Appendix: Bayesian model of COVID-19 mortality risk in HCT volunteers

A-1 | INTRODUCTION

This short document is a technical appendix to the paper discussing COVID-19 risks in human challenge trial. Here, we show how a Bayesian model can synthesise information on many infection fatality risks (IFRs) into a single estimate. This estimate is specific to certain age groups and can be further adjusted by e.g. co-morbidity status. The analysis presented here is a form of Bayesian meta-analysis , in that our primary objective is to weigh sources of evidence in a way that captures both variability (here, heterogeneity in real IFRs across different settings) and uncertainty (here, the fact that we do not know the IFRs in each setting precisely).

The ultimate objective of this model is to characterise risk in a way that is useful for design of HCTs. Therefore, as a minimum, we want to incorporate variability across different populations into our prediction. Even better would be to understand how different factors can drive heterogeneity: *a priori* we hypothesise that the three main drivers of differences in IFRs are time-specific, population-specific and otherwise country-specific.[A1]

To characterise differences in observed IFRs we first develop a Bayesian model and apply it to publicly available summary data on IFRs from multiple countries and contexts, with particular focus on the impact of age. This is covered by Section 2. We then use a simple model to hypothesise reduction in risk that may be achieved by screening individuals for comorbidities; this is Section 3. We summarise all results in Section 4.

A-2 | AGE-SPECIFIC RISK OF COVID-19 MORTALITY

A-2.1 | Bayesian evidence synthesis model

What follows is an adaptation of typical methods of Bayesian evidence synthesis to analysis of IFRs. IFR is the ratio of deaths to infections in a given population. Early estimates of COVID-19 mortality risk, e.g. by Verity, Okell, et al. (2020a), placed it at over 0.6%; however, it was also evident from data that IFR could be orders of magnitude higher in particular high risk groups, especially in the elderly, than in the general population.[A2]

By definition, our data on IFRs is a combination of data on deaths with data on infections. Typically, these are disjoint samples, in that numbers of infections are estimated (typically very imprecisely) on select subpopulation, while deaths are recorded in the general population (at a level of country, administrative region etc.). There are clear reasons to believe that IFRs will differ across studies (e.g. due to age, comorbidity status, time, genetic factors, quality of healthcare etc.). To address this, we will use a Bayesian hierarchical modelling framework to assume that the setting-specific estimates of IFR_k can differ from each other but are linked through some common parameters. (By k's we denote different populations; note that sometimes we may have multiple IFR's from different age groups in the same location.)

The most straight-forward and "canonical" way to implement such a Bayesian model is by modelling log odds of the event.[A3] Deeks (2002) present a general treatment of such approach in medical statistics. Note, that for very rare

[A1] The role of time may be due to new treatments, improvements over time in our ability to treat COVID-19 or selection pressures which may lead to more benign versions of the virus. Country-specific or location-specific factors in IFR data may be driven by under-reporting, health care factors (including access to health care services) or underlying distributions of known risk factors. Additionally, some unknown risk factors (e.g. genetic) may also be operating, in which case controlling for age and co-morbidities will be not sufficient to account for cross-location differences.

[A2] Various estimates published since suggested that the relationship of mortality risk to age is consistent across different countries.

[A3] It is also possible to work with IFR_k parameters and treat them as derived from Beta distribution with some "hyperparameters" α and β of

events the odds of mortality are very similar to probability of mortality, but we model events on odds scale as a good "generic" approach to modelling binary data (in this case death following infections).[A4]

Basic models for this type of analysis of binary data can be implemented using existing statistical analysis packages; see, for example, *metafor* package in R or *baggr* by Więcek and Meager (2020). Such analysis would treat IFR as a logit-normal parameter to meta-analyse. However, note that when no deaths are observed, analysis of IFR (equal to observed deaths divided by modelled infections) is problematic. Therefore we propose a "custom" model that built in Stan which treats deaths and *prevalences* (rather than the IFRs) as data.

Let d_k denote observed deaths for data point k and assume that logit of corresponding prevalence estimate is p_k is a parameter (typically obtained from a statistical modelling papers, government reports etc.). Total population in k-th setting is n_k. Total number of estimates is K. Then the model likelihood is as follows:

$$d_k \sim \text{Binomial}(n_k, p_k IFR_k) \tag{5}$$

$$\text{logit}(p_k) \sim \mathcal{N}(\mu_k^{(p)}, (\sigma_k^{(p)})^2) \tag{6}$$

where $\sigma_k^{(p)}$ and $\mu_k^{(p)}$ are parameters obtained from the literature (or converted from these parameters – see next section). The k data points collected can span many locations (studies); we denote them by loc_k and the total number of locations by K_{loc} (with $K_{loc} < K$).

In this model we can also account for various covariates impacting the IFRs (let's denote their total number by N_p), such as age groups (which we identify with median age of the population being studied, MedianAge_k). We code them in a design matrix X. To center our X at the value of interest in our model (risk in 20-30 year olds), we use a transformation `MedianAge/10 - 2.5` to construct our matrix X. We denote all of the covariates using a design matrix X and denote by N_p the number of columns in X. We assume the impact on IFR is on logit scale, same as in the "canonical" logistic models of binary data that we mentioned above:

$$\text{logit}(IFR_k) = \theta_{\text{loc}_k} + X\beta, \tag{7}$$

$$\theta_j \sim \mathcal{N}(\tau, \sigma^2), \text{where } j = 1, \ldots, K_{loc}. \tag{8}$$

This means θ spans location-specific (random) effects on IFR while β is N_p dimensional vector of (fixed) covariate effects.

We implement our model in Stan and assume very weakly informative priors on all parameters, with prior for τ centered at 1 death per 10,000 cases.

```
model {
  //Uncertain prevalence estimates (mu^p and sigma^p above):
  logit_prevalence ~ normal(mean_prevalence, sd_prevalence);
  //Likelihood of mortality (d_k = obs_deaths, p_k = prevalence):
```

Beta distribution, as done by e.g. Carpenter (2016). That approach, however, does not offer an easy way of modelling impact of covariates (e.g. age and co-morbidities) on the rates.

[A4] Another advantage of such a model is that it can use either individual-level or summary data and work with covariates (such as gender, age, time of the study, co-morbidities), captured as odds ratios or risk ratios. If only summary data are available, covariates can be defined as study level distributions (e.g. % male)

```
  obs_deaths ~ binomial(population, prevalence .* ifr);
  //Hierarchical component of the model (location-specific theta):
  //logit_ifr = theta_k[loc] + to_vector(X*beta);
  theta_k ~ normal(tau, sigma);

  //Priors:
  tau   ~ normal(logit(.0001), 5);
  sigma ~ normal(0, 10);
  beta  ~ normal(0, 10);
}
```

A-2.2 | Data

We used estimates originally collected by Levin, Cochran, and Walsh (2020) to construct the first version of analysis dataset, which we then supplemented with more values extracted from other studies.

We included all relevant data points from the Levin et al. study This meant sometimes including values that the original study omitted. For example, Levin et al. exclude data points with seroprevalence indistinguishable from zero; our model retains them. We systematically went through the included studies listed in Appendix I and Appendices H.4 and H.5 of Levin et al. to ensure complete inclusion of all their data. In addition, in the course of this process we because aware of either updates to the studies Levin et al. used or entirely new relevant studies. We have reviewed and included them in our model as well.

The input data into our model consists of deaths (treated as known) and prevalences (treated as logit-distributed parameter with known mean and SD) in all reported age groups in all studies[A5].

All of input data are given in Table 4. The analysis dataset contains 167 data points from 34 studies, each containing between 2 and 11 different age groups. We made only minimal modifications to source data, by 1) imputing the values from the Italian fatality data based on a seroprevalence survey, 2) imputing population size in Maranhao (as ratio of the reported number of infections and the mean infection risk) which were not reported and 3) assuming that uncertainty in prevalence 0-29 age group in Iceland is same as in the 30-39 age group since data were missing.

As mentioned, our model treats number of COVID-attributable deaths as measured without error (due to lack of data) but accounts for uncertainty in infection risks, which are always model-based estimates extracted from various available data sources. In preparing data, we assumed that logits of prevalence estimates from available studies are normally distributed, which seems to reproduce majority of data very well, see Figure 7 in the Supplement. There are some discrepancies with studies that allowed for prevalence estimates to be 0, something that our logit model does not allow.

For each study we construct a median age scalar defined by the average of the endpoints of each age range, rather than attempting to calculate a population-weighted mean.

Our approach of regressing on the median age and use of all available data (rather than the subset of data available in younger adults only) is necessitated by data limitations: out of 167 data points comprising age-specific estimates of prevalence (or IFR) and counts of deaths, 37 contain individuals aged 20-30 who are of primary interest to us. However, the populations are mixed with regards to age, with typical age groupings such as 19-49, 20-49, 20-39, 0-49 used instead. In fact, we find only one estimate out of 37 that is entirely specific to the 20-29 age group (Brazilian state of Maranhao), while one more has median age falling between 20 and 30 but is not specific to that age group.

[A5] This basic approach potentially exaggerates uncertainty, as we treat different 95% intervals reported in the study as uncorrelated.

