Physics of Life Data Epidemiology

Lect 3: Infection time scales

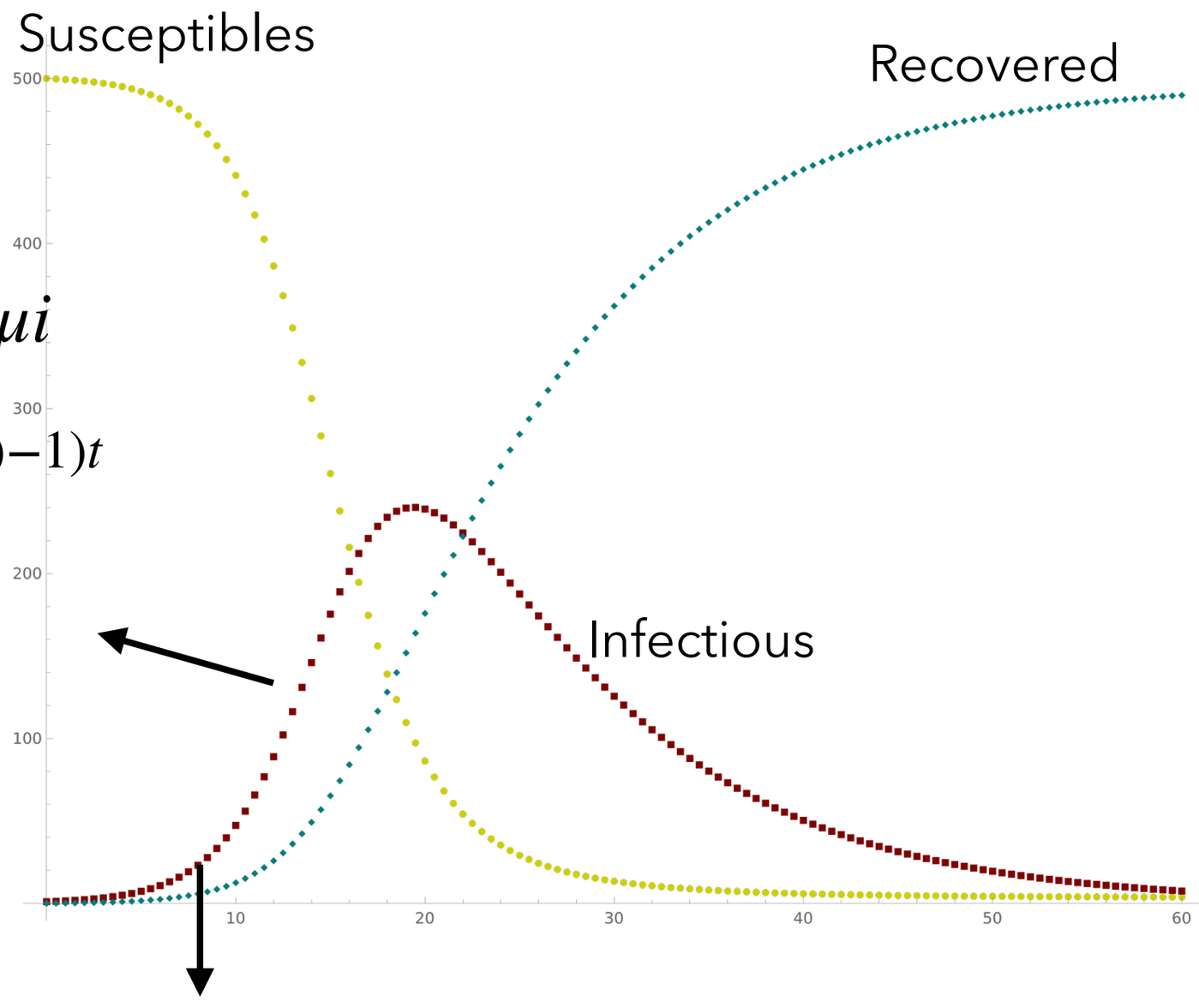
Chiara Poletto
mail: chiara.poletto@unipd.it
web: chiara-poletto.github.io
bsky: @chpoletto.bsky.social

SIR Model

evolution of $s(t)$, $i(t)$ and $r(t)$ for $R_0 > 1$ and initial conditions, $s_0 \simeq 1$, $r_0 = 0$ and $i_0 \ll 1$

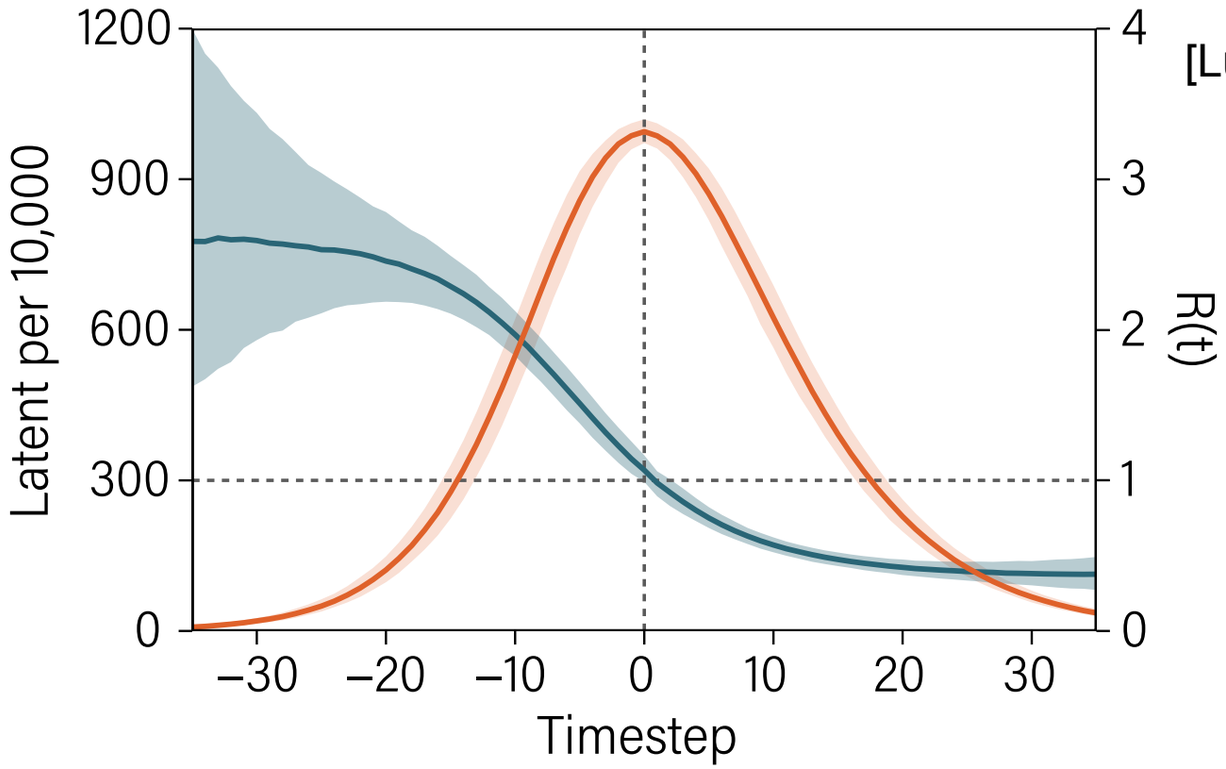
$$\frac{di}{dt} = \beta\langle k\rangle si - \mu i$$

