Submitted to the Annals of Applied Statistics

ESTIMATING GLOBAL AND COUNTRY-SPECIFIC EXCESS MORTALITY DURING THE COVID-19 PANDEMIC

BY VICTORIA KNUTSON[1], SERGE ALESHIN-GUENDEL[1], ARIEL KARLINSKY[2], WILLIAM MSEMBURI[3], JON WAKEFIELD[1,4]

[1]*Department of Biostatistics, University of Washington, Seattle, USA*

[2]*Hebrew University, Jerusalem, Israel*

[3]*World Health Organization, Geneva, Switzerland*

[4]*Department of Statistics, University of Washington, Seattle, USA*

Abstract: Estimating the true mortality burden of COVID-19 for every country in the world is a difficult, but crucial, public health endeavor. Attributing deaths, direct or indirect, to COVID-19 is problematic. A more attainable target is the "excess deaths", the number of deaths in a particular period, relative to that expected during "normal times", and we estimate this for all countries on a monthly time scale for 2020 and 2021. The excess mortality requires two numbers, the total deaths and the expected deaths, but the former is unavailable for many countries, and so modeling is required for these countries, and the expected deaths are based on historic data and we develop a model for producing expected estimates for all countries. We allow for uncertainty in the modeled expected numbers when calculating the excess. We describe the methods that were developed to produce World Health Organization (WHO) excess death estimates. To achieve both interpretability and transparency we developed a relatively simple overdispersed Poisson count framework, within which the various data types can be modeled. We use data from countries with national monthly data to build a predictive log-linear regression model with time-varying coefficients for countries without data. For a number of countries, subnational data only are available, and we construct a multinomial model for such data, based on the assumption that the fractions of deaths in specific sub-regions remain approximately constant over time. Our inferential approach is Bayesian, with the covariate predictive model being implemented in the fast and accurate `INLA` software. The subnational modeling was carried out using MCMC in `Stan` or in some non-standard data situations, using our own MCMC code. Based on our modeling, the 95% interval estimate for global excess mortality, over 2020–2021, is 13.3–16.6 million.

1. Introduction. The World Health Organization (WHO) has been tracking the impact of COVID-19 as the pandemic has evolved over time. Aggregate case and death numbers are being reported to the WHO and the data have been made publicly available at `https://covid19.who.int/`. For a number of reasons, these reported data neither provide a complete picture of the health burden attributable to COVID-19, nor of how many lives have been lost, directly and indirectly, due to the pandemic. Some deaths that are attributable to COVID-19 have not been certified as such because tests had not been conducted prior to death. Deaths may also be mistakenly certified as COVID-19, though this is less likely. It does not affect our estimates of excess mortality, based on all-cause mortality (ACM) data, however, only causing the resultant ratio of excess mortality to reported COVID-19 deaths to be lower than if such mistaken certification did not occur. There have also been variations in the death certification rules countries have applied in

Keywords and phrases: Bayesian inference, Global health, Poisson framework, Subnational modeling.

regards to COVID-19 (Garcia *et al.*, 2021; Riffe and Acosta, 2021). The impact of the pandemic is far reaching. Beyond the deaths directly attributable to it are those that can be linked to the conditions that have prevailed since the pandemic began and have led to some health systems being overwhelmed or some patients avoiding healthcare. In countries where COVID-19 spread was limited, due to lockdown measures or otherwise, some potential causes of death have decreased, such as those attributable to air pollution, or traffic accidents, or from other communicable diseases such as influenza like illness, resulting in negative excess or deficit deaths (Kung *et al.*, 2020; Karlinsky and Kobak, 2021). In light of the challenges posed by using reported COVID-19 data, excess mortality is considered a more objective and comparable (across countries) measure of the mortality impact of COVID-19 (Leon *et al.*, 2020). The WHO defines excess mortality as, "the mortality above what would be expected based on the non-crisis mortality rate in the population of interest" (`https://www.who.int/hac/about/definitions/en/`). Knowledge of the excess deaths not only paints a clearer picture of the pandemic, but can also aid in implementing public health initiatives.

