

# Demography, social contact patterns and the COVID-19 burden in different settings of Ethiopia: a modeling study

Filippo Trentini<sup>1</sup>, Giorgio Guzzetta<sup>1</sup>, Margherita Galli<sup>1,2</sup>, Agnese Zardini<sup>1,3</sup>, Fabio Manenti<sup>4</sup>, Giovanni Putoto<sup>4</sup>, Valentina Marziano<sup>1</sup>, Worku Nigussa Gamshie<sup>5</sup>, Ademe Tsegaye<sup>5</sup>, Alessandro Greblo<sup>4</sup>, Alessia Melegaro<sup>6,7</sup>, Marco Ajelli<sup>8,9</sup>, Stefano Merler<sup>1</sup>, Piero Poletti<sup>1,\*</sup>

\* corresponding author: [poletti@fbk.eu](mailto:poletti@fbk.eu)

<sup>1</sup> Bruno Kessler Foundation, Trento, Italy

<sup>2</sup> University of Udine, Udine, Italy

<sup>3</sup> University of Trento, Trento, Italy

<sup>4</sup> Doctors with Africa CUAMM, Padova, Italy

<sup>5</sup> Doctors with Africa CUAMM, Ethiopia, Italy

<sup>6</sup> Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University, Milan, Italy

<sup>7</sup> Department of Social and Political Sciences, Bocconi University, Milan, Italy

<sup>8</sup> Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA

<sup>9</sup> Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA USA

## Abstract

**Background** COVID-19 spread may have a dramatic impact in countries with vulnerable economies and limited availability of, and access to, healthcare resources and infrastructures. However, in sub-Saharan Africa a low prevalence and mortality have been observed so far.

**Methods** We collected data on individuals' social contacts in Ethiopia across geographical contexts characterized by heterogeneous population density, work and travel opportunities, and access to primary care. We assessed how socio-demographic factors and observed mixing patterns can influence the COVID-19 disease burden, by simulating SARS-CoV-2 transmission in remote settlements, rural villages, and urban neighborhoods, under the current school closure mandate.

**Results** From national surveillance data, we estimated a net reproduction number of 1.62 (95%CI 1.55-1.70). We found that, at the end of an epidemic mitigated by school closure alone, 10-15% of the overall population would have been symptomatic and 0.3-0.4% of the population would require mechanical ventilation and/or possibly result in a fatal outcome. Higher infection attack rates are expected in more urbanized areas, but the highest incidence of critical disease is expected in remote subsistence farming settlements.

**Conclusions** The relatively low burden of COVID-19 in Ethiopia can be explained by the estimated mixing patterns, underlying demography and the enacted school closures. Socio-demographic factors can also determine marked heterogeneities across different geographical contexts within the same country. Our findings can contribute to understand why sub-Saharan Africa is experiencing a relatively lower attack rate of severe cases compared to high income countries.

## Introduction

Despite limited access to healthcare<sup>1,2</sup> and relatively milder social distancing restrictions compared to those imposed in most high-income countries<sup>3,4</sup>, COVID-19 mortality rates have been relatively low throughout Africa<sup>5</sup>. As of November 5, 2020, the World Health Organization (WHO) reports 44,012 deaths out of 1,838,867 diagnosed cases in the continent, for a case-fatality ratio (CFR) of 2.4%, compared to about 3% in the rest of the world<sup>5</sup>. However, SARS-CoV-2 transmission dynamics have been highly heterogeneous across different African countries in terms of timing and implemented interventions<sup>6</sup>.

In sub-Saharan Africa, Ethiopia is second only to South Africa in terms of number of recorded cases and deaths, with an overall CFR of about 1.5%<sup>5</sup>. The first COVID-19 case was confirmed on March 13, 2020 and, less than a month later, the Ethiopian Prime Minister declared a state of emergency in the country on April 8, 2020<sup>7</sup>. Since then, borders and school closure were implemented, public institutions and firms operated at minimum capacity or under complete closure, and people were advised to stay at home<sup>8</sup>. Despite these restrictions, several hundred cases have been reported in all the 12 regions of Ethiopia<sup>9</sup>. The possible spread of SARS-CoV-2 in rural areas of the country is especially dangerous because of the sparse presence of well-resourced health facilities implying long travel distances for remote populations, which is an important barrier to universal access to primary care<sup>2</sup>. Moreover, the healthcare workforce in Ethiopia is 5 times lower than the minimum threshold defined by the WHO for Sustainable Development Goals health targets<sup>10</sup>, and far below the African average<sup>11</sup>.

Recent modeling studies investigated the impact of control measures, such as self-isolation and temporary lockdowns, in a number of sub-Saharan African countries, highlighting the difficulties in defining effective, feasible and sustainable strategies for suppression or mitigation of COVID-19 epidemics<sup>12,13,14,15</sup>. In this work, we aim to assess how demographic factors and age-specific mixing patterns can influence COVID-19 burden in Ethiopia across different geographical contexts characterized by different levels of access to healthcare.

## Methods

**Study design.** Between November and December 2019, we conducted a survey based on individual interviews to estimate age-specific mixing patterns in four districts (*woreda*) of the South West Shewa Zone (SWSZ) of the Oromia Region, Ethiopia. About 40% of the SWSZ population is below 15 years of age and about 68% lives in remote rural settlements, 18% in rural villages, and 14% in the largest town of the area (Woliso Town, 53,065 inhabitants). The districts targeted by our study encompass a population of 449,460 inhabitants and represent the main catchment area of the St. Luke Hospital located in Woliso Town, a well-resourced health facility acting as the referral hospital for the entire Zone<sup>2</sup>.

