Serial Interval
Reproduction Number

By:
S.S Hashemi Nazari
Department of Epidemiology, School of Public Health and Safety,
Shahid Beheshti University of Medical Sciences, Tehran, Iran

Stages of Infection and Disease

- During the course of an infection, different stages or time periods can be distinguished regarding the infectivity and manifestations of symptoms in the infected individual.

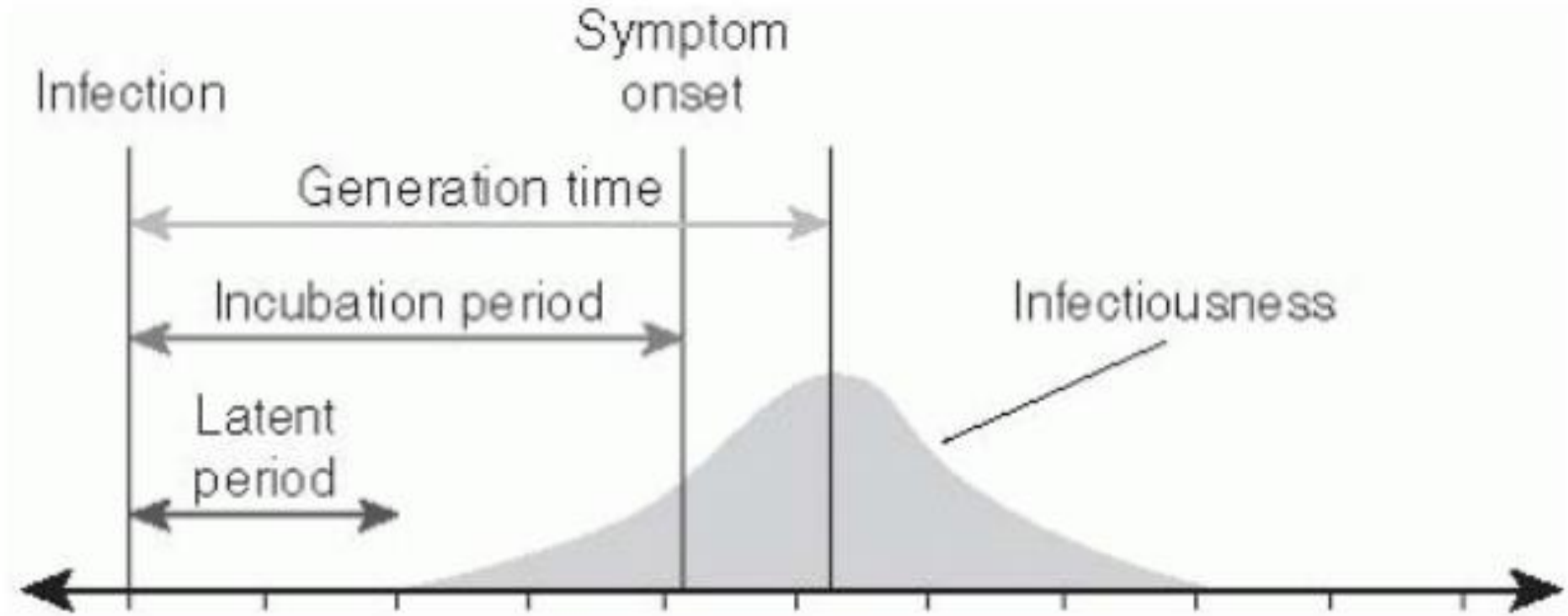
Incubation Period

- The incubation period is the time between infection with a disease and the development of symptoms.
- It plays an important role in the dynamics of disease transmission because it dictates when cases will be detected relative to individuals' time of infection.
- This delay must be accounted for when determining the true infectious burden and evaluating the effect of control measures based on symptomatic surveillance.

- While statements of a mean incubation period or range for the incubation period are common in the infectious disease literature, they are often too vague for use in dynamic models of disease transmission.
- Typically, in dynamic models, incubation periods are assumed to follow some statistical distribution.
- Common choices for this distribution are exponential, log-normal, Weibull, and gamma distributions.

Latent Period

- The latent period is the time between when an individual is infected and when he or she becomes infectious
- While in some cases the latent period and the incubation period have the same length, for many diseases this is not the case.
- Perhaps the most dramatic example of such a discrepancy involves HIV/AIDS, for which the incubation period can last years longer than the latent period.
- Care should be taken when reading the literature, as not all authors are precise in distinguishing the incubation period from the latent period, often referring to the latter as the former.


Infection
Symptom onset
Generation time
Incubation period
Latent period
Infectiousness

- When exploring the dynamics of epidemics, it is usually the latent period, rather than the incubation period, in which we are interested because the latent period has the more profound effect on the generation time, and hence epidemic growth.
- However, the incubation period may also be important, especially when attempting to interpret observed case counts, which may not include all infections that have already occurred.
- The distinction between the incubation period and the latent period is especially important when evaluating the effect of control methods based on symptomatic surveillance.
-

- If infectiousness proceeds symptom onset (i.e., the latent period is shorter than the incubation period), as in HIV, then interventions based on targeted controls toward symptomatic individuals are unlikely to be effective.
- Diseases where symptom onset is coincident with, or even proceeds, infectiousness (e.g., smallpox) are more likely to be controlled using methods based on symptomatic surveillance.
- Hence, the proportion of transmission that takes place before the onset of symptoms may be an important metric of controllability.[26]
- As with the incubation period, the latent period is usually modeled as following an exponential, lognormal, Weibull, or gamma distribution.

