Strength, speed and quickness: an approach to evaluating models of disease spread

Jonathan Dushoff

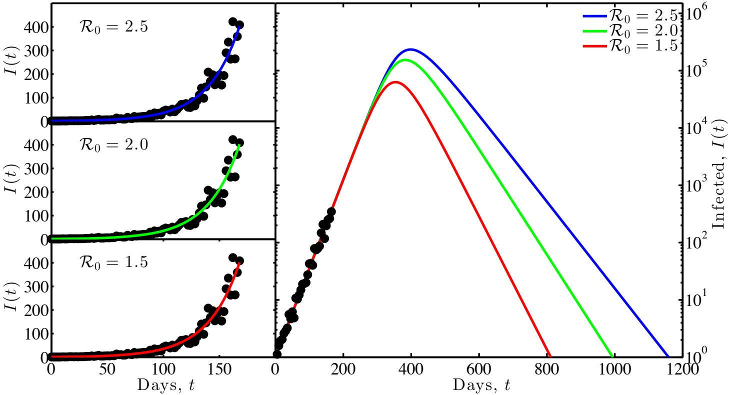
Mathematics of COVID-19
Fields Institute
Feb 2020

Outline

Speed: r

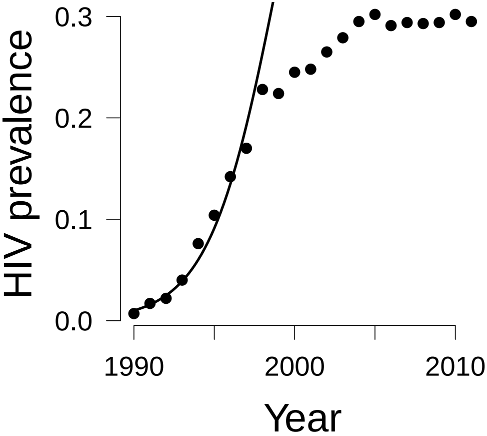
- We measure epidemic speed using little r:
  - The ratio of the *change* in disease impact to the *amount* of disease impact
  - *Units*: [1/time]
  - Disease increases like e^{rt}
- Time scale is $C = 1/r$

Ebola outbreak



$C \approx 1\,\mathrm{month}$. Sort-of fast.

HIV in sub-Saharan Africa



$C \approx 18\,\mathrm{month}$. Horrifyingly fast.

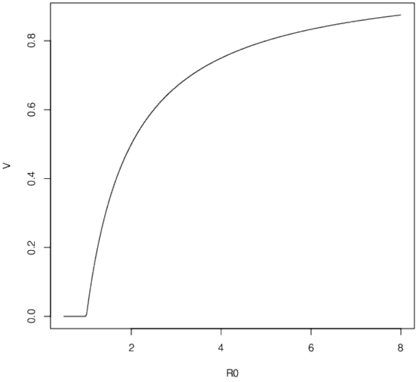
Strength: $\mathcal{R}$

- We describe epidemic strength with big $\mathcal{R}$
- Number of potential new cases per case
  - Not accounting for proportion susceptible
- To eliminate disease, we must:
  - Reduce effective reproduction by a factor of $\mathcal{R}$

$\mathcal{R}$ and equilibrium

- If we have $\mathcal{R}$ new cases per case when everyone is susceptible
- And 1 case per case (on average) at equilibrium:
  - Proportion susceptible at equilibrium is $S = 1/\mathcal{R}$
  - Proportion affected at equilibrium is $V = 1 - 1/\mathcal{R}$

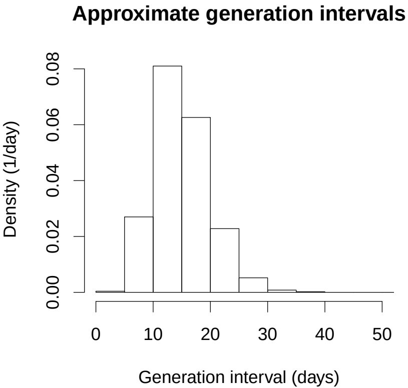
$\mathcal{R}$ and control



Coronavirus

- What we see clearly is r
- What we rush to calculate is $\mathcal{R}$
- How do we do this?
- *Why* do we do this?

Quickness: $g(\tau)$



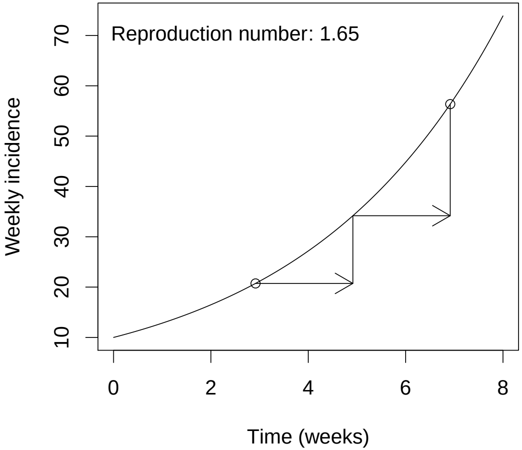
- The generation distribution measures generations of the disease
  - Interval between "index" infection and resulting infection
- Do fast disease generations mean more danger or less danger?

Outline

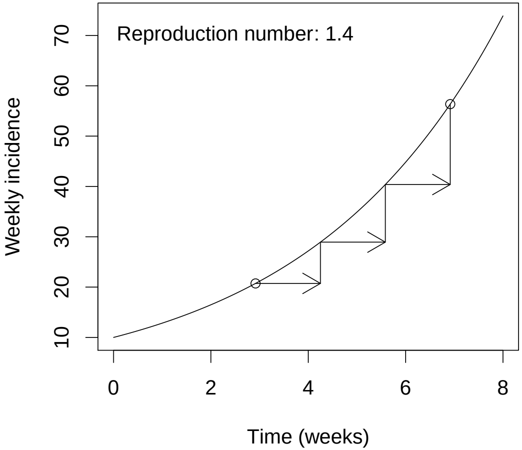
Conditional effect of quickness

- *Given* the reproductive number $\mathcal{R}$
  - quicker disease means faster growth rate r
  - More danger
- *Given* the growth rate r
  - quicker disease means *smaller* $\mathcal{R}$
  - Less danger