The excess mortality in country c, ACM counts in month t for 2020 and 2021 are denoted by $Y_{c,t}$. These counts, in addition to the contribution from expected deaths, are assumed to be a result of the direct effects of COVID-19 (i.e., deaths attributable to it) and the indirect knock-on effects on health systems and society, along with deaths that were averted. The choice of a monthly time scale gives sufficient temporal resolution for most public health purposes. The hypothetical or "counterfactual" no-COVID-19 scenario uses the expected death numbers $E_{c,t}$, which have been forecasted to month t, using historic (prior to the pandemic) deaths data, usually over 2015–2019. Excess deaths are defined as:

$$\delta_{c,t} = Y_{c,t} - E_{c,t} \tag{1}$$

for country c where $c = 1, \ldots, 194$, and in month t where $t = 1, \ldots, 24$, represent months in 2020 and 2021.

The exercise of determining excess deaths for all countries is non-trivial, because the required ACM counts $Y_{c,t}$ are currently unavailable for many country/month combinations. Routine mortality data is often received by the WHO a year or more after the year of death. In addition, differential reporting capacity and variable data quality across countries has resulted in many nations lacking the systems to provide good quality routine data even historically (Mikkelsen *et al.*, 2015; Adair and Lopez, 2018; GBD, 2020; UNSD, 2021; Karlinsky, 2021). Correspondingly, these countries lack the capacity required to monitor ACM during the unprecedented COVID-19 pandemic. Hence, a number of countries are unable to contribute to the centralized systematic mortality surveillance that would be needed to measure global, regional and country level excess mortality by the WHO.

In this paper we describe our ongoing methods development to produce the WHO excess mortality estimates. In Section 2 we discuss data sources, before describing models for estimation of the expected numbers in Section 3. Section 4 describes our national covariate model and in Section 5 we outline the models we used for countries with subnational monthly data, national annual data, or a combination. Section 6 provides main results, with more extensive summaries appearing in the Supplementary Materials. Two other sets of global estimates of excess deaths have been produced by The Economist and the Institute for Health Metrics and Evaluation (IHME) with the latter being described in Wang *et al.* (2022). We fully describe and critique these methods in Section 7. The paper concludes with a discussion in Section 8.

2. Data Sources.

2.1. *Mortality Data.* Excess mortality cannot be directly measured for all countries due to many not having the required ACM data. The WHO usually receives routine mortality data on an annual basis in the year after the year of death or perhaps after an even greater lag. Civil registration and vital statistics (CRVS) systems differ greatly across countries with varying timelines and quality control measures for compiling unit record cause-of-death numbers into aggregates identified by cause, age, sex, place, and period of death. In addition, differential reporting coverage, the absence of electronic surveillance systems in some locations and limited investments in CRVS systems has resulted in many nations lacking the structures necessary to provide good quality routine data, even before the COVID-19 pandemic. This lack of capacity and the data required to monitor ACM has been exacerbated during the unprecedented pandemic. Therefore, many countries are unable to contribute to a centralized systematic mortality surveillance that would be needed to measure global, regional and country level excess mortality by the WHO.

Region	Full National	Partial National	Mixed Data	No Data	Total	Proportion Population
AFRO	4	2	0	41	47	0.13
AMRO	11	12	4	9	36	0.90
EMRO	4	5	0	12	21	0.32
EURO	46	5	1	1	53	0.89
SEARO	1	1	3	6	11	0.04
WPRO	6	3	2	16	27	0.18
Global	72	27	10	85	194	0.33

TABLE 1

Country data availability summary for 2020 and 2021. Full national countries have data over all 24 months and partial national have data for less than 24 months; for example, 83 countries have data for at least the first 18 months, and 96 countries have data for at least the first 12 months. Mixed data refers to countries with subnational monthly data for some period (4 countries), national annual data (5 countries) or a combination (China). WHO regions: African Region (AFRO), Region of the Americas (AMRO), Eastern Mediterranean Region (EMRO), European Region (EURO), South-East Asian Region (SEARO), Western Pacific Region (WPRO). The proportion of the population that are observed column is calculated at the country-month level. The Supplementary Materials include a table that lists the type of data available for each country.

All countries report their official COVID-19 death count, but we would not expect this to be accurate, and for many countries we would expect serious underestimation, for the reasons already outlined and for political reasons. However, the official count does provide an interesting summary for comparison with the estimated excess, and the COVID-19 death rate is used as a covariate in our ACM estimation model.

