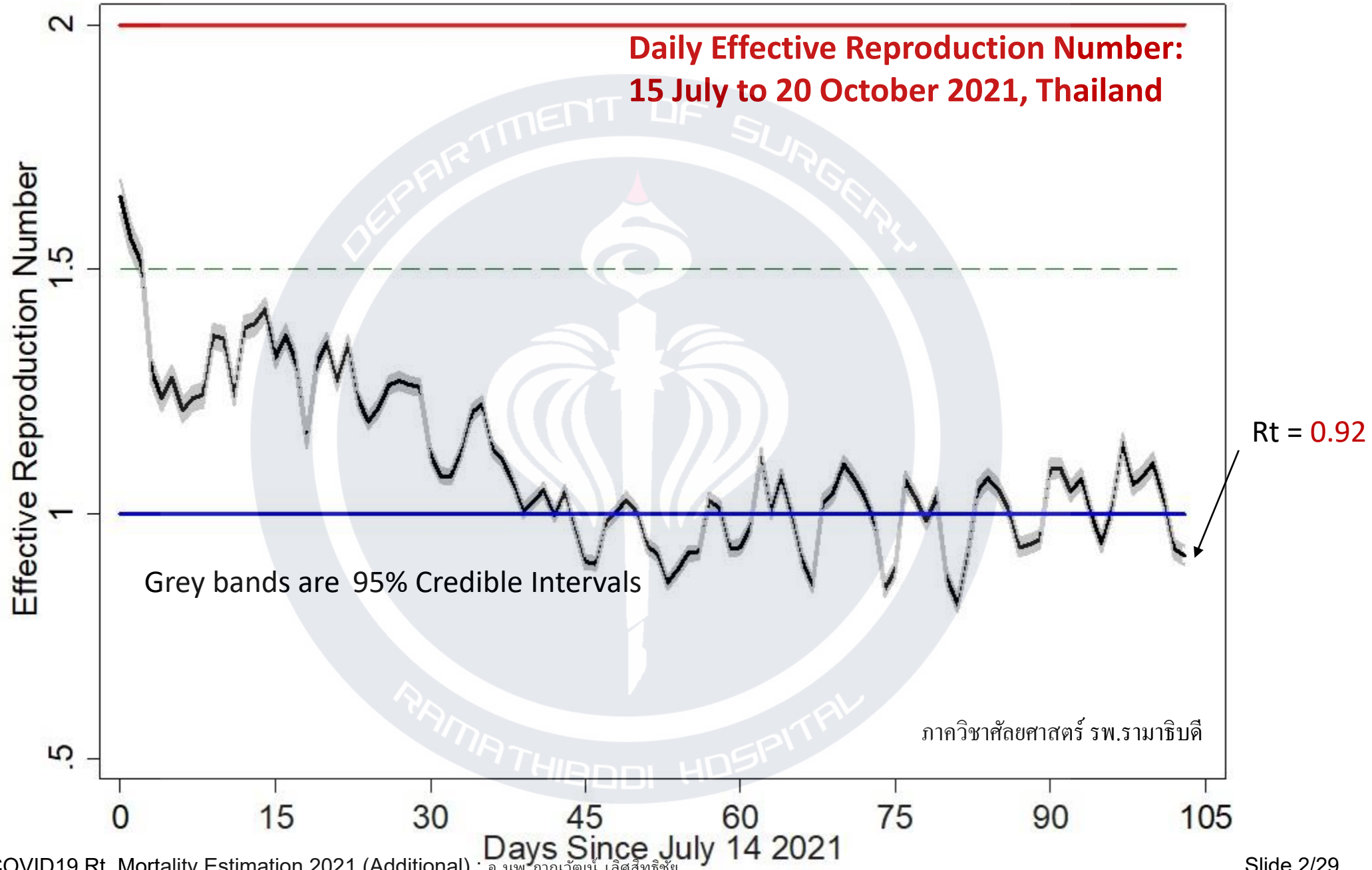
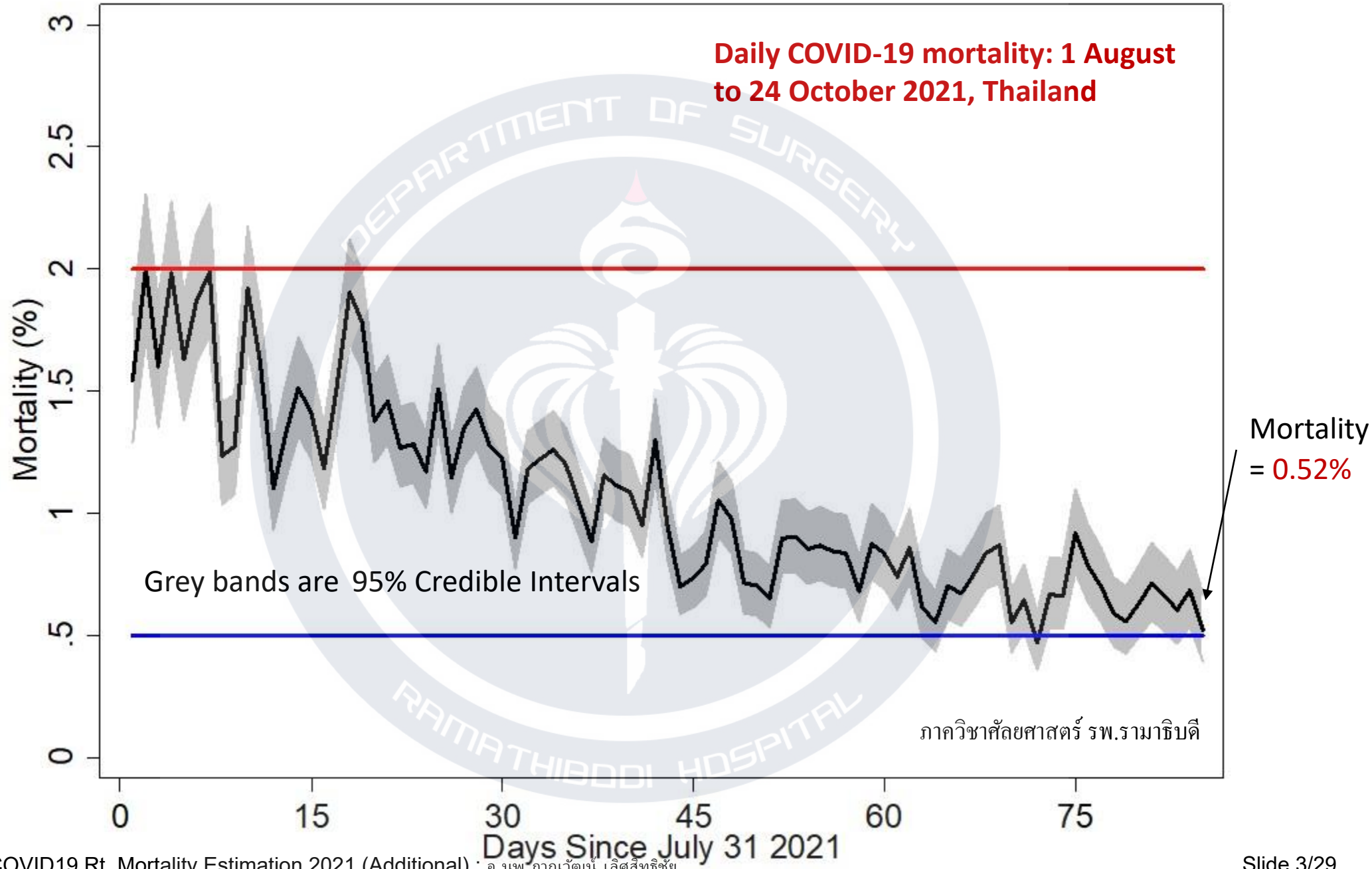
Model of COVID-19 Rt & Mortality Estimation: Additional Material