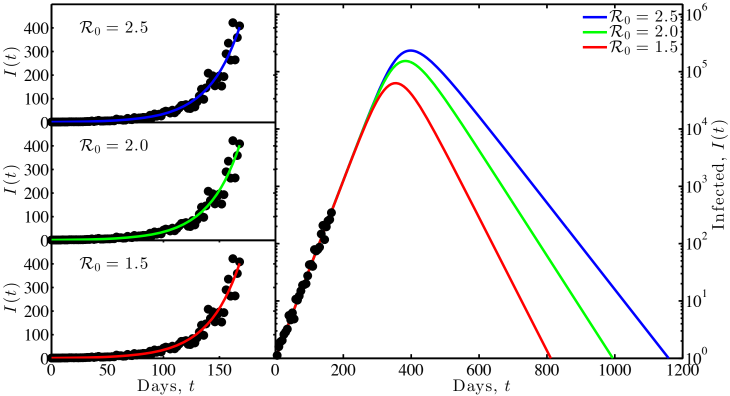
Generations and $\mathcal{R}$



Generations and $\mathcal{R}$

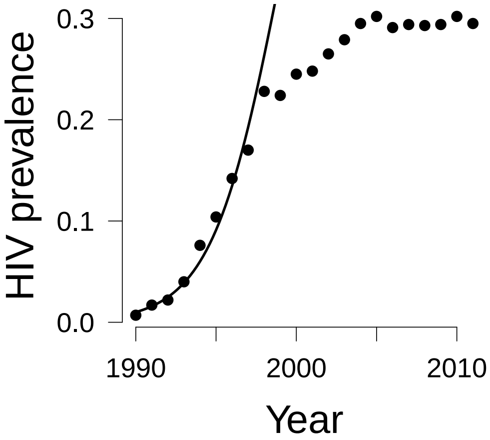


Ebola outbreak



$C \approx 1\,\text{month},\ G \approx 2\,\text{week}$

HIV in sub-Saharan Africa



$C \approx 18\,\mathrm{month},\ G \approx 4\,\mathrm{years}$

Linking framework

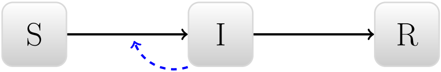
- Epidemic speed (r) is a *product*:
  - quickness $\times$
  - epidemic strength
- WRONG

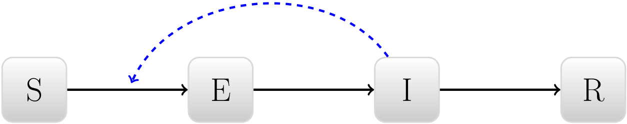
Linking framework

- Epidemic speed (r) is a *product*:
  - (something to do with) quickness $\times$
  - (something to do with) epidemic strength
- Strength ($\mathcal{R}$) is therefore (sort-of) a quotient
  - More quickness implies less strength
  - ...if speed is known

Outline

Box models



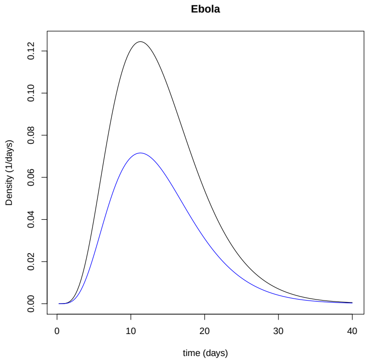


Renewal equation

- A broad framework that covers a wide range of underlying models
-
$$i(t) = S(t) \int k(\tau) i(t - \tau)\, d\tau$$
  - $i(t)$ is the *rate* of new infections (per-capita incidence)
  - $S(t)$ is the proportion of the population susceptible
  - $k(\tau)$ measures how infectious a person is (on average) at time τ after becoming infected
- For invasion, treat S as constant

Infection kernel

- $k(\tau)$ is the expected rate at which you infect at time τ after being infected
- $\int_\tau k(\tau)d\tau$ is the expected number of people infected:
  - $\mathcal{R}$ the effective reproductive number
- $k(\tau)/\mathcal{R}$ is a distribution:
  - $g(\tau)$, the *intrinsic* generation distribution

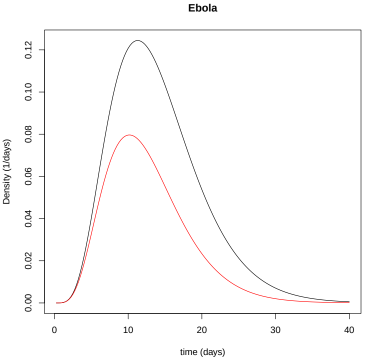


Euler-Lotka equation

- If we neglect S, we expect exponential growth
- $1 = \int k(\tau) \exp(-r\tau)\, d\tau$
  - i.e., the total of *discounted* contributions is 1
- $1/\mathcal{R} = \int g(\tau) \exp(-r\tau)\, d\tau$
- Note that $b(\tau) = k(\tau) \exp(-r\tau)$ is also a distribution
  - The initial "backwards" generation interval

Interpretation: generating functions

- $1/\mathcal{R} = \int g(\tau) \exp(-r\tau)\, d\tau$
- *J Wallinga, M Lipsitch; DOI: 10.1098/rspb.2006.3754*



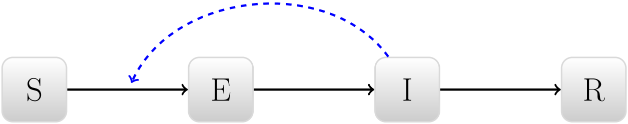
Interpretation: “effective” generation times

- *Define* $\hat{G}$
  - $$\mathcal{R} = \exp(r\hat{G})$$
- Then:
  - $$1/\mathcal{R} = \int g(\tau)\exp(-r\tau)\,d\tau$$
  - $$\exp(-r\hat{G}) = \langle \exp(-r\tau) \rangle_g .$$
  - A filtered mean:
    - The discounted value of $\hat{G}$ is the expectation of the discounted values across the distribution

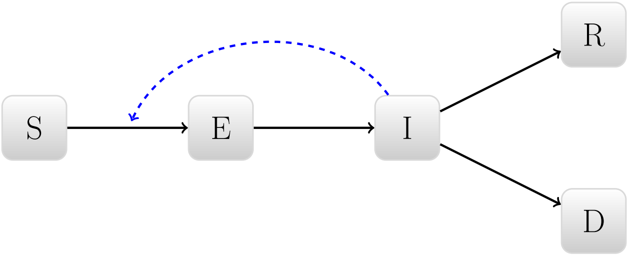
Example: Post-death transmission and safe burial

- How much Ebola spread occurs before vs. after death
- Highly context dependent
  - Funeral practices, disease knowledge
- *Weitz and Dushoff Scientific Reports 5:8751.*

