

Time between infections vs time between symptom onset in COVID-19: implications for estimating the reproduction number

Jesse Knight^{1,2} Sharmistha Mishra^{1,2,3,4}

¹ Institute of Medical Science, University of Toronto

² MAP Centre for Urban Health Solutions, Unity Health Toronto

³ Division of Infectious Disease, Department of Medicine, University of Toronto

⁴ Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto

2021 June 21

CAIMS Mini-Symposium

Mathematical modeling of COVID-19 transmission and mitigation strategies: efforts to end the pandemic

Outline

- ▶ Reproduction Number R
- ▶ Estimating R
- ▶ Estimating Generation Time Distribution $G(\tau)$
- ▶ Results
- ▶ Implications

Reproduction Number: Defined

Average number of secondary infections generated by each infectious individual

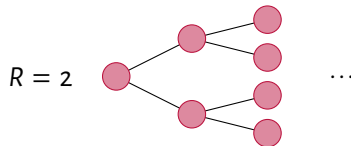
Reproduction Number: Defined

Average number of secondary infections generated by each infectious individual

$$R = 1 \quad \text{●} \text{---} \text{●} \text{---} \text{●} \quad \dots$$

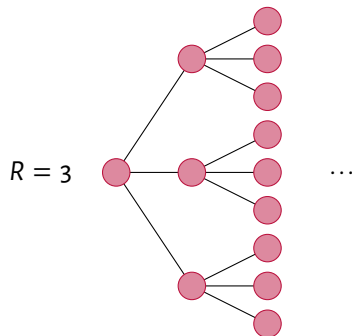
Reproduction Number: Defined

Average number of secondary infections generated by each infectious individual



Reproduction Number: Defined

Average number of secondary infections generated by each infectious individual



Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

- β = probability of transmission per contact

Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

- ▶ β = probability of transmission per contact
- ▶ C = contact rate

Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

- ▶ β = probability of transmission per contact
- ▶ C = contact rate
- ▶ D = duration of infectiousness

Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

$$R_e(t) = \beta_t \cdot C_t \cdot D_t \cdot \hat{S}(t)$$

- ▶ β = probability of transmission per contact
- ▶ C = contact rate
- ▶ D = duration of infectiousness

Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

$$R_e(t) = \beta_t \cdot C_t \cdot D_t \cdot \hat{S}(t)$$

- ▶ β = probability of transmission per contact
- ▶ C = contact rate
- ▶ D = duration of infectiousness
- ▶ \hat{S} = proportion who are susceptible

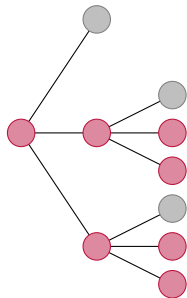
Basic R_0 and Effective $R_e(t)$

$$R_0 = \beta \cdot C \cdot D$$

$$R_e(t) = \beta_t \cdot C_t \cdot D_t \cdot \hat{S}(t)$$

- ▶ β = probability of transmission per contact
- ▶ C = contact rate
- ▶ D = duration of infectiousness
- ▶ \hat{S} = proportion who are susceptible

$$R_e(t) = 3 \cdot \frac{2}{3} = 2$$



R_0 for Common Diseases

Disease	R_0	Transmission
Measles	12–18	Aerosol
Chickenpox	10–12	Aerosol
Polio	5–7	Fecal-Oral
Smallpox	3–6	Droplet
HIV	1–5	Body Fluids
Seasonal Flu	1.5–2.5	Droplet
COVID-19	1.9–6.5 ¹	Droplet & Aerosol

- ▶ Without interventions
- ▶ Context-specific: $\beta \cdot C \cdot D$

¹Park et al. 2020

Applications of R

Applications of R

- ▶ Measure of epidemic “potential”

Applications of R

- ▶ Measure of epidemic “potential”
- ▶ Monitor interventions

Applications of R

- ▶ Measure of epidemic “potential”
- ▶ Monitor interventions
- ▶ Required vaccine coverage:

$$V = 1 - \frac{1}{R_0}$$

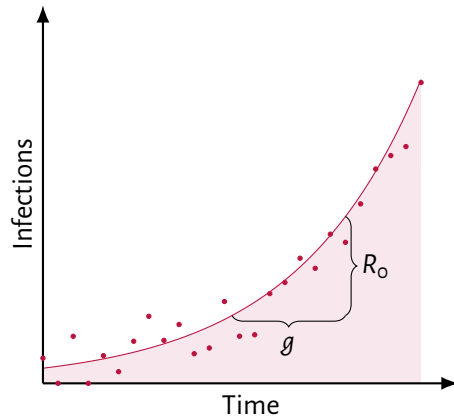
Estimating R : 3 Methods

- ▶ Static exponential: R_0
- ▶ Renewal equation: $R_e(t)$
- ▶ Mechanistic models (not covered)