Infectious Period

- The infectious period is clearly one of the most important determinants of infectious disease dynamics.
- The infectious period for an infection can range from days (e.g., influenza) to years (e.g., HIV) and plays an important role in determining the reproductive number for that disease.

- The dynamic effects of infectious disease control measures are often best understood by considering their effects on the infectious period.
- The primary effect of treatment, in terms of epidemic dynamics, is to decrease the infectious period. Similarly, case isolation can be seen as decreasing the effective infectious period of those isolated.
- In some circumstances a treatment may increase the infectious period—for instance, supportive care that prevents death but does nothing to prevent transmission.
- The infectious and latent periods play a primary role in determining the generation time of an infection.
- If infectiousness in evenly distributed across the infectious period, then the mean generation time will be equal to the mean latent period plus one-half the mean infectious period.
- In general, infectiousness may not be uniform during the infectious period, though the mean timing of infection may still be determined through more sophisticated techniques

Generation Time

- The generation time (or generation interval) of an infectious disease is the time separating the onset of infection in an individual (the infector) and the time that person transmits to another individual (the infectee).
- While the reproductive number dictates the number of infections produced by each generation of transmission, the generation time specifies a time scale at which these transmissions accumulate.
- The generation time is also referred to as the serial interval, a term coined by Hope-Simpson that is often used to refer to the time between symptom onset in successive generations of infection.[21**]
- The generation time is not fully observable, as the precise moment of infection is difficult, if not impossible, to detect.
- A proxy that is often used to estimate the generation time is the time separating the onset of symptoms in infector and infectee (i.e., the serial interval).

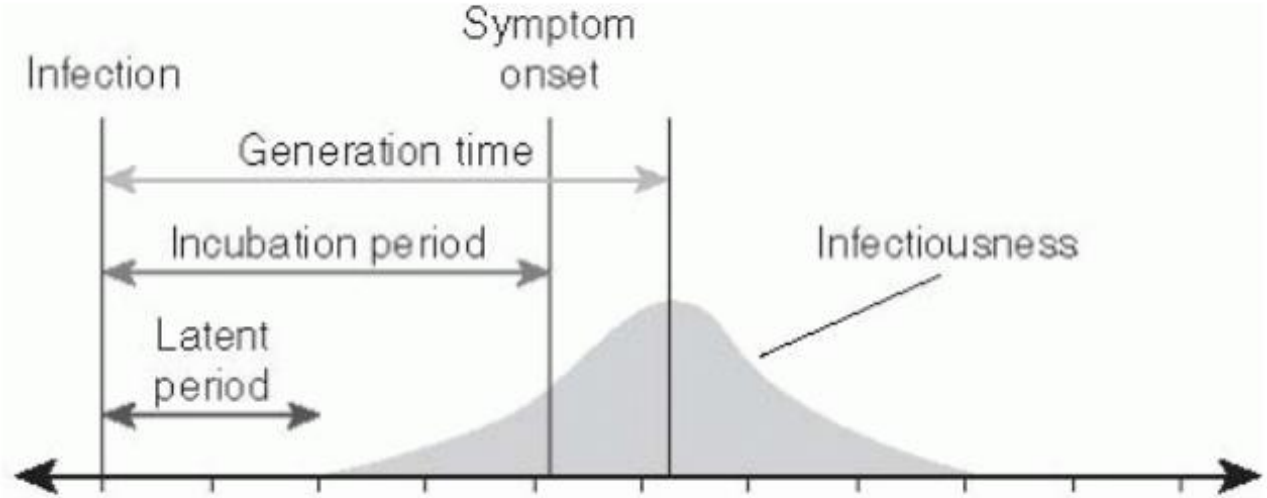
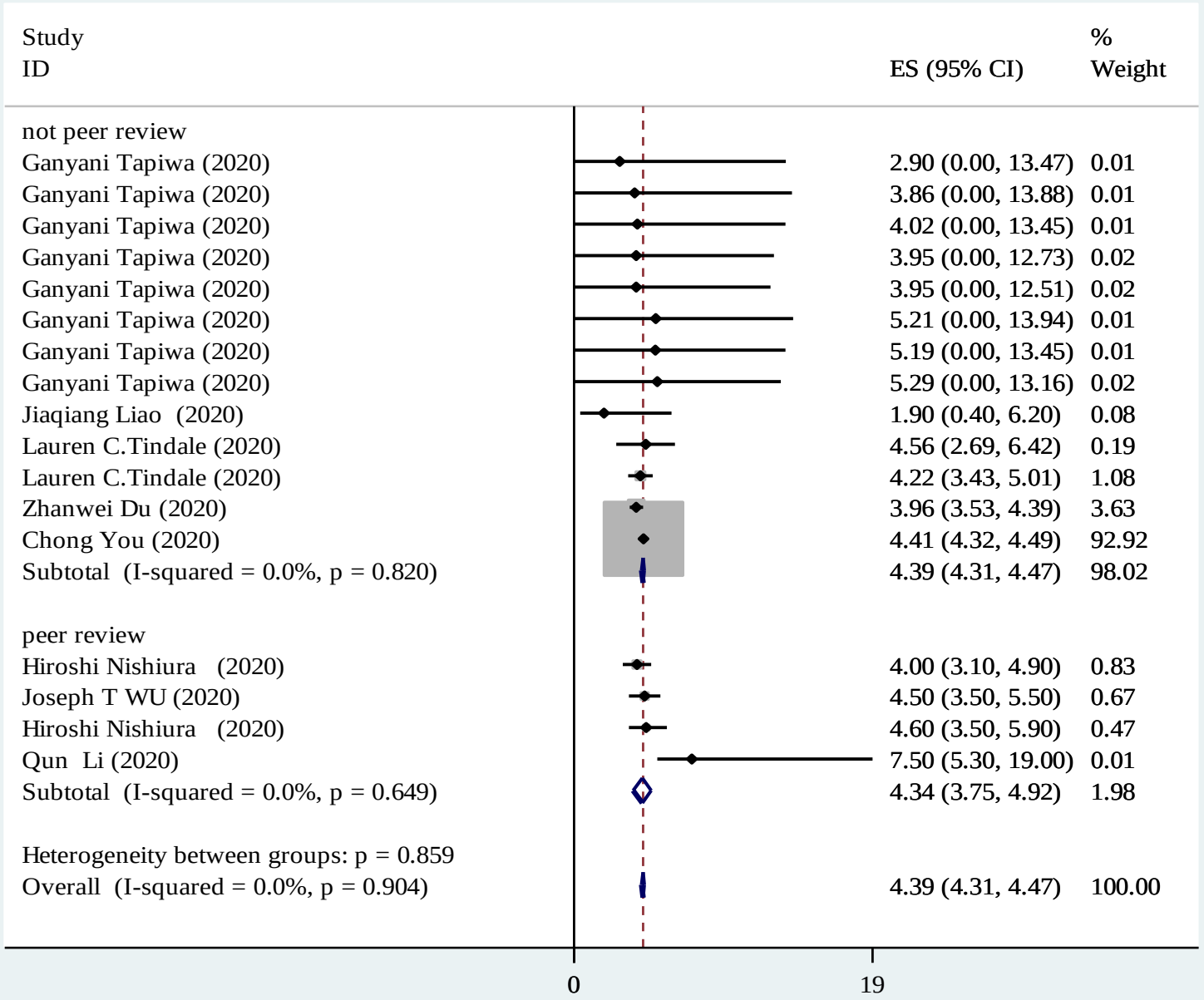