For this study, our main sources of data are reports of ACM as collected and reported by countries' relevant institutions – from national statistics offices, ministries of health, population registries, etc. These have been collected in several repositories such as the data routinely shared with WHO as part of its standing agreement with member states, Eurostat, The Human Mortality Database (HMD) as part of the Short-Term Mortality Fluctuations (STMF) project (Németh *et al.*, 2021) and the World Mortality Dataset (WMD), as described in Karlinsky and Kobak (2021). Monthly data are included after accounting for delayed registration either by adjusting for registration delay (Australia, Brazil, United States) or by not-including highly incomplete months.

In this paper we report the current state of data at our disposal. This project is ongoing and data is added as soon as available. Table 1 shows the breakdown of data availability by WHO region. Just over a half (99) of the 194 countries provide monthly national data

from at least some of the pandemic period, while 10 other countries provide subnational monthly data, national annual data, or a combination of the two (this includes Argentina which has partial national and subnational data, so could fit in with partial or mixed data). It is immediately clear that there is a huge regional imbalance in data availability, with the EURO region being very well represented, the AMRO region having data from 75% of the countries, and other regions being more poorly represented. For example, in the AFRO region we only have data from 6 out of 47 countries. For those countries with data in month t, we assume that the ACM part of the excess $\delta_{c,t}$, as defined in (1), is known exactly. Hence, we do not account for inaccuracies in the reported deaths (beyond the aforementioned accounting for delayed registration). For all countries we do, however, account for uncertainty in the expected numbers.

2.2. *Covariate Data.* For countries with no data, we predict the ACM count using a log-linear covariate model. A range of covariates were considered, including a high income country binary indicator, COVID-19 test positivity rate, COVID-19 death rate, temperature, population density, socio-demographic index (SDI), human development index (HDI), stringency (index for lockdown restrictions and closures, overall government response, economic (including measures such as income support and debt relief), containment (this index combines "lockdown" restrictions and closures with measures such as testing policy and contact tracing, short term investment in healthcare, as well investments in vaccines – it is calculated using all ordinal containment and closure policy indicators and health system policy indicators, for further details see Hale *et al.* (2020)), historic (from 2019) non-communicable disease rates, historic cardiovascular disease rate, historic HIV rate, historic diabetes prevalence, life expectancy, proportion of the population under-15, proportion of the population over-65. Some of the covariates are time-varying (COVID-19 test positivity rate, COVID-19 death rate, temperature, stringency, overall government response, containment), while the remainder are constant over time. A number of the covariates were not available by month for all countries and so their values were imputed. Specifically, (WHO) regional medians were used for countries with missing data. Details are given in the Supplementary Materials.

3. Expected Mortality Modeling. A key component of the excess mortality calculation is the ACM count that would be expected in non-pandemic times, for each country and month. We describe models for two types of countries: those that have historic monthly ACM data, and those that have historic annual ACM data only – 100 countries have historic monthly data and 94 have historic annual data. In terms of the period upon which we base the expected numbers, it is usually 2015–2019 for countries with monthly historical data, and is usually 2000–2019 for countries with annual historical data.

3.1. *Countries with Monthly Data.* We consider first those countries with monthly ACM data over multiple years (usually 2015–2019). For country c, $Y_{c,t}$ represents the ACM count for country c and month t, for $t = 1, \ldots, M_c$, where M_c is the number of historic months for which we have data. We assume the sampling model for $Y_{c,t}$ is,

$$Y_{c,t}|\mu_{c,t} \sim \text{NegBin}(\mu^{\text{E}}_{c,t}, \phi^{\text{E}}_c),$$

parametrized in terms of the mean, $\mu^{\text{E}}_{c,t}$, and the overdispersion parameter, ϕ^{E}_c, such that $\text{var}(Y_{c,t}|\mu^{\text{E}}_{c,t}, \phi^{\text{E}}_c) = \mu^{\text{E}}_{c,t}(1 + \mu^{\text{E}}_{c,t}/\phi^{\text{E}}_c)$, with the Poisson model being recovered as $\phi^{\text{E}}_c \to \infty$. We let $v[t]$ index the year in which month t occurred (for example, labeled $1, \ldots, 5$ when data are available for 2015–2019) and $m[t]$ be the month (labeled $1, \ldots, 12$), so that given v, m we can find t as $t = 12(v-1) + m$. The mean is modeled as,

