

SUPPLEMENTARY MATERIALS

1. MODEL FOR COVID-19 INFECTIONS

Let $R_{\ell,A}^*$ be the number of individuals who tested seropositive in age group A at location ℓ , and $n_{\ell,A}$ give the number of individuals tested in that age group for this location. We model the number of individuals with a positive serology test in the study as

$$R_{\ell,A}^* \sim \text{Binomial}(n_{\ell,A}, p_{\ell,A}), \text{ where} \quad (1)$$

$$p_{\ell,A,j} = \text{sens}_{t_\ell} \pi_{\ell,A} + (1 - \text{spec}_{t_\ell})(1 - \pi_{\ell,A}). \quad (2)$$

To account for the error rates of the test, the test positivity probability, $p_{\ell,A}$, is defined as a function of test sensitivity (sens_{t_ℓ}), test specificity (spec_{t_ℓ}), and the true seroprevalence ($\pi_{\ell,A}$) for the associated location and age group at the time of the study. For many studies, we did not have seropositivity by age, in which case A represented all ages.

To account for uncertainty in the test characteristics, we model the lab validation data directly. Let $n_{\text{sens},t}$ denote the number of positive specimens tested with test t , and $x_{\text{sens},t}$ the number of positive specimens that correctly tested positive. Similarly, let $n_{\text{spec},t}$ and $x_{\text{spec},t}$ denote the number of negative specimens tested and the number of negative specimens that correctly tested negative with test t , respectively. We model these quantities as follows:

$$x_{\text{sens},t} \sim \text{Binomial}(n_{\text{sens},t}, \text{sens}_t) \quad (3)$$

$$x_{\text{spec},t} \sim \text{Binomial}(n_{\text{spec},t}, \text{spec}_t) \quad (4)$$

2. MODEL FOR COVID-19 DEATHS

Let $D_{\ell,A}^*$ give the number of recorded COVID-19 deaths, for age group A at location ℓ . Note that if only a single death record is available, then A represents the entire age range. We model the recorded COVID-19 deaths as

$$D_{\ell,A}^* \sim \text{Poisson}(N_{\ell,A} \times \pi_{\ell,A} \times \text{IFR}_{\ell,A}) \quad (5)$$

where $N_{\ell,A}$ gives the number of individuals at location ℓ in age group A . Then $N_{\ell,A} \times \pi_{\ell,A}$ gives the expected number of infected individuals, and $\text{IFR}_{\ell,A}$ is the infection fatality rate for location ℓ and age group A , representing the probability an individual dies from COVID-19, given the individual had COVID-19. Note, the Poisson distribution reflects the relative rarity of a COVID-19 fatality relative to the entire population.

3. ACCOUNTING FOR DATA COLLECTED IN VARYING AGE BINS

Notice that the models above for deaths and infections in (1) and (5) are functions of prevalence and IFR, respectively, defined on discrete age bins. However, the discrete age bins are not necessarily the same for the death data and the seroprevalence studies. The following adjustments were made to match serology and death age bins:

- **Death bins nested within a serology bin:** We aggregate deaths for each location to match the respective serology age bins to avoid placing assumptions about the variability of prevalence across ages within a single serology age bin.
- **Serology bins nested within a death bin:** The average seroprevalence for the death age bin is calculated as an average of the serology age bins, weighted by the percent of the population in each age bin.
- **Bin endpoints slightly off:** When age bins were within one or two years of matching, serology age bins were adjusted to match the corresponding death age bins.

All modifications to age bins are documented in a spreadsheet in the data folder.

3.1. Population age distribution

Let $f_\ell(a)$ denote the number of individuals of age a at location ℓ for $a \in \{0, 1, \dots, 84+\}$. Note, if population age structure is only available in 5 year age bins, then define $f_\ell(a) = \sum_{b \in \{0, 5, \dots, 80\}} \frac{f_\ell([b, b+5]) \mathbb{I}_{[b, b+5)}(a)}{5}$ where $f_\ell([b, b+5))$ is the proportion of the population ages $[b, b+5)$.

