



# Morbidity and mortality burden of COVID-19 in rural Madagascar: results from a longitudinal cohort and nested seroprevalence study

Andres Garchitorena, Lova Tsikiniaina Rasoloharimanana, Rado JI Rakotonanahary, Michelle V Evans, Ann C Miller, Karen E Finnegan, Laura F Cordier, Giovanna Cowley, Benedicte Razafinjato, Marius Randriamanambintsoa, et al.

## ► To cite this version:

Andres Garchitorena, Lova Tsikiniaina Rasoloharimanana, Rado JI Rakotonanahary, Michelle V Evans, Ann C Miller, et al.. Morbidity and mortality burden of COVID-19 in rural Madagascar: results from a longitudinal cohort and nested seroprevalence study. *International Journal of Epidemiology*, inPress, 10.1093/ije/dyad135 . hal-04330452

**HAL Id: hal-04330452**

**<https://hal.science/hal-04330452>**

Submitted on 11 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Morbidity and mortality burden of COVID-19 in rural Madagascar: results from a longitudinal cohort and nested seroprevalence study

Andres Garchitorena<sup>1,2,3\*</sup>, Lova Tsikiniaina Rasoloharimanana<sup>2</sup>, Rado JL Rakotonanahary<sup>3,4</sup>, Michelle V Evans<sup>1</sup>, Ann C. Miller<sup>4</sup>, Karen E. Finnegan<sup>3,4</sup>, Laura F. Cordier<sup>3</sup>, Giovanna Cowley<sup>3</sup>, Benedicte Razafinjato<sup>3</sup>, Marius Randriamanambintsoa<sup>5</sup>, Samuel Andrianambinina<sup>5</sup>, Stephen J Popper<sup>6</sup>, Raphaël Hotahiene<sup>7</sup>, Matthew H. Bonds<sup>3,4†</sup>, Matthieu Schoenhals<sup>2†</sup>

<sup>1</sup> MIVEGEC, Université de Montpellier, CNRS, IRD, Montpellier, France

<sup>2</sup> Institut Pasteur de Madagascar, Antananarivo, Madagascar

<sup>3</sup> NGO Pivot, Ifanadiana, Madagascar

<sup>4</sup> Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA

<sup>5</sup> Direction de la Démographie et des Statistiques Sociales, Institut National de la Statistique, Antananarivo, Madagascar

<sup>6</sup> Division of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, CA, USA

<sup>7</sup> Ministère de la Santé Publique, Antananarivo, Madagascar

\* Corresponding author:

Address: 911 Avenue Agropolis, 34394 Montpellier, France

Email: [andres.garchitorena@ird.fr](mailto:andres.garchitorena@ird.fr)

† Equal contribution

Word count: 4,138

## Abstract

**Introduction:** Three years into the pandemic, there remains significant uncertainty about the true infection and mortality burden of COVID-19 in the World Health Organization Africa region. High quality, population-representative studies in Africa are rare and tend to be conducted in national capitals or large cities, leaving a substantial gap in our understanding of the impact of COVID-19 in rural, low-resource settings. Here, we estimated the spatio-temporal morbidity and mortality burden associated with COVID-19 in a rural health district of Madagascar until the first half of 2021.

**Methods:** We integrated a nested seroprevalence study within a pre-existing longitudinal cohort conducted in a representative sample of 1600 households in Ifanadiana District, Madagascar. Socio-demographic and health information was collected in combination with dried blood spots for about 6500 individuals of all ages, which were analysed to detect IgG and IgM antibodies against four specific proteins of SARS-CoV2 in bead-based multiplex immunoassay. We evaluated spatio-temporal patterns in COVID-19 infection history and its associations with several geographic, socio-economic and demographic factors via logistic regressions.

**Results:** Eighteen percent of people had been infected by April-June 2021, with seroprevalence increasing with individuals' age. COVID-19 primarily spread along the only paved road and in major towns during the first epidemic wave, subsequently spreading along secondary roads during the second wave to more remote areas. Wealthier individuals and those with occupations such as commerce and formal employment were at higher risk of being infected in the first wave. Adult mortality increased in 2020, particularly for older men for whom it nearly doubled up to nearly 40 deaths per 1000. Less than 10% of mortality in this period would be directly attributed to COVID-19 deaths if known infection fatality ratios are applied to observed seroprevalence in the district.

**Conclusion:** Our study provides a very granular understanding on COVID-19 transmission and mortality in a rural population of sub-Saharan Africa and suggests that the disease burden in these areas may have been substantially underestimated.

**Keywords:** SARS-CoV-2 seroprevalence; COVID-19 risk factors; COVID-19 mortality; sub-Saharan Africa

## Key messages

- Three years after the start of the COVID-19 pandemic, the disease burden in areas that are traditionally most vulnerable - rural populations in the developing world - is still unclear.
- Despite low density and connectivity, about 1 in 5 people were infected by the period of April-June 2021 in a rural district of Madagascar. COVID-19 primarily spread along the transportation network and transmission was shaped by socio-economic level.
- Adult mortality substantially increased in 2020, particularly for older men, but the majority of excess mortality during this period could not be directly attributed to COVID-19 death.
- The true burden of COVID-19 in poor rural areas of sub-Saharan Africa may be larger than recognized

## Introduction

Soon after the initial outbreak of COVID-19 in Wuhan, China, scientists estimated the key epidemiological properties that determined the spread and impact of the disease <sup>1–3</sup> and, as the pandemic progressed, the role of age, gender, and comorbidities as risk factors became clearer <sup>4</sup>. However, key aspects of the epidemic remained uncertain for areas of the world where high-quality data was not available, leading to considerable debate about the expected burden in these areas <sup>5,6</sup>. This was especially pronounced in the World Health Organization (WHO) Africa region. At the onset of the pandemic, there was substantial concern that the region could be especially vulnerable for the same reasons it suffers from high burdens of other infectious diseases, low access to health care and limited capacity for treatment <sup>7,8</sup>. As the pandemic continued, low rates of reported cases and deaths corresponded to a growing chorus suggesting that perhaps Africa was at lower risk for COVID-19 morbidity and mortality due to genetic, environmental, or immunological differences <sup>9,10</sup>.

Three years on, there remains significant uncertainty about the true infection and mortality burden of COVID-19 in Africa. Low capacity for routine testing has undermined the use of government statistics for understanding COVID-19 spread. Recent seroprevalence studies have demonstrated that patterns of infection in Africa are consistent with patterns observed elsewhere, ranging between 40 and 60% positivity by mid-2021 <sup>11</sup>. In fact, sub-Saharan Africa could be the region with the highest infection rates globally <sup>12</sup>. However, high quality, representative studies in Africa continue to be rare. Studies tend to rely on opportunistic data collection, such as from blood donors <sup>13–18</sup>, healthcare workers <sup>19–24</sup>, or patients coming to health facilities for reasons other than COVID-19 <sup>25–30</sup>, none of which are representative of the general population. Where population-representative studies do exist, they tend to be conducted in national capitals or large cities <sup>11</sup>.

A second challenge is estimating the mortality burden associated with COVID-19 in Africa from existing sources. Indeed, the limited availability of vital registration records means that deaths are not systematically reported <sup>31</sup>, and higher rates of background mortality in these populations can obscure the characterization of excess deaths from COVID-19 <sup>32</sup>. In addition, the indirect death toll of an epidemic can be substantial due to health system disruptions, lower access to healthcare and other factors, as observed during the 2014–2015 Ebola epidemic <sup>33,34</sup>. Consequently, estimates of the COVID-19 mortality burden for the African Region vary widely: while the WHO estimates that less than 500,000 died from COVID-19 in 2020–2021 <sup>35</sup>, other studies suggest this number could be three to four times higher <sup>12,36,37</sup>. All these challenges are exacerbated in poor, rural areas of sub-Saharan Africa, where there is virtually no population-representative information on COVID-19.

Here, we take advantage of a pre-existing longitudinal cohort study in a representative sample of a rural district of Madagascar to estimate the morbidity and mortality burden associated with COVID-19 in this population until April-June of 2021. We evaluated patterns in SARS-CoV-2 infection history across space and time in the context of several socio-economic and demographic factors. This resulted in the most granular data on COVID-19 for a rural population of the WHO Africa Region that we are aware of.