$$i(t) \simeq i_0 e^{\mu(R_0 s(t) - 1)t}$$



$$i(t) \simeq i_0 e^{\mu(R_0 - 1)t}$$

SIR Model

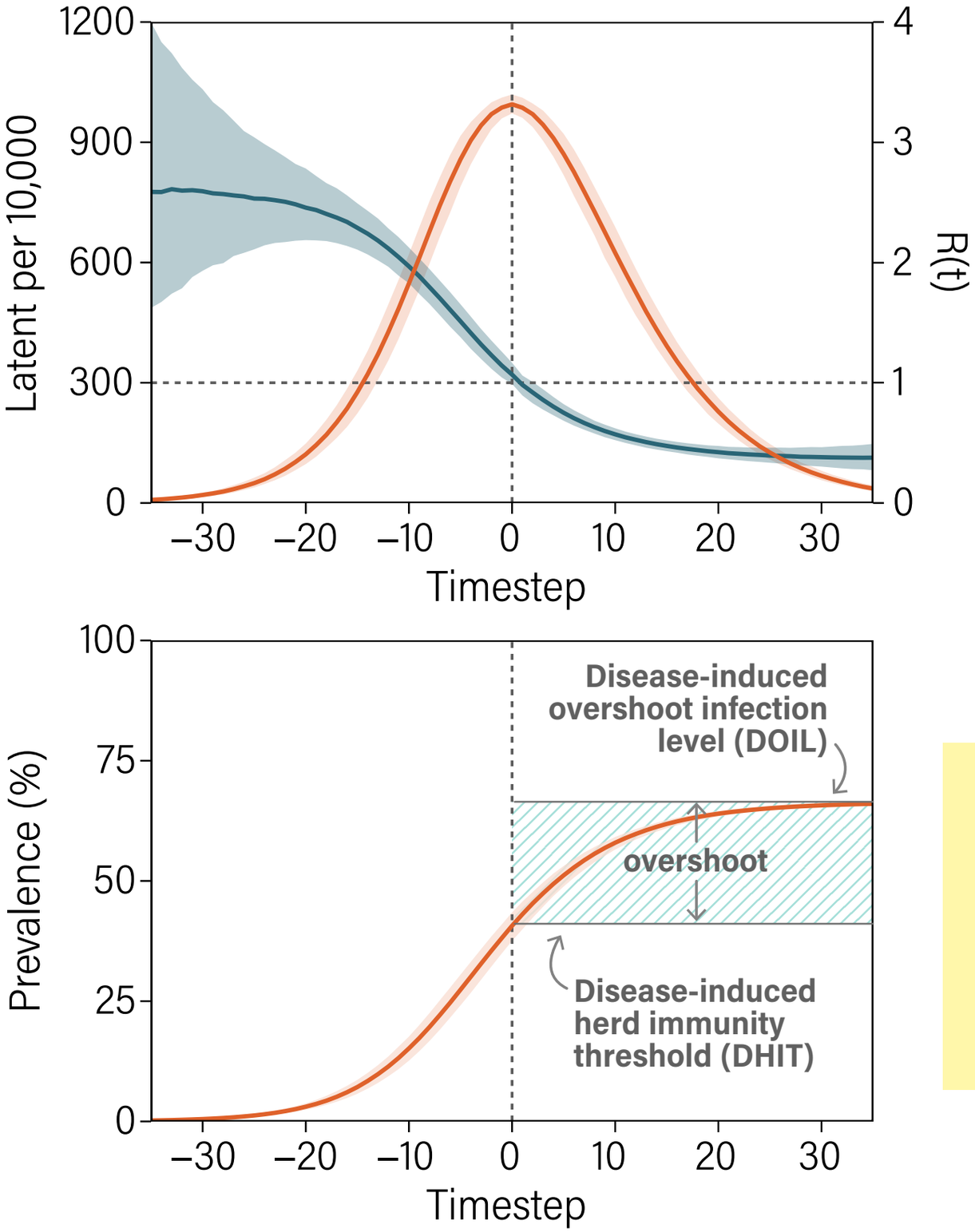


[Lu et al doi.org/10.1101/2021.03.19.21253974]

- Basic reproductive ratio, $R_0 = \beta\langle k\rangle\tau$: # of cases generated by a case in fully susceptible population
- **Reproductive ratio** $R(t)$: # of cases generated by a case in a *partially* susceptible population. If :
  - the virus transmissibility does not change ($\beta = \text{const}$)
  - individual behaviour does not change ($\langle k\rangle_t = \text{const}$)
  - infectious period does not change ($\tau = \text{const}$)

  then $R(t) = R_0\, s(t)$, it decreases due to depletion of susceptible

SIR Model



[Lu et al doi.org/10.1101/2021.03.19.21253974]

herd immunity threshold: the fraction of immune in the population, r_{HIT}, such that

$$R = R_0(1 - r_{\mathrm{HIT}}) = 1$$

- when $r(t) \geq r_{\mathrm{HIT}}$ the epidemic starts to decline
- if $r_0 \geq r_{\mathrm{HIT}}$ and I seed a small number of infectious individuals in the population ($i_0 \ll 1$) the epidemic does not start.

If I have a perfect vaccine and I can distribute it before the start of the epidemic r_{HIT} gives the proportion of the population I need to vaccine to avoid the epidemic

assumptions of SIR Model

$R(t) = R_0\, s(t) = \beta \langle k \rangle \tau s(t)$

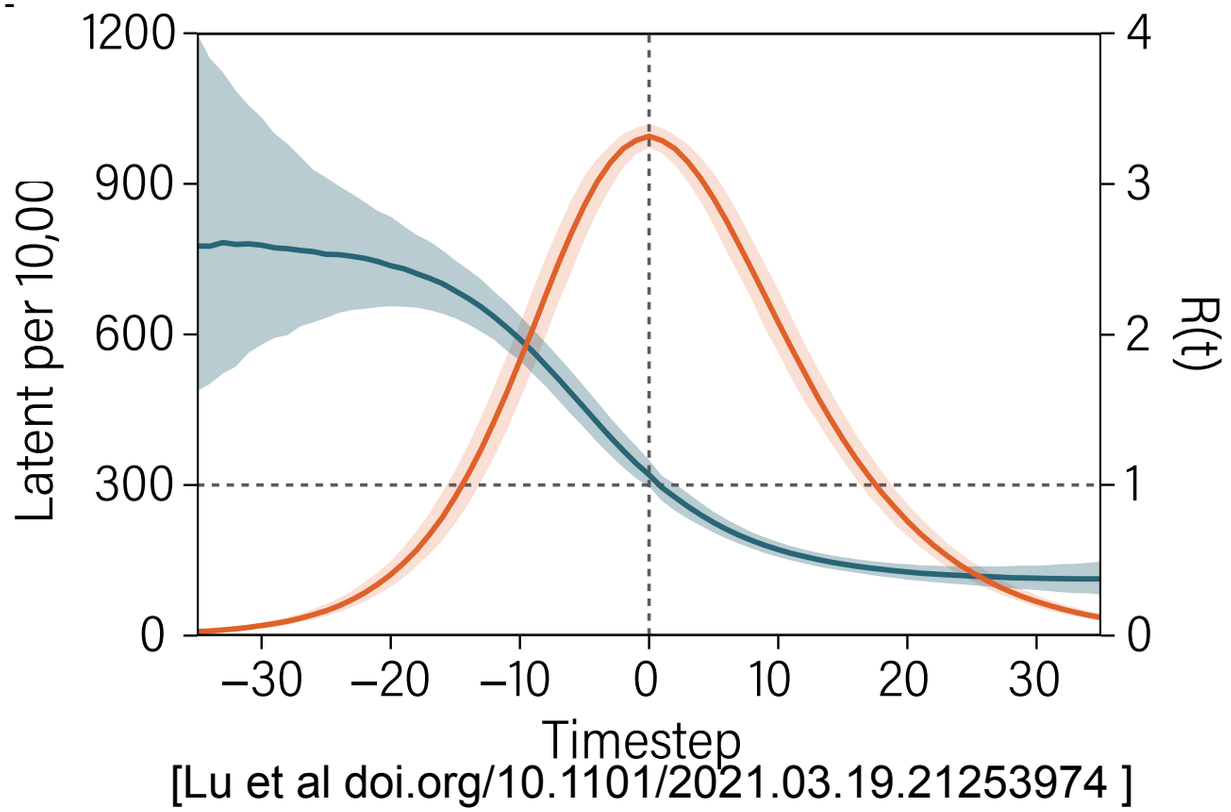
when β = const, $\langle k \rangle_t$ = const, τ = const

However

A priori $R(t) = \beta(t) \langle k \rangle_t \tau(t)\ s(t)$

- $\beta(t)$ depending on environmental variables (humidity, temperature, time spent indoor/outdoor, aeration), interventions/recommendation (masks, 1 meter distance, etc.)
- $\langle k \rangle_t$ depending on people seasonal behaviour (schools open/close), interventions (lockdowns, closure of restaurants, etc.)
- $\tau(t)$ depending on interventions (case isolation, more rapid case identification and hospitalisation, drugs that reduce the infectious period)

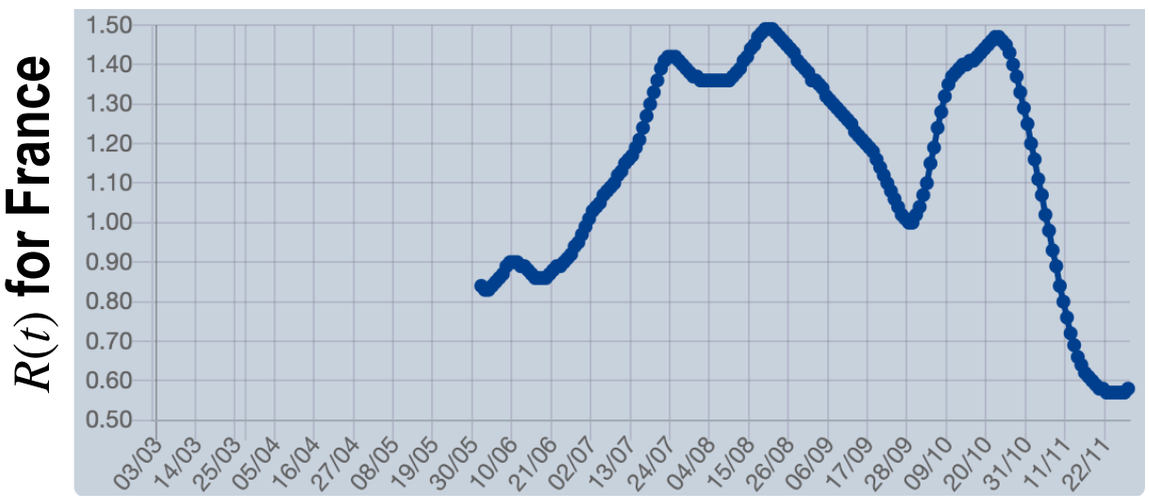
assumptions of SIR Model



[Lu et al doi.org/10.1101/2021.03.19.21253974]