A-2.3 | Results

There were no issues with convergence of the Bayesian model. We set number of iterations to 5,000 and used 4 chains, with `max_treedepth` option set to 15. There were no divergent transitions and effective sample size was greater than 2310 for all of 538 modeled parameters (this number includes fitted prevalences, IFRs, θ's and their transformations into/from logit scales). For the three main parameters in the model we obtained the following:

```
## Inference for Stan model: ifr_with0.
## 4 chains, each with iter=5000; warmup=2500; thin=1;
## post-warmup draws per chain=2500, total post-warmup draws=10000.
##
##          mean se_mean   sd  2.5%   25%   50%   75%   98% n_eff Rhat
## tau     -8.80       0 0.12 -9.05 -8.88 -8.80 -8.72 -8.56 10954    1
## sigma    0.66       0 0.09  0.51  0.59  0.65  0.72  0.87 11075    1
## beta[1]  1.12       0 0.01  1.10  1.11  1.12  1.12  1.13  2556    1
##
## Samples were drawn using NUTS(diag_e) at Wed Feb 17 17:43:34 2021.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

The mean coefficient of beta 1.12 corresponds to 3.06-fold increase in mortality risk following an infection per each extra decade of age (95% uncertainty interval is 3.01-3.11).

From these parameters we can predict average risks for subjects of any given age x, by using the posterior distribution of $\tau + (10x + 2.5)\beta$ (where 2.5 and 10 refer to the transformation that we applied to `MedianAge` inputs).

A-2.4 | Average infection fatality risk in young subjects

Since we centered our `MedianAge` at 25 years in constructing our matrix X, we can now obtain model-estimated risk for a typical HCT population (aged 20 to 30, with median 25) by ignoring the β coefficient and examining τ and σ only. We find that the average IFR for this group (equal to $\frac{\exp(\tau)}{\exp(1+\tau)}$) is 1.51×10^{-4} (with 95% interval from 1.18×10^{-4} to 1.92×10^{-4}).

A-2.4.1 | Heterogeneity in IFRs

However, there is a considerable variability in IFRs across different locations/dataset that we should consider. To take into account parameter σ, we can generate draws from the $\mathcal{N}(\tau, \sigma^2)$ distribution, corresponding to a hypothetical IFR in a new source of data. 95% interval for such model runs from 3.94×10^{-5} to 5.79×10^{-4}. Since the model works a logistic scale, another way of interpreting the across-dataset variability is reporting the fold-impact of σ on the mean IFR; here, we obtain on average a 3.82-fold increase (decrease) in IFR per 2σ increase (decrease).

The lower end of the 95% interval, 3.94×10^{-5}, is not extreme given input data, where the "crude" mean IFR (based on mean prevalence only) is below 7 per 10,000 for all data except for South Florida, and as low as 0 for some countries that did not record deaths (Belgium, New Zealand, Korea, Iceland) in various age groups including 20-29 year olds or 1.4 per 10,000 in Utah, in the population aged 19-44. (Please refer to Table 4 for complete list of inputs.)

TABLE 1 Main parameters in the sensitivity analysis models using subsets of data.

Included data	beta	sigma	tau	IFR, 20-29 year olds (per 1,000)
All data (main model)	-8.80 (0.12)	0.65 (0.09)	0.15 (0.02)	1.12 (0.01)
Median age between 10 and 75	-8.72 (0.14)	0.72 (0.11)	0.17 (0.02)	1.07 (0.01)
Median age over 10	-8.79 (0.12)	0.66 (0.09)	0.15 (0.02)	1.11 (0.01)
Median age under 75	-8.75 (0.13)	0.70 (0.11)	0.16 (0.02)	1.08 (0.01)

We can assess this heterogeneity by inspecting the distribution of random effects in the model transformed into IFRs, i.e. the inverse logit transformation θ parameters. The largest (posterior mean) IFR value of θ is 5.09×10^{-4} in Castiglione d'Adda. The smallest posterior mean for 20-29 year olds is 4.67×10^{-5} in France.

A-2.4.2 | Predictive checks for the model

We constructed posterior predictive distributions for number of deaths in each of the inputs by using the `generated quantities` functionality of Stan. Figure 3 compares the posterior means and 95% intervals with observed deaths. Out of 167 observations that were used to fit the model, 157 were within 95% intervals of the posterior predictive distributions. We observed the largest discrepancies occurred in Spanish data. Overall, we conclude that the simple binomial model we used here is flexible enough to capture both age-specific risk increases and heterogeneity in IFRs across settings/countries.

A-2.5 | Sensitivity analyses

As a sensitivity analysis, we also considered the impact on the main model parameters of dropping some data from our analysis. The result is summarised in a table containing parameters σ, β, τ and the mean IFR for the 20-29 year olds age group.

We considered excluding data for the youngest and the oldest individuals, as well as excluding both at once. Our hypothesis was that at extreme ends of age the assumption of log-linearity of IFRs may not hold and potentially lead to a biased estimate of the IFR in the 20-29 and 20-39 age groups. However, as shown in the accompanying table, we find no substantial effect of excluding data on the main model paramters.