$$\eta_{c,t} = \log(\mu_{c,t}) = f^{\text{y}}_c(v[t]) + f^{\text{m}}_c(m[t]) \tag{2}$$

where $f_c^{\mathrm{y}}(\cdot)$ models the *annual trend*, and $f_c^{\mathrm{m}}(\cdot)$ is a smooth function of time t which accounts for *within-year* seasonal variation. The yearly trend is modeled with a thin-plate spline and within-year variation with a cyclic cubic spline (Rivera *et al.*, 2020). In both cases we use the `gam` function in the `mgcv` package with generalized cross-validation (Wood, 2017, Section 4.5.3) used to select smoothing parameters. The spline model is fitted separately for each country. Algeria, Iraq and Sri Lanka have less than three years of historical data, and so a linear term is used for modeling yearly variation. This model is used to obtain predictions of the expected deaths $\mu_{c,t}^{\mathrm{E}}$ for all t in 2020 and 2021, with both a point estimate and a standard error being produced.

3.2. *Countries with Annual Data.* For countries with only annual historic data, the goal is to predict expected numbers by month t for $t = 1, \ldots, 24$. We summarize our strategy for producing expected numbers for countries with annual data only:

1. Fit a negative binomial spline model to the countries with annual counts only. Use the spline to predict the total annual ACM for 2020 and 2021, for these countries.
2. In a separate exercise, fit the multinomial model to all of the countries with monthly data, with deaths being attributed via the log-linear temperature model (3). This produces an estimate $\widehat{\beta}$.
3. Combine the spline model with the multinomial model using monthly temperature apportionment to obtain expected numbers for the countries without monthly data.

The annual trend can be estimated for each country using the method we described in the previous section minus the monthly term, i.e., a spline in year. To apportion the yearly totals to the months, we use the fact that a collection of Poisson random variables conditioned on their sum produce a multinomial distribution with within-year variation modeled using temperature, which is acting as a surrogate for seasonality. This relationship is learned from countries with historic monthly data. We use a smooth series of monthly temperatures since 2015. Let $\boldsymbol{Y}_{c,v} = \{Y_{c,v,m}, m = 1, \ldots, 12\}$ be the vector that contains the ACM counts by month in year v, $v = 1, \ldots, 5$. Suppose each of the 12 constituent counts are Poisson with mean $\zeta_{c,v,m}$, for $m = 1, \ldots, 12$. Then, within the year, conditional on the total ACM,

$$\boldsymbol{Y}_{c,v} | Y_{c,v}^{+}, \boldsymbol{p}_{c,v} \sim \text{Multinomial}(Y_{c,v}^{+}, \boldsymbol{p}_{c,v}),$$

where $Y_{c,v}^{+}$ is the national count in country c and year v and $\boldsymbol{p}_{c,v} = \{p_{c,v,m}, m = 1, \ldots, 12\}$ with

$$p_{c,v,m} = \frac{\zeta_{c,v,m}}{\sum_{m'=1}^{12} \zeta_{c,v,m'}},$$

We assume,

$$\log(\zeta_{c,v,m}) = z_{c,v,m}\beta \tag{3}$$

where $z_{c,v,m}$ is the temperature and β is the associated log-linear coefficient. The multinomial model can be fitted in `INLA` using the Poisson trick (Baker, 1994) which involves fitting the Poisson model for the data in country c, month m:

$$Y_{c,v,m} | \lambda_{c,v} \sim \text{Poisson}(\ \lambda_{c,v} \mathrm{e}^{z_{c,v,m}\beta}\ ),$$

where the $\lambda_{c,v}$ parameters are given (improper) priors $\pi(\lambda_{c,v}) \propto 1/\lambda_{c,v}$. Further details may be found in the Supplementary Materials.

The estimated expected counts are shown in blue in Figure 1, for selected countries.

