SUPPLEMENTARY INFORMATION

A. Additional examples of mortality-ratio evolutions

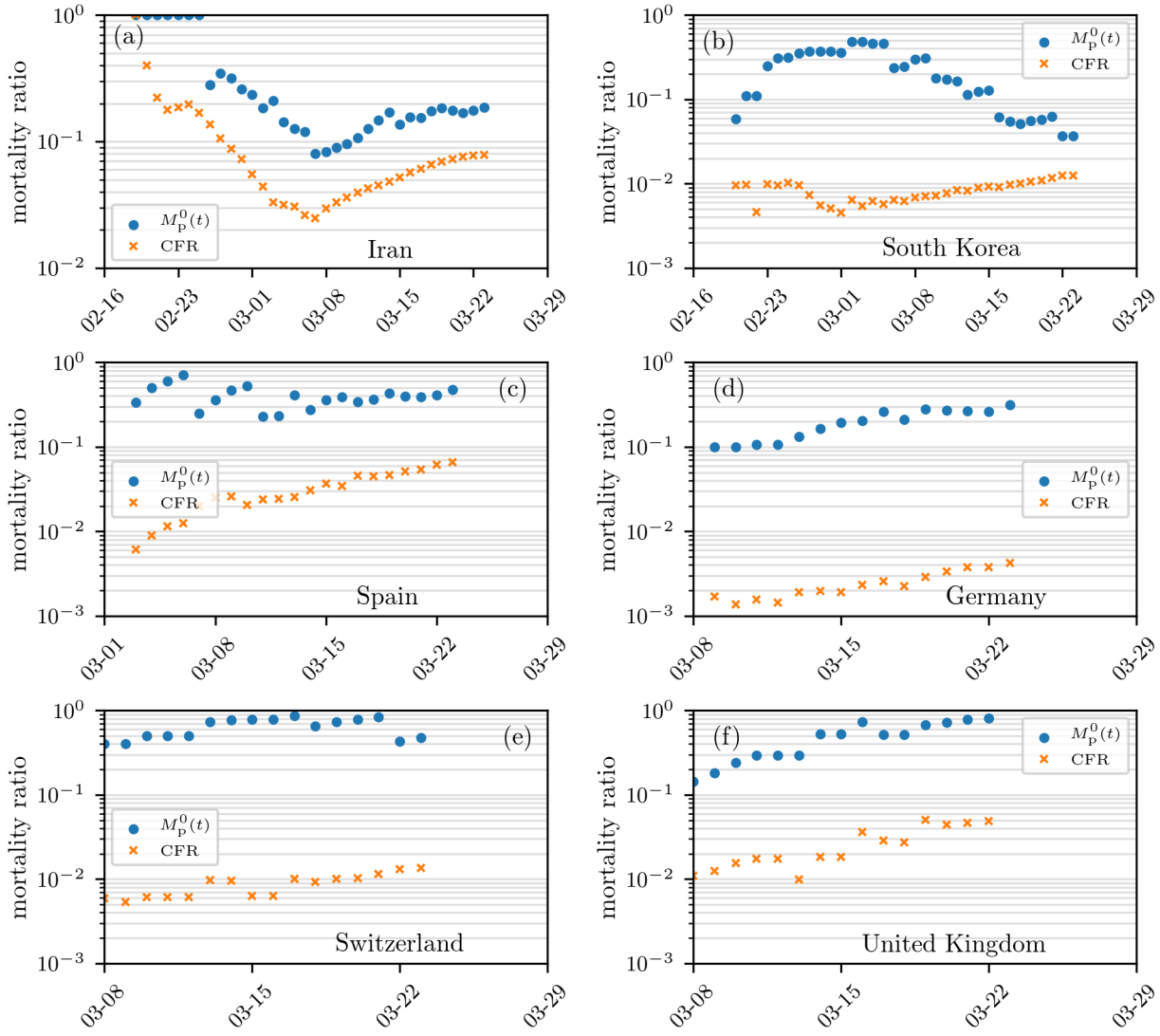

FIG. S1. **Mortality ratio estimates**. Estimates of mortality ratios (see Eqs. (8) and (14) in the main text) of SARS-CoV-2 infections in different countries. The case fatality ratio, CFR, corresponds to the number of deaths to date divided by the total number of cases to date. Another population-based mortality ratio is $M_p^0(t)$, the number of deaths divided by the sum of deaths and recovereds, up to time t. The data are derived from Ref. [5].

