Supplementary Information:

The timing of COVID-19 transmission

Luca Ferretti[1], Alice Ledda[2], Chris Wymant[1], Lele Zhao[1], Virginia Ledda[3], Lucie Abeler-Dörner[1], Michelle Kendall[1], Anel Nurtay[1], Hao-Yuan Cheng[4], Ta-Chou Ng[5], Hsien-Ho Lin[5], Rob Hinch[1], Joanna Masel[6], A. Marm Kilpatrick[7], Christophe Fraser[1]

[1] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Medicine, University of Oxford, UK

[2] MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, UK

[3] Liverpool University Hospitals NHS Foundation Trust, UK

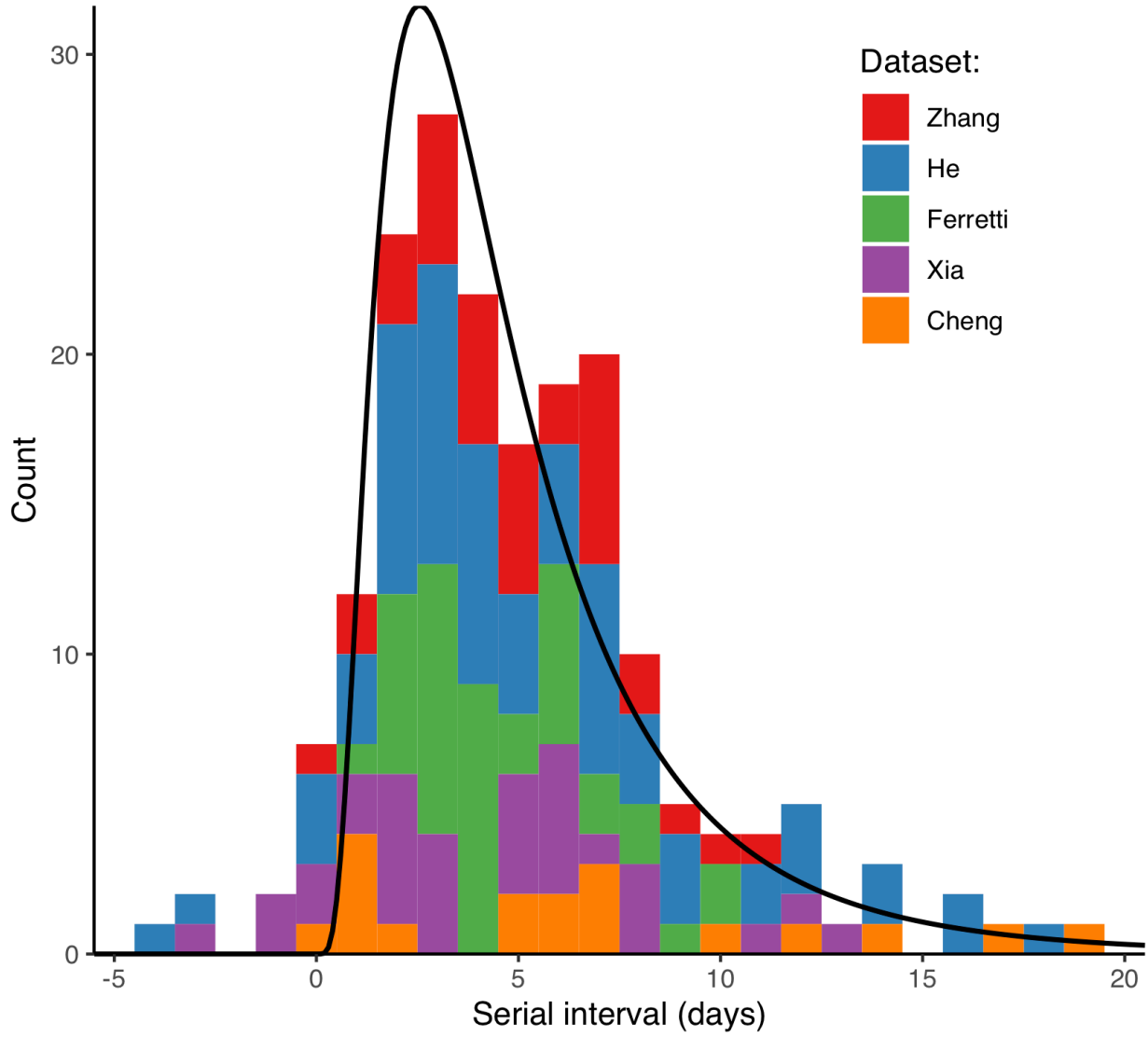
[4] Epidemic Intelligence Center, Taiwan Centers for Disease Control, Taipei, Taiwan

[5] Institute of Epidemiology and Preventive Medicine, National Taiwan University College of Public Health; Global Health Program, National Taiwan University College of Public Health, Taipei, Taiwan