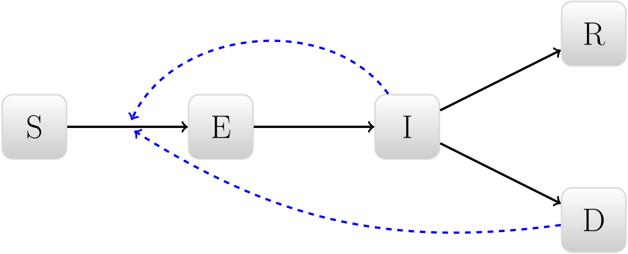
Standard disease model



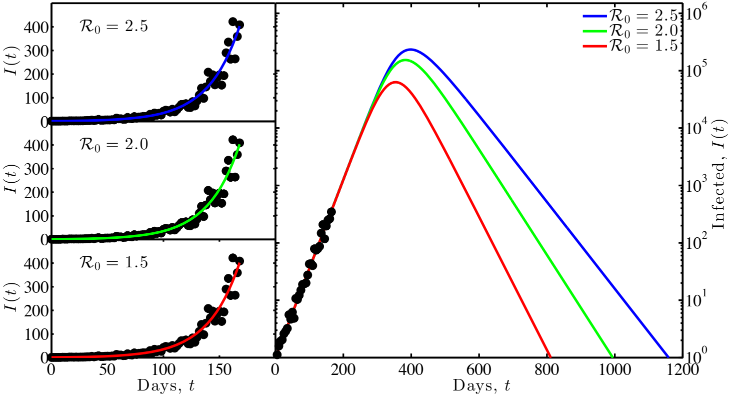
Disease model including post-death transmission



Disease model including post-death transmission



Scenarios



$\mathcal{R}_0 = 2.5$
$\mathcal{R}_0 = 2.0$
$\mathcal{R}_0 = 1.5$
$I(t)$
0
100
200
300
400
0
50
100
150
Days, t
200
400
600
800
1000
1200
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Infected, $I(t)$

Conclusions

- Different parameters can produce indistinguishable early dynamics
- More after-death transmission implies
  - Higher $\mathcal{R}_0$
  - Larger epidemics
  - Larger importance of safe burials

Outline

Limitations of effective time

- The filtered mean has nice theoretical and intuitive properties
- Practically, the effective generation time can be confusing
- How is
  - $\mathcal{R} = \exp(r\hat{G})$
- Consistent with the result from ODEs
  - $\mathcal{R} = 1 + r\bar{G}$?
- $\hat{G}$ changes with r, sometimes a lot

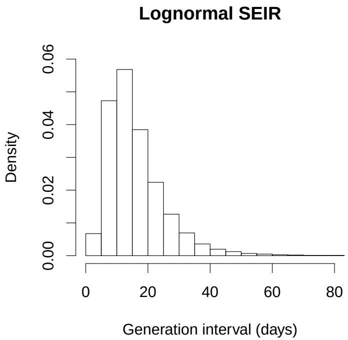
Gamma approximation

- If g has a gamma distribution, then:
- $\mathcal{R} \approx (1 + r\kappa\bar{G})^{1/\kappa}$
- κ is the *dispersion* (the squared coefficient of variation of the generation distribution)
- How good is the approximation?
- *Park et al., Epidemics DOI:10.1101/312397*

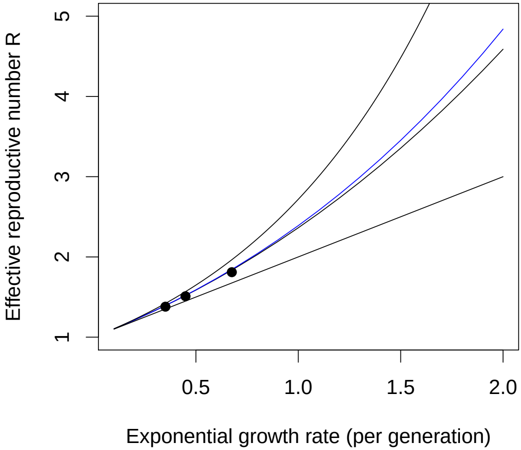
Fitting to Ebola

- Simulate generation intervals based on data and approach from WHO report
- Use both lognormals and gammas
  - WHO used gammas
  - Lognormals should be more challenging

Approximating the distribution



Approximating the curve



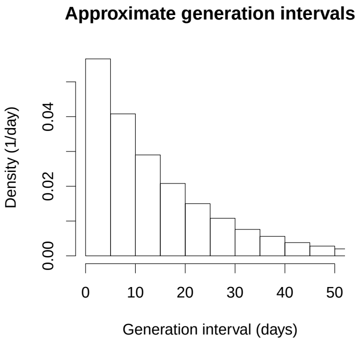
Effective dispersion

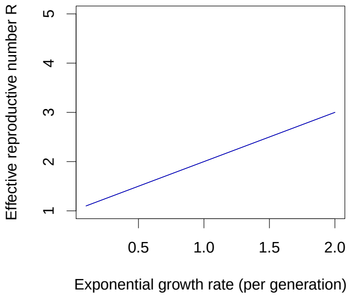
- *Define* $\bar{\kappa}$:
- $\mathcal{R} = (1 + r\hat{\kappa}\bar{G})^{1/\hat{\kappa}} \equiv X(r\bar{G}; 1/\hat{\kappa})$
- Conceptually more confusing than effective generation time
- Practically straightforward
  - For *many* applications $\bar{\kappa}$ changes very little with r
  - Doesn't work where you wouldn't expect it to:
    - syphilis, sexual transmission of Ebola

Compound-interest interpretation

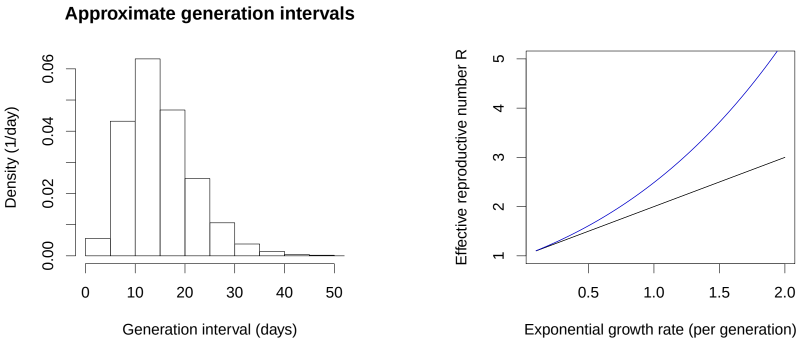
- $\mathcal{R} = (1 + r\kappa\bar{G})^{1/\kappa} \equiv X(r\bar{G}; 1/\kappa)$
- X is the compound-interest approximation to the exponential
  - Linear when $\kappa = 1$ (i.e., when g is exponential)
  - Approaches exponential as $\kappa \to 0$