ภาณุวัฒน์ เลิศสิทธิชัย

ภาควิชาศัลยศาสตร์ รพ.รามาธิบดี

๒๔ ต.ค. ๒๕๖๔


Daily Effective Reproduction Number:
15 July to 20 October 2021, Thailand
Effective Reproduction Number
2
1.5
1
.5
Grey bands are 95% Credible Intervals
Rt = 0.92
ภาควิชาศัลยศาสตร์ รพ.รามาธิบดี
0
15
30
45
60
75
90
105
Days Since July 14 2021


Daily COVID-19 mortality: 1 August to 24 October 2021, Thailand
Mortality (%)
0
.5
1
1.5
2
2.5
3
0
15
30
45
60
75
Days Since July 31 2021
Grey bands are 95% Credible Intervals
Mortality = 0.52%
ภาควิชาศัลยศาสตร์ รพ.รามาธิบดี

Why a Model?

- The previous graphs show **2 measures** of COVID-19: the ability to pass on the infection (*Rt*) and its severity (mortality “rate”)
- These measures have precise meaning or definitions only within context of some mathematical model
- To obtain these measures unambiguously requires the use of some appropriate theoretical framework
- Here we present a simple model, which might capture the essence of the pandemic in Thailand, and arrive at simple statistical estimates of these 2 measures

A Model of *Rt* Estimation

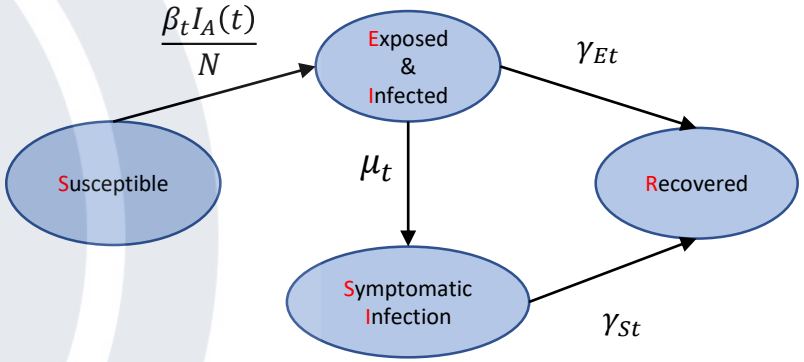
Suppose, in a population of size *N*, the SIR model below is approximately valid