## Methods

### Survey data collection and serological analyses

The study was conducted in Ifanadiana, a rural health district of approximately 200,000 people in South-eastern Madagascar (see the Supplementary material, Section S1 for details). A seroprevalence survey was added to an existing longitudinal cohort study initiated in 2014 (the Ifanadiana Health Outcomes and Prosperity longitudinal Evaluation, or IHOPE) <sup>38</sup> to obtain demographic, health and socio-economic information from a representative sample of 1,600 households in Ifanadiana District over time. Questionnaires in the cohort were mostly adapted from the Demographic and Health Survey <sup>39</sup>. The Madagascar National Institute of Statistics (INSTAT) was responsible for data collection, survey coordination, training and oversight. A two-stage sample stratified the district by the HSS intervention's initial catchment area. Eighty clusters, half from each stratum, were selected at random from enumeration areas mapped during the 2009 census, and households were then mapped within each cluster. Twenty households were selected at random from each cluster.

Four waves of data collection have been conducted in 2014, 2016, 2018 and 2021, in which the original 1600 households were revisited. Response rates were about 95% for each wave <sup>38</sup>. Individual face-to-face interviews were conducted with all women aged 15 to 49 years and men aged 15 to 59 years (usual residents or visitors). Data collected in the questionnaires included, among others, household composition (size, genders, age); indicators of socio-economic status (education, employment, household durable assets); adult, maternal and child mortality. To learn more about the impact of COVID-19, the 2021 wave of data collection (April 22<sup>nd</sup> to June 20<sup>th</sup>) included, for each household member, questions on COVID-related symptoms in the previous 6 months. For all consenting individuals of all ages, a dried blood spot (DBS) was obtained by finger prick using a single-use lancet needle by trained nurses, with 1–5 DBS collected on Whatman 903 Protein Caver Card filter papers. To ensure the safety and avoid transmission of COVID-19 during the survey, field protocols were adapted based on guidance from the SMART initiative <sup>40</sup>. This included frequent testing of survey teams before, during and after the survey; initial quarantine of all teams on site before beginning the

survey; immediate quarantine of any COVID-positive staff from the survey; use of masks and other protective equipment for survey teams and participants during interviews.

Using methods previously described <sup>41</sup>, DBS samples were processed using a multiplex bead assay on the Luminex platform (Magpix™) for antibodies against 4 antigens of SARS-CoV2: Spike S1, Spike S2, Spike RBD and NP. Magnetic beads (Luminex, Magplex™ MC100XX-01) were coupled to these four antigens using xMAP Antibody coupling kit (Luminex, 40-50016). Cut-off limits for determining positive antibodies for SARS-CoV-2 were estimated based on receiver operator characteristics (ROC) for the median fluorescence intensity. Further details on serological analyses are available in the Supplementary material, Section S1.

### Data analyses

#### *Seroprevalence of recent and past SARS-CoV2 infections*

To obtain seroprevalence estimates, we carried out two sets of analyses. First, normalised and protein-corrected values of mean fluorescence intensity (MFI) for each of the eight SARS-CoV2 antibodies were compared with their corresponding positivity threshold <sup>42</sup> to determine whether the sample was positive for that particular antibody. Second, to reduce the number of dimensions of the serological data and obtain discrete consistent groups we used k-means clustering to classify individuals' seropositivity. Data were visually inspected for outliers, and 134 individuals (outside the range of mean + 3 sd) were removed. We preconditioned the data via a principal components analysis (PCA) to reduce its dimensionality <sup>43</sup>, and used the first two principal components, which explained over 60% of the variance, in the subsequent cluster analysis. We then estimated the optimal number of clusters via the average silhouette width. We used the Hartigan-Wong algorithm to perform k-means clustering using 999 starting sets of centroids.

Clusters were assigned meaningful sero-positivity labels based on the component loadings of the initial PCA and the clusters' locations relative to these loadings (Figure 1). IgG and IgM antibodies had a strong positive effect on the first principal component (PC1), representing the gradient between overall positive and negative sero-positivity. The second principal component had strong positive loadings for IgM antibodies, which are indicative of recent infection <sup>44</sup>, and therefore represented the difference between past and recent infections.

#### *Trends and factors associated with SARS-CoV2 seroprevalence*

Statistical analyses were carried out to understand local patterns of COVID-19 seroprevalence, including geographic trends and associations with socio-economic and demographic factors (see the Supplementary material Section S1 for details). Associations were modelled individually using

univariate and multivariate logistic regressions in generalized linear models. Sampling weights that adjusted for unequal probability of selection due to stratification and non-response were calculated. Estimates (totals, proportions, odds ratios) were obtained using survey commands available in R-package *survey* and applicable sampling weights <sup>45</sup>. All analyses were done for each individual serological marker and for the composite indicators. All analyses were performed with R software, version 4.2.1 <sup>46</sup>, and R packages “survey”, “gstat”, “rgdal”, “stats”, and “ggplot2”.

#### *Excess mortality and infection fatality rates associated with SARS-CoV2*

Adult mortality was estimated from the IHOPE cohort using the synthetic life-table method for DHS surveys <sup>47</sup>. First, six-year averages of adult mortality per 1000 population, split by age group and sex, were estimated for each wave of the cohort in order to obtain robust estimates over time that are comparable to standard DHS methods, before and after the COVID-19 epidemic. Because the period for these six-year estimates overlaps, mortality rates per year were then estimated from the 2021 wave of the cohort only, for the ten years prior to the survey. From this, a ten-year average was estimated (2012-2021) and excess mortality occurring in the years 2020 and 2021 was estimated as the difference between these years' mortality and the ten-year average. Not all excess mortality in 2020-2021 can be assumed to be directly the result of COVID-19 deaths. To estimate expected excess mortality associated directly with SARS-CoV2 infections in our population, infection fatality rate (IFR) estimates per year of age (including lower and upper bounds for these estimates) were obtained from a recent study by the COVID-19 Forecasting Team <sup>48</sup>. These age-specific IFR values were then combined with the observed age-specific number of SARS-CoV2 cases and age distribution in our population to obtain an expected excess mortality by age group. Observed excess mortality in our cohort was then compared with expected excess mortality.

## Results

### SARS-CoV-2 seroprevalence trends and associated factors

Overall, 6496 individuals were included in the seroprevalence analyses, nearly half of whom were children under 15 years (Table 1). Only one out of ten individuals reported COVID-related symptoms in the 6 months prior to the survey (Table 2), with the most common symptoms being fever (6.5%) and respiratory problems (4.1%). District seroprevalence ranged from 5.1% (Spike RBD) to 43.8% (Spike S2) for IgG antibodies, and from 7.2% (Spike S1) to 17.9% (Spike S2) for IgM antibodies. Positivity to Spike S2 was higher than to any of the other markers, both for IgG and IgM antibodies. After clustering via PCA and k-means, overall seroprevalence was estimated to be 18%, with 10.1% having a predominantly IgG response suggestive of past infection and 7.9% having a predominantly IgM response suggestive of recent infection.

The spatial distribution of past infections suggests that COVID-19 cases during the first wave accumulated predominantly in proximity to the paved road, with the exception of a few clusters in remote areas in the north of the district where seroprevalence reached nearly 30% (Figure 2). In contrast, recent infections during the second wave were more evenly distributed, with lower prevalence in clusters located along the main road. Similarly, the factors associated with SARS-CoV-2 seroprevalence varied substantially between past and recent infections (Figure 3 and Table 3). In both recent and past infections, seroprevalence was similar for males and females, increased with age and decreased with distance to a major town. However, seroprevalence for past infections was higher for individuals in the wealthiest household quantile and those whose occupation was not agriculture, with opposite associations for recent infections. In addition, seroprevalence of recent infections doubled for individuals reporting a COVID-19 symptom in the previous 6 months compared to those not reporting symptoms, while no association was observed for past infections. Spatial distributions and associations for each of the eight SARS-CoV-2 serological markers are available in the Supplementary material, Section S2.

