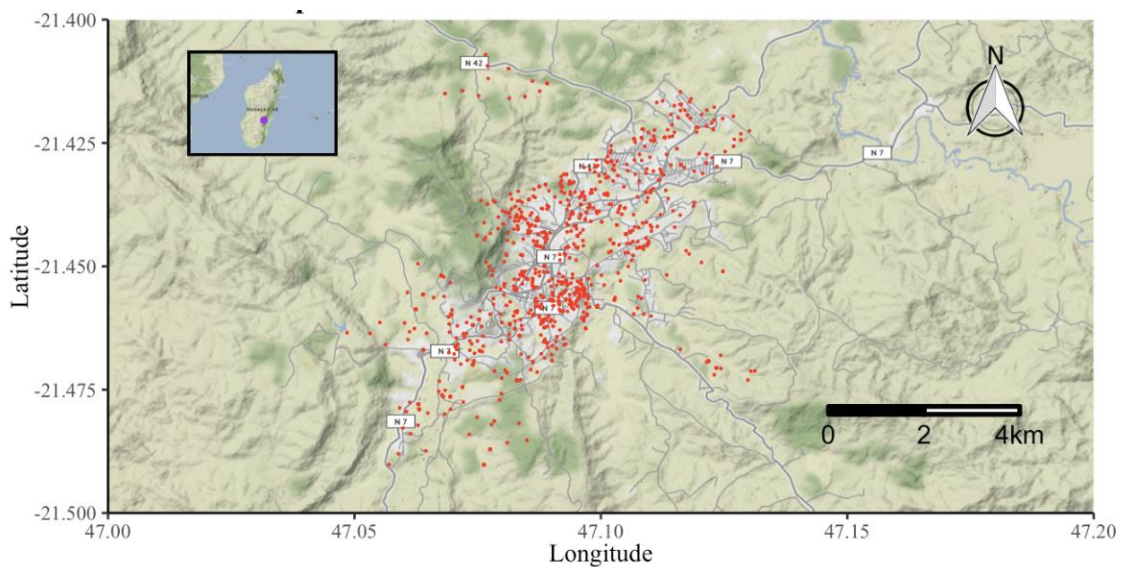
**Figure 12**  
The figure illustrates the official number of daily reported new COVID-19 cases during SeroCoV sampling period in 2021.

Source: <https://ourworldindata.org/coronavirus>



During data collection, a second wave of COVID-19 hit Madagascar. Figure 12 presents the official daily reported cases during the sampling period. The first and second spike of reported cases with more than 250 and 1000 cases, respectively, were caused by a change of reporting practice from weekly to daily notifications (Mandimbisoa 2021). From March 15<sup>th</sup> onwards, daily new cases of COVID-19 increased to a peak of 859 on April 15<sup>th</sup>. From then on, official daily cases dropped gradually to the end of the study period.

### Spatial Distribution of Included Households

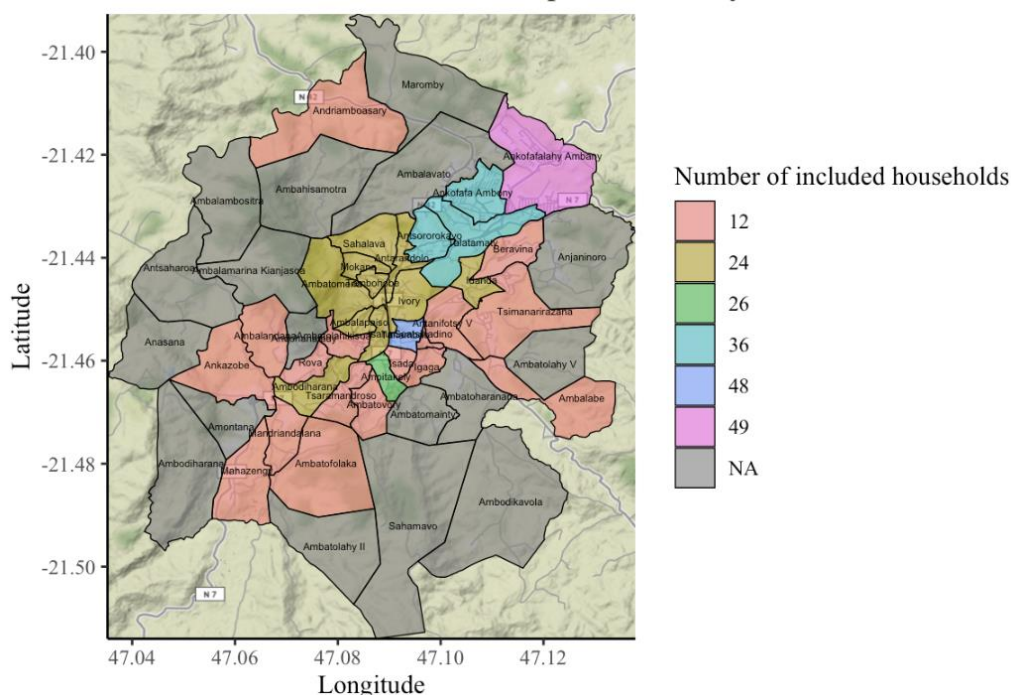


**Figure 13**

The map presents the GPS coordinates (red dots) of included households. The highway 7 ('Route nationale' 7) connects Fianarantsoa with the Malagasy capital Antananarivo. An overview map of the island state Madagascar is provided in the top-left corner and shows the location of Fianarantsoa in the southern highlands (purple dot). Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

Spatial distribution of included households in Fianarantsoa is visualized in Figure 13. Prior to the launch of the study, intended numbers of households to be sampled per Fokontany were calculated based on the two-stage population sampling strategy. Figure 14 indicates the household numbers that were sampled per Fokontany. Yet, in Fokontans Ampitakely and Ankofafalahy Ambany two respectively one household were even too much included.

### Number of Households included per Fokontany



**Figure 14**  
The map shows the number of included households per Fokontany. The excluded Fokontany are colored in grey. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

In total, 1121 participants living in 674 households were included in the survey. Among them, 420 Participants (37.5%) were tested positive for anti-SARS-CoV-2 IgG antibodies. Among the 701 negative tested participants (62.5%), three surveyed participants had an ambiguous test result, which represents 0.3% of all samples (Tab. 2).

**Table 2.** Participants ELISA Test Results - SeroCoV Madagascar

CHARACTERISTIC	NUMBER OF PARTICIPANTS	N (%) <sup>1</sup>	95% CI <sup>2</sup>
<b>IgG ELISA</b>	1,121		
positive		420.0 (37.5%)	[35; 40%]
negative <sup>3</sup>		701.0 (62.5%)	[60; 65%]

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

<sup>3</sup>incl. 3 ambiguous ELISA Test Results

The targeted stratification goals per age group and sex were achieved in all subgroups (Tab. 3, Appendix A). Data were collected either on a household or an individual level, which is summarized in Table 4 (Appendix A).

The surveyed study population had a median age of 29 years old (Tab. 5.1). The largest age groups were those of 10 to 19 and 20 to 29 years, with both groups together representing 50% (561) of the total population. Participants who were older than 59 years were grouped together but represented only 8.9% of the total population. Seropositivity rate among all age groups ranges from 30% to 45%. Interestingly, the age group from 30 to 39 years had the lowest seropositivity (30%). Slightly more women (52%) than men (48%) were included in the study. No significant difference regarding the seropositivity stratified by sex was detected (f: 38 % vs. m: 36%, Tab. 5.1). The median age of SARS-CoV-2 seropositive ELISA test result varied by the participant sex (Tab. 5.2, Appendix A). Female participants with a positive ELISA test result had a higher median age of 32 (IQR: 20, 47) compared to men with a median age of 26 (IQR: 17, 41).

Across female age groups, the SeroCoV study observed the highest seropositivity in women aged  $50 \leq 59$  years (47% [34; 60%]). In contrast, among male participants, the youngest subgroup (i.e.,  $10 \leq 19$  years old) had the highest seropositivity (46% [37; 54]). The lowest seroprevalence was found in participants aged  $30 \leq 39$  years in both age groups (f: 32% [24; 43%] and m: 27% [19; 37%], Tab. 5.2).

**Table 5.1** Demographics of Participants

CHARACTERISTIC	N	ELISA Test Result			
		positive, N = 420 <sup>f</sup>	95% CI <sup>2</sup>	negative, N = 701 <sup>f</sup>	95% CI <sup>2</sup>
<b>Age</b>	1,121				
Median (IQR)	29 (20, 43)				
<b>Age groups</b>	1,121				
10 ≤ 19 years	278 (25%)	118 (42%)	[37; 49%]	160 (58%)	[51; 63%]
20 ≤ 29 years	283 (25%)	98 (35%)	[29; 41%]	185 (65%)	[59; 71%]
30 ≤ 39 years	212 (19%)	63 (30%)	[24; 36%]	149 (70%)	[64; 76%]
40 ≤ 49 years	151 (13%)	60 (40%)	[32; 48%]	91 (60%)	[52; 68%]
50 ≤ 59 years	97 (8.7%)	44 (45%)	[35; 56%]	53 (55%)	[44; 65%]
≥ 60 years	100 (8.9%)	37 (37%)	[28; 47%]	63 (63%)	[53; 72%]
<b>Sex</b>	1,121				
Female	584 (52%)	224 (38%)	[34; 42%]	360 (62%)	[58; 66%]
Male	537 (48%)	196 (36%)	[32; 41%]	341 (64%)	[59; 68%]
<b>Education</b>	671				
No degree	62 (9.2%)	22 (35%)	[24; 49%]	40 (65%)	[51; 76%]
Primary	179 (27%)	52 (29%)	[23; 36%]	127 (71%)	[64; 77%]
Secondary school	295 (44%)	113 (38%)	[33; 44%]	182 (62%)	[56; 67%]
Baccalauréat	60 (8.9%)	25 (42%)	[29; 55%]	35 (58%)	[45; 71%]
University student	25 (3.7%)	10 (40%)	[22; 61%]	15 (60%)	[39; 78%]
License	19 (2.8%)	9 (47%)	[25; 71%]	10 (53%)	[29; 75%]
Master's degree	14 (2.1%)	4 (29%)	[9.6; 58%]	10 (71%)	[42; 90%]
Other	17 (2.5%)	4 (24%)	[7.8; 50%]	13 (76%)	[50; 92%]
Unknown	3	1		2	
<b>Employment status<sup>3</sup></b>	965				
Yes	786 (81%)	286 (36%)	[33; 40%]	500 (64%)	[60; 67%]
No	109 (11%)	46 (42%)	[33; 52%]	63 (58%)	[48; 67%]
Other	70 (7.3%)	32 (46%)	[34; 58%]	38 (54%)	[42; 66%]
Unknown	41	10		31	
<b>Profession<sup>3</sup></b>	965				
Farming	53 (5.5%)	14 (26%)	[16; 41%]	39 (74%)	[59; 84%]
Employed	41 (4.2%)	14 (34%)	[21; 51%]	27 (66%)	[49; 79%]

Self-employed	485 (50%)	165 (34%)	[30; 38%]	320 (66%)	[62; 70%]
Governmental employee	17 (1.8%)	10 (59%)	[33; 81%]	7 (41%)	[19; 67%]
Cleaning lady	38 (3.9%)	16 (42%)	[27; 59%]	22 (58%)	[41; 73%]
House keeper/ homemaker	84 (8.7%)	34 (40%)	[30; 52%]	50 (60%)	[48; 70%]
Student/ Intern/ Apprentice	152 (16%)	67 (44%)	[36; 52%]	85 (56%)	[48; 64%]
Retired	36 (3.7%)	18 (50%)	[34; 66%]	18 (50%)	[34; 66%]
Unemployed (health reasons)	2 (0.2%)	1 (50%)	[9.5; 91%]	1 (50%)	[9.5; 91%]
Unemployed (other reasons)	23 (2.4%)	11 (48%)	[27; 69%]	12 (52%)	[31; 73%]
Not applicable (under 18)	23 (2.4%)	11 (48%)	[27; 69%]	12 (52%)	[31; 73%]
Other	11 (1.1%)	3 (27%)	[7.3; 61%]	8 (73%)	[39; 93%]
Unknown	41	10		31	
<b>Main source of income<sup>3</sup></b>	<b>848</b>				
Agriculture	55 (6.5%)	13 (24%)	[14; 37%]	42 (76%)	[63; 86%]
Transportation	19 (2.2%)	6 (32%)	[14; 57%]	13 (68%)	[43; 86%]
Building and construction	29 (3.4%)	10 (34%)	[19; 54%]	19 (66%)	[46; 81%]
Business, trade, retail	470 (55%)	164 (35%)	[31; 39%]	306 (65%)	[61; 69%]
Health care	8 (0.9%)	6 (75%)	[36; 96%]	2 (25%)	[4.5; 64%]
Education	139 (16%)	65 (47%)	[38; 55%]	74 (53%)	[45; 62%]
Landlords or public officials	17 (2.0%)	8 (47%)	[24; 71%]	9 (53%)	[29; 76%]
Other	111 (13%)	45 (41%)	[31; 50%]	66 (59%)	[50; 69%]
Unknown	158	57		101	

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

<sup>3</sup>Only participants older than 14 years were included.

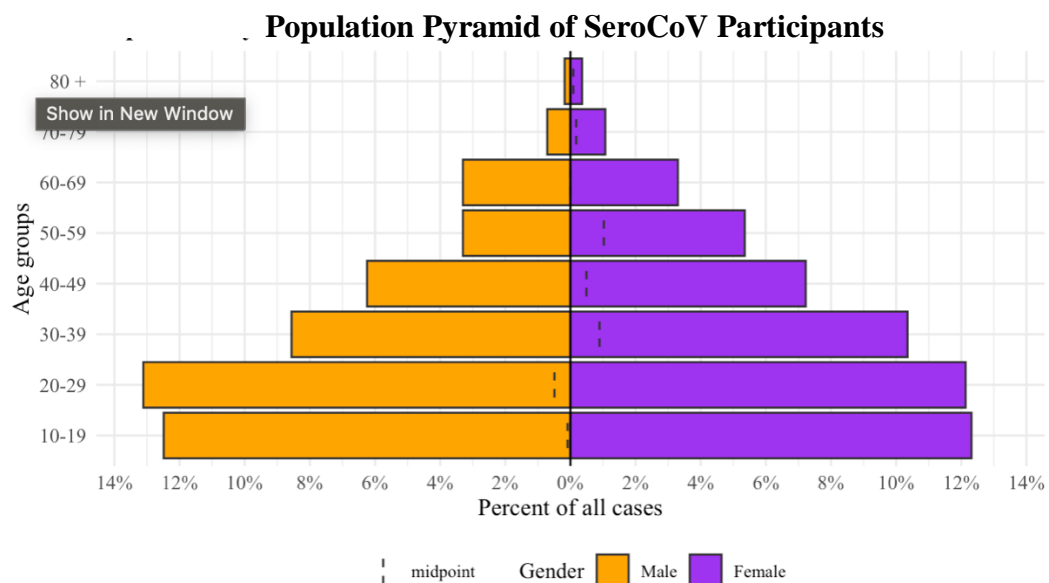
The level of education was assessed on a household level (n=674). Most of the participants had a secondary school degree (44%), followed by completion of primary school (Tab. 5.1). Less than one tenth (9.2%) had no degree. Likewise, 8.9% of participants accomplished a *baccalauréat* permitting to apply for university. A small proportion of participants (3.9%) were currently studying at a university at the time of the survey, or already had an academic degree (2.1%). Comparing SARS-CoV-2 serostatus with respect to education showed that the highest seropositivity was detected among *license* (47%) and *baccalauréat* graduates (42%), as well as university students (42%). Lowest seropositivity was observed among the small ‘Other’ subgroup (24%). No significant differences in seropositivity were found across levels of education. The different Malagasy levels of education are presented in Table 5.3 (Appendix A) with their respective English and German equivalence.

All participants older than 14 years old were included in the employment status. Hence, the employment status was assessable for 956 participants, of whom 81 % were employed. The employment status categorized as ‘Other’ included mainly retired participants (7.3%). For 41 participants (4.1%), the employment status was not collected. Unemployed participants (11%) had a higher seropositivity (42%).

The majority of employed participants were self-employed (50%) which includes shop – owners and craftsmen. The second-largest group were working in the educational sector (16%), followed by housekeeper/homemaker (8.7%), farming (5.5%), and cleaner

(3.9%). Looking at seropositivity within occupational groups, farmers had the lowest seropositivity (26%) and governmental employees the highest (59%). The levels of education within the different professions are shown in Table 5.4 (Appendix A).

The main source of income was indicated by 848 participants, of whom only participants older than 14 years old were included. Consistent with the high rate of self-employed participants, most of the participants (55%) named business, trade, or retail as their main source of income. Working in the educational sector (e.g., teacher) was the second most common occupation (16%). Looking at seropositivity across different occupational groups showed the highest rate among health care workers (75%) and the lowest in participants involved in agriculture (24%). However, the group sizes were comparatively small (n = 8 for health care and 13 for agriculture).



**Figure 15**

The population pyramid illustrates the age distribution of participants by 10 years age groups and sex.

The population pyramid presents the age distribution of the included study participants in percent, stratified by 10-year groups and sex. The largest subgroup was men aged 20 to 29 years old. As for women, the age group 10 - 19 was the largest (Fig. 15).

The majority (97%) of participants were not tested for acute SARS-CoV-2 (PCR on nasal swab) prior to the study (Tab. 6.1). Confirmed exposure to an acute SARS-CoV-2 infection was defined by contact to a suspected or confirmed case within the household or outside the household, or a positive PCR test. This was the case for 189 participants (17%). Seropositivity did not differ significantly between the two groups (39% and 37%, respectively). In addition, 545 participants (49%) reported to have experienced typical

COVID-19 symptoms over the past 12 months, e.g., cough, loss of smell and taste, fever, shortness of breathing, respiratory problems, as well as fatigue. Of the 545 participants with COVID-19 symptoms in the past year, most indicated that they visited a health care facility (42%). While traditional self-medication (23%) was more common than street pharmacy (19%), modern self-medication (12%) or private pharmacy, only 3 participants (0.6%) consulted a traditional healer (Tab. 6.2, Appendix A).

Chronic health concerns in the surveyed population were rare. Of those study participants who were chronically ill (n = 104, 9.3%), the majority suffered from high blood pressure (n = 51) or lung diseases (n = 17) (Tab. 6.3, Appendix A). Although tobacco smoking was generally rare among participants (16%), there was a highly significant difference by sex. Only 10 smokers were female (5.6%) compared to 167 male smokers (94%) (Tab. 6.4, Appendix A)

**Table 6.1.** Clinical and Epidemiologic Characteristics of Participants

Characteristic	N	ELISA TEST RESULT			
		positive, N = 420 <sup>1</sup>	95% CI <sup>2</sup>	negative, N = 701 <sup>1</sup>	95% CI <sup>2</sup>
<b>Previous SARS-CoV-2 testing</b>	1,119				
Yes	31 (2.8%)	15 (48%)	[31; 67%]	16 (52%)	[33; 69%]
No	1,087 (97%)	404 (37%)	[34; 40%]	683 (63%)	[60; 66%]
Do not know	1 (<0.1%)	0 (0%)	[0.00; 95%]	1 (100%)	[5.5; 100%]
Unknown	2	1		1	
<b>Exposure confirmed</b>	1,121				
Yes	189 (17%)	74 (39%)	[32; 47%]	115 (61%)	[53; 68%]
No	927 (83%)	344 (37%)	[34; 40%]	583 (63%)	[60; 66%]
Do not know	5 (0.4%)	2 (40%)	[7.3; 83%]	3 (60%)	[17; 93%]
<b>COVID-19 symptoms over the past 12 months</b>	1,121				
Yes	545 (49%)	207 (38%)	[34; 42%]	338 (62%)	[58; 66%]
No	569 (51%)	209 (37%)	[33; 41%]	360 (63%)	[59; 67%]
Do not know	7 (0.6%)	4 (57%)	[20; 88%]	3 (43%)	[12; 80%]
<b>At least one underlying chronic health concern</b>	1,120				
Yes	104 (9.3%)	46 (44%)	[35; 54%]	58 (56%)	[46; 65%]
No	1,014 (91%)	373 (37%)	[34; 40%]	641 (63%)	[60; 66%]
Do not know	2 (0.2%)	1 (50%)	[9.5; 91%]	1 (50%)	[9.5; 91%]
Unknown	1	0		1	
<b>Daily contacts</b>	1,111				
less than 5	21 (1.9%)	8 (38%)	[19; 61%]	13 (62%)	[39; 81%]
5 to less than 10	144 (13%)	56 (39%)	[31; 47%]	88 (61%)	[53; 69%]
10 to less than 50	714 (64%)	263 (37%)	[33; 41%]	451 (63%)	[59; 67%]
50 or more	232 (21%)	86 (37%)	[31; 44%]	146 (63%)	[56; 69%]
Unknown	10	7		3	
<b>Household size</b>	1,121				
1 to 2	81 (7.2%)	25 (31%)	[21; 42%]	56 (69%)	[58; 79%]
3 to 5	430 (38%)	145 (34%)	[29; 38%]	285 (66%)	[62; 71%]
≥ 6	610 (54%)	250 (41%)	[37; 45%]	360 (59%)	[55; 63%]

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

The number of daily contacts were categorized into 5 subgroups (Tab. 6.1). Many participants had 10 to less than 50 contacts per day (64%), followed by 50 or more

contacts (21%). Only 21 participants (1.9%) had less than 5 contacts per day. Among the above-mentioned clinical and epidemiological characteristics, no significant differences regarding the ELISA test result were detected.

Household size was grouped into 3 sizes: 1 to 2 household members, 3 to 5 household members, and households with equal or more than 6 household members. Households with only one or two members were generally rare (7.2%), compared to households of 3 to 5 persons (38%) and large households of equal to more than 6 household members (54%). Positive ELISA test results increased with the household size (households of 1 to 2 members: 31% [21; 42%] vs. household with 6 and more members 41% [37; 45%]).

The potential exposure to the coronavirus on the way to school or work, and at school or work was overall assessable for more than 800 participants, of whom the great majority (92%) walked to their destination (Tab. 6.5). This subset had also a lower seropositivity, compared to using the bus (41%) or a car (54%). Other forms of transportation were only used very scarcely. Participants with the fewest contacts on the way to school or work (i.e., less than 5 person) had the lowest seropositivity, but no trend was evident as the seropositivity decreased from the second-fewest contacts (38%) to the subset with the most contacts on the way (34%). Regarding the environment at school or work, participants who spent most of their time outdoors had a significantly lower seropositivity (33% [29; 37%]), than those staying indoors (43% [38; 48%]). Only two participants reported working in hospitals, none of them were positive. Working and living in the same house was only practiced by a minority (18%), yet the ELISA was more often positive (40% vs. 26%).

**Table 6.5** Coronavirus Exposure on the Way to School or Work, and at School or Work

Characteristic	N	ELISA TEST RESULT			
		positive, N = 420 <sup>l</sup>	95% CI <sup>2</sup>	negative, N = 701 <sup>l</sup>	95% CI <sup>2</sup>
<b>How do you get to school or work every day?</b>	813				
Walk	746 (92%)	267 (36%)	[32; 39%]	479 (64%)	[61; 68%]
Bus	32 (3.9%)	13 (41%)	[24; 59%]	19 (59%)	[41; 76%]
Car	26 (3.2%)	14 (54%)	[34; 73%]	12 (46%)	[27; 66%]
Bicycle	2 (0.2%)	1 (50%)	[9.5; 91%]	1 (50%)	[9.5; 91%]
Taxi	1 (0.1%)	0 (0%)	[0.00; 95%]	1 (100%)	[5.5; 100%]
Motorcycle	1 (0.1%)	0 (0%)	[0.00; 95%]	1 (100%)	[5.5; 100%]
Other	5 (0.6%)	0 (0%)	[0.00; 54%]	5 (100%)	[46; 100%]
Unknown	308	125		183	
<b>How many people do you come in contact with on your way to school or work?</b>	836				
less than 5	94 (11%)	26 (28%)	[19; 38%]	68 (72%)	[62; 81%]
5 to less than 10	308 (37%)	117 (38%)	[33; 44%]	191 (62%)	[56; 67%]
10 to less than 50	387 (46%)	145 (37%)	[33; 43%]	242 (63%)	[57; 67%]

50 or more	47 (5.6%)	16 (34%)	[21; 49%]	31 (66%)	[51; 79%]
Unknown	285	116		169	
<b>Where do you spend most of your class or working time?</b>	924				
Indoor	376 (41%)	161 (43%)	[38; 48%]	215 (57%)	[52; 62%]
Outdoor	535 (58%)	176 (33%)	[29; 37%]	359 (67%)	[63; 71%]
Hospital	2 (0.2%)	0 (0%)	[0.00; 80%]	2 (100%)	[20; 100%]
Other	11 (1.2%)	3 (27%)	[7.3; 61%]	8 (73%)	[39; 93%]
Unknown	197	80		117	
<b>Do you work and live in the same house?</b>	912				
Yes	161 (18%)	65 (40%)	[33; 48%]	96 (60%)	[52; 67%]
No	751 (82%)	269 (36%)	[32; 39%]	482 (64%)	[61; 68%]
Unknown	209	86		123	

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

To evaluate the economic consequences caused by the pandemic, the included households were asked about employment status and financial setting for the year preceding the survey (Tab. 7.1). In the majority of households (35%) at least one household member was unable to work due to the pandemic, and 33% (219 households) stated that 2 to 5 household members could not work, while 210 participants (31%) did not suffer from such consequences.

**Table 7.1** Economic Consequences for Included Households

CHARACTERISTIC	N = 674 <sup>1</sup>
<b>How many household members could not work during the Corona pandemic?</b>	
>5	6 (0.9%)
2-5	219 (33%)
1	235 (35%)
0	210 (31%)
Unknown	4
<b>Percentage of household members who could not work during the pandemic</b>	
At least 75%	26 (3.9%)
Between 50% and 74%	119 (18%)
Between 25% and 49%	193 (29%)
Not more than 24%	121 (18%)
No impact	210 (31%)
Unknown	5
<b>Did the pandemic reduce the household income?</b>	
Yes	451 (68%)
No	208 (31%)
Do not know	4 (0.6%)
Unknown	11
<b>Did your household borrow money from other people or took a bank loan?</b>	
Yes	78 (12%)
No	595 (88%)
Unknown	1
<b>Did your household lend money to other households?</b>	
Yes	156 (23%)
No	512 (76%)
Do not know	3 (0.4%)
Unknown	3

<sup>1</sup>n (%)



In relation to the total household size, it meant that in 193 households (29%) between 25% to 49% of the household members were unable to work during the first year of the pandemic. Furthermore, in 26 households (3.9%) the pandemic caused that at least 75% were not unable to work. Income reduction was frequent among participants (68%), nonetheless only 12% (78 households) needed to borrow money or take a bank loan. In contrast, lending money to other households was more frequent (23%). Among all participants, 729 (66%) missed either work (n = 485) or school (n = 243) due to the pandemic or lockdowns (Tab. 7.2). Interestingly, women were less likely than men to miss school (f: 44% vs. m: 56%), but more likely to miss work (f: 53% vs. m: 47%). To evaluate the impact, the missed days were classified into five subgroups, ranging from less than 30 to more than 301 days missed. The majority of participants (39%) missed 91 to 180 days of work or school, followed by 31 to 90 days (31%) and 181 to 300 days (21%). Few participants missed less than 31 days (6.1%), or more than 300 days (2.5%). Only among the small subgroup (n = 18) with more than 300 days missed, participants differed significantly by sex (f: 22% vs. m: 78%).

**Table 7.2** Pandemics Impact on Work and School by Sex

CHARACTERISTIC	N	FEMALE, N = 584 <sup>1</sup>	95% CI <sup>2</sup>	MALE, N = 537 <sup>1</sup>	95% CI <sup>2</sup>
<b>Did you miss work or school due to the pandemic/ lockdowns?</b>	1,112				
Yes	729 (66%)	364 (50%)	[46; 54%]	365 (50%)	[46; 54%]
No	383 (34%)	215 (56%)	[51; 61%]	168 (44%)	[39; 49%]
Unknown	9	5		4	
<b>What did you miss?</b>	728				
School	243 (33%)	107 (44%)	[38; 51%]	136 (56%)	[49; 62%]
Work	485 (67%)	257 (53%)	[48; 57%]	228 (47%)	[43; 52%]
Unknown	1	0		1	
<b>How many days have you missed?</b>	727				
≥ 301	18 (2.5%)	4 (22%)	[7.4; 48%]	14 (78%)	[52; 93%]
181 ≤ 300	152 (21%)	71 (47%)	[39; 55%]	81 (53%)	[45; 61%]
91 ≤ 180	286 (39%)	145 (51%)	[45; 57%]	141 (49%)	[43; 55%]
31 ≤ 90	227 (31%)	119 (52%)	[46; 59%]	108 (48%)	[41; 54%]
≤ 30	44 (6.1%)	24 (55%)	[39; 69%]	20 (45%)	[31; 61%]
Unknown	2	1		1	

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

A great majority of all participants were either worried (47%) or very worried (45%) about the health of their relatives or friends due to the pandemic. Fewer participants were a little worried (1.7%) or not worried at all (5.3%, Tab. 8.1).

**Table 8.1** Participants Mental Health by Sex

CHARACTERISTIC	N	FEMALE, N = 584 <sup>1</sup>	95% CI <sup>2</sup>	MALE, N = 537 <sup>1</sup>	95% CI <sup>2</sup>
<b>Were you worried about the health of family/ friends during the time of the SARS-Cov-2 outbreak?</b>	1,120				
Very worried	509 (45%)	256 (50%)	[46; 55%]	253 (50%)	[45; 54%]
Worried	525 (47%)	283 (54%)	[50; 58%]	242 (46%)	[42; 50%]
A little worried	19 (1.7%)	9 (47%)	[25; 71%]	10 (53%)	[29; 75%]
Not worried at all	59 (5.3%)	32 (54%)	[41; 67%]	27 (46%)	[33; 59%]
Do not know	8 (0.7%)	3 (38%)	[10; 74%]	5 (63%)	[26; 90%]
Unknown	1	1		0	
<b>WHO-5 Well-Being Score<sup>3</sup></b>	1,112				
91 - 100	460 (41%)	216 (47%)	[42; 52%]	244 (53%)	[48; 58%]
81 - 90	147 (13%)	79 (54%)	[45; 62%]	68 (46%)	[38; 55%]
71 - 80	394 (35%)	218 (55%)	[50; 60%]	176 (45%)	[40; 50%]
61 - 70	9 (0.8%)	5 (56%)	[23; 85%]	4 (44%)	[15; 77%]
51 - 60	74 (6.7%)	40 (54%)	[42; 66%]	34 (46%)	[34; 58%]
≤ 50 <sup>4</sup>	28 (2.5%)	20 (71%)	[51; 86%]	8 (29%)	[14; 49%]
Unknown	9	6		3	

<sup>1</sup>n (%)<sup>2</sup>CI = Confidence Interval<sup>3</sup>WHO-5 score ranges from worst imaginable well-being (0 P) to best imaginable well-being (100 P).<sup>4</sup>Cut-score ≤ 50 points indicates screening for depression.Reference: Topp C, W, Østergaard S, D, Søndergaard S, Bech P: The WHO-5 Well-Being Index: A Systematic Review of the Literature. *Psychother Psychosom* 2015; 84:167-176. doi: 10.1159/000376585

Despite the concerns, in the surveyed population only 28 participants (2.5%) scored equal or lower than 50 points in WHO-5 Well-Being Index, which serves as an indicator for depression. Yet, women (71%) were significantly more frequent than men (29%) among this vulnerable subgroup. The majority of participants scored between 91 and 100 (41%) or 71 – 80 (35%) points. Here, men were represented more often than women as well, but no significant (f: 47% vs. m: 53%). All participants were asked to describe changes in their living conditions since the outbreak of COVID-19 to assess emotional and economic coping strategies (Tab. 8.2).

**Table 8.2** Changes in Living Conditions since the Pandemic by Sex

SINCE THE OUTBREAK OF THE COVID-19 PANDEMIC, ...	N	FEMALE, N = 584 <sup>1</sup>	95% CI <sup>2</sup>	MALE, N = 537 <sup>1</sup>	95% CI <sup>2</sup>
<b>have you changed your eating habits?</b>	1,119				
Yes, totally	305 (27%)	163 (53%)	[48; 59%]	142 (47%)	[41; 52%]
Yes, a little	571 (51%)	292 (51%)	[47; 55%]	279 (49%)	[45; 53%]
No, not much	4 (0.4%)	4 (100%)	[40; 100%]	0 (0%)	[0.00; 60%]
No, not at all	232 (21%)	119 (51%)	[45; 58%]	113 (49%)	[42; 55%]
Do not know	7 (0.6%)	5 (71%)	[30; 95%]	2 (29%)	[5.1; 70%]
Unknown	2	1		1	
<b>have you made any food deviations?</b>	1,121				
Yes, totally	302 (27%)	160 (53%)	[47; 59%]	142 (47%)	[41; 53%]
Yes, a little	576 (51%)	295 (51%)	[47; 55%]	281 (49%)	[45; 53%]
No, not much	6 (0.5%)	4 (67%)	[24; 94%]	2 (33%)	[6.0; 76%]

No, not at all	230 (21%)	120 (52%)	[46; 59%]	110 (48%)	[41; 54%]
Do not know	7 (0.6%)	5 (71%)	[30; 95%]	2 (29%)	[5.1; 70%]
<b>what type of change did you make?</b>	965				
Increased quantity of food during meals	59 (6.1%)	28 (47%)	[34; 61%]	31 (53%)	[39; 66%]
Decreased quantity of food during meals	810 (84%)	420 (52%)	[48; 55%]	390 (48%)	[45; 52%]
Reduced the number of meals per day	2 (0.2%)	2 (100%)	[20; 100%]	0 (0%)	[0.00; 80%]
Cut out certain foods	10 (1.0%)	9 (90%)	[54; 99%]	1 (10%)	[0.52; 46%]
Other to specify	11 (1.1%)	4 (36%)	[12; 68%]	7 (64%)	[32; 88%]
Do not know	73 (7.6%)	38 (52%)	[40; 64%]	35 (48%)	[36; 60%]
Unknown	156	83		73	
<b>how was the access to basic necessities?</b>	1,120				
Much better	1 (<0.1%)	1 (100%)	[5.5; 100%]	0 (0%)	[0.00; 95%]
Good access or same as usual	196 (18%)	110 (56%)	[49; 63%]	86 (44%)	[37; 51%]
Worse than usual	920 (82%)	470 (51%)	[48; 54%]	450 (49%)	[46; 52%]
Much worse than usual	3 (0.3%)	3 (100%)	[31; 100%]	0 (0%)	[0.00; 69%]
Unknown	1	0		1	
<b>have you drunk more alcoholic beverages than usual?</b>	1,119				
Yes, totally	33 (2.9%)	1 (3.0%)	[0.16; 18%]	32 (97%)	[82; 100%]
Yes, a little	299 (27%)	77 (26%)	[21; 31%]	222 (74%)	[69; 79%]
No, not much	16 (1.4%)	10 (63%)	[36; 84%]	6 (38%)	[16; 64%]
No, not at all	770 (69%)	494 (64%)	[61; 68%]	276 (36%)	[32; 39%]
Do not know	1 (<0.1%)	1 (100%)	[5.5; 100%]	0 (0%)	[0.00; 95%]
Unknown	2	1		1	

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

Regardless of sex, most participants confirmed that their eating habits had changed at least a little (51%) to even totally (27%). In contrast, for 232 participants (21%) eating habits did not change. Therefore, food deviations were also common among participants with either “Yes, a little” (51%), “Yes, totally” (27%) or “No, not at all” (21%). Although 810 participants had to decrease the quantity of food during meals (84%), only 2 participants (0.2%) left out complete meals per day. In contrast, 59 participants increased (6.1%) the quantity of food during meals. Nonetheless, most participants described the access to basic necessities including water, food and medication as “worse than usual” (82%) and for only 196 participants (18%) the access did not differ from pre-pandemic settings. Regarding food deviations, food adaptations and access to basic necessities no significant difference by sex was detected. Most of the participants (69%) did not drink more alcoholic beverages since the outbreak of COVID-19. Yet, in the subgroups of participants who consumed more, sex played a crucial role. Men reported an increase in consumption significantly more often than women (“Yes, totally” f: 1 (3%) vs. m: 32 (97%) and “Yes, a little” f:77 (26%) vs. 222 (74%), Tab. 8.2).

Main source of information to learn about SARS-CoV-2 transmission was assessed by level of education. Overall, classic platforms such as radio (69%) and TV (62%) were most frequently mentioned among all participants and across different education levels. (Tab. 9.1) In contrast, newspapers were only used by 8.9% (n = 60) of participants. Family, and friends (18%) as well as social media (14%) were named third and fourth most utilized sources of information. Social media was mentioned more often with increasing education level (from 1.6% of participants without a degree to 54% of participants with a higher education). Regardless of the level of education, health workers (4.3%), information sites provided by the Ministry of Health (0.9%) or other webpages on SARS-CoV-2 (0.1%) were rarely used as sources of information. Consistent with these findings, participants also had their highest trust in the classic platform TV (56%) and radio (30%) across all education levels (Tab. 9.2, Appendix A).