(1) $$\frac{dS(t)}{dt} = -\frac{\beta_t\{I_E(t)+I_S(t)\}}{N}S(t) \equiv -\frac{\beta_t I_A(t)}{N}S(t)$$

(2) $$\frac{dI_E(t)}{dt} = \frac{\beta_t S(t)}{N}\{I_E(t) + I_S(t)\} - \mu_t I_E(t) - \gamma_{Et} I_E(t)$$

(3) $$\frac{dI_S(t)}{dt} = \mu_t I_E(t) - \gamma_{St} I_S(t)$$

(4) $$\frac{dR(t)}{dt} = \gamma_{Et} I_E + \gamma_{St} I_S; \quad S(t) + I_E(t) + I_S(t) + R(t) = N$$



From eqs. (2) & (3) we can write, where *A* refers to *all active infections,*

(5.1) $dI_A'(t) \equiv \frac{\beta_t dt S(t)}{N} I_A(t)$; i.e., the *increase in all infections within* $[t - dt, t]$ **ignoring recoveries** (this increase is symbolized with an apostrophe ’)

A Model of *Rt* Estimation

Equation (5.1) can also be written

(5.2) $$dI_A{}'(t) = \frac{\beta_t dt S(t)}{N} I_A(t) = \frac{\beta_t dt S(t)}{N} \int_{0+}^{t} w(u) dI_A{}'(t-u)$$

- Where $w(t)$ is called the ***infectiousness*** **distribution density function** (see later)*, *which takes into account the recovery from infection as well*, and is specifically defined such that the above equation is true

Technical note: $\int_{0+}^{t} w(u) dI_A{}'(t-u) = \int_{0}^{t-} w(t-u) dI_A{}'(u) = \int_{0}^{t-} w(t-u) \frac{dI_A{}'(u)}{du} du$; which is approximately $\sum_{j=1}^{n} \left(W_j - W_{j-1}\right) \frac{\Delta I_j{}'}{\Delta t}$, for n small intervals Δt, and where $W_j = \int_{0}^{t-t_j} w(u) du$

*The infectiousness distribution can be assumed similar to that given by He X, et al. *Temporal dynamics in viral shedding and transmissibility of COVID-19*. Nat Med, 2020

There is a **Conjecture Here**

I have yet to prove that

$$dI_A{}'(t) = \frac{\beta_t dt S(t)}{N} I_A(t) = \frac{\beta_t dt S(t)}{N} \int_{0+}^{t} w(u) dI_A{}'(t-u)$$

actually exists!*

- That is, there exists a convolution

$$I_A(t) = \int_0^t w(u) f(t-u) du \text{ ; with } f(t) = \frac{\beta_t}{N} S(t) I_A(t)$$

*See the appendix

The **Discrete Version**

Rewrite eq. (5.2) in discrete form for **daily infection incidence**, using the *discrete infectiousness function** w_i (omitting apostrophes):

(6.1) $\Delta I_{At} = \frac{\beta_t \Delta t S_t}{N} \sum_{i=1}^{t} w_i \Delta I_{At-i}$ where I_{At} begins at $t = 0$

Since all infections are all I_{Et} initially, as stated in eq. (2), we can also write $\Delta I_{Et} = \Delta I_{At}$, ignoring recoveries; and so it must be that

(6.2) $\Delta I_{Et} = \frac{\beta_t \Delta t S_t}{N} \sum_{i=1}^{t} w_i \Delta I_{Et-i}$ as well

Note on notation: ΔI_{At} refers to ΔI_A at day t; so, for example at day 3, ΔI_{At} would be written ΔI_{A3}

*The infectiousness distribution can be discretized in a way described in the previous slide, or as given by Cori A, et al. Am J Epidemiol, 2013

The Discrete Version

The increase in I_S according to equation (3) is

- $dI_S{}'(t) \equiv \mu_t dt I_E(t) \equiv \mu_t dt \int_{0+}^{t} v(u) dI_E{}'(t-u)$; where the increase with an apostrophe means recoveries are ignored, and $v(t)$ is the distribution function of days until the appearance of symptoms

The **discrete analogue** is (again omitting apostrophes)

(7) $\Delta I_{St} = \mu_t \Delta t \sum_{i=1}^{t} v_i \Delta I_{Et-i}$; v_i is the discretized version of $v(t)$

The Discrete Version

If we assume that symptomatic infection occurs at a lag exactly 3 days after exposure, thereby assuming a sharp peak* for v_i at $i = 3$, then $v_i = 1$ at $i = 3$ and thus

(8) $$\Delta I_{St} = \mu_t \Delta t \Delta I_{Et-3} \equiv \vartheta_t \Delta I_{Et-3}$$