The study consists in a cross-sectional survey with two-stage stratified random sampling by location and age group. The survey was conducted in eight different sites, choosing two neighborhoods (*kebele*) for each district under study, in such a way to capture contact patterns in areas characterized by different population

density, work and travel opportunities, and access to the healthcare infrastructure. Three types of geographical contexts were considered: remote settlements (consisting of scattered subsistence farming settlements), rural villages (consisting of concentrated clusters of households served by a main road, and better access to main public services), and urban neighborhoods inside Woliso Town (significantly higher population density and full access to public services<sup>16</sup>).

For each site, a target sample size of 105 study participants was set on the basis of findings from previous contact surveys<sup>17,18</sup> to provide the desired precision in the mean number of contacts (see Appendix S1). Households and study participants were randomly sampled using predefined quotas for each site, sex, and age group. A household was defined as a group of individuals living under the same roof and sharing the same kitchen on a daily basis. One individual per household was interviewed. If the study participant was a student, additional shorter interviews were performed to complement the data with information about close contacts occurring at school.

**Data collection** Participants were asked to recall information on the frequency, location, type of social encounters from the day preceding their interview, providing the age (or age range when exact age was unknown) and their relationship for each listed contact. A contact was defined as an interaction between two individuals, either physical (when involving skin-to-skin contact), or non-physical (when involving a two-way conversation with five or more words in the physical presence of another person, but no skin-to-skin contact)<sup>17,18</sup>. The participants' age, sex, education and occupational status were recorded along with details on their household composition. In the SWSZ, schools may host up to 100 students within a single class. To avoid inaccurate reporting of the number of school contacts, participants were only asked to count the total number of physical contacts they had at school in the previous day, without further details. Information on the age of students attending the targeted schools for different grades was also collected. Schools were regularly open during the survey period.

**Contact patterns and data analysis.** For each type of geographical context, we computed the mean number of contacts reported by respondents after grouping by age (six 10-year age groups from 0 to 59 years and one age group for individuals aged 60 years or older) and by contact setting (households, schools, and the general community). Since for many study participants it was difficult to distinguish encounters occurred because of their job from other random contacts, all social interactions occurring outside family and schools were aggregated with contacts occurring in the general community. Age-specific contact matrices were computed considering both physical and non-physical contacts. Variability due to sampling of study participants was explored by computing 1,000 bootstrapped contact matrices, where each bootstrap consisted in sampling with replacement a number of interviews equal to the original sample size, choosing the age of the participant with probability proportional to the Ethiopian age distribution<sup>19</sup>. The proportions of the SWSZ population living in remote settlements, rural villages and in urban neighborhoods were used as sampling weights to compute an average contact matrix for the entire SWSZ. Full details about the study design, data collection and the analysis of contact patterns are provided in the Appendix S1.

**Transmission model.** We simulated SARS-CoV-2 spread in the different geographical contexts of the SWSZ, using an age-structured SIR compartmental model with three consecutive stages of infectiousness, in such a way to reproduce a gamma-distributed generation time of mean 6.6 days<sup>20</sup>. The model was run separately for each geographical context, using estimates of the population age structure and of the age-specific contact matrix computed from survey data (see Appendix). Because school closure in all of Ethiopia was mandated much before the exponential growth of reported COVID-19 cases, we included only data on household and community contacts in the contact matrices. However, we included school contacts to estimate the theoretical SARS-CoV-2 transmission potential in the absence of a school closure mandate. We considered susceptibility to SARS-CoV-2 infection to vary with age. We adopted posterior distributions for the relative probability of developing infection upon effective exposure to an infectious case as previously estimated in Zhang et al.<sup>19</sup>, where the age-group 15-64 years is taken as a reference; an average relative susceptibility of 0.33 (95%CI: 0.24-0.47) was estimated for children under 15 years of age, and of 1.47 (95%CI: 1.16-2.06) for older adults (above 65 years)<sup>19</sup>. These estimates are aligned with other independent studies (reviewed in Viner et al.<sup>21</sup>). We assumed the same infectiousness across individuals of different ages. Two separate sensitivity analyses were carried out assuming 1) homogeneous susceptibility by age; 2) a lower infectiousness of children (see Appendix).

We estimated the SARS-CoV-2 net reproduction number from the curve of reported cases in Ethiopia during the phase of exponential growth<sup>5,22</sup> (see Appendix). We then set a per-contact transmission rate in the model such that the reproduction number, calculated using the next-generation matrix approach<sup>23</sup> matches the obtained national estimate (see Appendix). A sensitivity analysis was conducted to explore the effect of a 20% increase or a 20% decrease of the reproduction number on COVID-19 burden.

We computed projections of the number of SARS-CoV-2 infections, cases with respiratory symptoms or fever, and COVID-19 critical cases (either requiring mechanical ventilation or resulting in a fatal outcome), based on available estimates of the age-specific risks<sup>24</sup>.

## Results

**Social contact data.** A total of 938 study participants were interviewed with 43% of them living in rural remote settlements, 35% in rural villages, and 22% from urban neighborhoods (Table 1). 227 participants were students, 22.9% of whom were between 5 and 9 years of age, 71.8% between 10 and 19 years, and 4.9% older. School attendance rates among the study participants aged 5-18 years was 67%, 80% and 77% in remote, rural and urban sites, respectively. The median class size ranged from 70 children per class in rural villages to 90 in remote settlements. Only 27% of our study participants reported travels outside their village in the last month; 87.3% reported they were never admitted to the local hospital (see Appendix S1).