In cases where the location specific age structure is only available in large bins, but the national age structure is available in 5 year age bins, we leverage the national age structure to inform the location specific age structure as follows. Let A denote an interval the location specific age structure is available for (e.g., $[0, 18)$). If $f(A)$ is the proportion of the population at location ℓ with an age in A and $f_n(a)$ is the proportion of the population aged a at the national level, then we estimate $f_\ell(a)$, the proportion at location ℓ that is age a , as

$$f_\ell(a) = f(A) \frac{f_n(a)}{\sum_{b \in A \cap \mathbb{N}} f_n(b)}. \quad (6)$$

Essentially, we rescale $f_n(a)$ such that the total mass in A matches the observed total mass in A at location ℓ , $f(A)$. Since we model seroprevalence as constant past age 85, we let $f_\ell(85)$ represent the proportion of the population aged 85 or older, rather than just the proportion aged 85.

Define the population age density for age bin A as

$$f_{\ell, A}(a) = \frac{f_\ell(a)}{\sum_{b \in A \cap \mathbb{N}} f_\ell(b)}, \quad a \in \{0, 1, \dots, 84+\} \quad (7)$$

in order to truncate $f_\ell(a)$ to age bin A .

The prevalence for age bin $B = \cup_{A \in \mathcal{A}} A$ is then defined

$$\pi_{\ell, B} = \sum_{A \in \mathcal{A}} \left[\pi_{\ell, A} \sum_{b \in A \cap \mathbb{N}} f_{\ell, A}(b) \right]. \quad (8)$$

For the locations where we only have serology study information with no corresponding fatality data, the proportion of all study participants that were in a given age bin was assumed representative of the proportion of the population in each age bin since the studies were designed to have representative samples. For the locations with both serology and fatality data, population data were recorded in the Population Distributions tabs with citations.

4. PRIORS

For the seroprevalence parameters, $\pi_{\ell, A}$, we used independent, weakly informative priors:

$$\pi_{\ell, A} \sim \text{Beta}(2, 6) \quad \text{for all } \ell, A. \quad (9)$$

We also used independent priors for the test sensitivities and specificities. Because there are infinitely many combinations of prevalence, sensitivity, and specificity, that can result in the same test positivity rate, similar to Gelman and Carpenter (2020), we used informative priors to avoid a multimodal posterior. For each test assay t , the priors on the sensitivity and specificity were

$$\text{sens}_t \sim \text{Beta}(10, 1) \quad (10)$$

$$\text{spec}_t \sim \text{Beta}(50, 1). \quad (11)$$

To further narrow the seroprevalence, sensitivity, and specificity combinations, we used independent, mildly informative priors for each IFR parameter based on expert knowledge. IFR for COVID-19 is known to increase with age. We also expect IFR to be more extreme (smaller than average or larger than average) when the age bin is small. For example, we would expect an age bin from 20-80 to look similar to the country average, but we would expect an age bin from 70-80 to be much higher than the country average. To formulate a prior that reflects these characteristics, we modeled

$$\text{IFR}_{\ell, A} \sim \text{Beta}(1, \text{IFR}_{\ell, A}^{\text{prior}}) \quad (12)$$

where

$$\text{IFR}_{\ell,A}^{\text{prior}} = 30 - 20 \left[\frac{U_{\ell,A} - 50}{50} \left(1 - \frac{U_{\ell,A} - L_{\ell,A}}{100} \right) \right]. \quad (13)$$

The lower and upper bounds of age bin A at location ℓ are given by $L_{\ell,A}$ and $U_{\ell,A}$, respectively. For open ended upper ages, we set $U_{\ell,A} = 100$. As an example, $\text{IFR}_{\ell,[0,100]}^{\text{prior}} = 30$, while $\text{IFR}_{\ell,[80,100]}^{\text{prior}} = 14$, allowing for larger IFR estimates when focusing on the older individuals.