Figure 6-3 Schematic of an infected individual's progression in relation to time of infection. Gray indicates infectiousness over time. Incubation period is defined by the time from infection to symptom onset. The latent period is defined by the time from infection to non-zero infectiousness. The generation time is the mean time from infection to time of infection in those individuals that this person infects. We expect this to be equal to the latent period plus the time from the start of non-zero infectiousness to the probability weighted center of the infectiousness distribution.

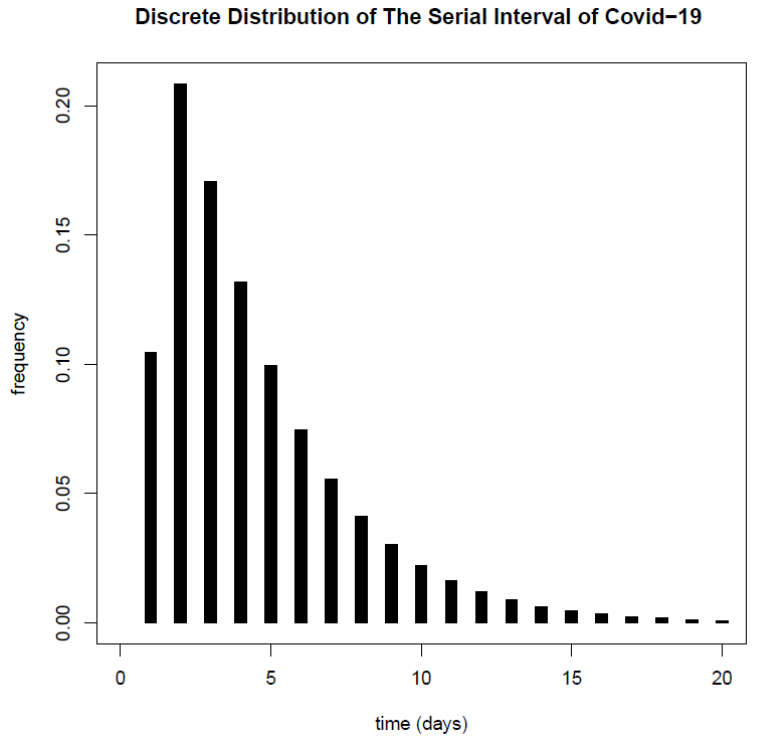
- The serial interval is the duration between symptom onset of a secondary case and that of its primary case.
- For infections in which cases can be infectious before symptom onset, it is possible that the serial interval attains negative values because some of the secondary cases might develop symptoms before their primary case did so.