In Fig. S1, we show additional examples of mortality-ratio estimates for Iran, South Korea, Spain, Germany, Switzerland, and the United Kingdom. As in Fig. 1 in the main text, we observe that, by definition, the population-based mortality ratio $M_p^0(t)$ is significantly larger than the corresponding CFR in all cases.

B. Solutions for τ_1-averaged probabilities

Using the method of characteristics, we find the formal solution to Eq. (1):

$$P(\tau,t|\tau_1) = \delta(\tau-t-\tau_1)e^{-\int_0^t(\mu(\tau-t+s,s|\tau_1)+c(\tau-t+s,s|\tau_1))\mathrm{d}s}, \tag{S1}$$

which can be used to construct the death and cure probabilities

$$P_{\rm d}(t|\tau_1) = \int_0^t {\rm d}t'\, \mu(\tau_1+t',t') e^{-\int_0^{t'}(\mu(\tau_1+s,s)+c(\tau_1+s,s)){\rm d}s}$$
$$P_{\rm r}(t|\tau_1) = \int_0^t {\rm d}t'\, c(\tau_1+t',t') e^{-\int_0^{t'}(\mu(\tau_1+s,s)+c(\tau_1+s,s)){\rm d}s}. \tag{S2}$$

If we now invoke the functional forms of μ and c given in Eq. (4), we find explicitly

$$P_{\rm d}(\tau,t|\tau_1) = \begin{cases} \dfrac{\mu_1}{\mu_1+c}\left(1-e^{-(\mu_1+c)t}\right) & \tau > t+\tau_{\rm inc} \\ 0 & \tau_{\rm inc} \geq \tau > \tau_1 \\ \dfrac{\mu_1 e^{-c(\tau_{\rm inc}-\tau_1)}}{\mu_1+c}\left(1-e^{-(\mu_1+c)(\tau-\tau_{\rm inc})}\right) & \tau > \tau_{\rm inc} \geq \tau_1 \end{cases} \tag{S3}$$

and

$$P_{\rm r}(\tau,t|\tau_1) = \begin{cases} \dfrac{c}{\mu_1+c}\left(1-e^{-(\mu_1+c)t}\right) & \tau > t+\tau_{\rm inc} \\ 1-e^{-ct} & \tau_{\rm inc} \geq \tau > \tau_1 \\ 1-e^{-c(\tau_{\rm inc}-\tau_1)} + \dfrac{ce^{-c(\tau_{\rm inc}-\tau_1)}}{\mu_1+c}\left(1-e^{-(\mu_1+c)(\tau-\tau_{\rm inc})}\right) & \tau > \tau_{\rm inc} \geq \tau_1. \end{cases} \tag{S4}$$