#### Mortality and infection fatality rates associated with COVID-19

Estimates of 6-year adult mortality in the 2021 wave of the IHOPE cohort collection were higher than in all previous waves of data collection (Figure 4). The largest increase was observed for older men (35-49 years), who experienced a nearly 100% increase in mortality in reports from the 2021 cohort wave as compared to the 2018 cohort wave (from 10.6 to 19.3 per 1000). For women and young men, mortality had been declining in the previous three waves, and then increased by 20-50% in the 2021 cohort wave to a level higher than baseline. Similar trends were observed using data from the 2021 cohort wave only, where analyses of annual mortality rates showed that mortality in 2020 was substantially higher than average, especially for older men (Figure 4). However, these time series were more stochastic given lower sample sizes and, for women and young men, the peak in mortality observed in 2020 was not higher than other peaks observed in previous years. Overall, excess mortality for 2020-2021 was estimated at 1.61 per 1000 for individuals 15-34 years and 4.82 per 1000 for individuals 35-49 years. Given previously estimated COVID-19 IFRs<sup>48</sup> and observed seroprevalence by age group, expected excess mortality associated directly with COVID-19 infections in our cohort would be 0.05 (range 0.04-0.08) per 1000 for individuals 15-34 years and 0.54 (range 0.39-0.85) per 1000 for individuals 35-49 years (Table 4). This suggests that if IFRs in Ifanadiana were consistent with those previously estimated, only 3.1% (range 2.5-5.0%) and 8.9% (range 6.5-14.1%) of observed excess mortality for individuals 15-34 years and 35-49 years, respectively, would be directly associated with COVID-19 deaths.

## Discussion

Nearly three years after the start of the COVID-19 pandemic, the burden of the disease in areas that are traditionally most vulnerable - rural populations in the developing world - is still unclear. SARS-CoV-2 seroprevalence surveys have been essential for understanding COVID-19 transmission, but quality studies have rarely focused on rural areas of sub-Saharan Africa <sup>11,48</sup>. Using a population-representative cohort and a nested seroprevalence survey of nearly 6500 people of all ages, we provide a fine-scale account of COVID-19 spread and burden in a rural district of Madagascar during its first two epidemic waves. Our results suggest that despite low density and connectivity in the majority of the district, about one in five people had been infected by April-June 2021. COVID-19 primarily spread along the only paved road during the first wave, and then spread along secondary roads during the second wave to more remote areas. Adult mortality increased in 2020, particularly for older men, but the majority of excess mortality during this period could not be directly attributed to COVID-19 deaths given previously estimated IFRs. This suggests that for populations living in rural, low-resource settings, COVID-19 could have significant health impacts, either because of higher IFRs than previously estimated or because of substantial indirect impacts on health care.

Our results reveal that seroprevalence was lower in this rural district of Madagascar than in the nearby city of Fianarantsoa (1h30 drive), where seroprevalence was about 20% by November 2020 <sup>14</sup> and over 40% by February-June 2021 <sup>49</sup>. This is consistent with studies conducted in other African settings, which found that seroprevalence tends to be lower in rural populations <sup>11,50–56</sup>. The lower seroprevalence in Ifanadiana could also be explained by delays in epidemic spread within Madagascar. While April-June 2021 was the middle of the second epidemic wave according to national data, the vast majority of reported cases come from Antananarivo due to diagnostic challenges in the rest of the country <sup>8</sup>. Given that there was a lag of nearly two months between the first epidemic wave in Antananarivo and in other cities <sup>14</sup>, it is possible that our data collection occurred at the beginning of the second wave in Ifanadiana, which could explain the low rates of IgM seroprevalence. However, our study revealed substantial heterogeneity in the spatio-temporal patterns of these rural infections. Populations in the district living within 5km to roads or large towns had comparable seroprevalence rates to those from urban Fianarantsoa <sup>14,49</sup>, which highlights the major role played by population density and road connectivity in the spread of COVID-19 even in rural areas of the developing world, where both factors are significantly lower.

Seroprevalence rates differed across demographic and socio-economic groups in our population. Seroprevalence in both epidemic waves increased with age, especially for those older than 30, but risk did not differ between men and women. Although associations with demographic factors can be context-specific and vary across settings, a recent review of seroprevalence studies found similar average trends for Africa <sup>11</sup>. Socio-economic factors also modified individuals' risk of infection

in this largely impoverished population, where the primary occupation is subsistence agriculture. Wealthier individuals and those with certain occupations such as commerce and formal employment were at higher risk of being infected in the first wave but at lower risk of being infected in the second. It is well known that individuals with high social connectivity are at higher risk of infection and can contribute disproportionately to the spread of diseases such as COVID-19<sup>57–59</sup>. While research on at-risk occupational activities in Africa has mostly focused on healthcare workers<sup>24</sup>, a better understanding of the role played by other socio-economic groups with high mobility and social connections could open new possibilities for disease control<sup>60</sup>. The inverse relationship found from the first wave to the second for some risk factors (geographic, socio-economic) was unexpected given low seroprevalence rates. We found that rates of cluster seroprevalence during the second wave were consistently lower when seroprevalence during the first wave was 30–50% (Supplementary Figure S10), which suggests that in this rural context with a potentially lower COVID-19 effective reproduction number, a seroprevalence level below 50% could have offered some level of protection, at least in the initial epidemic progression of the second wave.

In this poor rural setting where mortality rates were already high prior to the pandemic, our results suggest that the COVID-19 epidemic was associated with a substantial increase in adult mortality. The increase observed for Ifanadiana was similar to that found in Sudan's capital<sup>61</sup>, where a 67% rise was observed. However, about three quarters of deaths in the Sudan study were among individuals 50 years or older, a vulnerable population group that was not assessed here due to study design limitations. The excess mortality in individuals aged 15–49 years in Ifanadiana was substantially higher than what could have been expected based on observed seroprevalence and known infection fatality rates for these age groups in other parts of the world<sup>48</sup>. COVID-19 is known to have indirect impacts on mortality, such as through the effects on healthcare access and health system disruptions<sup>62,63</sup>, and these effects could have been larger here than in other settings<sup>64</sup>. Interestingly, the only other outlying year for mortality in this ten-year period was 2017, when Madagascar's largest plague epidemic in recent history occurred even though no plague deaths were reported from this area<sup>65</sup>. For both 2017 and 2020, it is also possible that other factors unrelated to COVID-19 were associated with higher mortality in this period. For instance, lower precipitation in 2016 due to El Niño Southern Oscillation led to lower agricultural yields and higher rice prices in 2017, which could have impacted rural populations' nutrition and illustrates the complexity of attributing excess mortality to specific events. Importantly, the lack of seroprevalence studies in rural areas of the developing world means that global age-specific IFRs used here as reference<sup>48</sup> could be substantially underestimated for our population given higher rates of poverty and lower rates of healthcare access. It is plausible that our

population had significantly higher IFRs than global estimates and that the direct contribution of COVID-19 to excess mortality was higher than reported here.

Our study had several limitations. First, although we used robust clustering methods to classify infections into past and recent, there is uncertainty around such classification. For instance, IgM titres can remain high in past symptomatic infections and IgG NP titres can increase early in the infection<sup>41,66</sup>. However, the fact that the first epidemic wave occurred nearly a year prior and that results are consistent with known patterns of COVID-19 spread (e.g. initial spread along better connected populations) suggests that potential misclassification biases had little impact. It is also possible that lower sensitivity of the assay for IgM markers (Supplementary Table S1) could have resulted in a slight underestimation of recent infections, meaning that both the overall seroprevalence and the direct contribution of COVID-19 to excess mortality could have been higher than reported here. Second, low sample sizes for analyses of adult mortality could have impacted our estimates of excess mortality, especially because annual rates in the 10-year period were not stable and had considerable uncertainty (Supplementary Table S2). Despite this, results of these analyses were consistent with 6-year averages (which include a much larger sample size but overlap over time). Third, our survey mirrored a DHS design, where mortality estimates are based on information for siblings provided by men and women of reproductive age. As a result, even though individuals older than 50 years are the most likely to be affected, we could not assess the impact of COVID-19 on this age group due to low sample sizes. Fourth, the survey was conducted in the middle of the second wave of COVID-19, which prevented us from fully capturing the impact of this wave. This could also have affected the trends observed if the timing of the survey in different parts of the district had an impact on their corresponding seroprevalence, but complementary analyses suggest this was unlikely (Supplementary Figure S9). Fifth, fear and stigma around COVID-19 could have biased participants' responses and resulted in an underestimation of COVID-19 symptoms in the previous 6 months. For instance, rates of illness reporting in children had been decreasing in previous survey years, but this decrease was particularly acute for respiratory infections in 2021 (Supplementary Figure S11). Finally, as is the case with any local-scale survey, the results of this study do not necessarily represent the COVID-19 situation in other parts of Madagascar or sub-Saharan Africa.

