THE LANCET
Global Health

Supplementary appendix 3

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: 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 Health* 2021; **9:** e1517–27.

Supplementary Material

1. Prevalence and Incidence Estimation

Supplementary methods

A two-way, repeated measures ANOVA model was used to examine the effect of population group and time point on the number of seropositive individuals. The variability explained by the model is divided into two factors: Group (Jimma vs. Addis Ababa and HCW vs Community; between-subjects factor indicating population group) and Round (1, 2, 3; within-subjects factor denoting time point of the serology test); and an interaction term Group:Round testing whether the effect of Round and Group jointly influences the seropositive count, i.e. if some groups have a differential effect in specific rounds.

$$y = \alpha + \beta_1 Group + \beta_2 Round + \beta_3 Group{:}Round + \varepsilon$$

In this case, y refers to the count of seropositives (pos) within each group and survey round and we have the following equations to estimate p_i which is the probability of positive in a group and round.

$$pos_i \sim Binomial(n_i, p_i)$$

$$logit(p_i) = \alpha + \beta_1 Group + \beta_2 Round + \beta_3 Group{:}Round$$

$$\alpha, \beta_i \sim Normal(0{,}10)$$

In addition to considering binomial outcomes, we also examined the count of seropositives assuming a Poisson outcome distribution. The equations are similar to the binomial distribution except for the need to have an offset variable adjusting for the denominator to estimate the rate λ_i for being positive.

$$pos_i \sim Normal(\lambda_i)$$

$$log(\lambda_i) = \alpha + \beta_1 Group + \beta_2 Visit + \beta_3 Group{:}Visit + offset(n_i)$$

$$\alpha, \beta_i \sim Normal(0{,}10)$$

Estimates of the counts along with the 95% Credible Intervals were obtained using non-informative priors (normal distribution with mean zero and standard deviation 10) with 5000 warm-up samples followed by 5000 MCMC samples for the posterior outcome of a generalized linear model using the brms (Bayesian Regression Models using 'Stan') package in R.[1,2,3] The prevalence estimates are obtained by dividing the estimated count for positives by the observed samples. In case of the sero-incidence measures we have the count of new positives instead of positives and there is no component of round. Instead, the denominator is person-weeks of being observed within the study. The above models were also used to estimate the contrasts to check group wise and/or round wise differences. We published the code and tables used in this paragraph at Zenodo.[4]

Supplementary results for incidence and prevalence estimation

Table S1: SARS-CoV-2 seroincidence rates per person-weeks for HCW at Jimma Medical Center and St. Paul's Hospital, and communities from Jimma and Addis Ababa

	New seropositives (N)	Person-weeks	Seroincidence rates per 100,000 person-weeks (95% CI)
HCW Jimma Medical Center	111	2913	3810 (3149, 4540)
HCW St· Paul's Hospital	90	4051	2223 (1785, 2696)
Jimma Community Combined	44	2556	1720 (1258, 2258)
Jimma Rural	23	1261	1824 (1157, 2727)
Jimma Urban	21	1295	1622 (1004, 2479)
Addis Ababa Community Combined*	46	1017	4535 (3372, 5906)
Yeka sub-city	24	557	4309 (2761, 6412)
Addis Ketema sub-city	19	409	4646(2797, 7255)

**New seropositives and person-weeks from Yeka and Addis Ketema sub-cities do not add up due to missing data for sub-city.*
CI – Credible Interval; HCW – Healthcare worker

Between cohorts, we observed statistically significant differences for seroincidence and seroprevalence during different survey periods (Table S2). For seroprevalence over time, we do not see much difference between Round 1 and Round 2 except for Addis HCW, which is by design and expected. However, the difference to Round 3 is statistically significant (Table S3).

Table S2: Difference in the seroincidence and seroprevalence during survey periods between communities and health care workers (HCW) observed in Addis Ababa and Jimma.