Note again that ΔI_{St} is the **reported number of new infections per day** on day t, which is assumed to be all symptomatic

*Technically, a sharp peak for the distribution density function $v(t)$

The Discrete Version

Interpreting $\vartheta_t \equiv \mu_t \Delta t$ as the **fraction** of ΔI_{Et-3} (e.g., the daily increase in all infections 3 days prior) **becoming symptomatic** within the interval $[t - \Delta t, t]$, then

$$\Delta I_{St} = \vartheta_t \Delta I_{Et-3} = \vartheta_t \frac{\beta_{t-3} \Delta t S_{t-3}}{N} \sum_{i=1}^{t-3} w_i \Delta I_{Et-3-i} = \frac{\beta_{t-3} \Delta t S_{t-3}}{N} \sum_{i=1}^{t-3} w_i \left(\frac{\vartheta_t}{\vartheta_{t-i}} \right) \Delta I_{St-i}$$

- *Note* that $\beta_t \Delta t$ is interpreted as the number of transmissions of infection within the interval $[t - \Delta t, t]$ (e.g. within a day); and
- A detailed knowledge of v_i (the distribution of days till symptoms) is not needed if an exact 3-day lag is assumed

Estimating *Rt*: Use of Observed Incidence

With equation (9), below

(9) $\Delta I_{St} = \frac{\beta_{t-3}\Delta t S_{t-3}}{N} \sum_{i=1}^{t-3} w_i \left(\frac{\vartheta_t}{\vartheta_{t-i}}\right) \Delta I_{St-i}$, and **defining**

$R_t = \frac{\beta_t \Delta t S_t}{N}$, then, assuming ϑ_t to be *approximately constant*,

as well as assuming w_i to *approach zero for large times*, we finally find

(10) $\Delta I_{St} = R_{t-3} \sum_{i=1}^{t} w_i \Delta I_{St-i}$

Hence, a **consistently observed daily incidence ΔI_{St} can be used to estimate R_t**, *though at a lag*, in a long series of observations – **so it is not necessary to observe all infections** (i.e., ΔI_{Et} is not needed)

Estimating *Rt*: **Statistical Version**

In terms of statistical theory, we can **interpret**

- $\frac{\beta_t dt I_A(t)}{N}$ as the probability of infection within $[t-dt, t]$, and $\frac{\beta_t dt I_A(t)}{N} S(t)$ as the expected (average) number of new infections within the same interval,
- $\mathrm{E}[dI_A{}'(t)] = \frac{\beta_t dt I_A(t)}{N} S(t)$; similarly, eq.(10) in statistical terms is
- $\mathrm{E}[\Delta I_{St}] = R_{t-3} \sum_{i=1}^{t} w_i \Delta I_{St-i}$ (given ΔI_{St-i}'s) and suitable statistical models can be found for ΔI_{St} and R_t

Estimating *Rt*: **Statistical Model**

The **Bayesian statistical approach** can be used to estimate R_t as, e.g., in the statistical model proposed by Cori, *et al**:

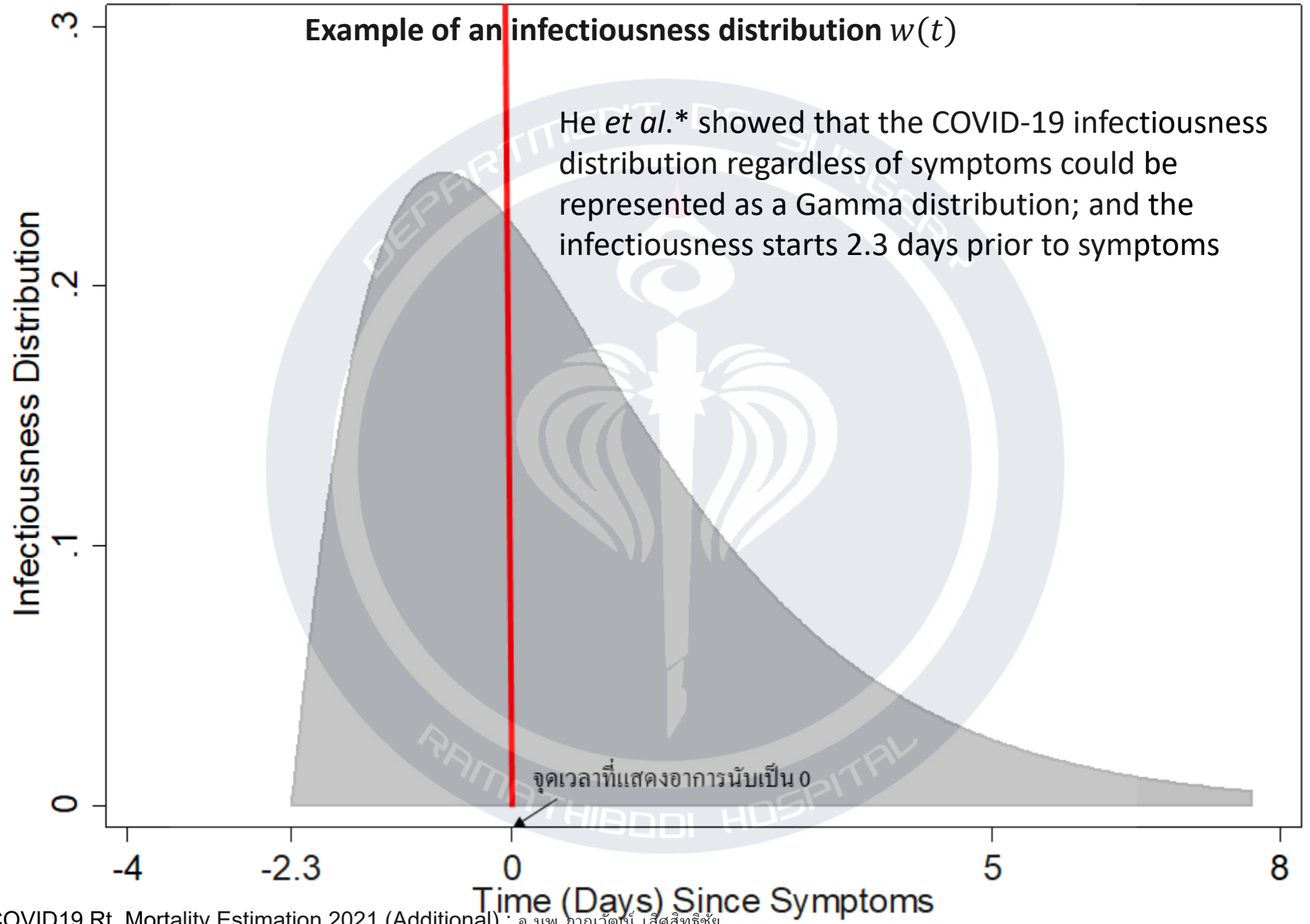
- *Prior distribution*: $R_{t,prior} \sim Gamma(a, b)$; with shape & scale parameters a, b
- *Likelihood*: $\Delta I_{St} | R_t, \Delta I_{St-1}, \ldots, \Delta I_{S0} \sim Poisson(R_t \Lambda_{wt})$;

  where $\Lambda_{wt} = \sum_{i=1}^{t} w_i \Delta I_{St-i}$, thus
- *Posterior distribution*: $R_{t,posterior} \sim Gamma\left(a + \Delta I_{st}, \frac{1}{\frac{1}{b} + \Lambda_{wt}}\right)$

Without taking into account the uncertainties associated with daily infections or infectiousness functions or lag times, the 95% credible intervals will be too narrow; these estimates are used in all *Rt* graphs

*Cori A, et al. Am J Epidemiol. 2013

Example of an infectiousness distribution $w(t)$



He *et al*.* showed that the COVID-19 infectiousness distribution regardless of symptoms could be represented as a Gamma distribution; and the infectiousness starts 2.3 days prior to symptoms