SIR fairly good for, e.g., a localised outbreaks of flu, when $\beta = \text{const}$, $\langle k \rangle_t = \text{const}$, $\tau = \text{const}$

- uncontrolled epidemic
- constant environmental factors

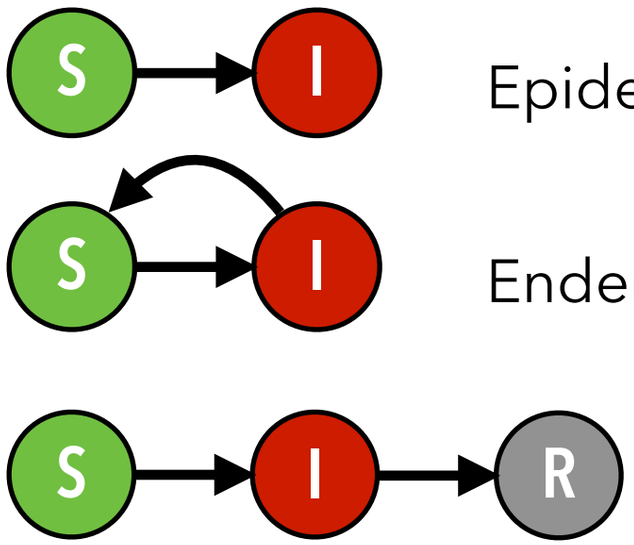


[https://www.gouvernement.fr/info-coronavirus/carte-et-donnees]

COVID-19 spread

- variable mask wearing
- variable social distancing
- variable test-trace-isolate

SI, SIS, SIR models



Epidemic expansion

Endemic circulation

Outbreak

SIR: Widely used in outbreak analysis because it captures all properties of an outbreak: initial exponential growth, peak, extinction before all individuals get infected

key simplifying assumptions:

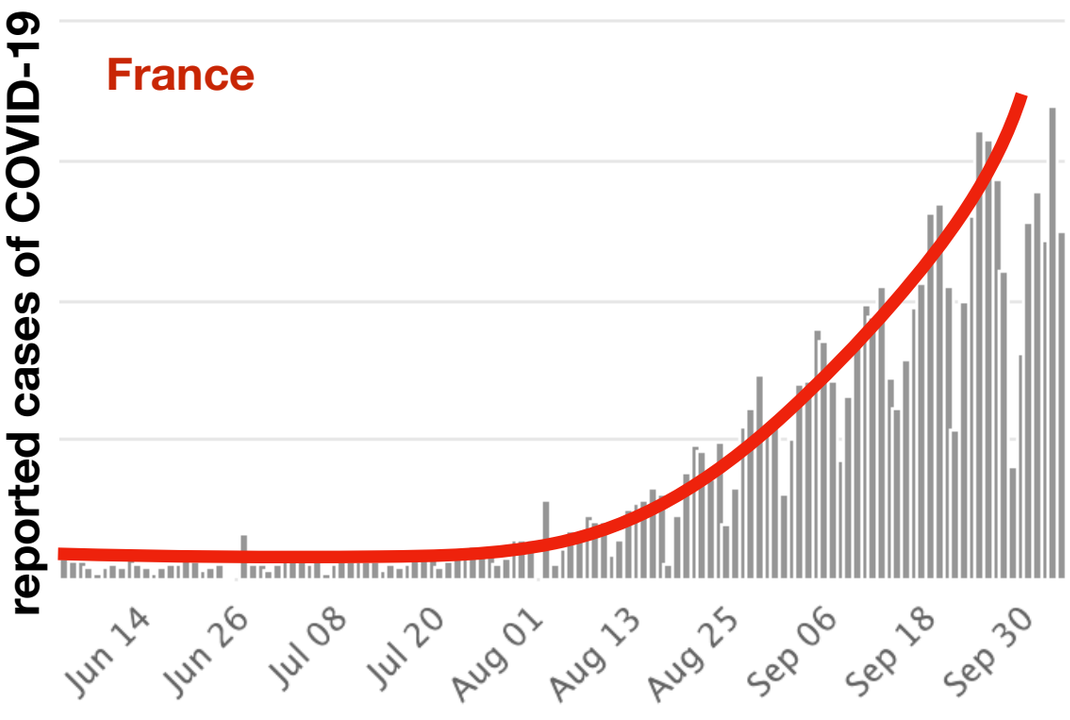
-simplified disease natural history: only one infectious stage is considered; infectivity is constant form infectious to recovery; constant rate of transition from infectious to recovery

-homogenous mixing

role of models

how many cases will we have in 2 weeks? what shall we expect for the next months if we don't do anything? what if I vaccinate? what if I enforce a lockdown?

- final attack rate and strength of the intervention to contain the epidemic is function on R_0
- Initial incidence growth depends on both R_0 and disease natural history



Ex: same epidemic profile can due to:

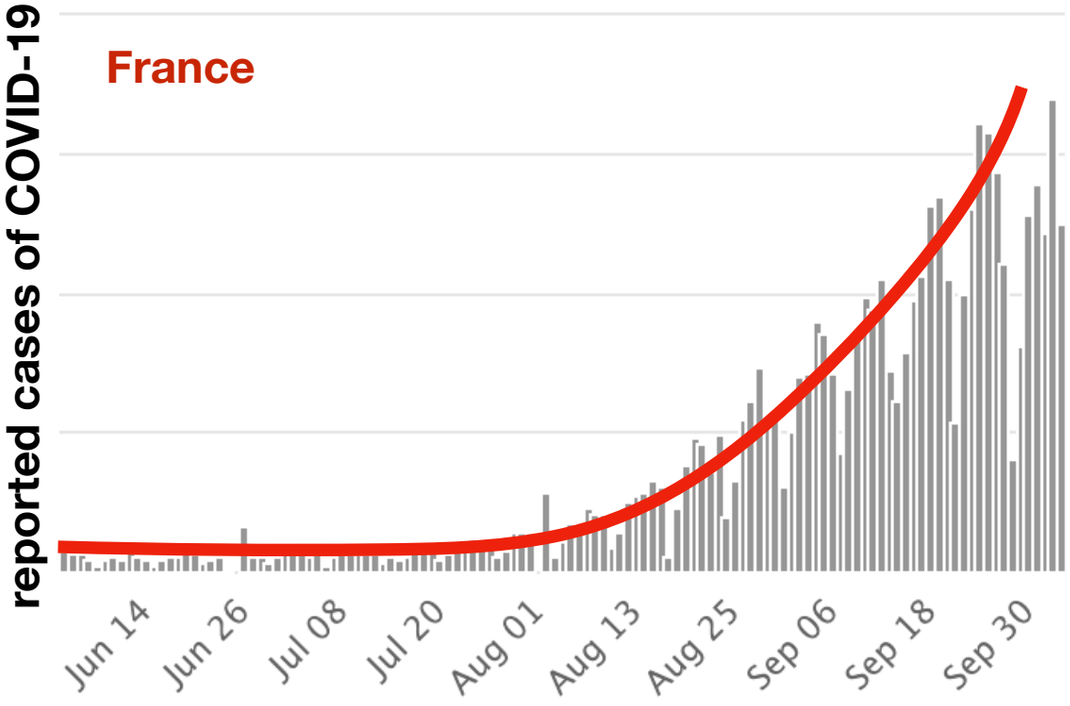
- a slow infection with high R_0
- a fast infection with small R_0

Fit incidence data with the wrong values of infectious or exposed period or the wrong model (**i.e. neglecting the exposed state or not well capturing the infectivity profile**) provide biased estimates of R_0

need to properly capture disease natural history

how many cases will we have in 2 weeks? what shall we expect for the next months if we don't do anything? what if I vaccinate? what if I enforce a lockdown?

- final attack rate and strength of the intervention to contain the epidemic is function on R_0
- Initial incidence growth depends on both R_0 and disease natural history



Ex: same epidemic profile can due to:

- a slow infection with high R_0
- a fast infection with small R_0

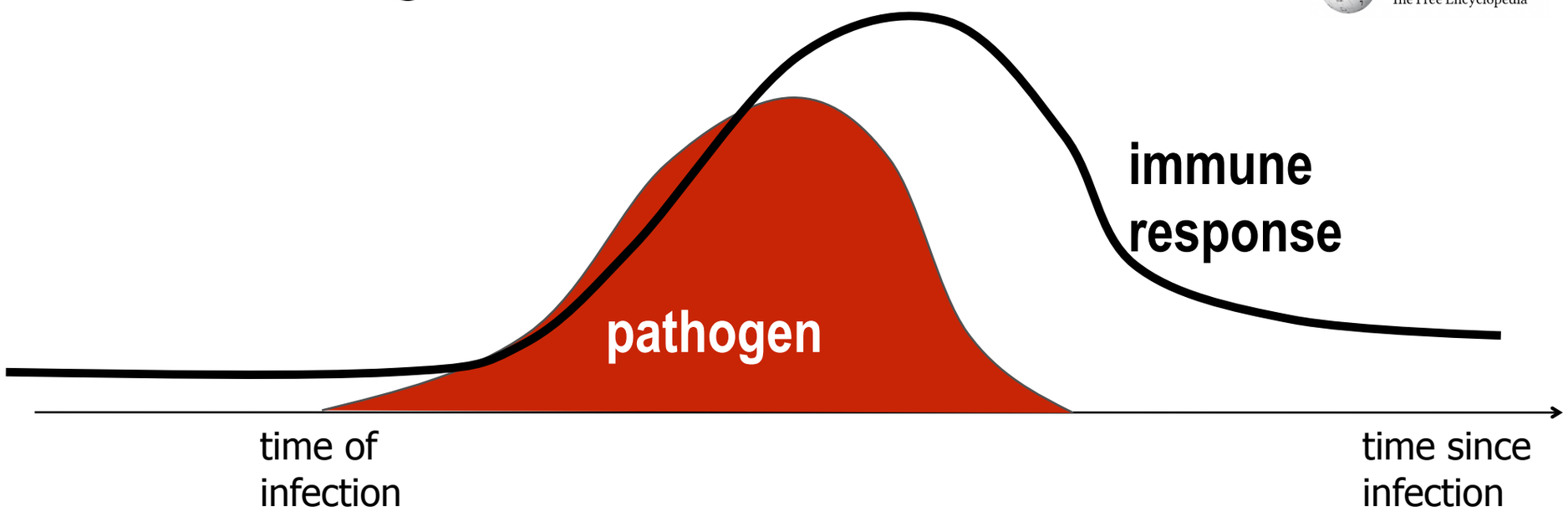
Fit incidence data with the wrong values of infectious or exposed period or the wrong model (**i.e. neglecting the exposed state or not well capturing the infectivity profile**) provide biased estimates of R_0

Incubation period, and infectious period (or better the generation time) are the very first quantities I need to estimate in a case of an emerging virus outbreak

disease natural history: incubation period, latency period, infectious period

Natural history of disease: the course a disease takes in individual people from its pathological onset ("inception") until its resolution, either through recovery or death

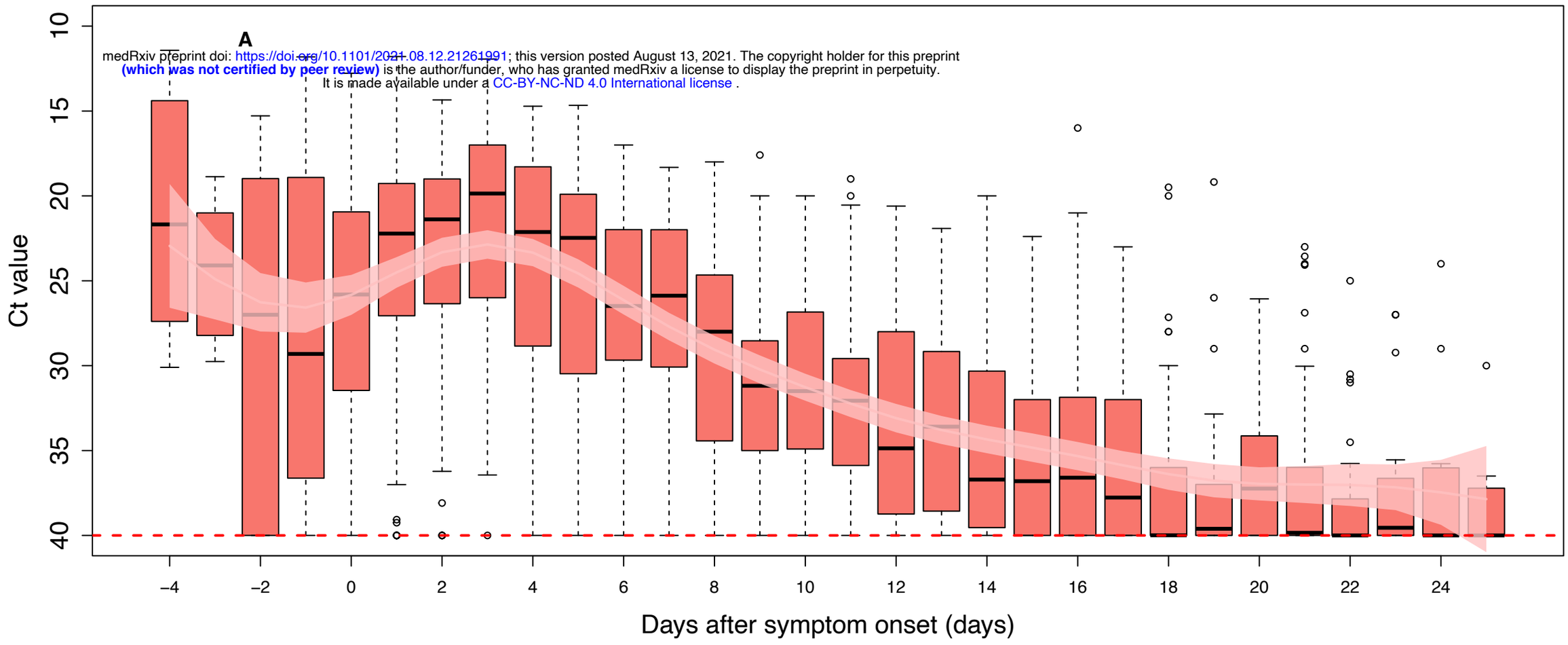
The inception of a disease is not a firmly defined concept. **The natural history of a disease is sometimes said to start at the moment of exposure to causal agents.**



[Keeling & Rohani, Modeling Infectious Diseases (2008)]

viral dynamics

viral load quantified by serial cycle threshold (Ct) values



[Kang et al medRxiv doi.org/10.1101/2021.08.12.21261991]

Not so easy to interpret:

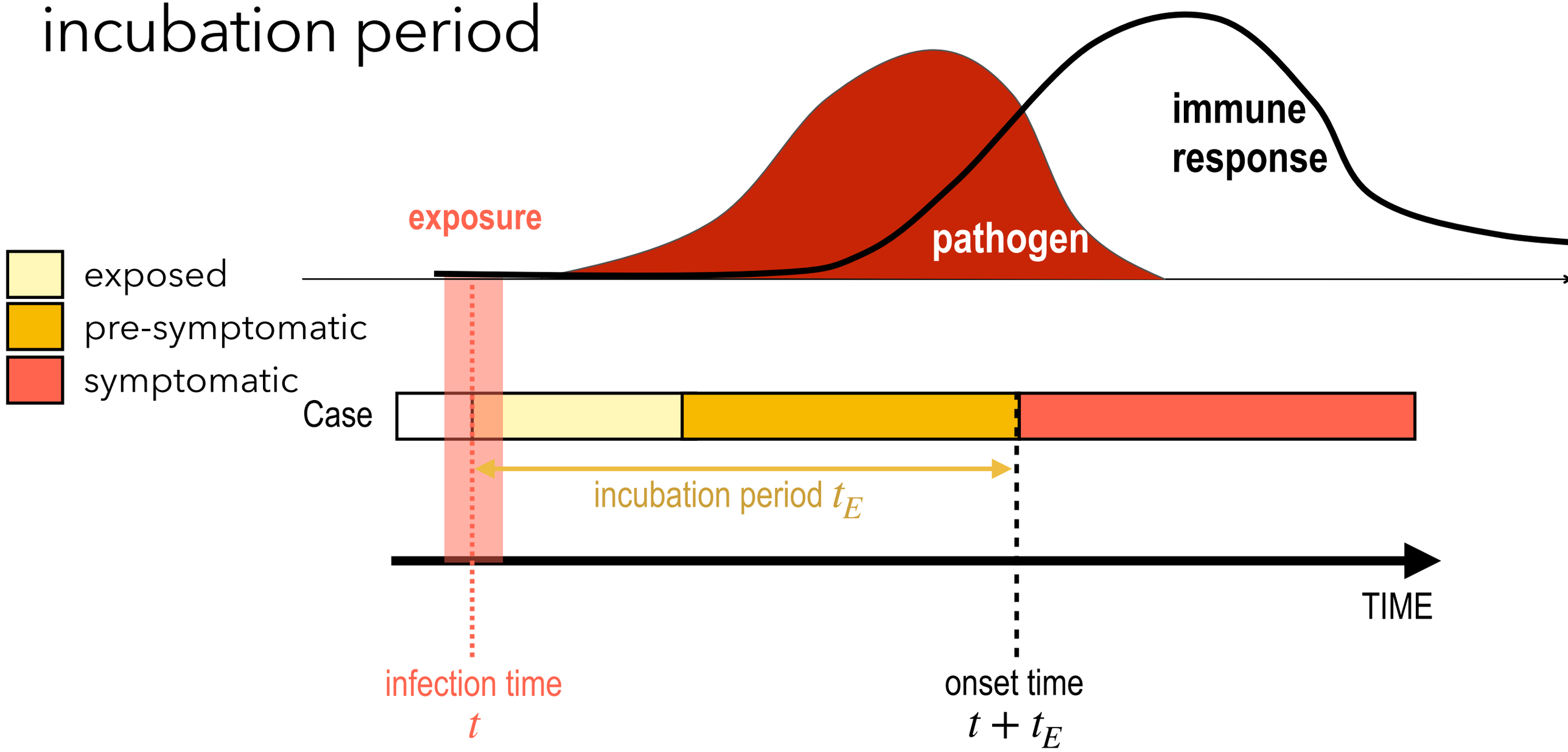
viral load ≠ symptoms score

viral load ≠ viral shading (what's the minimum viral load that makes you infectious?