**Table 9.1** Source of Information by Level of Education

What is your source of information about the coronavirus?	N	No Degree, N = 62 <sup>†</sup>	Primary, N = 179 <sup>†</sup>	Secondary School, N = 355 <sup>†</sup>	Higher, N = 58 <sup>†</sup>	Other, N = 17 <sup>†</sup>
<b>Social media (e.g., Facebook)</b>	671					
Mentioned	96 (14%)	1 (1.6%)	8 (4.5%)	55 (15%)	31 (53%)	1 (5.9%)
Not mentioned	575 (86%)	61 (98%)	171 (96%)	300 (85%)	27 (47%)	16 (94%)
<b>Webpages (e.g., Google)</b>	671					
Mentioned	1 (0.1%)	0 (0%)	0 (0%)	0 (0%)	1 (1.7%)	0 (0%)
Not mentioned	670 (100%)	62 (100%)	179 (100%)	355 (100%)	57 (98%)	17 (100%)
<b>Ministry of Health</b>	671					
Mentioned	6 (0.9%)	1 (1.6%)	0 (0%)	1 (0.3%)	4 (6.9%)	0 (0%)
Not mentioned	665 (99%)	61 (98%)	179 (100%)	354 (100%)	54 (93%)	17 (100%)
<b>Radio</b>	671					
Mentioned	463 (69%)	53 (85%)	122 (68%)	241 (68%)	34 (59%)	13 (76%)
Not mentioned	208 (31%)	9 (15%)	57 (32%)	114 (32%)	24 (41%)	4 (24%)
<b>TV</b>	671					
Mentioned	417 (62%)	32 (52%)	92 (51%)	247 (70%)	43 (74%)	3 (18%)
Not mentioned	254 (38%)	30 (48%)	87 (49%)	108 (30%)	15 (26%)	14 (82%)
<b>Newspaper</b>	671					
Mentioned	59 (8.8%)	1 (1.6%)	6 (3.4%)	45 (13%)	6 (10%)	1 (5.9%)
Not mentioned	612 (91%)	61 (98%)	173 (97%)	310 (87%)	52 (90%)	16 (94%)
<b>Health workers</b>	671					
Mentioned	29 (4.3%)	4 (6.5%)	5 (2.8%)	10 (2.8%)	8 (14%)	2 (12%)
Not mentioned	642 (96%)	58 (94%)	174 (97%)	345 (97%)	50 (86%)	15 (88%)
<b>Family and friends</b>	671					
Mentioned	117 (17%)	11 (18%)	36 (20%)	52 (15%)	10 (17%)	8 (47%)
Not mentioned	554 (83%)	51 (82%)	143 (80%)	303 (85%)	48 (83%)	9 (53%)
<b>Other</b>	671					
Mentioned	2 (0.3%)	0 (0%)	1 (0.6%)	0 (0%)	1 (1.7%)	0 (0%)
Not mentioned	669 (100%)	62 (100%)	178 (99%)	355 (100%)	57 (98%)	17 (100%)

<sup>†</sup>n (%)

Across all levels of education, almost every household mentioned that they took measures to protect household members (97%) in the form of wearing masks (94%), washing hands more frequently (89%), or avoiding contact to other households (18%, Tab. 10.1). The latter was most frequently mentioned by participants with no degree (34%). Other ways to reduce contacts, such as avoiding crowded places (0.5%) or eating together (1.8%) were only mentioned by a minority of participants regardless of their educational level. Regarding ELISA seropositivity, no significant difference was observed among non-pharmaceutical measures (Tab. 10.2, Appendix A).

**Table 10.1** Non-pharmaceutical Measures by Level of Education

Characteristic	N	No degree, n = 62 <sup>l</sup>	Primary, n = 179 <sup>l</sup>	Secondary school, n = 355 <sup>l</sup>	Higher, n = 58 <sup>l</sup>	Other, n = 17 <sup>l</sup>
<b>Did you take measures to protect household members?</b>	669					
Yes	652 (97%)	61 (100%)	175 (98%)	343 (97%)	56 (97%)	17 (100%)
No	17 (2.5%)	0 (0%)	4 (2.2%)	11 (3.1%)	2 (3.4%)	0 (0%)
Unknown	2	1	0	1	0	0
<b>Masks for every household member</b>	671					
Mentioned	631 (94%)	60 (97%)	170 (95%)	331 (93%)	53 (91%)	17 (100%)
Not mentioned	40 (6.0%)	2 (3.2%)	9 (5.0%)	24 (6.8%)	5 (8.6%)	0 (0%)
<b>Masks for household member with symptoms</b>	671					
Mentioned	6 (0.9%)	0 (0%)	2 (1.1%)	4 (1.1%)	0 (0%)	0 (0%)
Not mentioned	665 (99%)	62 (100%)	177 (99%)	351 (99%)	58 (100%)	17 (100%)
<b>Masks for vulnerable household members</b>	671					
Not mentioned	671 (100%)	62 (100%)	179 (100%)	355 (100%)	58 (100%)	17 (100%)
<b>Not taking meal together</b>	671					
Mentioned	12 (1.8%)	0 (0%)	2 (1.1%)	10 (2.8%)	0 (0%)	0 (0%)
Not mentioned	659 (98%)	62 (100%)	177 (99%)	345 (97%)	58 (100%)	17 (100%)
<b>Not sharing personal objects</b>	671					
Mentioned	1 (0.1%)	0 (0%)	0 (0%)	1 (0.3%)	0 (0%)	0 (0%)
Not mentioned	670 (100%)	62 (100%)	179 (100%)	354 (100%)	58 (100%)	17 (100%)
<b>Washing hands more regularly</b>	671					
Mentioned	594 (89%)	55 (89%)	160 (89%)	314 (88%)	53 (91%)	12 (71%)
Not mentioned	77 (11%)	7 (11%)	19 (11%)	41 (12%)	5 (8.6%)	5 (29%)
<b>Avoiding contacts with other households members</b>	671					
Mentioned	120 (18%)	21 (34%)	31 (17%)	54 (15%)	13 (22%)	1 (5.9%)
Not mentioned	551 (82%)	41 (66%)	148 (83%)	301 (85%)	45 (78%)	16 (94%)
<b>Avoiding market and other crowded place</b>	671					

Mentioned	5 (0.7%)	2 (3.2%)	1 (0.6%)	2 (0.6%)	0 (0%)	0 (0%)
Not mentioned	666 (99%)	60 (97%)	178 (99%)	353 (99%)	58 (100%)	17 (100%)
<b>Cleaning households objects more regularly</b>	671					
Mentioned	15 (2.2%)	3 (4.8%)	1 (0.6%)	9 (2.5%)	2 (3.4%)	0 (0%)
Not mentioned	656 (98%)	59 (95%)	178 (99%)	346 (97%)	56 (97%)	17 (100%)

<sup>1</sup>n (%)

Participants also confirmed that maintaining social distance (55%) was the most difficult for them across all 10-years age groups (Tab. 10.3, Appendix A). Although, it was more frequent among participants between 10 and 19 years old (66%) than participants older or equal to 60 years (49%). Wearing face masks was the second most frequently mentioned response (29%). Interestingly, only a few participants (4.2%) answered that the inability to work was their greatest difficulty (4.2%). Many participants replied that other nearby persons respected advised measures (76%), with the lowest percentage among 20–29 years old participants and the highest percentage among the oldest age groups (81%). Nearly all participants (96%) reported that other people should comply to these measures, but a small number of participants among the youngest age groups were unsure ( $10 \leq 19$ : 6.6% and  $20 \leq 29$ : 11%).

To evaluate the knowledge of SARS-CoV-2 risk factors among study participants, they were asked about those are supposedly most vulnerable to suffer from or die of COVID-19. On a household level, participants mainly named elderly people (79%) and people with chronic diseases, such as diabetic (62%), cardiac (47%) or pulmonary (45%) diseases and immunosuppressed conditions (46%) as risk factors for infection. These were most frequently mentioned by participants without a degree compared to the higher educated. Other factors mentioned included sex (m: 11% and w: 10%) age (children being vulnerable, 12%) or pregnant women (28%) across all levels of education (Tab. 11, Appendix A).

The socio-economic score (SES) was calculated using household assets and grouped into five quintiles – lowest, second, middle, fourth and highest according to their wealth. Before performing a varimax-rotated Principal Component Analysis (PCA), a correlation matrix of the selected items was calculated (Fig. 15). Only items with more than 1% of positive replies were included. Thereby, items landline phones, electric or gas stoves as well as motorbikes were excluded from the analysis. According to the presence of assets from the wealth index, households were divided into “wealthier” and “poorer” (Tab. 12).

**Table 12. Wealth Index Assets**

ASSET	WEALTHIER	POORER
<b>Housing quality</b>		
Type of housing	Unique courtyard/ Villa/ Other	Apartment/ Common courtyard (concession)
Nature of walls	Hard/ Semi-hard/ Improved brick	Brick/ Straw
Nature of floor	Tiles/ Cement/ Finished floor	Stamped earth/ Sand/ Natural soil/ Rough floor
Crowding	5 or less people per sleeping room	6 or more people per sleeping room
<b>Access to services</b>		
Type of toilets	Flush/ Ventilated latrine	Ordinary latrine/ none
Main source of lighting energy	Electricity/ Gas/ Solar lamp	Storm lamp/ Accumulator with batteries/ Candles/ Other
Source of water in rainy season	Tab water/ Protected borehole	Ordinary well/ Rainwater/ River, lake, dam
Source of water in dry season	Tab water/ Protected borehole	Ordinary well/ Rainwater/ River, lake, dam
Solar lamp	Yes	No
Oil lamp	Yes	No
<b>Consumer durables</b>		
Radio	Yes	No
TV	Yes	No
Video/CD/DVD Player	Yes	No
Cell phone	Yes	No
Smart phone	Yes	No
Fridge/ Freezer	Yes	No
Bicycle	Yes	No
Laptop	Yes	No
Bed nets	Yes	No

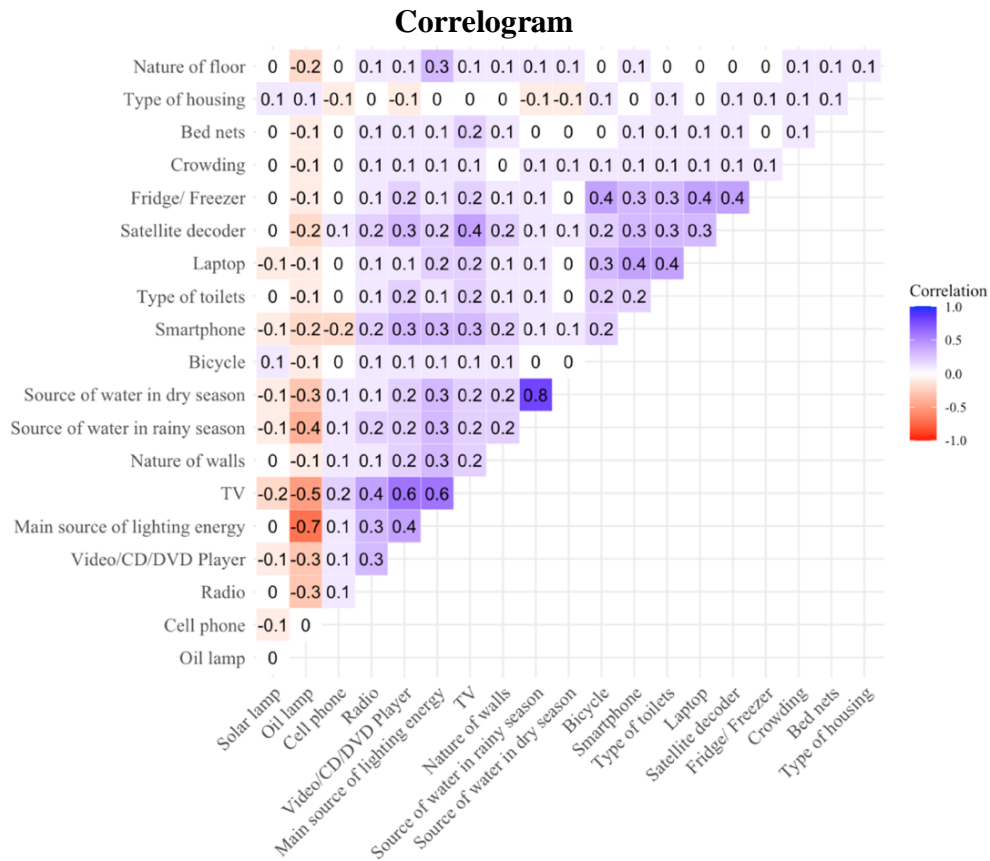
For varimax-rotated PCA, assets were assigned into wealthier = 1 and poorer = 0.

Assets with frequencies between 5% and 95% among all households were included in the analysis.

Reference: Hjelm, Lisa, et al. Creation of a wealth index. World Food Program (2017).

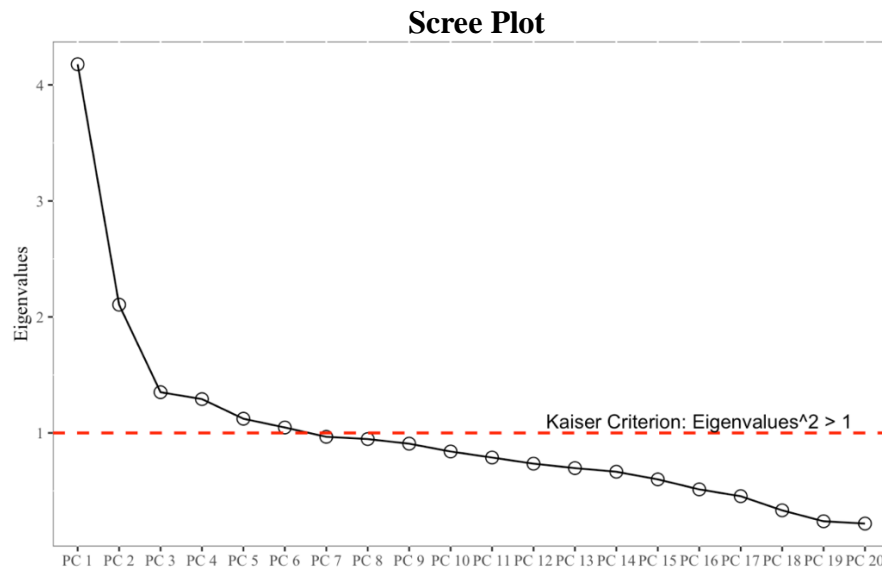
<https://www.wfp.org/publications/creation-wealth-index-june-2017>

The correlogram (Fig. 16) shows that sources of water in dry and rainy season were strongly positively correlated, all well as electronic devices such as TV, DVD player to the main source of lighting energy. The latter in turn was positively correlated to an improved type of floor and roof. Fridges/ freezers and type of toilets were positively correlated with mobility items, such as the possession of a car or a bicycle.



**Figure 16**  
The correlogram shows a correlation matrix of included wealth index assets.

To identify the necessary number of Rotated Components (RC) for varimax-rotated Principal Component Analysis, a scree plot was drawn (Fig. 17).



**Figure 17**  
The scree plot shows the number of principal components to retain for varimax-rotated PCA. PC 1 to PC 6 fulfill the Kaiser criterion.



The first 6 RCs fulfilled the Kaiser-criterion with Eigenvalues  $\lambda^2 > 1$ , and were retained for the analysis (Fig. 17 and Tab. 13) (Field, Miles, and Field 2012).

**Table 13.** Eigenvalues of PC1 to PC6

PRINCIPAL COMPONENT	EIGENVALUE <sup>1</sup>
PC 1	4.178
PC 2	2.106
PC 3	1.351
PC 4	1.292
PC 5	1.123
PC 6	1.047

<sup>1</sup>The eigenvalues indicate their contribution to the variance of the 19 variables.

Eigenvalues were calculated with R® package psych.

The varimax rotation is an orthogonal rotation method to maximize the variance of components, and the standard for the wealth index calculation (Costello and Osborne 2014; Hjelm et al. 2017). RC1 and RC2 together explain 28 % of the wealth variance (Tab. 14 in red).

**Table 14.** Results of the varimax-rotated PCA

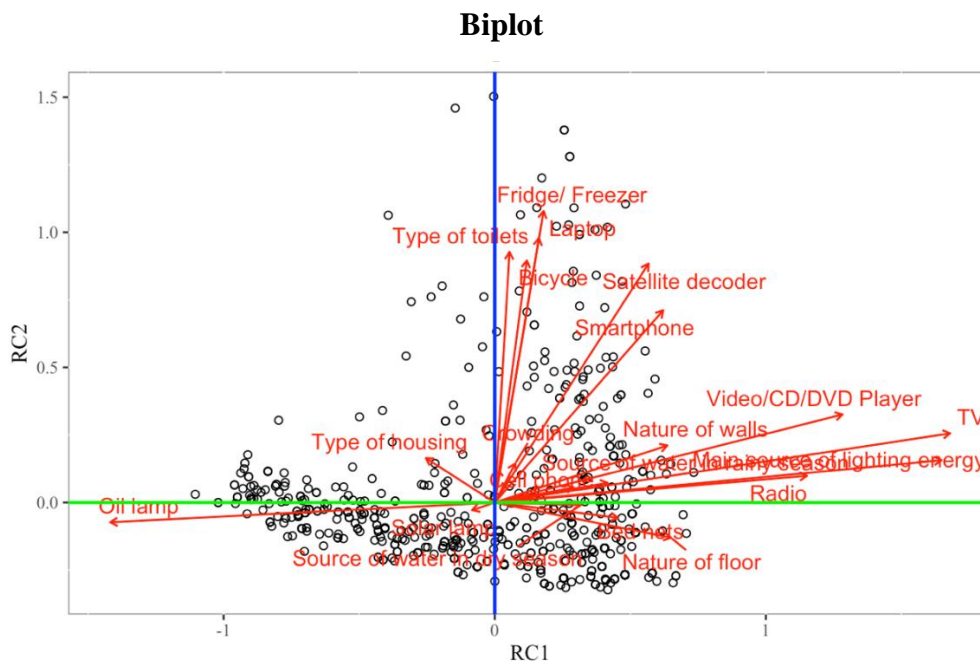
	RC1	RC2	RC3	RC6	RC4	RC5
SS loadings <sup>1</sup>	3.06	2.57	1.90	1.21	1.19	1.18
Proportion Variance	0.15	0.13	0.09	0.06	0.06	0.06
Cumulative Variance	0.15	<b>0.28</b>	0.38	0.44	0.50	0.55
Proportion Explained	0.28	0.23	0.17	0.11	0.11	0.11
Cumulative Proportion	0.28	0.51	0.68	0.79	0.89	1.00

<sup>1</sup>Sum of squared loadings

Varimax-rotated Principal Component Analysis was performed, using R® package psych.

Compound scores of both RCs (x-axis = RC 1, y-axis = RC 2) are presented in the score plot (Fig. 18.1, Appendix B), showing the included households (n= 674) as numbered black circles. Observations respective households clustering together are more likely to have a similar level of wealth. Thus, wealthier households are on the upper right side of the graph, while poorer households are on the lower left side. Households located near the origin are closer to the average.

The loading plot (Fig. 18.2, Appendix B) presents the contribution of each wealth index item to the variance of RC1 and RC2. Items that cluster together are likely to be correlated to each other, indicating that they co-occur in households. The extent of contribution of a variable/ item to either RC 1 or RC 2 is characterized by its distance to the origin. Similar to the correlogram (Fig. 16), the loading plot points out that the main source of lightning energy was associated with possession of electronic devices such as TV or radio, as well as with a higher quality roof. Mobility items such as cars and bicycles were associated with further electronic devices in the form of laptops and fridges. The type of toilets was also associated with the presence of a fridge or freezer and mobility assets such as a car or a bicycle. Both plots together are presented in the biplot (Fig. 18.3).



**Figure 18.3**  
The Biplot illustrated the component scores of sampled households (black circles, n = 674). The vectors (red arrows) represent the included wealth index assets.

In a biplot, the loading scores of each wealth index item are presented as vector (red arrows). Vectors pointing into the same direction are likely to be correlated with each other, whereas perpendicular vectors indicate no association. An opposite vector indicates a negative correlation. The length of a vector indicates on how good the variance of a variable is explained by the rotated components. This means that long vectors are well represented by the RCs (e.g., TV, radio and fridge), while shorter ones are not (e.g., computer or crowding) (Hartmann, Krois, and Waske 2018).

Included households were predominantly grouped within the fourth (32%) and highest wealth quintiles (25%) (Tab. 15). The third-largest group was assigned as second quintile

(16%), and 53 households (7.9%) belonged to the lowest quintile. Seropositivity across wealth ranged from 25% (poorest) to 42% (richest) with no significant differences between the quintiles.

**Table 15.** Wealth Index Quintiles by ELISA Test Result

Characteristic	N	ELISA TEST RESULT			
		positive, N = 240 <sup>1</sup>	95% CI <sup>2</sup>	negative, N = 434 <sup>1</sup>	95% CI <sup>2</sup>
<b>Wealth quintiles</b>	674				
highest	171 (25%)	71 (42%)	[34; 49%]	100 (58%)	[51; 66%]
fourth	218 (32%)	78 (36%)	[29; 43%]	140 (64%)	[57; 71%]
middle	126 (19%)	45 (36%)	[28; 45%]	81 (64%)	[55; 72%]
second	106 (16%)	33 (31%)	[23; 41%]	73 (69%)	[59; 77%]
lowest	53 (7.9%)	13 (25%)	[14; 39%]	40 (75%)	[61; 86%]

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

The possession of wealth index assets per wealth quintile is shown in Table 16 (Appendix A). Among the retained wealth index assets, most households had better walls (87%) and a cell phone (86%). Additionally, households of the highest quintile frequently owned a TV (100%), radio (99%) and Video/CD/DVD Player (91%), whereas the lowest did not. Further, the lowest (87%) and the second quintile (44%) used oil lamps for lighting, whereas almost all households in the fourth (99%) and all households in the highest quintile (100%) used an improved main source of lighting, such as electricity or gas. Overall, only few households possessed “wealthier” toilets (7.9%), fridge/ freezers (6.4%), bicycles (5.5%) and solar lamps (5.2%). Most of the households owned bed nets (70%), which gradually increased from the lowest (49%) to the highest (82%) quintile.

## 4. Discussion

The objective of this thesis was to determine the SARS-CoV-2 seroprevalence overall and in subgroups, to identify possible associations between participant characteristics (e.g. occupation, risk perception) and seropositivity, and to evaluate the representativeness of the sampling strategy as a means of enrolling a typical urban Malagasy population.

### 4.1 Major findings

The major findings of this doctoral thesis were that the SARS-CoV-2 seroprevalence among participants varied depending on their demographic, occupational and socioeconomic background in urban Malagasy settings, with a particularly high seropositivity among health care workers and a low seropositivity among participants working outside. While the surveyed households were mostly classified among the wealthier quintiles, the pandemic had a large impact on participants' food and economic security. Although, the WHO-5 Well-Being Score detected only a few potential cases of depression, sex seems to play an important role in coping with the pandemic. Across all levels of education, participants mainly used and trusted the radio and the TV to receive information about the COVID-19 pandemic. Regarding non-pharmaceutical interventions to mitigate the spread of COVID-19, the use of face masks and handwashing were frequently mentioned. Social distancing among participants was rare.

### 4.2 SARS-CoV-2 seroprevalence among subgroups

The SeroCoV study unraveled a crude seropositivity for IgG SARS-CoV-2 of 37.5% [35; 40] for the study population (n = 1121) in the site Fianarantsoa, which was similar to the previously published seroprevalence on a household level (Tab. 2) (Struck et al. 2022). Prior to the SeroCoV study, a multi-location cross-sectional study among Malagasy blood donors with a similar age distribution detected a lower crude seroprevalence of 18.5% in Fianarantsoa after the first COVID-19 wave in September 2020 (Schoenhals et al. 2021). A higher seroprevalence among the SeroCoV study participants was likely due to the sampling time period during the second COVID-19 wave in spring 2021 (Fig. 11 and 12), which led to a re-declaration of national health emergency by the Malagasy government on April 2<sup>nd</sup>, 2021 (Rakotonanahary et al. 2021). Additionally, the higher seroprevalence could also be explained by the applied two-stage sampling strategy, allowing to enroll a more representative study population than the

Malagasy blood donors who are mainly fit, young and male (Schoenhals et al. 2021; Rakotoniana et al. 2019). Furthermore, sickness in the past weeks before donation was an exclusion criterion. It has been stated previously that the unequal distribution of men vs. women blood donors is a general weakness of blood donor studies in SSA, which limits their validity as a representation for the entire population (Uyoga et al. 2021; Adetifa et al. 2021). On the other hand, household-based surveys in SSA tend to enroll more women than men due to the absence of men during the day, which was also the case for the SeroCoV survey. However, the difference between recruited men and women (w: 52% vs. m: 48%, Tab. 5.1) was smaller compared to other similar surveys in SSA (Wiens et al. 2021; Negera et al. 2013). Yet, women also outnumbered men in all ‘*Grand Centre Urbain*’ (CGU), including Fianarantsoa, regardless of age groups according to the official Malagasy census (Institut National de la Statistique 2021).

In May 2021, a follow-up SARS-CoV-2 IgG seroprevalence testing of blood samples among blood donors in Antananarivo unraveled a higher seroprevalence of 47.6%, which was likely caused by the beta variant of the Coronavirus (Razafimahatratra et al. 2021; Tegally et al. 2021). Likewise, a study of Malawian blood donors detected a seropositivity of 64.9 % in May 2021, which was also suggested to be due to the spread of the beta Coronavirus variant (Mandolo et al. 2021).

Zimbabwe, which is in geographical proximity to Madagascar and Malawi, had experienced a second COVID-19 wave driven by the beta Coronavirus variant in December 2020, and a third one driven by the delta variant during the time of sampling for the SeroCoV (Fryatt et al. 2021; Mlcochova et al. 2021). While the importance of the beta variant for the second COVID-19 wave is evident, the role of the delta Coronavirus variant remains uncertain during the sampling period, since the spread of the delta variant was tracked only scarcely and first detected in Madagascar in October 2021 (Gangavarapu et al. 2023). Due to very limited Coronavirus sequencing capacities, conclusions about these findings should be interpreted with great caution.

As of April 2023, based on a PubMed® with a defined search term mentioned above and the COVID-19 serotracker® website, no other household based seroprevalence sampling study had been published in Madagascar (Arora et al. 2021). Many seroprevalence studies in SSA assessed the occupational exposure to the Coronavirus and were conducted among health care workers (HCW), such as a cross-sectional survey in five Kenyan counties (Etyang et al. 2022). A systematic review and meta-analysis (SRMA) across multiple African countries also assessed a higher seroprevalence among HCWs (Hajissa et al.

2022). Similar to these findings, SeroCoV participants working in health care settings, had a higher rate of seropositivity 75% [36; 96%], pointing towards an increased risk for infection in that occupational sector (Tab. 5.1). However, the detected seroprevalence was in sharp contrast to a previously published lower seroprevalence among HCWs in Antananarivo during the first wave (Ratovoson et al. 2022). Further, inferences from the SRMA for Madagascar were limited, as no Malagasy study and almost exclusively studies conducted in 2020 were included. Regarding the small number of included HCWs (n = 8), daytime sampling could be an explanation, since HCWs had to work during the pandemic and were consequently absent from their households. Yet compared to official data for Madagascar, the SeroCoV study included a much higher proportion of HCWs (SeroCoV: 8 HCWs/ 846 participants (0.9 %) vs. WHO data for Madagascar: 3 Nurses or midwives/ 10,000 people (0.03%) and 2 doctors/ 10,000 people (0.02%)) (World Health Organization 2022a). This difference was probably due to the urban setting with a university hospital. Given that other studies mainly recruited special groups (e.g., blood donors, HCWs), the general exposure to SARS-CoV-2 in most SSA populations remains uncertain.

A SARS-CoV-2 IgG seroprevalence survey with a similar sampling strategy was conducted in Kinshasa, Democratic Republic Kongo, during their second COVID-19 wave (Munyeku-Bazitama et al. 2023). The study showed a comparable median age of 30 [18-46] years old, but they observed a lower seroprevalence among the poorest and the wealthiest quintile of participants, which differed from the findings of the SeroCoV study, where seropositivity gradually increased with households wealth (Tab. 15). The surveyed population in Kinshasa had the highest seroprevalence among participants with a university education. This was also the case for SeroCoV results, where participants with a bachelor's degree (*license*) had the highest seropositivity (Tab. 5.1). Further, both cross-sectional studies observed lower seroprevalences among participants with no or only primary education (Tab. 5.1). While the underlying reasons remain uncertain, it may suggest that outdoor occupations, such as farming, could possibly pose a lower risk for infection, and were predominantly performed by participants with no or primary education (Tab. 5.4). Considering that the seropositivity was also significantly lower among participants spending more time outside than inside while attending school or work (Tab. 6.5), SeroCoV results confirm the observation that a SARS-CoV-2 infection is less likely to occur outdoors, which is consistent with the results of systematic review (Bulfone et al. 2021).

Two different household-based cross-sectional studies investigating the SARS-CoV-2 seroprevalence were conducted in Cameroon's capital Yaoundé (Nwosu et al. 2021; Ndongo et al. 2022). Both had a comparable distribution of age and levels of education to the SeroCov participants. The first study that recruited in October and November 2020 examined the SARS-CoV-2 seroprevalence and observed a higher seroprevalence among men compared to women, which was different to SeroCoV findings (Tab. 5.1). The study also assessed the seroprevalence in different age groups by the participant's sex. In contrast to the SeroCoV study, the highest seropositivity was detected in the oldest age groups of both sexes (Tab. 5.2). Regarding the seropositivity, this study was only comparable to a limited extent, because it included younger participants (i.e.; equal or older than 5 years), used different age group sizes and performed an additional IgM ELISA. Other findings were consistent with the SeroCoV results, such as a gradual increase in seropositivity from smaller to larger households (Tab. 6.1), and a similar often performed traditional self-medication (23%) in cases of COVID-19 symptoms (Tab. 6.2). While the first included more women than men, the second survey had an equal sex distribution and detected a sharp increase of SARS-CoV-2 IgG seropositivity from 18.6% [15.7; 21.7%] in January/ and February 2021 to 51.3% [48.3; 54.2%] in April and May 2021. Even if these sub-Saharan African urban populations shared similar demographic characteristics, Fianarantsoa is only comparable to a limited extent to them. First, Fianarantsoa is much smaller than Kinshasa (15 million inhabitants) or Yaoundé (4 million inhabitants). Second, its population density is much smaller (Fianarantsoa: ~2,300 persons/km<sup>2</sup> vs. Kinshasa: ~ 15,000 p./km<sup>2</sup> and Yaoundé: ~10,000 p./km<sup>2</sup>) (Institut National de la Statistique 2021; Ritchie and Roser 2018b; Communauté Urbaine de Yaoundé 2023). The lower population density may have mitigated the spread of SARS-CoV-2 among inhabitants, which probably explains the lower SARS-CoV-2 seroprevalence. In addition, a systematic review and meta-analysis, evaluating the SARS-CoV-2 seroprevalence among UN African subregions per quartal, described a higher SARS-CoV-2 seroprevalence in Eastern Africa in 2<sup>nd</sup> quartal 2021, as well (Lewis et al. 2022). However, their results evaluated only 5 studies, of whom none were conducted in Madagascar.

#### 4.3 Food and economic insecurities

Early predictions that contingency measures to mitigate the COVID-19 pandemic in LMICs may lead to a variety of indirect consequences, such as missing school in the

case of children, or job loss in the case of adults, were also observed in Fianarantsoa (Tab. 7.2) (Zar et al. 2020; Paci 2021). While most of the households were classified among the fourth and highest quintile, using the wealth index (Tab. 15), many participants reported that they experienced an income reduction (Tab. 7.1), or food insecurities due to the pandemic (Tab. 8.2). Still, most households did not need to take a bank loan, and some households were even able to lend money to others, which probably indicated the wealthier settings of the included households (Tab. 7.1). Furthermore, unlike several SSA countries described in a multi-country survey that examined food insecurity in 2020, only two SeroCoV households had to skip meals (Tab. 8.2) (Dasgupta and Robinson 2021). A household-based survey among SSA countries observed a strong initial increase, followed by a gradual decline of food insecurity from April 2020 to July 2021 (Rudin-Rush et al. 2022). Given that the risk of malnutrition was already high for Malagasy children before the pandemic, more than one year after the outbreak, 86% of households reported that they had to face food insecurities by decreasing food quantity during meals (Tab. 8.2). These findings assume that the alimentary situation for children may have even worsened during the pandemic (Roser and Ritchie 2019).

Given that contingency measures in particular led to these grave ‘side effects’, future studies need to monitor these consequences. Further, they should evaluate which measures were not dispensable and which ones could be adapted to reduce the impact on the Malagasy society in future pandemics, focusing also on rural areas and regions in the further south of Madagascar that are currently particularly vulnerable to famine (World Food Programme 2023; ACAPS 2022).

#### 4.4 Socioeconomic score

The wealth index was calculated using the first varimax-rotated component (RC1) and can only explain 15% of the total wealth variance (Tab. 14). While the surveyed population was less wealthy than the census urban residences, they were wealthier than the findings for the ‘Haute-Matsiatra’ region, which may explain the wealth in the middle range of census results (Institut National de la Statistique and ICF 2022). Interestingly, according to the DHS Wealth Index, owning an oil lamp predicted wealthier living conditions (Tab. 12), but it turned out to be an indicator of less wealth among the SeroCoV households, as it was often negatively correlated with owning other ‘wealthier’ households assets (Fig. 16). In contrast, a ‘wealthier’ main source of lighting was associated with the presence of ‘wealthier’ housing quality, access to services and



consumer durables. This finding presumably indicated the general presence of electricity in the households, enabling the use of electronic devices. Given that the main source of water was highly correlated in dry or rainy season, a distinction seemed unnecessary in urban Malagasy settings.

According to official data from the International Labor Organization (ILO), the overall employment among SSA countries was 62% in 2021 (International Labour Organization 2023). While the employment rate of 81% among SeroCoV participants (Tab. 5.1) was comparable to official data for Madagascar, the majority of households reported that at least one of or even more household members could not work due to the pandemic (Tab. 7.1). The context of these ambiguous results remains unclear. One potential explanation could be that household members initially lost their employment, but were able to resume employment soon after. For instance, the tourism sector was particularly affected by the pandemic due to internationally closed borders. A survey examining the economic consequences also observed an initial increase of job loss in four SSA countries in 2020, followed by a rapid re-employment (Contreras-Gonzalez et al. 2022). Future studies will be needed to assess whether the Madagascar's employment situation remains as 'robust'.

#### 4.5 Mental health and well-being

Since the outbreak of the pandemic, a great majority of SeroCoV participants were either worried or very worried about the health of family and friends (Tab. 8.1). To some extent, this high level of worries may be caused by the aggravating second wave during the sample period (Borgia 2021; African Defense Forum 2021). Consistent with SeroCoV findings, a web-based survey across multiple anglophone SSA countries also observed a high-level of concerns during the first lockdowns in April and Mai 2020 (Osugwu et al. 2021). Future surveys of the COVID-19 pandemic will need to evaluate, whether other concerns, such as economic insecurities, may have influenced the perception.

Regarding the onset of clinical depression, only 2.5 % of participants scored lower than 51 points in the Well-Being Index, which suggests that depressions were rare among the SeroCoV participants. Yet, this finding should be interpreted with caution. A systematic review of mental health examined a pooled depression prevalence of 30% [22; 40%] among African countries in February 2021 (Chen et al. 2021). A systematic review of mental health in non-SSA countries also indicated a higher prevalence of depression of 20 % [18; 21%] (Dragioti et al. 2022). However, no Malagasy population was included in the review and other test/ scores were applied than the WHO-5 Well-Being Score.

Regarding mental health and depression in Madagascar, few publications exist. The latest WHO mental health atlas of Madagascar underlined this scarcity of treatment facilities and staff (e.g., Madagascar: 0.06 psychiatrist/ 100,000 persons vs. Germany: 27 psychiatrist/ 100,000 p.), as well as the lack of promotion and prevention campaigns (World Health Organization 2022b; Eurostat 2020). A retrospective analysis of psychiatric cases in the capital Antananarivo investigated possible reasons for this shortcoming, and highlighted the strong influence of religious beliefs on diagnostic and treatment of mental diseases among the Malagasy population (Bakohariliva et al. 2018). The observation that more women were at risk for a depression (Tab. 8.1) compared to men was consistent with findings of a multinational cross-sectional COVID-19 mental health survey in LMICs, and a small (n = 64) mental health survey among Malagasy adults, assessing depression symptoms with the Beck Depression Inventory II (Cénat et al. 2021; Foubert et al. 2021). The surprisingly good well-being, despite lockdown measures in the first COVID-19 waves, were also observed in a meta-analysis (Prati and Mancini 2021). Yet, this meta-analysis did not include any African studies. However, since stigmatization of mental illness is still common among Malagasy health care workers, participants may have erroneously overreported in the WHO-5 Well-Being Score (R. F. Randrianarivo et al. 2023).

While Malagasy alcohol consumption was at a low level, with only 1.97 Liter of pure alcohol per person compared to Germany with 12.97 Liter of pure alcohol per person in 2018, SeroCoV findings indicated that the pandemic appeared to have negatively affected this figure (Ritchie and Roser 2018a). Even though no baseline assessment of alcohol consumption was performed, nearly 30 % of participants reported that they increased their consumption, since the start of the pandemic (Tab. 8.2). As expected, SeroCoV findings indicate that male participants have a higher alcohol consumption, which was also observed in a pre-pandemic study, investigating cardiovascular diseases in four urban settings of SSA countries, including Madagascar (Kingue et al. 2017). In contrast to SeroCov results, a survey in Nigeria, Burkina Faso and Ethiopia found out that the alcohol drinking consumption had only increased in very few cases since the start of the pandemic (Assefa et al. 2021). Yet this study is only comparable to a limited extent, enrolling only health care workers.

Since few data has been published about the general mental health burden and psychoactive substance abuse before and during the pandemic in Madagascar, future studies will need to target these issues by establishing a baseline. In addition, awareness

campaigns for health care workers, among others, should address the stigmatization of mental health disorders.