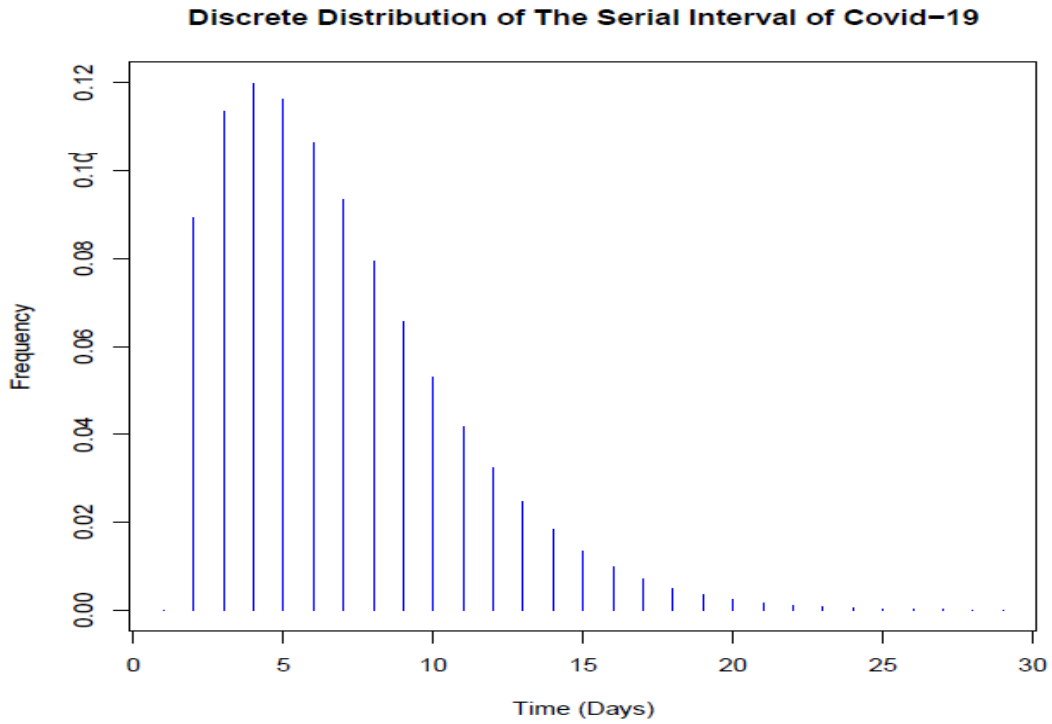
- **Figure 2. The overall serial interval (SI) for COVID-19 by peer review status**

- For estimation of the serial interval, we use data of the confirmed cases of COVID-19 outbreak in the Qom, Iran beginning on February 20, 2020. Information about 51 index cases with laboratory-confirmed COVID-19 and their 318 close contacts was used. Confirmed cases were selected from the first cases of the COVID-19 outbreak in Qom and tried to select with the maximum variety of age, sex, and severity of the disease. Anyone who has been in contact with a confirmed case (less than 2 meters away) during his/her symptomatic period, including 4 days before symptom onset, considered as close contact.

- To determine the serial interval several distributions were fitted on the time interval between primary cases and secondary cases and the best fitting model was a gamma distribution with a mean of 4.55 days and a standard deviation of 3.30 days (Figure 2).



- Close contact between 21 patients (21 infector-infectee pairs), including 12 primary cases and 21 secondary cases were confirmed. The Weibull distribution provides the best fit for the serial interval of the COVID-19 outbreak in Kermanshah. The mean (μ) and standard deviation (SD) of the SI were estimated 5.71 and 3.89 days, respectively (**Fig 2**).



library(R0)

Find the best-fitting GT distribution for a series of serial interval

est.GT(infector.onset.dates = NULL, infectee.onset.dates = NULL, serial.interval = NULL, request.plot = FALSE, ...)

```
# Data taken from traced cases of H1N1 viruses.
data(H1N1.serial.interval)
est.GT(serial.interval=H1N1.serial.interval)

## Best fitting GT distribution is a gamma distribution with mean = 3.039437 and sd = 1.676551 .
## Discretized Generation Time distribution
## mean: 3.070303 , sd: 1.676531
## [1] 0.0000000000 0.1621208802 0.2704857362 0.2358751176 0.1561845680 0.0888997193 0.0459909903
## 0.0222778094 0.0102848887 0.0045773285 0.0019791984 0.0008360608 0.0003464431 0.0001412594
```