Seroincidence (HCW)	**RR (95% CI)†**
Addis community versus Jimma community	2·6 (1·6; 3·8)*
Addis HCW versus Jimma HCW	0·6 (0·4; 0·7)*
Addis community versus Addis HCW	2·0 (1·4; 2·8)*
Jimma community versus Jimma HCW	0·4 (0·3; 0·6)*
Seroprevalence (Addis Ababa)	**OR (95% CI)†**
December 2020/January 2021	
Addis Ketema (R1) versus Addis Yeka (R1)	1·8 (1·2; 2·6)*
Addis Ketema (R1) versus Addis HCW (R2)‡	1·5 (1·1; 2·1)*
Addis Yeka (R1) versus Addis HCW (R2)‡	0·8 (0·6; 1·2)
February 2021/March 2021	
Addis Ketema (R2) versus Addis Yeka (R2)	1·6 (1·0; 2·5)*
Addis Ketema (R2) versus Addis HCW (R3)‡	1·2 (0·8; 2·1)
Addis Yeka (R2) versus Addis HCW (R3)‡	0·7 (0·4; 1·1)
April 2021	
Addis Ketema (R3) versus Addis Yeka (R3)‡	2·2 (1·3; 3·3)*
Seroprevalence Jimma	
November 2020/December 2021	
Jimma City (R1) versus Jimma Rural (R1)	2·2 (1·4; 3·2)*
Jimma City (R1) versus HCW (R1)	1·1 (0·7; 1·4)
Jimma Rural (R1) versus HCW (R1)	0·5 (0·3; 0·7)*
January 2021/February 2021	
Jimma City (R2) versus Jimma Rural (R2)	1·9 (1·1; 3·0)*
Jimma City (R2) versus Jimma HCW (R2)	0·8 (0·6; 1·1)
Jimma Rural (R2) versus Jimma HCW (R2)	0·4 (0·3; 0·6)*
February 2021/March	
Jimma City (R3) versus Jimma Rural (R3)	1·8 (0·9; 2·9)
Jimma City (R3) versus Jimma HCW (R3)	0·6 (0·4; 0·9)*
Jimma Rural (R3) versus Jimma HCW (R3)	0·3 (0·2; 0·5)*

†Estimate –ratio for the comparison of the contrasts, RR=risk ratio for seroincidence, OR=odds ratio for seroprevalence
* Statistically significant; R= survey round

Note: in order to compare seroprevalences between cohorts, we applied periods instead of round. This distinction was made as in Addis Ababa survey rounds in HCW did not match those of the corresponding communities in terms of time periods (initiated with ‡). In April, no matching HCW information from Addis was available.

Table S3: Difference in the seroprevalence over the different rounds for the overall population and by cohort

Effects	**Odds Ratio***	**Lower 95%CI**	**Upper 95% CI**	**Statistically significant difference**
Intercept	1·403	1·020	1·937	-
Yeka Sub-city	0·597	0·379	0·935	Yes
Jimma City	0·490	0·316	0·755	Yes
Jimma Rural	0·252	0·152	0·414	Yes
Jimma Medical Center	0·598	0·410	0·865	Yes
St· Paul's Hospital Addis	0·549	0·371	0·814	Yes
Overall Round 1	0·840	0·551	1·259	No
Yeka Sub-cityRound1 (interaction)	0·935	0·523	1·704	No
Jimma City Round1 (interaction)	0·825	0·470	1·443	No
Jimma Rural Round1 (interaction)	0·737	0·386	1·433	No
Jimma Medical Center Round1 (interaction)	0·631	0·387	1·031	No
St· Paul's Hospital Addis Round1 (interaction)	0·187	0·108	0·323	Yes
Overall Round 3	1·918	1·213	3·047	Yes
AddisYeka Round3 (interaction)	0·755	0·403	1·414	No
Jimma City Round3 (interaction)	0·624	0·337	1·161	No
Jimma Rural Round3 (interaction)	0·657	0·318	1·359	No
Jimma Medical Center Round3 (interaction)	0·798	0·464	1·369	No
St· Paul's Hospital Addis Round3 (interaction)	0·788	0·420	1·460	No

*Round 2 is reference category; Addis Ketema is reference site

In the above table, we see that the interaction effects are not significantly different except for the Round 1 at St. Paul's Hospital (Addis Ababa), which is a design effect. Overall, ignoring the interaction effect, we observed no significant difference between Round 1 and Round 2; however, Round 3 compared to Round 2 had an overall increase (OR 1·918 with 95% Credible Interval (1·213-3·047)). We also observe that within Round 2, Addis Ketema sub-city had the highest seroprevalence as compared all the other cohort groups.