#### 4.6 Sources of information

SeroCoV survey provided in-depth information about the participants' knowledge of COVID-19 (Tab. 9.1) and their most trusted sources of information (Tab. 9.2). The important role of radio as a primary source of information about COVID-19 was also observed in a scoping review among SSA countries (Workneh et al. 2021). Nevertheless, relying mostly on online questionnaires limited the generalization of this review, possibly excluding illiterate participants or people without access to the internet. According to official sources, the Malagasy rate of literacy has continuously increased from 59 % in 2005 to 77% in 2021 (UNESCO Institute for Statistics 2023). Thus, classic platforms, such as radio and TV will remain crucial, since they can provide information to illiterate people, as well. Interestingly, unlike other SSA countries with even lower literacy rates, social media and webpages were scarcely mentioned by SeroCoV households (Reuben et al. 2021; Brown Wilson et al. 2021). Especially participants with no degree did not use social media or webpages, but listened to the radio instead, which emphasizes the cultural importance of orally distributed information (Tab. 9.1).

Regarding the level of trust in sources of information, similar to the results of a nationwide representative survey conducted in Ghana in 2020, SeroCoV households mentioned TV and radio as most trustworthy (Tab. 9.2) (Kwabla et al. 2022). However, unlike in the Ghanaian survey, family and friends were only rarely mentioned by SeroCoV households. Other official sources like the Malagasy Ministry of Health were scarcely reported as well, which were also different to the findings reported for Liberia and Kenya (Brown Wilson et al. 2021; Austrian et al. 2020).

Given that these study results, even though slightly varied, showed a commonality in the importance of classic media platforms, policymakers should be encouraged to use them in future crisis situations. Yet, to ensure a fast and reliable communication in up-coming (health) crisis, studies and awareness campaigns will need to evaluate, to which extent main sources of information will need to be adapted. While the many SSA online surveys provided a good glimpse of the sources of information and were easier to perform throughout contingency measures, they were prone to have a selection bias towards wealthier and literate people (Kwabla et al. 2022; Ditekemena et al. 2021; Siewe Fodjo et al. 2021). Even if no SeroCoV household mentioned COVID-19 related SMS as main

sources of information, the strategy of the GIZ ‘CallVsCorona’, providing written and oral content about the pandemic, seemed to be a good approach, since mobile phones were also very frequent among SeroCov households (Tab. 16) (Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH 2021). Consistent with these observations that participants primarily used classic sources of information, they mainly consulted classic health care facilities for COVID-19 symptoms, as well (Tab. 6.2).

#### 4.7 Non-pharmaceutical protection measures

Similar to a scoping review of COVID-19 prevention measures in SSA countries, handwashing was frequently performed but social distancing was rare among SeroCoV participants (Tab. 10.1) (KB Matovu et al. 2022). In contrast to the scoping review results, wearing face masks was also very frequent among SeroCov households. As the scoping review mainly included studies after the first wave of COVID-19 in 2020, surveyed participants may have already relaxed the stringency of wearing face masks, whereas the SeroCov sampling was conducted during the second wave (Fig. 11 and 12). Thus, participants might have been more cautious and alarmed by the rapidly increasing cases. Especially during the early months of the pandemic, wearing less face masks could have been caused by the unprecedented global need for personal protection equipment (PPE) (Burki 2020). As SSA countries largely relied on import of PPEs from other countries, such as China, this shortage strongly affected SSA countries, which might have limited the regular use of face masks, too (Hakobyan and Cherif 2021). Regarding the use of face masks, the participants’ responses might be biased towards wearing them more often, because it was mandatory and probably socially desired, as well. However, participants seemed to reply honestly, since social distancing was rare. SeroCoV findings underlined the importance of age regarding the conduct of social distancing, which was also observed in other SSA countries (Júnior et al. 2021; Siewe Fodjo et al. 2021). Younger age groups seemed to have greater problems respecting them (Tab. 10.3), yet the majority of participants supported and respected preventive measures against COVID-19. Interestingly, participants within the age group of 20 to 29 years old varied from other age groups with lower approval rates. Since most households indicated that elderly people were particularly threatened by a severe COVID-19 infection (Tab. 11), the age group of 20 to 29 years old may have not feared a severe infection, but other consequences, such as economic or educational throwbacks.

Unlike an online survey conducted in the Democratic Republic of the Congo (DRC), lower levels of education among SeroCoV households were not associated with less adherence to preventive measures (Tab. 10.1) (Ditekemena et al. 2021). In fact, levels of education were quite homogeneously distributed regarding preventive measures except for social distancing to other households, which was mostly performed by participants without a degree. Further, participants without a degree were most likely to identify risk factors for a severe COVID-19 infection, challenging our assumption that lower education was associated with non-adherence to NPIs and lower knowledge about the coronavirus (Tab. 11) (Centers for Disease Control and Prevention 2023).

#### 4.8 Strengths and limitations

One of the major strengths of this study was its execution; laboratories were equipped, staff was trained and all required reagents shipped. Household sampling occurred during lockdown measures and samples were handled appropriately despite logistical challenges. The two-staged sampling strategy of households performed by local teams, not only allowed the assessment of the SARS-CoV-2 seroprevalence, but also identified the sociodemographic characteristics of an urban Malagasy society, providing a more holistic picture of the pandemic in this SSA country. This stands in contrast to the many online conducted surveys, which, for that reason, had the above-mentioned selection bias. Further, while the survey can not be representative for the entire country Madagascar, SeroCoV findings indicated that the participants were similar to official census data of an urban Malagasy population, allowing inferences about them. Lastly, by addressing neglected topics due to stigmatization in Madagascar, e.g., mental health of the population or alcohol consumption, the survey provided first insights and hopefully inspiration for future investigations.

Limitations were the lack of enrolled children. Madagascar has a typical age pyramid for developing countries (Global South), and enrolling participants older or equal to 10 years old resulted in an older median age of 29 years, and the neglect of more than 28 % of the Malagasy population, compared to the census data. However, the exclusion of children younger than 10 years old was discussed prior to the execution of the study, and recommended by the local PI. In addition and similar to other household-based studies, the SeroCoV survey was not free from a selection bias. Excluding community accommodations, such as dormitories, or homeless people might have falsified results towards a wealthier population. According to the census and anecdotal observations by

team members, homeless people are quite frequently in urban ‘Haute-Matsiatra’, having the second-highest percentage of homeless people after the capital region Analamanga (Institut National de la Statistique 2021). Further studies will be needed to assess the extent of this group. SeroCoV results represented urban Malagasy settings, which is the case for the minority of Malagasy, and limits its generalizability for the entire country. Interestingly, other SSA studies performed in border countries observed a higher seropositivity in rural areas possibly due to border travel activities. To which extent these observations are applicable to an island country like Madagascar, will need to be investigated in future studies (Niama et al. 2022; Iroungou et al. 2022). Since antibodies have shown to wane, it may have been possible that a fraction of infections was missed due to a false negative serostatus. The three ambiguous test results could be explained by seroconversion (Tab. 2); either by waning or early stages of antibody production. Other studies performed an RT-PCR to detect active infections in addition to seroprevalence, which would have allowed to survey the COVID-19 pandemic in greater detail (Alber et al. 2022). Lastly, while data collection was conducted accurately by the local teams, the level of education was assessed only at a household level and not at a level of individual household members, limiting its comparability with census data (Tab. 5.1).

## 5. Conclusion

In conclusion, this doctoral thesis aimed to describe the SARS-CoV-2 seroprevalence in Fianarantsoa, identify associations for SARS-CoV-2 seropositivity, and evaluate the household-based sampling strategy. Based on the IgG SARS-CoV-2 ELISA test results, it can be concluded that more than one-third (37.5% [35; 40 %]) of the participants living in Fianarantsoa had been exposed to the virus since the outbreak of the pandemic. Given that either testing for acute SARS-CoV-2 or the number of confirmed infections was rare among study participants, the SeroCoV serological findings exceeded the official figures for Madagascar and highlighted the need for more testing capacities in SSA. The results showed that sex had little effect on seropositivity, and that the seroprevalence varied by age group. Yet, no clear trend was observed. In contrast, the seropositivity increased steadily with household size. Regarding the occupational risk of coronavirus exposure, as expected, HCWs had the highest seroprevalence of all subgroups, whereas agricultural workers had the lowest. Higher seropositivity in larger households, together with the observation that the seroprevalence was lower among participants who spent more time outdoors, confirms the current understanding about the virus transmission. Contrary to the expectation that less educated or less wealthy people would be at a higher risk of an infection, these results reveal a trend for seropositivity to increase from the lowest to the highest quintile; more often, more educated participants were found to be seropositive. No significant differences in SARS-CoV-2 seroprevalence were found among the differently used NPIs. Further, the SeroCoV results showed that wearing a face mask and washing hands more regularly were practiced, regardless of education level. However, social distancing was notably less practiced in this setting. Especially, younger participants indicated that it was the most difficult protection measure to apply. These results are consistent with those of other SSA surveys. The SeroCoV results clearly showed that participants across all levels of education primarily used and trusted classic media platforms, such as TV and radio, as their main source of information about the COVID-19 pandemic. Social media also played a role but was mainly used by more educated participants. Based on these observations, it is recommended that policy makers and researchers continue to use the radio or TV as a communication and information platform to reach the majority of the urban population in the event of a future outbreak. It is important to note that all of these observations reflect the situation of individuals living in actual houses, as the homeless people or nomads, were not considered in this sampling strategy, thus limiting the generalizability. However,

this research study demonstrated that the cross-sectional two-stage cluster population-based study design is still a good and feasible method to select a representative study population in terms of sex distribution, education, occupation, and wealth quintiles. The unexpectedly high levels of well-being despite the prevailing food and economic insecurities, and the worries expressed in the survey, should be followed up in future studies. These include baseline assessment of mental health disorders and substance-use. The burden of chronic diseases was rare among participants. While quantitative analysis was used to describe the study population, future studies should also consider using qualitative analysis to identify contextual features that may not have been specifically asked about. This could provide more detailed information on keeping up a good well-being during the pandemic. This thesis contributed to the field of public health and infectious disease prevention by exploring the contextual influences of participant characteristics on the SARS-CoV-2 seroprevalence in Madagascar. The results confirmed previous observations from other studies performed in SSA, that official testing for acute infections did not cover the true extent of the spread of SARS-CoV-2. Further, food and economic security were severely affected during the pandemic. However, the results challenged the assumptions that participants' well-being was reduced as a result, and that wealthier or more educated people were less exposed to the virus.



## **6. Abstract**

### **Background:**

In December 2019, SARS-CoV-2 emerged in the Wuhan region, China, leading to a pandemic. Despite limited resources, the pandemic appeared to have taken a milder course in sub-Saharan African countries. To assess the underlying SARS-CoV-2 seroprevalence, the cross-sectional SeroCoV study was conducted among households in three African countries. This thesis focuses on the urban Malagasy study site of Fianarantsoa.

### **Methods and Analysis:**

A two-stage cluster population-based study design with randomly distributed GPS – coordinates was used to identify eligible households. The sample size was 650 households stratified by sex and three age groups. Inclusion criteria were being a household member and at least 10 years of age, providing informed consent, and allowing venous blood sampling. An ELISA against SARS-CoV-2 IgG antibodies was performed to determine the seroprevalence. A questionnaire was administered to the participants to collect background information. Data collection was conducted from February to June 2021 using REDCap®. Statistical analysis and visualization were performed in R®.

### **Results:**

Among the 1,121 participants, 37.5 % CI [35; 40 %] were seropositive. Seroprevalence was higher among health care workers (75% CI [36; 96%]), wealthier quintiles (42% CI [34; 49%]), and larger households (41% CI [37; 45%]), but significantly lower among participants spending more time outdoors (33% CI [29; 37]). Due to the pandemic, 68% of households experienced an income reduction, 84 % of participants had to decrease the quantity of food during meals, yet only 28 (2.5%) participants were positively screened for depression. Radio and TV were the main sources of information about the pandemic, regardless of education level.

### **Conclusion:**

The SARS-CoV-2 seroprevalence was higher than officially reported, particularly influenced by occupation, and living conditions. Despite the various economic and food insecurities, depression seemed to be rare among participants. In future outbreaks in Madagascar, policymakers should use the TV and radio to inform the population.

## **7. Zusammenfassung**

### **Hintergrund:**

Im Dezember 2019 trat in der Region Wuhan in China das neuartige SARS-CoV-2 auf und führte zu einer Pandemie. Trotz begrenzter Ressourcen schien die Pandemie in sub-Sahara Afrika einen milderen Verlauf zu nehmen. Um die zugrunde liegende SARS-CoV-2-Seroprävalenz zu ermitteln, wurde die Querschnittsstudie SeroCoV in Haushalten von drei afrikanischen Ländern durchgeführt. Der Fokus dieser Dissertation ist der urbane Studienort Fianarantsoa in Madagaskar.

### **Methoden und Analyse:**

Die einzuschließenden Haushalte wurden mittels eines zweistufigen bevölkerungsbasiertes Cluster-Studiendesign und zufällig verteilten GPS-Koordinaten ausgewählt. Die Stichprobengröße betrug 650 Haushalte, geschichtet nach Geschlecht und drei Altersgruppen. Die Einschlusskriterien waren Haushaltszugehörigkeit, ein Mindestalter von 10 Jahren, eine Einverständniserklärung und die Bereitschaft zur venösen Blutentnahme. Die Seroprävalenz wurde mit einem ELISA gegen SARS-CoV-2 IgG-Antikörper bestimmt. Ein Fragebogen diente der Erfassung von Hintergrundinformationen. Die Daten wurden von Februar bis Juni 2021 gesammelt und in REDCap® gespeichert. Die statistische Analyse und Visualisierung erfolgten in R®.

### **Ergebnisse:**

Die SARS-CoV-2 Seroprävalenz unter den 1.121 Teilnehmern betrug 37,5 % CI [35; 40 %]. Häufiger seropositiv waren Teilnehmer in medizinischen Berufen (75 % CI [36; 96 %]), im wohlhabendsten Quintil (42 % CI [34; 49 %]) und in größeren Haushalten (41 % CI [37; 45 %]). Die Seroprävalenz war signifikant niedriger bei Teilnehmern, die mehr Zeit im Freien verbrachten (33 % CI [29; 37]). Infolge der Pandemie litten 68 % der Haushalte unter Einkommenseinbußen, 84 % der Teilnehmer mussten die Menge ihrer Mahlzeiten reduzieren, aber nur 28 (2,5 %) der Befragten litten wahrscheinlich an einer Depression. Unabhängig vom Bildungsniveau waren Radio und Fernsehen die wichtigsten Informationsquellen über die Pandemie.

### **Schlussfolgerung:**

Die SARS-CoV-2-Seroprävalenz war höher als offiziell angegeben, wobei vor allem Beruf und Lebensbedingungen diese beeinflussten. Trotz der verschiedenen wirtschaftlichen und Ernährungsunsicherheiten, schienen unter den Teilnehmern Depressionen selten zu sein. Bei künftigen Ausbrüchen in Madagaskar sollten Fernsehen und Radio genutzt werden, um die Bevölkerung zu informieren.

## 8. List of abbreviations

AC	<i>Agent Communautaire</i>	NPI	Non-pharmaceutical Intervention
BEP	Blood EDTA Plasma	OCHA	UN Office for the Coordination of Humanitarian Affairs
BNITM	Bernhard Nocht Institute for Tropical Medicine	OD	Optical Density
BSL-2	Biosafety Level 2	ODCO	Optical Density cut-off
CERBM	Comité d'Éthique de la Recherche Biomédicale	OsmAnd®	OpenStreetMap Automated Navigation Directions
CGU	Centre Hospitalier Universitaire	PC	Principal Component
CI	Confidence Interval	PCA	Principal Component Analysis
CICM	Centre d'Infectiologie Charles Mérieux	PCR	Polymerase Chain Reaction
COVID-19	Coronavirus Disease 2019	PI	Principal Investigator
CRA	Clinical Research Assistant	PPE	Personal Protection Equipment
CRCO	Régional de Commandement Opérationnel COVID-19	RC	Rotated Components
DHS	Demographic and Health Survey	REDCap®	Research Electronic Data Capture
EDTA	Ethylenediaminetetraacetic acid	RT-PCR	Real-Time Polymerase Chain Reaction
ELISA	Enzyme-linked Immunosorbent Assay	SARS-CoV-1	Severe Acute Respiratory Syndrome Coronavirus type 1
GCP	Good Clinical Practice	SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus type 2
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH	SeroCoV	Seroprevalence of SARS-CoV-2 in urban settings in three sub-Saharan African countries
GNI	Gross National Income	SID	Screening Identifier
GPS	Global Positioning System	SOP	Standard Operating Procedure
HCW	Health Care Workers	SRMA	Systematic Review and Meta-Analysis
HID	Household Identifier	SSA	Sub-Saharan Africa
HRP	Horseradish Peroxidase	TMB	Tetramethylbenzidine
IgG	Immunoglobulin G	TV	Television
ILO	International Labor Organization	UN	United Nations
IQR	Interquartile Range	WHO	World Health Organization
LMIC	Low- and Middle-Income Country		
MERS-CoV	Middle East Respiratory Syndrome Coronavirus		
MID	Member Identifier		
NCP	Nucleocapsid Protein		

## 9. References

- ACAPS. 2022. 'Madagascar: Food Insecurity Crisis in the Grand Sud Regions'. [https://www.acaps.org/search?search\\_query=madagascar](https://www.acaps.org/search?search_query=madagascar).
- Adams, Janica, Mary J. MacKenzie, Adeladza Kofi Amegah, Alex Ezeh, Muktar A. Gadanya, Akinyinka Omigbodun, Ahmed M. Sarki, et al. 2021. 'The Conundrum of Low COVID-19 Mortality Burden in Sub-Saharan Africa: Myth or Reality?' *Global Health: Science and Practice* 9 (3): 433–43. <https://doi.org/10.9745/GHSP-D-21-00172>.
- Adetifa, Ifedayo M. O., Sophie Uyoga, John N. Gitonga, Daisy Mugo, Mark Otiende, James Nyagwange, Henry K. Karanja, et al. 2021. 'Temporal Trends of SARS-CoV-2 Seroprevalence during the First Wave of the COVID-19 Epidemic in Kenya'. *Nature Communications* 12 (1): 3966. <https://doi.org/10.1038/s41467-021-24062-3>.
- African Defense Forum. 2021. 'U.S. Sends COVID-19 Aid To Madagascar'. 13 April 2021. <https://adf-magazine.com/2021/04/u-s-sends-covid-19-aid-to-madagascar/>.
- Alber, Dagmar, Fadima Cheick Haidara, Juho Luoma, Laura Adubra, Per Ashorn, Ulla Ashorn, Henry Badji, et al. 2022. 'SARS-CoV-2 Infection and Antibody Seroprevalence in Routine Surveillance Patients, Healthcare Workers and General Population in Kita Region, Mali: An Observational Study 2020–2021'. *BMJ Open* 12 (6): e060367. <https://doi.org/10.1136/bmjopen-2021-060367>.
- Arora, Rahul K, Abel Joseph, Jordan Van Wyk, Simona Rocco, Austin Atmaja, Ewan May, Tingting Yan, et al. 2021. 'SeroTracker: A Global SARS-CoV-2 Seroprevalence Dashboard'. *The Lancet Infectious Diseases* 21 (4): e75–76. [https://doi.org/10.1016/S1473-3099\(20\)30631-9](https://doi.org/10.1016/S1473-3099(20)30631-9).
- Arya, Ravindra, Belavendra Antonisamy, and Sushil Kumar. 2012. 'Sample Size Estimation in Prevalence Studies'. *The Indian Journal of Pediatrics* 79 (11): 1482–88. <https://doi.org/10.1007/s12098-012-0763-3>.
- Assefa, Nega, Abdramane Soura, Elena C. Hemler, Michelle L. Korte, Dongqing Wang, Yasir Y. Abdullahi, Bruno Lankoande, et al. 2021. 'COVID-19 Knowledge, Perception, Preventive Measures, Stigma, and Mental Health Among Healthcare Workers in Three Sub-Saharan African Countries: A Phone Survey'. *The American Journal of Tropical Medicine and Hygiene* 105 (2): 342–50. <https://doi.org/10.4269/ajtmh.20-1621>.
- Austrian, Karen, Jessie Pinchoff, James B. Tidwell, Corinne White, Timothy Abuya, Beth Kangwana, Rhoune Ochako, et al. 2020. 'COVID-19 Related Knowledge, Attitudes, Practices and Needs of Households in Informal Settlements in Nairobi, Kenya'. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3576785>.
- Bakohariliva, Hasina Andrianarivony, Imisanavalona Hanitrinihaja Rafehivola, Evah Norotiana Raobelle, Adeline Raharivelo, and Bertille Hortense Rajaonarison. 2018. 'Etat Des Lieux Des Soins de Premier Recours Des Malades Mentaux à Antananarivo: Étude Rétrospective'. *Pan African Medical Journal* 29. <https://doi.org/10.11604/pamj.2018.29.1.11168>.
- Baquero, Oswaldo Santos. 2019. *Ggsn: North Symbols and Scale Bars for Maps Created with 'ggplot2' or 'Ggmap'*. <https://CRAN.R-project.org/package=ggsn>.
- BBC. 2020. 'Coronavirus: Africa Could Be next Epicentre, WHO Warns'. 2020. <https://www.bbc.com/news/world-africa-52323375>.
- Borgia, Gaëlle. 2021. 'Covid-19: l'oxygène manque à Madagascar, submergé par une seconde vague de malades'. 27 April 2021.

- <https://www.france24.com/fr/afrique/20210427-covid-19-l-oxygène-manque-à-madagascar-submergé-par-une-seconde-vague-de-malades>.
- Britannica, The Information Architects of Encyclopaedia. 2022. 'Madagascar'. 2022. <https://www.britannica.com/facts/Madagascar>.
- Brown Wilson, Josiah, Andreas Deckert, Rupal Shah, Nicholas Kyei, Laretta Copeland Dahn, Rosalita Doe-Rogers, Augustine Bowuo Hinneh, et al. 2021. 'COVID-19-Related Knowledge, Attitudes and Practices: A Mixed-Mode Cross-Sectional Survey in Liberia'. *BMJ Open* 11 (7): e049494. <https://doi.org/10.1136/bmjopen-2021-049494>.
- Bulfone, Tommaso Celeste, Mohsen Malekinejad, George W Rutherford, and Nooshin Razani. 2021. 'Outdoor Transmission of SARS-CoV-2 and Other Respiratory Viruses: A Systematic Review'. *The Journal of Infectious Diseases* 223 (4): 550–61. <https://doi.org/10.1093/infdis/jiaa742>.
- Burki, Talha. 2020. 'Global Shortage of Personal Protective Equipment'. *The Lancet Infectious Diseases* 20 (7): 785–86. [https://doi.org/10.1016/S1473-3099\(20\)30501-6](https://doi.org/10.1016/S1473-3099(20)30501-6).
- Cabore, Joseph Waogodo, Humphrey Cyprian Karamagi, Hillary Kipchumba Kipruto, Joseph Kyalo Mungatu, James Avoka Asamani, Benson Droti, Regina Titi-ofei, et al. 2022. 'COVID-19 in the 47 Countries of the WHO African Region: A Modelling Analysis of Past Trends and Future Patterns'. *The Lancet Global Health* 10 (8): e1099–1114. [https://doi.org/10.1016/S2214-109X\(22\)00233-9](https://doi.org/10.1016/S2214-109X(22)00233-9).
- Cénat, Jude Mary, Pari-Gole Noorishad, Cyrille Kossigan Kokou-Kpolou, Rose Darly Dalexis, Saba Hajizadeh, Mireille Guerrier, Lewis Ampidu Clorméus, et al. 2021. 'Prevalence and Correlates of Depression during the COVID-19 Pandemic and the Major Role of Stigmatization in Low- and Middle-Income Countries: A Multinational Cross-Sectional Study'. *Psychiatry Research* 297 (March): 113714. <https://doi.org/10.1016/j.psychres.2021.113714>.
- Centers for Disease Control and Prevention. 2023. 'People with Certain Medical Conditions'. 10 February 2023. <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html>.
- Chen, Jiyao, Nusrat Farah, Rebecca Kechen Dong, Richard Z. Chen, Wen Xu, Jin Yin, Bryan Z. Chen, et al. 2021. 'Mental Health during the COVID-19 Crisis in Africa: A Systematic Review and Meta-Analysis'. *International Journal of Environmental Research and Public Health* 18 (20): 10604. <https://doi.org/10.3390/ijerph182010604>.
- Communauté Urbaine de Yaoundé. 2023. 'Présentation de la ville de Yaoundé'. 2023. <https://yaounde.cm/presentation-de-la-ville-de-yaounde/>.
- Contreras-Gonzalez, Ivette, Gbemisola Oseni, Amparo Palacios-Lopez, Janneke Pieters, and Michael Weber. 2022. 'Inequalities in Job Loss and Income Loss in Sub-Saharan Africa during the COVID-19 Crisis'.
- Costello, AB, and Jason Osborne. 2014. 'Best Practices in Exploratory Factor Analysis'. *Create Space Independent Publishing* 10: 33.
- Cui, Jie, Fang Li, and Zheng-Li Shi. 2019. 'Origin and Evolution of Pathogenic Coronaviruses'. *Nature Reviews Microbiology* 17 (3): 181–92. <https://doi.org/10.1038/s41579-018-0118-9>.
- Dasgupta, Shouro, and Elizabeth J. Z. Robinson. 2021. 'Food Insecurity, Safety Nets, and Coping Strategies during the COVID-19 Pandemic: Multi-Country Evidence from Sub-Saharan Africa'. *International Journal of Environmental Research and Public Health* 18 (19): 9997. <https://doi.org/10.3390/ijerph18199997>.
- Deschermeier, Christina, Christa Ehmen, Ronald von Possel, Carolin Murawski, Ben

- Rushton, John Amuasi, Nimako Sarpong, et al. 2022. 'Fcγ-Receptor-Based Enzyme-Linked Immunosorbent Assays for Sensitive, Specific, and Persistent Detection of Anti-SARS-CoV-2 Nucleocapsid Protein IgG Antibodies in Human Sera'. Edited by Yi-Wei Tang. *Journal of Clinical Microbiology* 60 (6): e00075-22. <https://doi.org/10.1128/jcm.00075-22>.
- Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH. 2021. 'Tackling COVID-19 by phone: using mobile phones to respond to the pandemic'. 16 September 2021. <https://www.giz.de/en/mediacenter/101099.html>.
- Ditekemena, John D., Hypolite M. Mavoko, Michael Obimpeh, Stijn Van Hees, Joseph Nelson Siewe Fodjo, Dalau M. Nkamba, Antoinette Tshefu, Wim Van Damme, Jean Jacques Muyembe, and Robert Colebunders. 2021. 'Adherence to COVID-19 Prevention Measures in the Democratic Republic of the Congo, Results of Two Consecutive Online Surveys'. *International Journal of Environmental Research and Public Health* 18 (5): 2525. <https://doi.org/10.3390/ijerph18052525>.
- Dragiotti, Elena, Han Li, George Tsitsas, Keum Hwa Lee, Jiwoo Choi, Jiwon Kim, Young Jo Choi, et al. 2022. 'A Large-scale Meta-analytic Atlas of Mental Health Problems Prevalence during the COVID-19 Early Pandemic'. *Journal of Medical Virology* 94 (5): 1935–49. <https://doi.org/10.1002/jmv.27549>.
- Etyang, Anthony O, Ruth Lucinde, Henry Karanja, Catherine Kalu, Daisy Mugo, James Nyagwange, John Gitonga, et al. 2022. 'Seroprevalence of Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2 Among Healthcare Workers in Kenya'. *Clinical Infectious Diseases* 74 (2): 288–93. <https://doi.org/10.1093/cid/ciab346>.
- European Centre for Disease Prevention and Control. 2023. 'SARS-CoV-2 Variants of Concern as of 23 March 2023'. 24 March 2023. <https://www.ecdc.europa.eu/en/covid-19/variants-concern>.
- European Medicines Agency. 2023. 'COVID-19 Vaccines: Authorised'. 27 March 2023. <https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/vaccines-covid-19/covid-19-vaccines-authorised>.
- Eurostat. 2020. 'Number of Psychiatrists: How Do Countries Compare?' 6 May 2020. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20200506-1>.
- Field, Andy P., Jeremy Miles, and Zoë Field. 2012. *Discovering Statistics Using R*. London ; Thousand Oaks, Calif: Sage.
- Foubert, Laurent, Yvonnick Noël, Chandler M. Spahr, and George M. Slavich. 2021. 'Beyond WEIRD: Associations between Socioeconomic Status, Gender, Lifetime Stress Exposure, and Depression in Madagascar'. *Journal of Clinical Psychology* 77 (7): 1644–65. <https://doi.org/10.1002/jclp.23131>.
- Fryatt, Arun, Victoria Simms, Tsitsi Bandason, Nicol Redzo, Ioana D. Oлару, Chiratidzo E Ndhlovu, Hilda Mujuru, et al. 2021. 'Community SARS-CoV-2 Seroprevalence before and after the Second Wave of SARS-CoV-2 Infection in Harare, Zimbabwe'. *EClinicalMedicine* 41 (November): 101172. <https://doi.org/10.1016/j.eclinm.2021.101172>.
- Gangavarapu, Karthik, Alaa Abdel Latif, Julia L. Mullen, Manar Alkuzweny, Emory Hufbauer, Ginger Tsueng, Emily Haag, et al. 2023. 'Outbreak.Info Genomic Reports: Scalable and Dynamic Surveillance of SARS-CoV-2 Variants and Mutations'. *Nature Methods*, February. <https://doi.org/10.1038/s41592-023-01769-3>.
- Guan, Wei-jie, Zheng-yi Ni, Yu Hu, Wen-hua Liang, Chun-quan Ou, Jian-xing He, Lei Liu, et al. 2020. 'Clinical Characteristics of Coronavirus Disease 2019 in China'.

- New England Journal of Medicine* 382 (18): 1708–20.  
<https://doi.org/10.1056/NEJMoa2002032>.
- Hajissa, Khalid, Md Asiful Islam, Siti Asma Hassan, Abdul Rahman Zaidah, Nabilah Ismail, and Zeehaida Mohamed. 2022. ‘Seroprevalence of SARS-CoV-2 Antibodies in Africa: A Systematic Review and Meta-Analysis’. *International Journal of Environmental Research and Public Health* 19 (12): 7257.  
<https://doi.org/10.3390/ijerph19127257>.
- Hakobyan, Shushanik, and Reda Cherif. 2021. ‘Trade in Medical Goods: Challenges and a Way Forward for Sub-Saharan Africa’. *Special Series on COVID 19*.
- Hale, Thomas, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, et al. 2021. ‘A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker)’. *Nature Human Behaviour* 5 (4): 529–38. <https://doi.org/10.1038/s41562-021-01079-8>.
- Harris, Paul A., Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, and Jose G. Conde. 2009. ‘Research Electronic Data Capture (REDCap)—A Metadata-Driven Methodology and Workflow Process for Providing Translational Research Informatics Support’. *Journal of Biomedical Informatics* 42 (2): 377–81. <https://doi.org/10.1016/j.jbi.2008.08.010>.
- Hartmann, Kai, Joachim Krois, and Björn Waske. 2018. ‘E-Learning Project SOGA: Statistics and Geospatial Data Analysis’. *Department of Earth Sciences, Freie Universitaet Berlin* 33.
- Hjelm, Lisa, Astrid Mathiassen, Darryl Miller, and Amit Wadhwa. 2017. ‘Creation of a Wealth Index’. *World Food Program*.
- Hu, Ben, Hua Guo, Peng Zhou, and Zheng-Li Shi. 2021. ‘Characteristics of SARS-CoV-2 and COVID-19’. *Nature Reviews Microbiology* 19 (3): 141–54.  
<https://doi.org/10.1038/s41579-020-00459-7>.
- Iannone, Richard. 2022. *DiagrammeR: Graph/Network Visualization*. <https://CRAN.R-project.org/package=DiagrammeR>.
- Iannone, Richard, Joe Cheng, Barret Schloerke, Ellis Hughes, and JooYoung Seo. 2023. ‘Gt: Easily Create Presentation-Ready Display Tables’.  
<https://github.com/rstudio/gt>.
- Institut National de la Statistique. 2021. ‘Troisième Recensement Général de la Population et de l’Habitation (RGPH-3)’. BP 485 - Anosy Antananarivo - 101 Madagascar: INSTAT. <https://instat.mg/>.
- Institut National de la Statistique, and ICF. 2022. ‘Enquête Démographique et de Santé à Madagascar, 2021’. Antananarivo, Madagascar et Rockville, Maryland, USA: INSTAT et ICF. [https://dhsprogram.com/publications/publication-fr376-dhs-final-reports.cfm?csearch=638726\\_1](https://dhsprogram.com/publications/publication-fr376-dhs-final-reports.cfm?csearch=638726_1).
- International Labour Organization. 2023. ‘ILO Modelled Estimates Database’. 2023.  
<https://ilostat.ilo.org/data/>.
- Iroungou, Berthe A., Pamela Boundzanga Moussavou, Eric Elguero, Jamal Makran, Berthold Bivigou-Mboumba, Elise Wora, Jean R. Nzenze, Fabien Roche Niama, and Avelin F. Aghokeng. 2022. ‘Trend of Expansion of SARS-CoV-2 Infection and COVID-19 Burden in Gabon (Central Africa) in Mid-2021, Based on a Serological Survey’. *IJID Regions* 5 (December): 13–17.  
<https://doi.org/10.1016/j.ijregi.2022.08.006>.
- Júnior, António, Janeth Dula, Sérgio Mahumane, Olivier Koole, Sónia Enosse, Joseph Nelson Siewe Fodjo, and Robert Colebunders. 2021. ‘Adherence to COVID-19 Preventive Measures in Mozambique: Two Consecutive Online Surveys’. *International Journal of Environmental Research and Public Health* 18 (3): 1091. <https://doi.org/10.3390/ijerph18031091>.

- Kahle, David, and Hadley Wickham. 2013. 'Ggmap: Spatial Visualization with Ggplot2'. *The R Journal* 5 (1): 144–61.
- Kamvar, Zhian N. 2023. *Apyramid: Visualize Population Pyramids Aggregated by Age*. <https://CRAN.R-project.org/package=apyramid>.
- Kassambara, Alboukadel. 2022. *Ggcorrplot: Visualization of a Correlation Matrix Using 'Ggplot2'*. <https://CRAN.R-project.org/package=ggcorrplot>.
- KB Matovu, Joseph, Alex Mulyowa, Rogers Akorimo, and Daniel Kirumira. 2022. 'Knowledge, Risk-Perception, and Uptake of COVID-19 Prevention Measures in Sub-Saharan Africa: A Scoping Review'. *African Health Sciences* 22 (3): 542–60. <https://doi.org/10.4314/ahs.v22i3.59>.
- Kingue, Samuel, Solofonirina Rakotoarimanana, Nirina Rabearivony, and Francois Lepira Bompera. 2017. 'Prevalence of Selected Cardiometabolic Risk Factors among Adults in Urban and Semi-Urban Hospitals in Four Sub-Saharan African Countries'. *Cardiovascular Journal of Africa* 28 (3): 147–53. <https://doi.org/10.5830/CVJA-2016-072>.
- Kwabla, Mavis Pearl, Juliana Nyasordzi, Gideon Kye-Duodu, Mark Kwame Ananga, Gregory Kofi Amenuvegbe, Joseph Otoo, Dominic Demateh Nuerthey, et al. 2022. 'Factors Associated with COVID-19 Knowledge among Ghanaians: A National Survey'. Edited by Clement Ameh Yaro. *PLOS ONE* 17 (11): e0276381. <https://doi.org/10.1371/journal.pone.0276381>.
- Lehner, Bernhard, and Günther Grill. 2013. 'Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems: GLOBAL RIVER HYDROGRAPHY AND NETWORK ROUTING'. *Hydrological Processes* 27 (15): 2171–86. <https://doi.org/10.1002/hyp.9740>.
- Lewis, Hannah C, Harriet Ware, Mairead Whelan, Lorenzo Subissi, Zihan Li, Xiaomeng Ma, Anthony Nardone, et al. 2022. 'SARS-CoV-2 Infection in Africa: A Systematic Review and Meta-Analysis of Standardised Seroprevalence Studies, from January 2020 to December 2021'. *BMJ Global Health* 7 (8): e008793. <https://doi.org/10.1136/bmjgh-2022-008793>.
- Li, Jie, Daniel Q. Huang, Biyao Zou, Hongli Yang, Wan Zi Hui, Fajuan Rui, Natasha Tang Sook Yee, et al. 2021. 'Epidemiology of COVID-19: A Systematic Review and Meta-analysis of Clinical Characteristics, Risk Factors, and Outcomes'. *Journal of Medical Virology* 93 (3): 1449–58. <https://doi.org/10.1002/jmv.26424>.
- Lin, Alice V. 2015. 'Direct ELISA'. In *ELISA*, edited by Robert Hnasko, 1318:61–67. *Methods in Molecular Biology*. New York, NY: Springer New York. [https://doi.org/10.1007/978-1-4939-2742-5\\_6](https://doi.org/10.1007/978-1-4939-2742-5_6).
- Lorenz, Eva, Aurélia Souares, John Amuasi, Wibke Loag, Christina Deschermeier, Daniela Fusco, Dominik Benke, et al. 2021. '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'. *BMJ Open* 11 (12): e056853. <https://doi.org/10.1136/bmjopen-2021-056853>.
- Mandimbisoa, R. 2021. 'Bilan Covid-19: 29 décès et 1 327 cas en 15 jours'. *Madagascar-Tribune.com*, March. <https://www.madagascar-tribune.com/Bilan-Covid-19-29-deces-et-1-327-cas-en-15-jours.html>.
- Mandolo, Jonathan, Jacqueline Msefula, Marc Y. R. Henrion, Comfort Brown, Brewster Moyo, Aubrey Samon, Thandeka Moyo-Gwete, et al. 2021. 'SARS-CoV-2 Exposure in Malawian Blood Donors: An Analysis of Seroprevalence and Variant Dynamics between January 2020 and July 2021'. *BMC Medicine* 19 (1): 303. <https://doi.org/10.1186/s12916-021-02187-y>.