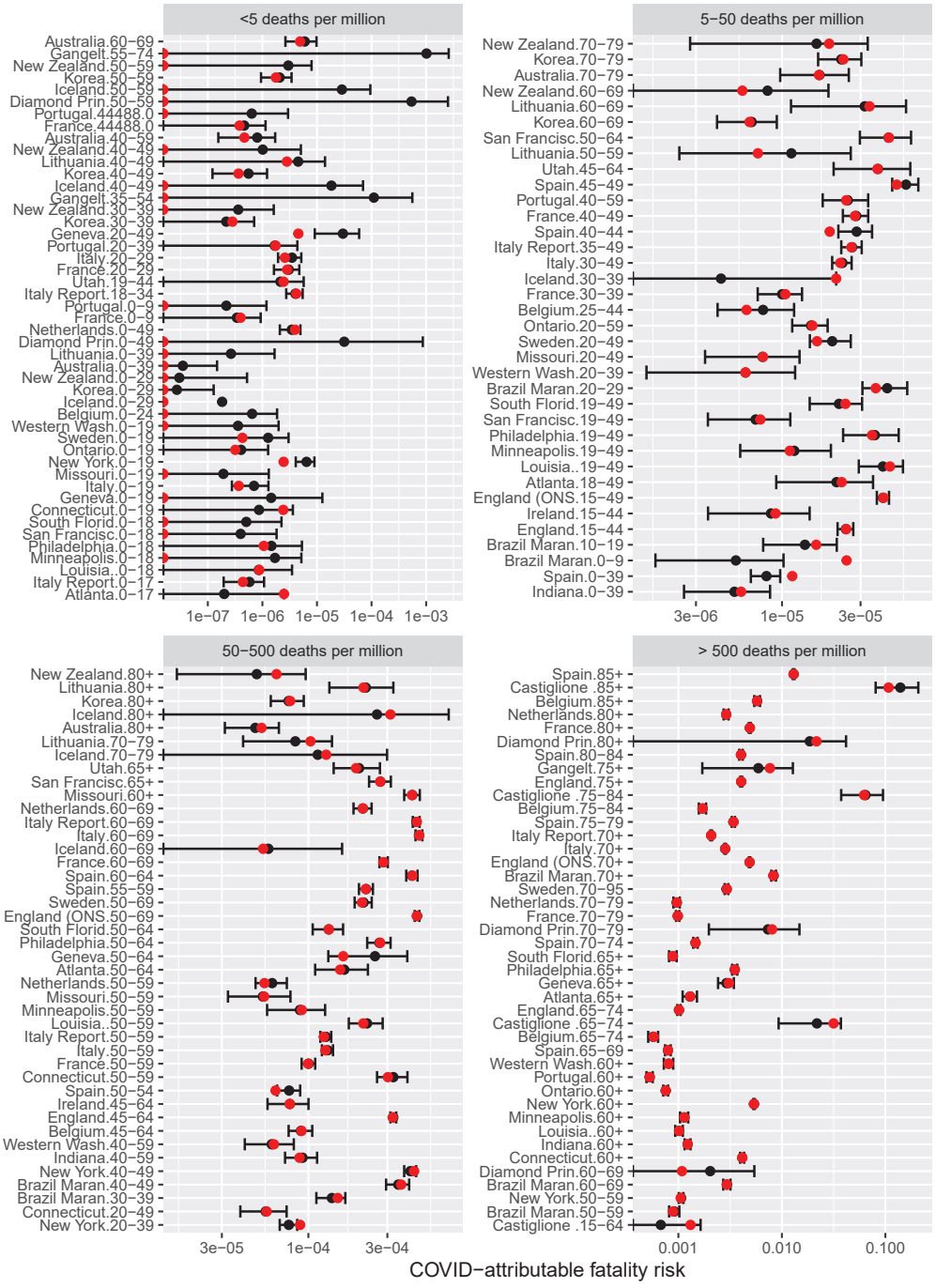

FIGURE 3 Comparison of model estimates (black) with data on observed fatality risk (FR, red), compared on logarithmic scale. FR is number of deaths divided by overall population size. Bars are 95% posterior interval; point is the mean. For better clarity, we grouped the plot into four panels according to observed FR X axes on each panel differ. For many low-risk populations (upper-left quadrant) no deaths were reported: we indicate this by plotting a red point on the left-hand side of the panel plot.

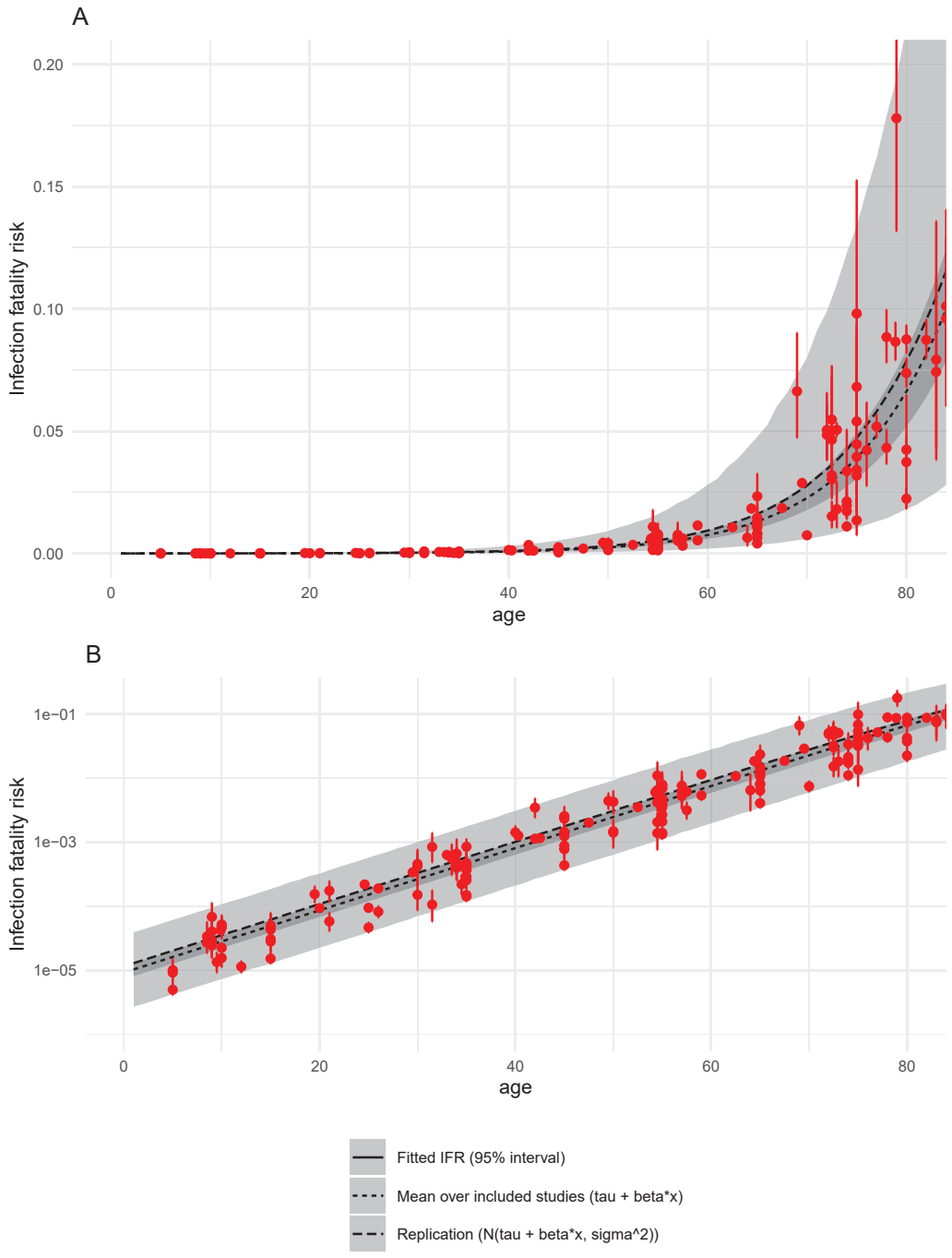

FIGURE 4 IFR as a function of age. Narrower ribbon corresponds to the 95% posterior interval of average across all included studies (tau parameter in the meta-analysis model), while the wider band takes into account heterogeneity (tau and sigma). Lines are means. Red points are model estimates of mean IFRs in partciular studies, with bars representing 95% posterior intervals. Panel A is untransformed data. Panel B shows the same data on log 10 scale.

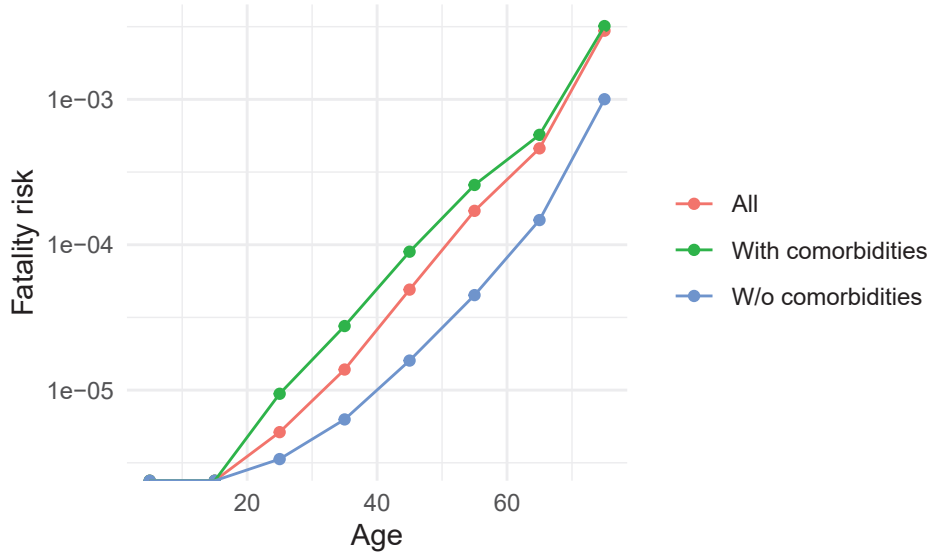

FIGURE 5 Fatality risks as a function of age in OpenSAFELY data. Fatality risk is zero in the 0-10 and 10-20 age groups.

A-3 | RISK REDUCTION IN HEALTHY INDIVIDUALS

We now turn our attention to the question of how much a human challenge trial designer could reduce the mortality risk by using simple screening methods, as discussed in the main paper.

Data for this section has been provided by OpenSAFELY (`https://opensafely.org/`) and was used by Williamson et al. (2020) to characterise COVID-19 mortality risk factors for 10,926 COVID-19 deaths in England. We group the total of 21,444,863 individuals into a total population and a lower-risk sub-population, defined as non-smoker, non-obese and without the comoribidities reported in the OpenSAFELY study[A6], most notably respiratory and cardiovascular diseases and type I diabetes. For brevity we refer to the population without one of the pre-defined comorbidities as "healthy". In contrast to the cited publication, we include records of individuals under 18 in our assessment. Counts grouped by age are presented in Table 2. Complete data (broken down by gender) are in Table 2 at the end of the document.

As shown in Table 2, for the age group of 20-29 the crude risk ratio (of general population vs the healthy subset only) is 1.53, but, due to low number of events in both healthy and general population, with a very wide 95% interval from 0.6 to 5.65.[A7] As data on relative risks in other age groups is clearly related to the relative risk in 20-29 age group, we use another meta-analysis model to improve our estimate. Additionally, relative risks are higher in women than in men – something that we can account for in our model too.

In our modelling we make a strong assumption that infection risks in population with comorbidities are the same

[A6] "asthma, other chronic respiratory disease, chronic heart disease, diabetes mellitus, chronic liver disease, chronic neurological diseases, common autoimmune diseases (Rheumatoid Arthritis (RA), Systemic Lupus Erythematosus (SLE) or psoriasis), solid organ transplant, asplenia, other immunosuppressive conditions, cancer, evidence of reduced kidney function, and raised blood pressure or a diagnosis of hypertension"

[A7] We obtain the interval using a simulation approach. Using normal approximation of log(RR) statistic we obtain a narrower 0.57 to 4.1, perhaps due to poor quality of approximation for rare events.

TABLE 2 Data from the OpenSAFELY database grouped by age.

	Population		Fatalities		FR per 100k		
Age group	All	Healthy	All	Healthy	All	Healthy	RR
0 to 9	2,160,958	2,007,997	0	0	0.00	0.00	NaN
10 to 19	2,428,494	2,000,761	0	0	0.00	0.00	NaN
20 to 29	2,530,792	1,788,907	13	6	0.51	0.34	1.5
30 to 39	2,960,611	1,909,227	41	12	1.38	0.63	2.2
40 to 49	2,849,984	1,565,935	140	25	4.91	1.60	3.1
50 to 59	3,051,110	1,243,728	522	56	17.11	4.50	3.8
60 to 69	2,392,392	622,357	1,101	92	46.02	14.78	3.1
70 to Inf	3,070,522	317,238	9,109	318	296.66	100.24	3.0

as in the general population. In other words, we assume that denominator for IFR is same in both populations. We then specify a generic partial pooling model of fatality risks (FR, defined as number of deaths in the entire population, without regards to infection status) such that

$$\text{logit}(\text{FR}_k) \sim \mathcal{N}(\alpha_{\text{age}_k} \text{comorb}_k + \beta \text{male}_k + \gamma_{\text{age}_k}), \tag{9}$$

$$\alpha_i \sim \mathcal{N}(\mu, \sigma) \text{ for all } i, \tag{10}$$

where, for k-th observation, age_k is the age group, and comorb_k and male_k are indicator variables. This means that each age group is assigned different "baseline" fatality risks.

Note that the assumption of FRs varying across age groups that we just mentioned differs from the model of age-specific IFRs in Section 2. This is because we hypothesised that infection risks (which are used as denominators in the IFR model of Section 2) will vary across age groups. This is borne out by Figure 5.

Summary of the main model parameters is as follows:

```
##             mean se_mean    sd   2.5%    25%    50%    75%    98% n_eff Rhat
## mu          1.43 0.00278 0.190   1.05   1.33   1.43   1.53   1.81  4674    1
## sigma       0.36 0.00462 0.212   0.12   0.22   0.31   0.44   0.90  2115    1
## alpha[3]    1.33 0.00440 0.329   0.60   1.14   1.34   1.53   1.95  5577    1
## beta        0.40 0.00019 0.019   0.36   0.39   0.40   0.42   0.44 10314    1
## gamma[3] -13.05 0.00415 0.342 -13.76 -13.27 -13.04 -12.82 -12.43  6795    1
```

We find that the mean risk ratio between population with comorbidities and healthy sub-population in 20-29 age group (exponent of α_3 above) is 3.97, with wide 95% uncertainty interval of 1.81 to 7.05.[A8]

[A8] Using simple models that assumed identical risk ratios in all age groups would lead to a mean RR of similar magnitude but a much less uncertain estimate, due to more rigid model assumptions; similarly, assuming some linear age structure on risks, such as in the main meta-analysis model above, would may lead to a different RR, but we do not think such an assumption is justified here. We do not include outputs of these models in this short write-up.

Next, using the posterior samples we calculate the event rate in total population (i.e. $\theta* = (\theta_1 n_1 + \theta_2 n_2)/(n_1 + n_2)$, where subscripts 1 and 2 are healthy and comorbid sub-populations) and then divide it by ratio in healthy population to obtain an estimate of risk reduction possible by selecting healthy volunteers only. The mean posterior value is 1.88, with 95% uncertainty interval from 1.24 to 2.79. Due to use of Bayesian hierarchical model over many age groups, the uncertainty interval is much narrower than on the risk reduction factor calculated on 20-29 age group only.

To validate the model, we conducted a simple posterior predictive check for numbers of deaths in different age groups and genders. The graphical check is presented in Figure 6. Overall we find that the simple model has no problem with reproducing observed data.

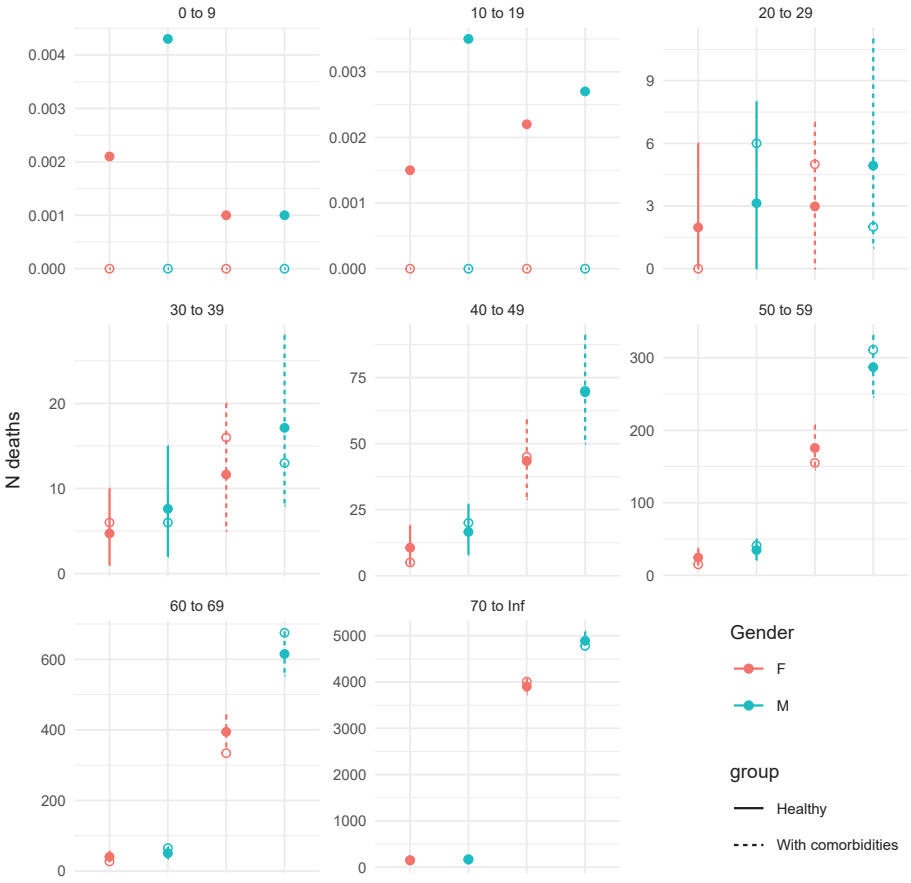

FIGURE 6 Comparison of posterior predictive numbers of deaths from the fitted Bayesian model (mean and 95% uncertainty intervals) with data inputs (circles). For each age grouping we have 4 estimates: male/female and healthy vs general population

A-4 | CONCLUSION AND SUMMARY OF RESULTS

In conclusion, the implications of the model for the risk in healthy young subjects are as follows:

- We find that average IFR in 20-29 age group for the studies included in this analysis is 1.51×10^{-4} with 95% interval from 1.18×10^{-4} to 1.92×10^{-4}.
  - It is feasible that the mean IFR can be decreased as much as 3.82-fold (2σ impact on the IFR according to hyper-SD parameter in the meta-analysis model).
  - It is easy to argue that a HCT designer would be able to achieve IFR at least as low as within any of the large-scale studies included in our sample of populations. The smallest posterior mean for 20-29 year olds is 4.67×10^{-5}, fitted to data from France.
  - Extending the HCT population to also include 30-39 year olds would lead to mean IFR of 2.65×10^{-4} with 95% interval from 2.06×10^{-4} to 3.35×10^{-4}. Lowest mean IFR would then be 8.2×10^{-5} (also in France).
- In the general population the risk rises by the factor or 3.06 per each decade of age, with 95% interval from 3.01 to 3.11.
- In healthy population (defined as lack of co-morbidities listed above), the average mortality risk in 20-29 year olds is 1.88 times lower than in the general population, with 95% uncertainty interval from 1.24 to 2.79.
  - Our 1.88 estimate is a bit higher than the mean "crude" risk ratio of 1.53 because we use a Bayesian hierarchical model that synthesises evidence across all age groups.
  - Expanding to 20-39 year olds, the risk in healthy sub-population would be 2.04 times lower than in the general population (95% interval from 1.24 to 2.79).
- Combining the smallest posterior IFR for 20-29 year olds with our estimated fold-reduction due to excluding individuals with co-morbidities from the population would lead to a mean infection fatality risk of 2.6×10^{-5} with 95% Bayesian interval from 1.62×10^{-5} to 3.89×10^{-5}.
  - In 20-39 year old subjects the risk would be 4.09×10^{-5} (2.88×10^{-5} to 5.55×10^{-5}).

A-5 | SUMMARY OF INCLUDED STUDIES AND COMPLETE INPUT DATA

We present two tables, one listing all study-level information and another breaking down information by age group. References to all included studies are given in a separate bibliography at the end of this appendix. Below the table, we provide a list that can be used to cross-reference study locations with their bibliographic references.

TABLE 3 Complete table of studies used by the meta-analysis model.

Study	End date	Age groups	Deaths	N	IR per 100
Atlanta	2020-05-03	0-17, 18-49, 50-64, 65+	366	1.8e+06	2.54
Australia	2020-06-12	0-39, 40-59, 60-69, 70-79, 80+	102	2.5e+07	0.06
Belgium	2020-05-16	0-24, 25-44, 45-64, 65-74, 75-84, 85+	4021	1.1e+07	6.10
Brazil Maranhao and Sao Luis	NA	0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70+	5026	7.0e+06	42.13
Castiglione d'Adda	2020-06-07	15-64, 65-74, 75-84, 85+	62	4.1e+03	23.21
Connecticut	2020-05-03	0-19, 20-49, 50-59, 60+	3867	3.6e+06	4.69
Diamond Princess	2020-02-20	0-49, 50-59, 60-69, 70-79, 80+	14	3.7e+03	16.72
England	2020-07-13	15-44, 45-64, 65-74, 75+	30174	4.6e+07	5.99
England (ONS)	2020-07-26	15-49, 50-69, 70+	50987	5.0e+07	5.30
France	2020-07-07	0-9, 44488, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+	28802	6.5e+07	6.73
Gangelt	2020-04-06	35-54, 55-74, 75+	9	7.8e+03	14.88