symptoms instrumental to transmission)

disease natural history defined in epidemiological terms

incubation period



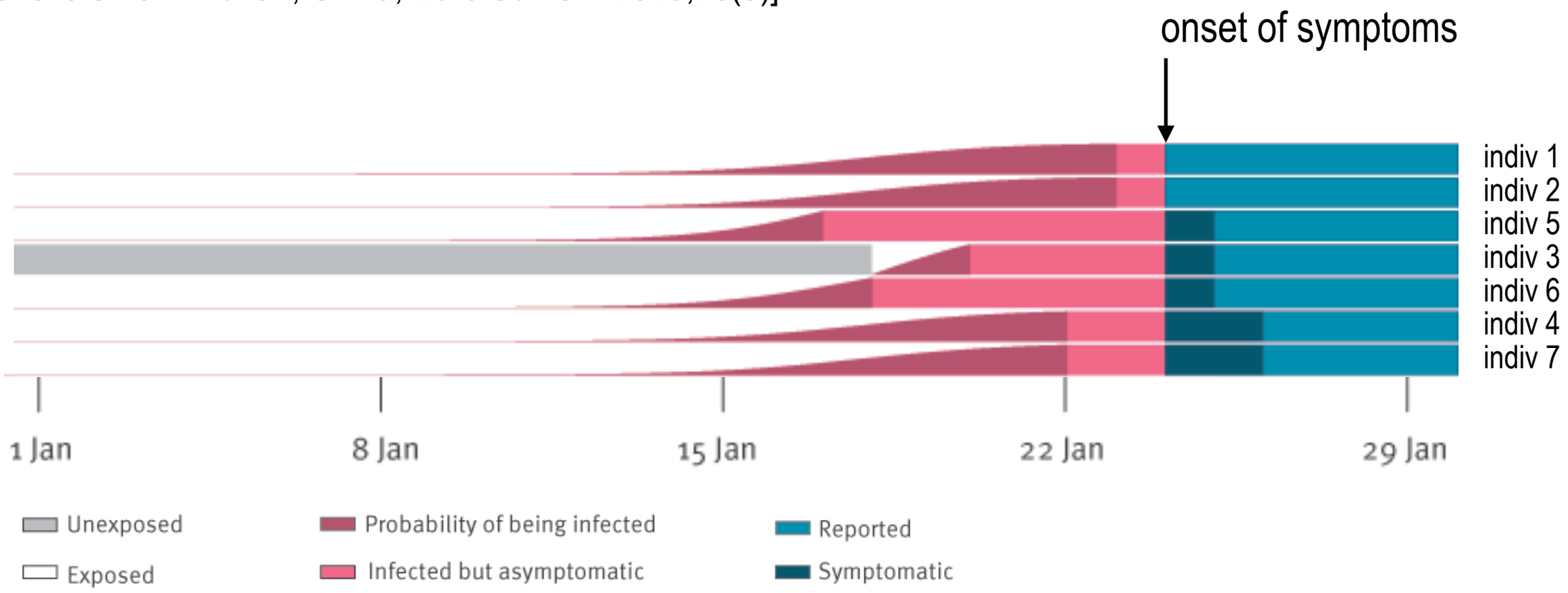
- contact tracing, case investigation. Once a case is confirmed, contacts are investigated. They are contacted, isolated and go through clinical and virological assessment. This allows collecting **infector-infected pairs**

- travelling cases

Onset time easier to collect than exposure time

incubation period

[Backer, Klinkenberg, Wallinga. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, Euro Surveill. 2020;25(5)]

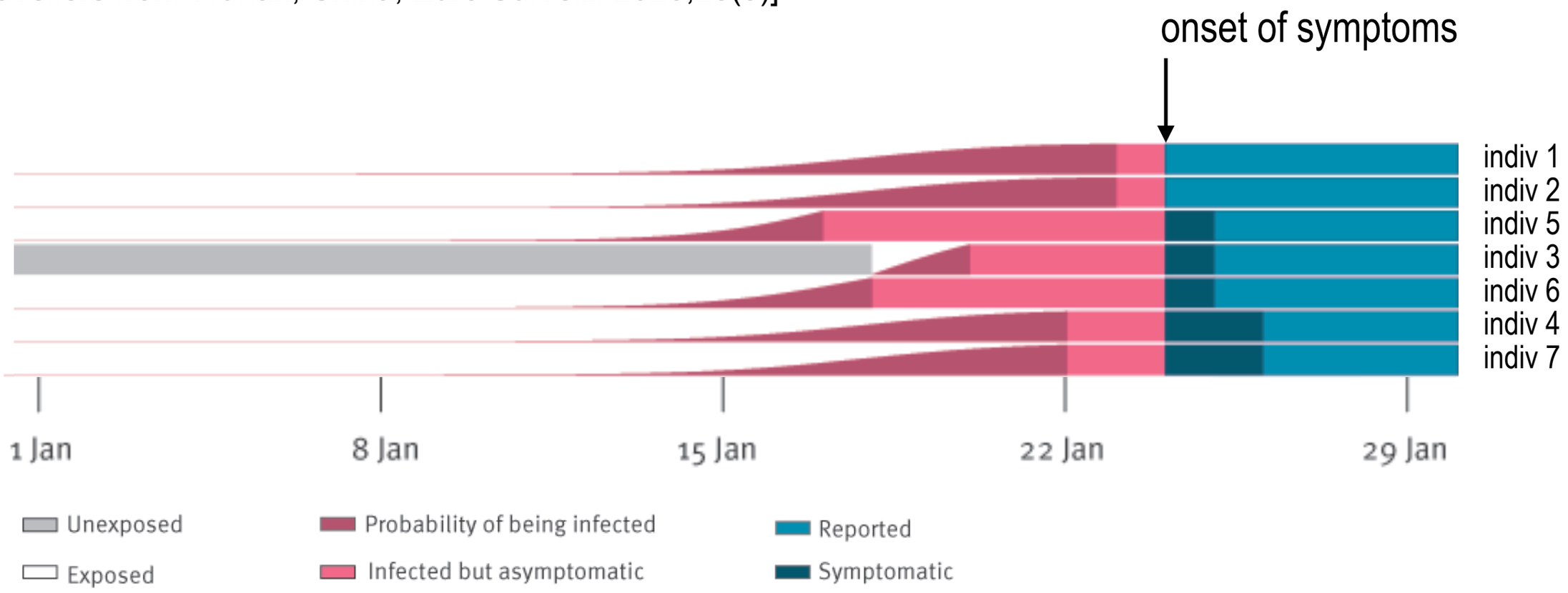


analysis of travelling cases: For 88 cases detected between 20 and 28 January, the travel history (to and) from Wuhan is known, as well as their symptom onset date

During this initial stage of the epidemic, it is most likely that these travellers were infected in Wuhan. Consequently, their time spent in Wuhan can be taken as the duration of exposure to infection.

incubation period

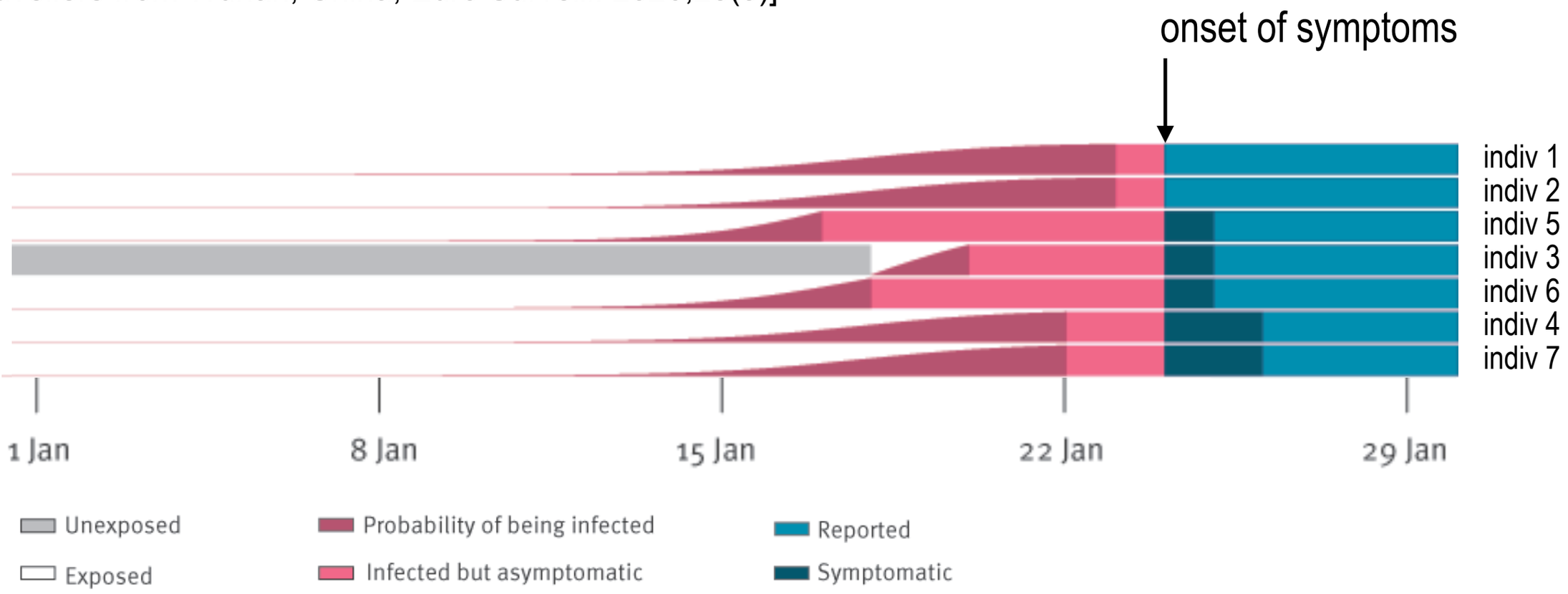
[Backer, Klinkenberg, Wallinga. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, Euro Surveill. 2020;25(5)]