- Meyerowitz, Eric A., Aaron Richterman, Rajesh T. Gandhi, and Paul E. Sax. 2021. 'Transmission of SARS-CoV-2: A Review of Viral, Host, and Environmental Factors'. *Annals of Internal Medicine* 174 (1): 69–79. <https://doi.org/10.7326/M20-5008>.
- Mlcochova, Petra, Steven A. Kemp, Mahesh Shanker Dhar, Guido Papa, Bo Meng, Isabella A. T. M. Ferreira, Rawlings Datir, et al. 2021. 'SARS-CoV-2 B.1.617.2 Delta Variant Replication and Immune Evasion'. *Nature* 599 (7883): 114–19. <https://doi.org/10.1038/s41586-021-03944-y>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Munyeku-Bazitama, Yannick, Gervais T. Folefack, Marc K. Yambayamba, Paul M. Tshiminyi, Benito M. Kazenza, John O. Otshudiema, Noe Tondri Guinko, et al. 2023. 'High SARS-CoV-2 Seroprevalence after Second COVID-19 Wave (October 2020–April 2021), Democratic Republic of the Congo'. *Emerging Infectious Diseases* 29 (1): 89–97. <https://doi.org/10.3201/eid2901.221009>.
- National Institutes of Health. 2023. 'Coronavirus Disease 2019 (COVID-19) Treatment Guidelines'. 6 March 2023. <https://www.covid19treatmentguidelines.nih.gov/>.
- Ndongo, Francis Ateba, Emilande Guichet, Eric Donald Mimbé, Justin Ndié, Raphael Pelloquin, Marie Varloteaux, Livo Esemu, et al. 2022. 'Rapid Increase of Community SARS-CoV-2 Seroprevalence during Second Wave of COVID-19, Yaoundé, Cameroon'. *Emerging Infectious Diseases* 28 (6). <https://doi.org/10.3201/eid2806.212580>.
- Negera, Assefa, Gebeyehu Abelti, Terefe Bogile, Tesfayi Gebrellassie, and Roger Person. 2013. 'An Analysis of the Trends, Differentials, and Key Proximate Determinates of Infant and under-Five Mortality in Ethiopia: Further Analysis of the 2000, 2005, and 2011 Demographic and Health Surveys'. DHS Further Analysis Reports No. 79. Calverton, Maryland, USA: ICF International. <http://dhsprogram.com/pubs/pdf/FA79/FA79.pdf>.
- Niama, Fabien R., Félix Koukouikila-Koussounda, Pembé Issamou Mayengue, Eric Elguero, Tarcisse Baloki Ngoulou, Victor Levier, Jamal Makran, Berthe A. Iroungou, and Avelin F. Aghokeng. 2022. 'Prevalence of SARS-CoV-2 Antibodies in the Republic of Congo in Mid-2021'. *IJID Regions* 5 (December): 121–23. <https://doi.org/10.1016/j.ijregi.2022.09.012>.
- Nwosu, Kene, Joseph Fokam, Franck Wanda, Lucien Mama, Erol Orel, Nicolas Ray, Jeanine Meke, et al. 2021. 'SARS-CoV-2 Antibody Seroprevalence and Associated Risk Factors in an Urban District in Cameroon'. *Nature Communications* 12 (1): 5851. <https://doi.org/10.1038/s41467-021-25946-0>.
- Ooms, Jeroen. 2023. *Magick: Advanced Graphics and Image-Processing in R*. <https://CRAN.R-project.org/package=magick>.
- 'OsmAnd'. 2023. 2023. <https://osmand.net>.
- Osuagwu, Uchechukwu L., Chundung A. Miner, Dipesh Bhattarai, Khathutshelo Percy Mashige, Richard Oloruntoba, Emmanuel Kwasi Abu, Bernadine Ekpenyong, et al. 2021. 'Misinformation About COVID-19 in Sub-Saharan Africa: Evidence from a Cross-Sectional Survey'. *Health Security* 19 (1): 44–56. <https://doi.org/10.1089/HS.2020.0202>.
- Paci, Pierella. 2021. 'How Livelihoods Deteriorated in Sub-Saharan Africa Due to Covid-19'. *Poverty and Equity Global Practice, World Bank*, January. <https://blogs.worldbank.org/africacan/how-livelihoods-deteriorated-sub-saharan-africa-due-covid-19>.
- Pebesma, Edzer. 2018. 'Simple Features for R: Standardized Support for Spatial Vector Data'. *The R Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.

- Prati, Gabriele, and Anthony D. Mancini. 2021. 'The Psychological Impact of COVID-19 Pandemic Lockdowns: A Review and Meta-Analysis of Longitudinal Studies and Natural Experiments'. *Psychological Medicine* 51 (2): 201–11. <https://doi.org/10.1017/S0033291721000015>.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- R. F. Randrianarivo, H. A. Bakohariliva, H.H. Ratobimanankasina, V.A. Rabiakely, E.N. Raobelle, D.K. Andrianarimanana, B.H. Rajaonarison, and A. Raharivelo. 2023. 'Mental Illness Stigmatization Factors: Survey among Caregivers in Mahajanga Madagascar'. *World Journal of Advanced Research and Reviews* 17 (1): 1083–89. <https://doi.org/10.30574/wjarr.2023.17.1.0195>.
- Rakotonanahary, Rado J. L., Herinjaka Andriambolamanana, Benedicte Razafinjato, Estelle M. Raza-Fanomezananahary, Vero Ramanandraitsiory, Fiainamirindra Ralaivavikoa, Andritiana Tsirinomen'ny Aina, et al. 2021. 'Integrating Health Systems and Science to Respond to COVID-19 in a Model District of Rural Madagascar'. *Frontiers in Public Health* 9 (July): 654299. <https://doi.org/10.3389/fpubh.2021.654299>.
- Rakotoniana, Andriamiarimbola Irène, Isabelle Ramandimbisoa, Andry Rasamindrakotroka, and Alson Rakoto. 2019. 'Blood Donors Profile at the Blood Transfusion Regional Center Tulear, Madagascar'. *Haematology International Journal* 3 (3). <https://doi.org/10.23880/HIJ-16000147>.
- Ratovoson, Rila, Mihaja Raberahona, Rado Razafimahatratra, Lova Randriamanantsoa, Emmanuel Harizaka Andriamasy, Perlinot Herindrany, Norosoa Razanajatovo, et al. 2022. 'SARS-CoV-2 Infection Rate in Antananarivo Frontline Health Care Workers, Madagascar'. *Influenza and Other Respiratory Viruses* 16 (6): 994–1003. <https://doi.org/10.1111/irv.13022>.
- Razafimahatratra, Solohery Lalaina, Mame Diarra Bousso Ndiaye, Lova Tsikiniaina Rasoloharimanana, Philippe Dussart, Paquerette Hanitriniala Sahondranirina, Zely Arivelo Randriamanantany, and Matthieu Schoenhals. 2021. 'Seroprevalence of Ancestral and Beta SARS-CoV-2 Antibodies in Malagasy Blood Donors'. *The Lancet Global Health* 9 (10): e1363–64. [https://doi.org/10.1016/S2214-109X\(21\)00361-2](https://doi.org/10.1016/S2214-109X(21)00361-2).
- Razanaboninahitra, Famatanantsoa. 2020. 'Being a Community Agent Means Effectively Supporting Routine Vaccination during COVID19'. *UNICEF*, October. <https://www.unicef.org/madagascar/en/stories/being-community-agent-means-effectively-supporting-routine-vaccination-during-covid19>.
- Reuben, Rine Christopher, Margaret M. A. Danladi, Dauda Akwai Saleh, and Patricia Ene Ejembi. 2021. 'Knowledge, Attitudes and Practices Towards COVID-19: An Epidemiological Survey in North-Central Nigeria'. *Journal of Community Health* 46 (3): 457–70. <https://doi.org/10.1007/s10900-020-00881-1>.
- Revelle, William. 2022. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. <https://CRAN.R-project.org/package=psych>.
- Rice, Benjamin L., Justin Lessler, Clifton McKee, and C. Jessica E. Metcalf. 2022. 'Why Do Some Coronaviruses Become Pandemic Threats When Others Do Not?' *PLOS Biology* 20 (5): e3001652. <https://doi.org/10.1371/journal.pbio.3001652>.
- Ritchie, Hannah, and Max Roser. 2018a. 'Alcohol Consumption'. *Our World in Data*.  
 ———. 2018b. 'Urbanization'. *Our World in Data*. 2018. <https://ourworldindata.org/urbanization#citation>.

- Roser, Max, and Hannah Ritchie. 2019. 'Hunger and Undernourishment'. *Our World in Data*.
- Rudin-Rush, Lorin, Jeffrey D. Michler, Anna Josephson, and Jeffrey R. Bloem. 2022. 'Food Insecurity during the First Year of the COVID-19 Pandemic in Four African Countries'. *Food Policy* 111 (August): 102306. <https://doi.org/10.1016/j.foodpol.2022.102306>.
- Rutstein, Shea O., and Kiersten Johnson. 2004. 'The DHS Wealth Index'. DHS Comparative Reports No. 6. Calverton, Maryland, USA: ORC Macro. <http://dhsprogram.com/pubs/pdf/CR6/CR6.pdf>.
- Schoenhals, Matthieu, Niry Rabenindrina, Jean Marius Rakotondramanga, Philippe Dussart, Rindra Randremanana, Jean-Michel Heraud, Soa Fy Andriamandimby, et al. 2021. 'SARS-CoV-2 Antibody Seroprevalence Follow-up in Malagasy Blood Donors during the 2020 COVID-19 Epidemic'. *EBioMedicine* 68 (June): 103419. <https://doi.org/10.1016/j.ebiom.2021.103419>.
- Siewe Fodjo, Joseph Nelson, Leonard Ngarka, Wepnyu Y. Njamnshi, Leonard N. Nfor, Michel K. Mengnjo, Edwige L. Mendo, Samuel A. Angwafor, et al. 2021. 'COVID-19 Preventive Behaviours in Cameroon: A Six-Month Online National Survey'. *International Journal of Environmental Research and Public Health* 18 (5): 2554. <https://doi.org/10.3390/ijerph18052554>.
- Sjoberg, Daniel D., Karissa Whiting, Michael Curry, Jessica A. Lavery, and Joseph Larmarange. 2021. 'Reproducible Summary Tables with the Gtsummary Package'. *The R Journal* 13 (1): 570–80. <https://doi.org/10.32614/RJ-2021-053>.
- Slowikowski, Kamil. 2023. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*. <https://CRAN.R-project.org/package=ggrepel>.
- Smith Rodgers, Lindsay. 2022. 'Why Confirming The Origin of COVID-19 Matters'. 15 August 2022. <https://publichealth.jhu.edu/2022/origins-of-sars-cov-2>.
- Smitham, Eleni, and Amanda Glassman. 2021. 'The next Pandemic Could Come Soon and Be Deadlier'. *Center for Global Development* (blog). 25 August 2021. <https://www.cgdev.org/blog/the-next-pandemic-could-come-soon-and-be-deadlier>.
- Struck, Nicole S., Eva Lorenz, Christina Deschermeier, Daniel Eibach, Jenny Kettenbeil, Wibke Loag, Steven A. Brieger, et al. 2022. 'High Seroprevalence of SARS-CoV-2 in Burkina-Faso, Ghana and Madagascar in 2021: A Population-Based Study'. *BMC Public Health* 22 (1): 1676. <https://doi.org/10.1186/s12889-022-13918-y>.
- Tegally, Houriiyah, Eduan Wilkinson, Marta Giovanetti, Arash Iranzadeh, Vagner Fonseca, Jennifer Giandhari, Deelan Doolabh, et al. 2021. 'Detection of a SARS-CoV-2 Variant of Concern in South Africa'. *Nature* 592 (7854): 438–43. <https://doi.org/10.1038/s41586-021-03402-9>.
- The Humanitarian Data Exchange. 2021a. 'Madagascar - Subnational Administrative Boundaries - Humanitarian Data Exchange'. 2021. <https://data.humdata.org/dataset/cod-ab-mdg>.
- . 2021b. 'Madagascar - Subnational Population Statistics - Humanitarian Data Exchange'. 2021. <https://data.humdata.org/dataset/cod-ps-mdg>.
- Topp, Christian Winther, Søren Dinesen Østergaard, Susan Søndergaard, and Per Bech. 2015. 'The WHO-5 Well-Being Index: A Systematic Review of the Literature'. *Psychotherapy and Psychosomatics* 84 (3): 167–76. <https://doi.org/10.1159/000376585>.
- UNESCO Institute for Statistics. 2023. 'UIS.Stat Bulk Data Download Service'. 2023. [apiportal.uis.unesco.org/bdds](http://apiportal.uis.unesco.org/bdds).
- United Nations, Department of Economic and Social Affairs, Population Division.

2022. 'World Population Prospects: The 2022 Revision'. 2022. <https://population.un.org/dataportal/data/indicators/67/locations/450/start/1990/end/2023/line/linetimeplot>.
- 'Université de Fianarantsoa'. 2022. <https://www.univ-fianarantsoa.mg/Etablissement/details/1>.
- Urbanek, Simon. 2022. *Jpeg: Read and Write JPEG Images*. <https://CRAN.R-project.org/package=jpeg>.
- Uyoga, Sophie, Ifedayo M. O. Adetifa, Mark Otiende, Christine Yegon, Ambrose Agweyu, George M. Warimwe, and J. Anthony G. Scott. 2021. 'Prevalence of SARS-CoV-2 Antibodies From a National Serosurveillance of Kenyan Blood Donors, January-March 2021'. *JAMA* 326 (14): 1436. <https://doi.org/10.1001/jama.2021.15265>.
- Wagner, Caroline S., Xiaojing Cai, Yi Zhang, and Caroline V. Fry. 2022. 'One-Year in: COVID-19 Research at the International Level in COVID-19 Data'. Edited by Alberto Baccini. *PLOS ONE* 17 (5): e0261624. <https://doi.org/10.1371/journal.pone.0261624>.
- WHO Regional Office for Europe. 1998. 'Use of Well-Being Measures in Primary Health Care - The DepCare Project. Health for All, Target 12, 1998'. <http://www.who.dk/document/e60246.pdf>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. 'Welcome to the Tidyverse'. *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wiens, Kirsten E., Pinyi Nyimol Mawien, John Rumunu, Damien Slater, Forrest K. Jones, Serina Moheed, Andrea Caflisch, et al. 2021. 'Seroprevalence of Severe Acute Respiratory Syndrome Coronavirus 2 IgG in Juba, South Sudan, 2020'. *Emerging Infectious Diseases* 27 (6): 1598–1606. <https://doi.org/10.3201/eid2706.210568>.
- Wilke, Claus O. 2020. *Cowplot: Streamlined Plot Theme and Plot Annotations for 'Ggplot2'*. <https://CRAN.R-project.org/package=cowplot>.
- Wilke, Claus O., and Brenton M. Wiernik. 2022. *Ggtext: Improved Text Rendering Support for 'Ggplot2'*. <https://CRAN.R-project.org/package=ggtext>.
- Workneh, Firehiwot, Dongqing Wang, Ouhohiré Millogo, Alemayehu Worku, Angela Chukwu, Bruno Lankoande, Nega Assefa, et al. 2021. 'Knowledge and Practice Related to COVID-19 and Mental Health among Adults in Sub-Saharan Africa'. *The American Journal of Tropical Medicine and Hygiene* 105 (2): 351–62. <https://doi.org/10.4269/ajtmh.21-0219>.
- World Bank. 2021a. 'GNI per Capita, Atlas Method (Current US\$) - Madagascar'. <https://data.worldbank.org/indicator/NY.GNP.PCAP.CD?locations=MG>.
- . 2021b. 'The World by Income and Region'. 2021. <https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html>.
- . 2021c. 'Urban Population (% of Total Population) - Madagascar'. 2021. <https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=MG>.
- World Food Programme. 2023. 'Southern Madagascar Emergency'. 2023. <https://www.wfp.org/emergencies/southern-madagascar-emergency>.
- World Health Organization. 2020a. 'WHO COVID-19 Dashboard'. 2020. <https://covid19.who.int/>.
- . 2020b. 'WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020'. 11 March 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media>

- briefing-on-covid-19---11-march-2020.
- . 2022a. *World Health Statistics 2022: Monitoring Health for the SDGs, Sustainable Development Goals*. Geneva.
- . 2022b. ‘Mental Health Atlas 2020 Country Profile: Madagascar’. World Health Organization. <https://www.who.int/publications/m/item/mental-health-atlas-mdg-2020-country-profile>.
- . 2023. ‘Statement on the Fourteenth Meeting of the International Health Regulations (2005) Emergency Committee Regarding the Coronavirus Disease (COVID-19) Pandemic’. 30 January 2023. [https://www.who.int/news/item/30-01-2023-statement-on-the-fourteenth-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-pandemic](https://www.who.int/news/item/30-01-2023-statement-on-the-fourteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic).
- Zar, Heather J., Jeanette Dawa, Gilberto Bueno Fischer, and Jose A. Castro-Rodriguez. 2020. ‘Challenges of COVID-19 in Children in Low- and Middle-Income Countries’. *Paediatric Respiratory Reviews* 35 (September): 70–74. <https://doi.org/10.1016/j.prrv.2020.06.016>.

## 10. Acknowledgment

First and foremost, I would like to thank my supervisor, Prof. Jürgen May, for his expertise, feedback, and support in enrolling me as a medical doctoral student at the University of Hamburg and for giving me the opportunity to do a research internship in Madagascar.

Further, I would like to thank Dr. Nicole Gilberger and Dr. Eva Lorenz for their experience and help along the way. This started with identifying my research questions, planning, and finally proofreading my manuscript. I would also like to thank Dr. Daniela Fusco, laboratory group leader at the Bernhard Nocht Institute for Tropical Medicine in Madagascar, for her close communication and coordination during my ten-month stay. Without her support and the expertise of the local Malagasy staff, the research project would not have been possible.

In particular, I would like to thank the Principal Investigator and Dean of the Medical University, Prof. Rivo Rakotoarivelo, as well as the Co-PIs Prof. Raphael Rakotozandriny of the University of Antananarivo, and Dr. Mandranto of the CICM. Thanks, are also due to the Malagasy SeroCoV field teams and laboratory personnel.

I am grateful to the Studienstiftung des deutschen Volkes for the financial support during my research internship.

My special thanks go to my BNITM colleagues in the field, with whom I experienced an inspiring and varied, sometimes challenging, daily life in Madagascar. Their humor and moral support, including numerous games of skat by candlelight during frequent power cuts, were important in keeping my spirits up. Lastly, I would like to thank my family and girlfriend for their support.

## **11. Curriculum vitae**

**Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt.**

## 12. Appendix(ces)

### 12.1 Appendix A: Additional tablets

**Table 3.** Stratification Goals - Fianarantsoa

AGE GROUP	MALE <sup>1</sup>	FEMALE <sup>1</sup>	TOTAL <sup>1</sup>
Adolescents 10 ≤ 19 years	140 /110	138 /110	278 /220
Adults 20 ≤ 44 years	282 /155	301 /155	583 /310
Adults ≥ 45 years	115 /60	145 /60	260 /120
Sum	537 /325	584 /325	1121 /650

<sup>1</sup>Included /Intended number of participants

**Table 4.** Overview Table of Collected Data

Group	Individual N = 1,121	Household N = 674
Demographic		
1	ELISA test result	
2	Age	
3	Sex	
Clinical		
4	Underlying health conditions	Confirmed or suspected cases of COVID-19
5	Past year COVID-19 symptoms	Non-pharmaceutical protection measures
6	Tobacco smoking	
7	Mental health	
8	Daily contacts	
9	Difficulties and compliance to protection measures	
Educational		
10	Missed work or school	Level of Education
11		Main source of information
12		Susceptible for severe COVID-19 infection
Socioeconomic		
13	Profession	DHS Wealth Index
14	Main source of income	Economic consequences
15	Changes in living conditions	

**Table 5.2** ELISA Test Result stratified by Sex and Age Groups

Characteristic	N	FEMALE			MALE					
		positive, N = 224 <sup>1</sup>	95% CI <sup>2</sup>	negative, N = 360 <sup>1</sup>	95% CI <sup>2</sup>	N	positive, N = 196 <sup>1</sup>	95% CI <sup>2</sup>	negative, N = 341 <sup>1</sup>	95% CI <sup>2</sup>
<b>Age</b>	584					537				
Median (IQR)		33 (20, 47)		30 (20, 42)			26 (17, 41)		29 (20, 41)	
<b>Age Group</b>	584					537				
10 ≤ 19		54 (39%)	[31; 48%]	84 (61%)	[52; 69%]		64 (46%)	[37; 54%]	76 (54%)	[46; 63%]
20 ≤ 29		47 (35%)	[27; 43%]	89 (65%)	[57; 73%]		51 (35%)	[27; 43%]	96 (65%)	[57; 73%]
30 ≤ 39		37 (32%)	[24; 41%]	79 (68%)	[59; 76%]		26 (27%)	[19; 37%]	70 (73%)	[63; 81%]
40 ≤ 49		36 (44%)	[34; 56%]	45 (56%)	[44; 66%]		24 (34%)	[24; 47%]	46 (66%)	[53; 76%]
50 ≤ 59		28 (47%)	[34; 60%]	32 (53%)	[40; 66%]		16 (43%)	[28; 60%]	21 (57%)	[40; 72%]
≥ 60		22 (42%)	[28; 56%]	31 (58%)	[44; 72%]		15 (32%)	[20; 47%]	32 (68%)	[53; 80%]

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval



**Table 5.3** Comparison of Educational Degrees

LEVEL OF EDUCATION	MALAGASY	ENGLISH	GERMAN
No degree	1 Pas de diplôme	No degree	Kein Abschluss
Primary	2 Primaire	Primary	Grundschule
Secondary	3 Secondaire	Middle school	Mittlere Schulabschluss
	4 Baccalauréat	High school	Abitur
Higher	5 Étudiant à l'université	University student	Student
	6 Licence	Bachelor's degree	Bachelor
	7 Master	Master	Master
	8 Doctorat	PhD	Doktor
Other	9 Autre	Other	Andere

Reference:

Madagascar - Public expenditure review 2015: education (English). Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/271911468185343596/Madagascar-Public-expenditure-review-2015-education>

**Table 5.4** Profession and Education Levels

CHARACTERISTIC	N	No degree, n = 55 <sup>1</sup>	Primary, n = 159 <sup>1</sup>	Secondary school, n = 343 <sup>1</sup>	Higher, n = 57 <sup>1</sup>	Other, n = 17 <sup>1</sup>
<b>Profession<sup>2</sup></b>	601					
Farming	9 (16%)	15 (9.9%)	12 (3.6%)	0 (0%)	1 (10%)	
Employed	2 (3.6%)	1 (0.7%)	22 (6.6%)	4 (7.5%)	0 (0%)	
Self-employed	31 (56%)	96 (63%)	177 (53%)	7 (13%)	6 (60%)	
Governmental employee	0 (0%)	0 (0%)	6 (1.8%)	7 (13%)	1 (10%)	
Cleaning lady	1 (1.8%)	8 (5.3%)	14 (4.2%)	1 (1.9%)	0 (0%)	
House keeper/ homemaker	7 (13%)	14 (9.2%)	29 (8.8%)	2 (3.8%)	1 (10%)	
Student/ Intern/ Apprentice	0 (0%)	7 (4.6%)	39 (12%)	26 (49%)	1 (10%)	
Retired	4 (7.3%)	2 (1.3%)	14 (4.2%)	3 (5.7%)	0 (0%)	
Unemployed (health reasons)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	
Unemployed (other reasons)	0 (0%)	2 (1.3%)	9 (2.7%)	2 (3.8%)	0 (0%)	
Not applicable (under 18)	1 (1.8%)	5 (3.3%)	7 (2.1%)	0 (0%)	0 (0%)	
Other	0 (0%)	2 (1.3%)	2 (0.6%)	1 (1.9%)	0 (0%)	
Unknown	0	7	12	4	7	

<sup>1</sup>n (%)<sup>2</sup>Only participants older than 14 years were included.**Table 6.2** Treatment of COVID-19 Symptoms

Which places did you seek after developing Covid-19 symptoms in the past year?	N = 545 <sup>1</sup>
<b>Traditional healer</b>	
Mentioned	3 (0.6%)
Not mentioned	542 (99%)
<b>Traditional self-medication</b>	
Mentioned	127 (23%)
Not mentioned	418 (77%)
<b>Modern self-medication</b>	
Mentioned	67 (12%)

Not mentioned	478 (88%)
<b>Private pharmacy</b>	
Mentioned	51 (9.4%)
Not mentioned	494 (91%)
<b>Street pharmacy</b>	
Mentioned	106 (19%)
<b>Not mentioned</b>	439 (81%)
<b>Health care facility</b>	
Mentioned	231 (42%)
Not mentioned	314 (58%)
<b>Other</b>	
Mentioned	12 (2.2%)
Not mentioned	533 (98%)

<sup>1</sup>n (%)

**Table 6.3** Chronic Diseases among Participants

CHARACTERISTIC	N	N = 1,121 <sup>1</sup>
<b>Do you have chronic diseases?</b>	1,118	
Yes		104 (9.3%)
No		1,014 (91%)
Unknown		3
<b>Lung</b>	104	17 (16%)
<b>Diabetes</b>	104	4 (3.8%)
<b>Heart</b>	104	10 (9.6%)
<b>Hypertension</b>	104	51 (49%)
<b>Other</b>	104	25 (24%)

<sup>1</sup>n (%)

**Table 6.4** Tobacco Smoking by Participants Sex

CHARACTERISTIC	N	FEMALE, N = 584 <sup>1</sup>	95% CI <sup>2</sup>	MALE, N = 537 <sup>1</sup>	95% CI <sup>2</sup>
<b>Do you smoke?</b>	1,117				
Yes	177 (16%)	10 (5.6%)	[2.9; 10%]	167 (94%)	[90; 97%]
No	940 (84%)	570 (61%)	[57; 64%]	370 (39%)	[36; 43%]
Unknown	4	4		0	

<sup>1</sup>n (%)

<sup>2</sup>CI = Confidence Interval

**Table 9.2** Trust in Source of Information

Which source of information did you trust the most during the pandemic?	N	No degree, n = 62 <sup>1</sup>	Primary, n = 179 <sup>1</sup>	Secondary school, n = 355 <sup>1</sup>	Higher, n = 58 <sup>1</sup>	Other, n = 17 <sup>1</sup>
	672					
Social media (e.g. Facebook)	19 (2.8%)	0 (0%)	2 (11%)	8 (42%)	9 (47%)	0 (0%)
Webpages (e.g. Google)	1 (0.1%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)
Ministry of Health	3 (0.4%)	0 (0%)	0 (0%)	1 (33%)	2 (67%)	0 (0%)
Radio	205 (31%)	25 (12%)	69 (34%)	93 (46%)	8 (3.9%)	9 (4.4%)
TV	372 (55%)	29 (7.8%)	83 (22%)	224 (60%)	32 (8.6%)	3 (0.8%)
Newspaper	7 (1.0%)	0 (0%)	0 (0%)	7 (100%)	0 (0%)	0 (0%)
Health workers	20 (3.0%)	3 (16%)	5 (26%)	5 (26%)	4 (21%)	2 (11%)
Family and friends	45 (6.7%)	5 (11%)	20 (44%)	16 (36%)	1 (2.2%)	3 (6.7%)
Other	0 (0%)	0 (NA%)	0 (NA%)	0 (NA%)	0 (NA%)	0 (NA%)
Unknown	2	0	0	1	1	0

<sup>1</sup>n (%)**Table 10.2** Non-pharmaceutical Protection Measures and ELISA Test Result

Characteristic	N	ELISA TEST RESULT			
		positive, N = 240 <sup>1</sup>	95% CI <sup>2</sup>	negative, N = 434 <sup>1</sup>	95% CI <sup>2</sup>
<b>Did you take measures to protect household members?</b>	672				
Yes	655 (97%)	232 (35%)	[32; 39%]	423 (65%)	[61; 68%]
No	17 (2.5%)	6 (35%)	[15; 61%]	11 (65%)	[39; 85%]
Unknown	2	2		0	
<b>Masks for every household member</b>	674				
Mentioned	634 (94%)	222 (35%)	[31; 39%]	412 (65%)	[61; 69%]
Not mentioned	40 (5.9%)	18 (45%)	[30; 61%]	22 (55%)	[39; 70%]
<b>Masks for household member with symptoms</b>	674				
Mentioned	6 (0.9%)	3 (50%)	[19; 81%]	3 (50%)	[19; 81%]
Not mentioned	668 (99%)	237 (35%)	[32; 39%]	431 (65%)	[61; 68%]
<b>Masks for vulnerable household members</b>	674				
Not mentioned	674 (100%)	240 (36%)	[32; 39%]	434 (64%)	[61; 68%]
<b>Not taking meal together</b>	674				
Mentioned	12 (1.8%)	7 (58%)	[29; 84%]	5 (42%)	[16; 71%]
Not mentioned	662 (98%)	233 (35%)	[32; 39%]	429 (65%)	[61; 68%]
<b>Not sharing personal objects</b>	674				
Mentioned	1 (0.1%)	0 (0%)	[0.00; 95%]	1 (100%)	[5.5; 100%]
Not mentioned	673 (100%)	240 (36%)	[32; 39%]	433 (64%)	[61; 68%]
<b>Washing hands more regularly</b>	674				
Mentioned	595 (88%)	212 (36%)	[32; 40%]	383 (64%)	[60; 68%]
Not mentioned	79 (12%)	28 (35%)	[25; 47%]	51 (65%)	[53; 75%]
<b>Avoiding contacts with other household members</b>	674				
Mentioned	121 (18%)	53 (44%)	[35; 53%]	68 (56%)	[47; 65%]
Not mentioned	553 (82%)	187 (34%)	[30; 38%]	366 (66%)	[62; 70%]
<b>Avoiding market and other crowded place</b>	674				
Mentioned	5 (0.7%)	1 (20%)	[1.1; 70%]	4 (80%)	[30; 99%]
Not mentioned	669 (99%)	239 (36%)	[32; 40%]	430 (64%)	[60; 68%]
<b>Cleaning household's objects more regularly</b>	674				
Mentioned	15 (2.2%)	7 (47%)	[22; 73%]	8 (53%)	[27; 78%]
Not mentioned	659 (98%)	233 (35%)	[32; 39%]	426 (65%)	[61; 68%]

<sup>1</sup>n (%)<sup>2</sup>CI = Confidence Interval

**Table 10.3** Difficulties and Compliance to COVID-19 Measures by Age Groups

Characteristic	N	AGE GROUP IN YEARS					
		10 ≤ 19, N = 138 <sup>1</sup>	20 ≤ 29, N = 140 <sup>1</sup>	30 ≤ 39, N = 301 <sup>1</sup>	40 ≤ 49, N = 282 <sup>1</sup>	50 ≤ 59, N = 145 <sup>1</sup>	≥ 60, N = 115 <sup>1</sup>
<b>Which measure was the most difficult to comply with?</b>	784						
Maintain social distance	434 (55%)	60 (66%)	58 (55%)	117 (55%)	109 (57%)	50 (48%)	40 (49%)
Wear face masks	228 (29%)	20 (22%)	37 (35%)	62 (29%)	48 (25%)	32 (31%)	29 (36%)
Frequently wash hands	43 (5.5%)	3 (3.3%)	4 (3.8%)	8 (3.8%)	12 (6.3%)	8 (7.7%)	8 (9.9%)
Socially isolate	38 (4.8%)	4 (4.4%)	5 (4.7%)	9 (4.3%)	10 (5.2%)	8 (7.7%)	2 (2.5%)
Not work	33 (4.2%)	2 (2.2%)	2 (1.9%)	13 (6.2%)	9 (4.7%)	5 (4.8%)	2 (2.5%)
Quarantine	4 (0.5%)	1 (1.1%)	0 (0%)	1 (0.5%)	1 (0.5%)	1 (1.0%)	0 (0%)
Other	4 (0.5%)	1 (1.1%)	0 (0%)	1 (0.5%)	2 (1.0%)	0 (0%)	0 (0%)
Unknown	337	47	34	90	91	41	34
<b>Do people in your immediate environment comply with the advised measures?</b>	1,111						
Yes	843 (76%)	108 (79%)	96 (70%)	229 (76%)	204 (73%)	113 (79%)	93 (81%)
No	261 (23%)	25 (18%)	40 (29%)	70 (23%)	75 (27%)	30 (21%)	21 (18%)
Do not know	7 (0.6%)	4 (2.9%)	1 (0.7%)	1 (0.3%)	0 (0%)	0 (0%)	1 (0.9%)
Unknown	10	1	3	1	3	2	0
<b>Do you think that people should comply with the advised measures?</b>	1,115						
Yes	1,069 (96%)	127 (93%)	120 (86%)	296 (99%)	274 (98%)	139 (97%)	113 (98%)
No	15 (1.3%)	1 (0.7%)	4 (2.9%)	2 (0.7%)	4 (1.4%)	2 (1.4%)	2 (1.7%)
Do not know	31 (2.8%)	9 (6.6%)	16 (11%)	1 (0.3%)	2 (0.7%)	3 (2.1%)	0 (0%)
Unknown	6	1	0	2	2	1	0

<sup>1</sup>n (%)

**Table 11.** Vulnerable persons by Level of Education

Who do you think is most vulnerable when it comes to getting sick with COVID-19?	N	No degree, n = 62 <sup>1</sup>	Primary, n = 179 <sup>1</sup>	Secondary school, n = 355 <sup>1</sup>	Higher, n = 58 <sup>1</sup>	Other, n = 17 <sup>1</sup>
<b>Men</b>	671					
Mentioned	75 (11%)	10 (16%)	18 (10%)	40 (11%)	5 (8.6%)	2 (12%)
Not mentioned	596 (89%)	52 (84%)	161 (90%)	315 (89%)	53 (91%)	15 (88%)
<b>Women</b>	671					
Mentioned	69 (10%)	11 (18%)	17 (9.5%)	36 (10%)	5 (8.6%)	0 (0%)
Not mentioned	602 (90%)	51 (82%)	162 (91%)	319 (90%)	53 (91%)	17 (100%)
<b>Elderly people</b>	671					
Mentioned	528 (79%)	48 (77%)	148 (83%)	283 (80%)	42 (72%)	7 (41%)
Not mentioned	143 (21%)	14 (23%)	31 (17%)	72 (20%)	16 (28%)	10 (59%)
<b>Children</b>	671					
Mentioned	84 (13%)	11 (18%)	21 (12%)	40 (11%)	9 (16%)	3 (18%)
Not mentioned	587 (87%)	51 (82%)	158 (88%)	315 (89%)	49 (84%)	14 (82%)
<b>Pregnant women</b>	671					
Mentioned	192 (29%)	17 (27%)	46 (26%)	107 (30%)	20 (34%)	2 (12%)

Not mentioned	479 (71%)	45 (73%)	133 (74%)	248 (70%)	38 (66%)	15 (88%)
<b>Diabetic</b>	671					
Mentioned	421 (63%)	44 (71%)	103 (58%)	230 (65%)	37 (64%)	7 (41%)
Not mentioned	250 (37%)	18 (29%)	76 (42%)	125 (35%)	21 (36%)	10 (59%)
<b>Cardiac</b>	671					
Mentioned	316 (47%)	43 (69%)	76 (42%)	163 (46%)	28 (48%)	6 (35%)
Not mentioned	355 (53%)	19 (31%)	103 (58%)	192 (54%)	30 (52%)	11 (65%)
<b>Pulmonary disease</b>	671					
Mentioned	302 (45%)	42 (68%)	74 (41%)	152 (43%)	28 (48%)	6 (35%)
Not mentioned	369 (55%)	20 (32%)	105 (59%)	203 (57%)	30 (52%)	11 (65%)
<b>Immunosuppressed people</b>	671					
Mentioned	312 (46%)	42 (68%)	72 (40%)	162 (46%)	31 (53%)	5 (29%)
Not mentioned	359 (54%)	20 (32%)	107 (60%)	193 (54%)	27 (47%)	12 (71%)
<b>Do not know</b>	671					
Mentioned	22 (3.3%)	2 (3.2%)	4 (2.2%)	6 (1.7%)	5 (8.6%)	5 (29%)
Not mentioned	649 (97%)	60 (97%)	175 (98%)	349 (98%)	53 (91%)	12 (71%)
<b>Other</b>	671					
Mentioned	5 (0.7%)	0 (0%)	3 (1.7%)	1 (0.3%)	1 (1.7%)	0 (0%)
Not mentioned	666 (99%)	62 (100%)	176 (98%)	354 (100%)	57 (98%)	17 (100%)

<sup>l</sup>n (%)

**Table 16.** Possession of Wealth Index Assets among Included Households

Wealth Index Assets	WEALTH INDEX QUINTILES					
	Improved household conditions, N = 674 <sup>l</sup>	lowest, N = 53 <sup>l</sup>	second, N = 106 <sup>l</sup>	middle, N = 126 <sup>l</sup>	fourth, N = 218 <sup>l</sup>	highest, N = 171 <sup>l</sup>
Type of housing	128 (19%)	15 (28%)	22 (21%)	24 (19%)	50 (23%)	17 (9.9%)
Nature of walls	584 (87%)	31 (58%)	76 (72%)	113 (90%)	203 (93%)	161 (94%)
Nature of floor	481 (71%)	14 (26%)	63 (59%)	97 (77%)	153 (70%)	154 (90%)
Type of toilets	53 (7.9%)	0 (0%)	3 (2.8%)	14 (11%)	25 (11%)	11 (6.4%)
Main source of lighting energy	522 (77%)	0 (0%)	22 (21%)	113 (90%)	216 (99%)	171 (100%)
Source of water in rainy season	519 (77%)	30 (57%)	60 (57%)	106 (84%)	189 (87%)	134 (78%)
Source of water in dry season	471 (70%)	28 (53%)	54 (51%)	99 (79%)	166 (76%)	124 (73%)
Radio	494 (73%)	8 (15%)	48 (45%)	86 (68%)	183 (84%)	169 (99%)
TV	404 (60%)	0 (0%)	1 (0.9%)	22 (17%)	210 (96%)	171 (100%)
Video/CD/DVD Player	227 (34%)	0 (0%)	0 (0%)	1 (0.8%)	71 (33%)	155 (91%)
Satellite decoder	111 (16%)	0 (0%)	0 (0%)	10 (7.9%)	54 (25%)	47 (27%)
Cell phone	578 (86%)	40 (75%)	86 (81%)	102 (81%)	192 (88%)	158 (92%)
Smartphone	187 (28%)	1 (1.9%)	6 (5.7%)	31 (25%)	76 (35%)	73 (43%)
Fridge/ Freezer	43 (6.4%)	0 (0%)	1 (0.9%)	3 (2.4%)	28 (13%)	11 (6.4%)
Solar lamp	35 (5.2%)	2 (3.8%)	6 (5.7%)	17 (13%)	7 (3.2%)	3 (1.8%)
Oil lamp	98 (15%)	46 (87%)	47 (44%)	5 (4.0%)	0 (0%)	0 (0%)
Bicycle	37 (5.5%)	1 (1.9%)	1 (0.9%)	6 (4.8%)	21 (9.6%)	8 (4.7%)
Laptop	73 (11%)	0 (0%)	4 (3.8%)	14 (11%)	35 (16%)	20 (12%)
Bed nets	475 (70%)	26 (49%)	64 (60%)	80 (63%)	164 (75%)	141 (82%)