2. The Models

We considered three different models for the analysis of the virus spread in Ethiopia: A simple SEIR model (which was applied separately to data for *healthcare workers (H)* or *community members (C)*), an extended SEIR model which simultaneously described the populations for healthcare worker and community members, and an SEIR model which

allows for the original virus (wt) and a virus variant (va). We chose SEIR models due to their widespread use for the study of the Covid-19 progression,[5–9] which facilitates a comparison to related work. Furthermore, we established earlier a comprehensive analysis pipeline for these types of models.[10] In all these models, the populations are split into *Susceptible (S), Exposed (E), Infectious (I)* and *Recovered (R).* To compare the model simulations to the observed seroprevalence, we compute the ratio of recovered to total population.

a) SEIR model

The model structure is depicted in Figure 4A and the corresponding ordinary differential equations (ODEs) for the time-dependent size of the compartments are:

$$\dot{S} = -\frac{\beta I}{N} S \qquad S(0) = 510$$
$$\dot{E} = \frac{\beta I}{N} S - \kappa E \qquad E(0) = 0$$
$$\dot{I} = \kappa E - \gamma I \qquad I(0) = I_0$$
$$\dot{R} = \gamma I \qquad R(0) = 0$$
$$N = S + E + I + R.$$

The parameters are listed in Table S4. This table includes the respective names in the PEtab model which we published at Zenodo.[4]

Table S4: Parameters of the SEIR model. Some depend on study site, i.e. Jimma and Addis Ababa.

Parameter	Description	Sampling result - Median (CI 95%)	Scale used for sampling	Prior (in scale)	Est. Start Sampling	Unit
β	Exp. rate	0·08 (0·06, 0·13)	log_{10}	$\mathcal{U}(-5, 1)$	0·09	$\frac{1}{\text{day}}$
κ^{-1}	Inc. period	5·6 (2·2, 13·6)	log	$\mathcal{N}(1 \cdot 63, 0 \cdot 50)$	5·0	days
γ^{-1}	Rec. time	19·3 (11·4, 28·9)	linear	$\mathcal{N}(15 \cdot 7, 6 \cdot 7)$	15·0	days
I_0	Initial inf.	J: 1·1 (0·3, 3·1) A: 1·2 (0·4, 2·9)	log_{10}	$\mathcal{U}(-1, 3)$	J: 0·74 A: 6·5	-

b) Extended SEIR model for two populations

In addition to the dynamics of the individual populations, we account for their interaction: Infectious healthcare workers can expose community members and vice versa. Virus transmission from community members to healthcare workers is supposed to be more probable, which is modeled by a factor $\alpha > 1$. The model structure can be seen in Figure 4C and the ODEs are:

$$\dot{S_H} = -\frac{\beta(I_H + \alpha I_C)}{N} S_H \qquad S_H(0) = 510$$
$$\dot{E_H} = \frac{\beta I_H}{N} S_H - \kappa E_H \qquad E_H(0) = 0$$
$$\dot{I_H} = \kappa E_H - \gamma I_H \qquad I_H(0) = 0$$
$$\dot{R_H} = \gamma I_H \qquad R_H(0) = 0$$
$$\dot{S_C} = -\frac{\beta(I_H + I_C)}{N} S_C \qquad S_C(0) = 100000$$
$$\dot{E_C} = \frac{\beta I_C}{N} S_C - \kappa E_C \qquad E_C(0) = 0$$
$$\dot{I_C} = \kappa E_C - \gamma I_C \qquad I_C(0) = I_0$$
$$\dot{R_C} = \gamma I_C \qquad R_C(0) = 0$$
$$N = S_H + E_H + I_H + R_H + S_C + E_C + I_C + R_C.$$

The parameters are listed in Table S5. This table includes the respective names in the PEtab model which we published at Zenodo.[4.] All initial states which are not mentioned in the table are set to 0.