Comparison between Gamma, Weibull and Lognormal distribution (2-parameter continuous distributions supported on a semi-infinite interval)

incubation period

[Backer, Klinkenberg, Wallinga. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, Euro Surveill. 2020;25(5)]

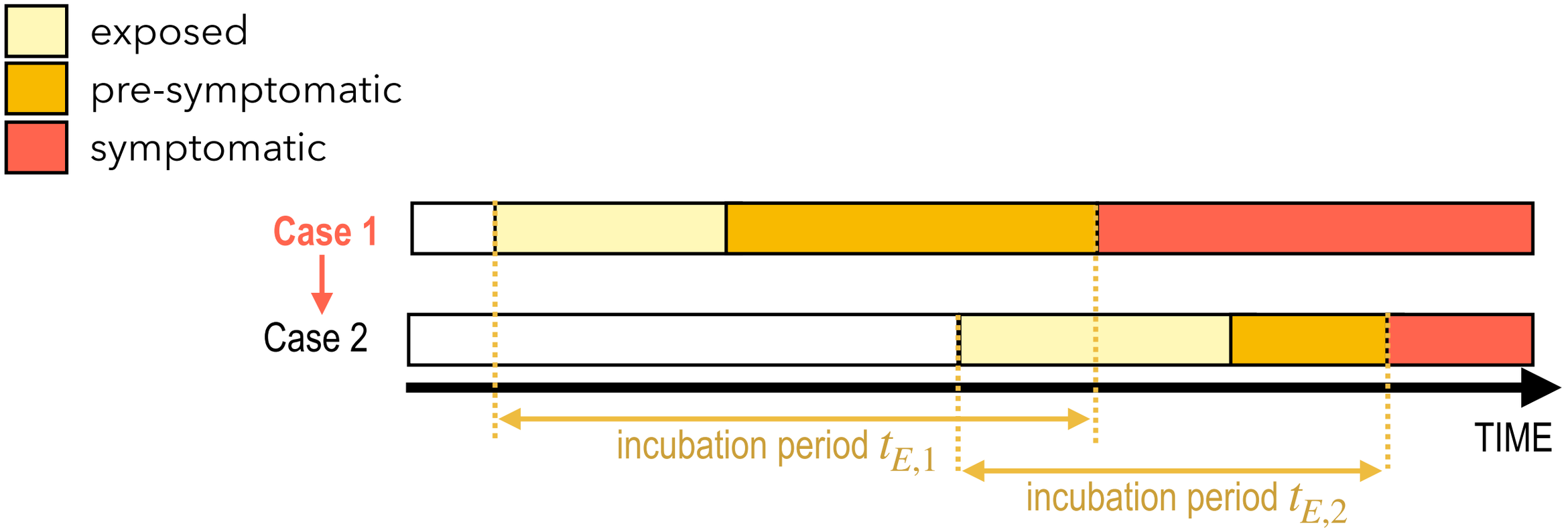


Comparison between Gamma, Weibull and Lognormal distribution (2-parameter continuous distributions supported on a semi-infinite interval)

Maximum Likelihood analysis to estimate the parameters for each distribution and compare the three assumptions

Result: incubation period = 6.4 days [2.1, 11.1] (other studies between 5, and 7)

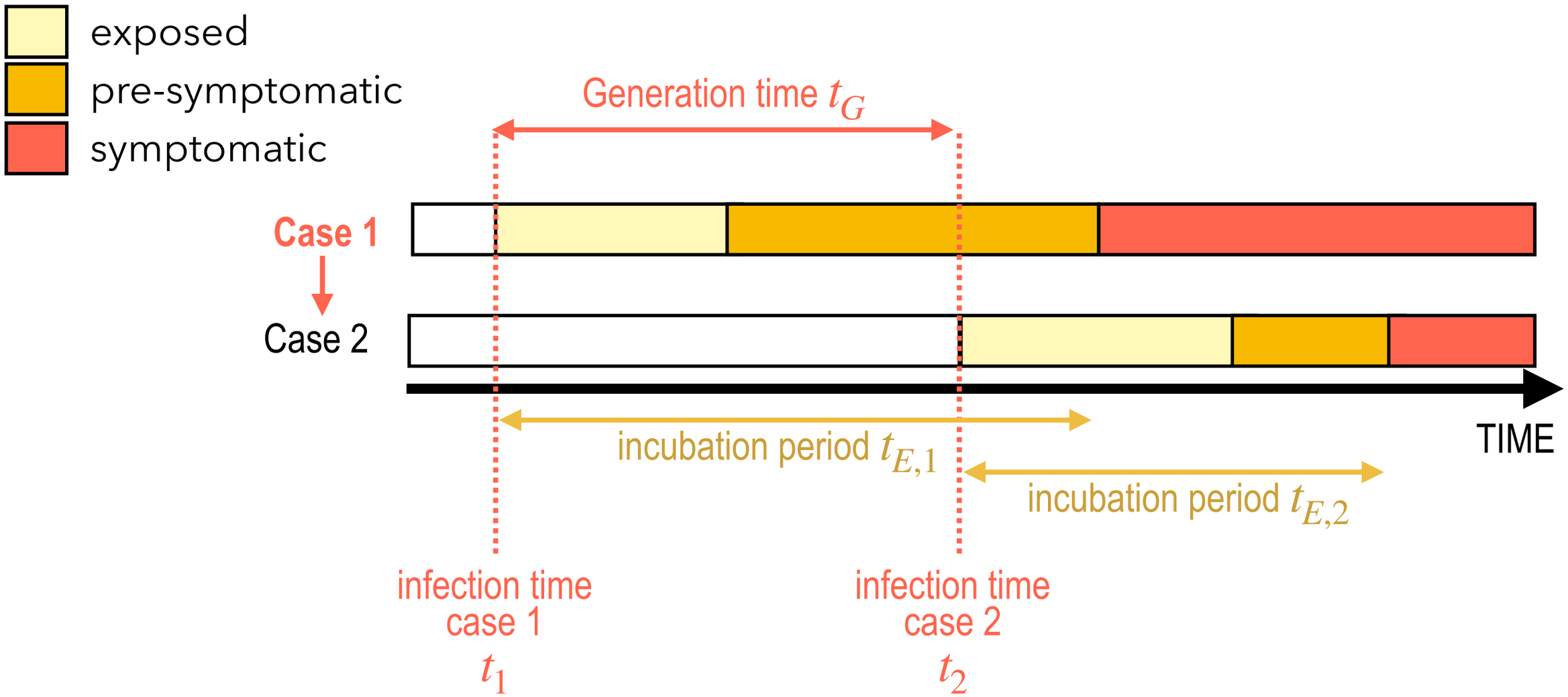
generation time & serial interval



t_E = incubation period, stochastic variable distributed according to $g(t_E)$ with average τ_E

t_I = infectious period, stochastic variable distributed according to $h(t_I)$ with average τ_I