<sup>l</sup>n (%)



## 12.3 Appendix C: Code

```
## Set up

## Install R® packages

install.packages(c("tidyverse", "here", "jpeg", "magick", "apyrmaid", "ggcorrplot", "cowplot", "DiagrammeR",
"ggtext", "ggrepel", "ggmap", "sf", "ggsn", "psych", "gtsummary", "gt"),
  repos = "http://cran.us.r-project.org")

## Load R® packages

library("tidyverse")
library("here")
library("haven")
library("readxl")
library("jpeg")
library("magick")
library("apyrmaid")
library("ggcorrplot")
library("cowplot")
library("DiagrammeR")
library("ggtext")
library("ggrepel")
library("ggmap")
library("sf")
library("ggsn")
library("psych")
library("gtsummary")
library("gt")

## Set file path

here()

## Methods
```

---

```
## Figure 1 -- Overview Map of Madagascar

## Map of Madagascar
map_madagascar_intro_overview <- get_stamenmap(
  bbox = c(left = 38.856, bottom = -26.687, right = 54.0, top = -10.833),
  maptype = "terrain-background",
  zoom = 6
)

## African rivers data set from hydroviz® (Source) and tutorial
shp_african_rivers <-
st_read("HydroRIVERS_v10_af_shp/HydroRIVERS_v10_af_shp/HydroRIVERS_v10_af.shp")

## Select the big rivers
big_rivers_africa <- shp_african_rivers %>%
  filter(ORD_FLOW < 6) %>%
  mutate(width = as.numeric(ORD_FLOW),
    width = case_when(
      width %in% c(2) ~ "1",
      width %in% c(3) ~ "0.8",
      width %in% c(4) ~ "0.6",
      width %in% c(5) ~ "0.4",
      TRUE ~ "0"
    )) %>%
  st_as_sf()
```

```

## Coordinates of cities, rivers and the région
df_location_fianar <- tribble(~ place, ~ longitude, ~ latitude,
  "Fianarantsoa", 47.0996, -21.4417)

df_location_tana <- tribble(~ place, ~ longitude, ~ latitude,
  "Antananarivo", 47.5282, -18.9148)

df_location_matsiatra_river <- tribble(~ place, ~ longitude, ~ latitude,
  "Matsiatra", 46.5284, -21.0567)

df_location_mangoky_river <- tribble(~ place, ~ longitude, ~ latitude,
  "Matsiatra", 45.3280, -21.6481)

df_mozambique_channel <- tribble(~ place, ~ longitude, ~ latitude,
  "Mozambique Channel", 41.1704, -18.6198)

df_indian_ocean <- tribble(~ place, ~ longitude, ~ latitude,
  "Indian Ocean", 50.7209, -19.9741)

sf_mada_regions <- st_read("mdg_adm_bngrc_ocha_20181031_shp")

sf_haute_matsiatra <- sf_mada_regions %>%
  filter(str_detect(ADM1_EN, "Haute Matsiatra"))

## Create the boundary box of Madagascar
crsLONGLAT <- "+proj=longlat +datum=WGS84 +no_defs"

bbox_mada <- st_sfc(st_polygon(list(cbind(
  c(38.856, 54.0, 54.0, 38.856, 38.856), # min and max longitude
  c(-10.833, -10.8333, -26.687, -26.687, -10.833) # min and max latitude
))),
  crs = crsLONGLAT
) %>%
  st_transform(4326) %>%
  st_bbox()

## Raw map
sf_intro_madagascar_1 <- ggmap(map_madagascar_intro_overview) +
  coord_cartesian() +
  north(x.min = 51.5, x.max = 52.5,
    y.min = -11.5, y.max = -13, scale = 1
  ) +
  scalebar(x.min = 49, x.max = 52,
    y.min = -24.0, y.max = -25,
    transform = TRUE, dist_unit = "km",
    dist = 200, model = "WGS84", height = 0.3,
    st.dist = 0.3, st.size = 3) +
  geom_sf(
    data = big_rivers_africa,
    aes(
      color = factor(ORD_FLOW), size = width,
      alpha = factor(ORD_FLOW)
    ),
    inherit.aes = FALSE) +
  coord_sf(
    crs = 4326,
    xlim = c(bbox_mada["xmin"], bbox_mada["xmax"]),
    ylim = c(bbox_mada["ymin"], bbox_mada["ymax"])
  ) +
  scale_color_manual(
    name = "",
    values = c(
      "#08306b", "#08519c", "#2171b5",
      "#6baed6", "orange"
    )
  ) +
  geom_sf(sf_haute_matsiatra,
    mapping = aes(fill = "orange",
      geometry = geometry),

```



```

    inherit.aes = FALSE,
    linewidth = 0.1,
    alpha = 0.4) +
geom_point(data = df_location_fianar,
  aes(x = longitude,
      y = latitude),
  color = "purple",
  size = 2) +
geom_point(data = df_location_tana,
  aes(x = longitude,
      y = latitude),
  color = "red",
  size = 2) +
geom_point(data = df_location_matsiatra_river,
  aes(x = longitude,
      y = latitude),
  color = "black",
  size = 2,
  shape = 1) +
geom_point(data = df_location_mangoky_river,
  aes(x = longitude,
      y = latitude),
  color = "black",
  size = 2,
  shape = 5)

## Additional details
sf_intro_madagascar_2 <- sf_intro_madagascar_1 +
  geom_text(df_mozambique_channel,
    mapping = aes(x = longitude,
                  y = latitude,
                  label = place),
    size = 4,
    hjust = -0.1,
    vjust = -0.2,
    angle = 55,
    color = "darkblue") +
  geom_text(df_indian_ocean,
    mapping = aes(x = longitude,
                  y = latitude,
                  label = place),
    size = 4,
    hjust = 0.3,
    angle = 0,
    color = "darkblue") +
labs(x = "Longitude",
     y = "Latitude") +
theme(
  text = element_text(family = "Times New Roman"),
  axis.line = element_line(),
  legend.position = "none")

ggsave(filename = "sf_intro_madagascar_2.png", sf_intro_madagascar_2)

```

---

## ## Figure 2 -- Fianarantsoa: Population Size per Fokontany

```

## Map of Fianarantsoa
fokontanys_ <- get_stamenmap(
  bbox = c(left = 47.0352, bottom = -21.51400, right = 47.13787, top = -21.39266),
  maptype = "terrain",
  zoom = 13)

## Final biweekly excel report from Dr. Tiana

df_rapport_hebdomadaire_final_modified <- read_excel("20210618_SeroCoV_REC-MeC-03_Rapport
hebdomadaire_v1.3 Final_modified_2.xlsx",
  sheet = "fokontany")

```

```

df_rapport_hebdomadaire_final_modified_v2 <- df_rapport_hebdomadaire_final_modified %>%
select(everything()) %>%
mutate(ADM3_EN = case_when(
  Fokotany %in% c("Idanda", "Beravina", "Antanifotsy V", "Tsimanarirazana", "Ambalabe")
  ~ "Andrainjato Avaratra",
  Fokotany %in% c("Soatsihadino", "Igaga")
  ~ "Andrainjato Sud",
  `Screening IDs` %in% c("127 - 155, 734 - 764") & Fokotany %in% c("Ambodiharana")
  ~ "Tanana Ambony",
  Fokotany %in% c("Ambatofolaka", "Mandriandalana", "Ambodiharana", "Mahazengy")
  ~ "Lalazana",
  Fokotany %in% c("Andriamboasary", "Ankofafalahy Ambany")
  ~ "Manolafaka",
  Fokotany %in% c("Ankofafa Ambony", "Sahalava", "Talatamaty", "Antsororokavo",
    "Mokana", "Antarandolo", "Tambohobe", "Ambatomena", "Ivory",
    "Tanambao", "Isada")
  ~ "Tanana Ambany",
  Fokotany %in% c("Ambatolahikisoa", "Ambalandapa", "Ambalapaiso", "Isaha",
    "Ampitakely", "Ambatovory", "Tsaramandroso", "Ambodiharana", "Rova")
  ~ "Tanana Ambony",
  Fokotany %in% c("Ankazobe")
  ~ "Vatosola")) %>%
mutate(ADM4_EN = Fokotany) %>%
filter(if_any(everything(), ~ !is.na(.))) %>%
filter(if_any(Fokotany, ~ !is.na(.))) %>%
filter(!c(ADM3_EN %in% c("Lalazana") & ADM4_EN %in% c("Ambodiharana"))) %>%
mutate(Screening_IDs_v2 = `Screening IDs`,
  number_of_GPS_coordinates_provided_with_backup = `nombre de coordonnées GPS fournies \r\n(30%
backup)`,
  number_of_households_to_be_included = `nombre de ménages à inclure`,
  number_of_households_included = `nombre de \r\nménages inclus`,
  entire_households_sampled = `nombre de ménage entier`,
  number_of_refusing_households = `nombre de \r\nménages refusant`,
  GPS_without_households = `GPS sans ménage`,
  number_of_households_included_v2 = case_when(
    number_of_households_included %in% c("12") ~ "12",
    number_of_households_included %in% c("24") ~ "24",
    number_of_households_included %in% c("26") ~ "26",
    number_of_households_included %in% c("36") ~ "36",
    number_of_households_included %in% c("48") ~ "48",
    number_of_households_included %in% c("49") ~ "49"),
  entire_households_sampled_v2 = case_when(
    entire_households_sampled %in% c("0") ~ "0",
    entire_households_sampled %in% c("1") ~ "1",
    entire_households_sampled %in% c("2") ~ "2",
    entire_households_sampled %in% c("3") ~ "3",
    entire_households_sampled %in% c("4") ~ "4",
    entire_households_sampled %in% c("5") ~ "5",
    entire_households_sampled %in% c("7") ~ "7",
    entire_households_sampled %in% c("8") ~ "8"),
  number_of_refusing_households_v2 = case_when(
    number_of_refusing_households %in% c("0") ~ "0",
    number_of_refusing_households %in% c("1") ~ "1",
    number_of_refusing_households %in% c("2") ~ "2",
    number_of_refusing_households %in% c("3") ~ "3",
    number_of_refusing_households %in% c("4") ~ "4",
    number_of_refusing_households %in% c("7") ~ "7"),
  GPS_without_households_v2 = case_when(
    GPS_without_households %in% c("0") ~ "0",
    GPS_without_households %in% c("1") ~ "1",
    GPS_without_households %in% c("2") ~ "2",
    GPS_without_households %in% c("3") ~ "3",
    GPS_without_households %in% c("5") ~ "5",
    GPS_without_households %in% c("9") ~ "9"))

## Official boundaries
sf_OCHA_Fokontany_adm4 <-
st_read("mdg_adm_bngrc_ocha_20181031_shp/mdg_admbnda_adm4_BNGRC_OCHA_20181031_shp")

```

```

sf_fokontany_Fianarantsoa_I_boundaries_v2 <- sf_OCHA_Fokontany_adm4 %>%
  select(., ADM2_EN, ADM3_EN, ADM4_EN, geometry) %>%
  filter(., ADM2_EN %in% c("Fianarantsoa I"))

sf_merged_final_rapport <- right_join(df_rapport_hebdomadaire_final_modified_v2,
sf_fokontany_Fianarantsoa_I_boundaries_v2)

## Official population size per Fokontany

df_population_ocha_fokontanys <- read_excel("mdg_admpop_2018_2.xls", sheet = "mdg_pop_adm4")

## Create district of Fianarantsoa subset:

df_population_ocha_fokontanys_fianar_I <- df_population_ocha_fokontanys %>%
  select(., ADM2_EN, ADM3_EN, ADM4_EN, T_TL) %>%
  filter(ADM2_EN %in% c("Fianarantsoa I"))

## Merge Fokontany boundaries with population
merged_df_fianar_I <- left_join(sf_fokontany_Fianarantsoa_I_boundaries_v2,
df_population_ocha_fokontanys_fianar_I)

merged_df_fianar_I_modified <- merged_df_fianar_I %>%
  select(everything()) %>%
  mutate(population_size = cut(T_TL,
    breaks = c(1000,
      2000,
      5000,
      7500,
      10000,
      15000,
      Inf),
    labels = fct_rev(c("1000 - 1999",
      "2000 - 4999",
      "5000 - 7499",
      "7500 - 9999",
      "10.000 - 15.000",
      "> 15.000"))))

gg_population_per_fokontany <- ggmap(fokontanys_) +
  geom_sf(merged_df_fianar_I_modified,
    mapping = aes(fill = population_size),
    inherit.aes = FALSE,
    color = "black",
    linetype = 1,
    lwd = 0.25,
    alpha = 0.6) +
  scale_fill_manual(values = c("1000 - 1999" = "#999999",
    "2000 - 4999" = "#56B4E9",
    "5000 - 7499" = "#0072B2",
    "7500 - 9999" = "#F0E442",
    "10.000 - 15.000" = "#E69F00",
    "> 15.000" = "#D55E00"),
    name = "Population size") +
  geom_sf_text(data = merged_df_fianar_I_modified,
    mapping = aes(label = ADM4_EN),
    inherit.aes = FALSE,
    show.legend = FALSE,
    size = 1.5) +
  theme(
    text = element_text(family = "Times New Roman"),
    legend.title = element_textbox(size= 7),
    legend.text= element_text(size= 7),
    legend.key.size = unit(0.3, 'cm'),
    legend.key.height = unit(0.3, 'cm'),
    axis.line = element_line()
  ) +

```

```
labs(
  x = "Longitude",
  y = "Latitude")
```

---

### ## Table 1 -- Stratification Goals - Fianarantsoa

```
stratification_goals_origin <- tribble(
  ~ age_group, ~ male_1, ~ female_1, ~ total_1,
  "adolescents_10_19", 110, 110, 220,
  "adults_20_44", 155, 155, 310,
  "adults_old_45", 60, 60, 120,
  "Sum", 325, 325, 650) %>%
mutate(age_group_2 = case_when(age_group %in% c("adolescents_10_19") ~ "Adolescents 10 ≤ 19 years",
  age_group %in% c("adults_20_44") ~ "Adults 20 ≤ 44 years",
  age_group %in% c("adults_old_45") ~ "Adults ≥ 45 years",
  age_group %in% c("Sum") ~ "Sum")) %>%
select(age_group_2, male_1, female_1, total_1)

gt_stratification_goals_origin <- stratification_goals_origin %>%
gt(rowname_col = "age_group_2") %>%
tab_stubhead(label = md("**Age Group**")) %>%
cols_label(male_1 = md("**Male**"),
  female_1 = md("**Female**"),
  total_1 = md("**Total**")) %>%
tab_caption(caption = md("**Table 1.** Stratification Goals - Fianarantsoa")) %>%
opt_table_font(
  font = "Times New Roman") %>%
tab_footnote(
  footnote = "Intended number of participants",
  locations = cells_column_labels(
    columns = c(male_1, female_1, total_1)))

gt_stratification_goals_origin %>%
gtsave(filename = "gt_stratification_goals_origin.docx")
```

---

### ## Figure 3 -- SeroCoV Staff Preparing for a Daily Field Visit

```
df_picture_jeep_raw <- tibble(x = 1:10,
  y = 1:10)

gg_picture_jeep_raw <- ggplot(df_picture_jeep_raw,
  aes(x, y)) +
  theme(
    text = element_text(family = "Times New Roman")
  ) +
  labs(x = NULL,
  y = NULL)

## Download jeep image

jeep_image <- readJPEG("IMG_7574.jpeg",
  native = TRUE)

gg_picture_jeep_better <- gg_picture_jeep_raw +
  inset_element(p = jeep_image,
    top = 1,
    right = 1,
    bottom = 0,
    left = 0,
    align_to = "plot")

ggsave("gg_picture_jeep_better.png", gg_picture_jeep_better)
```

---

### ## Figure 4 -- Household Identification in Fianarantsoa

```

df_picture_get_around_raw <- tibble(x = 1:10,
                                     y = 1:10)

gg_picture_get_around_raw <- ggplot(df_picture_get_around_raw,
                                   aes(x, y)) +
  theme(
    text = element_text(family = "Times New Roman")
  ) +
  labs(x = NULL,
       y = NULL)

## Download picture of the walking SeroCoV team

get_around_image <- readJPEG("/Users/leonardgunga/Desktop/Fotos/Madagascar/SeroCov/IMG_7563.jpeg",
                             native = TRUE)

gg_picture_get_around_better <- gg_picture_get_around_raw +
  inset_element(p = get_around_image,
               top = 1,
               right = 1,
               bottom = 0,
               left = 0,
               align_to = "plot")

ggsave("gg_picture_get_around_better.png", gg_picture_get_around_better)

```

---

#### ## Figure 5 -- Data collection in the Field

```

df_picture_tablet_raw <- tibble(x = 1:10,
                                y = 1:10)

gg_picture_tablet_raw <- ggplot(df_picture_tablet_raw,
                               aes(x, y)) +
  theme(
    text = element_text(family = "Times New Roman")
  ) +
  labs(x = NULL,
       y = NULL,)

## Download picture of study nurse with tablet

tablet_image <- readJPEG("/Users/leonardgunga/Desktop/Fotos/Madagascar/SeroCov/IMG_7565.jpeg",
                         native = TRUE)

gg_picture_tablet_better <- gg_picture_tablet_raw +
  inset_element(p = tablet_image,
               top = 1,
               right = 1,
               bottom = 0,
               left = 0,
               align_to = "plot")

ggsave("gg_picture_tablet_better.png", gg_picture_tablet_better)

```

---

#### ## Figure 6 -- Spatial Distribution of Included Fokontanys

```

sf_OCHA_Fokontany_adm4 <-
st_read("mdg_adm_bngrc_ocha_20181031_shp/mdg_admbnda_adm4_BNGRC_OCHA_20181031_shp")

sf_fokontany_Fianarantsoa_I_boundaries <- sf_OCHA_Fokontany_adm4 %>%
  filter(., ADM2_EN %in% c("Fianarantsoa I"))

## Get the boundary box (BBox) of Fokontanys
bbox_sf_fokontany_Fianarantsoa_I_boundaries <- sf_fokontany_Fianarantsoa_I_boundaries %>%
  filter(., ADM2_EN %in% c("Fianarantsoa I")) %>%
  st_bbox()

```

```

## Create a subset of Fokontany inclusion status
df_fokontanys <- sf_fokontany_Fianarantsoa_I_boundaries %>%
  select(ADM2_EN, ADM3_EN, ADM4_EN, geometry) %>%
  mutate(Fokontany_status = case_when(
    ADM4_EN %in% c("Ambahisamotra", "Ambalamarina Kianjasoa", "Ambalambositra", "Ambalavato",
      "Ambatoharanana", "Ambatolahy II", "Ambatolahy V", "Ambatomainty",
      "Ambodikavola", "Amontana", "Anasana", "Andohanatady",
      "Anjaninoro", "Antsaharoa", "Maromby", "Sahamavo")
      ~ "excluded",
    ADM3_EN %in% c("Lalazana") & ADM4_EN %in% c("Ambodiharana")
      ~ "excluded",
    ADM4_EN %in% c("Ambalabe", "Tsimanarirazana", "Beravina", "Antanifotsy V",
      "Idanda", "Soatsihadino", "Igaga", "Ambatofolaka", "Mandriandalana",
      "Mahazengy", "Andriamboasary", "Ankofafalahy Ambany",
      "Ankofafa Ambony", "Sahalava", "Talatamaty", "Antsororokavo", "Mokana",
      "Antarandolo", "Tambohobe", "Ambatomena", "Ivory", "Tanambao", "Isada",
      "Ambatolahikisoa", "Ambalandapa", "Ambalapaïso", "Isaha", "Ampitakely",
      "Ambatovory", "Tsaramandroso", "Ambodiharana", "Rova", "Ankazobe")
      ~ "included"))

## Number of excluded and included Fokontanys
number_ex_in_fokontanys <- df_fokontanys %>%
  group_by(Fokontany_status) %>%
  summarise(number = n())

## Map the data set
gg_in_ex_fokontanys <- ggmap(fokontanys_) +
  geom_sf(df_fokontanys, mapping = aes(fill = Fokontany_status),
    inherit.aes = FALSE,
    show.legend = FALSE,
    color = "black",
    linetype = 1,
    lwd = 0.25,
    alpha = 0.4) +
  scale_fill_manual(values = c("included" = "green",
    "excluded" = "red")) +
  geom_sf_text(data = sf_fokontany_Fianarantsoa_I_boundaries, mapping = aes(label = ADM4_EN),
    inherit.aes = FALSE,
    show.legend = FALSE,
    size = 1.5) +
  coord_sf() +
  theme(
    text = element_text(family = "Times New Roman"),
    legend.title = element_textbox(size = 7),
    legend.text = element_text(size = 7),
    legend.key.size = unit(0.3, 'cm'),
    legend.key.height = unit(0.3, 'cm'),
    axis.line = element_line()
  ) +
  labs(
    x = "Longitude",
    y = "Latitude")

ggsave("gg_in_ex_fokontanys.png", gg_in_ex_fokontanys)

```

---

### ## Figure 7 -- Flowchart of Household Identification with GPS Coordinates

```

flowchart_hh_identification <- mermaid(diagram = "
  graph TD
    A[GPS coordinates] --> B[Singe household]
    A --> C[No household]
    A --> D[Building with multiple households]

    B --> E[Household responsible <br> present]
    B --> F[No household <br> responsible present]
    C --> G[Identification of closest household <br> within 100 m]

```

D --> H[Recruitment of the household at first door <br> on the right side turning clockwise on <br> the ground floor]

E --> I[Recruitment]

E --> J[Recruitment denied]

F --> K[Return later or <br> next day]

G --> L[No household identifiable]

G --> M[Multiple households]

H --> N[No household identifiable]

J --> O[Proceed to the next <br> GPS coordinates]

L --> T[Proceed to the next <br> GPS coordinates]

M --> P[Select the first household on <br> the right after turning north <br> and 45 degrees clockwise]

N --> Q[Proceed to the next floor]

P --> U[No household identifiable]

Q --> R[No household identifiable]

U --> V[Proceed to the next floor]

R --> S[Proceed to the next <br> GPS coordinates]

V --> W[Proceed to the next <br> GPS coordinates]

)

```
df_figure_flowchart_raw <- tibble(x = 1:10,  
  y = 1:10)
```

```
gg_figure_flowchart_raw <- ggplot(df_figure_flowchart_raw,  
  aes(x, y)) +  
  theme(  
    text = element_text(family = "Times New Roman")  
  ) +  
  labs(x = NULL,  
    y = NULL)
```

## Take a manual screenshot of the flowchart and then proceed with the code below

```
magick_flowchart_hh_household_identification <-  
image_read("/Users/leonardgunga/Desktop/R_data/SeroCoV/sop_hh_screenshot.png")
```

```
trimmed_magick_flowchart_hh_household_identification <-  
image_trim(magick_flowchart_hh_household_identification)
```

```
scaled_trimmed_magick_flowchart_hh_household_identification <-  
image_scale(trimmed_magick_flowchart_hh_household_identification,  
  "1200")
```

```
gg_figure_flowchart_better <- gg_figure_flowchart_raw +  
  theme(  
    panel.background = element_blank()  
  ) +  
  draw_image(scaled_trimmed_magick_flowchart_hh_household_identification)
```

```
ggsave("gg_figure_flowchart_better.png", gg_figure_flowchart_better, width = 15, height = 10, dpi = 300)
```

---

## ## Figure 8 - Venous Sampling Preparation

```
df_picture_sampling_raw <- tibble(x = 1:10,  
  y = 1:10)
```

```
gg_picture_sampling_raw <- ggplot(df_picture_sampling_raw,  
  aes(x, y)) +  
  theme(  
    text = element_text(family = "Times New Roman")
```

```

) +
labs(x = NULL,
      y = NULL)

sampling_image <- readJPEG("/Users/leonardgunga/Desktop/Fotos/Madagaskar/SeroCov/IMG_7567.jpeg",
                           native = TRUE)

gg_picture_sampling_better <- gg_picture_sampling_raw +
  inset_element(p = sampling_image,
               top = 1,
               right = 1,
               bottom = 0,
               left = 0,
               align_to = "plot")

ggsave("gg_picture_sampling_better.png", gg_picture_sampling_better)

```

---

## ## Figure 9 -- SeroCoV Samples Transport

```

## Map of Madagascar
madagascar_sample_logistics <- get_stamenmap(
  bbox = c(left = 34.871, bottom = -26.647, right = 58.821, top = -10.790),
  mapttype = "terrain",
  zoom = 6)

## Storage places and routes:
storage_sokela <- tribble(~ place, ~ longitude, ~ latitude,
                        "Sokela Ambositra", 47.2481, -20.5157)
storage_chu_tambohobe <- tribble(~ place, ~ longitude, ~ latitude,
                                "CHU Tambohobe", 47.0869, -21.4444)
storage_cicm_tana <- tribble(~ place, ~ longitude, ~ latitude,
                             "CICM Tana", 47.5522, -18.9139)

transport_chu_sokela <- tribble(~ place, ~ longitude, ~ latitude,
                                "CHU Tambohobe", 47.0869, -21.4444,
                                "Sokela Ambositra", 47.2481, -20.5157)
transport_sokela_cicm <- tribble(~ place, ~ longitude, ~ latitude,
                                 "Sokela Ambositra", 47.2481, -20.5157,
                                 "CICM Tana", 47.5522, -18.9139)

## Final map
gg_sample_transport_mada_new <- ggmap(madagascar_sample_logistics) +
  geom_point(data = storage_chu_tambohobe,
            aes(x = longitude,
                y = latitude),
            shape = 21,
            size = 3,
            color = "black",
            fill = "purple") +
  geom_point(data = storage_sokela,
            aes(x = longitude,
                y = latitude),
            shape = 21,
            size = 3,
            color = "black",
            fill = "blue") +
  geom_point(data = storage_cicm_tana,
            aes(x = longitude,
                y = latitude),
            shape = 21,
            size = 3,
            color = "black",
            fill = "red") +
  geom_line(data = transport_chu_sokela,
            aes(x = longitude,
                y = latitude),
            color = "yellow",

```



```

    linewidth = 1.0,
    alpha = 0.8) +
geom_line(data = transport_sokela_cicm,
  aes(x = longitude,
    y = latitude),
  color = "orange",
  linewidth = 1.0,
  alpha = 0.6) +
coord_equal() +
north(x.min = 34.871, x.max = 58.821,
  y.min = -26.647, y.max = -11, scale = 0.2) +
scalebar(x.min = 40, x.max = 55,
  y.min = -24, y.max = -15,
  transform = TRUE, dist_unit = "km",
  dist = 200, model = "WGS84", height = 0.04,
  st.dist = 0.04, st.size = 3) +
theme(text = element_text(family = "Times New Roman"),
  axis.line = element_line()
  ) +
labs(
  x = "Longitude",
  y = "Latitude")

ggsave("gg_sample_transport_mada_new.png", gg_sample_transport_mada_new)

```

---

### ## Figure 10 -- The Principle of a ‘Sandwich’-ELISA

```

df_figure_principle_elisa_raw <- tibble(x = 1:10,
  y = 1:10)

gg_figure_elisa_principle_raw <- ggplot(df_figure_principle_elisa_raw,
  aes(x, y)) +
  theme(
    text = element_text(family = "Times New Roman")
  ) +
  labs(x = NULL,
    y = NULL)

## Take a manual screenshot of the example in the book.

magick_principle_elisa_image <- image_read("/Users/leonardgunga/Desktop/R_data/SeroCoV/sandwich_elisa.png")

size_elisa <- image_scale(magick_principle_elisa_image, "900")

gg_figure_elisa_principle_better <- gg_figure_elisa_principle_raw +
  theme(
    panel.background = element_blank()
  ) +
  draw_image(size_elisa)

ggsave("gg_figure_elisa_principle_better.png", gg_figure_elisa_principle_better)

```

## ## Results

---

```
## Download SeroCoV data set
SeroCoV_FI_dat1 <- read_dta("SeroCoV_FI_Leonard_20220222.dta")

## Create a reduced data set, excluding variables with more than 90% NA
SeroCoV_FI_dat1_reduced <- SeroCoV_FI_dat1 %>%
  select(screening_id, s_latitudeorigin_q, s_longitudeorigin_q, age_group_10yrs,
  age_group_5yrs, test_result, r_sample, everything(), h11_rooms_s,
  -c(h16_toiletother_s, h17_cookingother_s,
  h18_lightenergyother_c, h19_waterrainysother_c,
  h20_waterdrysother_c, h22_amentv_q,
  h23_amenplayer_q, h23_1_satelittedecoder_q,
  h24_amenlandline_q, h25_amenmobiletypeother_s,
  h26_amenfridgefreezer_q, h27_amenestove_q,
  h28_amengasstove_q, h29_amenoilstove_q,
  h30_amensolar_q, h31_amendsolarlamp_q,
  h32_amendoillamp_q, h33_gaslamp_q, h10_horsecart_d,
  h34_car_q, h35_amendmotorbike_q, h36_amendtricycle_q,
  h37_amendbicycle_q, h38_computer_q,
  h39_laptop_q, h40_bednets_q, h43_specialmeasuresother3_s,
  h42_measurehhoother_s, h42_measurehhoother2_s,
  h42_measurehhoother3_s, h43_vulnerableother1_s, h43_vulnerableother2_s, h43_vulnerableother3_s,
  h43_specialmeasuresother_s, h43_specialmeasuresother2_s,
  m04_noconsent_c, m04_noconsentwhy_s,
  m05_qualificationothermg_s, m07_salarydetail_d,
  m07_salaryday_d, m07_salary_q, m07_salaryper_d,
  m18_sharerroom_q, m26_coronatestwhere_c, m33_0_currency_c,
  m06_indivemploymentother_s, m12_transportother_s,
  m25_pastyartreatmen_v_3, m26_coronatestwhereother_s,
  m26_durationtestresultother_c, m33_treatmentcostcon_v_4,
  m34_treatmentcosthos_v_5, m42_sourceinfo_s,
  m43_trustinfooother_s, m44_4_changehabitother_c,
  m49_mostdifficultother_c))

## Further reduction
SeroCoV_FI_dat1_reduced2 <- SeroCoV_FI_dat1_reduced %>%
```

---

## ## Figure 11 -- Data Collection Over Time

```
gg_data_collection_over_time <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, s_latitudeorigin_q, s_longitudeorigin_q,
  test_result, r_sample, s_dateappointment, h02_date_s) %>%
  mutate(year = year(h02_date_s),
  month = month(h02_date_s),
  week = week(h02_date_s),
  date = day(h02_date_s)) %>%
  mutate(new_date = ymd(glue("2021-{:month}-{:date}"))) %>%
  ggplot(aes(x = new_date)) +
  geom_histogram(binwidth = 7,
  fill = "red",
  color = "#D0D3D4",
  alpha = 0.5) +
  scale_x_date(date_labels = "%b %d",
  date_breaks = "1 week") +
  scale_y_continuous(expand = c(0,0),
  breaks = seq(0, 100, 10),
  labels = seq(0, 100, 10),
  limits = c(0, 100)) +
  theme(
  text = element_text(family = "Times New Roman"),
  panel.background = element_blank(),
  panel.grid = element_blank(),
  axis.line = element_line(),
  axis.text.x = element_text(size=rel(1),
  angle = 60,
```

```

      vjust = 0.5,
      hjust = 0.5)
    ) +
  labs(x = NULL,
       y = "Enrolled Participants per Week")

## April 2nd - Good Friday
## April 4th - Eastern

gg_data_collection_over_time
ggsave("gg_data_collection_over_time.png", gg_data_collection_over_time)

```

---

### ## Figure 12 -- Development of COVID-19 Cases in Madagascar

```

## 'Our World in Data' COVID-19 cases per country

df_owid_covid_full <- read_excel("owid-covid-data-2_copy.xlsx")

gg_owid_new_cases_sampling_period <- df_owid_covid_full %>%
  select(iso_code, location, date, total_cases, new_cases, new_cases_smoothed,
         new_deaths, new_deaths_smoothed, reproduction_rate) %>%
  filter(., iso_code %in% c("MDG")) %>%
  mutate(year = year(date),
         month = month(date),
         week = week(date),
         date = day(date)) %>%
  mutate(new_date = ymd(glue("{year}-{month}-{date}"))) %>%
  select(new_date, new_cases) %>%
  ggplot(aes(x = new_date, y = new_cases)) +
  geom_line() +
  scale_x_date(labels = scales::label_date("%b %d"),
              date_breaks = "1 week",
              limits = as.Date(c("2021-02-26", "2021-06-18"))) +
  scale_y_continuous(expand = c(0,0),
                    breaks = seq(0, 1500, 250),
                    labels = seq(0, 1500, 250),
                    limits = c(0, 1500)) +
  theme(
    text = element_text(family = "Times New Roman"),
    panel.background = element_blank(),
    panel.grid = element_blank(),
    axis.line = element_line(),
    axis.text.x = element_text(size=rel(1),
                              angle = 60,
                              vjust = 0.5,
                              hjust = 0.5)
  ) +
  labs(x = NULL,
       y = "New cases per day")

ggsave("gg_owid_new_cases_sampling_period.png", gg_owid_new_cases_sampling_period)

```

---

### ## Figure 13 -- Spatial Distribution of Included Households

```

corrected_coordinates_all <- SeroCoV_FI_dat1 %>%
  select(screening_id, s_longitudeorigin_q, s_latitudeorigin_q) %>%
  mutate(corrected_s_longitudeorigin_q = case_when(
    s_longitudeorigin_q < 10 ~ (s_longitudeorigin_q*10),
    between(s_longitudeorigin_q, 10, 50) ~ s_longitudeorigin_q,
    between(s_longitudeorigin_q, (4.7*10^5), (5.0*10^5)) ~ (s_longitudeorigin_q/10000),
    between(s_longitudeorigin_q, (4.7*10^6), (5.0*10^6)) ~ (s_longitudeorigin_q/100000),
    s_longitudeorigin_q > (4.7*10^7) ~ (s_longitudeorigin_q/1000000))) %>%
  mutate(corrected_s_latitudeorigin_q = case_when(
    between(s_latitudeorigin_q, -30, -10) ~ s_latitudeorigin_q,
    between(s_latitudeorigin_q, (-3*10^4), (-2*10^4)) ~ (s_latitudeorigin_q/1000),
    between(s_latitudeorigin_q, (-2*10^6), (-2*10^5)) ~ (s_latitudeorigin_q/10000),

```

```

between(s_latitudeeorigin_q,(-2*10^7), (-2*10^6)) ~ (s_latitudeeorigin_q/100000),
between(s_latitudeeorigin_q,(-2*10^8), (-2*10^7)) ~ (s_latitudeeorigin_q/1000000)))

fianarantsoa_Sero_Cov <- get_stamenmap(
  bbox = c(left = 47.0, bottom = -21.5, right = 47.2, top = -21.4),
  maptype = "terrain",
  zoom = 13)
fianarantsoa_with_sampling_coord <- ggmap(fianarantsoa_Sero_Cov) +
  geom_point(data = corrected_coordinates_all,
    aes(x = corrected_s_longitudeorigin_q,
      y = corrected_s_latitudeeorigin_q),
    size = 0.1,
    color = "red") +
  coord_equal() +
  north(x.min = 47.0, x.max = 47.2,
    y.min = -21.405, y.max = -21.5,
    scale = 0.2) +
  scalebar(x.min = 47.10, x.max = 47.185,
    y.min = -21.480, y.max = -21.4,
    transform = TRUE,
    model = "WGS84",
    dist_unit = "km",
    dist = 2,
    height = 0.02,
    st.dist = 0.05,
    st.size = 4) +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.line = element_line()
  ) +
  labs(x = "Longitude",
    y = "Latitude",
    colour = "GPS Coordinates")

mada_overview <- ggmap(get_stamenmap(
  bbox = c(left = 34.871, bottom = -26.647, right = 58.821, top = -10.790),
  maptype = "terrain",
  zoom = 6)) +
  geom_point(aes(x = 47.08,
    y = -21.45),
    size = 0.8,
    color = "purple") +
  geom_rect(xmin = 34.871, xmax = 58.821, ymin = -26.647, ymax = -10.790, fill = NA,
    colour = "black", linewidth = 1) +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.title = element_blank(),
    axis.text = element_blank(),
    axis.ticks = element_blank(),
    panel.background = element_rect(fill = "transparent", colour = NA),
    plot.background = element_rect(fill = "transparent", colour = NA))

gg_final_map <- ggdraw() +
  draw_plot(fianarantsoa_with_sampling_coord) +
  draw_plot(mada_overview,
    x = 0.15,
    y = 0.70,
    width = 0.15,
    height = 0.15,
    scale = 1.0)

ggsave("gg_final_map.png", gg_final_map)

