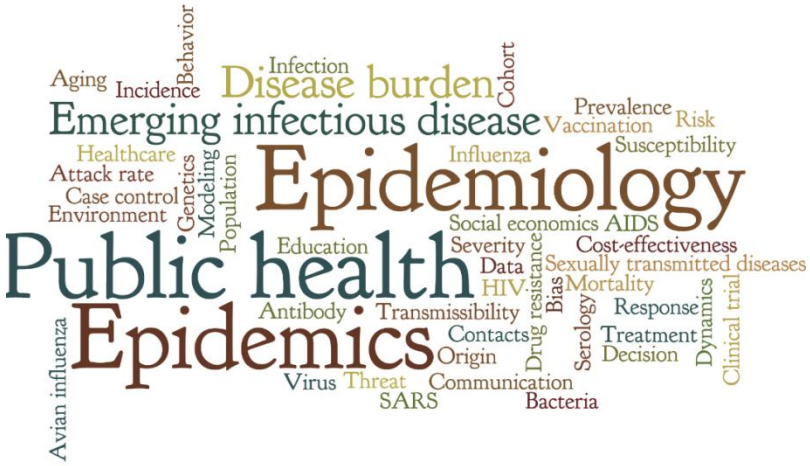
# Introduction to Infectious disease epidemiology

Dr Tim K. Tsang

WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong.


SCHOOL OF PUBLIC HEALTH
THE UNIVERSITY OF HONG KONG
香 港 大 學 公 共 衞 生 學 院

# Epidemiology is the basic science of public health

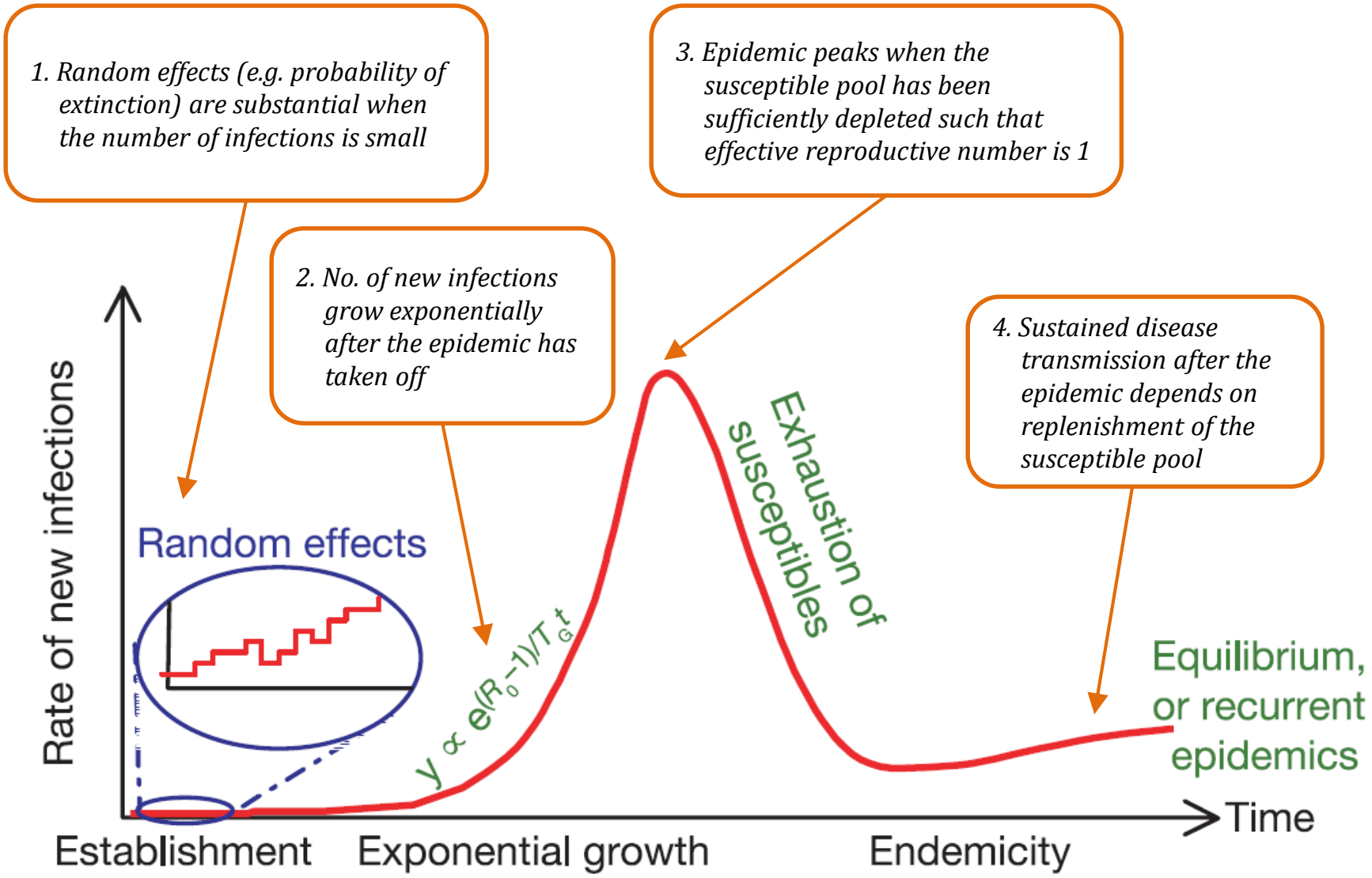



*Where and when does the disease of interest appear?*

*What is the burden of the disease in the population?*

*Do people who have the disease have any common characteristics that are not present among people who do not have the disease?*

# Anatomy of an epidemic

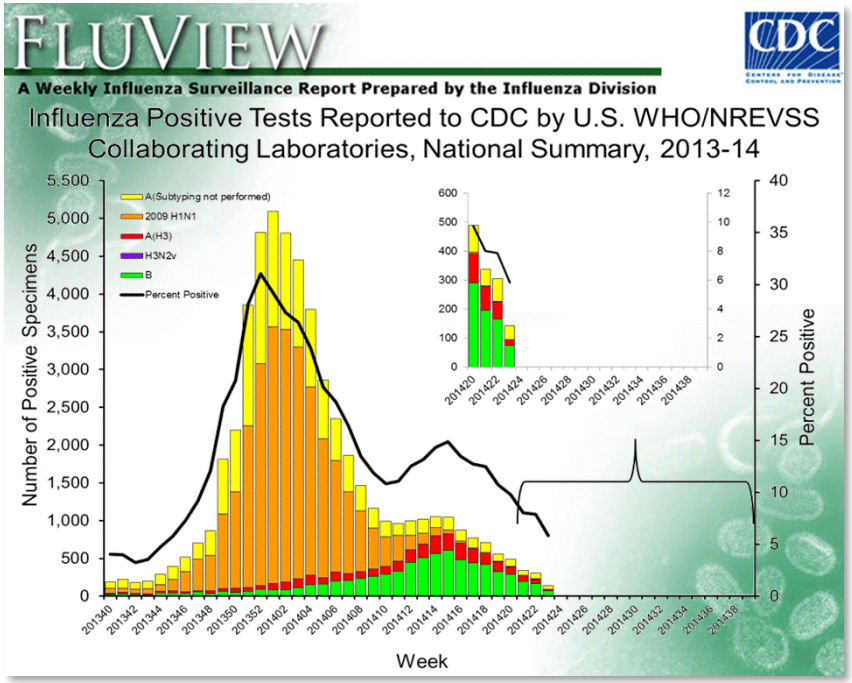


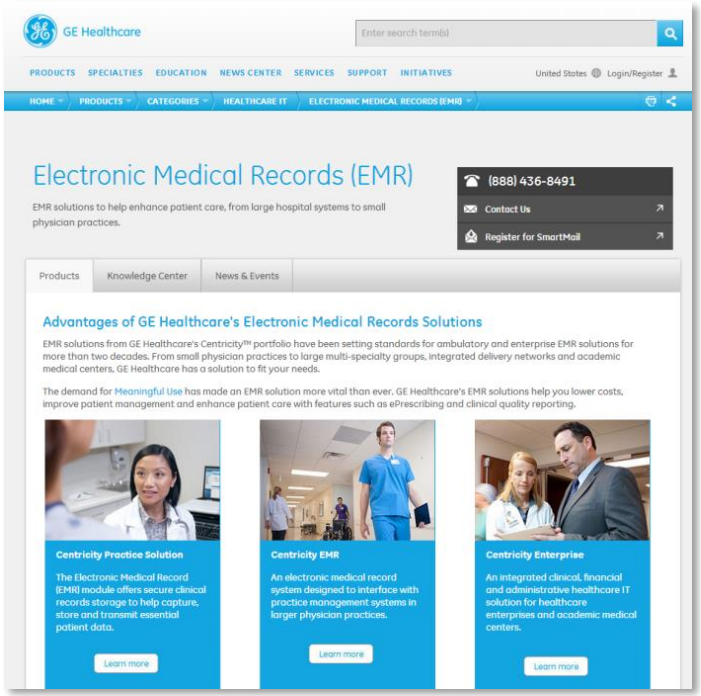

Ferguson et al *Nature* 2003

# Infectious disease epidemiology

Many epidemiologic studies of infectious diseases are observational

- Impossible to set up precisely the same set of conditions
- Releasing pathogens is unethical and prohibited
- Robust disease surveillance and meticulous medical records are essential for infectious disease studies