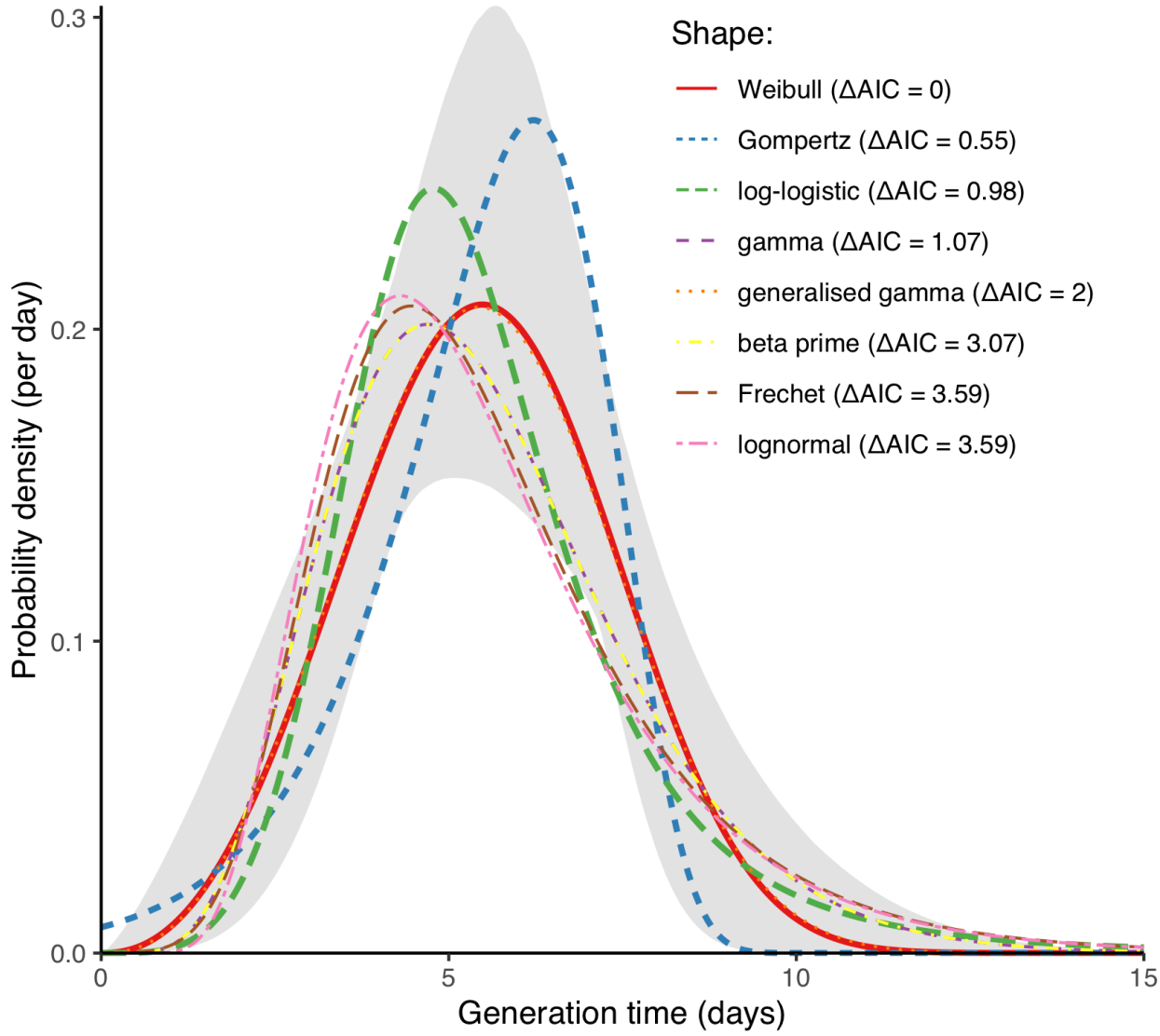
[6] Ecology & Evolutionary Biology, University of Arizona, USA