```

---

## ## Figure 14 -- Number of Households included per Fokontany

## Use the preparation of Figure 2 in methods (fokontanys\_ & sf\_merged\_final\_rapport)

```

## Mapping of included households per Fokontany
gg_number_included_households <- ggmap(fokontanys_) +
  geom_sf(sf_merged_final_rapport,
    mapping = aes(fill = number_of_households_included_v2,
      geometry = geometry),
    inherit.aes = FALSE,
    color = "black",
    linetype = 1,
    lwd = 0.25,
    alpha = 0.6) +
  geom_sf_text(data = sf_merged_final_rapport,
    mapping = aes(label = ADM4_EN, geometry = geometry),
    inherit.aes = FALSE,
    show.legend = FALSE,
    size = 1.5) +
  coord_sf() +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.line = element_line(),
    legend.title = element_markdown()) +
  labs(
    x = "Longitude",
    y = "Latitude",
    fill = "Number of included <br> households")

ggsave("gg_number_included_households.png", gg_number_included_households)

```

---

### ## Table 2 -- Participants ELISA Test Results - SeroCoV Madagascar

```

gt_overall_test_results <- SeroCoV_FI_dat1_reduced2 %>%
  select(test_result) %>%
  mutate(., test_result = case_when(
    test_result %in% c(1,2) ~ "negative",
    test_result %in% c("3") ~ "positive")) %>%
  mutate(., test_result = factor(test_result,
    levels = c("positive",
      "negative"))) %>%
  tbl_summary(digits = ~ 1,
    label = test_result ~ "IgG ELISA") %>%
  add_n() %>%
  bold_labels() %>%
  modify_header(stat_0 = "**n (%)**") %>%
  modify_header(n = "**Number of Participants**") %>%
  modify_table_styling(
    columns = label,
    rows = label == "negative",
  footnote = "incl. 3 ambiguous ELISA Test Results") %>%
  add_ci(statistic = list(all_categorical() ~
    "[{conf.low}; {conf.high}]")
  ) %>%
  as_gt() %>%
  tab_header(title = md("**Table 2.** Participants ELISA Test Results - SeroCoV Madagascar")) %>%
  opt_row_stripping() %>%
  opt_table_font(
    font = "Times New Roman")

gt_overall_test_results %>%
  gtsave(filename = "gt_overall_test_results.docx")

```

---

### ## Table 3 -- Stratification Goals – Fianarantsoa

```

stratification_goals_1 <- tribble(
  ~ age_group, ~ male_1, ~ female_1, ~ total_1,
  "adolescents_10_19", 110, 110, 220,
  "adults_20_44", 155, 155, 310,
  "adults_old_45", 60, 60, 120,

```

```

"Sum", 325, 325, 650)

stratification_goals_2 <- tribble(
  ~ age_group, ~ male_2, ~ female_2, ~ total_2,
  "adolescents_10_19", 140, 138, 278,
  "adults_20_44", 282, 301, 583,
  "adults_old_45", 115, 145, 260,
  "Sum", 537, 584, 1121)

stratification_goals_prep <- left_join(stratification_goals_1, stratification_goals_2)

stratification_goals_united <- stratification_goals_prep %>%
  unite("male_united", c(male_2, male_1), sep = "/",
    remove = FALSE) %>%
  unite("female_united", c(female_2, female_1), sep = "/",
    remove = FALSE) %>%
  unite("total_united", c(total_2, total_1), sep = "/",
    remove = FALSE) %>%
  mutate(age_group_2 = case_when(age_group %in% c("adolescents_10_19") ~ "Adolescents 10 ≤ 19 years",
    age_group %in% c("adults_20_44") ~ "Adults 20 ≤ 44 years",
    age_group %in% c("adults_old_45") ~ "Adults ≥ 45 years",
    age_group %in% c("Sum") ~ "Sum")) %>%
  select(age_group_2, male_united, female_united, total_united)

gt_stratification_goals_united <- stratification_goals_united %>%
  gt(rownames_col = "age_group_2") %>%
  tab_header(title = md("**Table 3.** Stratification Goals - Fianarantsoa")) %>%
  tab_stubhead(label = md("**Age Group**")) %>%
  cols_label(male_united = md("**Male**"),
    female_united = md("**Female**"),
    total_united = md("**Total**")) %>%
  tab_footnote(
    footnote = "Included /Intended number of participants",
    locations = cells_column_labels(
      columns = c(male_united, female_united, total_united))) %>%
  opt_row_stripping() %>%
  opt_table_font(
    font = "Times New Roman") %>%
  opt_align_table_header(align = "left")

gt_stratification_goals_united %>%
  gtsave(filename = "gt_stratification_goals_united.docx")

```

---

#### ## Table 4 -- Overview Table of Collected Data

```

## Download excel sheet with an overview of household and individual variables
df_household_vs_individual <- read_excel("Overview household_individual_variables.xlsx")

gt_household_vs_individual <- df_household_vs_individual %>%
  mutate(n_row = 1:n()) %>%
  gt(groupnames_col = "Group", rownames_col = "n_row") %>%
  tab_header(title = md("**Table 4.** Overview Table of Collected Data")) %>%
  tab_stubhead(label = md("**Group**")) %>%
  opt_row_stripping() %>%
  sub_missing(
    columns = 1:3,
    missing_text = ""
  ) %>%
  cols_label(Individual = md("**Individual** N = 1,121"),
    Household = md("**Household** N = 674")) %>%
  opt_table_font(
    font = "Times New Roman") %>%
  opt_align_table_header(align = "left")

gt_household_vs_individual %>%
  gtsave(filename = "gt_household_vs_individual.docx")

```

---

**## Table 5.1 -- Demographics of Participants**

```
# Demographics table
df_demographics_1 <- SeroCoV_FI_dat1_reduced2 %>%
  select(., m01_2_ageyears_q, age_group_10yrs, m01_1_sex_d,
         m05_qualificationmg_c, test_result, r_sample,
         screening_id) %>%
  mutate(., test_result = case_when(
    test_result %in% c(1,2) ~ "negative",
    test_result %in% c("3") ~ "positive"),
    test_result = factor(test_result,
      levels = c("positive",
                 "negative")),
    age_group_10yrs = case_when(age_group_10yrs %in% c(1) ~ "10 ≤ 19 years",
      age_group_10yrs %in% c(2) ~ "20 ≤ 29 years",
      age_group_10yrs %in% c(3) ~ "30 ≤ 39 years",
      age_group_10yrs %in% c(4) ~ "40 ≤ 49 years",
      age_group_10yrs %in% c(5) ~ "50 ≤ 59 years",
      age_group_10yrs %in% c(6) ~ "≥ 60 years"),
    age_group_10yrs = factor(age_group_10yrs,
      levels = c("10 ≤ 19 years",
                 "20 ≤ 29 years",
                 "30 ≤ 39 years",
                 "40 ≤ 49 years",
                 "50 ≤ 59 years",
                 "≥ 60 years")),
    m05_qualificationmg_c = case_when(m05_qualificationmg_c %in% c(1) ~ "No degree",
      m05_qualificationmg_c %in% c(2) ~ "Primary",
      m05_qualificationmg_c %in% c(4) ~ "Secondary school",
      m05_qualificationmg_c %in% c(5) ~ "Baccalauréat",
      m05_qualificationmg_c %in% c(6) ~ "License",
      m05_qualificationmg_c %in% c(7) ~ "Master's degree",
      # m05_qualificationmg_c %in% c(8) ~ "Doctorate",
      m05_qualificationmg_c %in% c(9) ~ "Other",
      m05_qualificationmg_c %in% c(10) ~ "University student"),
    m05_qualificationmg_c = factor(m05_qualificationmg_c,
      levels = c("No degree",
                 "Primary",
                 "Secondary school",
                 "Baccalauréat",
                 "University student",
                 "License",
                 "Master's degree",
                 # "Doctorate",
                 "Other")),
    m01_1_sex_d = case_when(
      m01_1_sex_d %in% c("0") ~ "Male",
      m01_1_sex_d %in% c("1") ~ "Female")
  )

df_demographics_2 <- SeroCoV_FI_dat1_reduced2 %>%
  select(m01_2_ageyears_q,
         m06_indivemployment_c,
         m07_mainsourceincome_c,
         test_result, r_sample,
         screening_id) %>%
  filter(m01_2_ageyears_q > 14) %>%
  mutate(m06_indivemployment_currently_employed = case_when(m06_indivemployment_c %in% c(1:5, 8) ~
    "Yes",
    m06_indivemployment_c %in% c(6, 9, 10) ~ "No",
    m06_indivemployment_c %in% c(7, 11, 12) ~ "Other"),
    m06_indivemployment_currently_employed = factor(m06_indivemployment_currently_employed,
      levels = c("Yes",
                 "No",
                 "Other")),
    m06_indivemployment_c = case_when(m06_indivemployment_c %in% c(1) ~ "Employed",
```

```

m06_indivemployment_c %in% c(2) ~ "Self-employed",
m06_indivemployment_c %in% c(3) ~ "Farming",
m06_indivemployment_c %in% c(4) ~ "Student/ Intern/ Apprentice",
m06_indivemployment_c %in% c(5) ~ "Cleaning lady",
m06_indivemployment_c %in% c(6) ~ "House keeper/ homemaker",
m06_indivemployment_c %in% c(7) ~ "Retired",
m06_indivemployment_c %in% c(8) ~ "Governmental employee",
m06_indivemployment_c %in% c(9) ~ "Unemployed (health reasons)",
m06_indivemployment_c %in% c(10) ~ "Unemployed (other reasons)",
m06_indivemployment_c %in% c(11) ~ "Not applicable (under 18)",
m06_indivemployment_c %in% c(12) ~ "Other"),
m06_indivemployment_c = factor(m06_indivemployment_c,
  levels = c("Farming",
    "Employed",
    "Self-employed",
    "Governmental employee",
    "Cleaning lady",
    "House keeper/ homemaker",
    "Student/ Intern/ Apprentice",
    "Retired",
    "Unemployed (health reasons)",
    "Unemployed (other reasons)",
    "Not applicable (under 18)",
    "Other")),
m07_mainsourceincome_c = case_when(m07_mainsourceincome_c %in% c(1) ~ "Health care",
  m07_mainsourceincome_c %in% c(2) ~ "Business, trade, retail",
  m07_mainsourceincome_c %in% c(3) ~ "Education",
  m07_mainsourceincome_c %in% c(4) ~ "Building and construction",
  m07_mainsourceincome_c %in% c(5) ~ "Transportation",
  m07_mainsourceincome_c %in% c(6) ~ "Agriculture",
  m07_mainsourceincome_c %in% c(7) ~ "Landlords or public officials",
  m07_mainsourceincome_c %in% c(8) ~ "Other"),
m07_mainsourceincome_c = factor(m07_mainsourceincome_c,
  levels = c("Agriculture",
    "Transportation",
    "Building and construction",
    "Business, trade, retail",
    "Health care",
    "Education",
    "Landlords or public officials",
    "Other")),
test_result = case_when(
  test_result %in% c(1,2) ~ "negative",
  test_result %in% c("3") ~ "positive"),
test_result = factor(test_result,
  levels = c("positive",
    "negative"))
)

# Percentage table
# 1
demo_table1 <- df_demographics_1 %>%
  select(m01_2_ageyears_q, age_group_10yrs, m01_1_sex_d) %>%
  tbl_summary(percent = "column",
    digits = NULL,
    label = list(age_group_10yrs ~ "Age groups",
      m01_2_ageyears_q ~ "Age",
      m01_1_sex_d ~ "Sex"),
    type = all_continuous() ~ "continuous2",
    statistic = list(all_continuous() ~ c("{median} ({p25}, {p75})",
      all_categorical() ~ c("{n} ({p}%)")
    )
  ) %>%
  bold_labels() %>%
  add_n(., statistic = "{n}")

demo_pos_neg_ind_1 <- df_demographics_1 %>%
  select(age_group_10yrs, m01_1_sex_d, test_result) %>%
  tbl_summary(by = "test_result",

```



```

    percent = "row",
    digits = NULL,
    label = list(age_group_10yrs ~ "Age groups",
                 m01_1_sex_d ~ "Sex")) %>%
add_ci(statistic = list(all_categorical() ~
                        "[{conf.low}; {conf.high}]%]")
    )
)

demo_merge1 <- tbl_merge(tbls = list(demo_table1, demo_pos_neg_ind_1))

# 2
demo_table2 <- df_demographics_1 %>%
  select(m05_qualificationmg_c, r_sample, screening_id) %>%
  filter(!duplicated(screening_id)) %>%
  select(m05_qualificationmg_c) %>%
  tbl_summary(percent = "column",
              digits = NULL,
              label = list(m05_qualificationmg_c ~ "Education")) %>%
  bold_labels() %>%
  add_n(., statistic = "{n}")

demo_pos_neg_ind_2 <- df_demographics_1 %>%
  select(m05_qualificationmg_c, test_result, screening_id) %>%
  filter(!duplicated(screening_id)) %>%
  select(m05_qualificationmg_c, test_result) %>%
  tbl_summary(by = "test_result",
              percent = "row",
              digits = NULL,
              label = list(m05_qualificationmg_c ~ "Education")) %>%
  add_ci(statistic = list(all_categorical() ~
                          "[{conf.low}; {conf.high}]%]")
    )

demo_merge2 <- tbl_merge(tbls = list(demo_table2, demo_pos_neg_ind_2))

#3
demo_table3 <- df_demographics_2 %>%
  select(m06_indivemployment_currently_employed, m06_indivemployment_c,
         m07_mainsourceincome_c) %>%
  tbl_summary(percent = "column",
              digits = NULL,
              label = list(m06_indivemployment_currently_employed ~ "Employment status",
                           m06_indivemployment_c ~ "Profession",
                           m07_mainsourceincome_c ~ "Main source of income")) %>%
  bold_labels() %>%
  add_n(., statistic = "{n}")

demo_pos_neg_ind_3 <- df_demographics_2 %>%
  select(m06_indivemployment_currently_employed, m06_indivemployment_c,
         m07_mainsourceincome_c, test_result) %>%
  tbl_summary(by = "test_result",
              percent = "row",
              digits = NULL,
              label = list(m06_indivemployment_currently_employed ~ "Employment status",
                           m06_indivemployment_c ~ "Profession",
                           m07_mainsourceincome_c ~ "Main source of income")) %>%
  add_ci(statistic = list(all_categorical() ~
                          "[{conf.low}; {conf.high}]%]")
    )

demo_merge3 <- tbl_merge(tbls = list(demo_table3, demo_pos_neg_ind_3))

# Final merge
demo_final_merge <- tbl_stack(tbls = list(demo_merge1, demo_merge2, demo_merge3)) %>%
  modify_header(stat_0_1 ~ ")") %>%
  modify_spanning_header((c(stat_1_2, ci_stat_1_2, stat_2_2, ci_stat_2_2) ~
                          "***ELISA Test Result***"),
                        (c(n_1, stat_0_1) ~ NA)) %>%

```

```

modify_footnote(stat_0_1 ~ NA)

gt_demo_final_merge <- demo_final_merge %>%
  as_gt()

gt_demo_final_merge_rdy <- gt_demo_final_merge %>%
  tab_header(title = md("***Table 5.1** Demographics of Participants")) %>%
  tab_footnote(
    footnote = "Only participants older than 14 years were included.",
    locations = cells_body(
      columns = label,
      rows = label %in% c("Employment status", "Profession", "Main source of income")
    )
  ) %>%
  opt_row_stripping() %>%
  opt_table_font(
    font = "Times New Roman" %>%
  )
  opt_align_table_header(align = "left")

gt_demo_final_merge_rdy %>%
  gtsave(filename = "gt_demo_final_merge_rdy.docx")

```

---

**## Table 5.2 -- ELISA Test Result stratified by Sex and Age Groups**

```

df_ten_age_group_sex_test_result <- SeroCoV_FI_dat1_reduced2 %>%
  select(m01_2_ageyears_q,
         age_group_10yrs,
         m01_1_sex_d,
         test_result,
         screening_id) %>%
  mutate(., test_result = case_when(
    test_result %in% c(1,2) ~ "negative",
    test_result %in% c("3") ~ "positive"),
    test_result = factor(test_result,
      levels = c("positive",
        "negative")),
    age_group_10yrs = case_when(age_group_10yrs %in% c(1) ~ "10 ≤ 19",
      age_group_10yrs %in% c(2) ~ "20 ≤ 29",
      age_group_10yrs %in% c(3) ~ "30 ≤ 39",
      age_group_10yrs %in% c(4) ~ "40 ≤ 49",
      age_group_10yrs %in% c(5) ~ "50 ≤ 59",
      age_group_10yrs %in% c(6) ~ "≥ 60"),
    age_group_10yrs = factor(age_group_10yrs,
      levels = c("10 ≤ 19",
        "20 ≤ 29",
        "30 ≤ 39",
        "40 ≤ 49",
        "50 ≤ 59",
        "≥ 60")),
    m01_1_sex_d = case_when(
      m01_1_sex_d %in% c("0") ~ "Male",
      m01_1_sex_d %in% c("1") ~ "Female")
  )

gt_ten_age_group_sex_test_result <- df_ten_age_group_sex_test_result %>%
  select(m01_1_sex_d, m01_2_ageyears_q, age_group_10yrs, test_result) %>%
  tbl_strata(
    strata = m01_1_sex_d,
    .tbl_fun =
      ~ .x %>%
    tbl_summary(by = test_result,
      percent = "row",
      missing = "no",
      label = list(
        m01_2_ageyears_q ~ "Age",
        age_group_10yrs ~ "Age Group"),

```

```

        type = all_continuous() ~ "continuous2",
        statistic = list(all_continuous() ~ c("{median} ({p25}, {p75})),
            all_categorical() ~ c("{n} ({p}%)"))
    ) %>%
    add_n() %>%
    add_ci(include = age_group_10yrs,
        statistic = list(all_categorical() ~ "[{conf.low}; {conf.high}%]")) %>%
    bold_labels(),
    .header = "***{strata}***"
) %>%
as_gt() %>%
tab_header(title = md("***Table 5.2** ELISA Test Result stratified by Sex and Age Groups")) %>%
opt_row_stripping() %>%
opt_table_font(
    font = "Times New Roman") %>%
tab_footnote(
    footnote = "In years",
    locations = cells_body(
        columns = label,
        rows = label %in% c("Age", "Age Group")
    ))

gt_ten_age_group_sex_test_result %>%
gtsave(filename = "gt_ten_age_group_sex_test_result.docx")

```

---

### ## Table 5.3 -- Comparison of Educational Degrees

```

df_education_comparison_2 <- tribble(~ Level_of_education, ~ Malagasy_in_fr, ~ English, ~ German,
    "No degree" , "Pas de diplôme", "No degree", "Kein Abschluss",
    "Primary" , "Primaire", "Primary", "Grundschule",
    "Secondary", "Secondaire", "Middle school", "Mittlere Schulabschluss",
    "Secondary", "Baccalauréat", "High school", "Abitur",
    "Higher", "Étudiant à l'université", "University student", "Student",
    "Higher", "Licence", "Bachelor's degree", "Bachelor",
    "Higher", "Master", "Master", "Master",
    "Higher", "Doctorat", "PhD", "Doktor",
    "Other", "Autre", "Other", "Andere") %>%
mutate(n_row = 1:n())

gt_education_comparison_2 <- df_education_comparison_2 %>%
gt(groupname_col = "Level_of_education", rowname_col = "n_row") %>%
tab_stubhead(label = md("***Level of Education***")) %>%
tab_header(title = md("***Table 5.3** Comparison of Educational Degrees")) %>%
cols_label(Malagasy_in_fr = md("***Malagasy***"),
    English = md("***English***"),
    German = md("***German***")) %>%
tab_source_note(
    source_note = md("Reference: Madagascar - Public expenditure review 2015 : education (English). Washington,
D.C. : World Bank Group. <br>
*[http://documents.worldbank.org/curated/en/271911468185343596/Madagascar-Public-expenditure-review-2015-education]*")
) %>%
opt_row_stripping() %>%
opt_table_font(
    font = "Times New Roman") %>%
opt_align_table_header(align = "left")

gt_education_comparison_2 %>%
gtsave(filename = "gt_education_comparison_2.docx")

```

---

### ## Table 5.4 -- Profession and Education Levels

```

df_employment_vs_education <- SeroCoV_FI_dat1_reduced2 %>%
select(screening_id, m01_2_ageyears_q, m05_qualificationmg_c, m06_indivemployment_c) %>%
filter(!duplicated(screening_id) &
    m01_2_ageyears_q > 14) %>%

```

```

mutate(
  m05_qualificationmg_c_modified = case_when(
    m05_qualificationmg_c %in% c(1) ~ "No degree",
    m05_qualificationmg_c %in% c(2) ~ "Primary",
    m05_qualificationmg_c %in% c(4,5) ~ "Secondary school",
    m05_qualificationmg_c %in% c(6,7,8,10) ~ "Higher",
    m05_qualificationmg_c %in% c(9) ~ "Other"),
  m05_qualificationmg_c_modified = factor(m05_qualificationmg_c_modified,
    levels = c("No degree",
      "Primary",
      "Secondary school",
      "Higher",
      "Other")
),
  m06_indivemployment_c_modified = case_when(m06_indivemployment_c %in% c(1) ~ "Employed",
    m06_indivemployment_c %in% c(2) ~ "Self-employed",
    m06_indivemployment_c %in% c(3) ~ "Farming",
    m06_indivemployment_c %in% c(4) ~ "Student/ Intern/ Apprentice",
    m06_indivemployment_c %in% c(5) ~ "Cleaning lady",
    m06_indivemployment_c %in% c(6) ~ "House keeper/ homemaker",
    m06_indivemployment_c %in% c(7) ~ "Retired",
    m06_indivemployment_c %in% c(8) ~ "Governmental employee",
    m06_indivemployment_c %in% c(9) ~ "Unemployed (health reasons)",
    m06_indivemployment_c %in% c(10) ~ "Unemployed (other reasons)",
    m06_indivemployment_c %in% c(11) ~ "Not applicable (under 18)",
    m06_indivemployment_c %in% c(12) ~ "Other"),
  m06_indivemployment_c_modified = factor(m06_indivemployment_c_modified,
    levels = c("Farming",
      "Employed",
      "Self-employed",
      "Governmental employee",
      "Cleaning lady",
      "House keeper/ homemaker",
      "Student/ Intern/ Apprentice",
      "Retired",
      "Unemployed (health reasons)",
      "Unemployed (other reasons)",
      "Not applicable (under 18)",
      "Other")
  )
) %>%
select(contains("_modified"))

gtsum_employment_vs_education <- df_employment_vs_education %>%
tbl_summary(by = "m05_qualificationmg_c_modified",
  label = m06_indivemployment_c_modified ~ "Profession") %>%
add_n() %>%
bold_labels()

gt_employment_vs_education <- gtsum_employment_vs_education %>%
as_gt() %>%
tab_header(title = md("***Table 5.4** Profession and Education Levels")) %>%
tab_footnote(footnote = "Only participants older than 14 years were included.",
  locations = cells_body(
    columns = label,
    rows = label %in% c("Profession")
  ))

gt_employment_vs_education %>%
gtsave(filename = "gt_profession_and_education_levels.docx")

```

---

### ## Figure 15 -- Population Pyramid of SeroCoV Participants

```

age_pyramid_serocov <- SeroCoV_FI_dat1_reduced2 %>%
select(m01_1_sex_d, m01_2_ageyears_q, age_group_5yrs, age_group_10yrs) %>%
mutate(m01_1_sex_d_modified = if_else(m01_1_sex_d %in% c("1"),
  "Female", "Male"),

```

```

m01_2_ageyears_q_modified = case_when(
  between(m01_2_ageyears_q, 10, 19) ~ "10-19",
  between(m01_2_ageyears_q, 20, 29) ~ "20-29",
  between(m01_2_ageyears_q, 30, 39) ~ "30-39",
  between(m01_2_ageyears_q, 40, 49) ~ "40-49",
  between(m01_2_ageyears_q, 50, 59) ~ "50-59",
  between(m01_2_ageyears_q, 60, 69) ~ "60-69",
  between(m01_2_ageyears_q, 70, 79) ~ "70-79",
  between(m01_2_ageyears_q, 81, 120) ~ "80 +",
  TRUE ~ "Unknown"),
m01_1_sex_d_modified = as_factor(m01_1_sex_d_modified),
m01_2_ageyears_q_modified = factor(m01_2_ageyears_q_modified,
  levels = c("10-19",
    "20-29",
    "30-39",
    "40-49",
    "50-59",
    "60-69",
    "70-79",
    "80 +"))

ap_population_serocov <- age_pyramid(data = age_pyramid_serocov,
  age_group = "m01_2_ageyears_q_modified",
  split_by = "m01_1_sex_d_modified",
  proportional = TRUE) +
  theme_minimal() +
  scale_fill_manual(
    values = c("orange", "purple"),
    labels = c("Male", "Female")) +
  theme(text = element_text(family = "Times New Roman"),
    legend.position = "bottom",
    axis.text = element_text(size = 10),
    axis.title = element_text(size = 12)
  ) +
  labs(y = "Percent of all cases",
    x = "Age groups",
    fill = "Gender")

ggsave("ap_population_serocov.png", ap_population_serocov)

```

---

### ## Table 6.1 -- Clinical and Epidemiologic Characteristics of Participants

```

clinical_epi_table <- SeroCoV_FI_dat1_reduced2 %>%
  select(h08_suspected_c, h08_confirmed_c, m17_knownconfcasehh_c, m17_contactconfcaseouthh_c,
    m19_underlyingcond_c, c(m24_pastyearfever_d:m24_pastyearrespiratoysymp_d), m26_coronatest_c,
    m14_contactday_c, s_membersno_q, test_result, r_sample) %>%
  mutate(., test_result = case_when(
    test_result %in% c(1,2) ~ "negative",
    test_result %in% c("3") ~ "positive"),
    test_result = factor(test_result,
      levels = c("positive",
        "negative")),
    across(c(h08_suspected_c, h08_confirmed_c,
      m17_knownconfcasehh_c,
      m17_contactconfcaseouthh_c), as.factor),
    exposure_confirmed = case_when(
      (h08_suspected_c %in% c("1") & (h08_confirmed_c %in% c(1, 2, 3))) |
      (m17_knownconfcasehh_c %in% c("1")) | (m17_contactconfcaseouthh_c %in% c("1"))) ~ 'Yes',
      h08_suspected_c %in% c("2") ~ "Do not know",
      h08_suspected_c %in% c("0") | (h08_suspected_c %in% c("1")) & (h08_confirmed_c %in% c("5")) ~ 'No'),
    exposure_confirmed = factor(exposure_confirmed,
      levels = c("Yes",
        "No",
        "Do not know")),
    Ever_test_for_Covid_19 = case_when(
      m26_coronatest_c %in% c(0) ~ "No",

```

```

m26_coronatest_c %in% c(1) ~ "Yes",
m26_coronatest_c %in% c(2) ~ "Do not know"),
Ever_test_for_Covid_19 = factor(Ever_test_for_Covid_19,
                                levels = c("Yes",
                                             "No",
                                             "Do not know")),
daily_contacts = case_when(
  m14_contactday_c %in% 1 ~ "less than 5",
  m14_contactday_c %in% 2 ~ "5 to less than 10",
  m14_contactday_c %in% 3 ~ "10 to less than 50",
  m14_contactday_c %in% 4 ~ "50 or more"),
daily_contacts = factor(daily_contacts,
                        levels = c("less than 5",
                                    "5 to less than 10",
                                    "10 to less than 50",
                                    "50 or more")),
householdsize = case_when(
  s_membersno_q %in% c(1:2) ~ "1 to 2",
  s_membersno_q %in% c(3:5) ~ "3 to 5",
  s_membersno_q %in% c(6:26) ~ "≥ 6"),
householdsize = factor(householdsize,
                       levels = c("1 to 2",
                                   "3 to 5",
                                   "≥ 6")),
m19_underlyingcond_c = case_when(m19_underlyingcond_c %in% 1 ~ "Yes",
                                  m19_underlyingcond_c %in% 0 ~ "No",
                                  m19_underlyingcond_c %in% 2 ~ "Do not know"),
m19_underlyingcond_c = factor(m19_underlyingcond_c,
                              levels = c("Yes",
                                          "No",
                                          "Do not know"))
) %>%
rowwise() %>%
mutate(past_year_covid_symptoms = case_when(
  1 %in% across(contains("m24_")) ~ "Yes",
  2 %in% across(contains("m24_")) ~ "Do not know",
  0 %in% across(contains("m24_")) ~ "No"),
past_year_covid_symptoms = factor(past_year_covid_symptoms,
                                  levels = c("Yes",
                                              "No",
                                              "Do not know"))
)

clin_epi1 <- clinical_epi_table %>%
select(Ever_test_for_Covid_19, exposure_confirmed,
       past_year_covid_symptoms,
       m19_underlyingcond_c, daily_contacts,
       householdsize) %>%
tbl_summary(percent = "column",
            digits = NULL,
            label = list(exposure_confirmed ~ "Exposure confirmed",
                          past_year_covid_symptoms ~
                          "COVID-19 Symptome over <br> the past 12 months",
                          Ever_test_for_Covid_19 ~ "Previous SARS-CoV-2 testing",
                          daily_contacts ~ "Daily contacts",
                          householdsize ~ "Household size",
                          m19_underlyingcond_c ~ "At least one underlying <br> chronic health concern")) %>%
bold_labels() %>%
add_n(., statistic = "{n}")

clin_epi_2 <- clinical_epi_table %>%
select(Ever_test_for_Covid_19,
       exposure_confirmed,
       past_year_covid_symptoms,
       m19_underlyingcond_c,
       daily_contacts,
       householdsize,
       test_result) %>%

```

```

tbl_summary(by = "test_result",
            percent = "row",
            digits = NULL,
            label = list(exposure_confirmed ~ "Exposure confirmed",
                          past_year_covid_symptoms ~
                            "COVID-19 Symptome over <br> the past 12 months",
                          Ever_test_for_Covid_19 ~ "Previous SARS-CoV-2 testing",
                          daily_contacts ~ "Daily contacts",
                          householdsize ~ "Household size",
                          m19_underlyingcond_c ~ "At least one underlying <br> chronic health concern")) %>%
add_ci(statistic = list(all_categorical() ~
                        "[{conf.low}; {conf.high}]")
)

clin_epi_final_merge <- tbl_merge(tbls = list(clin_epi1, clin_epi_2)) %>%
modify_header(stat_0_1 ~ ") %>%
modify_spanning_header((c(stat_1_2, ci_stat_1_2, stat_2_2, ci_stat_2_2) ~
                        "***ELISA Test Result***"),
                      (c(n_1, stat_0_1) ~ NA)) %>%
modify_footnote(stat_0_1 ~ NA)

clin_epi_final_merge <- clin_epi_final_merge %>%
as_gt() %>%
tbl_header(title = md("***Table 6.1** Clinical and Epidemiologic Characteristics of Participants")) %>%
opt_row_stripping() %>%
opt_table_font(font = "Times New Roman") %>%
opt_align_table_header(align = "left") %>%
gtsave(filename = "clin_epi_final_merge.docx")

```

---

## ## Table 6.2 -- Treatment of COVID-19 Symptoms

```

df_past_year_symptoms_treatment <- SeroCoV_FI_dat1_reduced2 %>%
select(screening_id, m01_1_sex_d, m01_2_stratagroup_c, test_result,
c(m24_pastyearfever_d:m24_pastyearrespiratoysymp_d),
  c(m25_pastytreatplace___th_d:m25_pastytreatplace___other_d)) %>%
rowwise() %>%
mutate(past_year_covid_symptoms = case_when(
  1 %in% across(contains("m24_")) ~ "Yes",
  2 %in% across(contains("m24_")) ~ "Do not know",
  0 %in% across(contains("m24_")) ~ "No"),
past_year_treatment = case_when(
  1 %in% across(contains("m25")) ~ "Yes",
  2 %in% across(contains("m25")) ~ "Do not know",
  0 %in% across(contains("m25")) ~ "No"))

df_past_year_treatment_types <- df_past_year_symptoms_treatment %>%
select(past_year_covid_symptoms, contains("m25_")) %>%
filter(str_detect("Yes", past_year_covid_symptoms)) %>%
mutate(across(c(contains("m25_")), ~ case_when(.x %in% c(1) ~ "Mentioned",
                                             .x %in% c(0) ~ "Not mentioned"),
          .names = "{col}_modified")
)

gtsm_treatment_past_year <- df_past_year_treatment_types %>%
select(contains("_modified")) %>%
tbl_summary(
  label = list(m25_pastytreatplace___th_d_modified ~ "Traditional healer",
              m25_pastytreatplace___tsm_d_modified ~ "Traditional self-medication",
              m25_pastytreatplace___msm_d_modified ~ "Modern self-medication",
              m25_pastytreatplace___pha_d_modified ~ "Private pharmacy",
              m25_pastytreatplace___street_d_modified ~ "Street pharmacy",
              m25_pastytreatplace___hcf_d_modified ~ "Health care facility",
              m25_pastytreatplace___other_d_modified ~ "Other")) %>%
bold_labels() %>%
modify_header(label = "***Which places did you seek after <br>
developing COVID-19 symptoms <br>
in the past year?***")

```

```

gt_treatment_past_year <- gtsum_treatment_past_year %>%
  as_gt()

gt_treatment_past_year %>%
  tab_header(title = md("***Table 6.2** Treatment of COVID-19 Symptoms")) %>%
  opt_row_stripping() %>%
  opt_table_font(
    font = "Times New Roman") %>%
  opt_align_table_header(align = "left") %>%
  gtsave(filename = "gt_treatment_past_year.docx")

```

### ## Table 6.3 -- Chronic Diseases among Participants

```

df_chronic_conditions <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, contains("m19_")) %>%
  mutate(across(c(2:11), ~ case_when(x %in% c(1) ~ "Yes",
    .x %in% c(0) ~ "No"),
    .names = "{col}_modified"),
    across(contains("_modified"), ~ factor(x,
      levels = c("Yes",
        "No")))
  )
) %>%
  select(screening_id, contains("_modified"))

gtsum_chronic_conditions_1 <- df_chronic_conditions %>%
  select(m19_underlyingcond_c_modified) %>%
  tbl_summary(label = m19_underlyingcond_c_modified ~ "Do you have chronic diseases?",
    type = m19_underlyingcond_c_modified ~ "categorical") %>%
  add_n() %>%
  bold_labels()

gtsum_chronic_conditions_2 <- df_chronic_conditions %>%
  filter(m19_underlyingcond_c_modified %in% c("Yes")) %>%
  select(m19_condition_c__lung_d_modified, m19_condition_c__dia_d_modified,
    m19_condition_c__heart__modified, m19_condition_c__hyperten_d_modified,
    m19_condition_c__other_d_modified) %>%
  tbl_summary(label = list(m19_condition_c__lung_d_modified ~ "Lung",
    m19_condition_c__dia_d_modified ~ "Diabetes",
    m19_condition_c__heart__modified ~ "Heart",
    m19_condition_c__hyperten_d_modified ~ "Hypertension",
    m19_condition_c__other_d_modified ~ "Other")
  ) %>%
  add_n() %>%
  bold_labels()

gtsum_chronic_conditions_stacked <- tbl_stack(tbls = list(gtsum_chronic_conditions_1,
  gtsum_chronic_conditions_2))

gtsum_chronic_conditions_stacked

gtsum_chronic_conditions_stacked %>%
  as_gt() %>%
  tab_header(title = md("***Table 6.3** Chronic Diseases among Participants")) %>%
  gtsave(filename = "gtsum_chronic_conditions_stacked.docx")

```

### ## Table 6.4 -- Tobacco Smoking by Participants Sex

```

df_smoke <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, m20_smoke_c, m01_1_sex_d) %>%
  mutate(
    m20_smoke_c = case_when(m20_smoke_c %in% c(1) ~ "Yes",
      m20_smoke_c %in% c(0) ~ "No"
    ),

```



```

m20_smoke_c = factor(m20_smoke_c,
  levels = c(
    "Yes",
    "No")),
m01_1_sex_d = case_when(m01_1_sex_d %in% c("0") ~ "Male",
  m01_1_sex_d %in% c("1") ~ "Female")
) %>%
select(m20_smoke_c, m01_1_sex_d)

gtsum_smoke_1 <- df_smoke %>%
select(m20_smoke_c) %>%
tbl_summary(
  percent = "column",
  type = m20_smoke_c ~ "categorical",
  label = m20_smoke_c ~ "Do you smoke?") %>%
add_n() %>%
bold_labels()

gtsum_smoke_2 <- df_smoke %>%
select(m20_smoke_c, m01_1_sex_d) %>%
tbl_summary(by = "m01_1_sex_d",
  percent = "row",
  type = m20_smoke_c ~ "categorical",
  label = m20_smoke_c ~ "Do you smoke?") %>%
add_ci(statistic = list(all_categorical() ~
  "[{conf.low}; {conf.high}]")
)
gtsum_smoke_merged <- tbl_merge(tbls = list(gtsum_smoke_1,
  gtsum_smoke_2)) %>%
modify_header(stat_0_1 = "") %>%
modify_spanning_header(everything() ~ NA) %>%
modify_footnote(stat_0_1 = NA)

gt_smoke_merged <- gtsum_smoke_merged %>%
as_gt() %>%
tab_header(title = md("**Table 6.4** Tobacco Smoking by Participants Sex")) %>%
opt_row_stripping() %>%
opt_table_font(
  font = "Times New Roman") %>%
opt_align_table_header(align = "left") %>%
gtsave(filename = "gt_smoke_merged.docx")