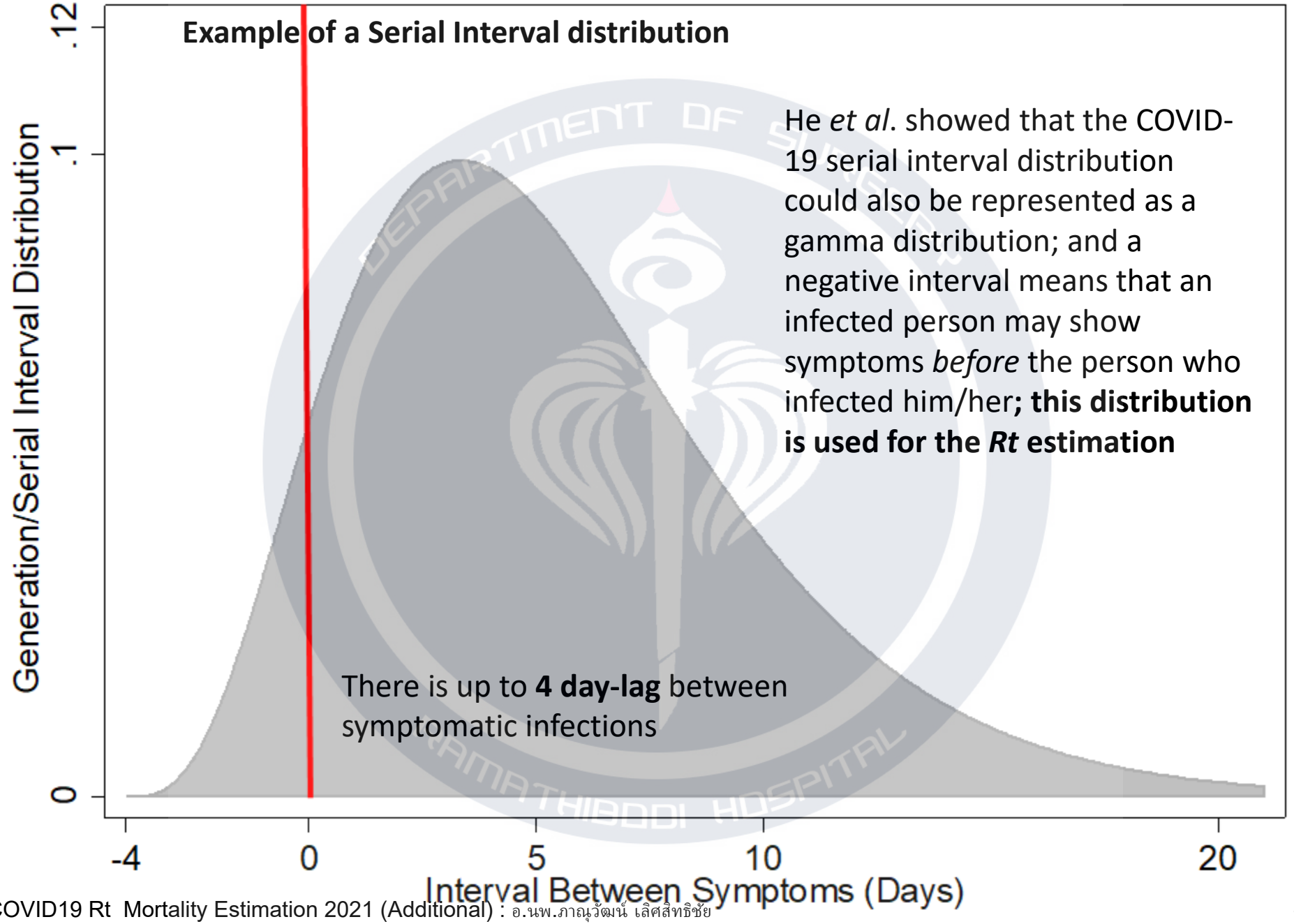
*He X, et al. *Temporal dynamics in viral shedding and transmissibility of COVID-19*. Nat Med, 2020

Some Theoretical Comments

- In fact, μ_t can **represent any observed fraction of all infections** (not necessarily symptomatic infections) – the conclusion will still be the same, as long as the assumption of *identical infectiousness distribution* (*identical* w) *for all infected persons* holds
- The effective reproduction number defined by $R_t = \frac{\beta_t \Delta t S_t}{N}$ depends on the *transmission parameter* β_t, which could be time dependent; as well as the *fraction of susceptibles* $\frac{S_t}{N}$, which decreases with time, and may appreciably change within the period of interest
- In particular, **vaccination** will significantly decrease the value of $\frac{S_t}{N}$

Notes on Statistical Estimation

- The actual estimated *Rt*'s as shown in these slides were obtained with a user-written *Stata*® program (available on request), where the *serial interval distribution replaced* the infectiousness distribution, such that there is a **lag of 4 days** (see the following figure)
- This replacement was partly because the serial interval distribution deals with symptomatic infections directly, without having to assume a theory linking presymptomatic to later symptomatic infection
- But the model in the program assumes all infections to become symptomatic, which is not true of COVID-19



He *et al*. showed that the COVID-19 serial interval distribution could also be represented as a gamma distribution; and a negative interval means that an infected person may show symptoms *before* the person who infected him/her**; this distribution is used for the *Rt* estimation**

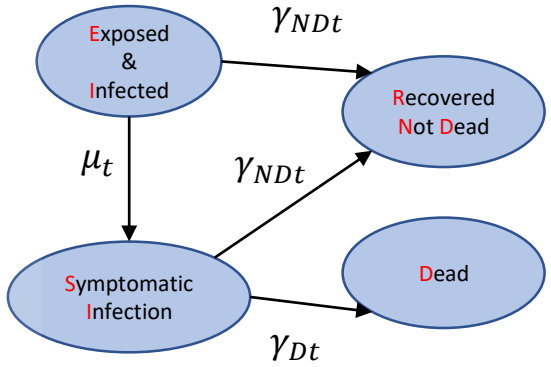
He X, et al. *Temporal dynamics in viral shedding and transmissibility of COVID-19*. Nat Med, 2020

Model for the Estimation of Mortality Rates

Assume that the recovery compartment R consists of those not dying (ND) and those who died (D)

That is, the equation for R is

(11) $$\frac{dR(t)}{dt} = \gamma_{Et} I_E + \gamma_{St} I_S = \gamma_{NDt} I_E + (\gamma_{Dt} + \gamma_{NDt}) I_S$$



We can define the rate of increase in no. of deaths as

(12) $$\frac{dD(t)}{dt} = \gamma_{Dt} I_S$$ for the compartment D

Where the required parameter to estimate is γ_{Dt}, **the mortality rate of symptomatic infections**

The Discrete Version

The discrete version of eq. (12), $dD(t) = \gamma_{Dt} dt \int_{0+}^{t} g(u) dI_S'(t-u)$, is