In conclusion, our study provides an unusually detailed picture of COVID-19 morbidity and mortality in a poor rural setting of sub-Saharan Africa, with important implications for similar settings. It suggests that the disease burden in these areas may have been substantially underestimated. Given known vulnerabilities to other infectious diseases, combined with the fragility of their health systems, more attention and quality research is needed to better understand the true burden of COVID-19 in

poor rural areas of sub-Saharan Africa and to devise appropriate responses to this and future pandemics.

## **ETHICS APPROVAL**

The study was approved by the Madagascar National Ethics Committee and Harvard Medical School IRB, including amendments for changes in 2021. All adults ( $\geq 15$  years) provided oral informed consent for the in-person interview and written informed consent for biological sample collection. Parents or guardians provided written consent for biological sample collection from children  $< 15$  years of age, and children 7-14 years provided written assent separately.

## **DATA AVAILABILITY**

Data are available upon request to the address [research@pivotworks.org](mailto:research@pivotworks.org)

## **SUPPLEMENTARY DATA**

Supplementary data are available at IJE online

## **AUTHOR CONTRIBUTIONS**

Conceived and designed the experiments: AG, ACM, KEF, MHB, MS. Performed laboratory analyses: LTR, MS. Performed statistical analyses: AG, ME. Contributed reagents/materials/data/analysis tools: AG, RJLR, ACM, MR, SA, LTR, MS. Wrote the initial draft of the manuscript: AG, MHB. Revised the manuscript and accepted it in its final form: AG, LTR, RJLR, ME, ACM, KEF, LFC, GC, BR, MR, SA, SP, RH, MHB, MS

## **FUNDING**

This work was supported by internal funding from Pivot, a grant from the Agence Nationale de la Recherche (Project ANR-19-CE36-0001-01), a grant from the Gates Medical Research Institute (#55989) and support from the Herrnstein Family Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## **ACKNOWLEDGEMENTS**

The authors acknowledge the contributions and support of Felana Ihantamalala and Mauricianot Randriamihaja. They are grateful to all of the staff at PIVOT for their field support and their remarkable work in strengthening the health system in Ifanadiana. Thanks are due to the Madagascar Ministry of Health, at both the district and the central levels, for their continuous support and valuable insights. The authors also thank the Institut National de la Statistique (INSTAT) field teams for their involvement in the district-wide population survey.

## **COMPETING INTERESTS**

Some authors are current or former employees of institutions discussed in this article, including the non-governmental organization Pivot and the Government of Madagascar. These affiliations are explicitly listed in the article.

## References

1. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* (80- ). 2020;**368**(6493):860–868.
2. Walker PGT, Whittaker C, Watson OJ, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- And middle-income countries. *Science* (80- ). 2020;**369**(6502):413–422.
3. Ferguson N, Laydon D, Nedjati-Gilani G, et al. COVID-19 report 9. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID- 19 mortality and healthcare demand. *Imp Coll COVID-19 response team* [Internet]. 2020;(March). Available from: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/>
4. Ghisolfi S, Almås I, Sandefur JC, Carnap T von, Heitner J, Bold T. Predicted COVID-19 fatality rates based on age, sex, comorbidities and health system capacity. *BMJ Glob Heal*. 2020;**5**(9):1–8.
5. Rosenthal PJ, Breman JG, Djimde AA, et al. COVID-19: Shining the light on Africa. *Am J Trop Med Hyg*. 2020;**102**(6):1145–1148.
6. El-Sadr WM, Justman J. Africa in the Path of Covid-19. *N Engl J Med*. 2020;**383**(3):e11.
7. Kapata N, Ihekweazu C, Ntoumi F, et al. Is Africa prepared for tackling the COVID-19 (SARS-CoV-2) epidemic. Lessons from past outbreaks, ongoing pan-African public health efforts, and implications for the future. *Int J Infect Dis*. 2020;**93**:233–236.
8. Evans MV, Garchitorena A, Rakotonanahary RJL, et al. Reconciling model predictions with low reported cases of COVID-19 in Sub-Saharan Africa: insights from Madagascar. *Glob Health Action* [Internet]. 2020;**13**(1). Available from: <https://doi.org/10.1080/16549716.2020.1816044>
9. Rice BL, Annapragada A, Baker RE, et al. Variation in SARS-CoV-2 outbreaks across sub-Saharan Africa. *Nat Med* [Internet]. Springer US; 2021;**27**(3):447–453. Available from: <http://dx.doi.org/10.1038/s41591-021-01234-8>
10. Mbow M, Lell B, Jochems SP, et al. COVID-19 in Africa: Dampening the storm? *Science* (80- ) [Internet]. 2020 Aug 7;**369**(6504):624–626. Available from: <https://www.science.org/doi/10.1126/science.abd3902>
11. Lewis HC, Ware H, Whelan M, et al. SARS-CoV-2 infection in Africa: A systematic review and meta-analysis of standardised seroprevalence studies, from January 2020 to December 2021. *BMJ Glob Heal*. 2022;**7**(8):1–15.
12. COVID-19 Cumulative Infection Collaborators. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. *Lancet (London, England)*. 2022;**399**(10344):2351–2380.
13. Sebastião CS, Galangue M, Gaston C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies and risk factors among healthy blood donors in Luanda, Angola. *BMC Infect Dis* [Internet]. BioMed Central; 2021;**21**(1):1–10. Available from: <https://doi.org/10.1186/s12879-021-06814-0>
14. Schoenhals M, Rabenindrina N, Rakotondramanga JM, et al. SARS-CoV-2 antibody seroprevalence follow-up in Malagasy blood donors during the 2020 COVID-19 Epidemic.

*EBioMedicine*. 2021;**68**.

15. Razafimahatratra SL, Ndiaye MDB, Rasoloharimanana LT, et al. Seroprevalence of ancestral and Beta SARS-CoV-2 antibodies in Malagasy blood donors. *Lancet Glob Heal* [Internet]. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license; 2021;**9**(10):e1363–e1364. Available from: [http://dx.doi.org/10.1016/S2214-109X\(21\)00361-2](http://dx.doi.org/10.1016/S2214-109X(21)00361-2)
16. Mandolo J, Msefula J, Henrion MYR, et al. SARS-CoV-2 exposure in Malawian blood donors: an analysis of seroprevalence and variant dynamics between January 2020 and July 2021. *BMC Med*. BMC Medicine; 2021;**19**(1):1–10.
17. Uyoga S, Adetifa IMO, Karanja HK, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors. *Science (80- )*. 2021;**371**(6524):79–82.
18. Adetifa IMO, Uyoga S, Gitonga JN, et al. Temporal trends of SARS-CoV-2 seroprevalence during the first wave of the COVID-19 epidemic in Kenya. *Nat Commun* [Internet]. Springer US; 2021;**12**(1):1–6. Available from: <http://dx.doi.org/10.1038/s41467-021-24062-3>
19. Gudina EK, Ali S, Girma E, et al. Seroepidemiology and model-based prediction of SARS-CoV-2 in Ethiopia: longitudinal cohort study among front-line hospital workers and communities. *Lancet Glob Heal* [Internet]. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license; 2021;**9**(11):e1517–e1527. Available from: [http://dx.doi.org/10.1016/S2214-109X\(21\)00386-7](http://dx.doi.org/10.1016/S2214-109X(21)00386-7)
20. Rusakaniko S, Sibanda EN, Mduluza T, et al. SARS-CoV-2 Serological testing in frontline health workers in Zimbabwe. *PLoS Negl Trop Dis* [Internet]. 2021;**15**(3):1–16. Available from: <http://dx.doi.org/10.1371/journal.pntd.0009254>
21. Somboro AM, Cissoko Y, Camara I, et al. High SARS-CoV-2 Seroprevalence among Healthcare Workers in Bamako, Mali. *Viruses*. 2022;**14**(1):1–11.
22. Ssuuna C, Galiwango RM, Kankaka EN, et al. Severe Acute Respiratory Syndrome Coronavirus-2 seroprevalence in South-Central Uganda, during 2019–2021. *BMC Infect Dis* [Internet]. BioMed Central; 2022;**22**(1):1–7. Available from: <https://doi.org/10.1186/s12879-022-07161-4>
23. Etyang AO, Lucinde R, Karanja H, et al. Seroprevalence of Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2 Among Healthcare Workers in Kenya. *Clin Infect Dis*. 2022;**74**(2):288–293.
24. Müller SA, Wood RR, Hanefeld J, El-Bcheraoui C. Seroprevalence and risk factors of COVID-19 in healthcare workers from 11 African countries: a scoping review and appraisal of existing evidence. *Health Policy Plan*. 2022;**37**(4):505–513.
25. Gignoux E, Athanassiadis F, Yarrow AG, et al. Seroprevalence of SARS-CoV-2 antibodies and retrospective mortality in a refugee camp, Dagahaley, Kenya. *PLoS One*. 2021;**16**(12 December):1–11.
26. Assefa N, Regassa LD, Teklemariam Z, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in women attending antenatal care in eastern Ethiopia: A facility-based surveillance. *BMJ Open*. 2021;**11**(11):1–6.
27. Kempen JH, Abashawl A, Suga HK, et al. SARS-CoV-2 serosurvey in Addis Ababa, Ethiopia. *Am J Trop Med Hyg*. 2020;**103**(5):2022–2023.
28. Seck SM, Mbow M, Kane Y, et al. Prevalence of SARS-CoV-2 antibodies in hemodialysis