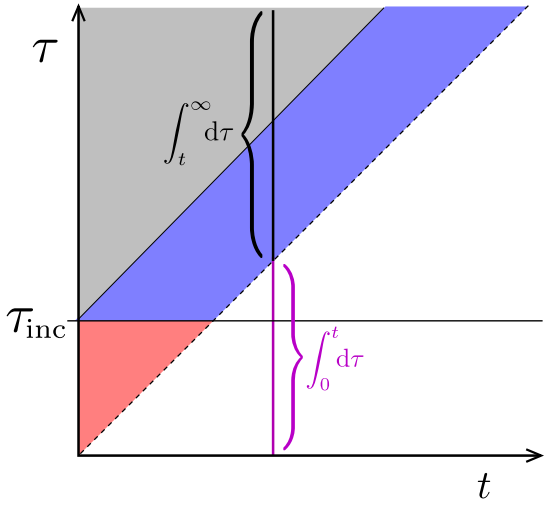

FIG. S2. **Phase plot for $P(\tau > t, t)$ and $I(\tau > t, t)$.** The regions delineating different forms for the solution (Eq. (S5)). Here, we have included an incubation time $\tau_{\rm inc}$ before which no death occurs. The solution for $\bar{P}(\tau,t)$ or $I(\tau,t)$ in the $\tau < t$ region must be self-consistently solved using the boundary condition Eq. (10). At any fixed time, the integral of $I(\tau,t)$ over $t < \tau \leq \infty$ captures only the initial population, excludes newly infecteds, and is used to compute $D_1(t), R_1(t)$, and $M_{\rm p}^1(t)$. To compute $D_0(t), R_0(t)$, and $M_{\rm p}^0(t)$, we integrate across all infecteds (including the integral over $t > \tau > \geq 0$ shown in magenta).

Finally, we can also find the τ_1-averaged probabilities for $\tau \geq t$ by weighting over $\rho(\tau_1; n, \gamma)$. For example,

$$\bar{P}(\tau,t) = \begin{cases} \rho(\tau-t;n,\gamma)e^{-(\mu_1+c)t} & \tau \geq t+\tau_{\rm inc} \\ \rho(\tau-t;n,\gamma)e^{-ct} & \tau_{\rm inc} \geq \tau > t \\ \rho(\tau-t;n,\gamma)e^{-ct}e^{-\mu_1(\tau-\tau_{\rm inc})} & t+\tau_{\rm inc} \geq \tau > \tau_{\rm inc} \end{cases}.$$