5. MODEL IMPLEMENTATION

The model was implemented in version 4.0.2 of the programming language R, and posterior samples were obtained via the software package Stan (version 2.21.1). We ran three chains for 10,000 iterations, where the first 5,000 iterations were discarded as warm-up samples. All parameters had an effective sample size greater than 2,800. Additionally, the \widehat{R} value was within 0.0034 of 1 for each parameter, suggesting convergence. Examination of traceplots also suggested convergence.

We compared plugin estimates for parameters to the posterior distribution for each parameter to check model fit. In each case there was good agreement, or the Bayesian estimate was superior. For example, in Figure 1, we compare the posterior distribution of the sensitivity and specificity estimates to the raw estimate. In most cases, the middle 50% of the posterior distribution contains the raw estimate. However, for test kit ID 11 (Qingdao Hightop Biotech IgM/IgG Duo), the raw estimate of specificity is outside the range of the posterior draws. In the case of this test, it was used in locations with extremely low prevalence, such that the Gladen-Rogan adjustment¹ results in a negative estimate of prevalence (meaning the expected number of false positives is greater than the number that tested positive.) Since this is unreasonable, the Bayesian model raises the specificity estimate, lowering the expected number of false positives.

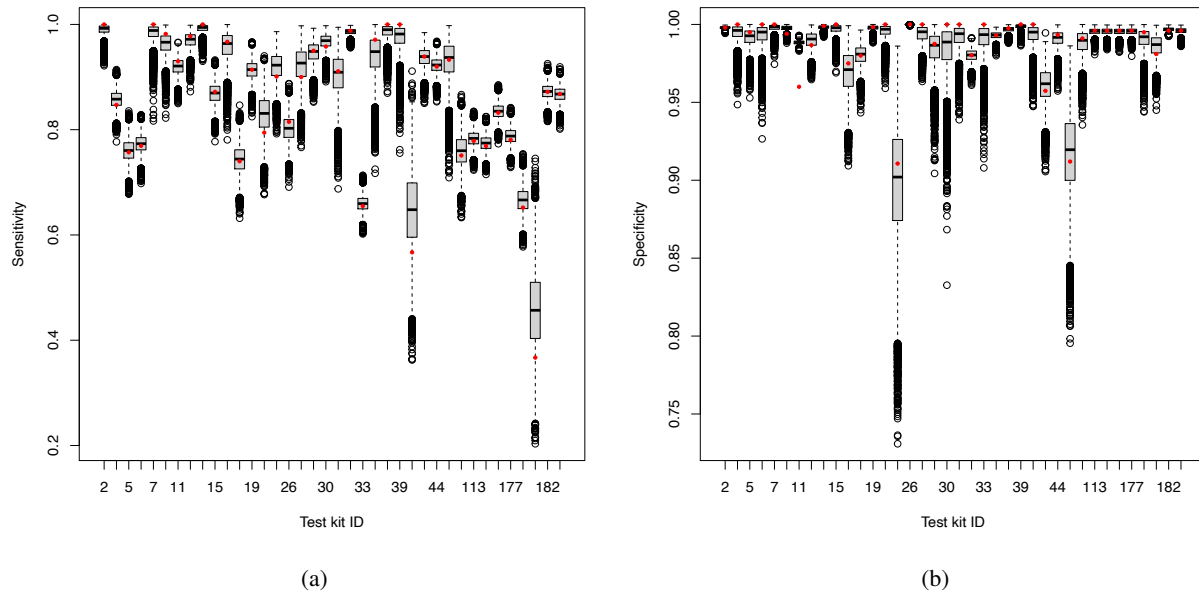


Figure 1. Boxplots showing the posterior distribution of the (a) sensitivity and (b) specificity for each test kit. Raw sensitivity/specificity based on the lab validation data is shown by a red point for each test kit.

¹Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. American Journal of Epidemiology. 1978;107(1):71-6. doi:10.1093/oxfordjournals.aje.a112510

6. MODEL OUTPUTS

For each model parameter, we use the posterior mean as the point estimate and produce 95% equal-tail credible intervals to describe uncertainty.