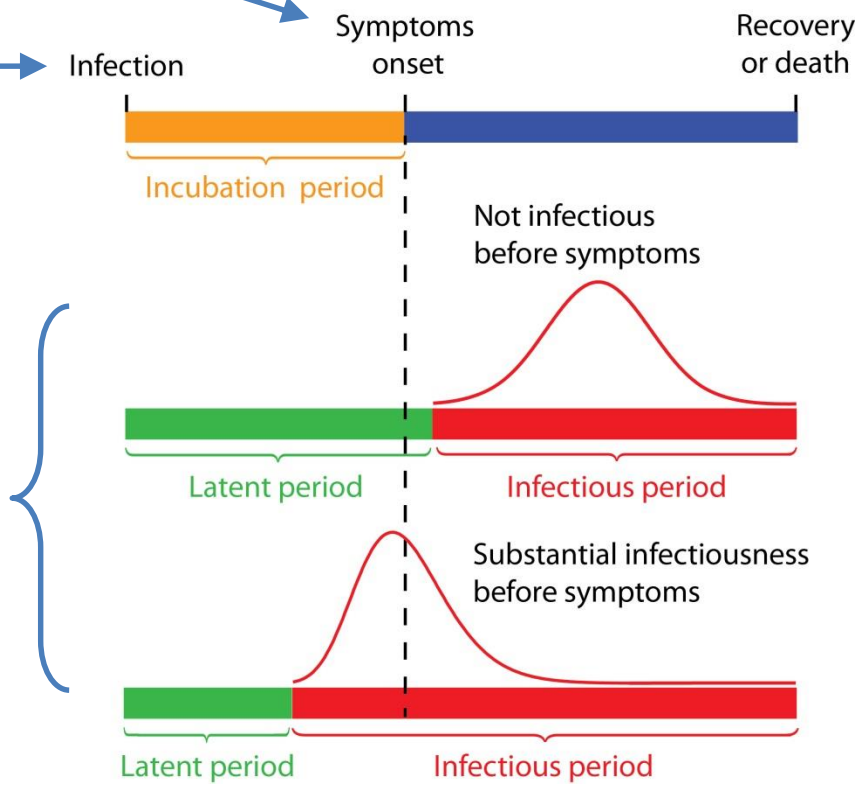



# Natural history

Incubation period: Time from infection to symptoms onset
Latent period: Time from infection to infectiousness onset
Infectious period: Time during which the infected person is infectiousness

Symptoms are readily observable or recallable

The time of infection is most often not directly observable, but might be inferred from exposure history




Diseases with higher pre-symptomatic infectiousness are generally more difficult to control (with symptoms-based interventions, e.g. isolation and contact tracing)

*"Factors that make an infectious disease outbreak controllable"*

Fraser et al *PNAS* 2004

# Incubation period

## The Incubation Period Distribution of Coronavirus Disease 2019: A Systematic Review and Meta-analysis

Hualei Xin,[1] Jessica Y. Wong,[1] Caitriona Murphy,[1] Amy Yeung,[1] Sheikh Taslim Ali,[1,2] Peng Wu,[1,2,©] and Benjamin J. Cowling[1,2]

[1]World Health Organization Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China; and [2]Laboratory of Data Discovery for Health Limited, Hong Kong Science Park, New Territories, Hong Kong Special Administrative Region, China

Incubation period is an important parameter to inform quarantine period and to study transmission dynamics of infectious diseases. We conducted a systematic review and meta-analysis on published estimates of the incubation period distribution of coronavirus disease 2019, and showed that the pooled median of the point estimates of the mean, median and 95th percentile for incubation period are 6.3 days (range, 1.8–11.9 days), 5.4 days (range, 2.0–17.9 days), and 13.1 days (range, 3.2–17.8 days), respectively. Estimates of the mean and 95th percentile of the incubation period distribution were considerably shorter before the epidemic peak in China compared to after the peak, and variation was also noticed for different choices of methodological approach in estimation. Our findings implied that corrections may be needed before directly applying estimates of incubation period into control of or further studies on emerging infectious diseases.

**Keywords.** COVID-19; SARS-CoV-2; incubation period; systematic review; meta-analysis.

SARS-CoV-2: Xin et al *Clin Infect Dis.* 2021

## Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: a population-based study of laboratory-confirmed cases

*Benjamin J Cowling*, Lianmei Jin*, Eric H Y Lau, Qiaohong Liao, Peng Wu, Hui Jiang, Tim K Tsang, Jiandong Zheng, Vicky J Fang, Zhaorui Chang, Michael Y Ni, Qian Zhang, Dennis K M Ip, Jianxing Yu, Yu Li, Liping Wang, Wenxiao Tu, Ling Meng, Joseph T Wu, Huiming Luo, Qun Li, Yuelong Shu, Zhongjie Li, Zijian Feng, Weizhong Yang, Yu Wang, Gabriel M Leung, Hongjie Yu*

**Summary**

**Background** The novel influenza A H7N9 virus emerged recently in mainland China, whereas the influenza A H5N1 virus has infected people in China since 2003. Both infections are thought to be mainly zoonotic. We aimed to compare the epidemiological characteristics of the complete series of laboratory-confirmed cases of both viruses in mainland China so far.

**Methods** An integrated database was constructed with information about demographic, epidemiological, and clinical variables of laboratory-confirmed cases of H7N9 (130 patients) and H5N1 (43 patients) that were reported to the Chinese Centre for Disease Control and Prevention until May 24, 2013. We described disease occurrence by age, sex, and geography, and estimated key epidemiological variables. We used survival analysis techniques to estimate the following distributions: infection to onset, onset to admission, onset to laboratory confirmation, admission to death, and admission to discharge.

**Findings** The median age of the 130 individuals with confirmed infection with H7N9 was 62 years and of the 43 with H5N1 was 26 years. In urban areas, 74% of cases of both viruses were in men, whereas in rural areas the proportions of the viruses in men were 62% for H7N9 and 33% for H5N1. 75% of patients infected with H7N9 and 71% of those with H5N1 reported recent exposure to poultry. The mean incubation period of H7N9 was 3·1 days and of H5N1 was 3·3 days. On average, 21 contacts were traced for each case of H7N9 in urban areas and 18 in rural areas, compared with 90 and 63 for H5N1. The fatality risk on admission to hospital was 36% (95% CI 26–45) for H7N9 and 70% (56–83%) for H5N1.

**Interpretation** The sex ratios in urban compared with rural cases are consistent with exposure to poultry driving the risk of infection—a higher risk in men was only recorded in urban areas but not in rural areas, and the increased risk for men was of a similar magnitude for H7N9 and H5N1. However, the difference in susceptibility to serious illness with the two different viruses remains unexplained, since most cases of H7N9 were in older adults whereas most cases of H5N1 were in younger people. A limitation of our study is that we compared laboratory-confirmed cases of H7N9 and H5N1 infection, and some infections might not have been ascertained.