[7] Ecology & Evolutionary Biology, University of California, Santa Cruz, USA

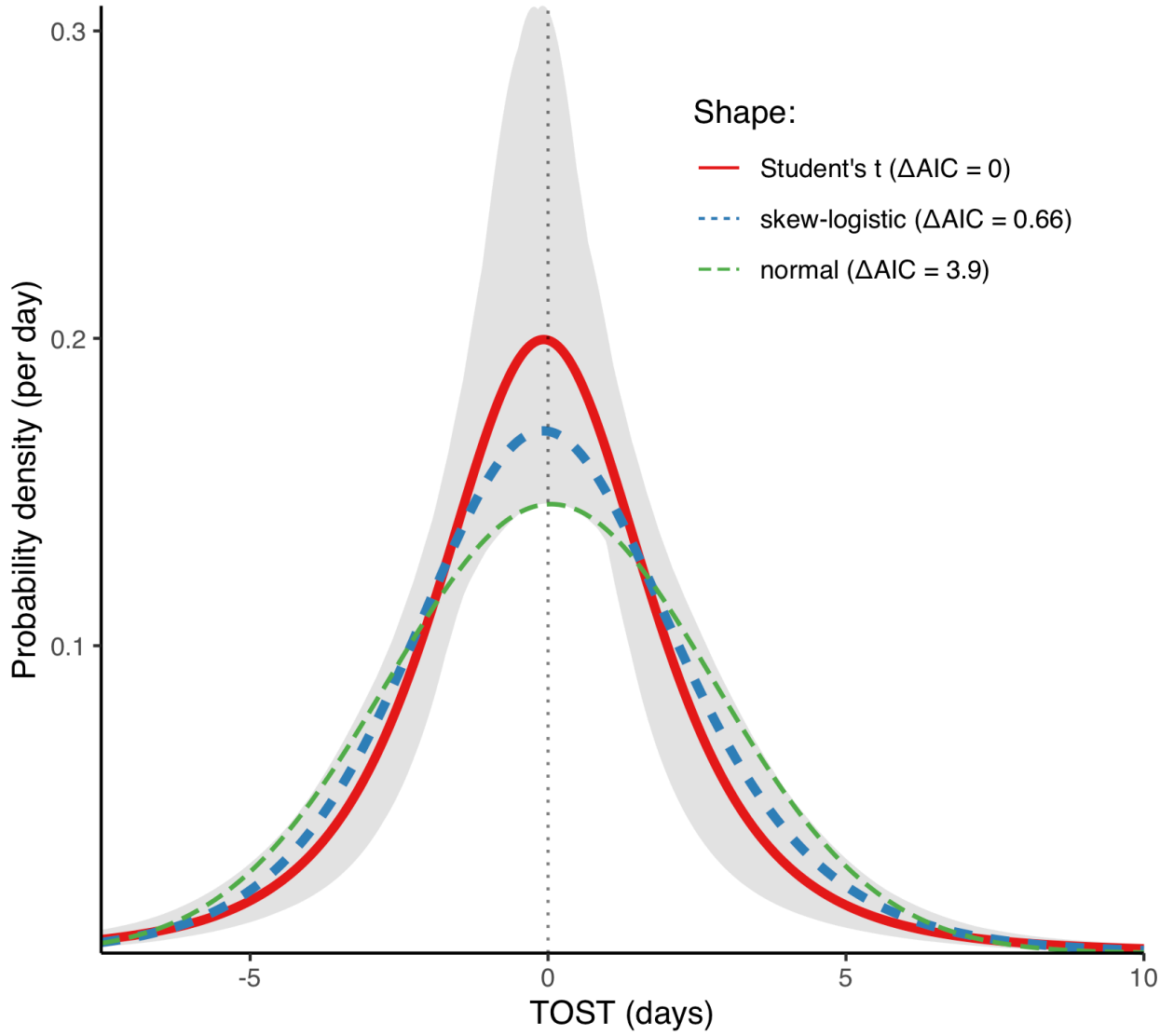
Supplementary Figures



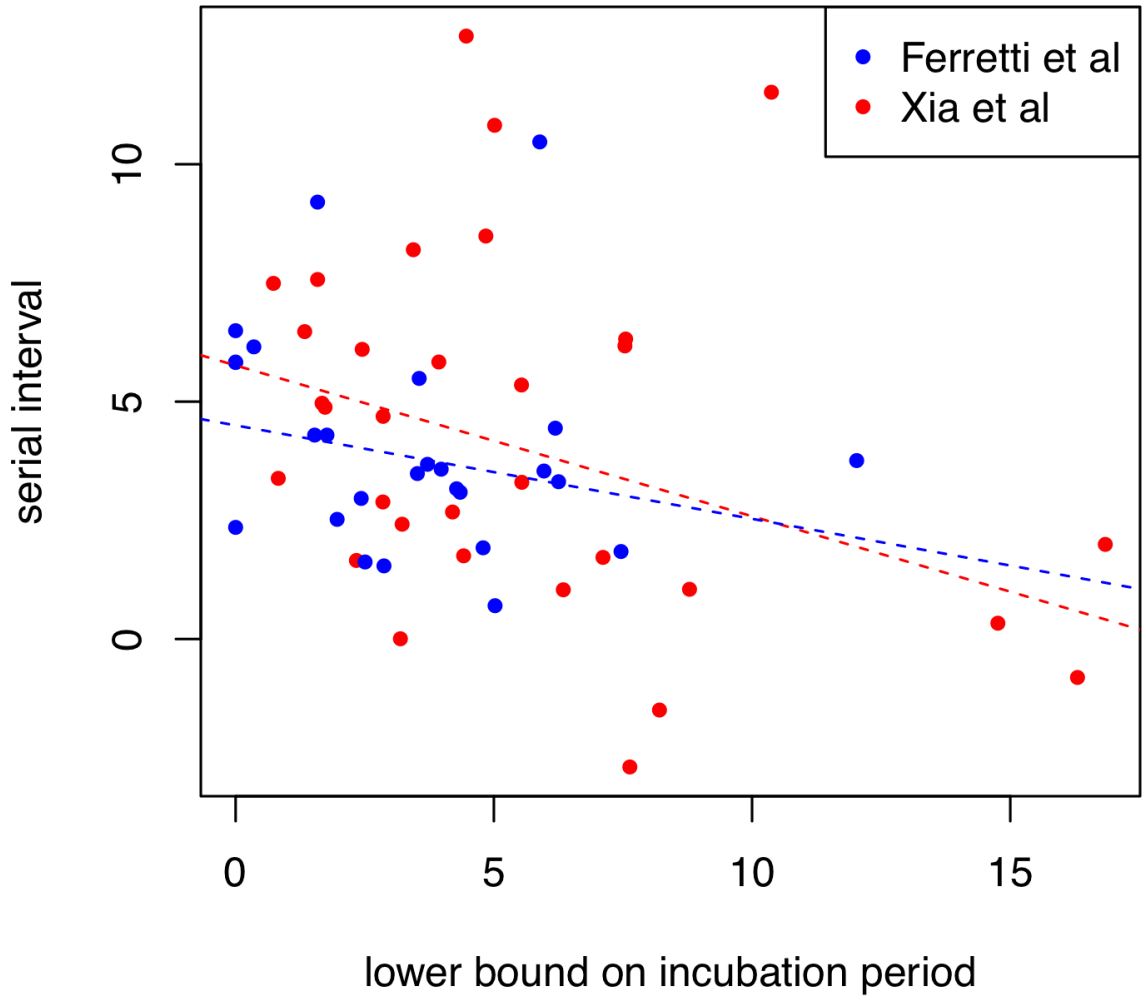
Supplementary Figure 1: *The empirical serial interval distribution from the five datasets shown with different colours. The solid black line shows a lognormal distribution with the same mean=5.1 and SD=3.8 as the empirical distribution, scaled to match the total number of counts binned in the histogram.*