6.1. Total Seroprevalence

Similar to calculating the average seroprevalence for a death age bin, we estimate total seroprevalence for a location by taking an average of the age bin seroprevalences, weighting by the population distribution at that location:

$$\pi_{\ell, [0, 100+)} = \sum_{A \in \mathcal{A}_\ell} \left[\pi_{\ell, A} \sum_{b \in A \cap \mathbb{N}} f_{\ell, A}(b) \right] \quad (14)$$

where \mathcal{A}_ℓ are the serology age bins associated with location ℓ .

6.2. Assessing uniformity of seroprevalence across age

We calculated total seroprevalence from approximately ages 18 to 59 and ages 40 to 59, compared to ages approximately 60 and older. The age bins used for each location were selected as follows:

- **18 to 59:** Any age bins such that $15 \leq \text{lower age} < 60$ and $20 < \text{upper age} \leq 65$ were included
- **40 to 59:** Any age bins such that $40 < \text{lower age} < 60$ and $40 < \text{upper age} \leq 65$ were included
- **60+:** Any age bins such that $60 \leq \text{lower age}$ were included.

This resulted in sets of bin where the 18-59 bins and the 40-59 bins did not overlap the 60+ bins.

We then calculated total seroprevalence from age a to b using appropriate age bins, \mathcal{A} , similar to equation (8)

$$\pi_{\ell, a-b} = \sum_{A \in \mathcal{A}} \pi_{\ell, A} \frac{n_{\ell, A}}{\sum_{B \in \mathcal{A}} n_{\ell, B}} \quad (15)$$

where $\frac{n_{\ell, A}}{\sum_{B \in \mathcal{A}} n_{\ell, B}}$ estimates the percent of the total population in that age bin, assuming representative age distributions in the serology studies.

For each draw from the posterior distribution, we calculated $\pi_{\ell, a-b}$ for each of our three age intervals of interest. We then calculated the ratios $\pi_{\ell, 60+} / \pi_{\ell, 18-59}$ and $\pi_{\ell, 60+} / \pi_{\ell, 40-59}$ for each draw.

6.3. Total IFR and Comparison to EJE Prediction

6.3.1 Total IFR

Suppose location ℓ has death age bins \mathcal{A}_ℓ . Let $\sum_{b \in A \cap \mathbb{N}} f_{\ell, A_i}(b) = \text{pop}_{\ell, A}$ for $A \in \mathcal{A}_\ell$. Then

$$\begin{aligned} \text{IFR}_{\text{total}} &= \frac{\text{number of deaths}}{\text{number of infections}} \\ &= \frac{\sum_{A \in \mathcal{A}_\ell} \text{IFR}_{\ell, A} \times \pi_{\ell, A} \times \text{pop}_{\ell, A}}{\sum_{B \in \mathcal{A}_\ell} \pi_{\ell, B} \times \text{pop}_{\ell, B}} \\ &= \sum_{A \in \mathcal{A}_\ell} \text{IFR}_{\ell, A} \times \left(\frac{\pi_{\ell, A} \times \text{pop}_{\ell, A}}{\sum_{B \in \mathcal{A}_\ell} \pi_{\ell, B} \times \text{pop}_{\ell, B}} \right) \end{aligned} \quad (16)$$

estimates the total IFR for location ℓ . By calculating $\text{IFR}_{\text{total}}$ for each posterior sample, we can then obtain posterior mean and credible intervals for $\text{IFR}_{\text{total}}$.

6.3..2 High income country benchmark

We compare the IFR estimate to a high income country benchmark based on results from Levin et. al.,² which found a log-linear relationship between age and IFR. Define

$$\text{HICB}_a = \begin{cases} \int_a^{a+1} \frac{1}{100} 10^{-3.27+0.0524a} da & \text{if } a < 85 \\ \frac{1}{100} 10^{-3.27+0.0524(85)} & \text{if } a \geq 85 \end{cases}. \quad (17)$$

Then HICB_a represents the IFR predicted by the high income countries line averaged over the interval $[a, a + 1)$ for ages less than 85 and assumes the high income countries line flattens out and becomes uniform for ages 85 and older.