generation time & serial interval

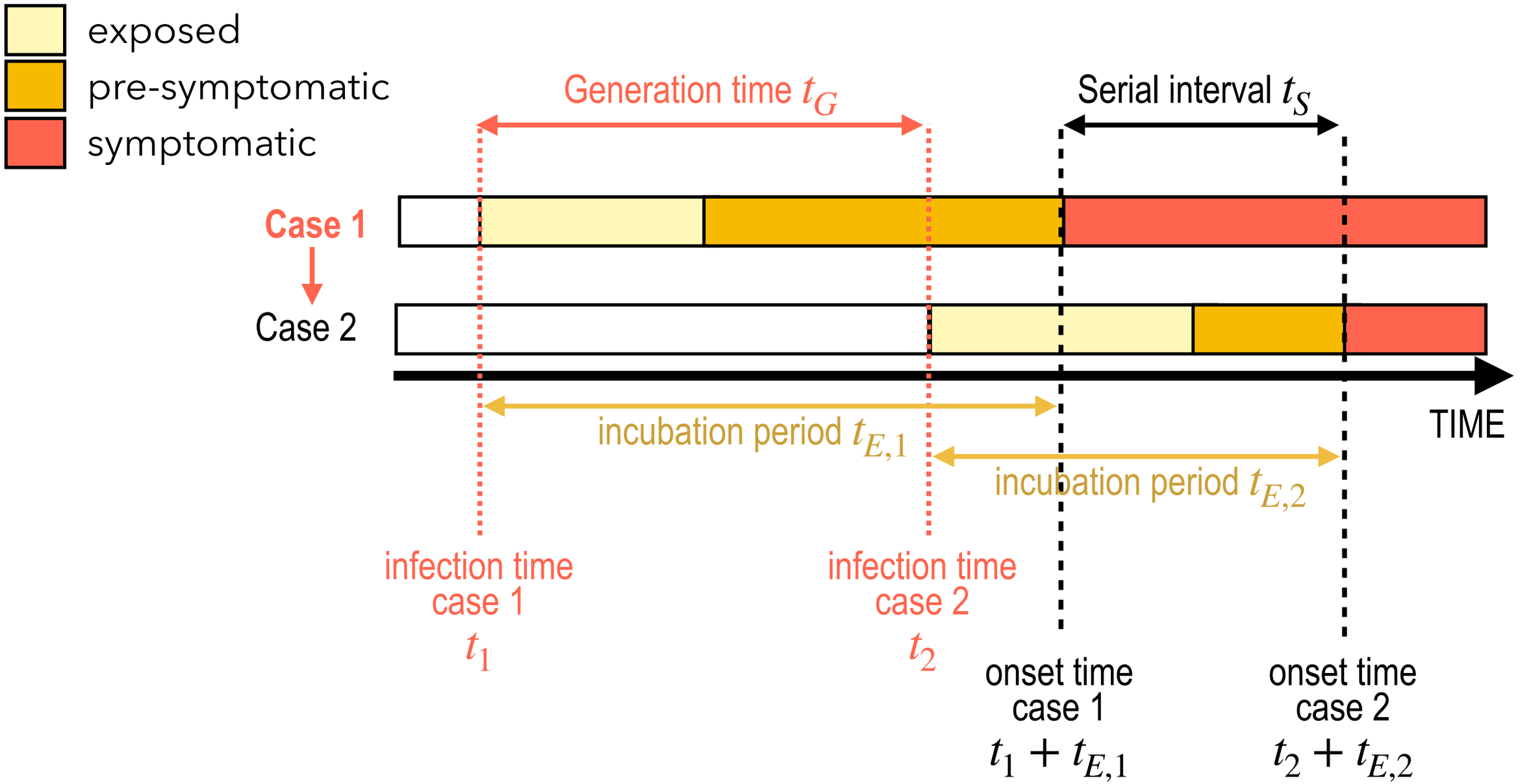


t_E = incubation period, stochastic variable distributed according to $g(t_E)$ with average τ_E

t_I = infectious period, stochastic variable distributed according to $h(t_I)$ with average τ_I

t_G = generation time, time from the infection of a case to the infection of the secondary case it generates

generation time & serial interval

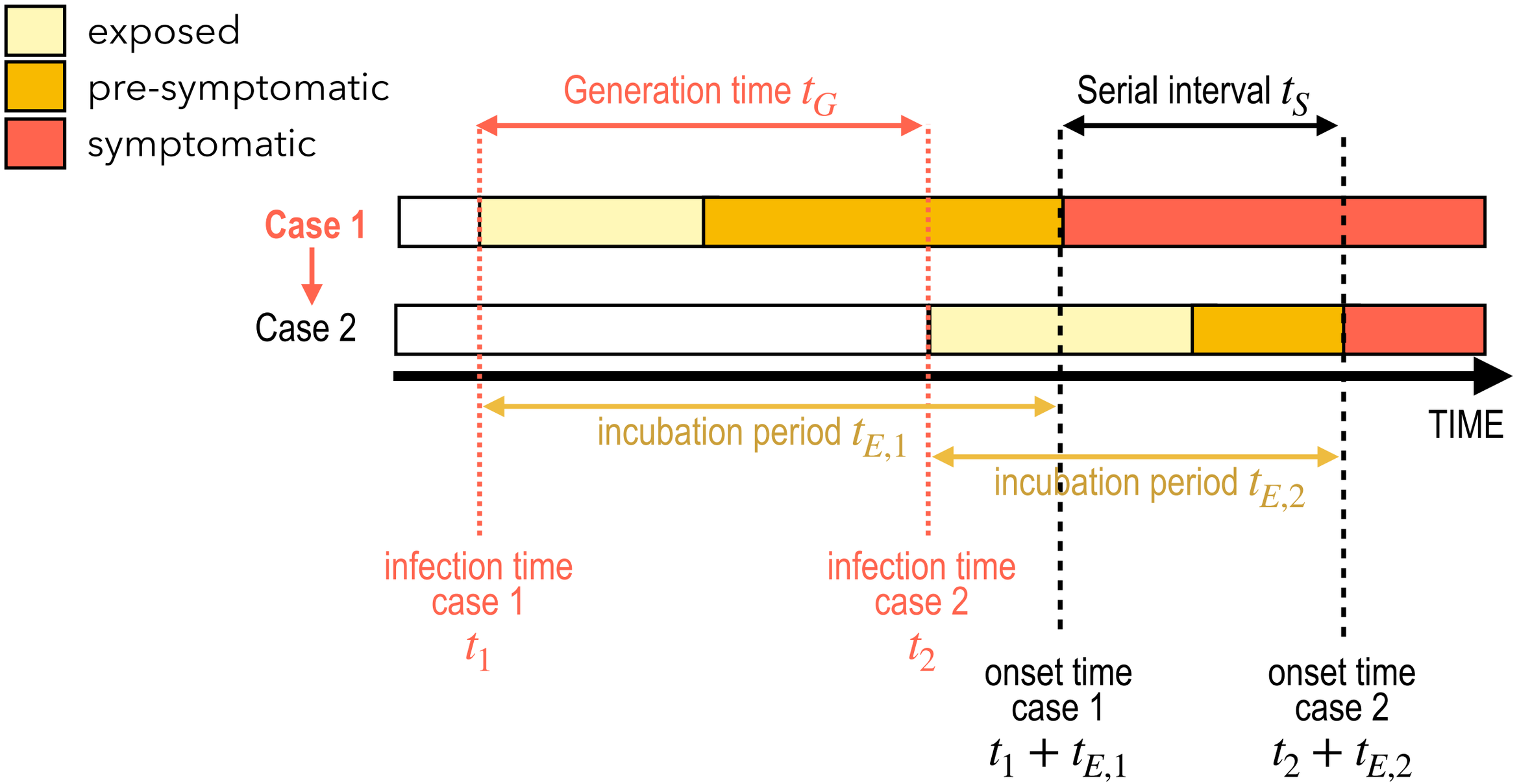


t_E = incubation period, stochastic variable distributed according to $g(t_E)$ with average τ_E

t_I = infectious period, stochastic variable distributed according to $h(t_I)$ with average τ_I

t_S = serial interval, time from the symptom onset of a case to the symptom onset of the secondary case it generates

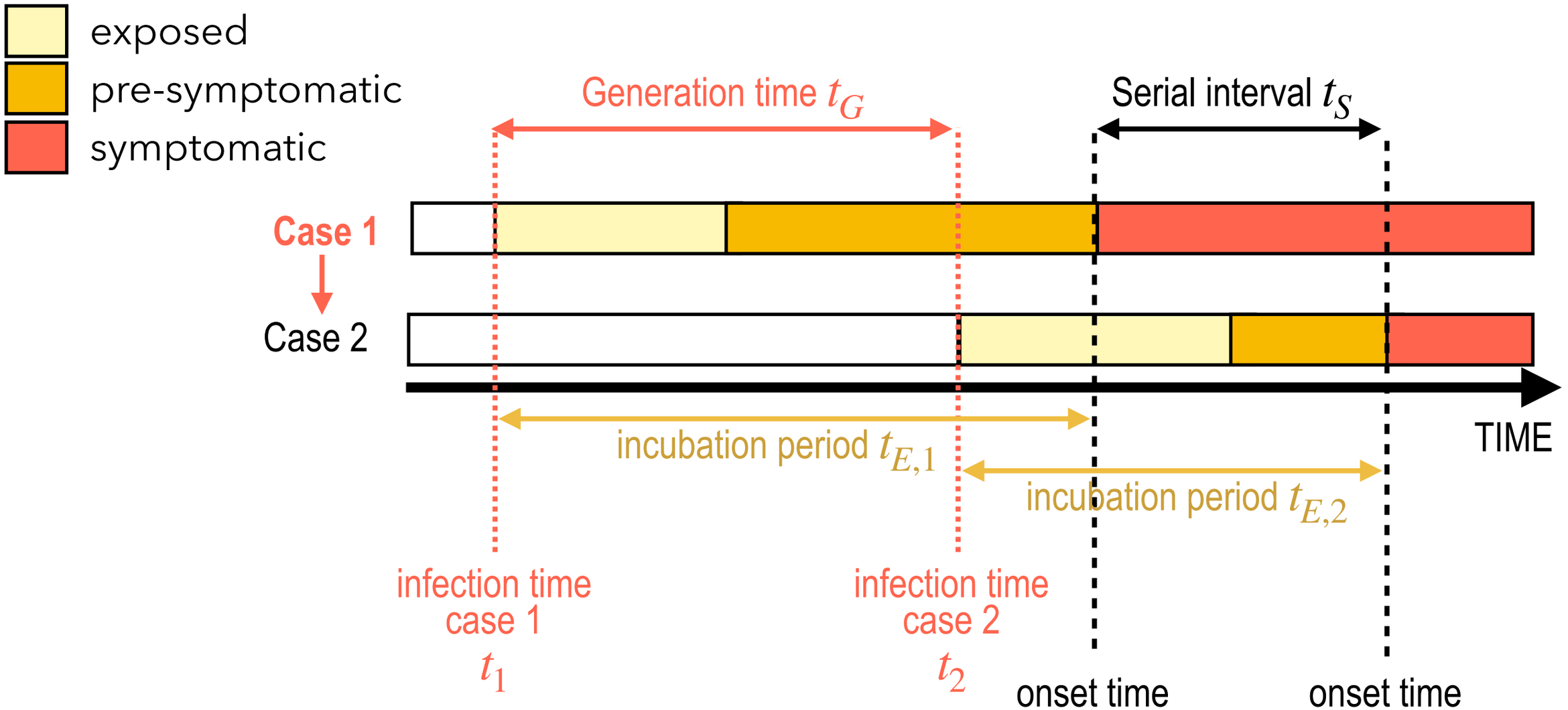
generation time & serial interval



t_E = incubation period, stochastic variable distributed according to $g(t_E)$ with average τ_E