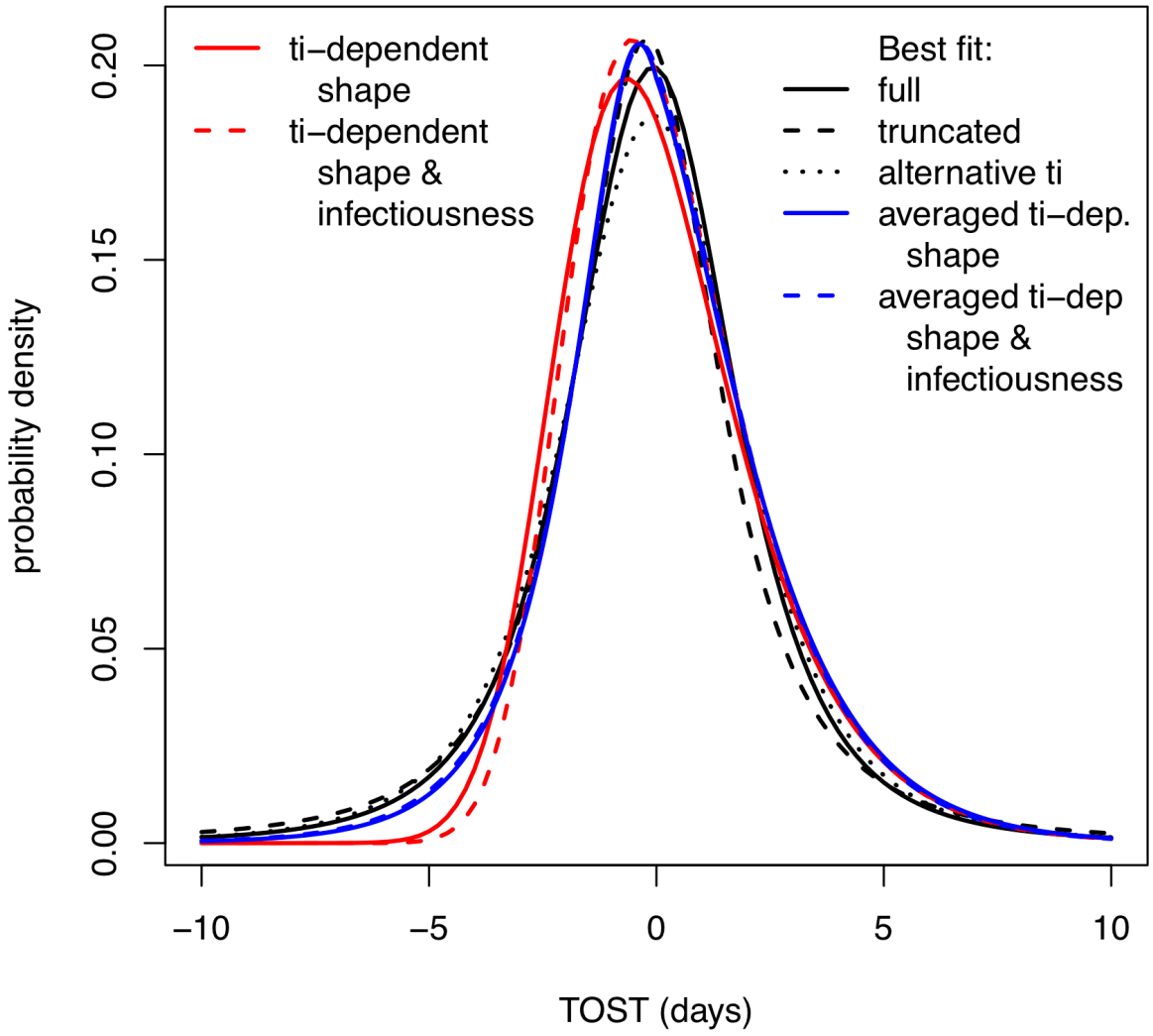
Supplementary Figure 2: *maximum-likelihood generation time distributions for different shapes (functional forms). All distributions were inferred from the full dataset (Ferretti & Wymant et al + Xia et al + He et al + Cheng et al). Shapes with $\Delta AIC < 1$ (Weibull, Gompertz and log-logistic) are shown with thicker lines. Two more shapes were tested but are not shown due to poor fits (see Supplementary Table 1). The pointwise 95% CI for the best fit distribution (Weibull) is shown in grey.*