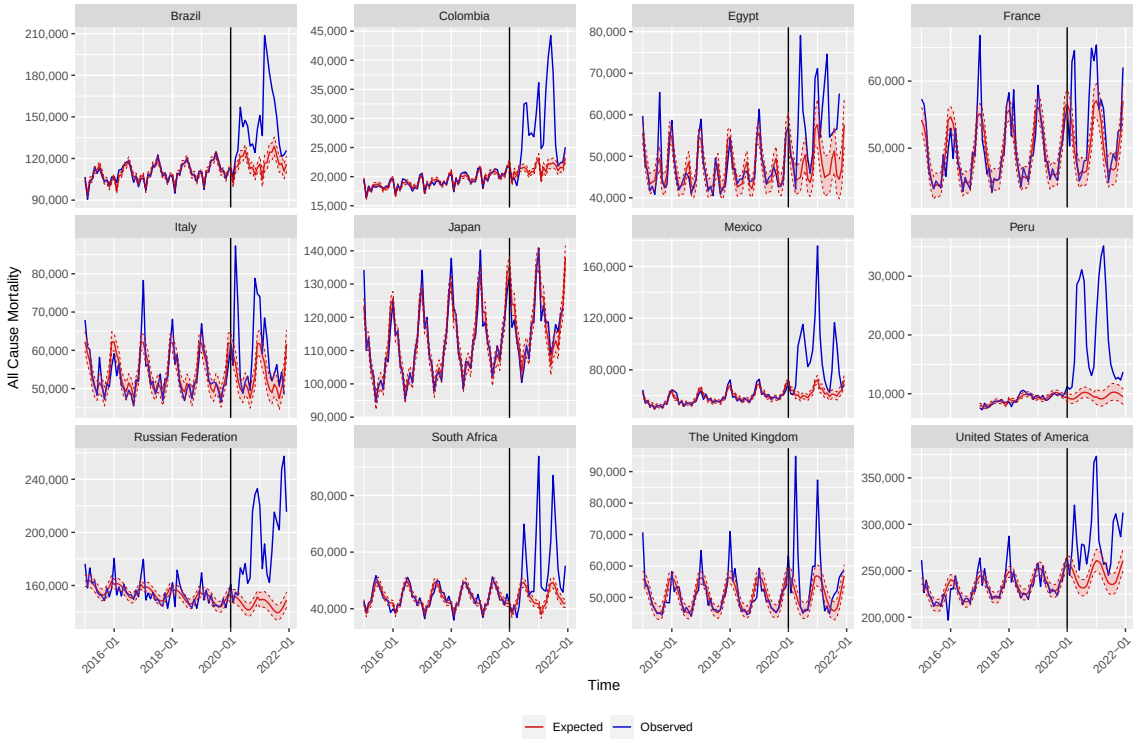

FIG 1. *Monthly time series of all cause mortality: expected counts in red and observed counts in blue, for selected countries. The black vertical line is drawn at the start of 2020. The dashed red bands denote 95% uncertainty intervals for the mean expected numbers. For these countries, ACM counts are available for all months apart from Egypt, for which the last month is missing.*

3.3. *Modeling Uncertainty in the Expected Numbers.* For all countries the expected numbers appear directly in the excess calculation, (1). In addition, for countries with no pandemic ACM data, the Poisson model we adopt for covariate modeling includes the expected number as an offset. For all countries and months, we obtain not just an estimate of the mean expected mortality but also a measure of the uncertainty (due to uncertainty in estimating the spline model) in this estimate. We now describe how the uncertainty in the mean expected count is accounted for in our modeling.

For countries with monthly data, we use the spline model to predict the log of the mean expected number of deaths. Asymptotically, the estimator for the log of the mean expected numbers is normally distributed. Let $\widehat{\eta}_{c,t'}$ and $\widehat{\sigma}^2_{c,t'}$ represent the mean and standard deviation of the prediction for pandemic months, labeled as $t' = 1, \ldots, 24$. We simulate S samples from the asymptotic normal sampling distribution with mean $\widehat{\eta}_{c,t'}$ and standard deviation $\widehat{\sigma}_{c,t'}$; denote these samples by $\eta^{(s)}_{c,t'}$, $s = 1, \ldots, S$. We then transform the samples so that we have samples for the expected numbers $E^{(s)}_{c,t'} = \exp(\eta^{(s)}_{c,t'})$, for $s = 1, \ldots, S$. We then use the method of moments to fit a gamma distribution to these S samples with shape $\tau_{c,t'}$ and rate $\tau_{c,t'}/E_{c,t'}$. In particular, letting $m_{c,t'}$ denote the sample mean, and $V_{c,t'}$ denote the sample variance, we set $\widehat{E}_{c,t'} = m_{c,t'}$ and $\widehat{\tau}_{c,t'} = m^2_{c,t'}/V_{c,t'}$. We approximate the distribution of the expected numbers as gamma, since this is conjugate to the Poisson, and so allows efficient inference with INLA (Rue *et al.*, 2009) using a negative binomial, as we describe in Section 4. Effectively, we are approximating the sampling distribution of the mean expected count by a gamma distribution.