Geneva	2020-06-02	0-19, 20-49, 50-64, 65+	274	4.8e+05	10.83
Iceland	2020-06-14	0-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+	10	3.4e+05	0.67
Indiana	2020-04-29	0-39, 40-59, 60+	2032	6.7e+06	2.76
Ireland	2020-07-16	15-44, 45-64	112	3.2e+06	1.39
Italy	2020-07-27	0-19, 20-29, 30-49, 50-59, 60-69, 70+	34142	6.0e+07	2.08
Italy Report	2020-08-01	0-17, 18-34, 35-49, 50-59, 60-69, 70+	26123	6.0e+07	2.47
Korea	2020-07-11	0-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+	289	5.1e+07	0.05
Lithuania	2020-06-18	0-39, 40-49, 50-59, 60-69, 70-79, 80+	76	2.7e+06	0.15
Louisia.	2020-04-08	0-18, 19-49, 50-59, 60+	1265	4.6e+06	5.71
Minneapolis	2020-05-12	0-18, 19-49, 50-59, 60+	993	3.9e+06	2.68
Missouri	2020-04-26	0-19, 20-49, 50-59, 60+	681	6.1e+06	2.67

TABLE 3 Complete table of studies used by the meta-analysis model. *(continued)*

Study	End date	Age groups	Deaths	N	IR per 100
Netherlands	2020-04-17	0-49, 50-59, 60-69, 70-79, 80+	4596	1.7e+07	3.54
New York	2020-04-28	0-19, 20-39	494	1.0e+07	14.60
New York	2020-05-23	40-49, 50-59, 60+	28166	9.5e+06	13.97
New Zealand	2020-07-09	0-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+	22	4.8e+06	0.07
Ontario	2020-06-30	0-19, 20-59, 60+	2723	1.5e+07	1.10
Philadelphia	2020-04-25	0-18, 19-49, 50-64, 65+	2716	4.1e+06	3.03
Portugal	2020-07-08	0-9, 10-19, 20-39, 40-59, 60+	1660	1.0e+07	2.19
San Francisco Bay	2020-04-27	0-18, 19-49, 50-64, 65+	424	7.7e+06	1.12
South Florida	2020-04-10	0-18, 19-49, 50-64, 65+	1290	6.3e+06	1.84
Spain	2020-07-15	0-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+	40486	4.7e+07	3.79
Sweden	2020-06-18	0-19, 20-49, 50-69, 70-95	5053	1.0e+07	5.41
Utah	2020-05-03	19-44, 45-64, 65+	98	2.2e+06	2.26
Western Washington	2020-04-01	0-19, 20-39, 40-59, 60+	777	4.3e+06	1.14

Atlanta: Biggs et al. (2020); Australia: A. D. of Health (2020); Belgium: Herzog et al. (2020); Belgium: Molenberghs et al. (2020); Brazil Maranhao and Sao Luis: Silva et al. (2020); Brazil Regional: Hallal et al. (2020); Cache County, UT: Project (2020); Castiglione d'Adda: Pagani et al. (2020); Connecticut: Havers et al. (2020); Connecticut: Mahajan et al. (2020); Diamond Princess: Mizumoto et al. (2020); England: Ward et al. (2020); England (ONS): England (2020); France: Carrat et al. (2020); France: F. P. Health (2020); Gangelt: Streeck et al. (2020); Geneva: Perez-Saez et al. (2020); Iceland: I. D. of Health (2020); Indiana: Menachemi et al. (2020); Ireland: HPSC (2020); Italy: Istat (2020); Italy Deaths: Group (2020); Italy Report: Istat (2020); Korea: Control and Agency (2020); Lithuania: Registry (2020); Lombardy: Paradisi and Rinaldi (2020); Louisia.: Havers et al. (2020); Minneapolis: Havers et al. (2020); Missouri: Havers et al. (2020); Netherlands: RIVM (2020); New York: Rosenberg et al. (2020); New York State Comorbidity: N. Y. S. D. of Health (2020); New Zealand: N. Z. M. of Health (2020); NYC JAMA: Richardson S (2020); Ontario: Health Protection and Health Ontario) (2020); Philadelphia: Havers et al. (2020); Portugal: Saúde (2020); San Francisco Bay: Havers et al. (2020); South Florida: Havers et al. (2020); Spain: Pastor-Barriuso et al. (2020); Sweden: Authority (2020); Utah: Havers et al. (2020); Utah: Project (2020); Verity et al.: Verity, Okell, et al. (2020b); Washington County, UT: Project (2020); Weber County, UT: Project (2020); West Salt Lake, UT: Project (2020); Western Washington: Havers et al. (2020)

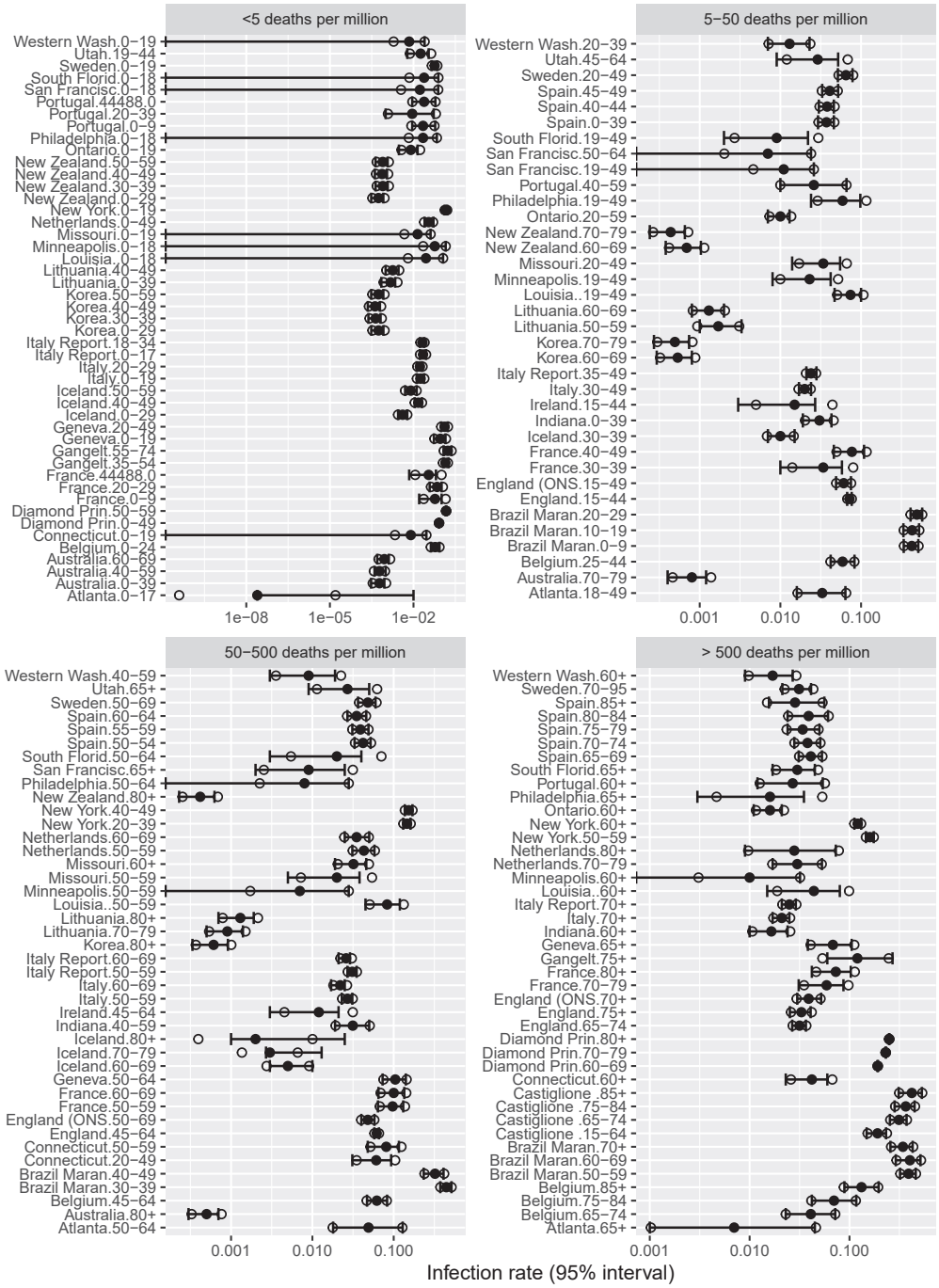

FIGURE 7 Comparison of model-estimated prevalences (95% CI's reported by modelling studies) collected by Levin, Cochran, and Walsh (2020) and our distributional assumptions: additional circles show 95% CIs recreated by assuming logit-normal distribution of prevalence. We group studies into 4 bands of mortality to mirror earlier figures. Please note that this approach produces discrepancies in a number of US estimates where the confidence intervals were skewed toward including 0. However, since we do not have access to source data, we decided to use the logit-normal assumption for all estimates. This assumption may have an effect of overestimating mortality risk in settings where prevalence was very low.

TABLE 4 Complete table of inputs used by the meta-analysis model and crude IFR's.

				Infection rate (%)			
Study location	Age	Deaths	N	Mean	2.5%	97.5%	IFR/100k
Brazil Maranhao and Sao Luis	0-9	29	1.2e+06	42.60	51.30	33.80	6.0e-02
France	0-9	3	7.5e+06	5.90	10.20	1.60	1.0e-02
Portugal	0-9	0	8.4e+05	2.20	6.00	0.90	0.0e+00
Atlanta	0-17	1	4.0e+05	0.00	1.00	0.00	1.0e+05
Italy Report	0-17	5	9.6e+06	2.20	2.80	1.70	2.0e-02
Louisia.	0-18	1	1.1e+06	2.80	11.50	0.00	3.0e-02
Minneapolis	0-18	0	9.7e+05	5.80	14.30	0.00	0.0e+00
Philadelphia	0-18	1	9.4e+05	2.20	6.90	0.00	5.0e-02
San Francisco Bay	0-18	0	1.6e+06	1.70	7.70	0.00	0.0e+00
South Florida	0-18	0	1.3e+06	2.40	7.80	0.00	0.0e+00
Missouri	0-19	0	1.5e+06	1.40	4.10	0.00	0.0e+00
Italy	0-19	4	1.1e+07	1.80	2.40	1.30	2.0e-02
Connecticut	0-19	2	8.3e+05	0.80	2.90	0.00	3.0e-01
Geneva	0-19	0	8.0e+04	9.17	14.06	5.40	0.0e+00
New York	0-19	12	4.9e+06	14.60	16.10	13.10	2.0e-02
Ontario	0-19	1	3.1e+06	0.80	1.40	0.30	4.0e-02
Sweden	0-19	1	2.3e+06	5.73	7.00	4.45	1.0e-02
Western Washington	0-19	0	1.0e+06	0.70	2.50	0.00	0.0e+00
Belgium	0-24	0	3.2e+06	6.00	8.60	4.20	0.0e+00
Brazil Maranhao and Sao Luis	10-19	21	1.3e+06	43.00	52.40	33.50	4.0e-02
France	44488	3	7.9e+06	3.50	6.40	0.70	1.0e-02
Iceland	0-29	0	1.4e+05	0.41	0.60	0.28	0.0e+00
Korea	0-29	0	1.6e+07	0.06	0.08	0.03	0.0e+00
New Zealand	0-29	0	1.9e+06	0.06	0.08	0.03	0.0e+00
Portugal	10-19	0	1.0e+06	2.40	6.10	0.90	0.0e+00