- $\Delta D_t = \gamma_{Dt} \Delta t \sum_{i=1}^{t} g_i \Delta I_{St-i}$, defined in a similar way as before

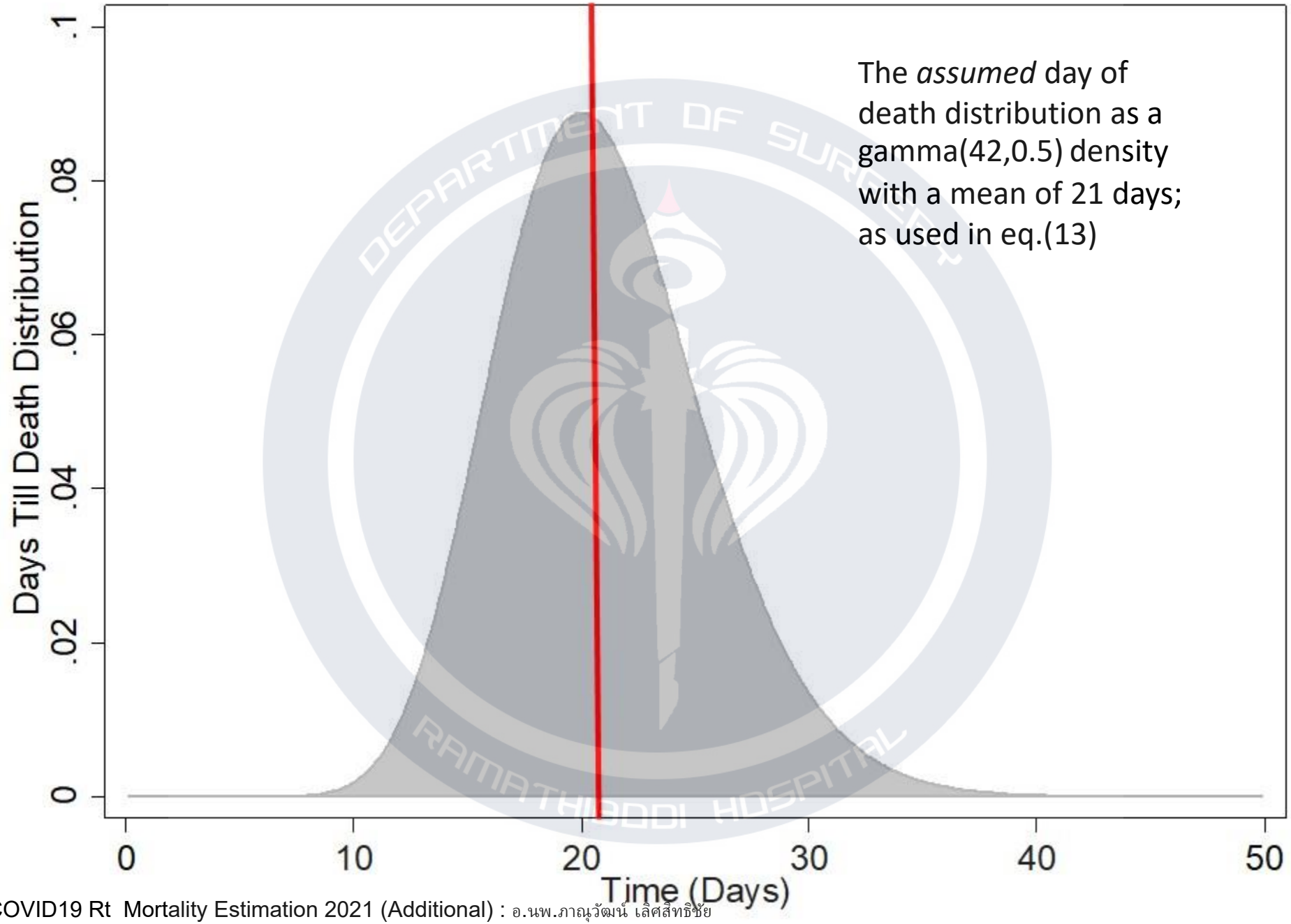
Writing the proportion of deaths in a day as $\delta_t = \gamma_{Dt} \Delta t$ we have

(13) $\Delta D_t = \delta_t \sum_{i=1}^{t} g_i \Delta I_{St-i}$;

where the **unknown distribution function** g_i can be a discretized version of an appropriate *gamma distribution* of the day of death of those with symptomatic infections, starting from 1st day of symptoms

- Assuming the average day of death to be 21 days*, with a relatively narrow distribution, one possibility could be (see next figure):

*Patel A, et al. JAMA Network Open 2021;doi:10.1001/jamanetworkopen.2021.12782



The *assumed* day of death distribution as a gamma(42,0.5) density with a mean of 21 days; as used in eq.(13)

Estimating δ_t: Statistical Model

The **Bayesian statistical approach** can once again be used (as per Cori, *et al**); via a statistical version of eq.(13) : $\mathrm{E}[\Delta D_t] = \delta_t \sum_{i=1}^{t} g_i \Delta I_{St-i}$