H7N9: Cowling et al *Lancet* 2004

## The Epidemiology of Severe Acute Respiratory Syndrome in the 2003 Hong Kong Epidemic: An Analysis of All 1755 Patients

Gabriel M. Leung, MD, MPH; Anthony J. Hedley, MD, FRCP; Lai-Ming Ho, PhD; Patsy Chau, MStat; Irene O.L. Wong, MPhil, MMedSc; Thuan Q. Thach, PhD; Azra C. Ghani, PhD; Christl A. Donnelly, ScD; Christophe Fraser, PhD; Steven Riley, DPhil; Neil M. Ferguson, DPhil; Roy M. Anderson, PhD; Thomas Tsang, MBBS, FHKAM; Pak-Yin Leung, MBBS, FFPH; Vivian Wong, MBBS, FHKAM; Jane C.K. Chan, MD, FHKAM; Eva Tsui, MStat; Su-Vui Lo, MBChB, FFPH; and Tai-Hing Lam, MD, FFPH

Background: **As yet, no one has written a comprehensive epidemiologic account of a severe acute respiratory syndrome (SARS) outbreak from an affected country.**

Objective: **To provide a comprehensive epidemiologic account of a SARS outbreak from an affected territory.**

Design: **Epidemiologic analysis.**

Setting: **The 2003 Hong Kong SARS outbreak.**

Participants: **All 1755 cases and 302 deaths.**

Measurements: **Sociodemographic characteristics; infection clusters by time, occupation, setting, and workplace; and geospatial relationships were determined. The mean and variance in the time from infection to onset (incubation period) were estimated in a small group of patients with known exposure. The mean and variance in time from onset to admission, from admission to discharge, or from admission to death were calculated. Logistic regression was used to identify important predictors of case fatality.**

Results: **49.3% of patients were infected in clinics, hospitals, or elderly or nursing homes, and the Amoy Gardens cluster accounted for 18.8% of cases. The ratio of women to men among infected individuals was 5:4. Health care workers accounted for 23.1% of all reported cases. The estimated mean incubation period was 4.6 days (95% CI, 3.8 to 5.8 days). Mean time from symptom onset to hospitalization varied between 2 and 8 days, decreasing over the course of the epidemic. Mean time from onset to death was 23.7 days (CI, 22.0 to 25.3 days), and mean time from onset to discharge was 26.5 days (CI, 25.8 to 27.2 days). Increasing age, male sex, atypical presenting symptoms, presence of comorbid conditions, and high lactate dehydrogenase level on admission were associated with a greater risk for death.**

Limitations: **Estimates of the incubation period relied on statistical assumptions because few patients had known exposure times. Temporal changes in case management as the epidemic progressed, unavailable treatment information, and several potentially important factors that could not be thoroughly analyzed because of the limited sample size complicate interpretation of factors related to case fatality.**

Conclusions: **This analysis of the complete data on the 2003 SARS epidemic in Hong Kong has revealed key epidemiologic features of the epidemic as it evolved.**

*Ann Intern Med.* 2004;141:662-673. www.annals.org

For author affiliations, see end of text.

SARS: Leung et al *Ann Intern Med* 2004

# Incubation period

The incubation period can vary substantially among individuals:

- Route and dose of transmission
- Host genetics, age, immunity
- Intervention (e.g. pharmacologic prophylaxis and treatment)

## Association between Severity of MERS-CoV Infection and Incubation Period

**Victor Virlogeux, Minah Park, Joseph T. Wu,[1] Benjamin J. Cowling[1]**

We analyzed data for 170 patients in South Korea who had laboratory-confirmed infection with Middle East respiratory syndrome coronavirus. A longer incubation period was associated with a reduction in the risk for death (adjusted odds ratio/1-day increase in incubation period 0.83, 95% credibility interval 0.68–1.03).

data, we assumed that their incubation time was 0–21 days because 21 days was the longest incubation period reported (*9,10*). Data for patients is provided in online Technical Appendix 1 (http://wwwnc.cdc.gov/EID/article/22/3/15-1437-Techapp1.xlsx).

To estimate incubation period distribution, we fitted a gamma distribution that enabled interval censoring (*6*) by using Markov Chain Monte Carlo methods in a Bayesian framework (online Technical Appendix 2, http://wwwnc.

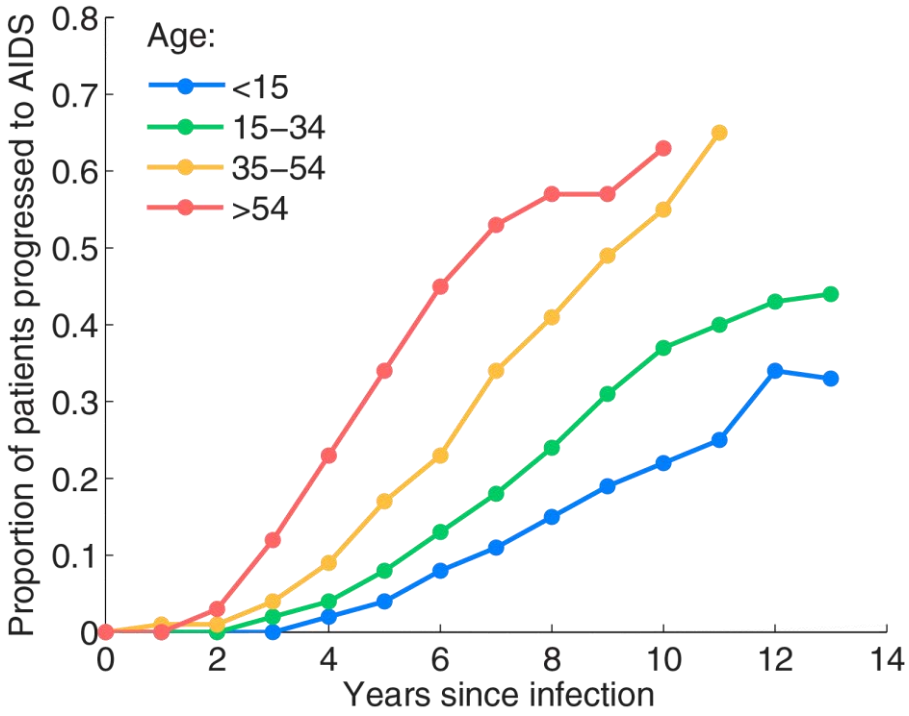
Viriogeux et al *Emerge Infect Dis 2016*

# Incubation period

The incubation period can vary substantially among individuals:

- Route and dose of transmission
- Host genetics, age, immunity
- Intervention (e.g. pharmacologic prophylaxis and treatment)

*Younger patients have longer incubation period for AIDS*



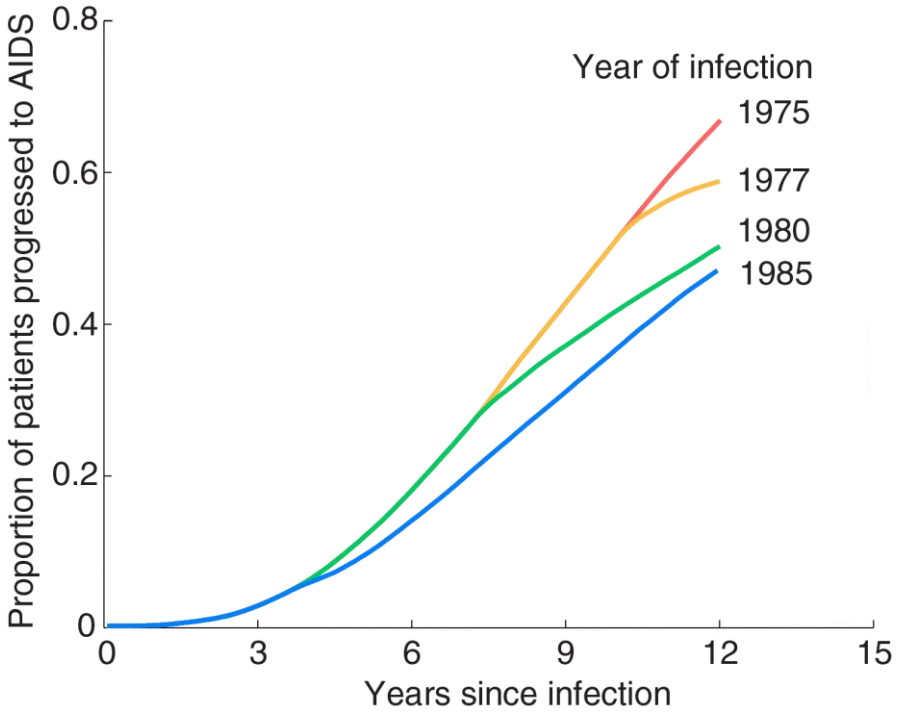

Darby et al *Lancet* 1996

# Incubation period

The incubation period can vary substantially among individuals:

- Route and dose of transmission
- Host genetics, age, immunity
- Intervention (e.g. pharmacologic prophylaxis and treatment)

*Antiretroviral therapies lengthen the incubation period of AIDS*

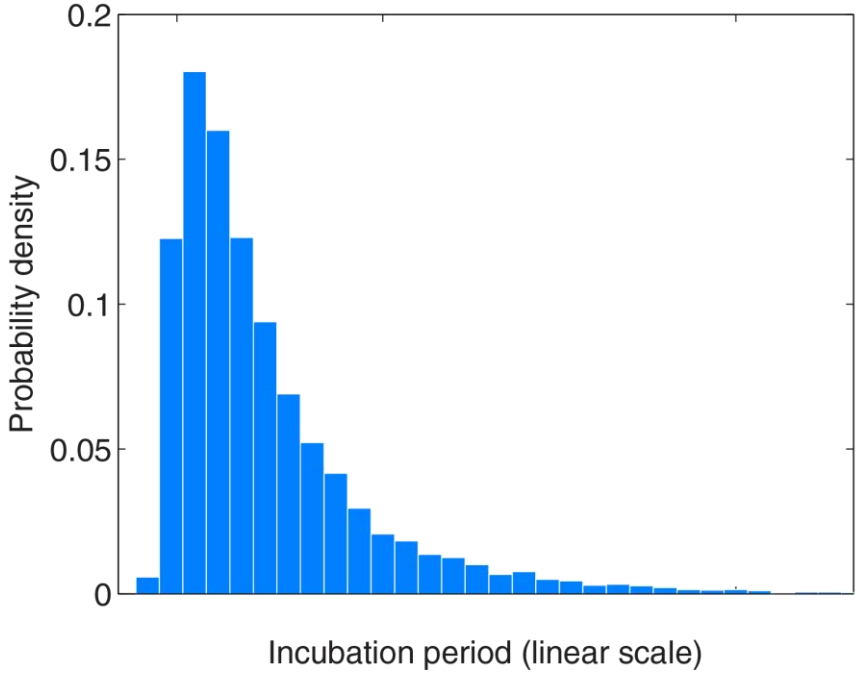


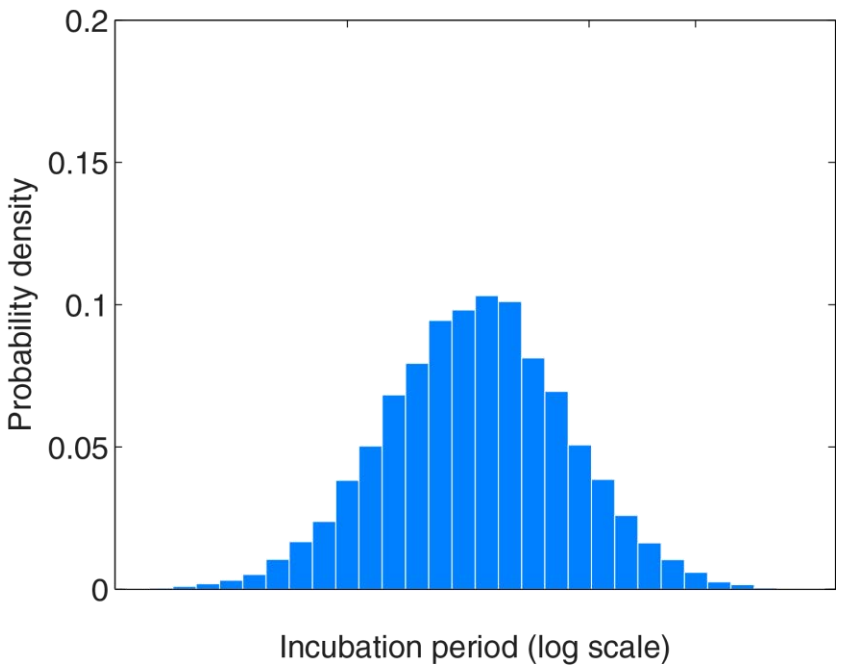

Broomeyer et al *Science* 1991

# Incubation period distribution

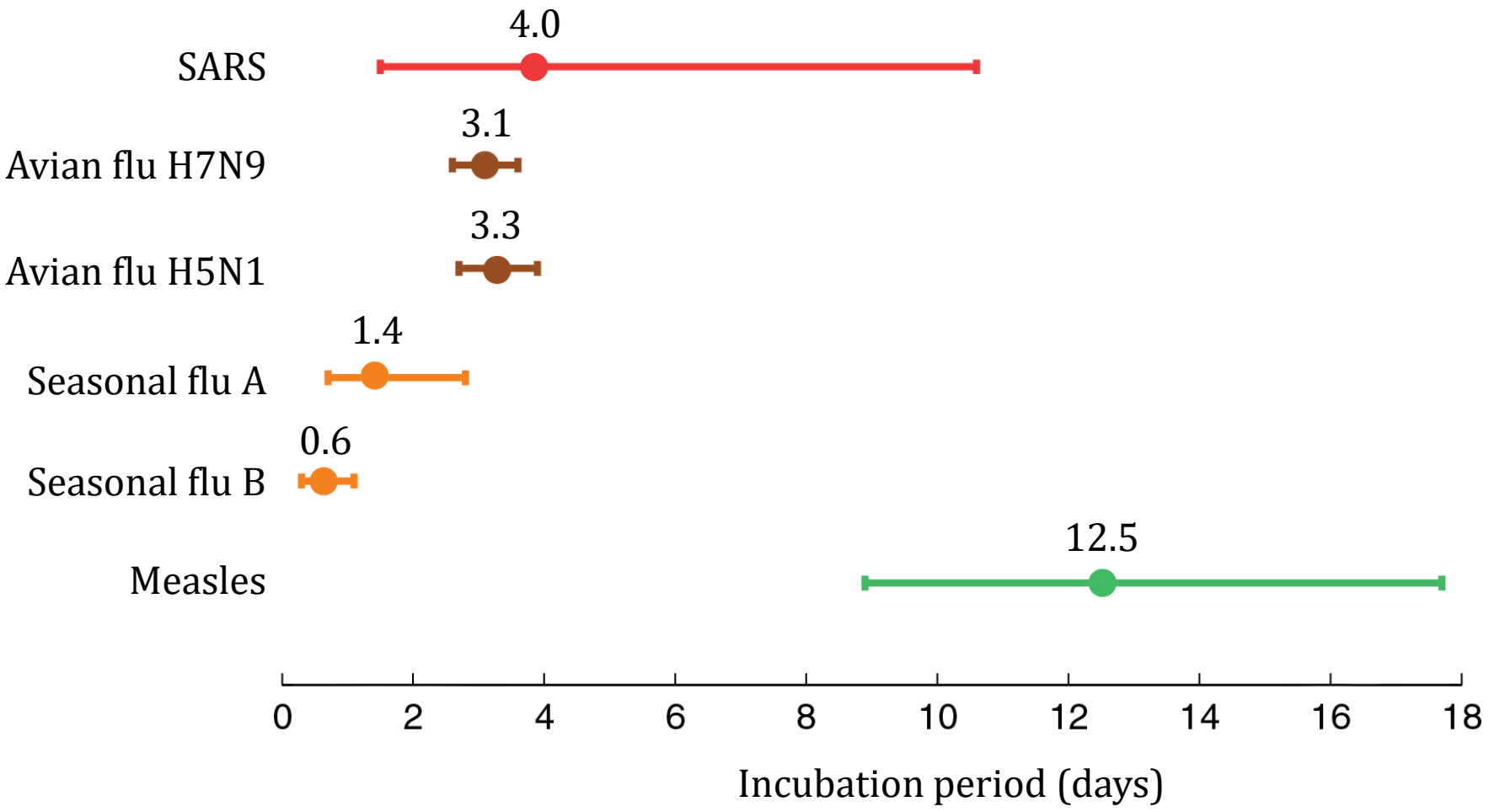
- Typically right-skewed, i.e. has a long right tail

- Often looks like a lognormal distribution (the log of incubation period is normally distributed)