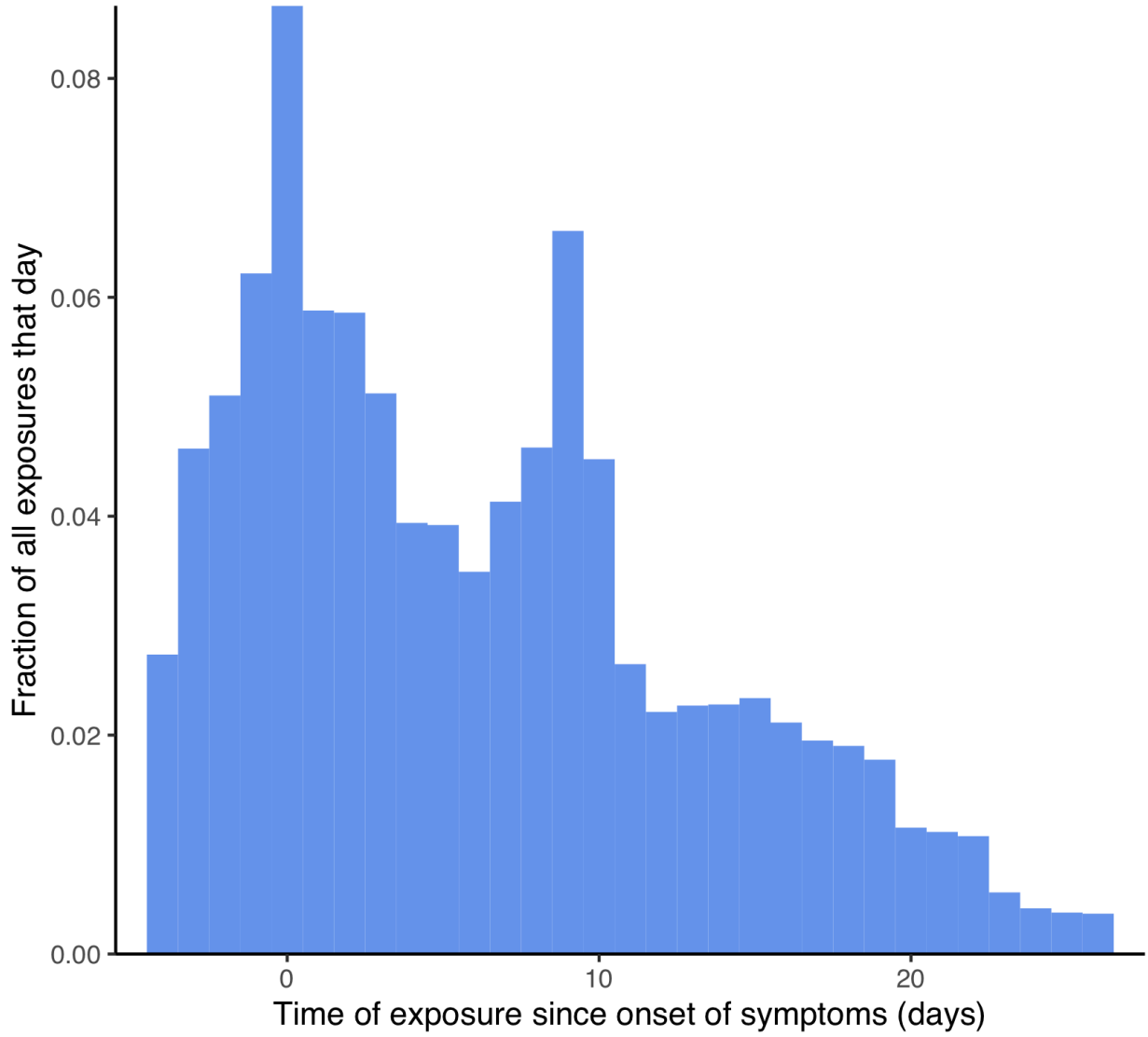
Supplementary Figure 3: *maximum-likelihood distributions for TOST using different shapes (functional forms). All distributions were inferred from the full dataset (Ferretti & Wymant et al + Xia et al + He et al + Cheng et al). Shapes with* $\Delta AIC < 1$ *(Student's* t*, skew-logistic) are shown with thicker lines. Two more shapes were tested but are not shown due to poor fits (see Supplementary Table YYY). The pointwise 95% CI for the best-fitting distribution (Student's* t*) is shown in grey. Note that this distribution should be interpreted as an average over all incubation periods; for a single individual, their infectiousness cannot precede the time they were infected, i.e. infectiousness is truncated where TOST equals minus the incubation period.*



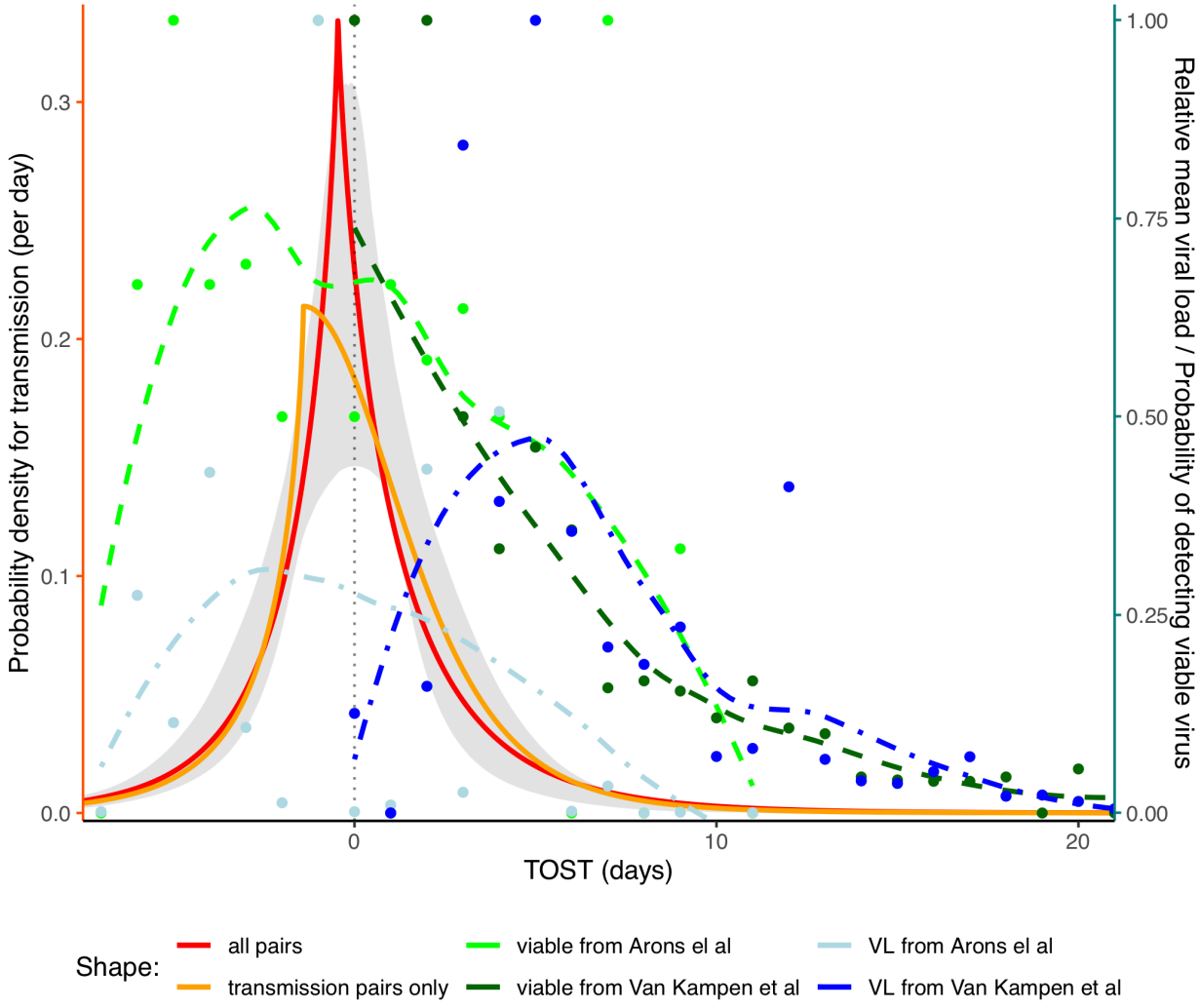
Supplementary Figure 4: *relation between the empirical serial interval and the lower bound on the incubation period of the source, i.e. the time between the last day of exposure (if known) and the day of onset of symptoms. Robust linear regressions (via iterated re-weighted least squares using the MASS package in R) are shown as dashed lines. Pearson's correlation across both datasets r=-0.26 (p=0.044).*