- *Prior distribution*: $\delta_{t,prior} \sim Gamma(a, b)$; with shape & scale parameters a, b
- *Likelihood*: $\Delta D_t | \delta_t, \Delta I_{St-1}, \ldots, \Delta I_{S0} \sim Poisson(\delta_t \Lambda_{gt})$;

  where $\Lambda_{gt} = \sum_{i=1}^{t} g_i \Delta I_{St-i}$, thus

- *Posterior distribution*: $\delta_{t,posterior} \sim Gamma\left(a + \Delta D_t, \frac{1}{\frac{1}{b} + \Lambda_{gt}}\right)$

Ignoring uncertainties associated with the number of daily deaths and day of death distribution functions will make the 95% credible intervals too narrow; all mortality graphs are those of δ_t

*Cori A, et al. Am J Epidemiol. 2013

Appendix: The Functions $w(t)$ and $\beta(t)$

I have yet to prove that

$$dI_A{}'(t) = \frac{\beta_t dt S(t)}{N} I_A(t) = \frac{\beta_t dt S(t)}{N} \int_{0+}^{t} w(u) dI_A{}'(t-u)$$

actually exists

- Is there a function $w(u)$ with the desired properties and satisfying the SIR equations **from a *point-wise* perspective**?
- According to the SIR model with constant $\beta_t = \beta$, $w(u) = 1 - \frac{\gamma N}{\beta S(t-u)}$ which makes no sense if $w(u)$ is interpreted as the infectiousness
- Perhaps we require β_t to *not* be a constant?

The Functions $w(t)$ and $\beta(t)$

If $w(t)$ is **arbitrary** and β no longer constant, from eq.s (2), (3), (5.2)

- $I_A(t) = \int_0^{t-} w(t-u) dI_A'(u) = \int_0^{t-} w(t-u) \frac{\beta(u)}{N} S(u) I_A(u)\, du =$
  $\int_0^t \left(\frac{\beta(u)}{N} S(u) - \gamma \right) I_A(u) du$; all γ's being equal

For *pointwise equivalence* of the integrals, and all non-zero functions

- $w(t-u) \frac{\beta(u)}{N} S(u) = \frac{\beta(u)}{N} S(u) - \gamma$; with $u \in [0, t]$, or
- $\frac{\beta(u)}{\gamma} = \frac{N}{S(u)\{1 - w(t-u)\}}$; but again, this does not make much sense

The Functions $w(t)$ and $\beta(t)$

- With $\frac{\beta(u)}{\gamma} = \frac{N}{S(u)\{1-w(t-u)\}}$, $\frac{\beta(u)}{\gamma}$ unrealistically depends on $w(t-u)$
- Also, $\frac{\beta}{\gamma}$ can not be smaller than 1 and must be 1 at $u = 0, t\ large$ and will always be larger than 1 at large times with $u = t$, and this can not be true in reality
- And $w(t-u)$ can not be a density, as we assumed it to be
- The **problem lies** with the equality $w(t-u)\frac{\beta(u)}{N}S(u) = \frac{\beta(u)}{N}S(u) - \gamma$ which brings together the infectiousness function and the SIR model **at each time point**

Does Not Work Either

- Another approach: in the equation $w(t-u)\frac{\beta\prime(u)}{N}S(u) = \frac{\beta}{N}S(u) - \gamma$ define another transmission parameter where $\beta'(u) \neq \beta$
- Then $\frac{\beta\prime(u)}{\gamma} = \frac{\frac{\beta}{\gamma} - \frac{N}{S(u)}}{w(t-u)}$; again, this makes no more sense than before

How to Resolve This?

- For $w(t-u)\frac{\beta(u)}{N}S(u) = \frac{\beta(u)}{N}S(u) - \gamma$, the LHS is always positive while the right can have negative values, but *their integrals are always positive or zero*
- **So the equivalence cannot be pointwise**
- For now convolution-like equations such as eq.5.2 are *pure conjecture*

A Possibility

- For the truncated convolution

$$I_A(t) = \int_0^t w(u) f(t-u) du \text{ ; with } f(t) = \frac{\beta_t}{N} S(t) I_A(t)$$

- If the following exists, with $\delta(u)$ a Dirac delta-like function,

$$I_A(t) = \frac{\beta_t}{N} \int_0^t \delta(u) I_A(t-u) du \text{ , where } \beta_t \text{ can be time-dependent}$$

- Then with $w(u) = \frac{\delta(u)}{S(t-u)}$ for non-zero $S(t)$ and a sharp-peaked $\delta(t)$ or something similar, such as distributions in slides 17 or 20, the convolution should exist

Notes

- The connection between the SIR models and the discrete model is not formal and is mainly for motivation purposes
- The convolution, whether or not it formally exists, is plausible
- The discrete version makes good statistical/probabilistic sense
- And some SIR-like equations will emerge from the discrete stochastic model in a continuum limit