We now consider a generic country c with yearly data only. In pandemic year v', we use the spline model to predict the log of the expected number of deaths. Let $\widehat{\eta}_{c,v'}$ and $\widehat{\sigma}^2_{c,v'}$ represent the mean and standard deviation of the prediction, for $v' = 1, 2$ (the two pandemic years). We then simulate S samples from a normal distribution with mean $\widehat{\eta}_{c,v'}$ and standard deviation $\widehat{\sigma}_{c,v'}$; denote these samples by $\eta^{(s)}_{c,v'}$, $s = 1, \ldots, S$. We then transform the samples so that we have samples for the expected numbers $E^{(s)}_{c,v'} = \exp(\eta^{(s)}_{c,v'})$, for $s = 1, \ldots, S$. We then apply the monthly temperature model to produce predictions of the proportion of deaths in each month in each year, i.e., for a given pandemic month m', we have S samples of the predicted proportion of deaths in month m' of year v', $p^{(s)}_{c,v',m'}$, for $s = 1, \ldots, S$. Converting to pandemic cumulative months $t' = 12(v' - 1) + m'$ we then produce samples of the expected number of deaths in month t', as $E^{(s)}_{c,t'} = E^{(s)}_{c,v'} \times p^{(s)}_{c,v',m'}$. We then use the method of moments to fit a gamma distribution to these S samples as for the countries with monthly data. To summarize, in both cases we have a distribution for $E_{c,t'}$ which is Gamma$(\widehat{\tau}_{c,t'}, \widehat{\tau}_{c,t'}/\widehat{E}_{c,t'})$. The Supplementary Materials provide comparisons of the true distribution of the mean expected counts and the approximating gamma distributions, and illustrates that the latter are accurate. We also experimented with including negative binomial sampling variability in the calculation of the expected numbers, but it made little additional contribution to the intervals for the excess.

In the next section we describe a Bayesian modeling of ACM in the pandemic, for countries without data. Inference for the expected numbers is frequentist, and we sample from the asymptotic normal distribution, but with flat priors, this will approximate a Bayesian analysis, and so when we combine the two components in the excess (1) we view the resultant inference as Bayesian.

We next describe how we model ACM – we have different models for different data scenarios but in each case the starting point is the Poisson distribution.

4. National Mortality Models for Countries with No Data. For countries with observed monthly national ACM data, $Y_{c,t}$, we use these directly in the excess calculation. In the countries with no data we need to estimate the ACM count. We follow a Bayesian approach so that for countries without data we obtain a predictive distribution over this count and this, when combined with the gamma distribution for the expected numbers, gives a distribution for the excess $\delta_{c,t}$.

In Figure 2 we plot the monthly counts for a range of countries with monthly ACM data, along with the reported COVID-19 deaths and the expected numbers. We see very different scenarios in different countries. In all countries but Japan there is a clear large difference between the observed and the expected, though within each country this difference shows large fluctuations over time. In Figure 3, again for countries with monthly ACM data, we plot the excess $\delta_{c,t} = Y_{c,t} - E_{c,t}$, as a function of month t (including uncertainty in the expected numbers), along with the reported COVID-19 deaths. As expected, $\delta_{c,t}$ is greater than the reported overall in general, but not for Japan, and for most countries displayed the difference between the excess and the reported shows a complex temporal pattern.

While complex models that attempt to pick up data nuances are desirable, given the idiosyncrasies of the different data sources described in Section 2, any modeling exercise is fraught with difficulties, and we resort to a relatively simple model in which we build an overdispersed Poisson log-linear regression model for the available monthly ACM data to predict the monthly ACM in those countries with no data. We cannot overemphasize the regional imbalance of the missing ACM data – in the AFRO region in particular, our estimates should be viewed with extreme caution, since they are predicted from data which overwhelmingly is from other regions.