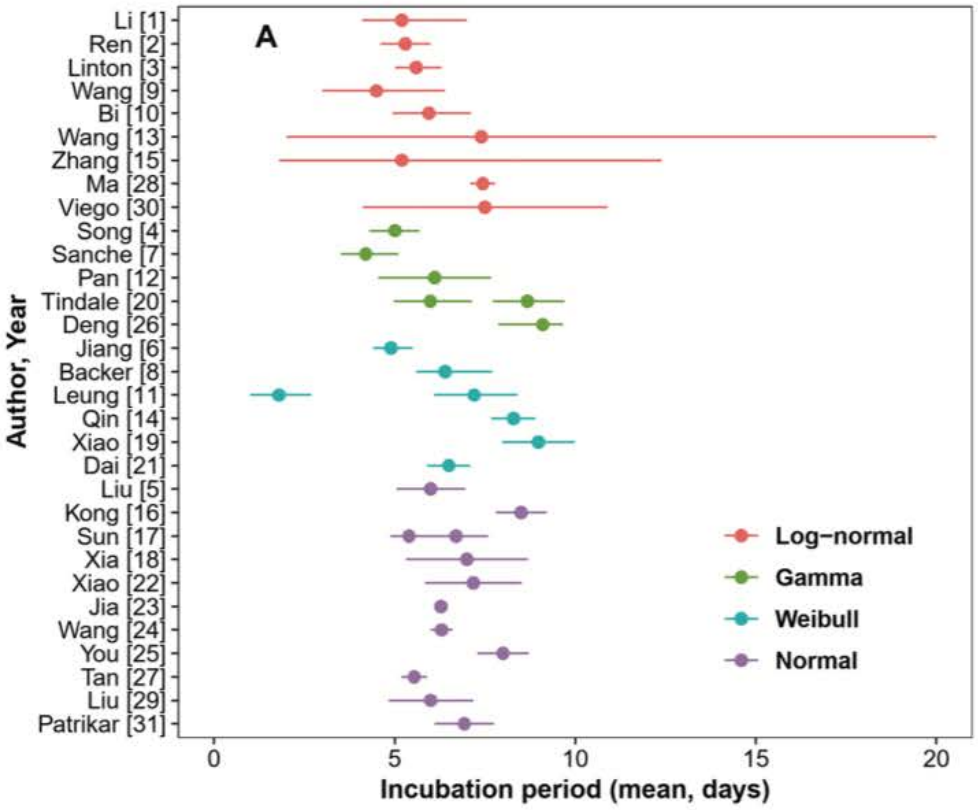



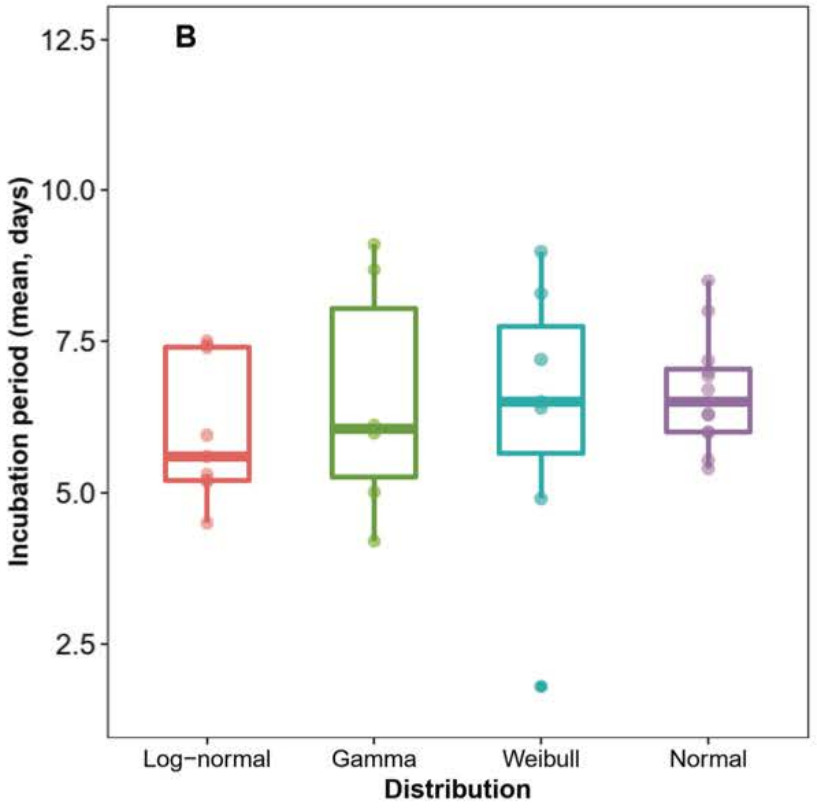
## Incubation periods of acute respiratory viral infections




Lessler et al *Lancet Inf Dis* 2009

# Incubation periods for COVID-19



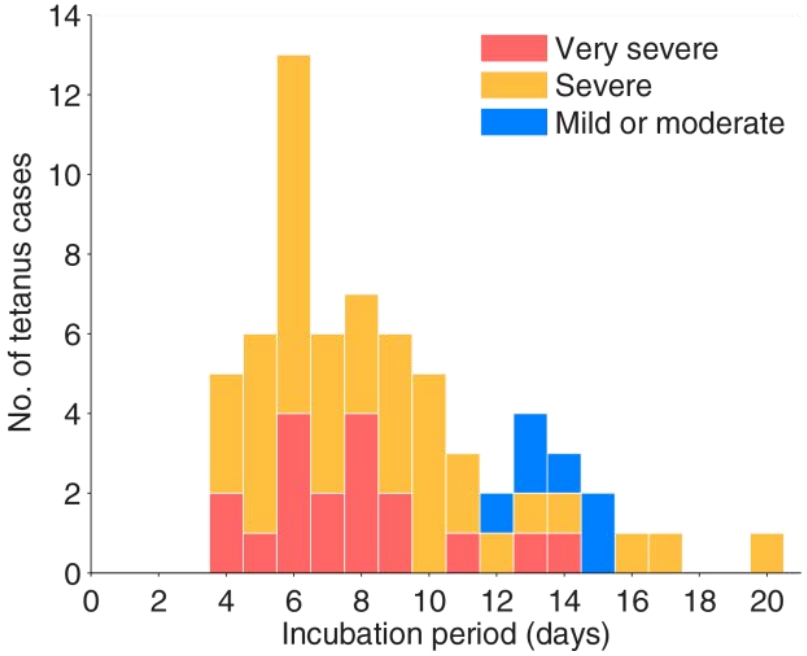

Xin et al *Clinic Inf Dis* 2021

# Why do we want to know about incubation period?

For clinical management:

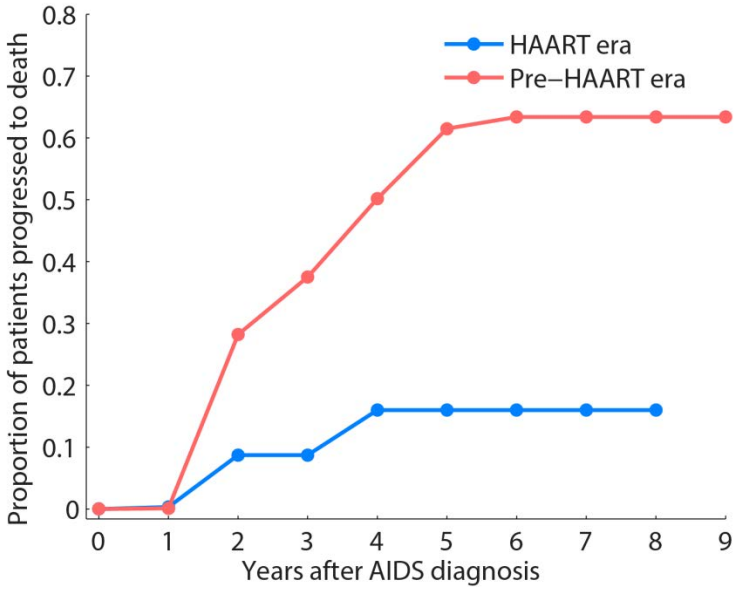
- To predict disease severity, e.g. shorter incubation time is associated with more severe outcome

65 tetanus ICU cases
1961-1977, Leeds, UK




Edmondson et al *Survey of Anesthesiology* 1980

*Risk of progression to death after AIDS diagnosis was reduced after widespread use of hyperactive antiretroviral therapies (HAART)*