Table S5: Parameters of the extended SEIR model. Some depend on study site, i.e. Jimma and Addis Ababa.

Parameter	Description	Sampling result - Median (CI 95%)	Scale used for sampling	Prior (in scale)	Est. Start Sampling	Unit
β	Exp. rate	0·08 (0·06, 0·10)	log_{10}	$\mathcal{U}(-5, 1)$	0·08	$\frac{1}{\text{day}}$
κ^{-1}	Inc. period	5·4 (2·6, 11·0)	log	$\mathcal{N}(1 \cdot 63, 0 \cdot 50)$	5·7	days
γ^{-1}	Rec. time	19·8 (14·9, 26·3)	linear	$\mathcal{N}(15 \cdot 7, 6 \cdot 7)$	18·5	days
α	Increased HCW risk	1·5 (1·3, 1·7)	log_{10}	$\mathcal{U}(-1, 3)$	1·5	-
I_0	Initial inf.	J: 131·4 (56·8, 293·3) A: 204·3 (96·7, 428·2)	log_{10}	$\mathcal{U}(-1, 3)$	J: 121·9 A: 189·9	-

c) Virus variant model

This model accounts for the possibility that a virus variant altered characteristics is present in Ethiopia. As sequencing data are missing, we assume the variant to appear at an unknown time t_0 and has a reproduction rate increased by a factor of 1·35, which is in the range of increase observed for variants such as B·1·1·7 and B·1·351. We account for the increase by reducing the recovery rate.[11] Moreover we assume previous variant infections make individuals immune to wild type infections but not the other way around.

The model structure is depicted Figure 5A and the ODEs are:

$$
\begin{aligned}
\dot{S} &= -\frac{\beta I_{wt}}{N} S - \frac{\beta (I_{va} + I_{va}^{wt})}{N} S & S(0) &= 510 \\
\dot{E_{wt}} &= \frac{\beta I_{wt}}{N} S - \kappa E_{wt} & E_{wt}(0) &= 0 \\
\dot{E_{va}} &= \frac{\beta (I_{va} + I_{va}^{wt})}{N} S - \kappa E_{va} & E_{va}(0) &= 0 \\
\dot{E_{va}^{wt}} &= \frac{\beta (I_{va} + I_{va}^{wt})}{N} R_{wt} - \kappa E_{va}^{wt} & E_{va}^{wt}(0) &= 0 \\
\dot{I_{wt}} &= \kappa E_{wt} - \gamma I_{wt} & I_{wt}(0) &= I_0 \\
\dot{I_{va}} &= \kappa E_{va} - \frac{\gamma}{1.35} I_{va} & I_{va}(t_0) &= 1 \\
\dot{I_{va}^{wt}} &= \kappa E_{va}^{wt} - \frac{\gamma}{1.35} I_{va}^{wt} & I_{va}^{wt}(0) &= 0 \\
\dot{R_{wt}} &= \gamma I_{wt} - \frac{\beta (I_{va} + I_{va}^{wt})}{N} R_{wt} & R_{wt}(0) &= 0 \\
\dot{R_{va}} &= \frac{\gamma}{1.35} I_{va} & R_{va}(0) &= 0 \\
\dot{R_{va}^{wt}} &= \frac{\gamma}{1.35} I_{va}^{wt} & R_{va}^{wt}(0) &= 0 \\
N &= S + E_{wt} + E_{va} + E_{va}^{wt} + I_{wt} \\
&\quad + I_{va} + I_{va}^{wt} + R_{wt} + R_{va} + R_{va}^{wt}.
\end{aligned}
$$

The parameters are listed in Table S6. This table includes the respective names in the PEtab model which we published at Zenodo.[4]

Table S6: Parameters of the virus variant model. Some depend on study site, i.e. Jimma and Addis Ababa.