Gamma approximation

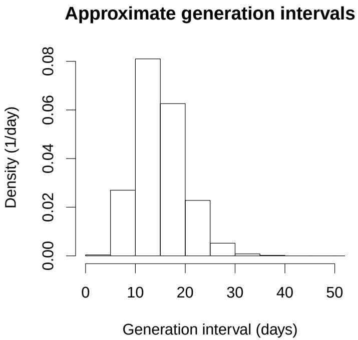


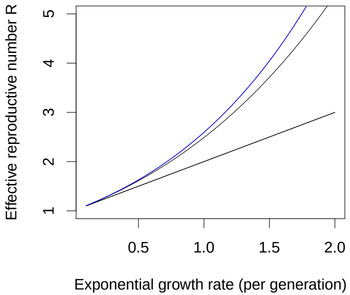


Gamma approximation

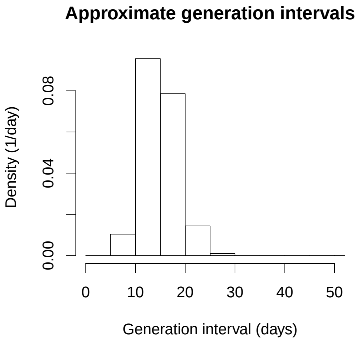


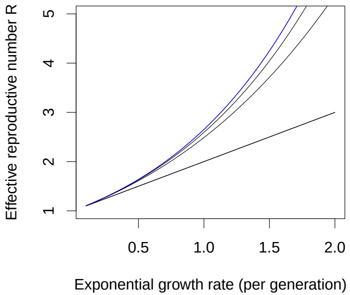
Gamma approximation





Gamma approximation





Qualitative response

- For a given value of $\bar{G}$, smaller values of κ mean:
  - less variation in generation interval
  - less compounding of growth
  - greater $\mathcal{R}$ required for a given r

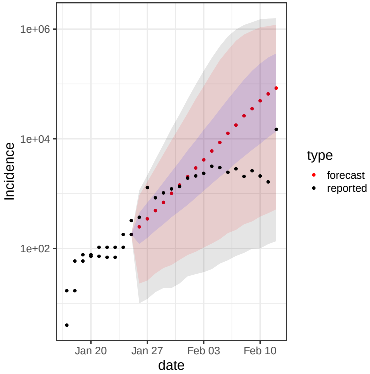
Linking framework

- Epidemic speed (r) is a *product*:
  - (something to do with) quickness $\times$
  - (something to do with) epidemic strength
- In particular:
  - $r \approx (1/\bar{G}) \times \ell(\mathcal{R}; \bar{\kappa})$
  - ℓ is the inverse of X

Evaluating

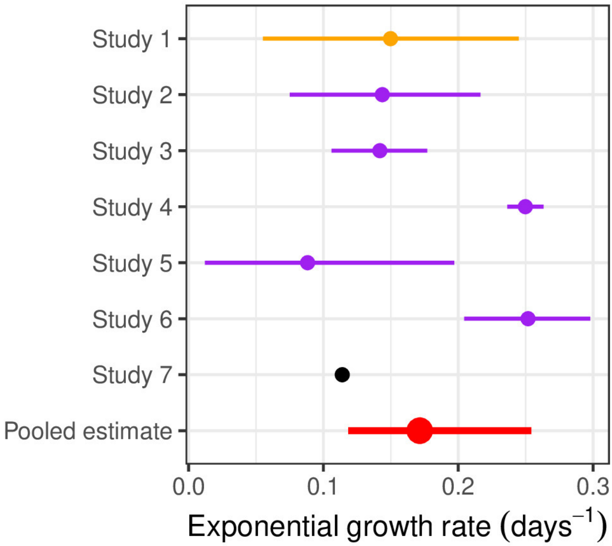
- Model fits to exponential case data are essentially estimating $\mathcal{R}$ using this quotient
- Can be evaluated and compared using (implicit or explicit) estimates of r, $\bar{G}$ and $\bar{\kappa}$

Park et al., DOI: 10.1101/2020.01.30.20019877 (preprint)