t_I = infectious period, stochastic variable distributed according to $h(t_I)$ with average τ_I

generation time & serial interval



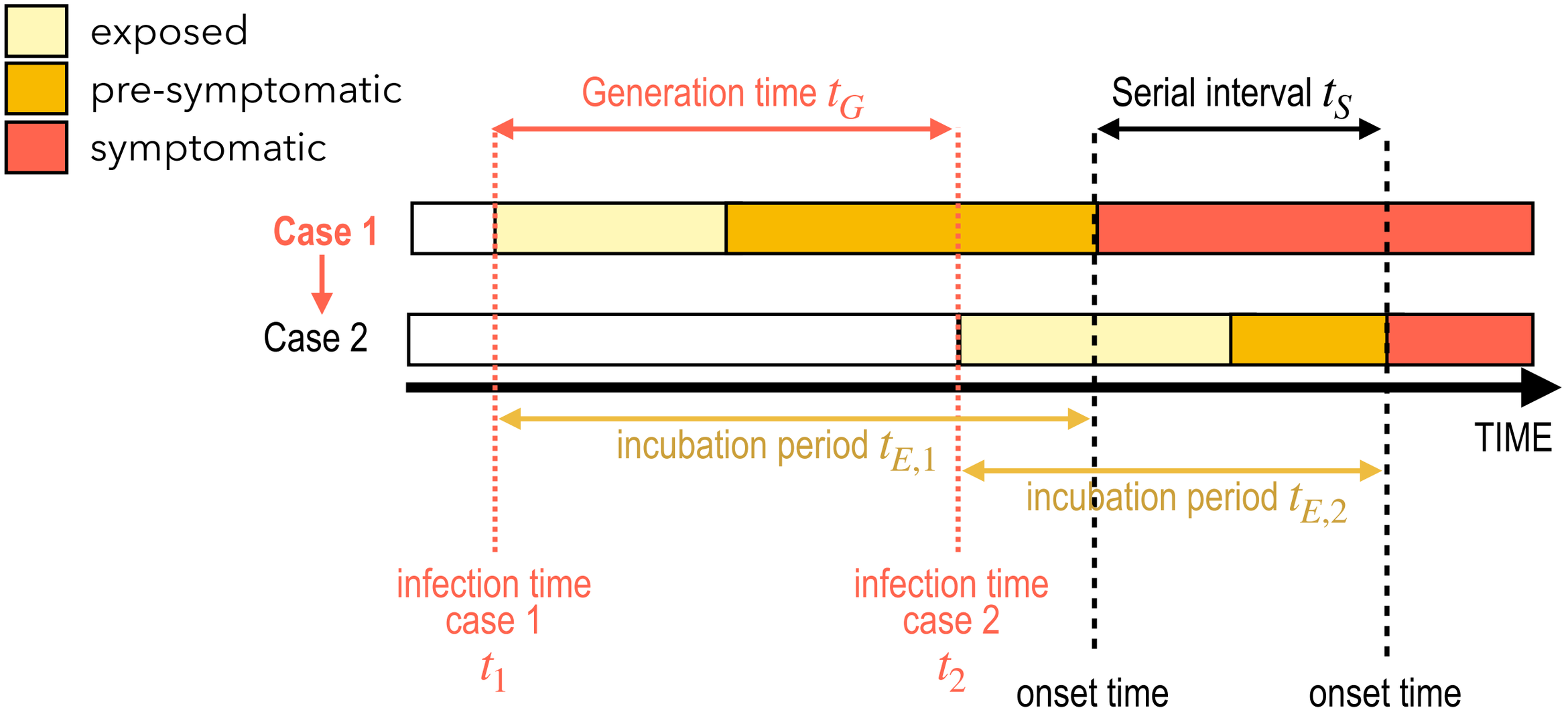
t_G distributed according to $w(t_G)$ with average τ_G

t_S distributed according to $f(t_S)$ with average τ_S

$$t_S = (t_2 + t_{E,2}) - (t_1 + t_{E,1}) = (t_2 - t_1) + (t_{E,2} - t_{E,1}) = t_G + (t_{E,2} - t_{E,1})$$

- f and w have same mean, $\tau_S = \tau_G \rightarrow$ serial interval used as a proxy for generation time

generation time & serial interval



t_G distributed according to $w(t_G)$ with average τ_G

t_S distributed according to $f(t_S)$ with average τ_S

$$t_S = (t_2 + t_{E,2}) - (t_1 + t_{E,1}) = (t_2 - t_1) + (t_{E,2} - t_{E,1}) = t_G + (t_{E,2} - t_{E,1})$$

- f and w have same mean, $\tau_S = \tau_G \rightarrow$ serial interval used as a proxy for generation time
- variance of f grater than variance of $w \rightarrow t_S$ can take negative values, signal that the pre-symptomatic phase is important for transmission

generation time & serial interval

[Ganyani, Kremer, Chen, Torneri, Faes, Wallinga, Hens. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, Euro Surveill. 2020;25(17):pii=2000257]

- 91 confirmed cases in Singapore + 135 confirmed cases in Tianjin
- Available information: cluster of a case, infector, traveling history
- the study relies on previous estimates of incubation period,
  $t_E \sim g(t_E|\theta_1) = \text{Gamma}(t_E|\theta_1)$

$$t_S = t_G + (t_{E,2} - t_{E,1})$$

$$x = t_{E,2} - t_{E,1}, x \sim h(x|\theta_1)$$

$$f(t_S|\theta_1,\theta_2) = \int_{-\infty}^{-\infty} w(t_G - x|\theta_2)h(x|\theta_1)dx$$

generation time & serial interval

[Ganyani, Kremer, Chen, Torneri, Faes, Wallinga, Hens. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, Euro Surveill. 2020;25(17):pii=2000257]

- 91 confirmed cases in Singapore + 135 confirmed cases in Tianjin
- Available information: cluster of a case, infector, traveling history
- the study relies on previous estimates of incubation period,
  $t_E \sim g(t_E|\theta_1) = \text{Gamma}(t_E|\theta_1)$

$$t_S = t_G + (t_{E,2} - t_{E,1})$$

$$x = t_{E,2} - t_{E,1}, x \sim h(x|\theta_1)$$

$$f(t_S|\theta_1,\theta_2) = \int_{-\infty}^{-\infty} w(t_G - x|\theta_2)h(x|\theta_1)dx = \mathscr{L}(\theta_2|\{t_{S,i}\},\theta_1)$$

generation time & serial interval

[Ganyani, Kremer, Chen, Torneri, Faes, Wallinga, Hens. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, Euro Surveill. 2020;25(17):pii=2000257]

- 91 confirmed cases in Singapore + 135 confirmed cases in Tianjin
- Available information: cluster of a case, infector, traveling history
- the study relies on previous estimates of incubation period,
  $t_E \sim g(t_E|\theta_1) = \text{Gamma}(t_E|\theta_1)$

$$t_S = t_G + (t_{E,2} - t_{E,1})$$

$$x = t_{E,2} - t_{E,1}, x \sim h(x|\theta_1)$$

$$f(t_S|\theta_1,\theta_2) = \int_{-\infty}^{-\infty} w(t_G - x|\theta_2)h(x|\theta_1)dx = \mathscr{L}(\theta_2|\{t_{S,i}\},\theta_1)$$

Once I have estimated θ_2 I can compute (numerically) $P(t_2 < t_1 + t_{E,1}) = P(t_S < t_{E,1})$:
Proportion of pre-symptomatic transmission

generation time & serial interval

[Ganyani, Kremer, Chen, Torneri, Faes, Wallinga, Hens. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, Euro Surveill. 2020;25(17):pii=2000257]

- 91 confirmed cases in Singapore + 135 confirmed cases in Tianjin
- Available information: cluster of a case, infector, traveling history
- the study relies on previous estimates of incubation period,
  $t_E \sim g(t_E|\theta_1) = \text{Gamma}(t_E|\theta_1)$

Results:

- Generation time 5.20 days [3.78–6.78] (Other studies between 5 and 7)
- Proportion of pre-symptomatic transmission ~50%