Indiana	0-39	20	3.5e+06	3.05	4.30	1.90	1.8e-01
Spain	0-39	225	1.9e+07	3.73	4.60	2.93	3.1e-01
Australia	0-39	0	1.4e+07	0.06	0.09	0.03	0.0e+00
Lithuania	0-39	0	1.2e+06	0.15	0.22	0.07	0.0e+00
Italy	20-29	16	6.2e+06	1.70	2.00	1.30	1.5e-01
Brazil Maranhao and Sao Luis	20-29	47	1.3e+06	49.20	57.30	41.10	8.0e-02
France	20-29	21	7.4e+06	7.00	10.20	3.80	4.0e-02
Italy Report	18-34	43	1.1e+07	2.10	2.40	1.70	1.9e-01
Netherlands	0-49	40	1.0e+07	3.50	5.20	2.50	1.1e-01
England	15-44	524	2.1e+07	7.20	7.66	6.73	3.4e-01
Ireland	15-44	18	2.0e+06	1.50	2.70	0.30	6.1e-01
New York	20-39	482	5.4e+06	14.60	16.10	13.10	6.1e-01

TABLE 4 Complete table of inputs used by the meta-analysis model and crude IFR's. *(continued)*

Study location	Age	Deaths	N	Mean	2.5%	97.5%	IFR/100k
Portugal	20-39	4	2.3e+06	0.90	5.30	0.10	1.9e-01
Western Washington	20-39	8	1.3e+06	1.30	2.30	0.70	4.6e-01
Philadelphia	19-49	51	1.4e+06	5.90	9.80	2.40	6.0e-01
Utah	19-44	3	1.2e+06	1.80	3.50	0.60	1.4e-01
England (ONS)	15-49	1035	2.5e+07	6.10	7.50	4.90	6.8e-01
Atlanta	18-49	20	8.7e+05	3.30	6.40	1.60	7.0e-01
Louisia.	19-49	85	1.9e+06	7.40	10.00	4.70	6.1e-01
Minneapolis	19-49	18	1.6e+06	2.30	4.20	0.80	4.8e-01
San Francisco Bay	19-49	25	3.4e+06	1.10	2.60	0.00	6.7e-01
South Florida	19-49	61	2.5e+06	0.90	2.20	0.20	2.7e+00
Missouri	20-49	18	2.3e+06	3.40	5.50	1.40	2.3e-01
Belgium	25-44	18	3.0e+06	5.90	8.30	4.20	1.0e-01
Brazil Maranhao and Sao Luis	30-39	163	1.1e+06	44.40	51.40	37.40	3.4e-01
Connecticut	20-49	75	1.3e+06	6.10	9.30	3.10	9.2e-01
Diamond Princess	0-49	0	1.2e+03	8.26	8.28	8.24	0.0e+00
France	30-39	84	8.0e+06	3.40	5.80	1.00	3.1e-01
Geneva	20-49	1	2.2e+05	13.12	17.00	9.75	3.0e-02
Iceland	30-39	1	4.7e+04	1.00	1.50	0.70	2.1e+00
Korea	30-39	2	7.1e+06	0.04	0.07	0.02	6.4e-01
New Zealand	30-39	0	6.2e+05	0.08	0.12	0.04	0.0e+00
Sweden	20-49	63	3.9e+06	6.50	7.84	5.16	2.5e-01
Ontario	20-59	122	8.0e+06	1.00	1.30	0.70	1.5e+00
Italy	30-49	369	1.6e+07	2.00	2.40	1.70	1.1e+00
Castiglione d'Adda	15-64	4	3.1e+03	19.10	23.24	14.86	6.9e+00
Italy Report	35-49	334	1.3e+07	2.40	2.80	2.10	1.1e+00
Spain	40-44	78	4.0e+06	3.80	4.60	3.00	5.1e-01
Brazil Maranhao and Sao Luis	40-49	290	8.0e+05	32.20	41.00	23.40	1.1e+00
France	40-49	231	8.3e+06	7.70	10.90	4.60	3.6e-01
Gangelt	35-54	0	3.6e+03	14.00	18.00	12.00	0.0e+00
Iceland	40-49	0	4.3e+04	1.50	2.00	1.10	0.0e+00
Korea	40-49	3	8.2e+06	0.04	0.06	0.02	8.8e-01
Lithuania	40-49	1	3.6e+05	0.18	0.29	0.10	1.6e+00
New York	40-49	1026	2.4e+06	15.30	17.00	13.70	2.8e+00
New Zealand	40-49	0	5.9e+05	0.07	0.11	0.04	0.0e+00
Spain	45-49	196	3.9e+06	4.10	5.20	3.30	1.2e+00
Indiana	40-59	148	1.7e+06	3.14	5.00	1.90	2.8e+00
Australia	40-59	3	6.4e+06	0.06	0.10	0.04	7.8e-01
Portugal	40-59	75	3.1e+06	2.60	6.60	1.00	9.4e-01

TABLE 4 Complete table of inputs used by the meta-analysis model and crude IFR's. *(continued)*

Study location	Age	Deaths	N	Mean	2.5%	97.5%	IFR/100k
Western Washington	40-59	69	1.1e+06	0.90	1.90	0.30	6.9e+00
Spain	50-54	230	3.6e+06	4.20	5.20	3.30	1.5e+00
Italy	50-59	1186	9.4e+06	2.70	3.10	2.30	4.7e+00
England	45-64	4657	1.4e+07	6.18	6.58	5.78	5.2e+00
Louisia.	50-59	126	5.9e+05	8.30	11.90	4.50	2.6e+00
Minneapolis	50-59	47	5.2e+05	0.70	2.80	0.00	1.3e+01
Missouri	50-59	43	8.0e+05	2.00	3.80	0.50	2.7e+00
Philadelphia	50-64	290	1.1e+06	0.80	2.80	0.00	3.4e+01
Utah	45-64	24	6.3e+05	2.90	5.20	0.90	1.3e+00
Belgium	45-64	280	3.1e+06	6.20	8.30	4.70	1.5e+00
Brazil Maranhao and Sao Luis	50-59	533	6.0e+05	39.10	46.10	32.10	2.3e+00
Connecticut	50-59	157	5.2e+05	8.10	11.60	4.80	3.7e+00
Diamond Princess	50-59	0	4.0e+02	14.82	14.90	14.70	0.0e+00
France	50-59	860	8.6e+06	9.70	13.10	6.40	1.0e+00
Iceland	50-59	0	4.2e+04	0.80	1.30	0.50	0.0e+00
Ireland	45-64	94	1.2e+06	1.20	2.10	0.30	6.4e+00
Italy Report	50-59	1196	9.6e+06	3.10	3.50	2.70	4.0e+00
Korea	50-59	15	8.5e+06	0.06	0.08	0.03	3.2e+00
Lithuania	50-59	3	4.2e+05	0.17	0.33	0.10	4.2e+00
Netherlands	50-59	137	2.5e+06	4.30	5.80	3.10	1.3e+00
New York	50-59	2764	2.6e+06	16.00	17.50	14.60	6.6e+00
New Zealand	50-59	0	6.3e+05	0.08	0.12	0.04	0.0e+00
Atlanta	50-64	51	3.3e+05	4.90	12.90	1.80	3.2e+00
San Francisco Bay	50-64	66	1.5e+06	0.70	2.40	0.00	6.4e+00
South Florida	50-64	169	1.3e+06	2.00	4.00	0.30	6.6e+00
Geneva	50-64	16	9.9e+04	10.45	14.11	7.31	1.6e+00
Spain	55-59	758	3.4e+06	3.90	4.90	3.10	5.7e+00
England (ONS)	50-69	7118	1.6e+07	4.80	5.80	4.00	9.4e+00
Sweden	50-69	504	2.4e+06	4.81	5.98	3.64	4.4e+00
Spain	60-64	1249	2.9e+06	3.50	4.60	2.70	1.2e+01
Gangelt	55-74	0	3.1e+03	17.00	23.00	12.00	0.0e+00
Italy	60-69	3433	7.3e+06	2.20	2.50	1.70	2.1e+01
Australia	60-69	13	2.7e+06	0.09	0.13	0.05	5.4e+00
Brazil Maranhao and Sao Luis	60-69	1155	3.9e+05	40.30	51.40	29.10	7.3e+00
Diamond Princess	60-69	1	9.2e+02	19.18	19.30	19.10	5.6e+00
France	60-69	2204	7.8e+06	10.00	13.50	6.50	2.8e+00
Iceland	60-69	2	3.8e+04	0.50	1.00	0.30	1.1e+01
Italy Report	60-69	3274	7.5e+06	2.60	2.90	2.10	1.7e+01

TABLE 4 Complete table of inputs used by the meta-analysis model and crude IFR's. *(continued)*

Study location	Age	Deaths	N	Mean	2.5%	97.5%	IFR/100k
Korea	60-69	41	6.5e+06	0.05	0.08	0.03	1.2e+01
Lithuania	60-69	12	3.5e+05	0.13	0.20	0.08	2.6e+01
Netherlands	60-69	454	2.1e+06	3.50	5.00	2.50	6.1e+00
New Zealand	60-69	3	5.2e+05	0.07	0.10	0.04	8.3e+00
Spain	65-69	1905	2.4e+06	4.10	5.30	3.10	1.9e+01
Castiglione d'Adda	65-74	17	5.4e+02	31.30	37.30	25.40	1.0e+02
England	65-74	5663	5.6e+06	3.16	3.66	2.67	3.2e+01
Belgium	65-74	663	1.1e+06	4.10	7.20	2.30	1.4e+01
Connecticut	60+	3633	8.8e+05	4.20	6.00	2.30	9.9e+01
Ontario	60+	2600	3.4e+06	1.60	2.10	1.10	4.7e+01
Indiana	60+	1864	1.5e+06	1.65	2.40	1.00	7.5e+01
Louisia.	60+	1053	1.0e+06	4.40	8.00	1.50	2.3e+01