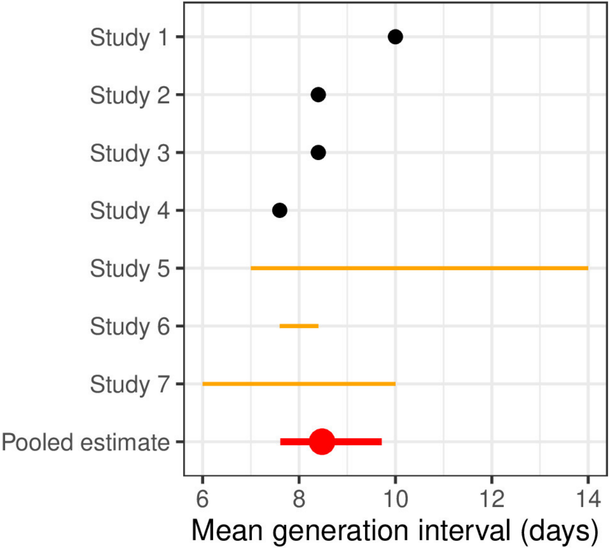


- *Mike Li*

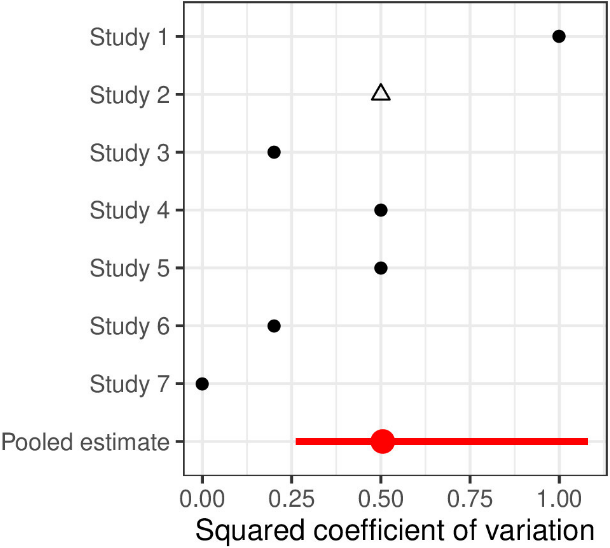
Assumptions



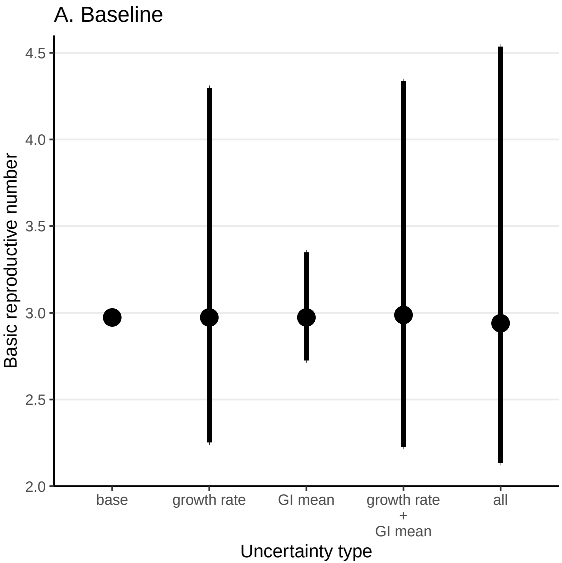
Assumptions



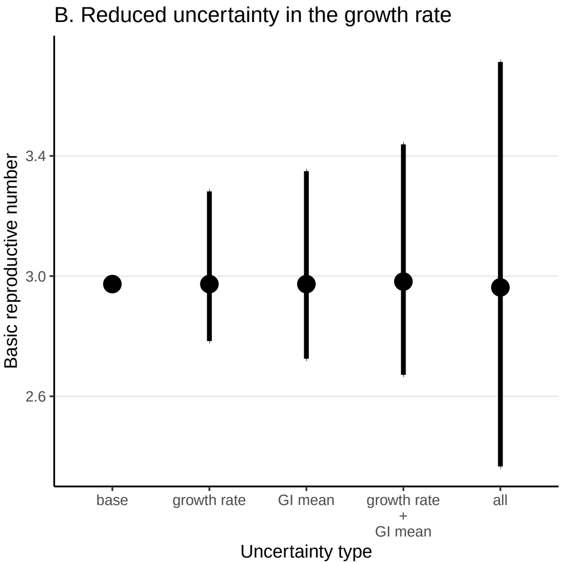
Assumptions



Propagating error



Propagating error



Outline

A false dichotomy

- Why are people scrambling to estimate $\mathcal{R}$ and mostly ignoring r?
  - History
  - Modelers gotta model

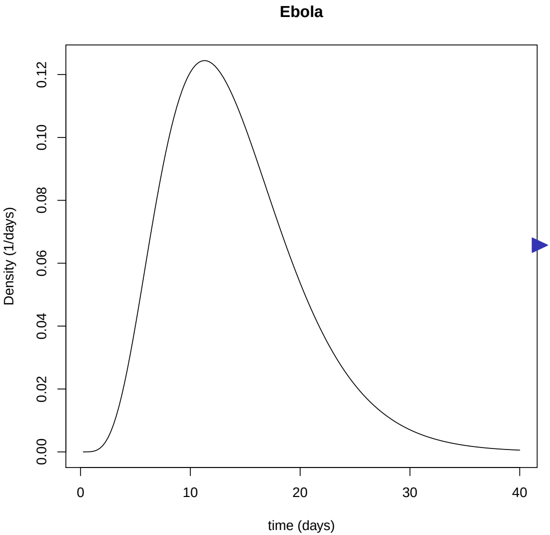
The strength paradigm

- $\mathcal{R} > 1$ is a threshold
- If we can reduce transmission by a constant *factor* of $\theta > \mathcal{R}$, disease can be controlled
- In general, we can define θ as a (harmonic) mean of the reduction factor over the course of an infection
  - weighted by the *intrinsic* generation interval
- Epidemic is controlled if $\theta > \mathcal{R}$
- More useful in long term (tells us about final size, equilibrium)

The speed paradigm

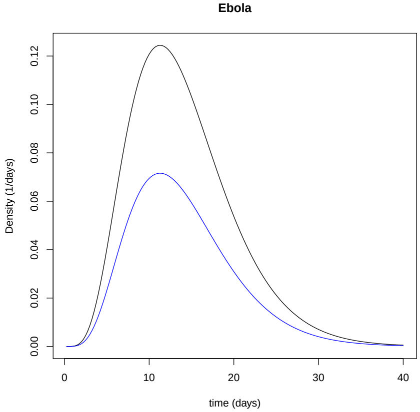
- $r > 0$ is a threshold
- If we can reduce transmission at a constant *hazard rate* of $\phi > r$, disease can be controlled
- In general, we can define ϕ as a (very weird) mean of the reduction factor over the course of an infection
  - weighted by the *backward* generation interval
- Epidemic is controlled if $\phi > r$
- More useful in short term (tells us about, um, speed)

Epidemic strength (present)



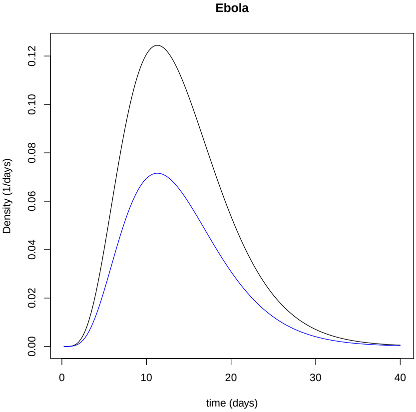
- $\mathcal{R}$, the epidemic strength, is the area under the curve.

Strength of intervention



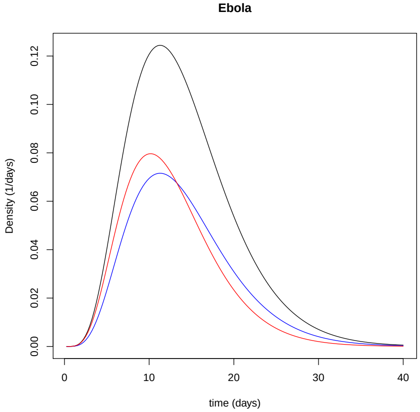
- . . . by what factor do I need to reduce this curve to eliminate the epidemic?