- patients in Senegal: a multicenter cross-sectional study. *BMC Nephrol* [Internet]. BioMed Central; 2021;**22**(1):1–8. Available from: <https://doi.org/10.1186/s12882-021-02582-w>
29. Batchi-Bouyou AL, Lobaloba Ingoba L, Ndounga M, et al. High SARS-CoV-2 IgG/IGM seroprevalence in asymptomatic Congolese in Brazzaville, the Republic of Congo. *Int J Infect Dis* [Internet]. International Society for Infectious Diseases; 2021;**106**:3–7. Available from: <https://doi.org/10.1016/j.ijid.2020.12.065>
  30. Okpala OCV, Dim CC, Ugwu CI, et al. Population seroprevalence of SARS-CoV-2 antibodies in Anambra State, South-East, Nigeria. *Int J Infect Dis* [Internet]. Elsevier Ltd; 2021;**110**(April 2020):171–178. Available from: <https://doi.org/10.1016/j.ijid.2021.07.040>
  31. Mikkelsen L, Phillips DE, Abouzahr C, et al. A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *Lancet* [Internet]. Elsevier Ltd; 2015;**386**(10001):1395–1406. Available from: [http://dx.doi.org/10.1016/S0140-6736\(15\)60171-4](http://dx.doi.org/10.1016/S0140-6736(15)60171-4)
  32. Rasambainarivo F, Rasoanomenjanahary A, Rabarison JH, et al. Monitoring for outbreak-associated excess mortality in an African city: Detection limits in Antananarivo, Madagascar. *Int J Infect Dis* [Internet]. International Society for Infectious Diseases; 2021;**103**:338–342. Available from: <https://doi.org/10.1016/j.ijid.2020.11.182>
  33. Parpia AS, Ndeffo-Mbah ML, Wenzel NS, Galvani AP. Effects of response to 2014–2015 ebola outbreak on deaths from malaria, HIV/AIDS, and tuberculosis, West Africa. *Emerg Infect Dis*. 2016;**22**(3):433–441.
  34. Sochas L, Channon AA, Nam S. Counting indirect crisis-related deaths in the context of a low-resilience health system: The case of maternal and neonatal health during the Ebola epidemic in Sierra Leone. *Health Policy Plan*. 2017;**32**:iii32–iii39.
  35. Cabore JW, Karamagi HC, Kipruto HK, et al. COVID-19 in the 47 countries of the WHO African region: a modelling analysis of past trends and future patterns. *Lancet Glob Heal* [Internet]. World Health Organization; 2022;**10**(8):e1099–e1114. Available from: [http://dx.doi.org/10.1016/S2214-109X\(22\)00233-9](http://dx.doi.org/10.1016/S2214-109X(22)00233-9)
  36. Bradshaw D, Dorrington R, Moultrie T, Groenewald P, Moultrie H. Underestimated COVID-19 mortality in WHO African region. *Lancet Glob Heal* [Internet]. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license; 2022;**10**(11):e1559. Available from: [http://dx.doi.org/10.1016/S2214-109X\(22\)00425-9](http://dx.doi.org/10.1016/S2214-109X(22)00425-9)
  37. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *Lancet (London, England)*. 2022;**399**(10334):1513–1536.
  38. Miller AC, Garchitorena A, Rabeza V, et al. Cohort Profile: Ifanadiana Health Outcomes and Prosperity longitudinal Evaluation (IHOPE). *Int J Epidemiol* [Internet]. 2018;**0**(0):1–7. Available from: [10.1093/ije/dyy099](https://doi.org/10.1093/ije/dyy099)
  39. ICF International. Survey Organization Manual for Demographic and Health Surveys. Calverton, Maryland USA; 2012.
  40. SMART Initiative. Interim guidance on restarting population level surveys and household level data collection in humanitarian situations during covid 19-pandemic [Internet]. 2020. Available from: [https://smartmethodology.org/wp-content/uploads/2020/10/Guidance-on-Household-Surveys-during-COVID-19\\_Final-version.pdf](https://smartmethodology.org/wp-content/uploads/2020/10/Guidance-on-Household-Surveys-during-COVID-19_Final-version.pdf)
  41. Ndiaye MDB, Rasoloharimanana LT, Razafimahatratra SL, et al. Using a Multiplex Serological

- Assay to Estimate Time Since SARS-CoV-2 Infection and Past Clinical Presentation in Malagasy Patients. *SSRN Electron J* [Internet]. 2022;**24**(9):5–7. Available from: <https://www.ssrn.com/abstract=4081990>
42. Ratovoson R, Razafimahatratra R, Randriamanantsoa L, et al. Household transmission of COVID-19 among the earliest cases in Antananarivo, Madagascar. *Influenza Other Respi Viruses*. 2022;**16**(1):48–55.
  43. Nasser A, Hamad D, Nasr C. K-means clustering algorithm in projected spaces. *2006 9th Int Conf Inf Fusion, FUSION*. 2006;(2).
  44. Iyer AS, Jones FK, Nodoushani A, et al. Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Sci Immunol*. 2020;**5**(52):1–13.
  45. Lumley T. Complex surveys : a guide to analysis using R. John Wiley; 2010.
  46. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.
  47. Rutstein SO, Rojas G. Guide to DHS Statistics. Demographic and Health Surveys Methodology. Calverton, Maryland USA; 2006.
  48. COVID-19 Forecasting Team. Variation in the COVID-19 infection–fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis. *Lancet* [Internet]. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license; 2022;**399**(10334):1469–1488. Available from: [http://dx.doi.org/10.1016/S0140-6736\(21\)02867-1](http://dx.doi.org/10.1016/S0140-6736(21)02867-1)
  49. Struck NS, Lorenz E, Deschermeier C, et al. High seroprevalence of SARS-CoV-2 in Burkina-Faso, Ghana and Madagascar in 2021: a population-based study. *BMC Public Health* [Internet]. BioMed Central; 2022;**22**(1):1–9. Available from: <https://doi.org/10.1186/s12889-022-13918-y>
  50. Kleynhans J, Tempia S, Wolter N, et al. SARS-CoV-2 seroprevalence in a rural and urban household cohort during first and second waves of infections, South Africa, July 2020–March 2021. *Emerg Infect Dis*. 2021;**27**(12):3020–3029.
  51. Kleynhans J, Tempia S, Wolter N, et al. SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa. *Emerg Infect Dis* [Internet]. 2022;**28**(5):1055–1058. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/35320701>
  52. Mutevedzi PC, Kawonga M, Kwatra G, et al. Estimated SARS-CoV-2 infection rate and fatality risk in Gauteng Province, South Africa: a population-based seroepidemiological survey. *Int J Epidemiol*. 2021;(November 2020):1–14.
  53. Madhi SA, Kwatra G, Myers JE, et al. Population Immunity and Covid-19 Severity with Omicron Variant in South Africa. *N Engl J Med*. 2022;**386**(14):1314–1326.
  54. Barrie MB, Lakoh S, Kelly JD, et al. SARS-CoV-2 antibody prevalence in Sierra Leone, March 2021: A cross-sectional, nationally representative, age-stratified serosurvey. *BMJ Glob Heal*. 2021;**6**(11):1–6.
  55. Sagara I, Woodford J, Kone M, et al. Rapidly Increasing Severe Acute Respiratory Syndrome Coronavirus 2 Seroprevalence and Limited Clinical Disease in 3 Malian Communities: A Prospective Cohort Study. *Clin Infect Dis*. 2022;**74**(6):1030–1038.
  56. Abdella S, Riou S, Tessema M, et al. Prevalence of SARS-CoV-2 in urban and rural Ethiopia:

- Randomized household serosurveys reveal level of spread during the first wave of the pandemic. *eClinicalMedicine*. 2021;**35**:1–7.
57. Firth JA, Hellewell J, Klepac P, et al. Using a real-world network to model localized COVID-19 control strategies. *Nat Med*. 2020;**26**(10):1616–1622.
  58. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature*. 2005;**438**(7066):355–359.
  59. Woolhouse MEJ, Dye C, Etard JF, et al. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proc Natl Acad Sci U S A*. 1997;**94**(1):338–342.
  60. Evans M V, Ramiadantsoa T, Kauffman K, et al. Sociodemographic Variables Can Guide Prioritized Testing Strategies for Epidemic Control in Resource-Limited Contexts. *J Infect Dis* [Internet]. 2023 Mar 24;(jiad076). Available from: <https://academic.oup.com/jid/advance-article/doi/10.1093/infdis/jiad076/7085693>
  61. Moser W, Fahal MAH, Abualas E, et al. SARS-CoV-2 Antibody Prevalence and Population-Based Death Rates, Greater Omdurman, Sudan. *Emerg Infect Dis*. 2022;**28**(5):1026–1030.
  62. Weiss DJ, Bertozzi-Villa A, Rumisha SF, et al. Indirect effects of the COVID-19 pandemic on malaria intervention coverage, morbidity, and mortality in Africa: a geospatial modelling analysis. *Lancet Infect Dis* [Internet]. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license; 2021;**21**(1):59–69. Available from: [http://dx.doi.org/10.1016/S1473-3099\(20\)30700-3](http://dx.doi.org/10.1016/S1473-3099(20)30700-3)
  63. Robertson T, Carter ED, Chou VB, et al. Early estimates of the indirect effects of the COVID-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study. *Lancet Glob Heal* [Internet]. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license; 2020;**8**(7):e901–e908. Available from: [http://dx.doi.org/10.1016/S2214-109X\(20\)30229-1](http://dx.doi.org/10.1016/S2214-109X(20)30229-1)
  64. Woolf SH, Chapman DA, Sabo RT, Zimmerman EB. Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. *JAMA* [Internet]. 2021 May 4;**325**(17):1786. Available from: <https://jamanetwork.com/journals/jama/fullarticle/2778361>
  65. Bonds M, Ouenzar M, Garchitorena A, et al. Madagascar can build stronger health systems to fight plague and prevent the next epidemic. *PLoS Negl Trop Dis* [Internet]. 2018;**12**(1):e0006131. Available from: <https://doi.org/10.1371/journal.pntd.0006131>
  66. Kurano M, Morita Y, Nakano Y, et al. Response kinetics of different classes of antibodies to SARS-CoV2 infection in the Japanese population: The IgA and IgG titers increased earlier than the IgM titers. *Int Immunopharmacol* [Internet]. Elsevier B.V.; 2022;**103**(December 2021):108491. Available from: <https://doi.org/10.1016/j.intimp.2021.108491>

## FIGURE CAPTIONS

### Figure 1. K-means clustering of SARS-CoV2 seroprevalence.

Results are based on normalized mean fluorescence intensity values for eight antibodies against SARS-CoV2 (names in white boxes). Colours represent the three clusters obtained, plotted along the axes of the first two components of a principal components analysis that explained over 60% of the variance in these antibodies.

PC1: First Principal Component; PC2: Second Principal Component; Var: Variance; RBD: Spike receptor-binding domain protein; S1: Spike S1 Protein; S2: Spike S2 Protein; NP: Nucleocapsid protein.

### Figure 2. Spatial distribution of SARS-CoV-2 seroprevalence in Ifanadiana District.

From left to right, maps show seroprevalence of all infections, past infections and recent infections based on k-clustering analyses, with colours ranging from light (low seroprevalence) to dark orange (high seroprevalence). Average seroprevalence and location of each of the 80 clusters in the survey are represented by circles, while the rest of the raster is based on inverse distance weighted interpolation. Location of major towns is represented by black points and the district capital is represented by a black square.

### Figure 3. Factors associated with SARS-CoV-2 seroprevalence in Ifanadiana District.

Horizontal bars show average seroprevalence per group, split into past infections (IgG, filled colour bars) and recent infections (IgM, translucent colour bars), with 95% confidence intervals as whiskers. Vertical lines represent average seroprevalence in Ifanadiana for past infections (dashed) and recent infections (dotted). Four confidence interval limits (Occupation variable) were removed to improve visualisation of results.

### Figure 4. Trends in adult mortality rates in the IHOPE cohort, Ifanadiana District.

Graphs show mortality per 1000 people before (green) and during (red) the COVID-19 pandemic. Left panels show changes in mortality rates across survey waves using the 6-year average prior to each survey wave. Right panels show changes in mortality rates per year in the 10 years prior to the 2021 survey. Dashed lines are the 10-year average for the period for each age and sex group. Note that the 2021 survey was conducted in April-June, so mortality estimates for this year only comprise part of the year.

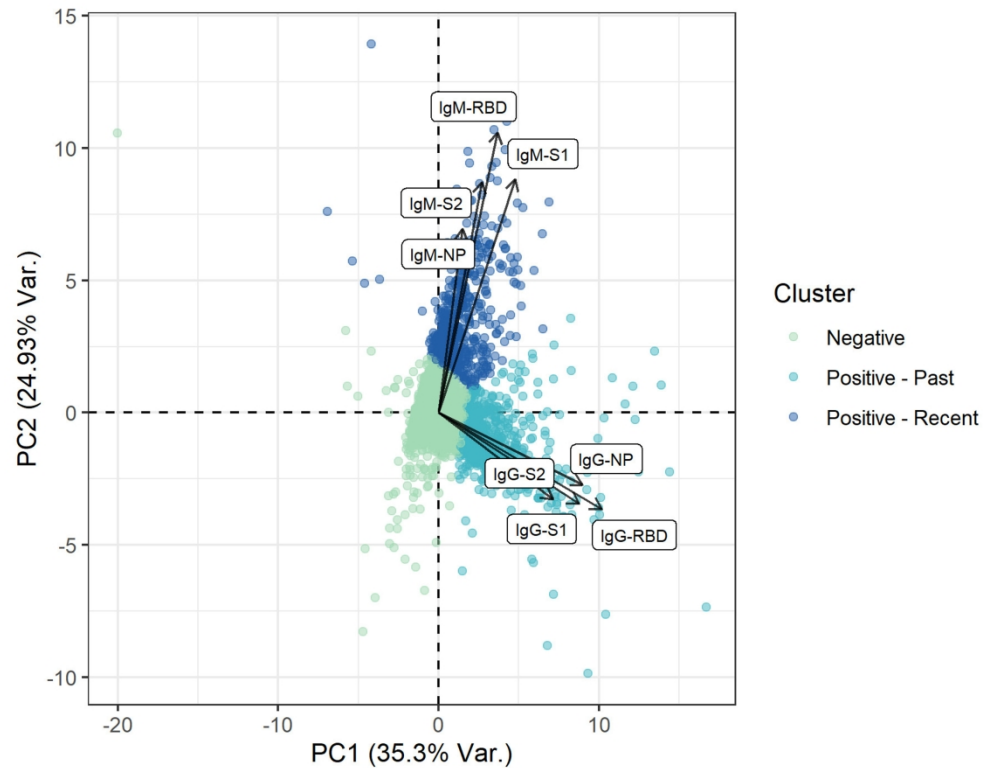


Figure 1. K-means clustering of SARS-CoV2 seroprevalence. Results are based on normalized mean fluorescence intensity values for eight antibodies against SARS-CoV2 (names in white boxes). Colours represent the three clusters obtained, plotted along the axes of the first two components of a principal components analysis that explained over 60% of the variance in these antibodies. PC1: First Principal Component; PC2: Second Principal Component; Var: Variance; RBD: Spike receptor-binding domain protein; S1: Spike S1 Protein; S2: Spike S2 Protein; NP: Nucleocapsid protein.