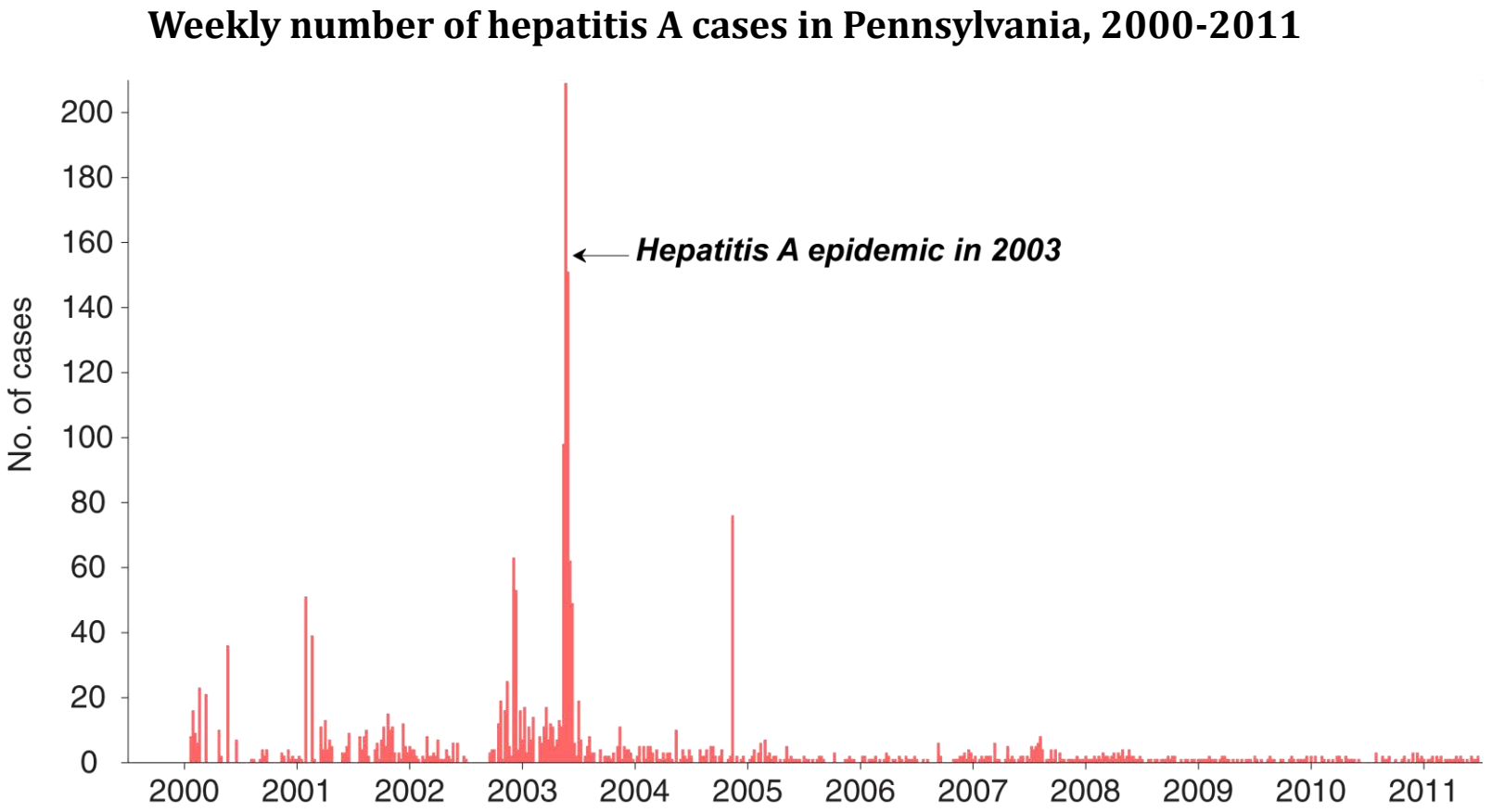



Wong et al *CID* 2004

# Why do we want to know about incubation period?

For public health control:

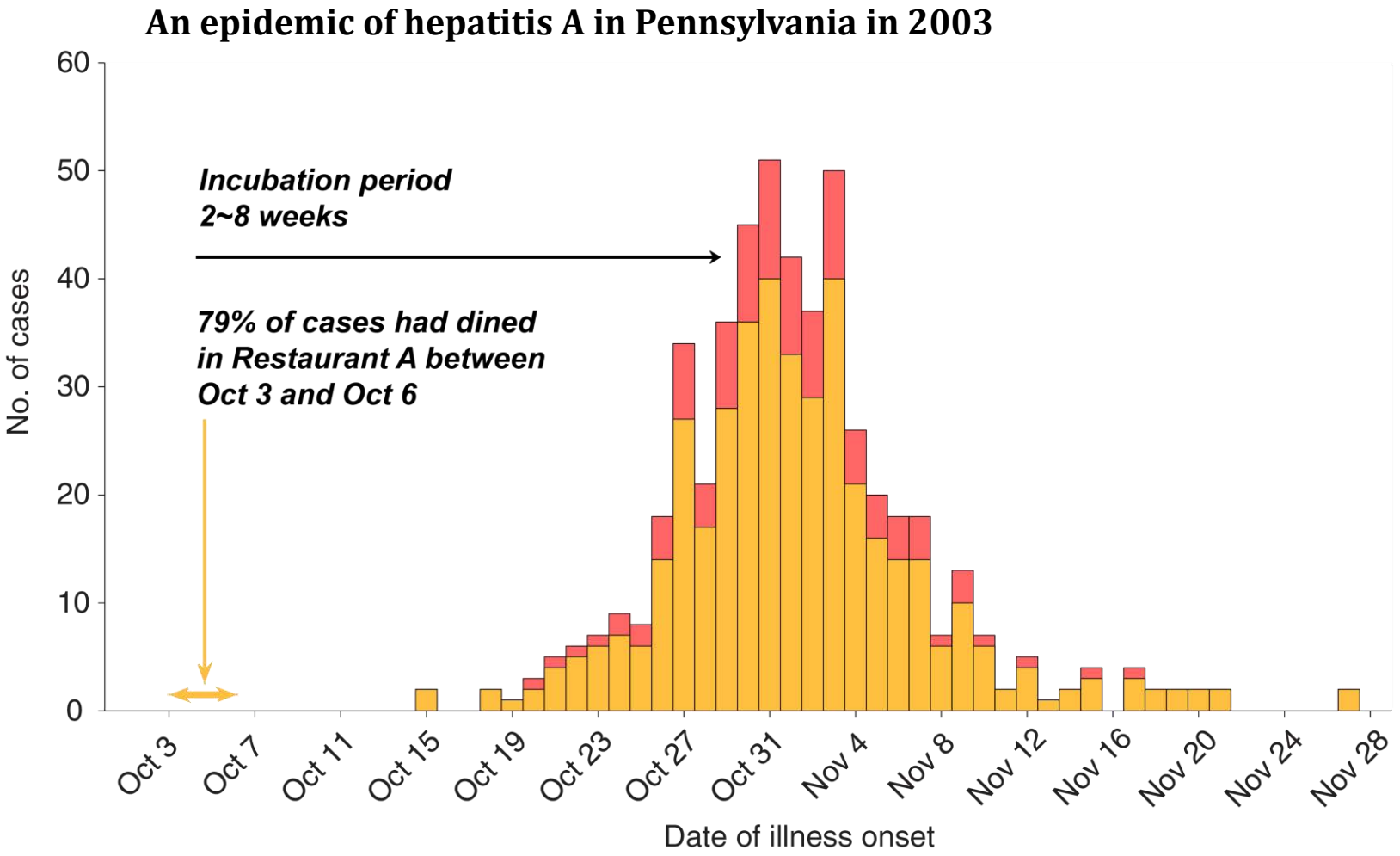
- To identify the origin of common-source outbreaks



# Why do we want to know about incubation period?

For public health control:

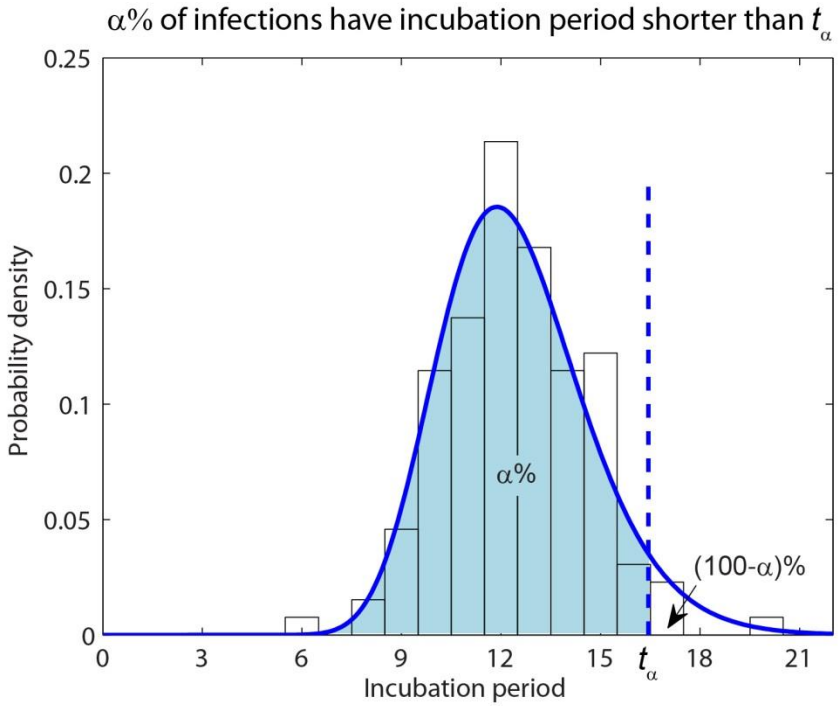
- To identify the origin of common-source outbreaks



# Why do we want to know about incubation period?

For public health control:

- To estimate the duration necessary for quarantining suspected or contacts of cases to ensure that they are not infected upon release
- Quarantine means isolating suspected cases and contacts from the community to prevent disease transmission in case they have been infected but do not yet have the clinical or virologic evidence of so.




Set $\alpha$ high enough so that upon released from quarantine, the individual is uninfected with high probability, e.g. $\alpha$ = 95 or 99

# Why do we want to know about incubation period?

For public health surveillance:

- Backcalculate *incidence of infection* from *incidence of clinical cases*

An example: HIV

The incubation period of AIDS had a median of around 9 years and could range from 3 to more than 12 years, so the observed AIDS cases represented only a small proportion of the total number of HIV infections in the population.

*"HIV infection rates are related to AIDS incidence through the incubation period distribution. The fundamental convolution equation is given by*

$$a(t) = \int_0^t I(s)F(t-s|s)ds$$

*where a(t) is the cumulative number of cases of AIDS diagnosed by year t, I(s) is the infection rate in year s, and F(t|s) is the incubation period distribution among individuals infected in year s."*