The Reproductive Number

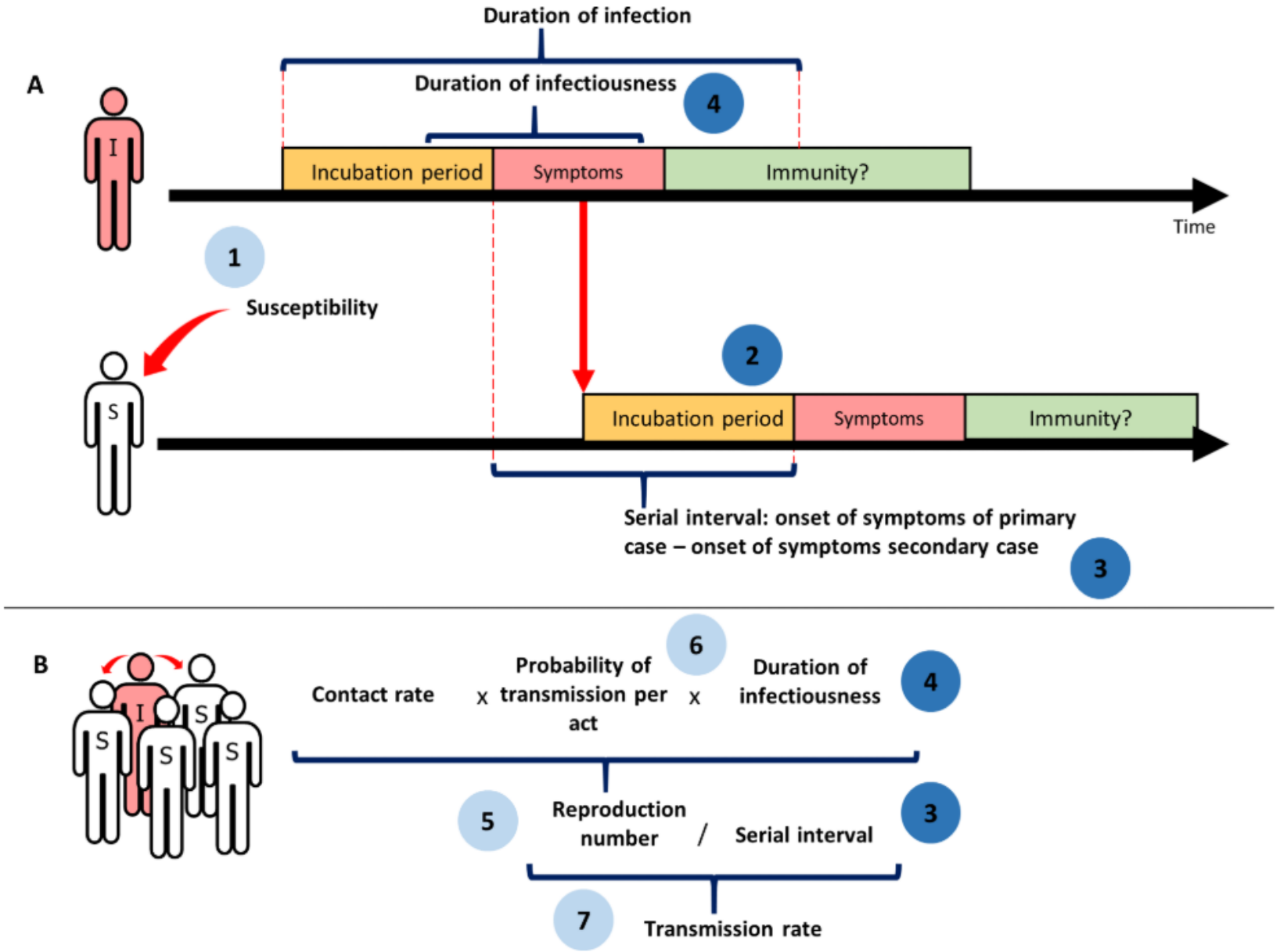
- The reproductive number, *R,* is the number of secondary cases expected to be caused by a single, typical infected individual in a population with some level of susceptibility.
- If the population is fully susceptible, this is termed the basic reproductive number and denoted as R_0
- The reproductive number is the primary metric used to quantify the transmission of a disease in infectious disease dynamics; it provides a measure of how fast an outbreak will grow across subsequent generations of transmission.
- For instance, for influenza $R_0 \approx 2$; hence we would expect a single infected case to cause 2 cases after one generation, 4 cases after two generations, 8 cases after three generations, and so on.

- For measles, where $R_0 \approx 11$, we would expect to see 11 cases in one generation, 121 cases in two generations, and 1331 cases in three generations.
- However, the observed speed of growth does not depend only on R_0, as the average time between generations of transmission varies by disease (generation time) and the number of people available to be infected decreases over time as the pool of susceptible individuals is depleted by infection.

- While R_0 refers to the number of cases caused by a typical individual in a completely susceptible population, *R* refers to the number of cases caused by a typical infectious individual given the proportion of individuals still available to be infected at the current time.
- *R* changes over time and is sometimes denoted R_t, whereas R_0 refers to a theoretical time 0 when the entire population is susceptible.
- *R* accounts for the reduction in transmission due to some individuals being immune .
- For instance, if R_0 is 4, but half of a primary case's contacts are immune due to previous infection or vaccination, then we would expect that case to infect only 2 individuals; in this scenario, *R* would be equal to 2.
- In general, if s_t is the proportion of the population still susceptible to infection at time *t*, then R_t is $R_0 s_t$.