```

---

**## Table 6.5 -- Coronavirus Exposure on the Way to School or Work, and at School or Work**

```

df_exposure_to_and_at_work <- SeroCoV_FI_dat1_reduced2 %>%
select(screening_id, m_memberbarcode_id,
  m09_workinginout_c, m10_worklocation_d,
  m12_transport_c, m13_contactway_c,
  test_result) %>%
mutate(m09_workinginout_c_modified = case_when(
  m09_workinginout_c %in% c(1) ~ "Indoor",
  m09_workinginout_c %in% c(2) ~ "Outdoor",
  m09_workinginout_c %in% c(3) ~ "Hospital",
  m09_workinginout_c %in% c(4) ~ "Other"),
  m09_workinginout_c_modified = factor(
  m09_workinginout_c_modified,
  levels = c("Indoor",
    "Outdoor",
    "Hospital",
    "Other")),
  m10_worklocation_d_modified = case_when(
  m10_worklocation_d %in% c(1) ~ "Yes",
  m10_worklocation_d %in% c(0) ~ "No"),
  m10_worklocation_d_modified = factor(
  m10_worklocation_d_modified,
  levels = c("Yes",
    "No")),

```

```

m12_transport_c_modified = case_when(
  m12_transport_c %in% c(1) ~ "Walk",
  m12_transport_c %in% c(2) ~ "Bus",
  m12_transport_c %in% c(3) ~ "Car",
  m12_transport_c %in% c(4) ~ "Bicycle",
  m12_transport_c %in% c(5) ~ "Taxi",
  m12_transport_c %in% c(7) ~ "Motorcycle",
  m12_transport_c %in% c(8) ~ "Other"),
m12_transport_c_modified = factor(
  m12_transport_c_modified,
  levels = c("Walk",
            "Bus",
            "Car",
            "Bicycle",
            "Taxi",
            "Motorcycle",
            "Other")),
m13_contactway_c_modified = case_when(
  m13_contactway_c %in% 1 ~ "less than 5",
  m13_contactway_c %in% 2 ~ "5 to less than 10",
  m13_contactway_c %in% 3 ~ "10 to less than 50",
  m13_contactway_c %in% 4 ~ "50 or more"),
m13_contactway_c_modified = factor(m13_contactway_c_modified,
  levels = c("less than 5",
            "5 to less than 10",
            "10 to less than 50",
            "50 or more")),
test_result = case_when(
  test_result %in% c(1,2) ~ "negative",
  test_result %in% c("3") ~ "positive"),
test_result = factor(test_result,
  levels = c("positive",
            "negative"))
) %>%
select(screening_id, m_memberbarcode_id,
  m09_workinginout_c_modified, m10_worklocation_d_modified,
  m12_transport_c_modified, m13_contactway_c_modified,
  test_result)

gtsum_exposure_to_and_at_work_1 <- df_exposure_to_and_at_work %>%
select(m12_transport_c_modified, m13_contactway_c_modified,
  m09_workinginout_c_modified, m10_worklocation_d_modified) %>%
tbl_summary(percent = "column",
  digits = NULL,
  type = m10_worklocation_d_modified ~ "categorical",
  label = list(m12_transport_c_modified ~ "How do you get to school <br> or work every day?",
    m13_contactway_c_modified ~ "How many people do you come in <br> contact with on your way
to school <br> or work?",
    m09_workinginout_c_modified ~ "Where do you spend most of <br> your class or working time?",
    m10_worklocation_d_modified ~ "Do you work and live <br> in the same house?")
  ) %>%
bold_labels() %>%
add_n(., statistic = "{n}")

gtsum_exposure_to_and_at_work_2 <- df_exposure_to_and_at_work %>%
select(m12_transport_c_modified, m13_contactway_c_modified,
  m09_workinginout_c_modified, m10_worklocation_d_modified,
  test_result) %>%
tbl_summary(by = "test_result",
  percent = "row",
  digits = NULL,
  type = m10_worklocation_d_modified ~ "categorical",
  label = list(m12_transport_c_modified ~ "How do you get to school <br> or work every day?",
    m13_contactway_c_modified ~ "How many people do you come in <br> contact with on your way
to school <br> or work?",
    m09_workinginout_c_modified ~ "Where do you spend most of <br> your class or working time?",
    m10_worklocation_d_modified ~ "Do you work and live <br> in the same house?")
  ) %>%
add_ci(statistic = list(all_categorical() ~

```

```

    "[{conf.low}; {conf.high}%]")
  )

gtsum_exposure_to_and_at_work_merge <- tbl_merge(tbls = list(gtsum_exposure_to_and_at_work_1,
gtsum_exposure_to_and_at_work_2)) %>%
  modify_header(stat_0_1 ~ ") %>%
  modify_spanning_header((c(stat_1_2, ci_stat_1_2, stat_2_2, ci_stat_2_2) ~
    "***ELISA Test Result***"),
    (c(n_1, stat_0_1) ~ NA)) %>%
  modify_footnote(stat_0_1 ~ NA)

gtsum_exposure_to_and_at_work_merge %>%
  as_gt() %>%
  tab_header(title = md("***Table 6.5** Coronavirus Exposure on the Way to School or Work, and at School or
Work")) %>%
  opt_row_stripping() %>%
  opt_table_font(
    font = "Times New Roman") %>%
  opt_align_table_header(align = "left") %>%
  gtsave(filename = "gt_exposure_to_and_at_work_final_merge.docx")

```

---

### ## Table 7.1 -- Economic Consequences for Included Households

```

df_not_working_new <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, h45_hheconomicconseq_s, h46_incomereduction_c,
    h47_borrowmoney_d, h48_lendmoney_c, s_membersno_q, r_sample) %>%
  filter(r_sample %in% c(1)) %>%
  mutate(., How_many_household_members_could_not_work_during_the_Corona_epidemic = case_when(
    h45_hheconomicconseq_s %in% c(0) ~ "0",
    h45_hheconomicconseq_s %in% c(1) ~ "1",
    h45_hheconomicconseq_s %in% c(2,3,4,5) ~ "2-5",
    h45_hheconomicconseq_s %in% c(6,7,9) ~ ">5"),
    How_many_household_members_could_not_work_during_the_Corona_epidemic = factor(
    How_many_household_members_could_not_work_during_the_Corona_epidemic,
    levels = c(">5",
      "2-5",
      "1",
      "0"))) %>%
  mutate(P_household_members_not_working_due_to_COVID = (h45_hheconomicconseq_s/s_membersno_q)*100,
    P_household_members_not_working_due_to_COVID =
  round(P_household_members_not_working_due_to_COVID, 0)) %>%
  relocate(., P_household_members_not_working_due_to_COVID, .after = h45_hheconomicconseq_s) %>%
  mutate(., PGroups_household_members_not_working_due_to_COVID = case_when(
    P_household_members_not_working_due_to_COVID %in% c(0) ~ "No impact",
    P_household_members_not_working_due_to_COVID %in% c(75:100) ~ "At least 75%",
    P_household_members_not_working_due_to_COVID %in% c(50:74) ~ "Between 50% and 74%",
    P_household_members_not_working_due_to_COVID %in% c(25:49) ~ "Between 25% and 49%",
    P_household_members_not_working_due_to_COVID %in% c(1:25) ~ "Not more than 24%"),
    PGroups_household_members_not_working_due_to_COVID =
  factor(PGroups_household_members_not_working_due_to_COVID,
    levels = c("At least 75%",
      "Between 50% and 74%",
      "Between 25% and 49%",
      "Not more than 24%",
      "No impact"))) %>%
  mutate(., Did_it_reduce_the_household_income = case_when(
    h46_incomereduction_c %in% c(1) ~ "Yes",
    h46_incomereduction_c %in% c(0) ~ "No",
    h46_incomereduction_c %in% c(2) ~ "Do not know"))) %>%
  mutate(Did_it_reduce_the_household_income = factor(
    Did_it_reduce_the_household_income,
    levels = c("Yes",
      "No",
      "Do not know"))) %>%
  mutate(., Borrow_money_from_people_or_bank_loan = as.factor(h47_borrowmoney_d)) %>%
  mutate(., Borrow_money_from_people_or_bank_loan = case_when(
    h47_borrowmoney_d %in% c(1) ~ "Yes",

```

```

h47_borrowmoney_d %in% c(0 ~ "No")) %>%
mutate(Borrow_money_from_people_or_bank_loan = factor(
  Borrow_money_from_people_or_bank_loan,
  levels = c("Yes",
    "No"))) %>%
mutate(., Did_you_lend_money_to_other_households = case_when(
  h48_lendmoney_c %in% c(1 ~ "Yes",
  h48_lendmoney_c %in% c(0 ~ "No",
  h48_lendmoney_c %in% c(2 ~ "Do not know"))) %>%
mutate(Did_you_lend_money_to_other_households = factor(
  Did_you_lend_money_to_other_households,
  levels = c("Yes",
    "No",
    "Do not know")))

gtsum_not_working_new <- df_not_working_new %>%
select(How_many_household_members_could_not_work_during_the_Corona_epidemic,
  PGroups_household_members_not_working_due_to_COVID, Did_it_reduce_the_household_income,
  Borrow_money_from_people_or_bank_loan, Did_you_lend_money_to_other_households) %>%
tbl_summary(percent = "column",
  digits = NULL,
  type = list(Borrow_money_from_people_or_bank_loan ~ "categorical"),
  label = list(How_many_household_members_could_not_work_during_the_Corona_epidemic ~
    "How many household members could not work \n during the Corona pandemic?",
  PGroups_household_members_not_working_due_to_COVID ~
    "Percentage of household members \n who could not work during the pandemic",
  Did_it_reduce_the_household_income ~
    "Did the pandemic reduce the household income?",
  Borrow_money_from_people_or_bank_loan ~
    "Did your household borrow money from other people \n or took a bank loan?",
  Did_you_lend_money_to_other_households ~
    "Did your household lend money to other households?")
) %>%
bold_labels()

gtsum_not_working_new %>%
as_gt() %>%
tab_header(title = md("***Table 7.1** Economic Consequences for Included Households")) %>%
gtsave(filename = "gtsum_not_working_new.docx")

```

---

### ## Table 7.2 -- Pandemics Impact on Work and School by Sex

```

df_missed_gender <- SeroCoV_FI_dat1_reduced2 %>%
select(m18_missedwork_d, m18_missoutwhat_c, m18_misshowmanydays_s, m18_misshowmanydays_s,
m01_1_sex_d, m01_2_stratagroup_c) %>%
mutate(m18_misshowmanydays_s = str_replace(m18_misshowmanydays_s, "120j", "120"),
  m18_misshowmanydays_s_modified = as.integer(m18_misshowmanydays_s),
  m18_misshowmanydays_s_modified = case_when(
    between(m18_misshowmanydays_s_modified, 0, 30) ~ "<= 30",
    between(m18_misshowmanydays_s_modified, 31, 90) ~ "31 <= 90",
    between(m18_misshowmanydays_s_modified, 91, 180) ~ "91 <= 180",
    between(m18_misshowmanydays_s_modified, 181, 300) ~ "181 <= 300",
    m18_misshowmanydays_s > c(300) ~ ">= 301"),
  m18_misshowmanydays_s_modified = factor(m18_misshowmanydays_s_modified,
    levels = rev(c("<= 30",
      "31 <= 90",
      "91 <= 180",
      "181 <= 300",
      ">= 301"))),
  m18_missedwork_d_modified = case_when(
    m18_missedwork_d %in% c("1") ~ "Yes",
    m18_missedwork_d %in% c("0") ~ "No"),
  m18_missedwork_d_modified = factor(m18_missedwork_d_modified,
    levels = c("Yes",
      "No")
  ),
  m18_missoutwhat_c_modified = case_when(

```

```

    m18_missoutwhat_c %in% c("1") ~ "School",
    m18_missoutwhat_c %in% c("2") ~ "Work"),
  m18_missoutwhat_c = factor(m18_missoutwhat_c,
    levels = c("School",
      "Work")
  ),
  m01_1_sex_d = case_when(m01_1_sex_d %in% c("0") ~ "Male",
    m01_1_sex_d %in% c("1") ~ "Female")
) %>%
select(m18_missedwork_d_modified, m18_missoutwhat_c_modified, m18_misshowmanydays_s_modified,
m01_1_sex_d, m01_2_stratagroup_c)

gtsum_missed_gender_1_missed <- df_missed_gender %>%
  select(m18_missedwork_d_modified, -m18_missoutwhat_c_modified, -m18_misshowmanydays_s_modified, -
m01_1_sex_d, -m01_2_stratagroup_c) %>%
  tbl_summary(
    percent = "column",
    type = m18_missedwork_d_modified ~ "categorical",
    label = m18_missedwork_d_modified ~ "Did you miss work or school <br> due to the pandemic/ lockdowns?")
%>%
  add_n() %>%
  bold_labels()

gtsum_missed_gender_1_what <- df_missed_gender %>%
  select(m18_missedwork_d_modified, m18_missoutwhat_c_modified, m18_misshowmanydays_s_modified, -
m01_1_sex_d, -m01_2_stratagroup_c) %>%
  filter(m18_missedwork_d_modified %in% c("Yes")) %>%
  select(m18_missoutwhat_c_modified, m18_misshowmanydays_s_modified) %>%
  tbl_summary(percent = "column",
    type = m18_missoutwhat_c_modified ~ "categorical",
    label = list(m18_missoutwhat_c_modified ~ "What did you miss?",
      m18_misshowmanydays_s_modified ~ "How many days have you missed?")
  ) %>%
  add_n() %>%
  bold_labels()

gtsum_missed_gender_2_missed <- df_missed_gender %>%
  select(m18_missedwork_d_modified, -m18_missoutwhat_c_modified, -m18_misshowmanydays_s_modified,
m01_1_sex_d, -m01_2_stratagroup_c) %>%
  tbl_summary(by = "m01_1_sex_d",
    percent = "row",
    type = m18_missedwork_d_modified ~ "categorical",
    label = m18_missedwork_d_modified ~ "Did you miss work or school <br> due to the pandemic/
lockdowns?") %>%
  add_ci(statistic = list(all_categorical() ~
    "[{conf.low}; {conf.high}]")
  ) %>%
  bold_labels()

gtsum_missed_gender_merged_1 <- tbl_merge(tbls = list(gtsum_missed_gender_1_missed,
gtsum_missed_gender_2_missed))

gtsum_missed_gender_2_what <- df_missed_gender %>%
  select(m18_missedwork_d_modified, m18_missoutwhat_c_modified, m18_misshowmanydays_s_modified,
m01_1_sex_d, -m01_2_stratagroup_c) %>%
  filter(m18_missedwork_d_modified %in% c("Yes")) %>%
  select(m18_missoutwhat_c_modified, m18_misshowmanydays_s_modified, m01_1_sex_d) %>%
  tbl_summary(by = "m01_1_sex_d",
    percent = "row",
    type = m18_missoutwhat_c_modified ~ "categorical",
    label = list(m18_missoutwhat_c_modified ~ "What did you miss?",
      m18_misshowmanydays_s_modified ~ "How many days have you missed?")
  ) %>%
  add_ci(statistic = list(all_categorical() ~
    "[{conf.low}; {conf.high}]")
  ) %>%
  bold_labels()

```

```

gtsum_missed_gender_merged_2 <- tbl_merge(tbls = list(gtsum_missed_gender_1_what,
gtsum_missed_gender_2_what))

gtsum_missed_gender_stack <-tbl_stack(tbls = list(gtsum_missed_gender_merged_1,
gtsum_missed_gender_merged_2)) %>%
  modify_header(stat_0_1 = "") %>%
  modify_spanning_header(everything() ~ NA) %>%
  modify_footnote(stat_0_1 = NA)

gtsum_missed_gender_stack %>%
  as_gt() %>%
  tab_header(title = md("***Table 7.2** Pandemics Impact on Work and School by Sex")) %>%
  gtsave(filename = "gtsum_missed_gender_stack.docx")

```

---

### ## Table 8.1 -- Participants Mental Health by Sex

```

Worries_WBI_gender <- SeroCoV_FI_dat1_reduced2 %>%
  select(m53_whoworries_s, screening_id,m52_1_whocheerful_s,
         m52_2_whocalm_s, m52_3_whoactive_s, m52_4_whofresh_s,
         m52_5_whointerest_s, test_result, -(r_sample), m01_1_sex_d) %>%
  rowwise() %>%
  mutate(., WBI_Score = (sum(c(m52_1_whocheerful_s,
                             m52_2_whocalm_s,
                             m52_3_whoactive_s,
                             m52_4_whofresh_s,
                             m52_5_whointerest_s)))*4) %>%
  relocate(., WBI_Score, .after = m52_5_whointerest_s) %>%
  mutate(., WBI_cut = cut(WBI_Score, breaks =
                          c(100, 90, 80, 70,
                             60, 50, 40, 30,
                             20, 10, 0)),
         WBI_cut = case_when(
           WBI_cut %in% c("(90,100]") ~ "91 - 100",
           WBI_cut %in% c("(80,90]") ~ "81 - 90",
           WBI_cut %in% c("(70,80]") ~ "71 - 80",
           WBI_cut %in% c("(60,70]") ~ "61 - 70",
           WBI_cut %in% c("(50,60]") ~ "51 - 60",
           WBI_cut %in% c("(40,50]", "(30,40]", "(20,30]", "(10,20]", "(0,10]") ~ "<= 50"),
         WBI_cut = factor(WBI_cut,
                          levels = c("91 - 100",
                                     "81 - 90",
                                     "71 - 80",
                                     "61 - 70",
                                     "51 - 60",
                                     "<= 50")),
         m53_whoworries_s_named = case_when(
           m53_whoworries_s %in% c(1) ~ "Worried",
           m53_whoworries_s %in% c(2) ~ "Very worried",
           m53_whoworries_s %in% c(3) ~ "A little worried",
           m53_whoworries_s %in% c(4) ~ "Not worried at all",
           m53_whoworries_s %in% c(5) ~ "Do not know"),
         m53_whoworries_s_named = factor(m53_whoworries_s_named,
                                         levels = c("Very worried",
                                                    "Worried",
                                                    "A little worried",
                                                    "Not worried at all",
                                                    "Do not know")
                                         ),
         m01_1_sex_d = case_when(m01_1_sex_d %in% c("0") ~ "Male",
                                m01_1_sex_d %in% c("1") ~ "Female")
         ) %>%
  select(m53_whoworries_s_named, WBI_cut, m01_1_sex_d)

gtsum_Worries_WBI_gender_1 <- Worries_WBI_gender %>%
  select(m53_whoworries_s_named, WBI_cut) %>%
  tbl_summary(percent = "column",
             digits = NULL,

```

```

label = list(m53_whoconcerns_s_named ~
  "Were you worried about the health <br>
  of family/ friends during the time <br>
  of the SARS-Cov-2 outbreak?",
  WBI_cut ~ "WHO-5 Well-Being Score") ) %>%
bold_labels() %>%
add_n(., statistic = "{n}")

gtsum_Worries_WBI_gender_2 <- Worries_WBI_gender %>%
select(m53_whoconcerns_s_named, WBI_cut, m01_1_sex_d) %>%
tbl_summary(by = "m01_1_sex_d",
  percent = "row",
  digits = NULL,
  label = list(m53_whoconcerns_s_named ~
    "Were you worried about the health <br>
    of family/ friends during the time <br>
    of the SARS-Cov-2 outbreak?",
    WBI_cut ~ "WHO-5 Well-Being Score") ) %>%
bold_labels() %>%
add_ci(statistic = list(all_categorical() ~
  "[{conf.low}; {conf.high}]")
)

gtsum_Worries_WBI_gender_merged <- tbl_merge(tbls = list(gtsum_Worries_WBI_gender_1,
gtsum_Worries_WBI_gender_2)) %>%
modify_header(stat_0_1 = "") %>%
modify_spanning_header(everything() ~ NA) %>%
modify_footnote(stat_0_1 = NA)

gtsum_Worries_WBI_gender_merged %>%
as_gt() %>%
tab_header(title = md("***Table 8.1** Participants Mental Health by Sex")) %>%
tab_footnote(
  footnote = "WHO-5 score ranges from worst imaginable well-being (0 P) to best imaginable well-being (100 P).",
  locations = cells_body(
    columns = label,
    rows = label %in% c("WHO-5 Well-Being Score")
  )
) %>%
tab_footnote(
  footnote = "Cut-score ≤ 50 points indicates screening for depression.",
  locations = cells_body(
    columns = label,
    rows = label %in% ("≤ 50")
  )
) %>%
tab_source_note(
  source_note = md("Reference: Topp C, W, Østergaard S, D, Søndergaard S, Bech P: The WHO-5 Well-Being
Index: A Systematic <br>
Review of the Literature. Psychother Psychosom 2015;84:167-176. doi: 10.1159/000376585 <br>
*[https://www.karger.com/Article/FullText/376585]*")
) %>%
gtsave(filename = "gtsum_Worries_WBI_gender_merged.docx")

```

---

### ## Table 8.2 -- Changes in Living Conditions since the Pandemic by Sex

```

df_since_outbreak_changes <- SeroCoV_FI_dat1_reduced2 %>%
select(screening_id, r_sample, test_result, m01_1_sex_d,
  m44_1_changeeatinghabbit_c,
  m44_2_changefooddeviation_c,
  m44_3_changemorealc_c, m44_4_changehabit_c,
  m44_5_changeaccessfood_c) %>%
mutate(across(c(m44_1_changeeatinghabbit_c:m44_3_changemorealc_c), ~ case_when(.x %in% c(1) ~ "Yes,
totally",
  .x %in% c(2) ~ "Yes, a little",
  .x %in% c(3) ~ "No, not much",
  .x %in% c(4) ~ "No, not at all",
  .x %in% c(5) ~ "Do not know"),

```

```

.names = "{col}_modified"),
across(c(m44_1_changeeatinghabit_c_modified:m44_3_changemorealc_c_modified), ~ factor(.x,
levels = c(
  "Yes, totally",
  "Yes, a little",
  "No, not much",
  "No, not at all",
  "Do not know")
)),
m44_4_changehabit_c_modified = case_when(
  m44_4_changehabit_c %in% c(1) ~ "Increased quantity of food during meals",
  m44_4_changehabit_c %in% c(2) ~ "Decreased quantity of food during meals",
  m44_4_changehabit_c %in% c(3) ~ "Reduced the number of meals per day",
  m44_4_changehabit_c %in% c(4) ~ "Cut out certain foods",
  m44_4_changehabit_c %in% c(5) ~ "Other to specify",
  m44_4_changehabit_c %in% c(6) ~ "Do not know"
),
m44_4_changehabit_c_modified = factor(m44_4_changehabit_c_modified,
levels = c("Increased quantity of food during meals",
  "Decreased quantity of food during meals",
  "Reduced the number of meals per day",
  "Cut out certain foods",
  "Other to specify",
  "Do not know")
),
m44_5_changeaccessfood_c_modified = case_when(
  m44_5_changeaccessfood_c %in% c(4) ~ "Much worse than usual",
  m44_5_changeaccessfood_c %in% c(3) ~ "Worse than usual",
  # m44_5_changeaccessfood_c %in% c(2) ~ "Same as usual",
  # m44_5_changeaccessfood_c %in% c(1) ~ "Good Access",
  # m44_5_changeaccessfood_c %in% c(5) ~ "Better than usual",
  m44_5_changeaccessfood_c %in% c(1, 2) ~ "Good access or same as usual",
  m44_5_changeaccessfood_c %in% c(6) ~ "Much better"),
m44_5_changeaccessfood_c_modified = factor(m44_5_changeaccessfood_c_modified,
levels = c(
  "Much better",
  # "Better than usual",
  # "Good Access",
  # "Same as usual",
  "Good access or same as usual",
  "Worse than usual",
  "Much worse than usual")
),
m01_1_sex_d = case_when(m01_1_sex_d %in% c("0") ~ "Male",
  m01_1_sex_d %in% c("1") ~ "Female")
) %>%
select(screening_id, r_sample, test_result, m01_1_sex_d,
  m44_1_changeeatinghabit_c_modified,
  m44_2_changefooddeviation_c_modified,
  m44_4_changehabit_c_modified,
  m44_5_changeaccessfood_c_modified,
  m44_3_changemorealc_c_modified)

gtsum_since_outbreak_changes_1 <- df_since_outbreak_changes %>%
select(everything(), -screening_id, -r_sample, -test_result, -m01_1_sex_d) %>%
tbl_summary(
  percent = "column",
  label = list(m44_1_changeeatinghabit_c_modified ~ "have you changed your eating habits?",
    m44_2_changefooddeviation_c_modified ~ "have you made any food deviations?",
    m44_3_changemorealc_c_modified ~ "have you drunk more alcoholic <br> beverages than usual?",
    m44_4_changehabit_c_modified ~ "what type of change did you make?",
    m44_5_changeaccessfood_c_modified ~ "how was the access to basic <br> necessities?")
) %>%
add_n() %>%
#italicize_labels()
bold_labels()

gtsum_since_outbreak_changes_2 <- df_since_outbreak_changes %>%
select(everything(), -screening_id, -r_sample, -test_result, m01_1_sex_d) %>%

```



```

tbl_summary(by = "m01_1_sex_d",
            percent = "row",
            label = list(m44_1_changeeatinghabbit_c_modified ~ "have you changed your eating habits?",
                        m44_2_changefooddeviation_c_modified ~ "have you made any food deviations?",
                        m44_3_changemorealc_c_modified ~ "have you drunk more alcoholic <br> beverages than usual?",
                        m44_4_changehabit_c_modified ~ "what type of change did you make?",
                        m44_5_changeaccessfood_c_modified ~ "how was the access to basic <br> necessities?")
            ) %>%
add_ci(statistic = list(all_categorical() ~
                        "[{conf.low}; {conf.high}]")
        ) %>%
#italicize_labels()
bold_labels()
#%>%
#add_p() %>%
#add_significance_stars()

gtsum_since_outbreak_changes_merged <- tbl_merge(tbls = list(
gtsum_since_outbreak_changes_1, gtsum_since_outbreak_changes_2)) %>%
modify_header(label = "***Since the outbreak of the COVID-19 <br>
pandemic, ...**",
stat_0_1 = "") %>%
modify_spanning_header(everything() ~ NA) %>%
modify_footnote(stat_0_1 = NA)

gtsum_since_outbreak_changes_merged %>%
as_gt() %>%
opt_row_stripping() %>%
opt_table_font(
font = "Times New Roman") %>%
opt_align_table_header(align = "left") %>%
tab_header(title = md("***Table 8.2** Changes in Living Conditions since the Pandemic by Sex")) %>%
gtsave(filename = "gtsum_since_outbreak_changes_merged.docx")

```

---

### ## Table 9.1 -- Source of Information by Level of Education

```

m42_sourceInfo_modified_household_edu <- SeroCoV_FI_dat1_reduced2 %>%
select(screening_id, contains("m42_"), m05_qualificationmg_c) %>%
filter(!duplicated(screening_id)) %>%
mutate(across(c(2:10), ~ if_else(.x %in% c(1),
" Mentioned",
" Not mentioned"),
.names = "{col}_modified"),
m05_qualificationmg_c_modified = case_when(
m05_qualificationmg_c %in% c(1) ~ "No degree",
m05_qualificationmg_c %in% c(2) ~ "Primary",
m05_qualificationmg_c %in% c(4,5) ~ "Secondary school",
m05_qualificationmg_c %in% c(6,7,8,10) ~ "Higher",
m05_qualificationmg_c %in% c(9) ~ "Other"),
m05_qualificationmg_c_modified = factor(m05_qualificationmg_c_modified,
levels = c("No degree",
"Primary",
"Secondary school",
"Higher",
"Other"))
) %>%
select(screening_id, contains("_modified"))

gtsum_media_education_v2 <- m42_sourceInfo_modified_household_edu %>%
select(c(2:11)) %>%
tbl_summary(by = m05_qualificationmg_c_modified,
            label = list(m42_sourceinfo__socialm_d_modified ~
                        "Social media (e.g. Facebook)",
                        m42_sourceinfo__web_d_modified ~
                        "Webpages (e.g. Google)",
                        m42_sourceinfo__ministry_d_modified ~
                        "Ministry of Health",

```

```

m42_sourceinfo__radio_d_modified ~
  "Radio",
m42_sourceinfo__tv_d_modified ~
  "TV",
m42_sourceinfo__newspaper_d_modified ~
  "Newspaper",
m42_sourceinfo__healthwork_d_modified ~
  "Health workers",
m42_sourceinfo__famfriends_d_modified ~
  "Family and friends",
m42_sourceinfo__other_d_modified ~
  "Other")) %>%
modify_header(label = "***What is your source of information <br> about the Coronavirus?***") %>%
bold_labels() %>%
add_n() %>%
add_overall() %>%
modify_header(stat_0 ~ "",
  stat_1 ~ "***No degree**", <br> N = 62",
  stat_2 ~ "***Primary**", <br> N = 179",
  stat_3 ~ "***Secondary school**", <br> N = 355",
  stat_4 ~ "***Higher**", <br> N = 58",
  stat_5 ~ "***Other**", <br> N = 17") %>%
modify_footnote(stat_0 ~ NA)

gtsum_media_education_v2 %>%
as_gt() %>%
tab_header(title = md("***Table 9.1** Source of Information by Level of Education")) %>%
gtsave(filename = "gtsum_media_education_v2.docx")

```

---

## ## Table 9.2 -- Trust in Source of Information

```

df_trust_most <- SeroCoV_FI_dat1_reduced2 %>%
  select(m43_trustinfo_c, screening_id, m05_qualificationmg_c, test_result, m01_2_ageyears_q, r_sample,
    m01_1_sex_d, m01_2_stratagroup_c) %>%
  mutate(m43_trustinfo_c_modified = case_when(
    m43_trustinfo_c %in% c(1) ~ "Social media (e.g. Facebook)",
    m43_trustinfo_c %in% c(2) ~ "Webpages (e.g. Google)",
    m43_trustinfo_c %in% c(3) ~ "Ministry of Health",
    m43_trustinfo_c %in% c(4) ~ "Radio",
    m43_trustinfo_c %in% c(5) ~ "TV",
    m43_trustinfo_c %in% c(6) ~ "Newspaper",
    m43_trustinfo_c %in% c(7) ~ "Health workers",
    m43_trustinfo_c %in% c(8) ~ "Family and friends",
    m43_trustinfo_c %in% c(9) ~ "Other"),
    m43_trustinfo_c_modified = factor(m43_trustinfo_c_modified,
      levels = c(
        "Social media (e.g. Facebook)",
        "Webpages (e.g. Google)",
        "Ministry of Health",
        "Radio",
        "TV",
        "Newspaper",
        "Health workers",
        "Family and friends",
        "Other")),
    m01_1_sex_d = case_when(m01_1_sex_d %in% c("0") ~ "Male",
      m01_1_sex_d %in% c("1") ~ "Female"),
    age_group_10yrs = case_when(
      m01_2_stratagroup_c %in% c(1) ~ "10 ≤ 19 years",
      m01_2_stratagroup_c %in% c(2) ~ "20 ≤ 29 years",
      m01_2_stratagroup_c %in% c(3) ~ "30 ≤ 39 years",
      m01_2_stratagroup_c %in% c(4) ~ "40 ≤ 49 years",
      m01_2_stratagroup_c %in% c(5) ~ "50 ≤ 59 years",
      m01_2_stratagroup_c %in% c(6) ~ "≥ 60 years"),
    age_group_10yrs = factor(age_group_10yrs,
      levels = c(
        "10 ≤ 19 years",

```

```

      "20 ≤ 29 years",
      "30 ≤ 39 years",
      "40 ≤ 49 years",
      "50 ≤ 59 years",
      "≥ 60 years")),
m05_qualificationmg_c_modified = case_when(
  m05_qualificationmg_c %in% c(1) ~ "No degree",
  m05_qualificationmg_c %in% c(2) ~ "Primary",
  m05_qualificationmg_c %in% c(4,5) ~ "Secondary school",
  m05_qualificationmg_c %in% c(6,7,8,10) ~ "Higher",
  m05_qualificationmg_c %in% c(9) ~ "Other"),
m05_qualificationmg_c_modified = factor(m05_qualificationmg_c_modified,
  levels = c("No degree",
    "Primary",
    "Secondary school",
    "Higher",
    "Other"))
) %>%
select(m43_trustinfo_c_modified, m05_qualificationmg_c_modified,
  m01_2_stratagroup_c, m01_1_sex_d, screening_id
)

gtsum_trust_most_1 <- df_trust_most %>%
  filter(!duplicated(screening_id)) %>%
  select(m43_trustinfo_c_modified) %>%
  tbl_summary(percent = "column",
    label = m43_trustinfo_c_modified ~ "Which source of information <br>
    did you trust the most <br>
    during the pandemic?") %>%
  add_n() %>%
  bold_labels()

gtsum_trust_most_2 <- df_trust_most %>%
  filter(!duplicated(screening_id)) %>%
  select(m43_trustinfo_c_modified, m05_qualificationmg_c_modified) %>%
  tbl_summary(by = "m05_qualificationmg_c_modified",
    percent = "row",
    label = m43_trustinfo_c_modified ~ "Which source of information <br>
    did you trust the most <br>
    during the pandemic?")

gtsum_trust_most_merged <- tbl_merge(tbls = list(gtsum_trust_most_1, gtsum_trust_most_2)) %>%
  modify_header(stat_0_1 ~ "",
    stat_1_2 ~ "***No degree**", <br> N = 62",
    stat_2_2 ~ "***Primary**", <br> N = 179",
    stat_3_2 ~ "***Secondary school**", <br> N = 355",
    stat_4_2 ~ "***Higher**", <br> N = 58",
    stat_5_2 ~ "***Other**", <br> N = 17") %>%
  modify_spanning_header(c(n_1:stat_5_2) ~ NA) %>%
  modify_footnote(stat_0_1 ~ NA)

gtsum_trust_most_merged %>%
  as_gt() %>%
  tab_header(title = md("***Table 9.2** Trust in Source of Information")) %>%
  gtsave(filename = "gtsum_trust_most_merged.docx")

```

---

### ## Table 10.1 -- Non-pharmaceutical Measures by Level of Education

```

measures_hh_test_named_v2 <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, h41_protecthmembr_c,
    contains("h42_"), test_result, r_sample,
    m05_qualificationmg_c) %>%
  filter(!duplicated(screening_id)) %>%
  mutate(h41_protecthmembr_c_modified = case_when(
    h41_protecthmembr_c %in% c(1) ~ "Yes",
    h41_protecthmembr_c %in% c(0) ~ "No"),
    h41_protecthmembr_c_modified = factor(h41_protecthmembr_c_modified,

```

```

        levels = c("Yes",
                  "No"),
across(c(3:12), ~ if_else(.x %in% c(1),
                          "Mentioned",
                          "Not mentioned"),
       .names = "{col}_modified"),
test_result = case_when(
  test_result %in% c(1,2) ~ "negative",
  test_result %in% c(3) ~ "positive"),
test_result = factor(test_result,
                    levels = c("positive",
                              "negative")),
m05_qualificationmg_c_modified = case_when(
  m05_qualificationmg_c %in% c(1) ~ "No degree",
  m05_qualificationmg_c %in% c(2) ~ "Primary",
  m05_qualificationmg_c %in% c(4,5) ~ "Secondary school",
  m05_qualificationmg_c %in% c(6,7,8,10) ~ "Higher",
  m05_qualificationmg_c %in% c(9) ~ "Other"),
m05_qualificationmg_c_modified = factor(m05_qualificationmg_c_modified,
                                       levels = c("No degree",
                                                 "Primary",
                                                 "Secondary school",
                                                 "Higher",
                                                 "Other")
)
) %>%
select(screening_id, contains("_modified"), test_result)

gtsum_npi_education <- measures_hh_test_named_v2 %>%
select(c(2:11), m05_qualificationmg_c_modified) %>%
tbl_summary(by = "m05_qualificationmg_c_modified",
            percent = "column",
            digits = NULL,
            type = list(h41_protecthhmemb_c_modified ~ "categorical"),
            label = list(h41_protecthhmemb_c_modified ~
              "Did you take measures to protect <br> household members?",
                h42_measurehh__maskall_d_modified ~
              "Masks for every household member",
                h42_measurehh__masksymp_d_modified ~
              "Masks for household member <br> with symptoms",
                h42_measurehh__maskvul_d_modified ~
              "Masks for vulnerable <br> household members",
                h42_measurehh__nomeal_d_modified ~
              "Not taking meal together",
                h42_measurehh__noshare_d_modified ~
              "Not sharing personal objects",
                h42_measurehh__handwash_d_modified ~
              "Washing hands more regularly",
                h42_measurehh__contact_d_modified ~
              "Avoiding contacts with other <br> households members",
                h42_measurehh__crowd_d_modified ~
              "Avoiding market and other <br> crowded place",
                h42_measurehh__clean_d_modified ~
              "Cleaning households objects <br> more regularly")) %>%
add_n(., statistic = "{n}") %>%
add_overall() %>%
bold_labels() %>%
modify_header(label = "***Characteristic**",
              stat_0 ~ "",
              stat_1 ~ "***No degree**", <br> N = 62",
              stat_2 ~ "***Primary**", <br> N = 179",
              stat_3 ~ "***Secondary school**", <br> N = 355",
              stat_4 ~ "***Higher**", <br> N = 58",
              stat_5 ~ "***Other**", <br> N = 17") %>%
modify_footnote(stat_0~ NA)

gtsum_npi_education %>%
as_gt() %>%
tab_header(title = md("***Table 10.1** Non-pharmaceutical Measures by Level of Education")) %>%