Exponential Model for R_0

$$I(t) = I_0 \cdot e^{\alpha t}$$

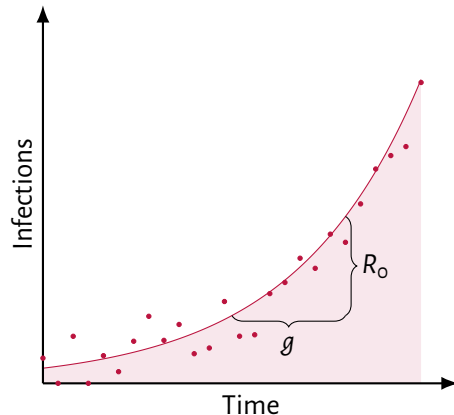
- ▶ I = infections
- ▶ α = exponential growth rate



Exponential Model for R_0

$$I(t) = I_0 \cdot e^{\alpha t}$$

- ▶ I = infections
- ▶ α = exponential growth rate
- ▶ g = generation time:
time between infections

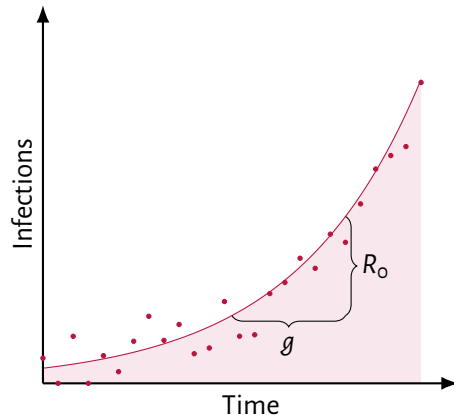


Exponential Model for R_0

$$I(t) = I_0 \cdot e^{\alpha t}$$

$$R_0 = \frac{I(t+g)}{I(t)} = e^{\alpha g}$$

- ▶ I = infections
- ▶ α = exponential growth rate
- ▶ g = generation time:
time between infections



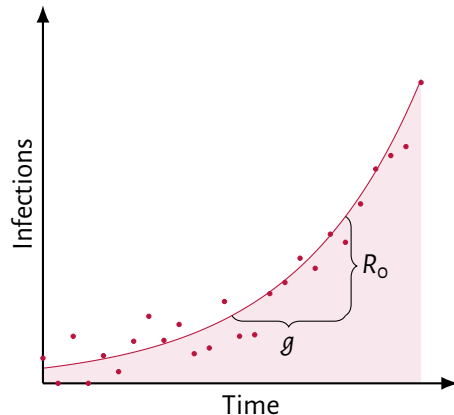
Exponential Model: Assumptions

Assumptions:

- ▶ Homogeneous susceptible population
- ▶ Constant R_0
- ▶ Reported cases \propto true cases

If longer generation time g :

- ▶ Expect slower epidemic growth
- ▶ Infer *greater* R_0



Renewal Equation for $R_e(t)$

Similar to exponential model, but with generation time *distribution*: $g \sim G(\tau)$

²Fraser 2007; Cori et al. 2013.

Renewal Equation for $R_e(t)$

Similar to exponential model, but with generation time *distribution*: $g \sim G(\tau)$

$$I(t) = R_e(t) \int_{t_0}^t I(t - \tau) G(\tau) d\tau$$

- ▶ $I(t - h)$ = infections τ days ago
- ▶ $G(\tau)$ = infectiousness τ days since infection
- ▶ $t_0 = t - w$, integration window

²Fraser 2007; Cori et al. 2013.

Renewal Equation for $R_e(t)$

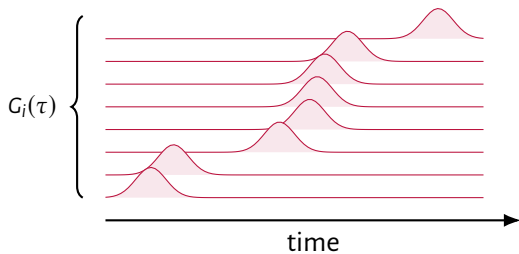
Similar to exponential model, but with generation time *distribution*: $g \sim G(\tau)$

$$I(t) = R_e(t) \int_{t_0}^t I(t - \tau) G(\tau) d\tau$$
$$R_e(t) = I(t) \left[\int_{t_0}^t I(t - \tau) G(\tau) d\tau \right]^{-1}$$