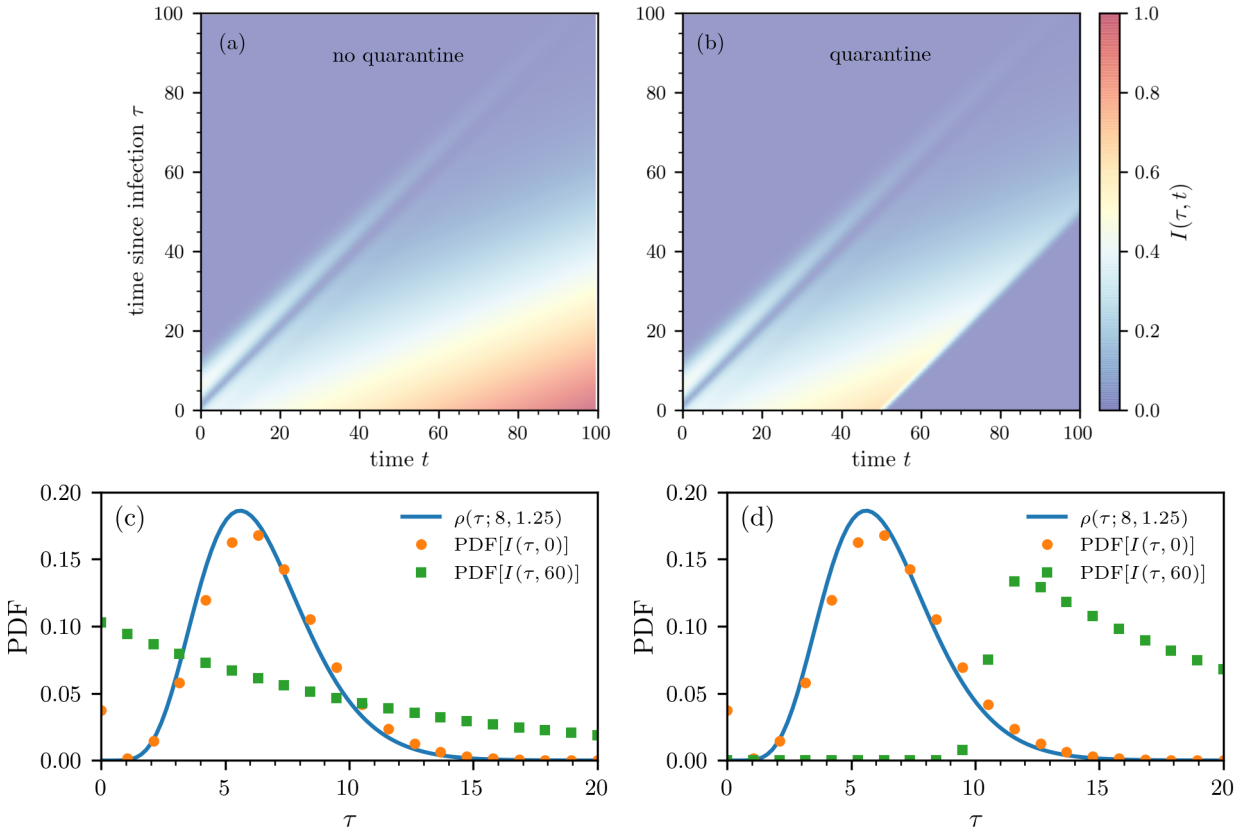

FIG. S3. **Density plots of $I(\tau, t)$ in the $t - \tau$ plane.** Numerical solution of the equation for $I(\tau, t)$ in Eqs. (9) under the assumption of a fixed susceptible size $\beta_1 S = 0.158$/day. (a) The density without quarantine monotonically grows with time t in the region $\tau < t$ as an unlimited number of susceptibles continually produces infecteds. (b) With quarantining after $t_q = 50$ days, we set $\beta_1 S = 0$ for $t > t_q$, which shuts off new infections. Both plots were generated using the same initial density $\rho(\tau_1)$ defined in Eq. (7). In both cases, the density $I(\tau > t)$ is identical to $P(\tau > t)$ if the same $\rho(\tau_1)$ is used and is independent of disease transmission, susceptible dynamics, etc. (c-d) Probability-density functions (PDFs) of the number of infected $I(\tau, t)$ for $t = 0, 60$ (b) without and (c) with quarantine. The blue solid line corresponds to the initial distribution $\rho(\tau; n = 8, \gamma = 1.25)$ (see Eq. (7)).

These solutions hold for the different regions shown in the phase plot of Fig. S2 and are equivalent to those for $I(\tau > t, t)$. Corresponding expressions for $\bar{P}_d(t)$ and $\bar{P}_r(t)$ can be found and used to construct $M_p^1(t)$. Fig. S3(a) shows the magnitude of $I(\tau, t)$ in the $t - \tau$ plane when we set $S(t) = S$ constant (so that the first equation in Eq. (18) does not apply) such that $\beta_1 SS \approx 0.158$/day. In this case, the epidemic continues to grow in time, but the mortality rates $M_p^{0,1}(t)$ nonetheless converge as $t \to \infty$. In Fig. S3(b), we set $\beta_1 S = 0$ for $t > t_q$ to model strict quarantining after $t_q = 50$ days. We observe no new infections after the onset of strict quarantine measures. In both cases (quarantine and no quarantine), we use $\rho(\tau; n = 8, \gamma = 1.25)$ (see Eq. (7) in the main text) to describe the initial distribution of infection times τ. As time progresses, more of the distribution of τ moves towards smaller values until quarantine measures take effect (see Fig. S3(c) and (d)).

C. Effects of undertesting

Note that $I(\tau, t)$ in the SIR equations determines the dynamics of the actual infected population. However, (i) typically only a fraction f of the total number of infecteds might be tested and confirmed positive and (ii) the testing of newly infecteds may also be delayed by a distribution $\rho(\tau; n, \gamma)$.

If positive tests represent only a fraction f of the total infected population, and the confirmation of newly infecteds occurs immediately, the known infected density is given by $I^*(\tau, t) = fI(\tau, t)$ where $I(\tau, t)$ is the true total infected population. If testing of newly infecteds occurs after a distribution $\rho(\tau; n, \gamma)$ of infection times, $I^*(\tau, t) = f \int_0^\tau I(t - \tau + \tau_1, t)\rho(\tau_1; n, \gamma) d\tau_1$.