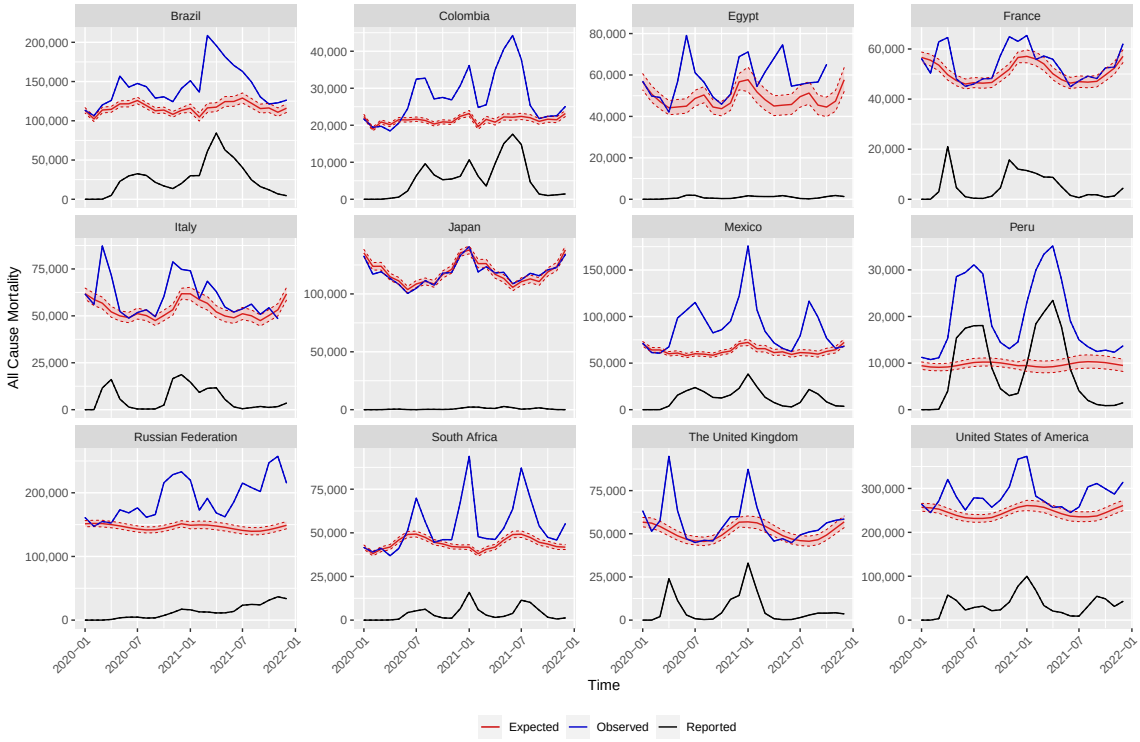

FIG 2. *Monthly time series of ACM counts, expected counts (with 95% interval estimates) and reported COVID-19 mortality counts, for selected countries. ACM counts are available for all months apart from Egypt, for which the last month is missing.*

The basic starting model is

$$Y_{c,t}|E_{c,t},\theta_{c,t} \sim \text{Poisson}(E_{c,t}\theta_{c,t}), \tag{4}$$

so that $\theta_{c,t} > 0$ is a relative rate parameter, with $\theta_{c,t} > 1$ / $\theta_{c,t} < 1$ corresponding to a higher/lower ACM rate than expected, based on historic data. Recall, from Section 3, that we model the distribution of the expected counts $E_{c,t}$ as Gamma$(\widehat{\tau}_{c,t}, \widehat{\tau}_{c,t}/\widehat{E}_{c,t})$. When combined with (4), we obtain the sampling model,

$$Y_{c,t}|\theta_{c,t} \sim \text{NegBin}(\widehat{E}_{c,t}\theta_{c,t}, \widehat{\tau}_{c,t})$$

with known overdispersion parameter $\widehat{\tau}_{c,t}$ to give var$(Y_{c,t}|\theta_{c,t}) = \widehat{E}_{c,t}\theta_{c,t}(1 + \widehat{E}_{c,t}\theta_{c,t}/\widehat{\tau}_{c,t})$. The mean is $\mathrm{E}[Y_{c,t}|\theta_{c,t}] = \widehat{E}_{c,t}\theta_{c,t}$. The relative rate parameter $\theta_{c,t}$ is modeled as