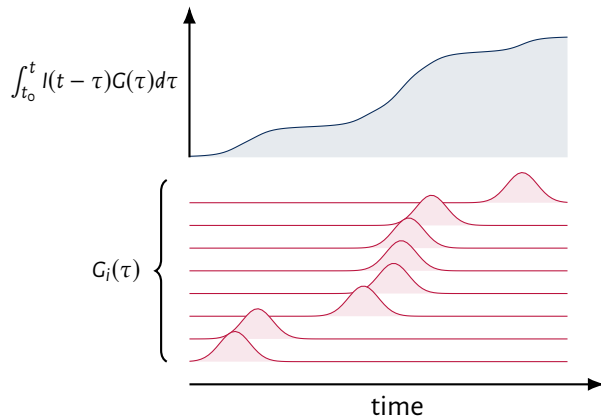
- ▶ $I(t - h)$ = infections τ days ago
- ▶ $G(\tau)$ = infectiousness τ days since infection
- ▶ $t_0 = t - w$, integration window

²Fraser 2007; Cori et al. 2013.

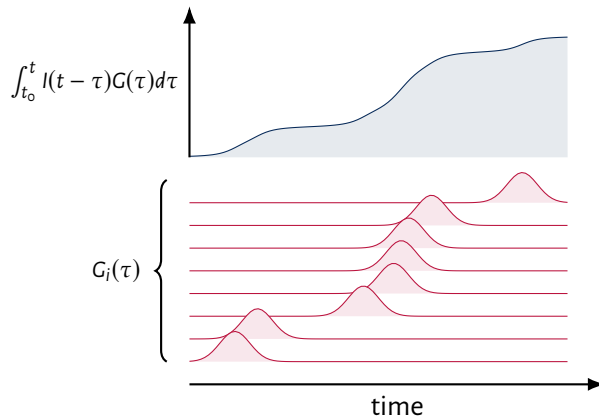
Renewal Equation: Illustrated



Renewal Equation: Illustrated



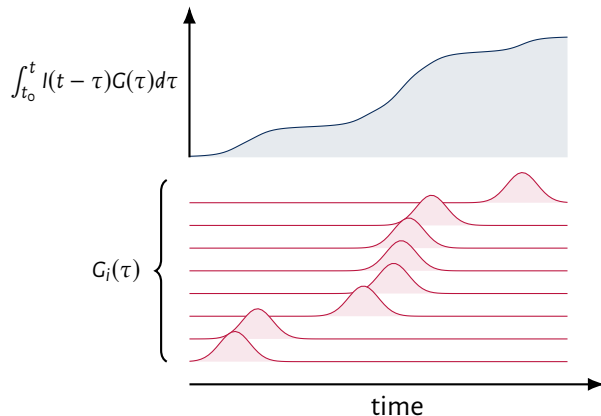
Renewal Equation: Illustrated



Assumptions:

- ▶ Same as exponential
- ▶ Constant R_e over each window

Renewal Equation: Illustrated

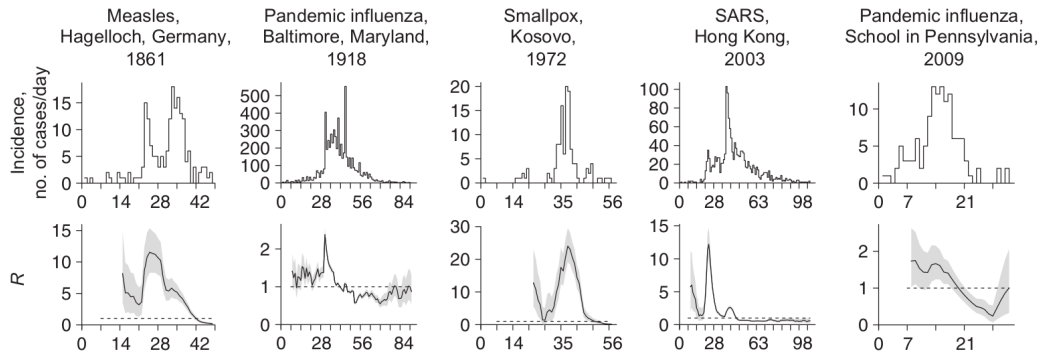


Assumptions:

- ▶ Same as exponential
- ▶ Constant R_e over each window

Moving window $\rightarrow R_e(t)$

$R_e(t)$: Examples



³Cori et al. 2013.