**Table 1.** Characteristics of study participants.

### Number of study participants

VARIABLE*	OVERALL n (%)	REMOTE n (%)	RURAL n (%)	URBAN n (%)
<b>Total</b>				
	938 (100.0)	400 (42.6)	326 (34.8)	212 (22.6)
<b>Age</b>				
<10y	382 (40.7)	160 (40)	137 (42)	85 (40.1)
10-19y	198 (21.1)	85 (21.2)	66 (20.2)	47 (22.2)
20-29y	92 (9.8)	40 (10)	32 (9.8)	20 (9.4)
30-39y	117 (12.5)	50 (12.5)	42 (12.9)	25 (11.8)
40-49y	59 (6.3)	26 (6.5)	18 (5.5)	15 (7.1)
50-59y	40 (4.3)	17 (4.2)	13 (4)	10 (4.7)
60y +	50 (5.3)	22 (5.5)	18 (5.5)	10 (4.7)
<b>Occupation</b>				
Student	226 (24.1)	85 (21.2)	87 (26.7)	54 (25.5)
Office/Shop worker	47 (5.0)	4 (1.0)	25 (7.7)	18 (8.5)
Agriculture	112 (11.9)	84 (21)	25 (7.7)	3 (1.4)
Unemployed/Retired	44 (4.7)	9 (2.3)	12 (3.7)	23 (10.8)
Other	509 (54.3)	218 (54.5)	177 (54.3)	114 (53.8)
<b>Sex</b>				
Female	478 (51)	206 (51.5)	170 (52.1)	102 (48.1)
Male	460 (49)	194 (48.5)	156 (47.9)	110 (51.9)

\* No missing data for any of the three listed variables.

Age and sex were also recorded for all the 4,635 household members of the 938 study participants. The mean household size in remote settlements was 5.5 (95% CI: 5.3-5.7), significantly larger (Tukey test  $p < 0.001$ ) than in rural villages (4.6, 95% CI: 4.4-4.8) and in urban neighborhoods (4.4, 95%CI: 4.2-4.6), while no significant difference in the household size was found between the latter two settings (Tukey test  $p = 0.48$ ).

Overall, 5,690 non-school contacts were reported by the 938 study participants (median 6 contacts per person, range 1-26, see Table 2). Of these, 79.9% were physical and 43.0% involved a single social interaction during the day.

**Table 2.** Mean number of recorded daily contacts, excluding contacts at school, by age, across different geographical contexts.

VARIABLE	Mean number of contacts per day (excluding school contacts)			
	OVERALL mean (95% CI)	REMOTE mean (95% CI)	RURAL mean (95% CI)	URBAN mean (95% CI)

<b>Total</b>					
		6.07 (5.88-6.26)	6.19 (5.87-6.51)	5.73 (5.44-6.02)	6.35 (5.96-6.73)
<b>Age</b>					
<10y		5.57 (5.32-5.83)	5.67 (5.23-6.12)	5.21 (4.84-5.58)	5.96 (5.46-6.47)
10-19y		6.48 (6.02-6.94)	6.33 (5.6-7.06)	6.30 (5.63-6.98)	7.00 (5.96-8.04)
20-29y		5.77 (5.28-6.26)	5.8 (5.04-6.56)	5.72 (4.84-6.6)	5.80 (4.82-6.78)
30-39y		6.99 (6.41-7.57)	6.84 (5.96-7.72)	7.05 (6.00-8.09)	7.20 (6.05-8.35)
40-49y		6.86 (6.08-7.65)	7.23 (5.85-8.61)	6.67 (5.51-7.82)	6.47 (5.01-7.92)
50-59y		5.80 (4.90-6.70)	5.76 (4.66-6.87)	5.77 (3.95-7.59)	5.90 (3.74-8.06)
60y +		5.84 (4.69-6.99)	7.73 (5.54-9.91)	3.56 (2.84-4.27)	5.80 (4.26-7.34)
<b>Sex</b>					
Male		6.15 (5.87-6.43)	6.15 (5.69-6.61)	6.02 (5.55-6.49)	6.34 (5.77-6.9)
Female		5.99 (5.73-6.24)	6.23 (5.79-6.67)	5.46 (5.11-5.82)	6.36 (5.84-6.89)
<b>Occupation</b>					
Student		6.57 (6.10-7.04)	6.51 (5.63-7.38)	6.24 (5.61-6.88)	7.20 (6.25-8.15)
Office/Shop worker		7.66 (6.57-8.75)	7.75 (3.97-11.53)	7.6 (5.93-9.27)	7.72 (6.16-9.28)
Agriculture		7.02 (6.35-7.68)	7.35 (6.53-8.16)	5.92 (4.86-6.98)	7.00 (5.04-8.96)
Unemployed/Retired		5.49 (5.29-5.70)	5.58 (5.25-5.92)	5.13 (4.81-5.45)	5.85 (5.42-6.27)
Others		5.84 (5.34-6.35)	5.91 (5.08-6.75)	5.95 (5.1-6.8)	5.63 (4.63-6.63)
<b>Setting</b>					
Household		2.8 (2.68-2.92)	2.94 (2.74-3.15)	2.48 (2.28-2.67)	3.02 (2.8-3.24)
Community		3.27 (3.09-3.45)	3.25 (2.95-3.54)	3.25 (2.97-3.53)	3.33 (2.98-3.67)
<b>Traveled to a different neighborhood in the month prior the interview</b>					
Yes		6.21 (5.83-6.59)	6.08 (5.34-6.81)	5.66 (5.16-6.16)	7.22 (6.46-7.97)
No		6.01 (5.79-6.23)	6.22 (5.86-6.57)	5.76 (5.4-6.13)	5.93 (5.51-6.35)
<b>Contacts outside neighborhood</b>					