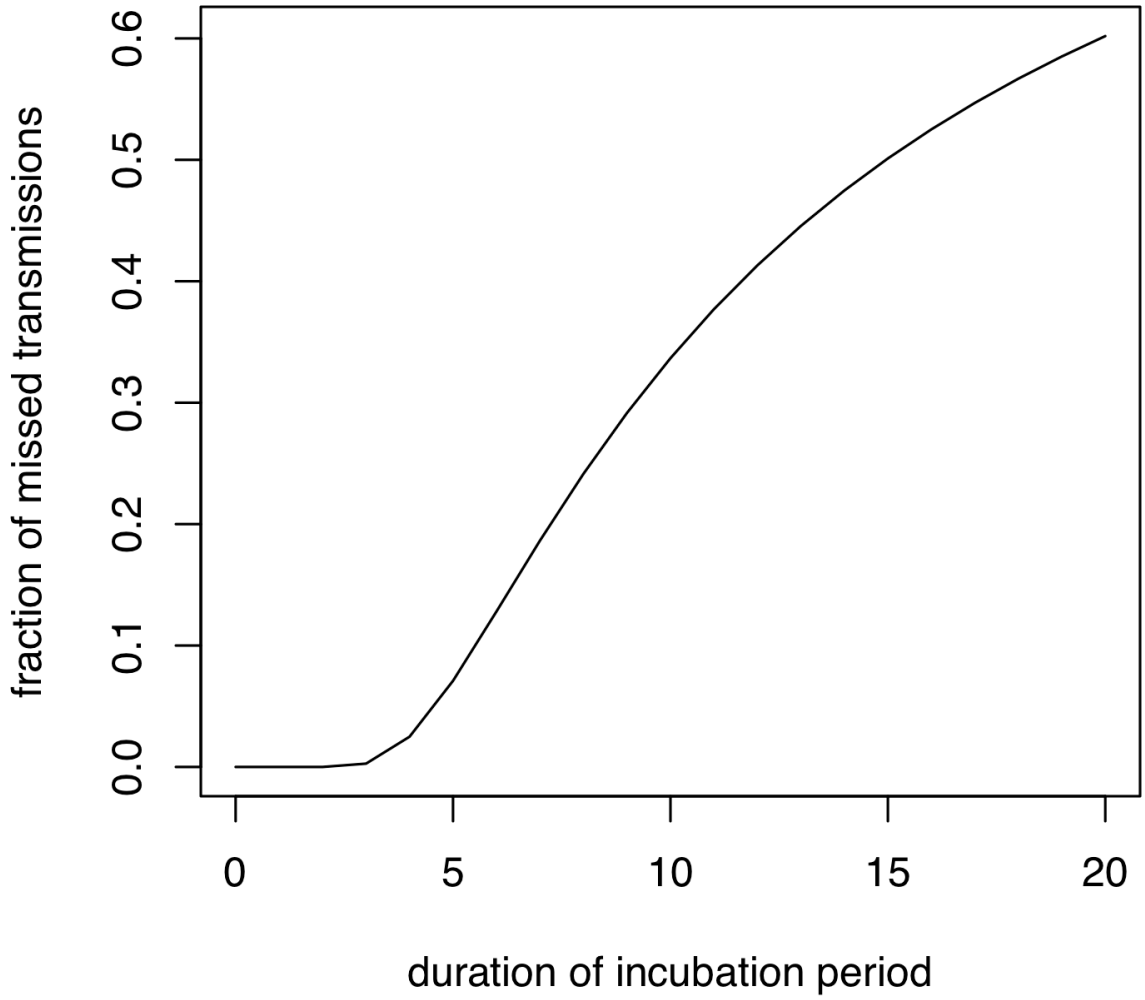
Supplementary Figure 5: *comparison between different distributions of TOST inferred for the full dataset (Ferretti & Wymant et al + Xia et al + He et al + Cheng et al). The "full" distribution is the same as the best-fit curve in Figure 1A and Supplementary Figure 3. The "truncated" distribution corresponds to the best fit when infectiousness is truncated before a time $-t_i$, where t_i is the length of the incubation period. The "alternative" best fit is inferred using the alternative distribution of the incubation period t_i (see Methods). The other curves show the distributions for the incubation period-dependent fits (Figure 3B), assuming that either it is only the shape that depends on the incubation period, or both shape and absolute infectiousness. The curves in red represent the shapes for the mean incubation period t_i=5.42, while the curves in blue represent the distribution of TOST averaged over all possible incubation periods.*