Parameter	Description	Sampling result - Median (CI 95%)	Scale used for sampling	Prior (in scale)	Est. Start Sampling	Unit
β	Exp. rate	0·08 (0·06, 0·10)	log_{10}	$\mathcal{U}(-5, 1)$	0·08	$\frac{1}{\text{day}}$
κ^{-1}	Inc. period	5·0 (2·4, 10·0)	log	$\mathcal{N}(1 \cdot 63, 0 \cdot 50)$	5·3	days
γ^{-1}	Rec. time	16·7 (12·9, 22·1)	linear	$\mathcal{N}(15 \cdot 7, 6 \cdot 7)$	17·2	days
t_0	Entry va	184·5 (152·6, 231·3)	linear	$\mathcal{U}(150, 360)$	170·3	days
s_{TPR}	Scaling nat. TPR	J: 2·3 (1·5, 3·6) A: 2·8 (1·7, 4·3)	log_{10}	$\mathcal{U}(-1, 3)$	J: 2·3 A: 2·7	-
I_0	Initial inf.	J: 1·8 (0·6, 4·9) A: 13·8 (3·6, 42·5)	log_{10}	$\mathcal{U}(-1, 3)$	J: 2·2 A: 16·2	-

d) Calibration workflow

The models were encoded using the Systems Biology Markup Language (SBML)[12] and the Parameter estimation problems were formulated using the Parameter Estimation table (PEtab)[13] standard. The two community standards allow for the direct reproduction of the result in various software tools.

For parameter estimation, the seroprevalence data for each site, round and study group was each split by month of their collection and then accumulated on the mean date respectively. Standard deviations were calculated assuming binomial distribution in a similar way as described in the paragraph *Prevalence and Incidence Estimation* of this section. The seroprevalence measurement is assumed to not distinguish between infection with original virus or variant. In addition to seroprevalence information, we used for the virus variant model also information about national test positivity rates (TPR). As over a long time the number of test and test strategies remained unchanged, we assumed that the TPR is roughly proportional to the sum of exposed and infectious individuals in the different groups and location. For incubation and recovery times we used priors from literature.[14,15]

Bayesian parameter estimation was performed using the adaptive Metropolis-Hastings algorithm methods implemented in the parameter estimation toolbox pyPESTO[16]. Selected results were confirmed using pyMC3. Simulation was performed using the simulation toolbox AMICI[17]. The sampling results were post-processed, e.g. by removing the burn-in, and convergence was assessed visually and using the Geweke test.

Supplementary Results for model prediction

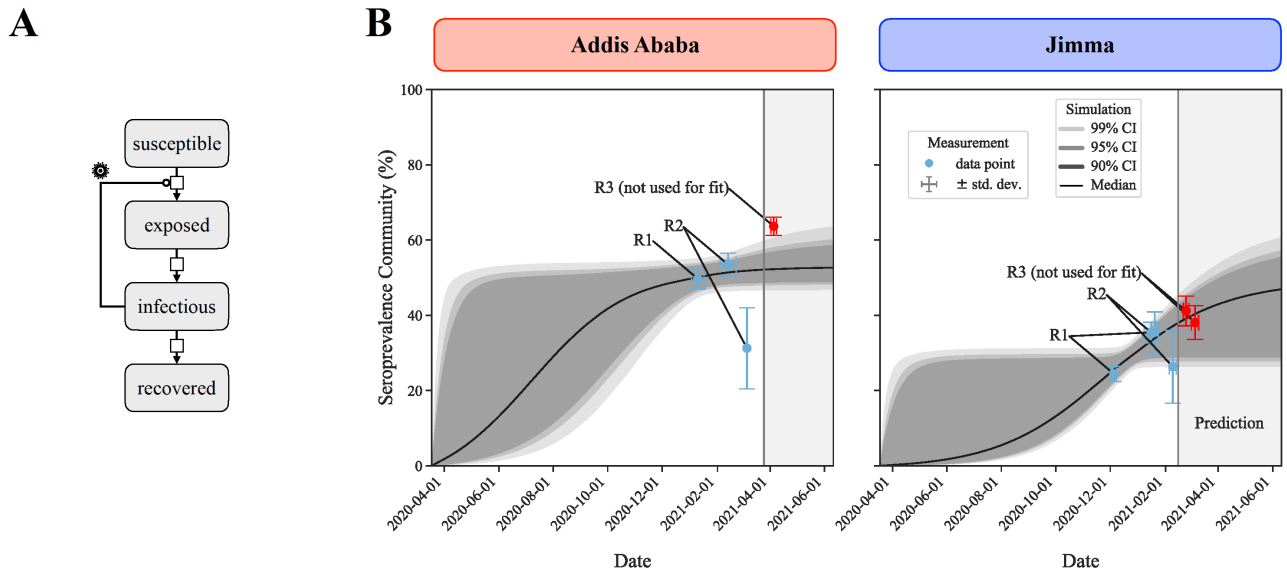
The parameter sampling for the SEIR model with healthcare workers data was performed with a sample size of 1e6. Convergence of parameters was achieved after a burn in of 5e4 samples.