(%)					
	0-14y	0.67% (0.4-1.04)	0.38% (0.14-0.94)	0.19% (0.03-0.77)	1.80% (1.03-3.09)
	15-59y	3.98% (3.23-4.89)	4.54% (3.36-6.07)	3.3% (2.19-4.9)	3.91% (2.47-6.07)
	60y+	2.74% (1.28-5.53)	1.18% (0.20-4.63)	0.00% (0.00-0.00)	10.34% (4.28-21.84)

For all sites, contacts outside school were predominantly reported between family members (46.1%), neighbors (25.2%), and other relatives outside the household (13.1%), while the remaining 15.5% of contacts occurred with friends, schoolmates outside school, or other unspecified categories. Individuals with a recent history of travel outside their neighborhood did not report an increased number of contacts, except for urban residents (t-test  $p=0.004$ ). The mean number of contacts (excluding school contacts) reported by participants was lower in rural villages (5.73, 95%CI 5.44-6.02) with respect to both urban neighborhoods (6.35, 95%CI 5.96-6.73) and remote settlements (6.19, 95%CI 5.87-6.51). In particular, the mean number of daily contacts reported by the elderly (60+ years old) was much higher in remote settlements and urban neighborhoods than in rural villages (7.7 and 5.8 vs 3.6, see Table 2).

Students reported 1,372 additional contacts in schools, resulting in a mean number of 6.07 (95%CI 4.98-7.16) daily physical contacts per child (median 3, range 0-50). There were limited differences in the mean number of school contacts across geographical contexts (6.31, 95%CI 4.13-8.50 in remote settlements; 5.70, 95%CI 4.19-7.21 in rural towns; 6.54, 95%CI 4.25-8.84 in urban neighborhoods).

The analysis of contacts by age clearly shows that subjects below 30 years of age tend to interact mostly with individuals of similar age (assortative mixing). The highest contact rates were found between school aged children (10-19 years), young adults (20-39 years) and between children below 10 years and their parents (Figure 1, and Appendix). However, a marked intergenerational mixing both within households and in the community was found, especially in remote settlements.

### **Effect of demography and age-specific contacts on COVID-19 epidemics.**

From the epidemic curve of reported cases, we estimated a net reproduction number  $R$  of 1.62 (95%CI 1.55-1.70) over approximately 6 weeks of exponential growth starting from May 1, 2020 when schools were closed in the entire country (see Appendix). We relied on this estimated of  $R$  to simulated COVID-19 epidemics in the SWSZ considering no school contacts. If school contacts are included, we estimate  $R$  to increase to 2.03 (95%CI 1.87-2.18, see Appendix), which is comparable, although slightly lower, than estimates of the basic reproduction number from other parts of the world<sup>25,26,27,28</sup>.

Our simulation results show that, should schools remain closed for the entire duration of the epidemic and no other interventions enacted, between 10% and 15% of the overall population would have developed respiratory symptoms or fever because of COVID-19. The fraction of critical cases (requiring mechanical ventilation and/or resulting in a fatal outcome) is estimated to be between 0.29% and 0.41% of

the overall population (Figure 2). The highest prevalence of critical cases (between 4.2% and 5.4% on average) is expected within subjects aged 60 years or older. This age segment represents only about 5% of the total population in Ethiopia but is expected to represent 7 to 14% of symptomatic cases and 45% to 63% of all critical cases.

Remote settlements are expected to suffer a higher overall burden of critical cases (0.40% of the total population, 95%CI: 0.37-0.41%) compared to rural villages (0.32%, 95%CI: 0.30-0.35%) and urban neighborhoods (0.31%, 95%CI 0.29-0.33%). This difference is explained by a higher proportion of the elderly in the population (Figure 3), but also by their higher number of daily contacts and the higher intergenerational mixing (Figure 1c) compared to the other settings, which results in a higher attack rate of infections, symptomatic cases, and critical disease in this age group (Figure 2). Urban neighborhoods, where highest contact rates at younger ages were recorded, are expected to have the highest attack rate of infections (58.8%, 95%CI: 50.6-67.3) and symptomatic cases (13.4%, 95%CI: 11.8-15.0). However, since a large proportion of the overall number of infections (81.8%, 95%CI: 79.5-84.2) is concentrated on children and younger adults (up to 40 years of age), this does not result in a high overall proportion of critical disease. Finally, rural villages have lower attack rates among the elderly because of the significantly lower number of contacts reported by that age group in this geographical context (Figure 1f, Table 2).

## Discussion

Our analysis explored the effect of demographics and social contact patterns on COVID-19 burden in the South West Shewa Zone of the Oromia region, Ethiopia. Data collected within an interview-based survey highlighted differences in demographic structure and in age-specific contacts between urban neighborhoods, rural villages, and remote settlements, and were used to inform an epidemic model simulating the transmission dynamic of SARS-CoV-2. On the basis of the trajectory of COVID-19 cases observed in the country up to June 12, 2020, we estimated that between 3.1 and 4.0 patients per 1,000 inhabitants may experience critical disease (i.e., requiring mechanical ventilation and/or resulting in a fatal outcome) at the end of an epidemic mitigated by school closure alone. Considering the low availability and accessibility of healthcare, especially in remote and rural settlements, and the lack of intensive care units to treat critical patients<sup>2,29</sup>, it is possible that a large fraction of those cases would result in a fatal outcome, adding up to the already high background mortality rate in the region (estimated at about 6.4 per 1,000 per year<sup>30</sup>).