Table 6-1 R_0 for Multiple Pathogens

Pathogen	R_0	Generation Time
Cholera	5.0*,[121] 2.6,[122] 4-15[123]	7.1-9.3 days,[124] 7-10 days[123]
Dengue	1.3-6.3[125]	19-22 days,[126] 24 days[127]
Influenza	1.5-2[57]	3.6 days (s.d. 1.6 days);[128] 2.3 days (H1N1, range 1.5-2.7 days),[30] 3.1 days (H3N2, range 2.2-4.0 days),[30] 2.7 days (H1N1 pdm)[73]
Malaria	1-10 low transmission areas/100-1000 high transmission areas,[123] 1-3000[129]	˜60-120 days,[123] > 200 days[129]


A
Duration of infection
Duration of infectiousness
4
I
Incubation period
Symptoms
Immunity?
Time
1
Susceptibility
S
2
Incubation period
Symptoms
Immunity?
Serial interval: onset of symptoms of primary case – onset of symptoms secondary case
3
B
I
S
S
S
S
6
Contact rate
x
Probability of transmission per act
x
Duration of infectiousness
4
5
Reproduction number
/
Serial interval
3
7
Transmission rate

- The basic reproduction number R0 is defined as the (average) number of new infections generated by one infected individual during the entire infectious period in a fully susceptible population.
- It can be also understood as the average number of infections caused by a typical individual during the early stage of an outbreak when nearly all individuals in the population are susceptible.
- The basic reproduction number reflects the ability of an infection spreading under no control.

- When the size of susceptible population is limited, the quantity, effective reproduction number R_e, is used instead of R_0 . Similarly, the quantity, controlled reproduction number R_c, should be used to describe the ability of disease spreading when interventions (such as quarantine, isolation, or traffic control) are taking place.

- Hence a good measure of any intervention is to reduce R_c . Note that the disease will decline and eventually die out if R_c <1

Table 1. The overall estimation of epidemiologic parameters for COVID-19

Parameters		No. of studies	Estimate	95% CI	P for Heterogeneity	I^2 (%)
basic reproductive number (R_0)	Overall	75	2.72	2.45-2.99	<0.001	99.3
	Korea	7	0.76	0-1.75	<0.001	99.2
	China	56	3.21	2.89-3.52	<0.001	99.1
	Singapore	7	1.15	1.02-1.27	<0.001	86.2
	Iran	2	3.6 [a]	3.1-4.09	0.99	-
	Japan	3	2.35	2.1-2.6	<0.001	80.1
	Peer Review	19	2.02	1.54-2.49	<0.001	99.4
	Not Peer Review	56	2.99	2.66-3.33	<0.001	99.2
Growth Rate (%)	Overall	5	0.38	0.2-0.55	<0.001	97.7
Mortality Rate (%)	Overall	5	5.45	3.477.43	<0.001	95.9
Symptom onset to Hospitalization (day)	Overall	6	5.09	2.15-8.02	0.03	53
Incubation Period (day)	Overall	22	4.24	3.03-5.44	0.02	35
	Peer Review	18	4.03	2.72-5.33	0.01	41
	Not Peer Review	4	5.82[a]	2.91-8.74	0.76	16

[a] Fixed effect model

- Library(R0)
- Library(earlyR)
- Library(epiestim)

- Estimation of reproduction numbers for disease outbreak, based on incidence data.
- The R0 package implements several documented methods.
- Depending on the methods requested by user, basic reproduction number (commonly denoted as R0) or real-time reproduction number (referred to as R(t)) is computed, along with a 95% Confidence Interval.
- Sensitivity analysis tools are also provided, and allow for investigating effects of varying Generation Time distribution or time window on estimates.

Estimate R from exponential growth rate

- est.R0.EG
- Estimate R from exponential growth rate.

- est.R0.EG(epid, GT, t = NULL, begin = NULL, end = NULL, date.first.obs = NULL, time.step = 1, reg.met = "poisson", checked = FALSE, ...)
- data(Germany.1918) mGT<-generation.time("gamma", c(3, 1.5))
- est.R0.EG(Germany.1918, mGT, begin=1, end=27)
- ## Reproduction number estimate using Exponential Growth ## R : 1.525895[1.494984 , 1.557779]

Estimate the reproduction number by maximum likelihood

- est.R0.ML

- Two maximum likelihood methods for estimatig the reproduction ratio.
- The first (and used by default in this package) assumes that the serial interval distirbution is known, and subsequently the likelihood is only maximised depending on the value of R.
- The second method can be used if the serial interval distribution is unknown: in that case, the generation time is set to follow a Gamma distribution with two parameters (size, shape), and the optimization routine finds the values of R, size and shape that maximize the likelihood

- However, the epidemic curve must be long enough to account for a whole generation. The authors showed that this is achieved when the cumulated amount of incident cases reaches 150.
- When using this method, the flag unknown.GT must be set to TRUE. GT must still be provided with a R0.GT-class object, however its mean and sd will be recycled as starting value for the optimization routine.
- data(Germany.1918)
- mGT<-generation.time("gamma", c(2.45, 1.38)) est.R0.ML(Germany.1918, mGT, begin=1, end=27, range=c(0.01,50))
- # Reproduction number estimate using Maximum Likelihood method.
- # R : 1.307222[1.236913 , 1.380156]

Estimate the time dependent reproduction number using a Bayesian approach

- Estimate the time dependent reproduction number using a Bayesian approach.
- All known data are used as a prior for next iteration

- data(Germany.1918)
- mGT <- generation.time("gamma", c(3,1.5))
- SB <- est.R0.SB(Germany.1918, mGT)
- ## Results will include "most likely R(t)" (ie. the R(t) value for which the computed probability is the highest), along with 95% CI, in a data.frame object
- SB # Reproduction number estimate using Real Time Bayesian method.
- # 0 0 2.02 0.71 1.17 1.7 1.36 1.53 1.28 1.43 ...