Different interventions (present)



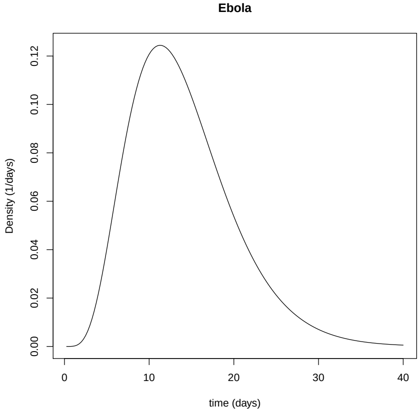
- idealized vaccination
- removes a fixed proportion of people

Different interventions (present)



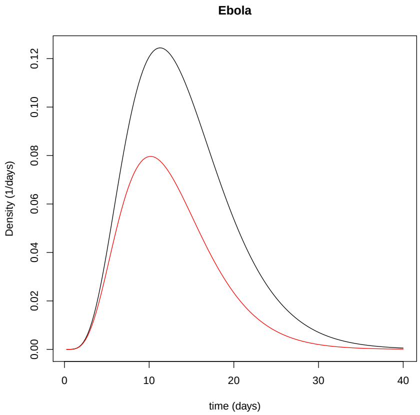
- idealized quarantine
- removes people at a fixed rate

Epidemic speed



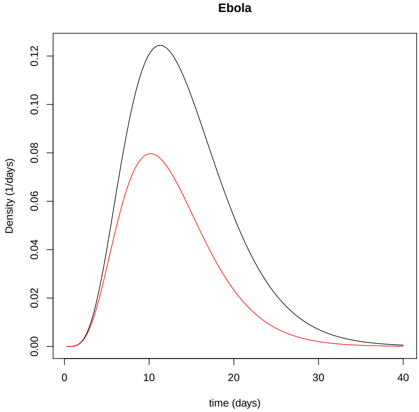
- r, the epidemic speed, is the "discount" rate required to balance the tendency to grow

Epidemic speed



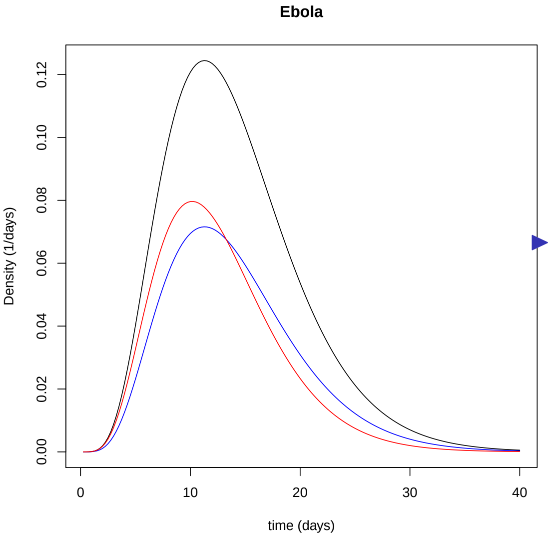
- $k(\tau) = \exp(r\tau)b(\tau)$, where $b(\tau)$ is the initial backward generation interval

Speed of intervention



- . . . how *quickly* do I need to reduce this curve to eliminate the epidemic?

Different interventions (present)



- Sometimes it's easier to estimate strength, sometimes speed

Measuring the intervention

Measuring the intervention

- We imagine an intervention with potentially variable effect over the course of infection, $L(\tau)$
- Assume the intervention takes
  - $k(\tau) \to \hat{k}(\tau) = k(\tau)/L(\tau)$

Measuring intervention strength

- Define intervention *strength* $\theta = \mathcal{R}/\hat{\mathcal{R}}$ – the proportional amount by which the intervention reduces transmission.

- $\theta = 1/\left\langle 1/L(\tau)\right\rangle_{g(\tau)}$

- θ is *the harmonic mean* of L, weighted by the generation distribution g.

- Outbreak can be controlled if $\theta > \mathcal{R}$

Measuring intervention speed

- Define intervention *speed* $\phi = r - \hat{r}$ – the amount by which the intervention slows down spread.
- We then have:
  - $1 = \left\langle \frac{\exp(\phi\tau)}{L(\tau)} \right\rangle_{b(\tau)}$
- ϕ is *sort of a mean* of the *hazard* associated with L
  - Because $L(t) = \exp(ht)$ when hazard is constant
- Averaged over the initial *backwards* generation interval
- Outbreak can be controlled if $\phi > r$.

The strength paradigm

- $k(\tau) = \mathcal{R} g(\tau)$
  - g is the intrinsic generation interval
  - $\mathcal{R}$ is the strength of the epidemic
- If $L(\tau) \equiv L$, then $\theta = L$ is the strength of the intervention
- In general, θ is a (harmonic) mean of L
  - weighted by $g(\tau)$, but not affected by $\mathcal{R}$.
- Epidemic is controlled if $\theta > \mathcal{R}$

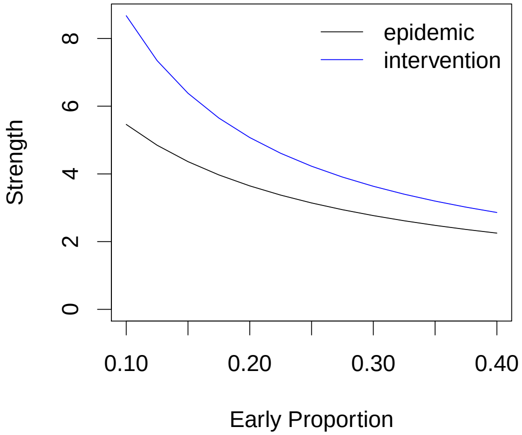
The speed paradigm

- $k(\tau) = \exp(r\tau)b(\tau)$,
  - r is the speed of the epidemic
  - b is the initial backward generation interval
- If $h(\tau) \equiv h$, then $\phi = h$ is the speed of the intervention
- In general, ϕ is a (weird) mean of h
  - weighted by $b(\tau)$, but not affected by r.
- Epidemic is controlled if $\phi > r$

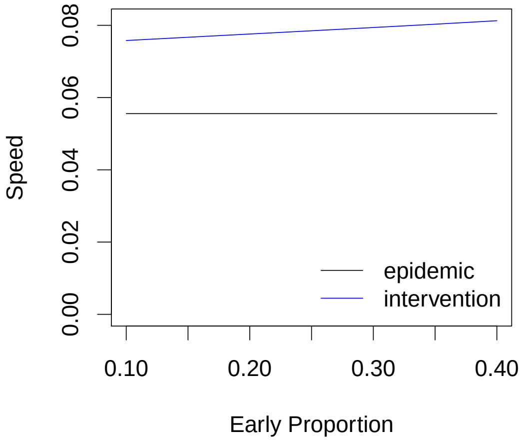
HIV

- The importance of transmission speed to HIV control is easier to understand using the speed paradigm
  - We know the speed of invasion
    - $\approx 0.7/\mathrm{yr}$
    - Characteristic scale $\approx 1.4\mathrm{yr}$
  - And can hypothesize the speed of intervention
    - Or aim to go fast enough

HIV test and treat



HIV test and treat



Coronavirus outbreak

- What we know well is the speed of the outbreak
- What do we think if the pathogen is actually quicker than we thought?
  - e.g., Nishiura et al.
- It has less strength (easier to control by vaccination)
- Control by identification and isolation may be a little harder?