Minneapolis	60+	928	8.1e+05	1.00	3.20	0.00	1.1e+02
Missouri	60+	620	1.5e+06	3.20	4.60	1.90	1.3e+01
Spain	70-74	3230	2.2e+06	3.80	5.10	2.80	3.9e+01
Western Washington	60+	700	8.7e+05	1.70	2.70	0.90	4.7e+01
New York	60+	24376	4.5e+06	12.10	13.10	11.20	4.4e+01
Portugal	60+	1581	3.0e+06	2.70	5.40	1.20	2.0e+01
Australia	70-79	31	1.8e+06	0.08	0.12	0.04	2.1e+01
France	70-79	5650	5.7e+06	5.90	8.70	3.10	1.7e+01
Geneva	65+	257	8.4e+04	6.82	10.53	3.83	4.5e+01
Netherlands	70-79	1539	1.6e+06	3.00	5.30	1.70	3.2e+01
New Zealand	70-79	7	3.6e+05	0.04	0.07	0.02	4.4e+01
Atlanta	65+	294	2.3e+05	0.70	4.50	0.10	1.9e+02
Diamond Princess	70-79	8	1.0e+03	23.05	23.20	23.00	3.5e+01
Iceland	70-79	3	2.3e+04	0.30	1.30	0.27	4.3e+01
Korea	70-79	84	3.6e+06	0.05	0.07	0.03	4.8e+01
Lithuania	70-79	23	2.2e+05	0.09	0.14	0.05	1.1e+02
Philadelphia	65+	2374	6.8e+05	1.60	3.50	0.30	2.2e+02
San Francisco Bay	65+	333	1.2e+06	0.90	2.50	0.20	3.0e+01
Utah	65+	71	3.7e+05	2.70	5.00	0.90	7.2e+00
South Florida	65+	1060	1.2e+06	3.00	4.50	1.70	2.9e+01
Spain	75-79	6175	1.8e+06	3.40	5.00	2.40	1.0e+02
England (ONS)	70+	42834	8.8e+06	3.90	5.20	3.00	1.3e+02
Sweden	70-95	4485	1.5e+06	3.12	4.12	2.13	9.4e+01
Italy	70+	29134	1.0e+07	2.10	2.50	1.70	1.3e+02
Castiglione d'Adda	75-84	25	4.0e+02	36.60	44.90	28.30	1.7e+02
Belgium	75-84	1182	6.9e+05	7.00	11.70	4.20	2.4e+01

TABLE 4 Complete table of inputs used by the meta-analysis model and crude IFR's. *(continued)*

Study location	Age	Deaths	N	Mean	2.5%	97.5%	IFR/100k
Brazil Maranhao and Sao Luis	70+	2788	3.4e+05	34.30	42.90	25.70	2.4e+01
England	75+	19330	4.8e+06	3.30	4.08	2.53	1.2e+02
Gangelt	75+	9	1.2e+03	12.00	27.00	6.00	6.4e+01
Italy Report	70+	21271	1.1e+07	2.50	2.90	2.10	8.1e+01
Spain	80-84	5192	1.3e+06	3.90	6.10	2.40	1.0e+02
Diamond Princess	80+	5	2.2e+02	25.00	25.10	24.90	8.6e+01
Iceland	80+	4	1.3e+04	0.20	2.50	0.10	1.6e+02
Korea	80+	144	1.9e+06	0.06	0.09	0.03	1.3e+02
New Zealand	80+	12	1.9e+05	0.04	0.06	0.02	1.5e+02
Lithuania	80+	37	1.7e+05	0.13	0.19	0.07	1.7e+02
Netherlands	80+	2426	8.4e+05	2.80	7.30	0.90	1.0e+02
Australia	80+	55	1.1e+06	0.05	0.07	0.03	1.0e+02
France	80+	19746	4.0e+06	7.30	10.30	4.20	6.7e+01
Spain	85+	21248	1.6e+06	2.85	5.60	1.56	4.6e+02
Castiglione d'Adda	85+	16	1.5e+02	42.10	53.10	31.10	2.6e+02
Belgium	85+	1878	3.3e+05	13.20	19.60	8.90	4.4e+01

Additional references for the Appendix

Authority, Sweden Public Health. 2020. "COVID-19 Report for Week 24 - Veckorapport Om Covid-19, Vecka 24." Stockholm: Folkhälsomyndigheten.

Biggs, Holly M, Jennifer B Harris, Lucy Breakwell, F Scott Dahlgren, Glen R Abedi, Christine M Szablewski, Jan Drobeniuc, et al. 2020. "Estimated Community Seroprevalence of Sars-Cov-2 Antibodies—Two Georgia Counties, April 28–May 3, 2020." *Morbidity and Mortality Weekly Report* 69 (29): 965.

Carpenter, Bob. 2016. "Hierarchical Partial Pooling for Repeated Binary Trials." https://mc-stan.org/users/documentation/case-studies/pool-binary-trials.html.

Carrat, Fabrice, Xavier de Lamballerie, Delphine Rahib, Helene Blanche, Nathanael Lapidus, Fanny Artaud, Sofiane Kab, et al. 2020. "Seroprevalence of Sars-Cov-2 Among Adults in Three Regions of France Following the Lockdown and Associated Risk Factors: A Multicohort Study." *medRxiv*. https://doi.org/10.1101/2020.09.16.20195693.

Control, Korea Disease, and Prevention Agency. 2020. "Weekly Report on the Covid-19 Situation in the Republic of Korea." Cheongju: KDCA.

Deeks, Jonathan J. 2002. "Issues in the Selection of a Summary Statistic for Meta-Analysis of Clinical Trials with Binary Outcomes." *Statistics in Medicine* 21 (11): 1575–1600. https://doi.org/10.1002/sim.1188.

England, Public Health. 2020. "Coronavirus (Covid-19) Infection Survey: Characteristics of People Testing Positive for Covid-19 in England, August 2020." London: Public Health England.

Group, SARS-CoV-2 Surveillance. 2020. "Characteristics of Sars-Cov-2 Patients Dying in Italy Report Based on Available Data on October 4th, 2020." Italy: Istituto Superiore di Sanita.

Hallal, Pedro C, Fernando P Hartwig, Bernardo L Horta, Gabriel D Victora, Mariângela F Silveira, Claudio J Struchiner, Luis Paulo Vidaleti, et al. 2020. "Remarkable Variability in Sars-Cov-2 Antibodies Across Brazilian Regions: Nationwide Serological Household Survey in 27 States." *medRxiv*. https://doi.org/10.1101/2020.05.30.20117531.

Havers, Fiona P, Carrie Reed, Travis Lim, Joel M Montgomery, John D Klena, Aron J Hall, Alicia M Fry, et al. 2020. "Seroprevalence of Antibodies to Sars-Cov-2 in 10 Sites in the United States, March 23-May 12, 2020." *JAMA Internal Medicine* 180 (12): 1576–86.

Health, France Public. 2020. "COVID-19 : Point épidémiologique Du 9 Juillet 2020." Paris: Santé publique France.

Health, Australia Department of. 2020. "Coronavirus (Covid-19) Current Situation and Case Numbers." Canberra: Australia Department of Health.

Health, Iceland Directorate of. 2020. "COVID-19 in Iceland – Statistics 28 Feb - 14 June 2020." Reykjavik: Iceland Directorate of Health.

Health, New York State Department of. 2020. "NYSDOH Covid-19 Tracker - Fatalities." https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-Fatalities?%3Aembed=yes&%3Atoolbar=no.

Health, New Zealand Ministry of. 2020. "COVID-19: Current Cases." Wellington: New Zealand Ministry of Health.

Health Protection, Ontario Agency for, and Promotion (Public Health Ontario). 2020. "COVID-19 Serosurveillance Summary - Seroprevalence in Ontario: March 27, 2020 to June 30, 2020." Toronto, ON: Queen's Printer for Ontario.