Infectiousness: Serial Interval $S(\tau)$

1. Known case-pairs:

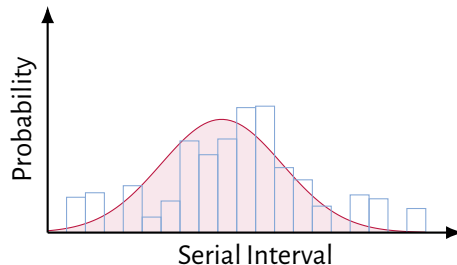
Pair	Symptom Onset	
	Infector	Infectee
1	2020-01-01	2020-01-07
2	2020-01-04	2020-01-14
3	2020-01-05	2020-01-09
4	2020-01-02	2020-01-15
5	2020-01-08	2020-01-12
6	2020-01-12	2020-01-23
⋮	⋮	⋮

Infectiousness: Serial Interval $S(\tau)$

1. Known case-pairs:

Pair	Symptom Onset	
	Infector	Infectee
1	2020-01-01	2020-01-07
2	2020-01-04	2020-01-14
3	2020-01-05	2020-01-09
4	2020-01-02	2020-01-15
5	2020-01-08	2020-01-12
6	2020-01-12	2020-01-23
⋮	⋮	⋮

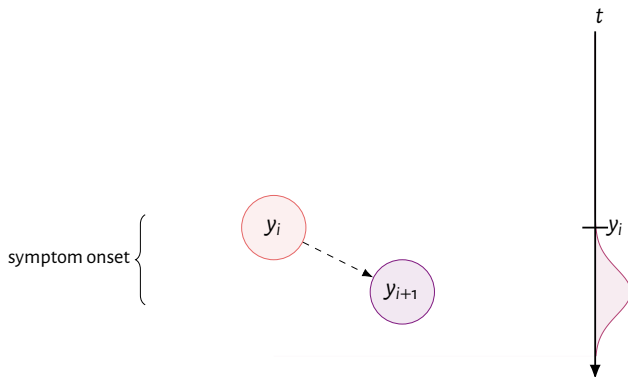
2. Fit a distribution:



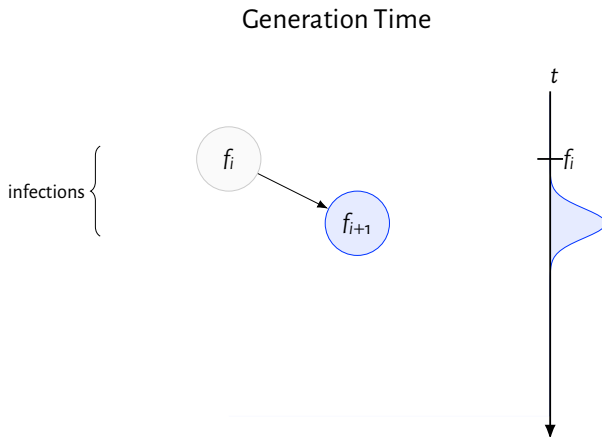
Infectiousness: Serial Interval vs Generation Time

Infectiousness: Serial Interval vs Generation Time

Serial Interval

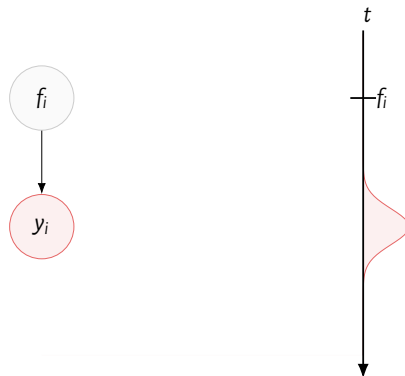


Infectiousness: Serial Interval vs Generation Time



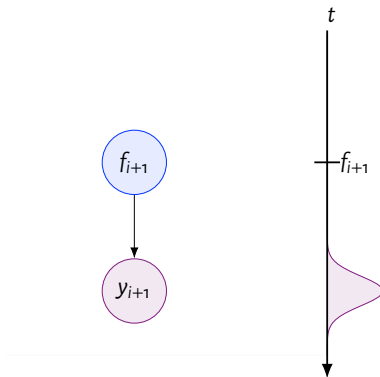
Infectiousness: Serial Interval vs Generation Time