The parameter sampling for the extended SEIR model for two populations with combined healthcare workers and community data was performed with a sample size of 1e5. Convergence of parameters was achieved without any burn.
The parameter sampling for the virus variant model with combined community and national TPR data was performed with a sample size of 1e5. Convergence of parameters was achieved after a burn in of 1e4 samples.

The parameter sampling for the SEIR model with combined community members data was performed with a sample size of 1e6. Since the parameters showed alternating behaviour between two models, we refrained from conducting prediction simulations based on this model-data combination.

The parameter sampling for the SEIR model with combined community members data was performed with a sample size of 1e6. Since the parameters showed alternating behaviour between two models, we refrained from conducting prediction simulations based on this model-data combination. For completeness we included these prediction results as Figure S1.

Figure S1: SEIR model of SARS-CoV-2 epidemic in Ethiopia.



(A) Compartments of the SEIR models and possible transition. **(B)** Model simulation for Community members in Jimma Medical Center and St. Paul's Hospital. Data from the 1st and 2nd round was used for model training. Later points, including the 3rd round, were predictions.

3. Information on missing data

The following tables describe the numbers and percentages of missing data between rounds (A. between Round 1 and Round 2; B (between Round 2 and Round 3); C. between Round 1 and Round 3) and for different cohorts (1. HCW Jimma, 2. urban and rural community combined for Jimma, C. HCW Addis Ababa, D. Addis community combined (Ketema and Yeka). Overall, dropout rates are higher, especially in Addis Ababa as compared to Jimma. However, dropout rates do not significantly differ between seropositive and seronegative population, which indicates that there was no sampling bias over the entire period of the study.

1. Jimma Health Care Workers (HCW) Missing Data by Result

A

	Round 2				
Round 1	**Negative**	**Positive**	**Missing**	**(all)**	**Round 2 Missing %**
Negative	235	66	52	353	14·73%
Positive	1	132	24	157	15·29%
(all)	236	198	76	510	

B

	Round 3				
Round 2	**Negative**	**Positive**	**Missing**	**(all)**	**Round 3 Missing %**
Negative	152	43	41	236	17·37%
Positive	3	162	33	198	16·67%
Missing	7	5	64	76	
(all)	162	210	138	510	

C

		Round 3				
Round 1	**Round 2**	**Negative**	**Positive**	**Missing**	**(all)**	**Round 3 Missing %**
Negative	Negative	151	43	41	235	17·45%
Negative	Positive	1	55	10	66	15·15%
Negative	Missing	7	2	43	52	
Negative	(all)	159	100	94	353	26·63%
Positive	Negative	1	0	0	1	
Positive	Positive	2	107	23	132	17·42%
Positive	Missing	0	3	21	24	
Positive	(all)	3	110	44	157	28·03%
(all)	(all)	162	210	138	510	