In our development of $M_{\rm p}^{0,1}(t)$ and $\mathrm{CFR}_{\rm d}(t,\tau_{\rm res})$ in the manuscript, we assumed the entire infected population was tested and confirmed. Thus, $M_{\rm p}^{0,1}(t)$ and $\mathrm{CFR}_{\rm d}(t,\tau_{\rm res})$ were computed using $f=1$ and more accurately represent the mortality ratios of the population *conditioned* on being tested positive.

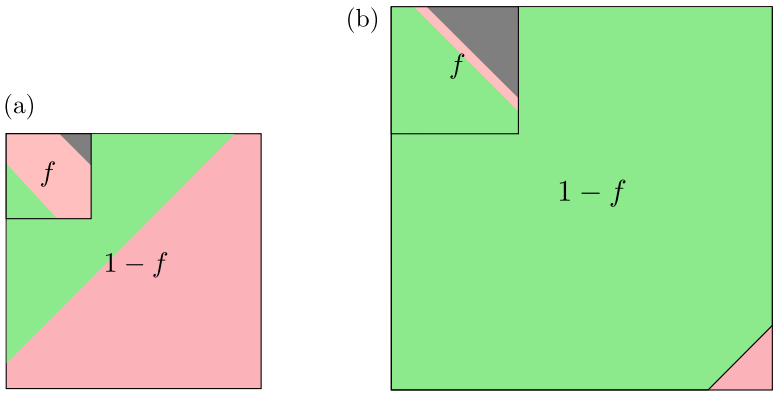

FIG. S4. **Fractional testing.** An example of fractional testing in which a fixed fraction f of the real total infected population is assumed to be tested. The remaining $1-f$ proportion of infecteds are untested. Equivalently, if the total tested fraction has unit population, then the total population of the untested pool is $1/f-1$. (a) At short times after an outbreak, the known tested infected population has not yet resolved and is composed of deaths (gray), recovereds (green), and infecteds (red). We assume that the untested fraction of infecteds (red) have mild or no symptoms, do not die, and can only recover (green). (b) At longer times, the infecteds further resolve. The $M_{\rm p}(t)$ and CFR metrics that are based on only the tested fraction will overestimate the true mortality fraction of all infected cases.

To estimate the mortality ratio of the population conditioned simply on being infected, we have to estimate the larger number of recovereds that went untested. As shown in Fig. S4, the untested recovered fraction can be estimated by assuming that the death rate for the untested infecteds is zero and by writing an SIR model without death for the untested pool of infecteds

$$\begin{aligned}
\frac{\mathrm{d}S(t)}{\mathrm{d}t} &= -S(t)\int_0^\infty \mathrm{d}\tau'\,\beta(\tau',t)(I^*(\tau',t)+I^{\rm u}(\tau',t)),\\
\frac{\partial I^*(\tau,t)}{\partial t}+\frac{\partial I^*(\tau,t)}{\partial \tau} &= -(\mu(\tau,t)+c(\tau,t))I^*(\tau,t),\\
\frac{\partial I^{\rm u}(\tau,t)}{\partial t}+\frac{\partial I^{\rm u}(\tau,t)}{\partial \tau} &= -c(\tau,t)I^{\rm u}(\tau,t),\\
\frac{\mathrm{d}R(t)}{\mathrm{d}t} &= \int_0^\infty \mathrm{d}\tau c(\tau,t)(I^*(\tau,t)+I^{\rm u}(\tau,t)),
\end{aligned} \tag{S5}$$

where $I(\tau,t)=I^*(\tau,t)+I^{\rm u}(\tau,t)$. The true mortality ratio is then straightforwardly defined by, for example,

$$\mathcal{M}_{\rm p}^0(t)=\frac{D_0^*}{D_0^*(t)+R_0^*(t)+R_0^{\rm u}(t)}, \tag{S6}$$

where

$$\begin{aligned}
D_0^*(t)=\int_0^\infty \mathrm{d}\tau\int_0^t \mathrm{d}t'\,\mu(\tau,t')I^*(\tau,t'),\quad R_0^*(t)=\int_0^\infty \mathrm{d}\tau\int_0^t \mathrm{d}t'\,c(\tau,t')I^*(\tau,t'),\\
\text{and}\quad R_0^{\rm u}(t)=\int_0^\infty \mathrm{d}\tau\int_0^t \mathrm{d}t'\,c(\tau,t')I^{\rm u}(\tau,t'),
\end{aligned} \tag{S7}$$