Supplementary Figure 6: *distribution of dates of exposure in relation to onset of symptoms for the contact pairs in Cheng et al.*



Supplementary Figure 7: *comparison between timing of transmission and infectiousness. Continuous lines: best fit of infectiousness profile based on either transmission pairs only as before, or both transmission pairs and case-contact pairs with no transmission. The grey area illustrates the pointwise 95% CI for the distribution of TOST previously inferred from transmission pairs only (Figure 1 and Supplementary Figure 3). Dashed lines: probability of detecting viable virus for a given time from onset of symptoms (curve from LOESS of the data points shown in the same color). Dashed-dotted lines: mean viral load for a given time from onset of symptoms, relative to the maximum viral load inferred from the same dataset (curve from LOESS of the data points shown in the same color).*



Supplementary Figure 8: *fraction of transmissions occurring more than 2 days before onset of symptoms, and that would therefore be missed by tracing contacts from 2 days before onset of symptoms.*

Supplementary Tables

AIC:	*Weibull*	*Gompertz*	*log-logistic*	*gamma*	*generalised*	*Beta'*	*Frechet*	*lognormal*	*inverse*	*Levy*

					gamma				*gamma*	
Ferretti	129	126.8	132.4	131.9	129.4	133.9	135.1	134.7	138.8	183.7
Xia	259.5	259.6	258.9	259.1	260.9	260.6	260.4	258.9	258.6	298.9
Ferretti+Xia	389.9	387.6	392.1	392.8	390.8	394.8	395.4	395.6	400.4	481.6
Ferretti+Xia +He+Cheng	1045.9	1046.5	1046.9	1047	1047.9	1049	1049.5	1049.5	1055	1215.1

Supplementary Table 1: *values of AIC for the ML fit of the generation time distribution for different datasets and functional forms. The best value for each dataset is shown in red.*

Shape	**AIC**	**Parameters**	**R example implementation**
Weibull	1045.94	shape = 3.2862, scale = 6.1244	`dweibull(x, shape, scale)`
Gompertz	1046.49	a = 0.0083, b = 0.7175	`extraDistr::dgompertz(x, a, b)`
log-logistic	1046.91	a = 5.2024, b = 4.8918	`z <- (x / a)^b; return((x > 0) * z * b / (x + 1e-10) / (1 + z)^2)`
gamma	1047	shape = 6.8004, rate = 1.2344	`dgamma(x, shape, rate)`
generalised gamma	1047.93	mu = 1.8064, sigma = 0.3098, Q = 0.9583	`(x > 0)*flexsurv::dgengamma(x, mu, sigma, Q)`
beta'	1049.01	shape1 = 6.7959, shape2 = 14347.2705, scale = 11628.6973	`extraDistr::dbetapr(x, shape1, shape2, scale)`
Frechet	1049.53	lambda = 46.9561, mu = -78.7973, sigma = 83.2902	`extraDistr::dfrechet(x, lambda, mu, sigma)`
lognormal	1049.53	meanlog = 1.6242, sdlog = 0.405	`dlnorm(x, meanlog, sdlog)`

inverse gamma	1055.01	shape = 6.1609, rate = 28.3546	`(x > 0) * dgamma(1 / x, shape, rate) / (x + 1e-10)^2`
Levy	1215.15	s = 2.291	`rmutil::dlevy(pmax(x, 0 + 1e-3), m = 0, s) * (x > 0)`

Supplementary Table 2: *different shapes (functional forms) considered for the generation time distribution, their AIC values, their maximum-likelihood parameter values, and the R code we used for implementing the distribution (to clarify any ambiguity in parameter definition). Shapes are ordered by AIC values for the full dataset: the best fits are at the top.*

AIC:	*Student's t*	*skew-logistic*	*normal*	*skew-normal*	*Cauchy*
Ferretti	114.9	119.7	121.4	120.4	112.9
Xia	249.1	248.8	247.6	249.1	248.2
Ferretti+Xia	366.3	368.3	368.8	369.9	365.2
Ferretti+Xia+He+Cheng	1039.1	1039.7	1043	1044.6	1046.5

Supplementary Table 3: *values of AIC for the ML fit of the TOST distribution for different datasets and functional forms. The best value for each dataset is shown in red.*

Shape	**AIC**	**Parameters**	**R example implementation**
Student's t	1039.07	shift = -0.0747, scale = 1.8567, df = 3.3454	`dt((x - shift) / scale, df) / scale`
skew-logistic	1039.72	mu = -0.2657, sd = 1.5235, a = 1.1248	`a * exp(-(x - mu) / sd) / (1 + exp(-(x - mu) / sd))^(a + 1) / sd`
normal	1042.97	mean = 0.0486, sd = 2.7315	`dnorm(x, mean, sd)`
skew-normal	1044.61	xi = 1.7442, omega = 3.1957, alpha = -0.8949	`dsn(x, xi, omega, alpha)`
Cauchy	1046.52	location = -0.2438, scale = 0.8294	`dcauchy(x, location, scale)`

Supplementary Table 4: *different shapes (functional forms) considered for the distribution for TOST, their AIC values, their maximum-likelihood parameter values, and the R code we used for implementing the distribution (to clarify any ambiguity in parameter definition). Shapes are ordered by AIC values for the full dataset: the best fits are at the top.*