Estimate the time dependent reproduction number

- est.R0.TD
- Estimate the time dependent reproduction number according to Wallinga & Teunis.

- data(Germany.1918)
- mGT<-generation.time("gamma", c(3, 1.5))
- TD <- est.R0.TD(Germany.1918, mGT, begin=1, end=126, nsim=100)
- # Warning messages:
- # 1: In est.R0.TD(Germany.1918, mGT) : Simulations may take several minutes.
- # 2: In est.R0.TD(Germany.1918, mGT) : Using initial incidence as initial number of cases.
- TD
- # Reproduction number estimate using Time-Dependent method.
- # 2.322239 2.272013 1.998474 1.843703 2.019297 1.867488 1.644993 1.553265 1.553317 1.601317 ...

Library(epiestim)

- We can run estimate_R on the incidence data to estimate the reproduction number R.
- For this, we need to specify
- i) the time window(s) over which to estimate R and
- ii) information on the distribution of the serial interval.

- For i), the default behavior is to estimate R over weekly sliding windows.

- For ii), there are several options, specified in the method argument.

- The simplest is the parametric_si method, where you only specify the mean and standard deviation of the SI.

Estimating R on sliding weekly windows, with a parametric serial interval

- In this example, we only specify the mean and standard deviation of the serial interval. In that case an offset gamma distribution is used for the serial interval.

- res_parametric_si <- estimate_R(Flu2009$incidence, method="parametric_si",

config = make_config(list(mean_si = 2.6, std_si = 1.5)))

```
#>   t_start t_end  Mean(R)    Std(R) Quantile.0.025(R) Quantile.0.05(R)
#> 1       2     8 1.735798 0.4091314         1.0287437         1.121933
#> 2       3     9 1.749168 0.3647267         1.1088223         1.195480
#> 3       4    10 1.537058 0.3074116         0.9947030         1.068694
#> 4       5    11 1.431839 0.2705921         0.9514466         1.017661
#> 5       6    12 1.422725 0.2515046         0.9731426         1.035808
#> 6       7    13 1.635373 0.2523436         1.1786332         1.243590
#>   Quantile.0.25(R) Median(R) Quantile.0.75(R) Quantile.0.95(R)
#> 1         2.458972  1.703761         2.458972         2.458972
#> 2         2.389121  1.723884         2.389121         2.389121
#> 3         2.075176  1.516613         2.075176         2.075176
#> 4         1.904047  1.414830         1.904047         1.904047
#> 5         1.860107  1.407932         1.860107         1.860107
#> 6         2.071372  1.622413         2.071372         2.071372
#>   Quantile.0.975(R)
#> 1          2.624781
#> 2          2.533119
#> 3          2.195540
#> 4          2.008849
#> 5          1.956336
#> 6          2.165745
```

Estimating R with a non parametric serial interval distributiond

- If one already has a full distribution of the serial interval, and not only a mean and standard deviation:
- res_non_parametric_si <- estimate_R(Flu2009$incidence, method="non_parametric_si", config = make_config(list(si_distr = Flu2009$si_distr))