with analogous expressions for $D_1^*(t)$, $R_1^*(t)$, and $R_1^{\rm u}(t)$. At long times, after resolution of all infecteds, the untested recovered population is

$$R_{0,1}^{\rm u}(\infty)=\left(\frac{1}{f}-1\right)(D_{0,1}^*(t)+R_{0,1}^*(t)), \tag{S8}$$

which yields the asymptotic true ratio $\mathcal{M}_{\rm p}^{0,1}(\infty) = f M_{\rm p}^{0,1}(\infty)$ as described in the Discussion and Summary. In this simple rescaling to account for untested populations, we have assumed that all deaths come from the tested pool and that the recovery rate c is the same in the tested and untested pools.

D. Influence of different transmission rates

In Fig. 3 of the main text, we observe that the population-level mortality ratio $M_{\rm p}^0(t)$ approaches a plateau during the initial exponential growth phase of an epidemic (*i.e.*, for $S(t) \approx S_0$). If the number of new infections decreases (*e.g.*, due to quarantine measures), $M_{\rm p}^0(t)$ starts growing until it reaches its asymptotic value $M_{\rm p}^0(\infty)$. Interestingly, the pre-asymptotic values of $M_{\rm p}^0(t)$ are smaller for larger infection rates β_1 (see Fig. S5(a)). This counter-intuitive effect arises because larger values of β_1 generate relatively larger numbers of new infected which have a lower chance of dying before $\tau_{\rm inc}$ (see Eq. (4) in the main text). A similar effect occurs for non-delayed transmission (*i.e.*, $\tau_\beta \approx 0$).

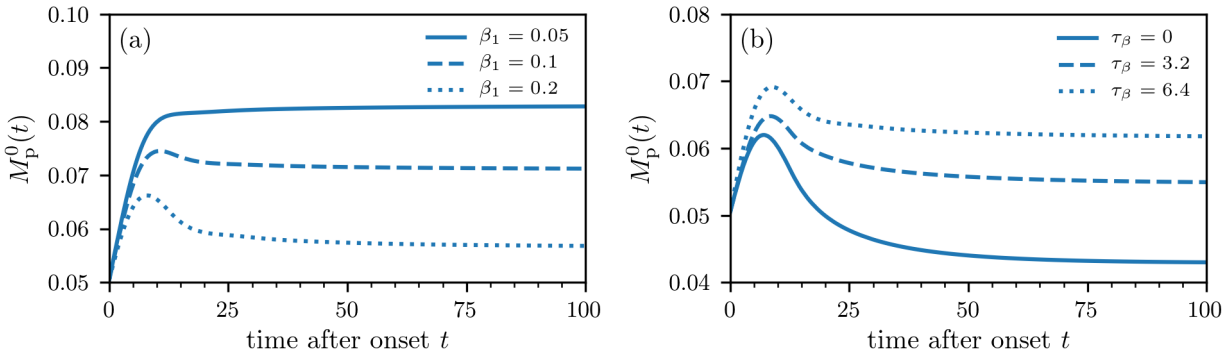

FIG. S5. **Population-level mortality for different infection rates.** (a) The population-level mortality ratio $M_{\rm p}^0(t)$ for different values of β_1 and an incubation time of $\tau_{\rm inc} = 6.4$ days. In the initial exponential growth phase of the epidemic (*i.e.*, $S(t) \approx S_0$), larger infection rates β_1 lead to smaller values of $M_{\rm p}^0(t)$. (b) We observe a similar effect for non-delayed transmissions (*i.e.*, $\tau_\beta \approx 0$). As long as $S(t) \approx S_0$, smaller transmission delays τ_β lead to larger relative numbers of new infections and smaller $M_{\rm p}^0(t)$.

As the transmission delay decreases, more secondary cases will result from one infection, leading to smaller values of $M_{\rm p}^0(t)$ in the initial exponential growth phase of an epidemic (see Fig. S5(b)).