304x237mm (300 x 300 DPI)

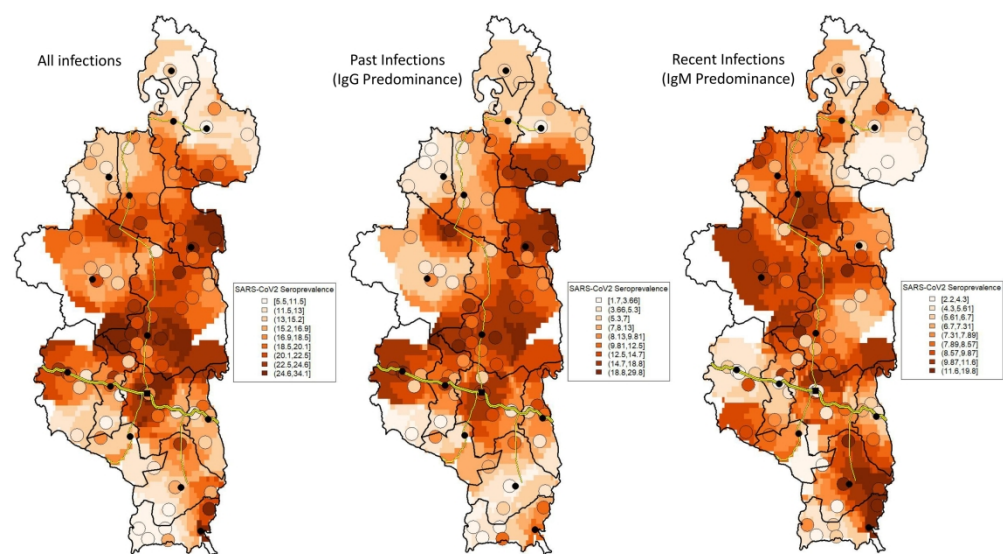


Figure 2. Spatial distribution of SARS-CoV-2 seroprevalence in Ifanadiana District. From left to right, maps show seroprevalence of all infections, past infections and recent infections based on k-clustering analyses, with colours ranging from light (low seroprevalence) to dark orange (high seroprevalence). Average seroprevalence and location of each of the 80 clusters in the survey are represented by circles, while the rest of the raster is based on inverse distance weighted interpolation. Location of major towns is represented by black points and the district capital is represented by a black square.

328x177mm (300 x 300 DPI)

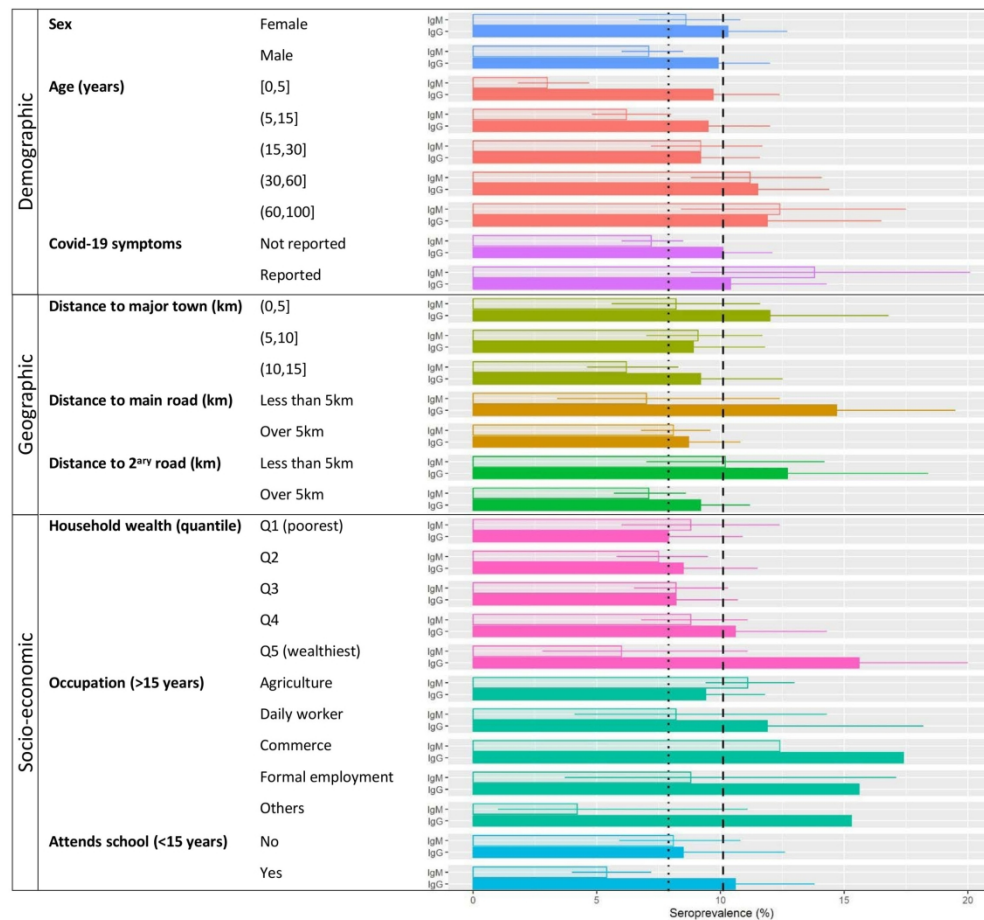


Figure 3. Factors associated with SARS-CoV-2 seroprevalence in Ifanadiana District. Horizontal bars show average seroprevalence per group, split into past infections (IgG, filled colour bars) and recent infections (IgM, translucent colour bars), with 95% confidence intervals as whiskers. Vertical lines represent average seroprevalence in Ifanadiana for past infections (dashed) and recent infections (dotted). Four confidence interval limits (Occupation variable) were removed to improve visualisation of results.

304x279mm (300 x 300 DPI)

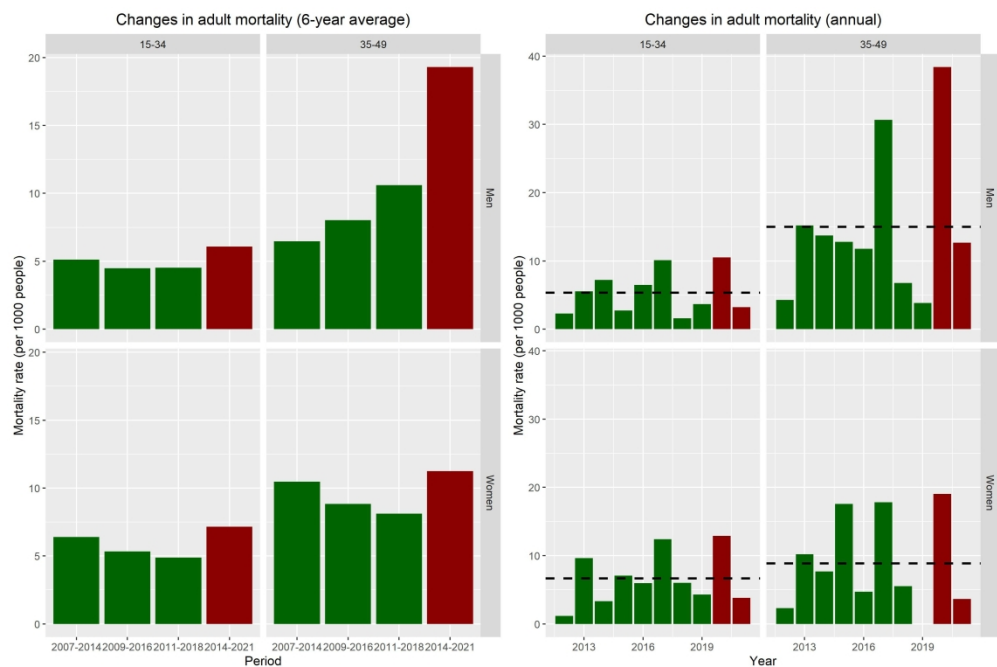


Figure 4. Trends in adult mortality rates in the IHOPE cohort, Ifanadiana District. Graphs show mortality per 1000 people before (green) and during (red) the COVID-19 pandemic. Left panels show changes in mortality rates across survey waves using the 6-year average prior to each survey wave. Right panels show changes in mortality rates per year in the 10 years prior to the 2021 survey. Dashed lines are the 10-year average for the period for each age and sex group. Note that the 2021 survey was conducted in April-June, so mortality estimates for this year only comprise part of the year.