Incubation Period i

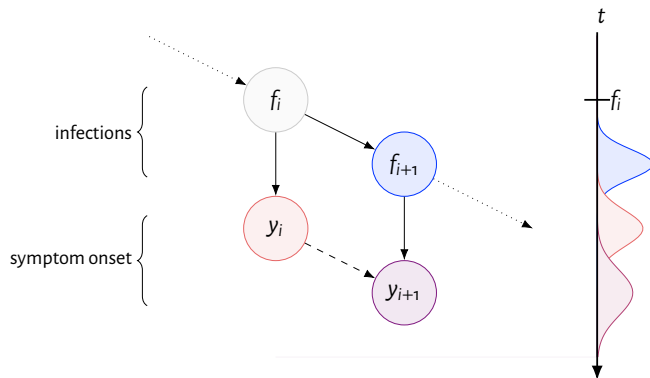


Infectiousness: Serial Interval vs Generation Time

Incubation Period $i + 1$



Infectiousness: Serial Interval vs Generation Time



Recovering the Generation Time Distribution

Random variables:

$$g_i = [f_{i+1} - f_i] \sim G(\tau)$$

$$s_i = [s_{i+1} - s_i] \sim S(\tau)$$

$$h_i = [s_i - f_i] \sim H(\tau)$$

⁴Hogg et al. 2005.

Recovering the Generation Time Distribution

Random variables:

$$\left. \begin{aligned} g_i &= [f_{i+1} - f_i] && \sim G(\tau) \\ s_i &= [s_{i+1} - s_i] && \sim S(\tau) \\ h_i &= [s_i - f_i] && \sim H(\tau) \end{aligned} \right\} s_i = g_i + h_{i+1} - h_i$$

⁴Hogg et al. 2005.

Recovering the Generation Time Distribution

Random variables:

$$\left. \begin{array}{ll} g_i = [f_{i+1} - f_i] & \sim G(\tau) \\ s_i = [s_{i+1} - s_i] & \sim S(\tau) \\ h_i = [s_i - f_i] & \sim H(\tau) \end{array} \right\} \begin{array}{l} s_i = g_i + h_{i+1} - h_i \\ \sim G(\tau) * H(\tau) * H(-\tau) \end{array}$$

where $*$ is convolution.

⁴Hogg et al. 2005.

Recovering the Generation Time Distribution

Random variables:

$$\left. \begin{aligned} g_i &= [f_{i+1} - f_i] && \sim G(\tau) \\ s_i &= [s_{i+1} - s_i] && \sim S(\tau) \\ h_i &= [s_i - f_i] && \sim H(\tau) \end{aligned} \right\} \begin{aligned} s_i &= g_i + h_{i+1} - h_i \\ &\sim G(\tau) * H(\tau) * H(-\tau) \end{aligned}$$

where $*$ is convolution. Thus:

$$G(\tau) = [S(\tau) *^{-1} H(\tau)] *^{-1} H(-\tau)$$

⁴Hogg et al. 2005.

Practical Deconvolution

- ▶ True deconvolution: unstable
- ▶ Solution: assume parametric $\hat{G}(\tau)$:

$$\hat{G}(\tau \mid \theta) \sim \text{Gamma}(\tau \mid \theta)$$

Practical Deconvolution

- ▶ True deconvolution: unstable
- ▶ Solution: assume parametric $\hat{G}(\tau)$:

$$\hat{G}(\tau \mid \theta) \sim \text{Gamma}(\tau \mid \theta)$$

- ▶ Estimate θ to minimize:

$$D_{KL}(S(\tau) \parallel \hat{S}(\tau \mid \theta)), \quad \hat{S}(\tau \mid \theta) = \hat{G}(\tau \mid \theta) * H(\tau) * H(-\tau)$$

Estimating $G(\tau)$: Data

What is the Generation Time distribution for COVID-19?

- ▶ Serial interval: Du et al. (2020), $N = 468$

$$S(\tau) \sim \text{Norm}(\mu = 3.96, \sigma = 4.75)$$

- ▶ Incubation period: Lauer et al. (2020), $N = 181$

$$H(\tau) \sim \text{Gamma}(\alpha = 5.81, \beta = 0.95)$$

COVID-19 & Pre-symptomatic Transmission

- ▶ Pre-symptomatic transmission $\rightarrow S(\tau)$ can be negative
- ▶ However: $G(\tau)$ must be non-negative

COVID-19 & Pre-symptomatic Transmission

- ▶ Pre-symptomatic transmission $\rightarrow S(\tau)$ can be negative
- ▶ However: $G(\tau)$ must be non-negative

Popular R package `EpiEstim`

- ▶ Estimates $R_e(t)$ by renewal equation
- ▶ Input: “serial interval”, but must be non-negative

Estimating $R_e(t)$: Data

What is the potential bias in $R_e(t)$ using $S(\tau)$ vs $G(\tau)$?

- ▶ $I(t)$: COVID-19 in GTA during March–April 2020

⁵Knight et al. 2020.