Herzog, Sereina, Jessie De Bie, Steven Abrams, Ine Wouters, Esra Ekinci, Lisbeth Patteet, Astrid Coppens, et al. 2020. "Seroprevalence of Igg Antibodies Against Sars Coronavirus 2 in Belgium a Serial Prospective Cross-Sectional Nationwide Study of Residual Samples." *medRxiv*. https://doi.org/10.1101/2020.06.08.20125179.

HPSC. 2020. "Epidemiology of Covid-19 in Ireland: Report Prepared by Hpsc on 04/08/2020 for National Public Health Emergency Team." Dublin: Health Protection Surveillance Centre.

Istat. 2020. “Primi Risultati Dell’indagine Di Sieroprevalenza Sul Sars-Cov-2.” Rome: Instituto Nazionale di Statistica.

Levin, Andrew T., Kensington B. Cochran, and Seamus P. Walsh. 2020. “ASSESSING THE AGE SPECIFICITY OF INFECTION FATALITY RATES FOR COVID-19: META-ANALYSIS &Amp; PUBLIC POLICY IMPLICATIONS.” *medRxiv*, July, 2020.07.23.20160895. https://doi.org/10.1101/2020.07.23.20160895.

Mahajan, Shiwani, Rajesh Srinivasan, Carrie A Redlich, Sara K Huston, Kelly M Anastasio, Lisa Cashman, Dorothy S Massey, et al. 2020. “Seroprevalence of Sars-Cov-2-Specific Igg Antibodies Among Adults Living in Connecticut: Post-Infection Prevalence (Pip) Study.” *The American Journal of Medicine*.

Menachemi, Nir, Constantin T Yiannoutsos, Brian E Dixon, Thomas J Duszynski, William F Fadel, Kara K Wools-Kaloustian, Nadia Unruh Needleman, et al. 2020. “Population Point Prevalence of Sars-Cov-2 Infection Based on a Statewide Random Sample—Indiana, April 25–29, 2020.” *Morbidity and Mortality Weekly Report* 69 (29): 960.

Mizumoto, Kenji, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell. 2020. “Estimating the Asymptomatic Proportion of Coronavirus Disease 2019 (Covid-19) Cases on Board the Diamond Princess Cruise Ship, Yokohama, Japan, 2020.” *Eurosurveillance* 25 (10): 2000180.

Molenberghs, Geert, Christel Faes, Johan Verbeeck, Patrick Deboosere, Steven Abrams, Lander Willem, Jan Aerts, et al. 2020. “Belgian Covid-19 Mortality, Excess Deaths, Number of Deaths Per Million, and Infection Fatality Rates (9 March—28 June 2020).” *medRxiv*.

Pagani, Gabriele, Federico Conti, Andrea Giacomelli, Dario Bernacchia, Rossana Rondanin, Andrea Prina, Vittore Scolari, et al. 2020. “Seroprevalence of Sars-Cov-2 Significantly Varies with Age: Preliminary Results from a Mass Population Screening.” *The Journal of Infection* 81 (6): e10–e12.

Paradisi, Matteo, and Gianluca Rinaldi. 2020. “An Empirical Estimate of the Infection Fatality Rate of Covid-19 from the First Italian Outbreak.” *Available at SSRN 3582811*. https://doi.org/10.2139/ssrn.3582811.

Pastor-Barriuso, Roberto, Beatriz Perez-Gomez, Miguel A Hernan, Mayte Perez-Olmeda, Raquel Yotti, Jesus Oteo-Iglesias, Jose L Sanmartin, et al. 2020. “Infection Fatality Risk for Sars-Cov-2 in Community Dwelling Population of Spain: Nationwide Seroepidemiological Study.” *Bmj* 371.

Perez-Saez, Javier, Stephen A Lauer, Laurent Kaiser, Simon Regard, Elisabeth Delaporte, Idris Guessous, Silvia Stringhini, Andrew S Azman, Serocov-POP Study Group, and others. 2020. “Serology-Informed Estimates of Sars-Cov-2 Infection Fatality Risk in Geneva, Switzerland.” *medRxiv*.

Project, The Utah HERO. 2020. “Utah Hero Project: Tracking Covid-19 to Inform the Return to Normal.” Salt Lake City, UT: University of Utah.

Registry, Lithuania Central. 2020. “Koronaviruso (Covid-19) Lietuvoje Statistika.” Vilnius: Lithuania Central Registry.

Richardson S, Narasimhan M, Hirsch JS. 2020. “Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized with Covid-19 in the New York City Area.” *JAMA*. https://doi.org/10.1001/jama.2020.6775.

RIVM. 2020. “Children and Covid-19.” Amsterdam: Rijksinstituut voor Volksgezondheid en Milieu.

Rosenberg, Eli S, James M Tesoriero, Elizabeth M Rosenthal, Rakkoo Chung, Meredith A Barranco, Linda M Styer, Monica M Parker, et al. 2020. “Cumulative Incidence and Diagnosis of Sars-Cov-2 Infection in New York.” *medRxiv*.

Saúde, Instituto Nacional de. 2020. “Relatório de Apresentação Dos Resultados Preliminares Do Primeiro Inquérito Serológico Nacional Covid-19.” Lisbon: Instituto Nacional de Saúde.

Silva, Antônio Augusto Moura da, Lidio Gonçalves Lima-Neto, Conceição de Maria Pedrozo e Silva de Azevedo, Léa Márcia Melo da Costa, Maylla Luana Barbosa Martins Bragança, Allan Kardec Duailibe Barros Filho, Bernardo Bastos Wittlin, et al. 2020. “Population-Based Seroprevalence of Sars-Cov-2 Is More Than Halfway Through the Herd

Immunity Threshold in the State of Maranhão, Brazil." *medRxiv*. https://doi.org/10.1101/2020.08.28.20180463.

Streeck, Hendrik, Bianca Schulte, Beate Kuemmerer, Enrico Richter, Tobias Höller, Christine Fuhrmann, Eva Bartok, et al. 2020. "Infection Fatality Rate of Sars-Cov-2 Infection in a German Community with a Super-Spreading Event." *Medrxiv*.

Verity, Robert, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, et al. 2020b. "Estimates of the Severity of Coronavirus Disease 2019: A Model-Based Analysis." *The Lancet Infectious Diseases* 20 (6): 669–77.

Ward, Helen, Christina J Atchison, Matthew Whitaker, Kylie E. C. Ainslie, Joshua Elliott, Lucy C Okell, Rozlyn Redd, et al. 2020. "Antibody Prevalence for Sars-Cov-2 in England Following First Peak of the Pandemic: REACT2 Study in 100,000 Adults." *medRxiv*. https://doi.org/10.1101/2020.08.12.20173690.

Więcek, Witold, and Rachael Meager. 2020. "Baggr: Bayesian Aggregate Treatment Effects Package." Zenodo. https://doi.org/10.5281/zenodo.3813443.

Williamson, Elizabeth J., Alex J. Walker, Krishnan Bhaskaran, Seb Bacon, Chris Bates, Caroline E. Morton, Helen J. Curtis, et al. 2020. "Factors Associated with COVID-19-Related Death Using OpenSAFELY." *Nature* 584 (7821): 430–36. https://doi.org/10.1038/s41586-020-2521-4.