Considering the extreme scenario where all critical cases would result in a fatal outcome, we obtain an estimate of the infection-fatality ratio (IFR) ranging between 0.53% in urban neighborhoods and 0.76% in remote settlements. Such estimates are generally lower than the IFR estimated from serological studies for higher income countries<sup>31,32</sup>. This difference is partially due to the younger age structure of the Ethiopian population, where only 5% of individuals are older than 60 years (compared to over 20% in most of Europe<sup>33</sup>). However, by simply adjusting the age-specific IFR to the local demographics, Ghisolfi et al.<sup>34</sup> estimated a four-fold reduction in the overall IFR in Eastern Africa with respect to European countries,



which is around 2 time lower than our estimates. In fact, our simulations not only account for the demography of the population, but also for its mixing patterns. Indeed, we found that in Ethiopia the effects of a younger population is partially compensated by high infection attack rates in the elderly, which derive from the intense intergenerational mixing and the larger number of contacts observed among the elderly. In particular, we show that these characteristics are especially marked in remote settlements, where the highest incidence of critical disease is expected to occur.

To properly interpret the results presented in our study, it is important to consider the following limitations. First, the target study population may be not representative of all Ethiopia, and in particular of epidemic patterns observed in highly urbanized areas such as the capital Addis Abeba. Second, the net reproduction number was estimated from national surveillance data<sup>9</sup>. This data reports cases aggregated at the country level and may suffer from a number of biases: it does not account for reporting delays; the growth over time in the number of cases may partly be ascribable to the increase in testing capacity; total cases represent the superimposition of different, asynchronous epidemics in multiple parts of the country, a majority of which coming from the highly urbanized Addis Abeba area<sup>9</sup>. However, we show that, assuming no restriction to school contacts, the reproduction number estimated by the model is in the range 1.87-2.18, just slightly lower than estimates of the SARS-CoV-2 basic reproduction number from other countries<sup>25,26,27,28</sup>. Moreover, our conclusions remain robust when considering a 20% increase or a 20% decrease of the reproduction number. In this case, we estimated an attack rate of critical cases ranging from 0.24 to 0.37 for rural villages and from 0.34 to 0.42 for remote settlements (see Appendix). Third, the model lacks of spatial structure. The finding from the survey that about 97% of recorded contacts have occurred within the participant's neighborhood of residence (Table 2) suggests that local containment or confinement of COVID-19 outbreaks in rural regions of Ethiopia may be favored by low human mobility. On the other hand, the observation of a large number of cases in all regions of Ethiopia<sup>9</sup> may imply that a significant widespread diffusion of the epidemic, possibly sustained by a high fraction of asymptomatic infections (Figure 2), is already ongoing. Fourth, the role played by children in the transmission of SARS-CoV-2 infections is still poorly understood and highly debated<sup>19,35</sup>. In the main analysis we assumed that the probability of transmission is homogeneous across all ages; however, an alternative assumption in which children are assumed half as infectious as adults would result in similar attack rates of critical cases (see Appendix S1). These results are also robust with respect to the assumption of a homogeneous susceptibility across age groups (see Appendix S1). Finally, in absence of direct data from sub-Saharan Africa, the age-specific susceptibility and proportions of infections resulting in symptomatic cases or critical disease were estimated from data from China or Europe<sup>19,24</sup>. However, the high prevalence of comorbidities which are uncommon in higher income countries (e.g., malnutrition<sup>36</sup>, tuberculosis, and malaria) and inequalities in the access to primary care represent additional vulnerabilities for African settings<sup>2</sup> and may result in an underestimation of the expected disease burden. Further studies are warranted regarding the severity of COVID-19 in sub-Saharan populations, and especially in relation to its understudied potential interaction with specific comorbidities.

## Conclusions

This study provides novel data on mixing patterns in rural Ethiopia and highlights the potential impact of COVID-19 epidemics in less urbanized regions of the country. We provide estimates on the potential burden of COVID-19 under the assumption of a mitigated, but not controlled epidemic. We conclude that, although the overall mortality might be generally lower in sub-Saharan Africa compared to high income settings, thanks to younger demographics<sup>34,37,38</sup>, this effect may be partially offset in rural areas by higher attack rates in elderly individuals, due to high rates of intergenerational mixing. The observed contact patterns suggest that elderly individuals in remote settlements may be even more exposed to the risk of infection (and thus of critical disease), which is especially worrisome in light of the major obstacles in access to healthcare for those populations<sup>2</sup>.

## **DECLARATIONS**

### **Ethics approval and consent to participate**

The study was approved by the Oromia Regional Health Bureau. Written informed consent was sought for literate individuals aged 18 years or more, and from a parent or carer for younger individuals. In addition, assent was sought from all participants under 18 years of age, illiterate participants and participants who cannot read in the Oromyffa language. A written confirmation of verbal assent by a witness elected by the study participant was sought.

### **Consent for publication**

Not applicable

### **Availability of data and material**

Data analysed during this study will be included in this published article as supplementary information files.

### **Competing interests**

MA has received research funding from Seqirus. The funding is not related to COVID-19. All other authors declare no conflict of interest.

### **Funding**

This work was supported by the Italian Ministry of Foreign Affairs and International Cooperation within the project entitled “Rafforzamento del sistema di sorveglianza e controllo delle malattie infettive in Etiopia”—AID 011330. The funders had no role in the study design, data collection and analysis, interpretation, or preparation of the manuscript.