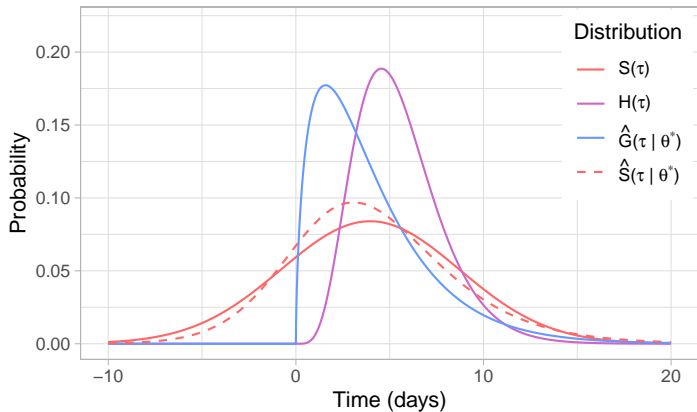
Estimating $R_e(t)$: Data

What is the potential bias in $R_e(t)$ using $S(\tau)$ vs $G(\tau)$?

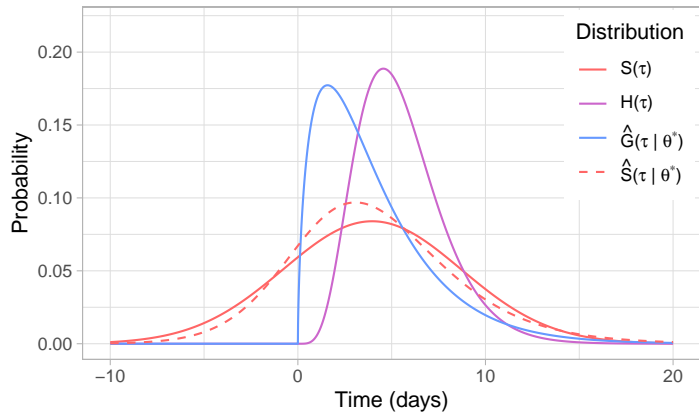
- ▶ $I(t)$: COVID-19 in GTA during March–April 2020
- ▶ $G(\tau)$: “recovered”
- ▶ $S(\tau)$: from literature:
 - ▶ Norm ($\mu = 3.96, \sigma = 4.75$) Du et al. (2020) Negative permitting
 - ▶ Gamma ($\alpha = 3.62, \beta = 1.42$) Zhang et al. (2020) Non-Negative
 - ▶ LogN ($\mu = 4.7, \sigma = 2.9$) Nishiura et al. (2020) Non-Negative

⁵Knight et al. 2020.