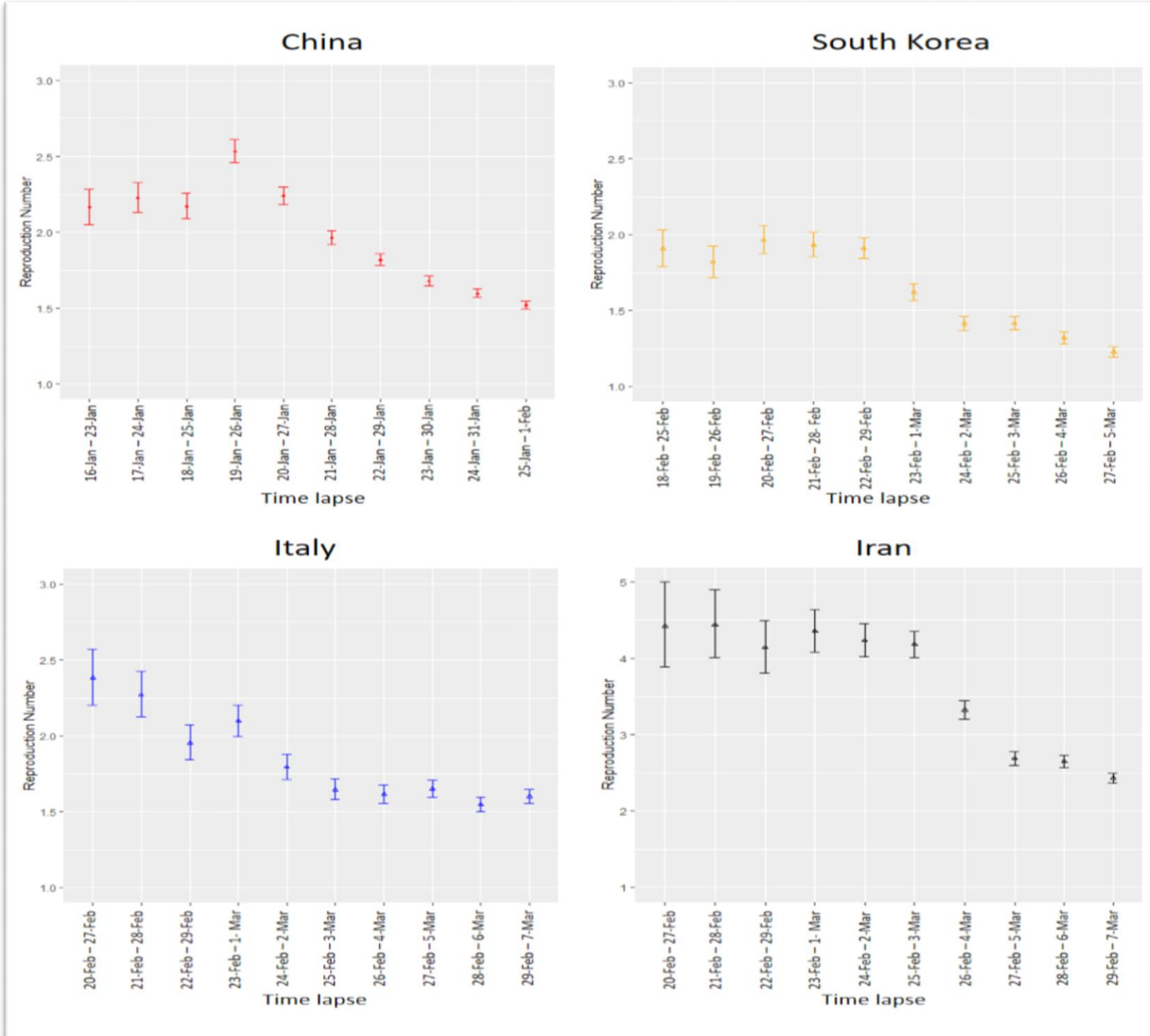
Estimating R accounting for uncertainty on the serial interval distribution

- config <- make_config(list(mean_si = 2.6, std_mean_si = 1, min_mean_si = 1, max_mean_si = 4.2, std_si = 1.5, std_std_si = 0.5, min_std_si = 0.5, max_std_si = 2.5))
- res_uncertain_si <- estimate_R(Flu2009$incidence, method = "uncertain_si", config = config)

Estimating R and the serial interval using data on pairs infector/infected

- In estimate_R, we now allow the serial interval distribution to be directly estimated, using MCMC, from interval censored exposure data. The reproduction number is then estimated using the posterior distribution of the SI, hence accounting for the uncertainty associated with this estimate.
- As the epidemic progresses, newly collected exposure data can be incorporated to update the serial interval estimate.

```
#>   EL ER SL SR type
#> 1  0  1  7  8    0
#> 2  0  1  2  3    0
#> 3  0  1  3  4    0
#> 4  0  1  2  5    0
#> 5  0  1  1  9    0
#> 6  0  1  2  4    0
```


China
Reproduction Number
3.0
2.5
2.0
1.5
1.0
16-Jan – 23-Jan
17-Jan – 24-Jan
18-Jan – 25-Jan
19-Jan – 26-Jan
20-Jan – 27-Jan
21-Jan – 28-Jan
22-Jan – 29-Jan
23-Jan – 30-Jan
24-Jan – 31-Jan
25-Jan – 1-Feb
Time lapse
South Korea
Reproduction Number
3.0
2.5
2.0
1.5
1.0
18-Feb – 25-Feb
19-Feb – 26-Feb
20-Feb – 27-Feb
21-Feb – 28- Feb
22-Feb – 29-Feb
23-Feb – 1-Mar
24-Feb – 2-Mar
25-Feb – 3-Mar
26-Feb – 4-Mar
27-Feb – 5-Mar
Time lapse
Italy
Reproduction Number
3.0
2.5
2.0
1.5
1.0
20-Feb – 27-Feb
21-Feb – 28-Feb
22-Feb – 29-Feb
23-Feb – 1- Mar
24-Feb – 2-Mar
25-Feb – 3-Mar
26-Feb – 4-Mar
27-Feb – 5-Mar
28-Feb – 6-Mar
29-Feb – 7-Mar
Time lapse
Iran
Reproduction Number
5
4
3
2
1
20-Feb – 27-Feb
21-Feb – 28-Feb
22-Feb – 29-Feb
23-Feb – 1- Mar
24-Feb – 2-Mar
25-Feb – 3-Mar
26-Feb – 4-Mar
27-Feb – 5-Mar
28-Feb – 6-Mar
29-Feb – 7-Mar
Time lapse