Outline

Measuring generation intervals

- Ad hoc methods
- Error often not propagated
- Importance of heterogeneity

Generations through time

- Generation intervals can be estimated by:
  - Observing patients:
    - How long does it take to become infectious?
    - How long does it take to recover?
    - What is the time profile of infectiousness/activity?
  - Contact tracing
    - Who (probably) infected whom?
    - When did each become infected?
    - — or ill (serial interval)?

Which is the real interval?

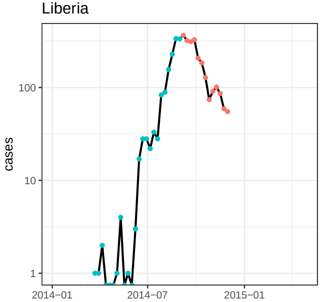
- Contact-tracing intervals look systematically different, depending on when you observe them.
- Observed in:
  - Real data, detailed simulations, simple model
- Also differ from intrinsic (infector centered) estimates

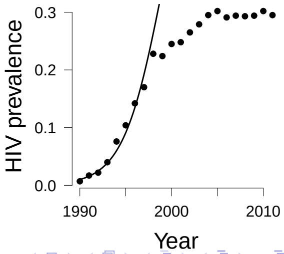
Types of interval

- Define:
  - *Intrinsic interval*: How infectious is a patient at time τ after infection?
  - *Forward interval*: When will the people infected today infect others?
  - *Backward interval*: When did the people who infected people today themselves become infected?
  - *Censored interval*: What do all the intervals observed up until a particular time look like?
    - Like backward intervals, if it's early in the epidemic

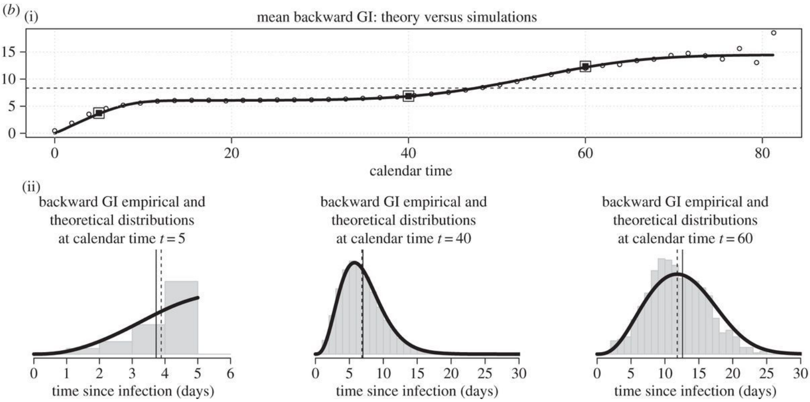
Growing epidemics

- Generation intervals look *shorter* at the beginning of an epidemic
  - A disproportionate number of people are infectious right now
  - They haven't finished all of their transmitting
  - We are biased towards observing faster events





Backward intervals

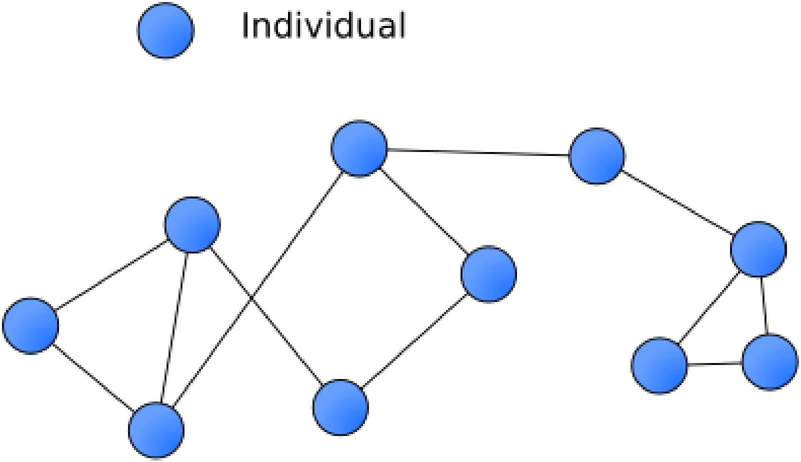


Champredon and Dushoff, 2015. DOI:10.1098/rspb.2015.2026

Outline

Generations in space

- How do local interactions affect realized generation intervals?



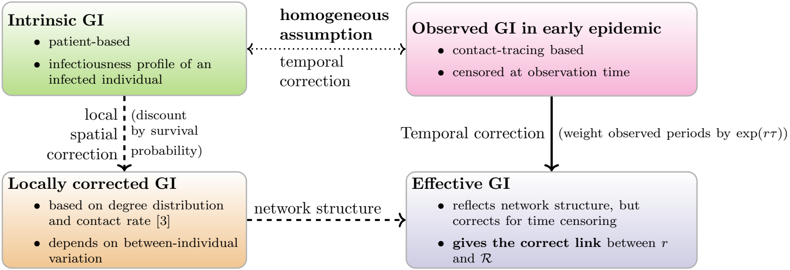
Surprising results

- We tend to think that heterogeneity leads to underestimates of $\mathcal{R}$, which can be dangerous.
- $\mathcal{R}$ on networks generally *smaller* than values estimated using r.
  - *Trapman et al., 2016. JRS Interface DOI:10.1098/rsif.2016.0288*

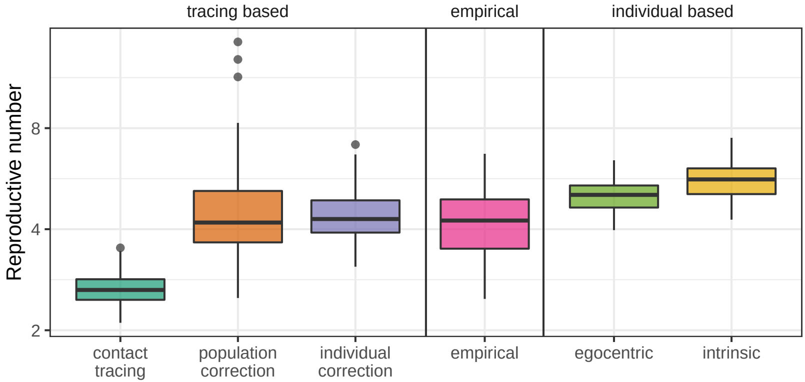
Generation-interval perspective

- Modelers don't usually question the intrinsic generation interval
- But spatial network structure does change generation intervals:
  - Local interactions
  - $\implies$ wasted contacts
  - $\implies$ shorter generation intervals
  - $\implies$ smaller estimates of $\mathcal{R}$.