Then if we assume uniform prevalence for a location, the total IFR estimate over age bin A for high income countries is

$$\text{HICB}_A = \frac{\sum_{a \in A} \text{EJE}_a \times f_\ell(a)}{\sum_{b \in A} f_\ell(b)}. \quad (18)$$

6.3..3 Subsetting to ages 18-65

To estimate the IFR between ages 18 and 65, we used the same strategy in picking age bins as we did when testing uniform prevalence. That is, we selected \mathcal{B}_ℓ to be the death age bins in \mathcal{A} such that the lower age of the bin is greater than or equal to 18 and the upper age of the bin is less than 66. We then applied equation (16), replacing \mathcal{A}_ℓ with \mathcal{B}_ℓ . We were not able to calculate the IFR between 18 and 65 for locations there were no age bins in \mathcal{B}_ℓ .

6.3..4 Baseline population

In order to compare the impact of the age specific IFR while controlling for population age distribution, we calculated the Total IFR substituting $f_\ell(a)$ for a baseline population age distribution, $\bar{f}(a)$ in (16). We calculated $\bar{f}_A(a)$ following (7). The baseline population was calculated as a median across locations for each age, then rescaled to sum to one:

$$\bar{f}_*(a) = \text{median} \{f_\ell(a) \mid \ell \text{ is one of the observed locations with fatality data}\} \quad (19)$$

$$\bar{f}(a) = \frac{\bar{f}_*(a)}{\sum_{a=0}^{85} \bar{f}_*(a)}. \quad (20)$$

We also considered taking the mean across locations for each age and taking the mean across locations for each age after removing the top five and bottom five locations for that age (a censored mean). All three approaches gave similar values as shown in Figure 2.

6.3..5 Country average

To get an average total IFR estimate for each country, we took a weighted average of the total IFR estimate of the locations within that country. We chose to weight by $\frac{1}{\sqrt{n_\ell}}$ in order to give more weight to locations with more certain seroprevalence estimates. In locations with multiple age bins, we took the average across $n_{\ell,A}$ as n_ℓ . We weighted by certainty in the seroprevalence estimates rather than the IFR estimates because larger IFR estimates tend to be due to small seroprevalence estimates and consequently have more uncertainty (i.e., small differences in the denominator, seroprevalence, can result in large changes in the IFR estimate when seroprevalence is small). We did not want to bias the average by down weight all of the larger IFR estimates unless there was a smaller sample size in the seroprevalence.

We followed the same process to estimate the country average IFR between 18 and 65, with the added step of removing any locations in the country where \mathcal{B}_ℓ was empty.

²Levin, A.T., Hanage, W.P., Owusu-Boaitey, N. et al. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol* 35, 1123–1138 (2020) <https://doi.org/10.1007/s10654-020-00698-1>

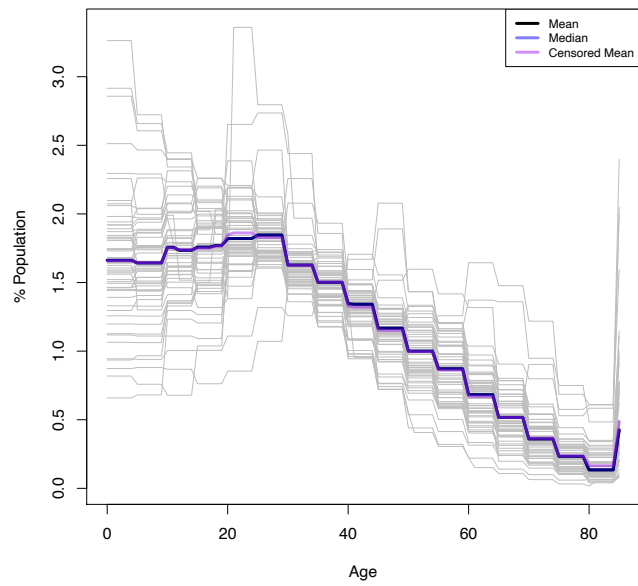


Figure 2. Various estimates of baseline age distribution (mean, median, and censored mean) plotted on age distribution for observed locations in gray.