### **Author's contribution**

FT, FM, GP, MA, SM and PP conceived and designed the study. FT, VM, WNG,AT, AG and PP collected the data. FT, MG, AZ and PP analysed the data. FT, GG, MG and AZ performed the experiments. FT, GG and PP drafted the first version of the manuscript. All authors contributed to the interpretation of the results and edited and approved the final manuscript.

### **Acknowledgments**

Not applicable.

## References

- 1 Gilbert M, Pullano G, Pinotti F, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *Lancet*. 2020 Mar 14;**395**(10227):871–7.
- 2 Poletti P, Parlamento S, Fayyisaa T, et al. The hidden burden of measles in Ethiopia: how distance to hospital shapes the disease mortality rate. *BMC Med*. 2018 Oct 18;**16**(1):177.
- 3 Massinga Loembé M, Tshangela A, Salyer SJ, Varma JK, Ouma AEO, Nkengasong JN. COVID-19 in Africa: the spread and response. *Nat Med*. 2020 Jul;**26**(7):999–1003.
- 4 International Monetary Fund. Policy responses to Covid-19. 2020. <https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19> (Accessed October 24, 2020).
- 5 World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. 2020. <https://covid19.who.int/WHO-COVID-19-global-data.csv> (Accessed October 24, 2020).
- 6 Makoni M. COVID-19 in Africa: half a year later. *Lancet Infect Dis*. 2020 Oct;**20**(10):1127.
- 7 World Health Organization. Covid-19 Response Bulletin Ethiopia. 2020. [https://www.afro.who.int/sites/default/files/2020-04/ETHIOPIA\\_COVID-19%20response%20bulletin\\_10APR2020%20%28002%29.pdf](https://www.afro.who.int/sites/default/files/2020-04/ETHIOPIA_COVID-19%20response%20bulletin_10APR2020%20%28002%29.pdf) (Accessed October 24, 2020).
- 8 UNICEF. Ethiopia COVID-19 Situation Report No. 3. 2020. <https://www.unicef.org/ethiopia/sites/unicef.org.ethiopia/files/2020-04/UNICEF%20Ethiopia%20COVID-19%20Situation%20Report%20No.%203%20-%2027%20March-3%20April%202020.pdf> (Accessed October 24, 2020).
- 9 Ministry of Health Ethiopia. COVID-19 pandemic preparedness and response in Ethiopia weekkly bullettin n. 23. 2020. [https://covid19.eph.gov.et/wp-content/uploads/2020/10/EPHI\\_PHEOC\\_COVID-19\\_Weekly\\_Bulletin\\_23\\_English\\_10052020.pdf](https://covid19.eph.gov.et/wp-content/uploads/2020/10/EPHI_PHEOC_COVID-19_Weekly_Bulletin_23_English_10052020.pdf) (Accessed October 24, 2020).
- 10 World Health organization. Health Workforce Requirements for Universal Health Coverage and the Sustainable Development Goals. Human Resource for Health Observers Series No. 17. WHO; 2016. <https://apps.who.int/iris/bitstream/handle/10665/250330/9789241511407-eng.pdf;jsessionid=1E48420C3A268CA43F1E9F9AB59BA0AE?sequence=1> (Accessed October 24, 2020).
- 11 Haileamlak A. How Can Ethiopia Mitigate the Health Workforce Gap to Meet Universal Health Coverage? *Ethiop J Health Sci*. 2018 May;**28**(3):249–250.
- 12 van Zandvoort K, Jarvis CI, Pearson CAB, et al. Response strategies for COVID-19 epidemics in African settings: a mathematical modelling study. *BMC Med*. 2020 Oct 14;**18**(1):324.
- 13 Quaife M, van Zandvoort K, Gimma A, et al. The impact of COVID-19 control measures on social contacts and transmission in Kenyan informal settlements. *BMC Med*. 2020 Oct 5;**18**(1):316.
- 14 Brand SP, Aziza R, Kombe IK, et al. Forecasting the scale of the COVID-19 epidemic in Kenya. *MedRxiv*. 2020; published online April 16. <https://doi.org/10.1101/2020.04.09.20059865> (preprint).
- 15 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*. 2020 Jul 24;**369**(6502):413-422.
- 16 United States Department of Agriculture. Economic Research Service. <https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/> (Accessed October 24, 2020).
- 17 Melegaro A, Del Fava E, Poletti P, et al. Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe. *PLoS One*. 2017 Jan 18;**12**(1):e0170459.
- 18 le Polain de Waroux O, Cohuet S, Ndazima D, et al. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: a survey in Southwest Uganda. *BMC Infect Dis*. 2018 Apr 11;**18**(1):172.
- 19 Zhang J, Litvinova M, Liang Y, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*. 2020 Jun 26;**368**(6498):1481–1486.
- 20 Cereda D, Tirani M, Rovida F, et al. The early phase of the COVID-19 outbreak in Lombardy, Italy. *arxiv* 2020; published online March 20. <http://arxiv.org/abs/2003.09320> (preprint)
- 21 Viner RM, Mytton OT, Bonell C, Melendez-Torres GJ, Ward J, Hudson L, Waddington C, Thomas J, Russell S, van der Klis F, Panovska-Griffiths J. Susceptibility to SARS-CoV-2 infection amongst children and adolescents compared with adults: a systematic review and meta-analysis. *JAMA Pediatr*. 2020; doi:10.1001/jamapediatrics.2020.4573