AIC:	Rescaling with incubation period				Datasets			
distribution independent of incubation period	*rescaled side*	*with respect to*	*type of scaling*	*normalised for each incubation period*	*Ferretti*	*Xia*	*Ferretti +Xia*	*Ferretti+Xia +He+Cheng*
generation time	-	-	-	-	126.8	258.6	387.6	1045.9
TOST	-	-	-	-	112.9	247.6	365.2	1039.1
TOST*	-	-	-	-	114.5	258.7	378.5	1074.7
TOST**	-	-	-	-	111.4	246.6	361.3	1030.2
joint	both	OS	linear	yes	113.7	237.3	356.0	1021.3
joint	left/right	loc.par.	linear	yes	113.6	240.2	358.1	1022.5
joint	left/right	mixed	nonlinear	yes	115.4	241.6	360.0	1025.3
joint	left only	OS	linear	yes	111.5	237.3	351.2	1014.7
joint	left only	loc.par.	linear	yes	111.4	238.8	352.0	1017.5
joint	left only	mixed	linear	yes	113.4	239.3	353.2	1016.7
joint	left only	OS	nonlinear	yes	115.3	239.9	353.2	1018.6
joint	left only	OS	linear	no	114.1	212.8	340.0	1005.0
joint	left only	loc.par.	linear	no	114.7	218.4	342.1	1007.8
joint	left only	mixed	linear	no	116.1	207.2	342.0	1007.0
joint	left only	OS	nonlinear	no	115.4	191.6	341.7	1009.0

* using alternative incubation period distribution

** truncated before time of infection

Supplementary Table 5: *best AIC values for different models and datasets. The models are described in Supplementary Methods. 'OS' = onset of symptoms; 'joint' = joint distribution of TOST and incubation period; 'loc.par.' = location parameter of the distribution; 'mixed' = rescaling with respect to a parameter interpolating between the OST and the location parameter of the distribution. The model corresponding to the fit considered in Figure 3 and in the analyses of pre-symptomatic and early symptomatic transmission is outlined in red.*

Correlations:	*Ferretti*	*Xia*	*Ferretti+Xia*	*Ferretti+Xia+He+Cheng*
incubation period vs generation time	0.97	0.98	0.96	0.69
incubation period vs TOST	-0.15	-0.64	-0.26	-0.38
generation time vs TOST	0.09	-0.48	0.01	0.40

Supplementary Table 6: *correlations between incubation period, TOST and generation time, inferred from linear rescaling of pre-symptomatic TOST with respect to incubation period and for different datasets.*

Supplementary Methods

Modelling dependence between TOST distribution and incubation period

Even if TOST is the main determinant of the peak of infectiousness, the width of the distribution could still depend on the incubation period. To obtain a joint distribution which would model this, we consider a distribution $w_{tost}(t|\mu)$ of TOST, where μ is the location parameter of the distribution (i.e. the parameter that controls the shift from a centered distribution) and we rescale it in several possible ways.

The most general form for rescaling on both sides is

$$w(t|t_i) \propto w_{tost}(c\mu + (t - c\mu)/s|\mu)\,, \qquad s = (k + t_i/5.42)/(k+1)$$

where $c$is a parameter that interpolates between 0 (corresponding to a rescaling with respect to onset of symptoms) and 1 (corresponding to a rescaling with respect to location parameter), and k is a positive parameter that controls the nonlinearity of the rescaling ($k = 0$ corresponds to a linear scaling).

The most general form for rescaling on the left side is

$$w(t|t_i) \propto [1 - \theta(t - c\mu)]w_{tost}(c\mu + (t - c\mu)/s|\mu) + \theta(t - c\mu)w_{tost}(t|\mu)\,,$$
$$s = (k + (t_i/5.42)^{\gamma})/(k+1)$$

where$\gamma$is another positive parameter that controls the exponent for the nonlinearity of the rescaling ($\gamma = 1, k = 0$ corresponds to a linear scaling).

Transmission versus infectiousness for symptomatic individuals

The right tail for the distribution of transmissions is subject to several epidemiological biases, since both interventions and self-isolation reduce it. The distribution of exposure events in relation to symptom onset is illustrated in Supplementary Figure 6 and can correct some of the abovementioned epidemiological biases. The distribution of exposure events of traced contacts is biased towards the likely period of transmission, being affected both by isolation behaviour and by the protocols for contact tracing. However it extends to later times compared to our infectiousness profile and it is informative about potential biases in our analysis.

We therefore included in our analysis the case-contact pairs from Cheng et al with known dates of exposure in which no transmission occurred, adding their contribution to the likelihood as described in the Methods.

We also considered both tails of the distribution separately to allow for a wider range of shapes of infectiousness with possibly marked asymmetry between pre- and post-symptomatic periods. We require the resulting distribution to be continuous. This means that the two curves are centered, rescaled to their value at the peak, combined, then shifted by a single location parameter and normalised by the integral of the resulting curve in order to obtain a probability density.

Finally, before combining the distributions, we included a multiplicative power-law factor $1/(1+t)^{\alpha}$ in the right side of the curve, in order to model an additional decay in the right tail of the distribution, induced by interventions. We consider a different parameter for exponent$\alpha$for each dataset, to allow for different epidemiological biases among datasets.

Hence, if $w_{left}(t), w_{right}(t)$ are the centered distributions, the final model is given by

$$w(t) = \frac{(1-\theta(t-\mu))w_{left}(t-\mu)/w_{left}(0) + \theta(t-\mu)(1+t-\mu)^{-\alpha}w_{right}(t-\mu)/w_{right}(0)}{\int_{-\infty}^{0} dx\, w_{left}(x)/w_{left}(0) + \int_{0}^{\infty} dx\, (1+x)^{-\alpha}w_{right}(x)/w_{right}(0)}$$

The best fit shown in Supplementary Figure 7 has $\alpha \approx 0$ for all datasets. Note that this does not mean that interventions could not be influencing on the timing of symptomatic transmissions, but rather that they influence it in the same way across all datasets.