```

```
gtsave(filename = "gtsum_npi_education.docx")
```

---

**## Table 10.2 -- Non-pharmaceutical Protection Measures and ELISA Test Result**

```
measures_hh_test_named <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, h41_protecthhmemb_c,
         contains("h42_"), test_result, r_sample) %>%
  filter(!duplicated(screening_id)) %>%
  mutate(h41_protecthhmemb_c_modified = case_when(
    h41_protecthhmemb_c %in% c(1) ~ "Yes",
    h41_protecthhmemb_c %in% c(0) ~ "No"),
         h41_protecthhmemb_c_modified = factor(h41_protecthhmemb_c_modified,
         levels = c("Yes",
                    "No")),
         across(c(3:12), ~ if_else(.x %in% c(1),
                                   "Mentioned",
                                   "Not mentioned"),
         .names = "{col}_modified"),
         test_result = case_when(
         test_result %in% c(1,2) ~ "negative",
         test_result %in% c(3) ~ "positive"),
         test_result = factor(test_result,
         levels = c("positive",
                   "negative"))) %>%
  select(screening_id, contains("_modified"), test_result)

measures_hh_test_1_v2 <- measures_hh_test_named %>%
  select(contains("_modified"), -h42_measurehh___other_d_modified) %>%
  tbl_summary(percent = "column",
             digits = NULL,
             type = list(h41_protecthhmemb_c_modified ~ "categorical"),
             label = list(h41_protecthhmemb_c_modified ~
               "Did you take measures to protect <br> household members?",
               h42_measurehh___maskall_d_modified ~
               "Masks for every household member",
               h42_measurehh___masksymp_d_modified ~
               "Masks for household member <br> with symptoms",
               h42_measurehh___maskvul_d_modified ~
               "Masks for vulnerable <br> household members",
               h42_measurehh___nomeal_d_modified ~
               "Not taking meal together",
               h42_measurehh___noshare_d_modified ~
               "Not sharing personal objects",
               h42_measurehh___handwash_d_modified ~
               "Washing hands more regularly",
               h42_measurehh___contact_d_modified ~
               "Avoiding contacts with other <br> households members",
               h42_measurehh___crowd_d_modified ~
               "Avoiding market and other <br> crowded place",
               h42_measurehh___clean_d_modified ~
               "Cleaning households objects <br> more regularly")) %>%
  add_n(., statistic = "{n}")

measures_hh_test_2_v2 <- measures_hh_test_named %>%
  select(contains("_modified"), -h42_measurehh___other_d_modified, test_result) %>%
  tbl_summary(by = "test_result",
             percent = "row",
             digits = NULL,
             type = list(h41_protecthhmemb_c_modified ~ "categorical"),
             label = list(h41_protecthhmemb_c_modified ~
               "Did you take measures to protect <br> household members?",
               h42_measurehh___maskall_d_modified ~
               "Masks for every household member",
               h42_measurehh___masksymp_d_modified ~
               "Masks for household member <br> with symptoms",
               h42_measurehh___maskvul_d_modified ~
```

```

      "Masks for vulnerable <br> household members",
h42_measurehh__nomeal_d_modified ~
      "Not taking meal together",
h42_measurehh__noshare_d_modified ~
      "Not sharing personal objects",
h42_measurehh__handwash_d_modified ~
      "Washing hands more regularly",
h42_measurehh__contact_d_modified ~
      "Avoiding contacts with other <br> households members",
h42_measurehh__crowd_d_modified ~
      "Avoiding market and other <br> crowded place",
h42_measurehh__clean_d_modified ~
      "Cleaning households objects <br> more regularly")) %>%
bold_labels() %>%
add_ci(statistic = list(all_categorical() ~
      "[{conf.low}; {conf.high}]%]")
)

gtsum_measures_hh_test <- tbl_merge(tbls = list(measures_hh_test_1_v2, measures_hh_test_2_v2)) %>%
modify_header(stat_0_1 ~ ") %>%
modify_spanning_header((c(stat_1_2, ci_stat_1_2,
      stat_2_2, ci_stat_2_2) ~
      "**ELISA Test Result**"),
      (c(n_1, stat_0_1) ~ NA)) %>%
modify_footnote(stat_0_1 ~ NA)

gtsum_measures_hh_test %>%
as_gt() %>%
tab_header(title = md("**Table 10.2** Non-pharmaceutical Protection Measures and ELISA Test Result")) %>%
gtsave(filename = "gtsum_measures_hh_test.docx")

```

---

### ## Table 10.3 -- Difficulties and Compliance to COVID-19 Measures by Age Groups

```

df_difficulties_compliance <- SeroCoV_FI_dat1_reduced2 %>%
select(screening_id, m49_mostdifficult_c,
      m50_generalcompliance_c, m51_shouldcomply_c,
      m01_1_sex_d, m01_2_stratagroup_c) %>%
mutate(n_row = 1:n(),
      m49_mostdifficult_c_modified = case_when(
        m49_mostdifficult_c %in% c(1) ~ "Maintain social distance",
        m49_mostdifficult_c %in% c(2) ~ "Wear face masks",
        m49_mostdifficult_c %in% c(3) ~ "Not work",
        m49_mostdifficult_c %in% c(4) ~ "Socially isolate",
        m49_mostdifficult_c %in% c(5) ~ "Quarantine",
        m49_mostdifficult_c %in% c(6) ~ "Frequently wash hands",
        m49_mostdifficult_c %in% c(7) ~ "Other"),
      m49_mostdifficult_c_modified = factor(m49_mostdifficult_c_modified,
        levels = c(
          "Maintain social distance",
          "Wear face masks",
          "Frequently wash hands",
          "Socially isolate",
          "Not work",
          "Quarantine",
          "Other")
      ),
      across(c(m50_generalcompliance_c, m51_shouldcomply_c), ~ case_when(.x %in% c(1) ~ "Yes",
        .x %in% c(0) ~ "No",
        .x %in% c(2) ~ "Do not know"),
      .names = "{col}_modified"),
      across(c(m50_generalcompliance_c_modified, m51_shouldcomply_c_modified), ~ factor(.x,
        levels = c("Yes",
          "No",
          "Do not know"))
      ),
      m01_1_sex_d = case_when(m01_1_sex_d %in% c("0") ~ "Male",
        m01_1_sex_d %in% c("1") ~ "Female"),

```

```

age_group_10yrs = case_when(
  m01_2_stratagroup_c %in% c(1) ~ "10 ≤ 19 years",
  m01_2_stratagroup_c %in% c(2) ~ "20 ≤ 29 years",
  m01_2_stratagroup_c %in% c(3) ~ "30 ≤ 39 years",
  m01_2_stratagroup_c %in% c(4) ~ "40 ≤ 49 years",
  m01_2_stratagroup_c %in% c(5) ~ "50 ≤ 59 years",
  m01_2_stratagroup_c %in% c(6) ~ "≥ 60 years"),
age_group_10yrs = factor(age_group_10yrs,
  levels = c(
    "10 ≤ 19 years",
    "20 ≤ 29 years",
    "30 ≤ 39 years",
    "40 ≤ 49 years",
    "50 ≤ 59 years",
    "≥ 60 years")
)
) %>%
select(screening_id, n_row, m49_mostdifficult_c_modified:m51_shouldcomply_c_modified, m01_1_sex_d,
  age_group_10yrs)

gtsum_difficulties_compliance_1_age <- df_difficulties_compliance %>%
select(everything(), -m01_1_sex_d, -screening_id, -n_row, -age_group_10yrs) %>%
tbl_summary(percent = "column",
  digits = NULL,
  label = list(m49_mostdifficult_c_modified ~
    "Which measure was the most <br>
    difficult to comply with?",
  m50_generalcompliance_c_modified ~
    "Do people in your immediate <br>
    environment comply with the <br>
    advised measures?",
  m51_shouldcomply_c_modified ~
    "Do you think that people <br>
    should comply with the <br>
    advised measures?")) %>%
bold_labels() %>%
add_n()

gtsum_difficulties_compliance_2_age <- df_difficulties_compliance %>%
select(everything(), -m01_1_sex_d, -screening_id, -n_row, age_group_10yrs) %>%
tbl_summary(
  by = "age_group_10yrs",
  percent = "column",
  label = list(m49_mostdifficult_c_modified ~
    "Which measure was the most <br>
    difficult to comply with?",
  m50_generalcompliance_c_modified ~
    "Do people in your immediate <br>
    environment comply with the <br>
    advised measures?",
  m51_shouldcomply_c_modified ~
    "Do you think that people <br>
    should comply with the <br>
    advised measures?")
)

gtsum_difficulties_compliance_age_merged <- tbl_merge(tbls = list(
  gtsum_difficulties_compliance_1_age, gtsum_difficulties_compliance_2_age)
) %>%
modify_header(stat_0_1 ~ "",
  stat_1_2 ~ "***10 ≤ 19**", <br> N = 138",
  stat_2_2 ~ "***20 ≤ 29**", <br> N = 140",
  stat_3_2 ~ "***30 ≤ 39**", <br> N = 301",
  stat_4_2 ~ "***40 ≤ 49**", <br> N = 282",
  stat_5_2 ~ "***50 ≤ 59**", <br> N = 145",
  stat_6_2 ~ "***≥ 60**", <br> N = 115") %>%
modify_spanning_header(c(n_1:stat_0_1) ~ NA,
  c(stat_1_2:stat_6_2) ~ "***Age Group in Years**") %>%
modify_footnote(stat_0_1 = NA)

```

```

gtsum_difficulties_compliance_age_merged %>%
  as_gt() %>%
  tab_header(title = md("***Table 10.3** Difficulties and Compliance to COVID-19 Measures by Age Groups"))
%>%
  gtsave(filename = "gtsum_difficulties_compliance_age_merged.docx")

```

---

### ## Table 11 -- Vulnerable persons by Level of Education

```

Vul_edu_1_v2 <- SeroCoV_FI_dat1_reduced2 %>%
  select(screening_id, contains("h43_w"), m05_qualificationmg_c, r_sample) %>%
  filter(!duplicated(screening_id)) %>%
  mutate(across(c(2:12), ~ if_else(.x %in% c(1),
    "Mentioned",
    "Not mentioned"),
    .names = "{col}_modified"),
    m05_qualificationmg_c_modified = case_when(
      m05_qualificationmg_c %in% c(1) ~ "No degree",
      m05_qualificationmg_c %in% c(2) ~ "Primary",
      m05_qualificationmg_c %in% c(4,5) ~ "Secondary school",
      m05_qualificationmg_c %in% c(6,7,8,10) ~ "Higher",
      m05_qualificationmg_c %in% c(9) ~ "Other"),
    m05_qualificationmg_c_modified = as.factor(m05_qualificationmg_c_modified),
    m05_qualificationmg_c_modified = factor(m05_qualificationmg_c_modified,
      levels = c("No degree",
        "Primary",
        "Secondary school",
        "Higher",
        "Other"))
  ) %>%
  select(screening_id, contains("_modified"))

gtsum_vulnerable_education <- Vul_edu_1_v2 %>%
  select(contains("_modified")) %>%
  tbl_summary(by = "m05_qualificationmg_c_modified",
    percent = "column",
    digits = NULL,
    label = list(h43_whoisvulnerable__m_d_modified ~
      "Men",
      h43_whoisvulnerable__w_d_modified ~
      "Women",
      h43_whoisvulnerable__old_d_modified ~
      "Elderly people",
      h43_whoisvulnerable__child_d_modified ~
      "Children",
      h43_whoisvulnerable__preg_d_modified ~
      "Pregnant women",
      h43_whoisvulnerable__diab_d_modified ~
      "Diabetic",
      h43_whoisvulnerable__card_d_modified ~
      "Cardiac",
      h43_whoisvulnerable__lung_d_modified ~
      "Pulmonary disease",
      h43_whoisvulnerable__immun_d_modified ~
      "Immunosuppressed people",
      h43_whoisvulnerable__v_0_modified ~
      "Do not know",
      h43_whoisvulnerable__other_d_modified ~
      "Other")) %>%
  add_n(., statistic = "{n}") %>%
  add_overall() %>%
  bold_labels() %>%
  modify_header(stat_0 ~ "",
    label = "***Who do you think is most vulnerable <br> when it comes to getting sick with <br> COVID-
19?***",

```



```

stat_1 ~ "***No degree**", <br> N = 62",
stat_2 ~ "***Primary**", <br> N = 179",
stat_3 ~ "***Secondary school**", <br> N = 355",
stat_4 ~ "***Higher**", <br> N = 58",
stat_5 ~ "***Other**", <br> N = 17") %>%
modify_footnote(stat_0~ NA)

gtsum_vulnerable_education %>%
as_gt() %>%
tab_header(title = md("***Table 11.** Vulnerable persons by Level of Education")) %>%
gtsave(filename = "gtsum_vulnerable_education.docx")

```

---

## ## PCA items preparation

```

dhs_wealth_items <- SeroCoV_FI_dat1 %>%
select(((c(o01_typehouse_c:o04_roofother_s,
h16_toilet_c:h39_laptop_q)) &
c(ends_with("_c"), ends_with("_d"))),
h40_bednets_d, h11_rooms_s, s_membersno_q,
screening_id, test_result) %>%
filter(!duplicated(screening_id)) %>%
mutate(., h11_rooms_s_modified = replace(h11_rooms_s,
h11_rooms_s %in% c("0"), "2"),
h11_rooms_s_modified = replace_na(h11_rooms_s_modified, "3"),
h11_rooms_s_modified = as.integer(h11_rooms_s_modified),
s_membersno_q_modified = as.integer(s_membersno_q),
Household_members_per_sleeping_room =
(s_membersno_q_modified/ h11_rooms_s_modified),
crowding_1 = if_else(Household_members_per_sleeping_room > 5,
true = "0", false = "1"),
crowding_1_modified = as.numeric(crowding_1),
h16_toilet_c_modified = case_when(
h16_toilet_c %in% c(1,2) ~ "1",
h16_toilet_c %in% c(3,4,5) ~ "0"),
h17_cooking_c_modified = case_when(
h17_cooking_c %in% c(1,2) ~ "1",
h17_cooking_c %in% c(3,4,5) ~ "0"),
h18_lightenergy_c_modified = case_when(
h18_lightenergy_c %in% c(1,2,4) ~ "1",
h18_lightenergy_c %in% c(3,5,6,7) ~ "0"),
h19_waterrainys_c_modified = case_when(
h19_waterrainys_c %in% c(1,2) ~ "1",
h19_waterrainys_c %in% c(3,4,5,6) ~ "0"),
h20_waterdrys_c_modified = case_when(
h20_waterdrys_c %in% c(1,2) ~ "1",
h20_waterdrys_c %in% c(3,4,5,6) ~ "0"),
o01_typehouse_c_modified = case_when(
o01_typehouse_c %in% c(2,4) ~ "1",
o01_typehouse_c %in% c(1,3) ~ "0"),
o03_floor_c_modified = case_when(
o03_floor_c %in% c(1,2,7) ~ "1",
o03_floor_c %in% c(3,4,5,6,8) ~ "0"),
o04_roof_c_modified = case_when(
o04_roof_c %in% c(1,2,3) ~ "1",
o04_roof_c %in% c(4,5,6) ~ "0"),
o02_walls_c_modified = case_when(
o02_walls_c %in% c(1,2,3) ~ "1",
o02_walls_c %in% c(4,5,6) ~ "0"),
test_result = case_when(
test_result %in% c(1,2) ~ "negative",
test_result %in% c("3") ~ "positive"),
test_result = factor(test_result,
levels = c("positive",
"negative"))
) %>%
select(contains("_modified"),
h40_bednets_d,

```

```

-h11_rooms_s_modified,
-s_membersno_q_modified,
c(h20_waterdrysother_c:h39_laptop_d),
-h25_amenmobile_d,
-h20_waterdrysother_c,
-h10_horsecart_d,
screening_id,
test_result) %>%
mutate(across(c(everything(), -screening_id, -test_result), ~as.numeric(.)),
        across(c(everything(), -screening_id, -test_result), ~replace_na(., 0))
        ) %>%
select(sort(names()))

## NA and 0 imputation

## 0:
zero_room_number_all_households <- SeroCoV_FI_dat1_reduced2 %>%
select(., screening_id, h11_rooms_s, s_membersno_q) %>%
filter(!duplicated(screening_id),
        s_membersno_q %in% c(2)) %>%
summarise(., mean = mean(h11_rooms_s), n=n())

## Scenario 1: Among all households with 2 household members, the mean room number
## is 1.52 ~ 2 . ( 1 observation = S20-S1281-MG)

## NA:
na_room_number_all_households <- SeroCoV_FI_dat1_reduced2 %>%
select(., screening_id, h11_rooms_s, s_membersno_q) %>%
filter(!duplicated(screening_id),
        s_membersno_q %in% c(7)) %>%
drop_na() %>%
summarise(., mean = mean(h11_rooms_s), n=n())

## Scenario 2: Among all households with 7 household members, the mean room number
## is 2.37 ~ 2 . (1 observation = S20-S0488-MG)

percentage_of_dhs_wealth_items <- dhs_wealth_items %>%
select(everything(), -screening_id, -test_result) %>%
pivot_longer(cols = everything(), names_to = "variable") %>%
group_by(variable, value) %>%
summarise(cnt = n()) %>%
mutate(percent = round(cnt / sum(cnt), 3)*100) %>%
filter(value %in% c(1) &
        (percent >= 5 & percent <= 95)
        ) %>%
arrange(variable)

```

---

**## Table 12 -- Wealth Index Assets**

```

df_wealth_index_items <- tribble( ~ variable, ~ variable_named, ~ Wealthier, ~ Poorer,
"o01_typehouse_c", "Type of housing", "Unique courtyard/ Villa/ Other", "Apartment/
Common courtyard (concession)",
"o02_walls_c_modified", "Nature of walls", "Hard/ Semi-hard/ Improved brick", "Brick/
Straw",
"o03_floor_c_modified", "Nature of floor", "Tiles/ Cement/ Finished floor", "Stamped earth/
Sand/ Natural soil/ Rough floor",
"h16_toilet_c_modified", "Type of toilets", "Flush/ Ventilated latrine", "Ordinary latrine/ none",
"h18_lightenergy_c_modified", "Main source of lighting energy", "Electricity/ Gas/ Solar
lamp", "Storm lamp/ Accumulator with batteries/ Candles/ Other",
"h19_waterrainys_c_modified", "Source of water in rainy season", "Tab water/ Protected
borehole", "Ordinary well/ Rainwater/ River, lake, dam",
"h20_waterdrys_c_modified", "Source of water in dry season", "Tab water/ Protected borehole",
"Ordinary well/ Rainwater/ River, lake, dam",
"h21_amenradio_d", "Radio", "Yes", "No",
"h22_amentv_d", "TV", "Yes", "No",

```

```

        "h23_amenplayer_d", "Video/CD/DVD Player", "Yes", "No",
        "h25_amenmobile__cell_d", "Cell phone", "Yes", "No",
        "h25_amenmobile__smart_d", "Smart phone", "Yes", "No",
        "h26_amenfridgefreezer_d", "Fridge/ Freezer", "Yes", "No",
        "h31_amendsolarlamp_d", "Solar lamp", "Yes", "No",
        "h32_amendoillamp_d", "Oil lamp", "Yes", "No",
        "h37_amendbicycle_d", "Bicycle", "Yes", "No",
        "h39_laptop_d", "Laptop", "Yes", "No",
        "h40_bednets_d", "Bed nets", "Yes", "No",
        "crowding_1", "Crowding", "5 or less people per sleeping room", "6 or more people per
sleeping room") %>%
mutate(n_row = 1:n())

gt_wealth_index_items <- df_wealth_index_items %>%
select(variable_named, Wealthier, Poorer) %>%
gt() %>%
cols_label(variable_named = md("**Asset**"),
            Wealthier = md("**Wealthier**"),
            Poorer = md("**Poorer**")) %>%
tab_header(title = md("**Table 12.** DHS Wealth Index Assets")) %>%
tab_source_note(
  source_note = md("For PCA, assets were assigned into wealthier = 1 and poorer = 0. <br>
Assets with frequencies between 5% and 95% among all households were included in the analysis.")
) %>%
tab_source_note(
  source_note = md("Reference: Hjelm, Lisa, et al. Creation of a wealth index. World Food Program (2017). <br>
*[https://www.wfp.org/publications/creation-wealth-index-june-2017]*")
) %>%
tab_row_group(
  label = "Housing quality",
  rows = c(1:3, 19)
) %>%
tab_row_group(
  label = "Access to services",
  rows = c(4,6,7,5, 15, 14)
) %>%
tab_row_group(
  label = "Consumer durables",
  rows = c(8:13, 16:18)
) %>%
row_group_order(groups = c("Housing quality", "Access to services",
"Consumer durables")) %>%
tab_style(
  cell_text(weight = "bold"),
  locations = cells_row_groups()) %>%
opt_row_stripping() %>%
opt_table_font(
  font = "Times New Roman") %>%
opt_align_table_header(align = "left")

gt_wealth_index_items %>%
gtsave(filename = "gt_wealth_index_items.docx")

```

---

### ## Figure 16 – Correlogram

```

dhs_wealth_items_named <- dhs_wealth_items %>%
select(o01_typehouse_c_modified,
       o02_walls_c_modified,
       o03_floor_c_modified,
       h16_toilet_c_modified,
       h18_lightenergy_c_modified,
       h19_waterrainys_c_modified,
       h20_waterdrys_c_modified,
       h21_amenradio_d,
       h22_amentv_d,
       h23_amenplayer_d,
       h23_1_satelittedecoder_d,

```

```

h25_amenmobile__cell_d,
h25_amenmobile__smart_d,
h26_amenfridgefreezer_d,
h31_amendsolarlamp_d,
h32_amendoillamp_d,
h37_amendbicycle_d,
h39_laptop_d,
h40_bednets_d,
crowding_1_modified,
screening_id,
test_result) %>%
rename("Type of housing" = o01_typehouse_c_modified) %>%
rename("Nature of walls" = o02_walls_c_modified) %>%
rename("Nature of floor" = o03_floor_c_modified) %>%
rename("Type of toilets" = h16_toilet_c_modified) %>%
rename("Main source of lighting energy" = h18_lightenergy_c_modified) %>%
rename("Source of water in rainy season" = h19_waterrainys_c_modified) %>%
rename("Source of water in dry season" = h20_waterdrys_c_modified) %>%
rename("Radio" = h21_amenradio_d) %>%
rename("TV" = h22_amentv_d) %>%
rename("Video/CD/DVD Player" = h23_amenplayer_d) %>%
rename("Satellite decoder" = h23_1_satelittedecoder_d) %>%
rename("Cell phone" = h25_amenmobile__cell_d) %>%
rename("Smartphone" = h25_amenmobile__smart_d) %>%
rename("Fridge/ Freezer" = h26_amenfridgefreezer_d) %>%
rename("Solar lamp" = h31_amendsolarlamp_d) %>%
rename("Oil lamp" = h32_amendoillamp_d) %>%
rename("Bicycle" = h37_amendbicycle_d) %>%
rename("Laptop" = h39_laptop_d) %>%
rename("Bed nets" = h40_bednets_d) %>%
rename("Crowding" = crowding_1_modified)

corplot_dhs_wealth_items <- round(cor(dhs_wealth_items_named[1:20]), 1)

gg_wealth_index_corrplot_hh <- corplot_dhs_wealth_items %>%
ggcorrplot(method = "square",
  hc.order = TRUE,
  lab = TRUE,
  type = "upper",
  outline.color = "white",
  legend.title = "Correlation",
  colors = c("red", "white", "blue")) +
  theme(text = element_text(family = "Times New Roman"))
)
ggsave("gg_wealth_index_corrplot_hh.png", gg_wealth_index_corrplot_hh)

```

---

## ## Figure 17 -- Scree plot

## and preparation of Table 13

## Scree plot calculated with psych® package (default scree plot not presented)

```

scree_plot <- scree(dhs_wealth_items_named[1:20],
  factors = FALSE)

```

## Presented scree plot

```

scree_plot_modified_1 <- scree_plot$pcv %>%
  as_tibble() %>%
  rownames_to_column() %>%
  mutate(PC = str_replace_all(rowname, pattern = "\\d", "PC \\1")) %>%
  mutate(PC = as_factor(PC)) %>%
  arrange(PC)
kaiser <- "Kaiser Criterion: Eigenvalues^2 > 1"

gg_scree_plot <- scree_plot_modified_1 %>%
ggplot(mapping = aes(x = PC, y = value, group = 1)) +
  geom_point(shape = 1,
    size = 3) +

```

```

geom_line() +
geom_hline(yintercept= 1,
           linetype="dashed",
           color = "red",
           linewidth=0.8) +
annotate("text", x
         = "PC 16",
         y=1.1,
         label= kaiser) +
theme(text = element_text(family = "Times New Roman"),
      panel.background = element_rect(fill = "white",
                                       colour = "grey50")
      ) +
labs(
  x = NULL,
  y = "Eigenvalues")

## Table 13 preparation
gt_eigenvalue <- scree_plot$pcv %>%
  as_tibble() %>%
  rownames_to_column() %>%
  mutate(PC = str_replace_all(rowname, pattern = "^(\d)", "PC \|1"),
         eigenvalue = round(value, digits = 3)
         ) %>%
  select(PC, eigenvalue) %>%
  filter(eigenvalue > 1) %>%
  rename("Eigenvalue" = eigenvalue,
        "Principal Component" = PC) %>%
  gt() %>%
  tab_style(
    locations = cells_column_labels(columns = everything()),
    style = cell_text(weight = "bold")
  ) %>%
  fmt_markdown(columns = everything()) %>%
  tab_header(title = md("***Table 13.** Eigenvalues of PC1 to PC6")) %>%
  tab_footnote(
    locations = cells_column_labels(columns = "Eigenvalue"),
    footnote = "The eigenvalues indicate their contribution to the variance of the 19 variables."
  ) %>%
  tab_source_note(md("Eigenvalues were calculated with R® package psych")) %>%
  opt_row_stripping() %>%
  opt_table_font(
    font = "Times New Roman") %>%
  opt_align_table_header(align = "left") %>%
  tab_style(
    style = cell_text(indent = "25px"),
    locations = cells_body(
      columns = 1,
      rows = c(1:6)))

## Scree plot:
gg_scree_plot
ggsave("gg_scree_plot.png", gg_scree_plot)

```

---

### ## Table 13 -- Eigenvalues of PC1 to PC6

```

## Presentation of Tab. 13
gt_eigenvalue %>%
  gtsave(filename = "gt_eigenvalue.docx")

```

---

### ## PCA Calculation with psych® package

```

## varimax-rotated PCA
psy_package_pca_only_hh <- psych::principal(dhs_wealth_items_named[1:20],
                                           rotate="varimax",
                                           nfactors=6,

```

```

covar=T,
scores=TRUE)

pca_values_hh <- psy_package_pca_only_hh$values
pca_loadings_hh <- psy_package_pca_only_hh$loadings

index4 <- psy_package_pca_only_hh$scores[,1]

nlab <- c(1,2,3,4,5)

dhs_wealth_items_named_with_quintiles <- dhs_wealth_items_named %>%
  mutate(., quintile = as.factor(cut(index4, breaks = 5, labels = nlab)))

dhs_wealth_index_with_quintiles_named <- dhs_wealth_items_named_with_quintiles %>%
  select(everything()) %>%
  mutate(., quintile = case_when(
    quintile %in% c(1) ~ "lowest",
    quintile %in% c(2) ~ "second",
    quintile %in% c(3) ~ "middle",
    quintile %in% c(4) ~ "fourth",
    quintile %in% c(5) ~ "highest"),
    quintile = factor(quintile,
      levels = c("highest",
        "fourth",
        "middle",
        "second",
        "lowest")))

household_wealth_index <- dhs_wealth_index_with_quintiles_named %>%
  group_by(quintile) %>%
  summarise(number = n())

```

---

#### ## Table 14 -- Results of the varimax-rotated PCA

```

gt_vaccounted <- psy_package_pca_only_hh$Vaccounted %>%
  as.data.frame() %>%
  rownames_to_column() %>%
  as_tibble() %>%
  mutate(across(c(2:7), ~ round(.x, digits = 2)),
    rowname = str_replace_all(rowname,
      pattern = "Var",
      replacement = "Variance")) %>%
  gt()

gt_vaccounted %>%
  tab_header(title = md("***Table 13.** Results of the varimax-rotated PCA")) %>%
  tab_footnote(
    footnote = "Sum of squared loadings",
    locations = cells_stub(rows = c(1))
  ) %>%
  tab_source_note(md("Varimax-rotated Principal Component Analysis performed with R@ package psych")) %>%
  gtsave(filename = "gt_vaccounted.docx")

```

---

#### ## Figure 18.1 -- Score Plot

```

df_score_data <- psy_package_pca_only_hh$scores[,1:2] %>%
  as_tibble() %>%
  mutate(n_row = 1:n())

gg_score_plot_wi <- df_score_data %>%
  ggplot(aes(x = RC1, y = RC2)) +
  geom_point(shape = 21) +
  geom_text(aes(label = n_row), color = "blue",
    size = 4,
    hjust = -0.3,
    vjust = -0.2) +

```

```

xlim(-1.25,1) +
ylim(-0.5, 1.75) +
geom_hline(yintercept= 0,
  #linetype="dashed",
  color = "green",
  linewidth=0.8) +
geom_vline(xintercept = 0,
  #linetype="dashed",
  color = "blue",
  linewidth=0.8) +
theme(text = element_text(family = "Times New Roman"),
  panel.background = element_rect(fill = "white",
  colour = "grey50")
)

ggsave("gg_score_plot_wi.png", gg_score_plot_wi)

```

---

### ## Figure 18.2 -- Loading Plot

```

df_loading_data <- psy_package_pca_only_hh$loadings[,1:2] %>%
as.data.frame() %>%
rownames_to_column(., var = "asset")

gg_loading_plot_wi <- df_loading_data %>%
ggplot(aes(x = RC1, y = RC2)) +
geom_point(shape = 2,
  size = 3) +
geom_text_repel(aes(label = asset), color = "red") +
theme(text = element_text(family = "Times New Roman"),
  panel.background = element_rect(fill = "white",
  colour = "grey50")
) +
geom_hline(yintercept= 0,
  #linetype="dashed",
  color = "green",
  linewidth=0.8) +
geom_vline(xintercept = 0,
  #linetype="dashed",
  color = "blue",
  linewidth=0.8)

ggsave("gg_loading_plot_wi.png", gg_loading_plot_wi)

```

---

### ## Figure 18.3 -- Biplot

```

pca_sd <- sqrt(psy_package_pca_only_hh$values)

gg_biplot_wi <- ggplot() +
geom_point(data = df_score_data, aes(x = RC1, y = RC2),
  shape = 21) +
geom_segment(data = df_loading_data, mapping = aes(x = 0, y = 0,
  xend = RC1 * pca_sd[1],
  yend = RC2 * pca_sd[2]),
  arrow = arrow(length = unit(0.15, "cm")), color = "red") +
geom_text_repel(data = df_loading_data, mapping = aes(x = RC1 * pca_sd[1],
  y = RC2 * pca_sd[2],
  label = asset),
  color = "red") +
geom_hline(yintercept = 0,
  color = "green",
  linewidth=0.8) +
geom_vline(xintercept = 0,
  color = "blue",
  linewidth=0.8) +
theme(text = element_text(family = "Times New Roman"),

```

```

        panel.background = element_rect(fill = "white",
                                         colour = "grey50")
    )
ggsave("gg_biplot_wi.png", gg_biplot_wi)

```

---

**## Table 15 -- Wealth Index Quintiles by ELISA Test Result**

```

WI_Test_1_hh <- dhs_wealth_index_with_quintiles_named %>%
  select(quintile) %>%
  tbl_summary(percent = "column",
              digits = NULL,
              label = list(quintile ~ "Wealth quintiles")) %>%
  bold_labels() %>%
  add_n(., statistic = "{n}")

WI_TEST_2_hh <- dhs_wealth_index_with_quintiles_named %>%
  select(test_result, quintile) %>%
  tbl_summary(by = "test_result",
              percent = "row",
              digits = NULL,
              label = list(quintile ~ "Wealth quintiles")) %>%
  add_ci(statistic = list(all_categorical() ~
                          "[{conf.low}; {conf.high}]")
        )

gtsum_WI_test_merged_hh <- tbl_merge(tbls = list(WI_Test_1_hh, WI_TEST_2_hh)) %>%
  modify_header(stat_0_1 ~ ") %>%
  modify_spanning_header((c(stat_1_2, ci_stat_1_2,
                             stat_2_2, ci_stat_2_2) ~
                           "***ELISA Test Result**"),
                        (c(n_1, stat_0_1) ~ NA)) %>%
  modify_footnote(stat_0_1 ~ NA)

gt_WI_test_merged_hh <- gtsum_WI_test_merged_hh %>%
  as_gt()

gt_WI_test_merged_hh %>%
  tab_header(title = md("***Table 15.** Wealth Index Quintiles by ELISA Test Result")) %>%
  gtsave(filename = "gt_WI_test_merged_hh.docx")

```

---

**## Table 16 -- Possession of Wealth Index Assets among Included Households**

```

gtsum_wealth_index_assets_per_quintile_hh <- dhs_wealth_index_with_quintiles_named %>%
  select(c(1:19), quintile) %>%
  mutate(quintile = factor(quintile,
                           levels = c("lowest",
                                       "second",
                                       "middle",
                                       "fourth",
                                       "highest"))) %>%
  tbl_summary(by = "quintile",
              percent = "column"
            ) %>%
  add_overall() %>%
  modify_header(label = "***Wealth Index Assets**",
                stat_0 = "***Improved household conditions**", <br> N = 674"
            ) %>%
  modify_spanning_header(c(stat_1: stat_5) ~ "***Wealth Index Quintiles**")
gt_wealth_index_assets_per_quintile_hh <- gtsum_wealth_index_assets_per_quintile_hh %>%
  as_gt() %>%
  tab_header(md("***Table 16.** Possession of Wealth Index Assets among Included Households"))

gt_wealth_index_assets_per_quintile_hh %>%
  gtsave(filename = "gt_wealth_index_assets_per_quintile_hh.docx")

```



## 12.4 Appendix D: Literature search terms

PubMed Search:

### 1. **SARS-CoV-2 in Madagascar in general**

Keywords:

("SARS-CoV-2" OR "COVID-19") AND "Madagascar"

MeSH terms:

((("SARS-CoV-2"[Mesh]) OR ("COVID-19"[Mesh])) AND ("Madagascar"[Mesh]))

### 2. **SARS-CoV-2 seroprevalence in sub-Saharan Africa**

Keywords:

("SARS CoV-2" OR "COVID-19") AND "seroprevalence" AND "sub-Saharan Africa";  
("SARS-CoV-2" OR "COVID-19") AND "prevalence" AND "Antibod\*" AND "sub-Saharan Africa"

MeSH Terms:

"Africa South of the Sahara"[Mesh]; "Madagascar"[Mesh]; "COVID-19"[Mesh];  
"SARS-CoV-2"[Mesh]; "Seroepidemiologic Studies"[Mesh]; "Cross-Sectional Studies"[Mesh]; "Antibodies, Viral"[Mesh]

Combination:

("Africa South of the Sahara"[Mesh] OR "Madagascar"[Mesh]) AND ("COVID-19"[Mesh] OR "SARS-CoV-2"[Mesh]) AND ("Seroepidemiologic Studies"[Mesh] OR "Cross-Sectional Studies"[Mesh]) AND "Antibodies, Viral"[Mesh];  
"Madagascar"[Mesh] AND ("COVID-19"[Mesh] OR "SARS-CoV-2"[Mesh]) AND ("Seroepidemiologic Studies"[Mesh] OR "Cross-Sectional Studies"[Mesh]) AND "Antibodies, Viral"[Mesh]

### 3. **Wealth Index in Madagascar and sub-Saharan Africa**

Keywords:

Madagascar, sub-Saharan Africa, DHS, Wealth Index, socioeconomic score, urban  
("Madagascar" OR "sub-Saharan Africa") AND ("DHS" OR "Wealth Index" OR "socioeconomic score") AND "urban"

MeSH terms:

(Focus on Madagascar) ("Socioeconomic Factors"[MeSH]) AND "Madagascar"[MeSH];  
("Socioeconomic Factors"[MeSH]) AND "Madagascar"[MeSH] AND "Urban Population"[Mesh]

### 4. **Depression in Madagascar**

Keywords: Madagascar, Depression, Mental Health

"Madagascar" AND ("Depression" OR "Mental Health")

MeSH terms:

"Madagascar"[Mesh] AND ("Mental Health"[Mesh] OR "Depression"[Mesh])

### 5. **Alcohol consumption in SSA or Madagascar:**

Keywords:

Madagascar, Sub-Saharan Africa, Alcohol

("Madagascar" OR "sub-Saharan Africa") AND "Alcohol" -> many results

"Madagascar" AND "Alcohol" -> scarce results

MeSH terms:

("Madagascar"[Mesh] AND "Alcohol-Related Disorders"[Mesh])

("Madagascar"[Mesh]) AND "Substance-Related Disorders"[Mesh]

## 6. **Food insecurity**

### Keywords:

Food insecurity, COVID-19, SARS-CoV-2, sub-Saharan Africa, Madagascar  
("Madagascar" OR "sub-Saharan Africa") AND "Food Insecurity" AND ("COVID-19"  
OR "SARS-CoV-2")

### MeSH terms:

("Food Insecurity"[Mesh] OR "Malnutrition"[Mesh]) AND "Madagascar"[Mesh] AND  
("COVID-19"[Mesh] OR "SARS-CoV-2"[Mesh])

## 7. **Knowledge of COVID-19**

### Keywords:

COVID-19, Knowledge, Madagascar [ in google scholar]

COVID-19, Knowledge, sub-saharan Africa (pubmed)

"COVID-19" AND "Knowledge" AND "sub-saharan Africa" -> many results

"COVID-19" AND "Knowledge" AND "Madagascar" -> few results

### MeSH terms:

("Health Knowledge, Attitudes, Practice"[Mesh]) AND (("COVID-19"[Mesh]) OR  
("SARS-CoV-2"[Mesh])) AND "Africa South of the Sahara"[Mesh]

## 8. **Social distancing by age group**

### Keywords:

social distancing, covid 19, age groups, sub-Saharan Africa

"Social Distancing" AND "COVID-19" AND "sub-Saharan Africa"

### MeSH terms:

("Physical Distancing"[Mesh]) AND (("COVID-19"[Mesh]) OR ("SARS-CoV-2"  
[Mesh])) AND ("Age Groups"[Mesh]) AND ("Africa South of the Sahara"[Mesh])

## 9. **Number of publications by countries' income on COVID-19**

### Keywords:

Number of publications by countries' income on COVID-19 [in google search]

## 10. **COVID-19 Characteristics**

### Keywords:

clinical characteristic\* covid-19

"Clinical Characteristic\*" AND "COVID-19"

### **13. Eidesstattliche Versicherung**

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: .....