2. Jimma Community (combined Jimma City and Jimma urban)

A

	Round 2				Round 2 Missing %
Round 1	**Negative**	**Positive**	**Missing**	**(all)**	
Negative	207	31	158	396	39·90%
Positive	4	82	53	139	38·13%
(all)	211	113	211	535	

B

	Round 3				Round 3 Missing %
Round 2	Negative	Positive	Missing	(all)	
Negative	124	6	81	211	38·39%
Positive	4	78	31	113	27·43%
Missing	32	22	157	211	
(all)	160	106	269	535	

C

		Round 3				Round 3 Missing %
Round 1	**Round 2**	**Negative**	**Positive**	**Missing**	**(all)**	
Negative	Negative	121	6	80	207	38·65%
Negative	Positive	1	19	11	31	35·48%
Negative	Missing	32	7	119	158	
Negative	(all)	154	32	210	396	53·03%
Positive	Negative	3	0	1	4	
Positive	Positive	3	59	20	82	24·39%
Positive	Missing	0	15	38	53	
Positive	(all)	6	74	59	139	42·45%
(all)	(all)	160	106	269	535	

3. Addis Health Care Workers (HCW) Missing Data by Result

A

	Round 2				
Round 1	**Negative**	**Positive**	**Missing**	**(all)**	**% Missing Round 2**
Negative	103	53	275	431	63·81%
Positive	5	22	25	52	48·08%
Missing	56	48	0	104	
(all)	164	123	300	587	

B

	Round 3				
Round 2	**Negative**	**Positive**	**Missing**	**(all)**	**% Missing Round 3**
Negative	28	27	109	164	66·46%
Positive	6	22	95	123	77·24%
Missing	18	13	269	300	
(all)	52	62	473	587	

C

		Round 3				
Round 1	**Round 2**	**Negative**	**Positive**	**Missing**	**(all)**	**% Missing Round 3**
Negative	Negative	19	12	72	103	69·90%
Negative	Positive	4	4	45	53	84·91%
Negative	Missing	17	11	247	275	89·82%
Negative	(all)	40	27	364	431	84·45%
Positive	Negative	0	1	4	5	
Positive	Positive	0	4	18	22	81·82%
Positive	Missing	1	2	22	25	88·00%
Positive	(all)	1	7	44	52	84·62%
Missing	Negative	9	14	33	56	
Missing	Positive	2	14	32	48	
Missing	(all)	11	28	65	104	
(all)	(all)	52	62	473	587	

4. Addis Community (combined for Ketema and Yeka)

A

	Round2				
Round1	**Negative**	**Positive**	**Missing**	**(all)**	**% Missing Round 2**
Negative	84	22	62	168	36·90%
Positive	5	92	68	165	41·21%
Missing	48	36	259	343	
(all)	137	150	389	676	

B

	Round3				
Round2	**Negative**	**Positive**	**Missing**	**(all)**	**% Missing Round 3**
Negative	11	10	116	137	84·67%
Positive	12	40	98	150	65·33%
Missing	112	185	92	389	
(all)	135	235	306	676	

C

		Round3				
Round1	**Round2**	**Negative**	**Positive**	**Missing**	**(all)**	**% Missing Round 3**
Negative	Negative	9	6	69	84	82·14%
Negative	Positive	0	5	17	22	77·27%
Negative	Missing	14	15	33	62	53·23%
Negative	(all)	23	26	119	168	70·83%
Positive	Negative	1	1	3	5	
Positive	Positive	8	29	55	92	59·78%
Positive	Missing	3	6	59	68	86·76%
Positive	(all)	12	36	117	165	70·91%
Missing	Negative	1	3	44	48	
Missing	Positive	4	6	26	36	
Missing	Missing	95	164	0	259	
Missing	(all)	100	173	70	343	
(all)	(all)	135	235	306	676	