- 22 Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc Biol Sci*. 2007 Feb 22; **274**(1609):599-604.
- 23 Diekmann O, Heesterbeek JA, Metz JA. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J Math Biol*. 1990; **28**(4):365–82.
- 24 Poletti P, Tirani M, Cereda D, et al. Probability of symptoms and critical disease after SARS-CoV-2 infection. *arxiv* 2020; published online June 22. <http://arxiv.org/abs/2003.09320> (preprint)
- 25 Riccardo F, Ajelli M, Andrianou X, et al. Epidemiological characteristics of COVID-19 cases in Italy and estimates of the reproductive numbers one month into the epidemic. *Euro Surveill*. 2020; *In press*
- 26 Munayco CV, Tariq A, Rothenberg R, et al. Early transmission dynamics of COVID-19 in a southern hemisphere setting: Lima-Peru: February 29th-March 30th, 2020. *Infect Dis Model*. 2020 May 12; **5**:338–345.
- 27 Muniz-Rodriguez K, Fung IC, Ferdosi SR, Ofori SK, Lee Y, Tariq A, Chowell G. Severe Acute Respiratory Syndrome Coronavirus 2 Transmission Potential, Iran, 2020. *Emerg Infect Dis*. 2020 Aug; **26**(8):1915–1917.
- 28 Park M, Cook AR, Lim JT, Sun Y, Dickens BL. A Systematic Review of COVID-19 Epidemiology Based on Current Evidence. *J Clin Med*. 2020 Mar **31**; **9**(4):967.
- 29 Murthy S, Leligdowicz A, Adhikari NK. Intensive care unit capacity in low-income countries: a systematic review. *PLoS One*. 2015 Jan 24; **10**(1):e0116949.
- 30 The World Bank. World Bank Open Data. <https://data.worldbank.org/indicator/SP.DYN.CDRT.IN?locations=ET> (Accessed October 24, 2020).
- 31 Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet*. 2020 Aug 22; **396**(10250):535–544.
- 32 Poletti P, Tirani M, Cereda D, et al. Age-specific SARS-CoV-2 infection fatality ratio and associated risk factors, Italy, February to April 2020. *Euro Surveill*. 2020 Aug; **25**(31):2001383.
- 33 Eurostat. Population Structure and Ageing. [https://ec.europa.eu/eurostat/statistics-explained/index.php/Population\\_structure\\_and\\_ageing#The\\_share\\_of\\_elderly\\_people\\_continues\\_to\\_increase](https://ec.europa.eu/eurostat/statistics-explained/index.php/Population_structure_and_ageing#The_share_of_elderly_people_continues_to_increase) (Accessed October 24, 2020).
- 34 Ghisolfi S, Almás I, Sandefur JC, von Carnap T, Heitner J, Bold T. Predicted COVID-19 fatality rates based on age, sex, comorbidities and health system capacity. *BMJ Glob Health*. 2020 Sep; **5**(9):e003094.
- 35 Bi Q, Wu Y, Mei S, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis*. 2020 Aug; **20**(8):911-919.
- 36 Endris N, Asefa H, Dube L. Prevalence of Malnutrition and Associated Factors among Children in Rural Ethiopia. *Biomed Res Int*. 2017; **2017**:6587853.
- 37 Dowd JB, Andriano L, Brazel DM, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci USA*. 2020 May 5; **117**(18):9696-9698.
- 38 Hilton J, Keeling MJ. Estimation of country-level basic reproductive ratios for novel Coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. *PLoS Comput Biol*. 2020 Jul 2; **16**(7):e1008031.

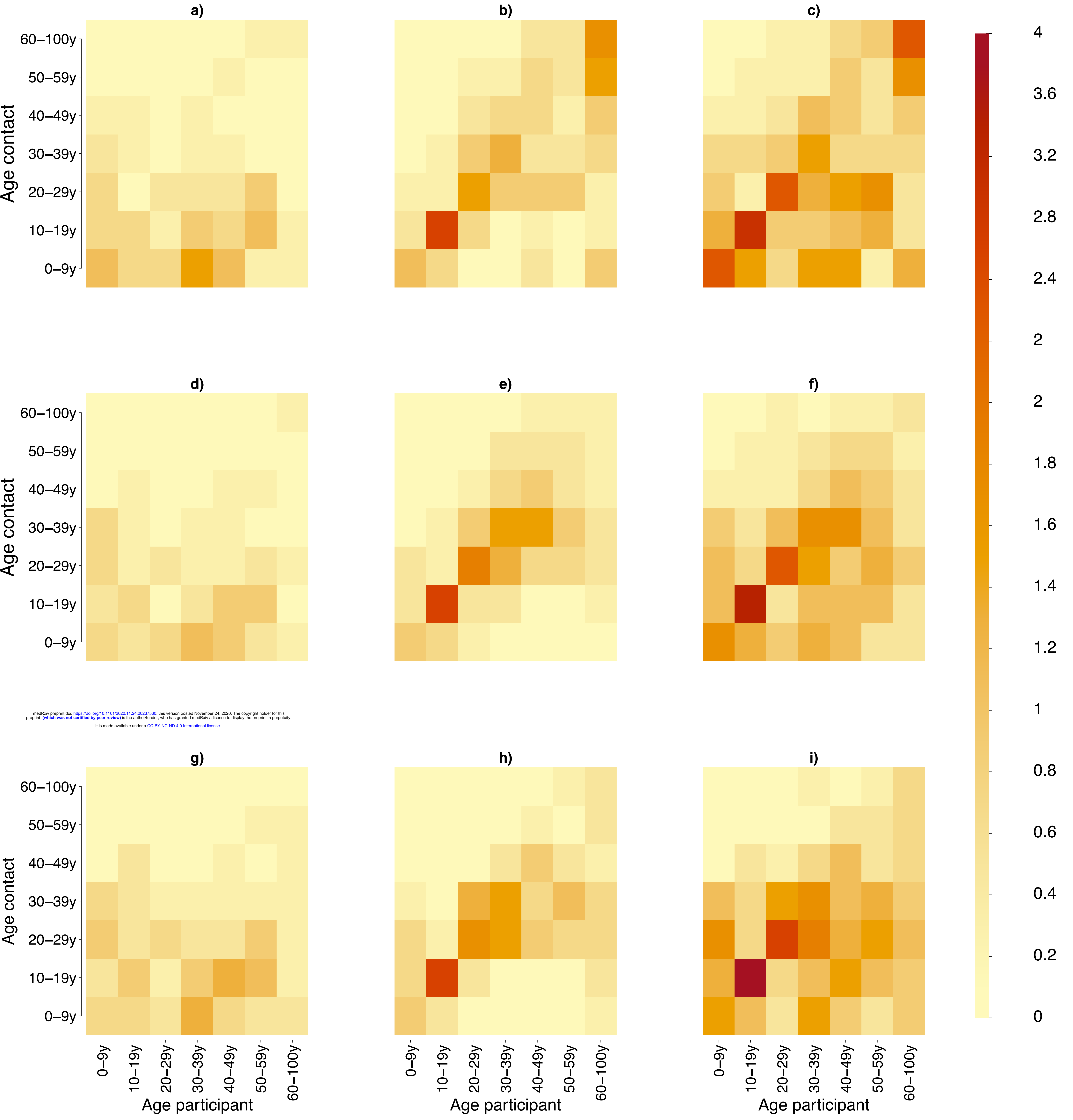
## Figures:

**Figure 1** Contact matrix representing the mean number of daily contacts reported by a participant in the age group  $i$  with individuals in the age group  $j$  in household (a), in the general community (b) and both (c) in remote settlements. Panels d-f and g-l show the same quantities estimated for rural villages and for the urban neighborhoods, respectively.

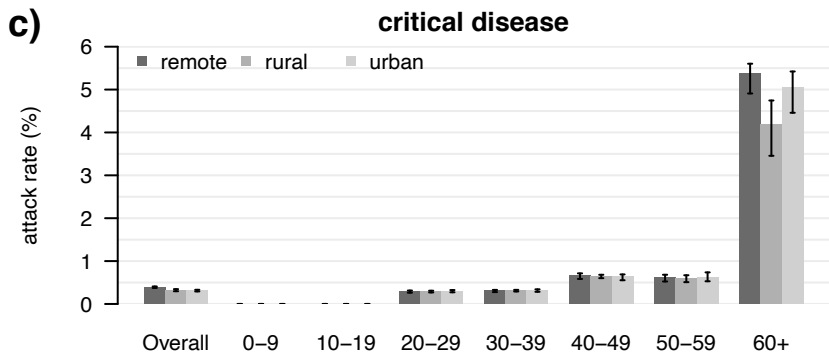
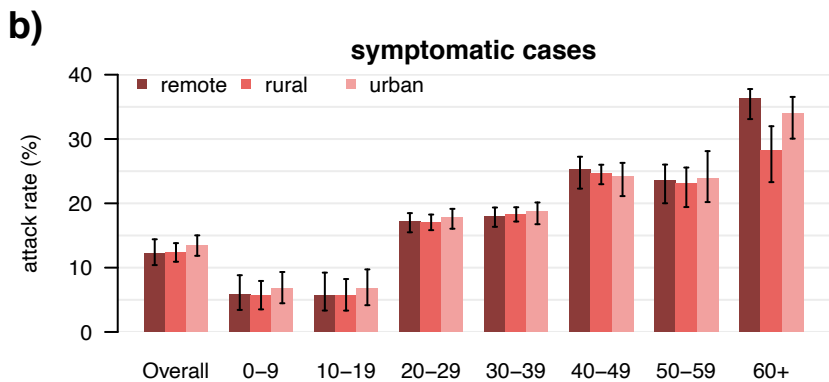
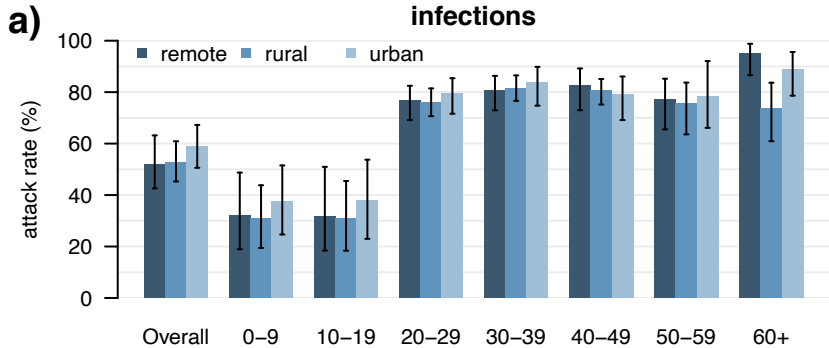
**Figure 2** Estimated attack rates of infection (a), symptomatic cases (b), and critical disease (c), overall and by age group in different geographical contexts. We considered hypothetical epidemics with school closure and a reproduction number of 1.62 (95%CI 1.55-1.70), estimated from surveillance data<sup>5</sup>.

**Figure 3** Age structure of the population in the South West Showa Zone (SWSZ), Oromia Region, Ethiopia, by geographical context, obtained from survey data.









% population