7. STAN CODE

```
data {
  int num_loc_ifr; // number of locations with fatality data
  int num_loc_all; // total number of locations (with or without fatality data)
  int length_dstar;
  int length_rstar;

  int study_match_pi[length_rstar]; // match test kit to pi_A

  int D_star[length_dstar]; // # deaths each age bin loc1, # deaths loc2, ...
  vector[length_dstar] N; // population each age bin/loc matching D_star
  int R_star[length_rstar]; // # tested positive each age bin loc1, loc2, ...
  int n[length_rstar]; // # tested each age bin loc1, loc2, ...

  int num_test;
  int<lower=0> sens_n[num_test]; // number of positive controls tested
  int<lower=0> sens_x[num_test]; // number of positive controls that tested positive
  int<lower=0> spec_n[num_test]; // number of negative controls tested
  int<lower=0> spec_x[num_test]; // number of negative controls that tested negative

  int min_match_sero_to_death[length_dstar]; // smallest R_star index that overlaps
                                              // corresponding D_star bin
  int max_match_sero_to_death[length_dstar]; // largest R_star index that overlaps
                                              // corresponding D_star bin
  int min_match_total_sero[num_loc_all]; // smallest serology age bin in corresponding
                                          // location
  int max_match_total_sero[num_loc_all]; // largest serology age bin in corresponding
                                          //location

  matrix[86,num_loc_ifr] f_expanded; // population age distribution with ages 0 to 85+
                                      // as rows and location index as columns
  int loc_ind[length_dstar]; // the location index for each death bin
  vector[length_rstar] pi_weights_total; // weights for calculating location total
                                          // seroprevalence
  vector[length_rstar] pi_weights_ag; // weights for aggregating seroprevalence age bins
                                       // to death age bins

  real ifr_prior[length_dstar]; // IFR_{\ell,A}^{\text{prior}}
}

parameters {
  // Seroprevalence
  vector<lower=0, upper=1>[length_rstar] pi_A; // seroprevalence for each location

  // Test characteristics
  vector<lower=0, upper=1>[num_test] sens; // sensitivity for each test
  vector<lower=0, upper=1>[num_test] spec; // specificity for each test

  // IFR
  vector<lower=0, upper=1>[length_dstar] ifr_A; // IFR for each location/age
}

model {
  vector[length_rstar] p_A; // test positivity
  vector[length_dstar] pi_ag; // seroprevalence aggregated to death age bins

  // Positivity //
  p_A = sens[study_match_pi] .* pi_A +
        (1 - spec[study_match_pi]) .* (1-pi_A);

  // Average seroprevalence to match death age bins
  for (i in 1:length_dstar) {
    pi_ag[i] = sum(pi_A[min_match_sero_to_death[i]:max_match_sero_to_death[i]].*
                  pi_weights_ag[min_match_sero_to_death[i]:max_match_sero_to_death[i]]);
  }
}
```

```

// Model //
D_star ~ poisson(N .* pi_ag .* ifr_A);
R_star ~ binomial(n, p_A);

// Priors //

// Prev
pi_A ~ beta(2,6);

// IFR
ifr_A ~ beta(1,ifr_prior);

// Test characteristics
sens_x ~ binomial(sens_n, sens);
spec_x ~ binomial(spec_n, spec);

sens ~ beta(10,1);
spec ~ beta(50, 1);
}

generated quantities {
  vector[length_dstar] pi_ag;
  vector[num_loc_all] pi_bar;

  // Average seroprevalence to match death age bins
  for (i in 1:length_dstar) {
    pi_ag[i] = sum(pi_A[min_match_sero_to_death[i]:max_match_sero_to_death[i]].*
      pi_weights_ag[min_match_sero_to_death[i]:max_match_sero_to_death[i]]);
  }

  // Get total location seroprevalence
  for (i in 1:num_loc_all) {
    pi_bar[i] = sum(pi_A[min_match_total_sero[i]:max_match_total_sero[i]] .*
      pi_weights_total[min_match_total_sero[i]:max_match_total_sero[i]]);
  }
}

```