Recovered Generation Time Distribution



Recovered Generation Time Distribution

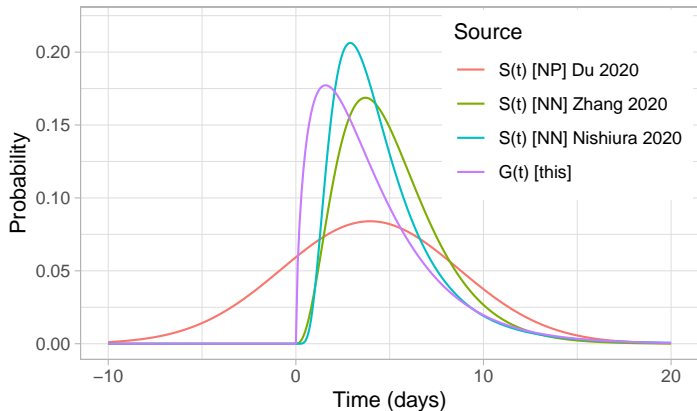


$$\hat{G}(\tau | \alpha, \beta) \sim \text{Gamma}(\alpha = 1.63, \beta = 2.50)$$

► mean: 4.08 \approx 3.98

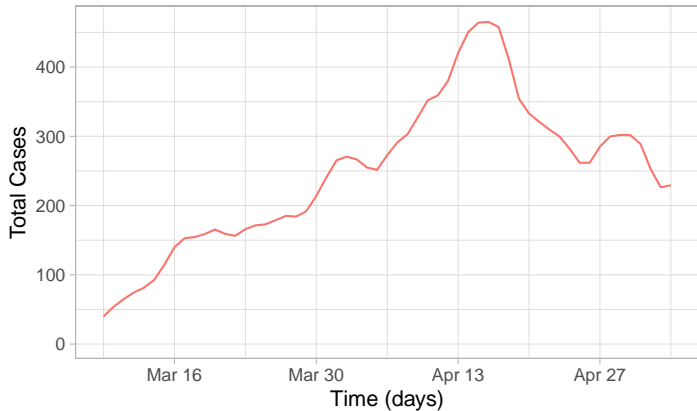
► SD: 3.19 $<$ 4.75

Generation Time vs Serial Interval Distributions

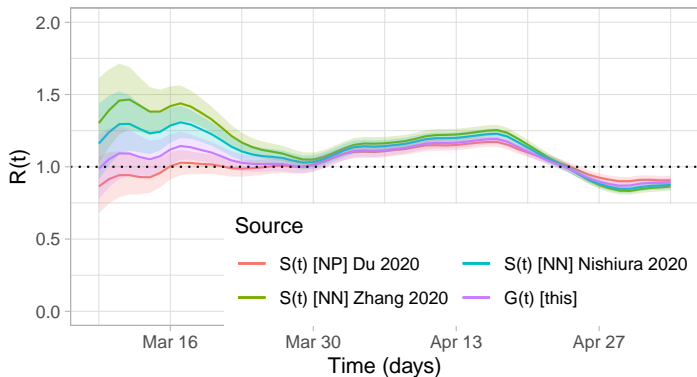


Ref.	Dist.	Mean	SD
Du	$S(\tau)$	3.96	4.75
Zhang	$S(\tau)$	5.1	2.7
Nishiura	$S(\tau)$	4.7	2.9
(this)	$G(\tau)$	4.08	3.19

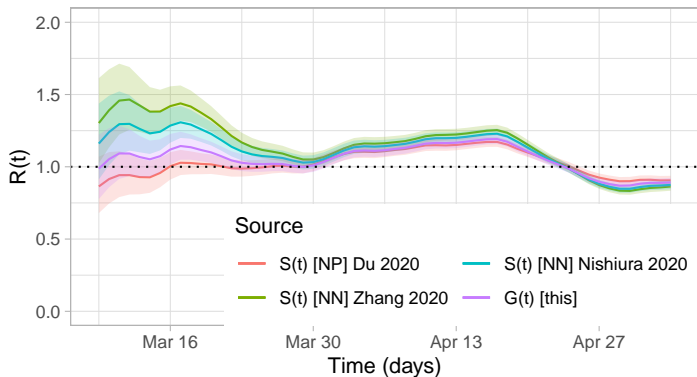
COVID-19 Infections $I(t)$ in GTA



$R_e(t)$ using $S(\tau)$ vs $G(\tau)$

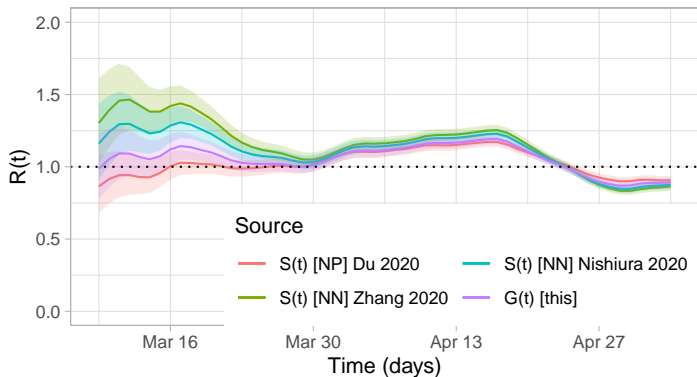


$R_e(t)$ using $S(\tau)$ vs $G(\tau)$



Compared to $G(\tau)$:

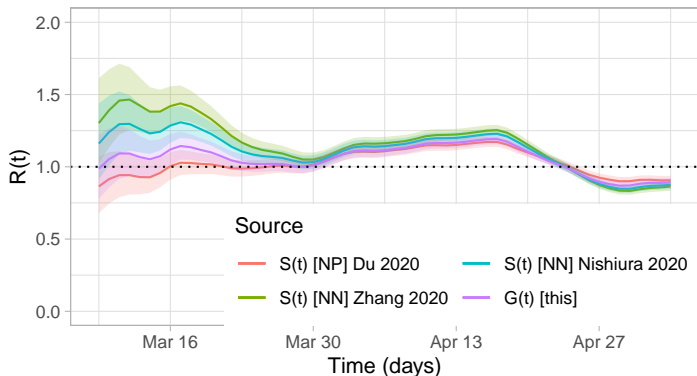
$R_e(t)$ using $S(\tau)$ vs $G(\tau)$



Compared to $G(\tau)$:

- **Non-Negative $S(\tau)$:**
overestimate $R_e(t)$
due to: larger mean

$R_e(t)$ using $S(\tau)$ vs $G(\tau)$



Compared to $G(\tau)$:

- ▶ **Non-Negative** $S(\tau)$:
overestimate $R_e(t)$
due to: larger mean
- ▶ **Negative-Permitting** $S(\tau)$:
underestimate $R_e(t)$
due to: larger variance

Implications

Implications

1. Characterized generation time distribution for COVID-19

- Generation time = distribution of infectiousness (after exposure)

Implications

1. Characterized generation time distribution for COVID-19

- ▶ Generation time = distribution of infectiousness (after exposure)

2. New method to infer $G(\tau)$ from $S(\tau)$

- ▶ Applicable to other infections

Implications

1. Characterized generation time distribution for COVID-19

- ▶ Generation time = distribution of infectiousness (after exposure)

2. New method to infer $G(\tau)$ from $S(\tau)$

- ▶ Applicable to other infections

3. Illustrated potential bias in $R_e(t)$ using $S(\tau)$ vs $G(\tau)$

- ▶ Non-negative $S(\tau)$: overestimate $R_e(t)$
- ▶ Negative-permitting $S(\tau)$: underestimate $R_e(t)$

Limitations

Limitations

- ▶ Assumed independent distributions $G(\tau)$, $S(\tau)$, $H(\tau)$

Limitations

- ▶ Assumed independent distributions $G(\tau)$, $S(\tau)$, $H(\tau)$
- ▶ Parametric approximations
However: useful for meta-analyses

Limitations

- ▶ Assumed independent distributions $G(\tau)$, $S(\tau)$, $H(\tau)$
- ▶ Parametric approximations
However: useful for meta-analyses
- ▶ No uncertainty analysis / confidence intervals

Thanks



Canada Graduate Scholarship—D



Ontario Early Researcher Award



2020 COVID-19 Research Award

Support

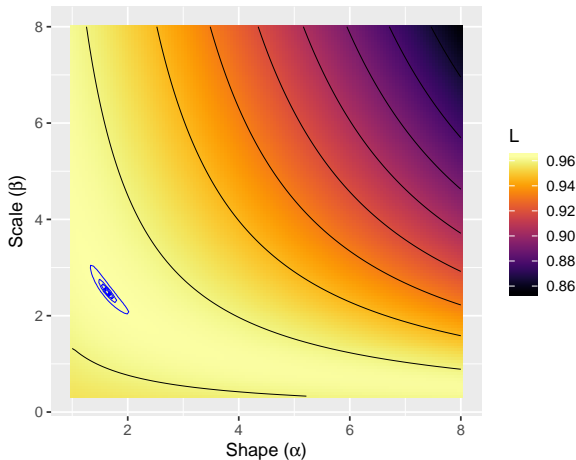


Kristy Yiu, Linwei Wang, Huiting Ma, David Landsman, *Unity Health Toronto*
David Champredon, *University of Western Ontario*
Elena Aruffo, Jude Dzevela Kong, *CAIMS 2021 Mini Symposium Organizers*

References

- Cori, Anne et al. (Nov. 2013). "A new framework and software to estimate time-varying reproduction numbers during epidemics". In: *American Journal of Epidemiology* 178.9, pp. 1505–1512.
- Du, Zhanwei et al. (Mar. 2020). "Serial Interval of COVID-19 among Publicly Reported Confirmed Cases". In: *Emerging infectious diseases* 26.6, pp. 1341–1343.
- Fraser, Christophe (Aug. 2007). "Estimating individual and household reproduction numbers in an emerging epidemic". In: *PLoS ONE* 2.8, e758.
- Hogg, Robert V, Joseph McKean, and Allen T Craig (2005). *Introduction to mathematical statistics*. Pearson Education.
- Knight, Jesse and Sharmistha Mishra (2020). "Estimating effective reproduction number using generation time versus serial interval, with application to COVID-19 in the Greater Toronto Area, Canada". In: *Infectious Disease Modelling* 5, pp. 889–896.
- Lauer, Stephen A. et al. (Mar. 2020). "The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application". In: *Annals of Internal Medicine* 172.9, pp. 577–582.
- Nishiura, Hiroshi, Natalie M. Linton, and Andrei R. Akhmetzhanov (Apr. 2020). "Serial interval of novel coronavirus (COVID-19) infections". In: *International Journal of Infectious Diseases* 93, pp. 284–286.
- Park, Minah et al. (2020). "A Systematic Review of COVID-19 Epidemiology Based on Current Evidence". In: *Journal of Clinical Medicine* 9.4, p. 967.
- Zhang, Juanjuan et al. (2020). "Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study". In: *The Lancet Infectious Diseases* 20.7, pp. 793–802.

Preliminary Uncertainty Analysis



Likelihood of α, β for $\hat{G}(\tau)$
from "deconvolution"