5. Seroprevalence among complete cases for Jimma

Complete cases	Round	Observed Individuals	Seropositivity	Estimated Seroprevalence reported in manuscript
Jimma HCW	1	360	30·60%	30·8% (26·9%, 34·8%)
	2	360	45·80%	45·6% (41·0%, 50·3%)
	3	360	56·90%	56·1% (51·1%, 61·1%)
Jimma Urban	1	132	38·60%	32·3% (27·0%, 37·9%)
	2	132	47·00%	40·8% (33·9%, 47·9%)
	3	132	47·00%	45·2% (37·7%, 52·7%)
Jimma Rural	1	80	17·50%	18·0% (13·5%, 23·2%)
	2	80	25·00%	26·3% (19·1%, 34·3%)
	3	80	27·50%	31·0% (22·3%, 40·3%)

Supplementary References

1 McElreath R. Statistical Rethinking: A Bayesian Course With Examples in R and Stan. Chapman & Hall/CRC, 2016 https://books.google.com/books/about/Statistical_Rethinking.html?hl=&id=d7fCsgEACAAJ.

2 Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software. 2017; **80**. DOI:10.18637/jss.v080.i01.

3 Bürkner P-C. Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal. 2018; **10**: 395–411.

4 Gudina EK, Ali S, Girma E, *et al.* Supplementary code and models for 'Silent spread of SARS-CoV-2 in Ethiopia: Longitudinal cohort study among frontline healthcare workers and community'. 2021; published online June 7. DOI:10.5281/zenodo.4885064.

5 Ahmad Z, Arif M, Ali F, Khan I, Nisar K S. A report on COVID-19 epidemic in Pakistan using SEIR fractional model. *Scientific Reports* 2020; **10**: 22268.

6 Annas S, Pratama M I, Rifandi M, Sanusi W, Side S. Stability analysis and numerical simulation of SEIR model for pandemic COVID-19 spread in Indonesia. *Chaos, Solitons & Fractals* 2020; **139**: 110072.

7 Grimm V, Mengel F, Schmidt M. Extensions of the SEIR model for the analysis of tailored social distancing and tracing approaches to cope with COVID-19. *Scientific Reports* 2021; **11**: 4214.

8 López L, Rodó X. A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: Simulating control scenarios and multi-scale epidemics. *Results in Physics* 2021; **21**: 103746.

9 Prem K, Liu Y, Russell T W, Kucharski A J, *et al*. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health* 2020; **5**: e261–e270.

10 Raimúndez E, Dudkin E, Vanhoefer J, *et al*. COVID-19 outbreak in Wuhan demonstrates the limitations of publicly available case numbers for epidemiological modeling. *Epidemics* 2021; **34**: 100439.
11 Kissler SM, Fauver JR, Mack C, *et al.* Densely sampled viral trajectories suggest longer duration of acute infection with B.1.1.7 variant relative to non-B.1.1.7 SARS-CoV-2. *medRxiv* 2021. DOI:10.1101/2021.02.16.21251535.
12 Keating SM, Waltemath D, König M, *et al.* SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol Syst Biol* 2020; **16**: e9110.
13 Schmiester L, Schälte Y, Bergmann FT, *et al.* PEtab—Interoperable specification of parameter estimation problems in systems biology. PLOS Computational Biology. 2021; **17**: e1008646.
14 Fang Z, Zhang Y, Hang C, Ai J, Li S, Zhang W. Comparisons of viral shedding time of SARS-CoV-2 of different samples in ICU and non-ICU patients. J. Infect. 2020; **81**: 147–78.
15 McAloon C, Collins Á, Hunt K, *et al.* Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open* 2020; **10**: e039652.
16 Schälte Y, Fröhlich F, Stapor P, *et al.* ICB-DCM/pyPESTO: pyPESTO 0.2.6. Zenodo, 2021; published online May DOI:10.5281/ZENODO.4768592.
17 Fröhlich F, Weindl D, Schälte Y, *et al.* AMICI: High-Performance Sensitivity Analysis for Large Ordinary Differential Equation Models. *Bioinformatics* 2021; published online April. DOI:10.1093/bioinformatics/btab227.