Observed and estimated intervals



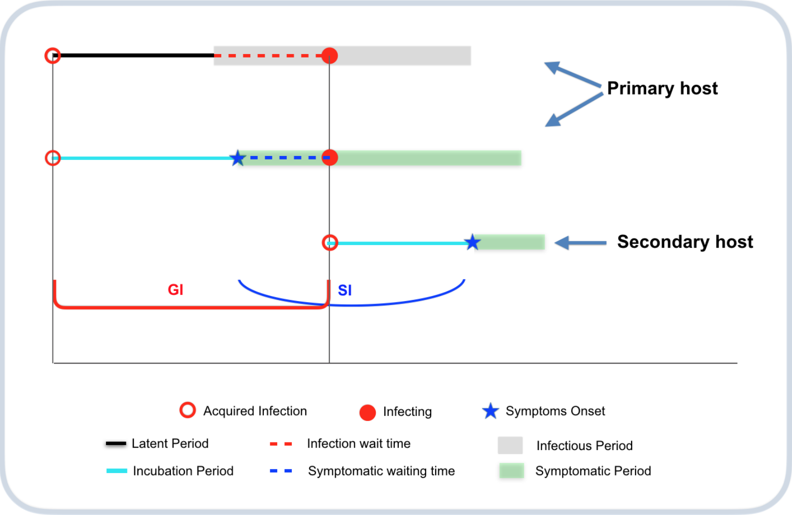
Outbreak estimation



Park et al. DOI: 10.1101/683326 (preprint)

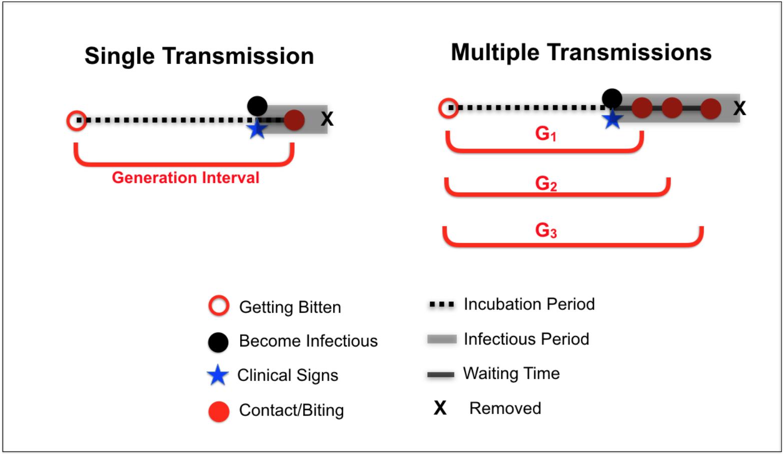
Outline

Serial intervals



Serial intervals

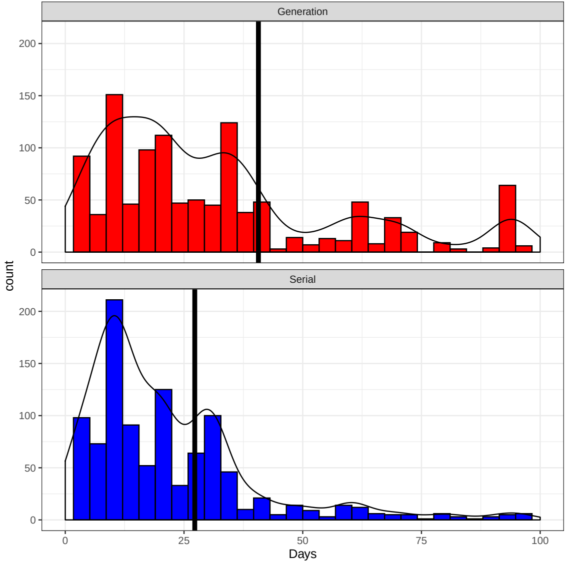
- Do serial intervals and generation intervals have the same distribution?
- It seems that they should: they describe generations of the same process
  - But serial intervals can even be very different
  - Even negative! You might report to the clinic with flu before me, even though I infected you
- For rabies, we *thought* that serial intervals and generation intervals should be the same
  - Symptoms are closely correlated with infectiousness



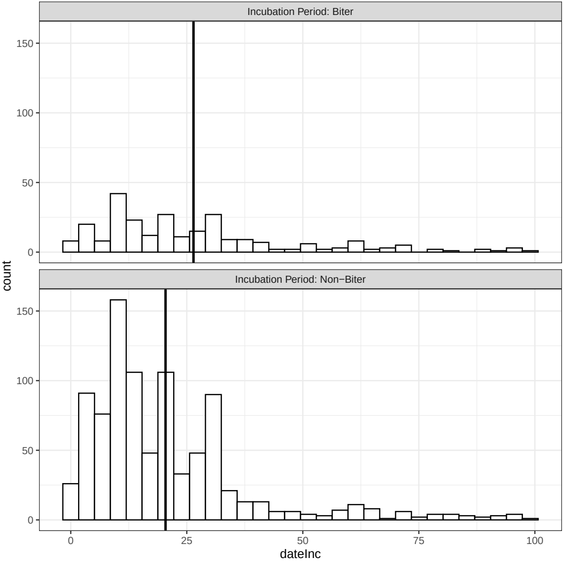
Single Transmission
X
Generation Interval
Multiple Transmissions
X
G1
G2
G3
Getting Bitten
Become Infectious
Clinical Signs
Contact/Biting
Incubation Period
Infectious Period
Waiting Time
X
Removed

Rabies

- If symptoms always start *before* infectiousness happens, then serial interval should equal generation interval:
  - incubation time + extra latent time + waiting time
  - extra latent time + waiting time + incubation time



Generation
Serial
count
200
150
100
50
0
0
25
50
75
100
Days


Incubation Period: Biter
Incubation Period: Non−Biter
count
150
100
50
0
0
25
50
75
100
dateInc

Thanks

- Fields
- Organizers
- Collaborators:
  - *especially*: Li, Park, Weitz
  - *proximately*: Bolker, Ma
- Funders: NSERC, CIHR