304x203mm (300 x 300 DPI)

## MANUSCRIPT TABLES

**Table 1.** Characteristics of individuals included in the COVID-19 seroprevalence survey

Variable	Observations	Percentage
<b>Demographic factors</b>		
<b>Sex</b>		
Female	3323	51,2%
Male	3173	48,8%
<b>Age (years)</b>		
[0,5]	1069	16,5%
(5,15]	2109	32,5%
(15,30]	1386	21,3%
(30,60]	1646	25,3%
(60,100]	285	4,4%
<b>COVID-19 symptoms within previous 6 months<sup>1</sup></b>		
Not reported	5837	89,9%
Reported	659	10,1%
<b>Geographic factors</b>		
<b>Distance to main town (km)</b>		
(0,5]	2358	36,3%
(5,10]	2074	31,9%
(10,15]	2064	31,8%
<b>Distance to main road</b>		
Less than 5km	1531	23,6%
Over 5km	4964	76,4%
<b>Distance to secondary road</b>		
Less than 5km	1704	26,2%
Over 5km	4791	73,8%
<b>Socio-economic factors</b>		
<b>Household wealth</b>		
Q1 (poorest)	1297	20,0%
Q2	1334	20,5%
Q3	1334	20,5%
Q4	1277	19,7%
Q5 (wealthiest)	1253	19,3%
<b>Occupation (15+ years only)</b>		
Agriculture	2353	82,0%
Daily worker	177	6,2%
Commerce	156	5,4%
Formal employment	83	2,9%
Others	102	3,6%
<b>Attends school (5-14 years only)</b>		
No	377	20,9%
Yes	1423	79,1%
<b>TOTAL</b>	<b>6496</b>	<b>100,0%</b>

<sup>1</sup> Includes any of the following symptoms: fever, cough or respiratory problems, headache, fatigue or pain in muscles/joints, diarrhea or nausea, loss of smell, sore throat. See Table 2 for details on each.

**Table 2.** Reported COVID-19 symptoms in the previous 6 months and estimated COVID-19 seroprevalence, all ages (N=6496)

COVID-19 symptom (last 6 months)	Percentage (95% CI)
Fever	6.46 (5.3-7.6)
Cough or resp. problems	4.1 (3.2-5.0)
Headache	3.9 (2.9-4.8)
Fatigue or pain in muscles/joints	3.3 (2.5-4.0)
Diarrhea or nausea	1.1 (0.8-1.5)
Loss of smell	0.8 (0.4-1.2)
Sore throat	0.4 (0.2-0.6)
SARS Cov-2 serological marker	Seroprevalence (95% CI)
<b>IgG</b>	
Spike S1	13.0 (10.6-15.4)
Spike S2	43.8 (40.9-46.7)
Spike RBD	5.1 (3.8-6.3)
Spike NP	16.4 (14.1-18.7)
<b>IgM</b>	
Spike S1	7.2 (6.3-8.2)
Spike S2	17.9 (15.8-19.9)
Spike RBD	9.1 (7.9-10.3)
Spike NP	9.1 (7.5-10.7)
<b>Composite (k-means clustering)</b>	
Infected vs. healthy	18.0 (15.9-20.1)
IgG Predominance (past infection)	10.1 (8.3-11.9)
IgM Predominance (recent infection)	7.9 (6.5-9.2)

**Table 3.** Logistic regression results for associations with SARS-CoV-2 seroprevalence in past and recent infections, based on IgG and IgM predominance

Variable	IgG (Past infection)		IgM (Recent infection)	
	Univariate Odds Ratio (95% CI)	Adjusted Odds Ratio (95% CI)	Univariate Odds Ratio (95% CI)	Adjusted Odds Ratio (95% CI)
<b>Demographic factors</b>				
<b>Sex (ref. female)</b>				
Male	0.95 (0.79-1.14)	-	0.82 (0.64-1.05)	-
<b>Age in years (ref. 0-5)</b>				
(5,15]	0.97 (0.74-1.28)	-	2.17 (1.32-3.58)**	2.24 (1.36-3.71)**
(15,30]	0.94 (0.71-1.24)	-	3.32 (1.97-5.59)***	3.45 (2.04-5.85)***
(30,60]	1.21 (0.9-1.62)	-	4.12 (2.51-6.75)***	4.07 (2.48-6.68)***
(60,100]	1.25 (0.86-1.81)	-	4.64 (2.38-9.04)***	4.34 (2.22-8.48)***
<b>COVID-19 symptoms (ref. not reported)</b>				
Reported	1.03 (0.75-1.42)	-	2.05 (1.35-3.12)**	1.91 (1.26-2.91)**
<b>Geographic factors</b>				
<b>Distance to health center (ref. 0-5km)</b>				
(5,10]	0.72 (0.44-1.17)	-	1.12 (0.72-1.75)	1.21 (0.86-1.7)
(10,15]	0.74 (0.45-1.23)	-	0.75 (0.47-1.18)	0.74 (0.52-1.05).
<b>Distance to main road (ref. 0-5km)</b>				
Over 5km	0.55 (0.37-0.81)**	0.7 (0.49-1.02).	1.18 (0.64-2.19)	-
<b>Distance to secondary road (ref. 0-5km)</b>				
Over 5km	0.69 (0.43-1.11)	-	0.67 (0.45-1).	0.62 (0.44-0.86)**
<b>Socio-economic factors</b>				
<b>Household wealth (ref. Q1 - poorest)</b>				
Q2	1.07 (0.74-1.57)	1.04 (0.71-1.53)	0.84 (0.56-1.27)	0.79 (0.53-1.17)
Q3	1.03 (0.69-1.54)	0.97 (0.65-1.45)	0.93 (0.64-1.35)	0.84 (0.6-1.17)
Q4	1.37 (0.89-2.12)	1.23 (0.79-1.92)	1 (0.65-1.54)	0.83 (0.56-1.22)
Q5 (wealthiest)	2.14 (1.4-3.27)***	1.69 (1.14-2.52)*	0.67 (0.31-1.44)	0.48 (0.26-0.88)*
<b>Occupation, individuals 15+ years (ref. Agriculture)<sup>1</sup></b>				
Daily worker	1.31 (0.78-2.18)	-	0.71 (0.39-1.31)	-
Commerce	2.03 (1.13-3.67)*	-	1.14 (0.22-5.97)	-
Formal employment	1.78 (1.04-3.07)*	-	0.77 (0.36-1.65)	-
Others	1.74 (0.87-3.47)	-	0.35 (0.12-1.01).	-
<b>Attends school, children 5-14 years (ref. No)<sup>1</sup></b>				
Yes	1.29 (0.8-2.06)	-	0.65 (0.44-0.97)*	-

.  $P < 0.1$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ <sup>1</sup> Variable only applicable to a population subgroup, not included in multivariate analyses

**Table 4.** Estimation of excess adult mortality and percentage that could be directly attributed to COVID-19 given observed seroprevalence and known infection fatality ratios by age.

Age Group	Seroprevalence (%)	Observed excess mortality 2020-2021 (per 1000)	Expected excess mortality directly from COVID-19 infections (per 1000)	% of excess mortality directly attributed to COVID-19 infections
15-34 Years	32.7	1.61	0.05 (0.04-0.08)	3.1 (2.5-5.0)
35-49 Years	35.8	6.04	0.54 (0.39-0.85)	8.9 (6.5-14.1)