Brookmeyer *Science* 1991

# From HIV infection to AIDS diagnoses

$$\sum_{u=1}^{t} AIDS(u) = \sum_{u=1}^{t} HIV(u) \times F(t-u)$$

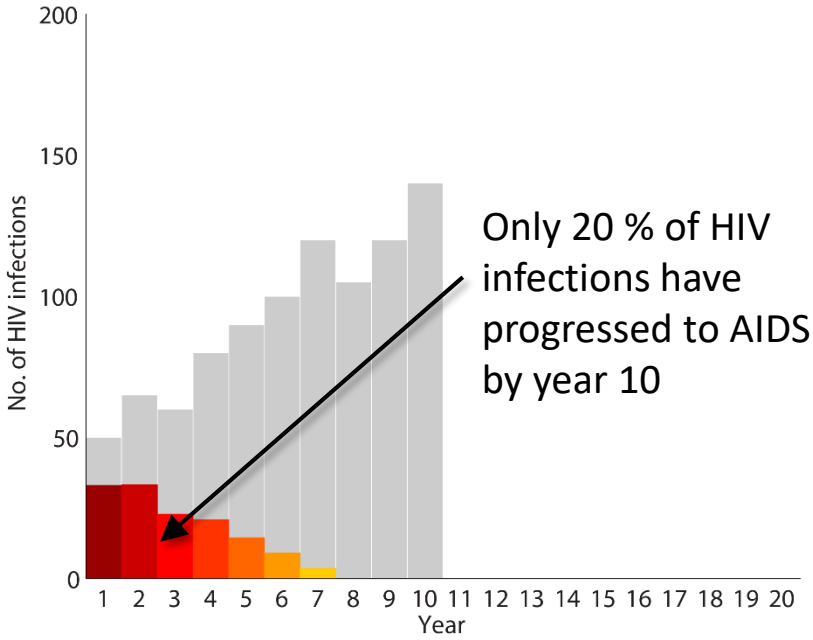
Cumulative number of AIDS in year $t$

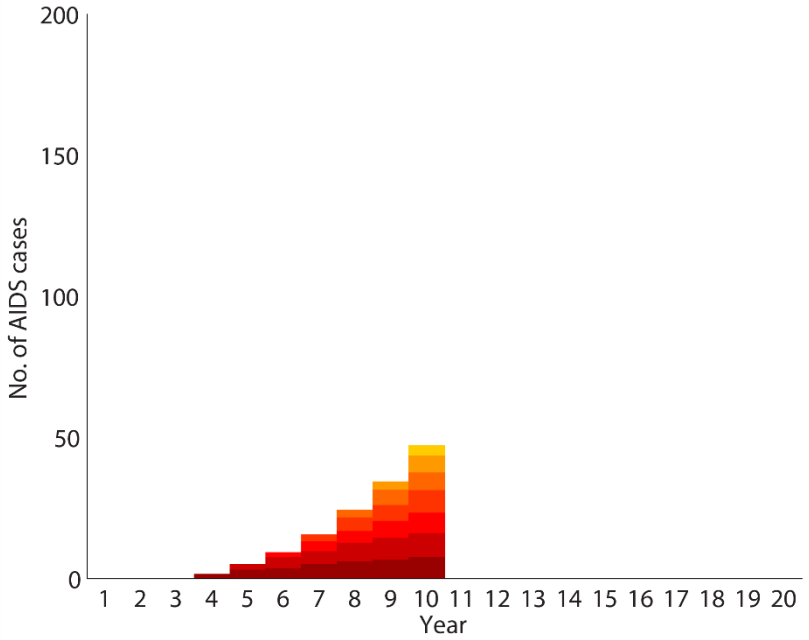
Number of HIV infections in year $u$

Proportion of infections with incubation period shorter than $t-u$ years

<u>Backcalculation</u>

Inferring HIV incidence from AIDS epi-curves and incubation period

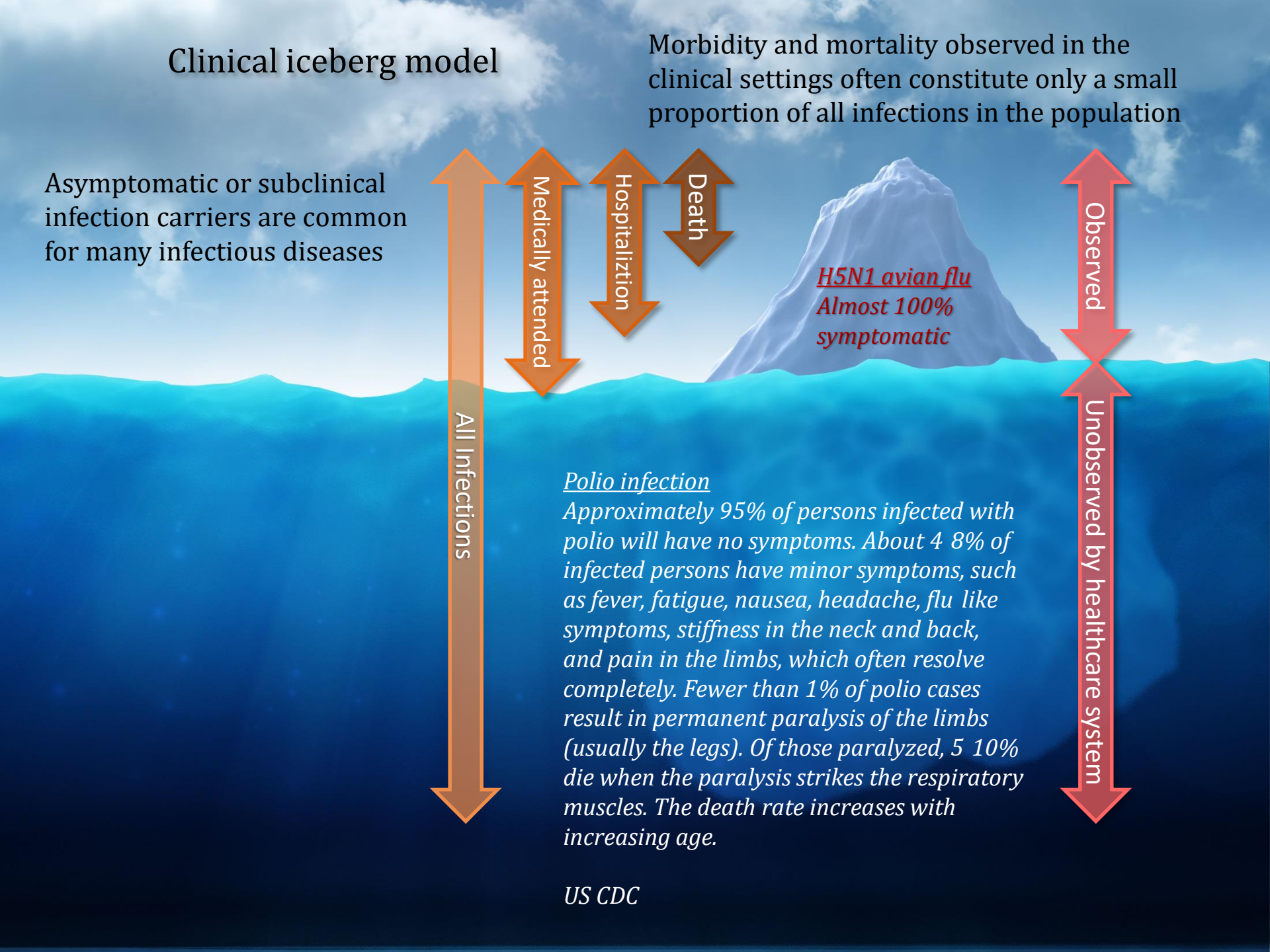


# Clinical iceberg model

Morbidity and mortality observed in the clinical settings often constitute only a small proportion of all infections in the population

Asymptomatic or subclinical infection carriers are common for many infectious diseases




*Polio infection*

*Approximately 95% of persons infected with polio will have no symptoms. About 4 8% of infected persons have minor symptoms, such as fever, fatigue, nausea, headache, flu like symptoms, stiffness in the neck and back, and pain in the limbs, which often resolve completely. Fewer than 1% of polio cases result in permanent paralysis of the limbs (usually the legs). Of those paralyzed, 5 10% die when the paralysis strikes the respiratory muscles. The death rate increases with increasing age.*

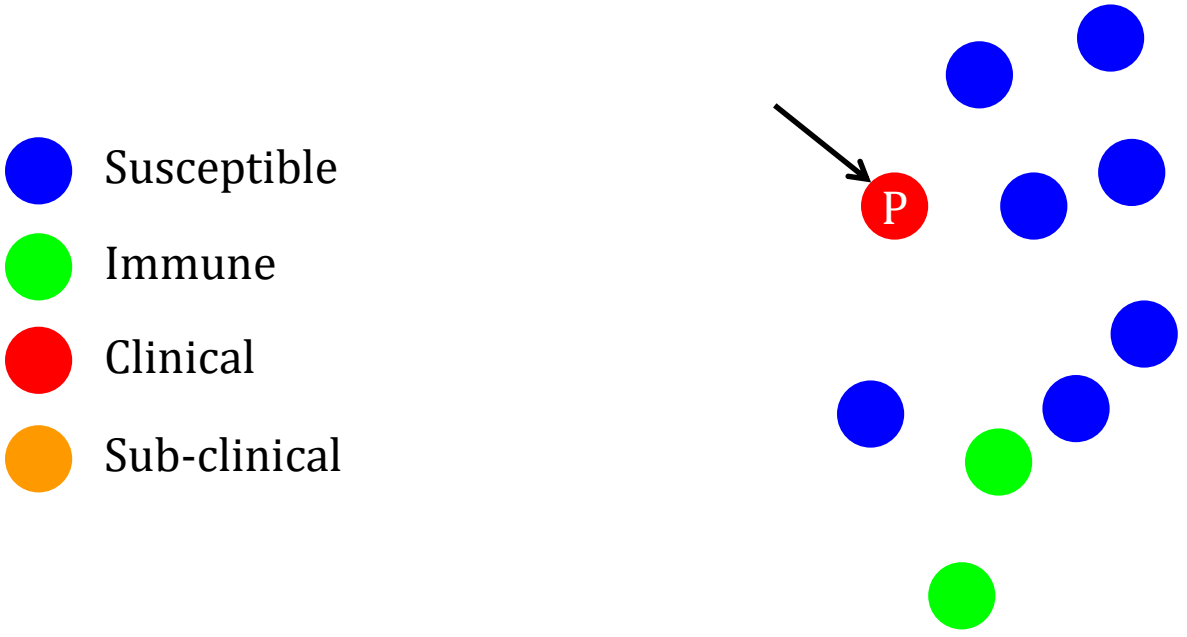
*US CDC*

# Measures for transmissibility

- Infection attack rate
  - The proportion of a population (subgroup) infected over the course of an epidemic
- Secondary (infection) attack rate
  - The proportion of individuals infected in a semi-closed setting (e.g. households, airplanes, military barracks) in an outbreak caused by an index case (ideally accounting for pre-existing immunity)
- Basic reproductive number
  - The average number of secondary cases generated by an index case when an epidemic begins in a completely susceptible population

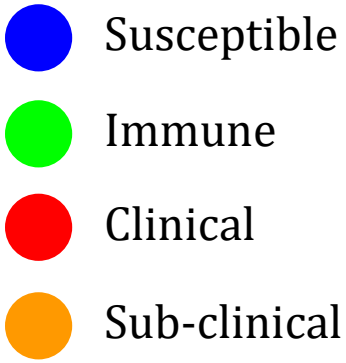
# Disease transmission

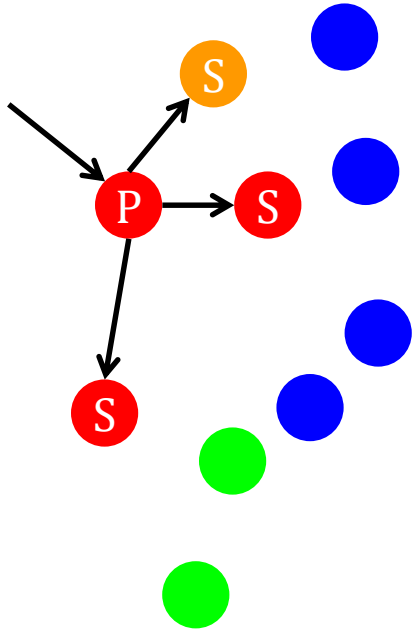
**Index case –** the first case identified
**Primary case –** the case bringing infection to a population



# Disease transmission

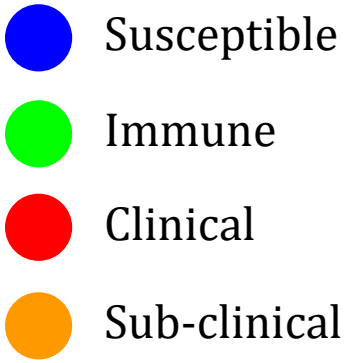
**Index case –** the first case identified
**Primary case –** the case bringing infection to a population
**Secondary case –** infected by a primary case

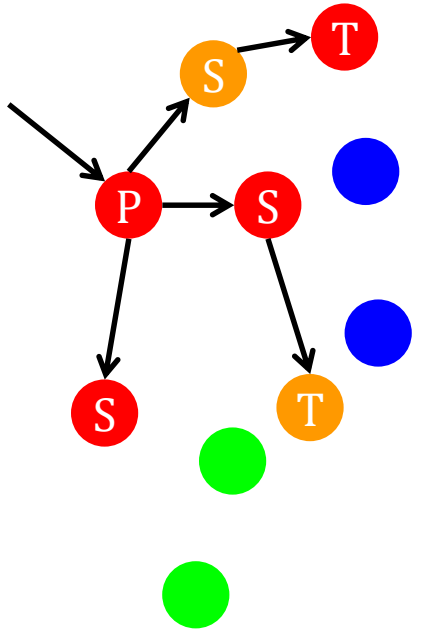


# Disease transmission

**Index case –** the first case identified
**Primary case –** the case bringing infection to a population
**Secondary case –** infected by a primary case
**Tertiary case –** infected by a secondary case



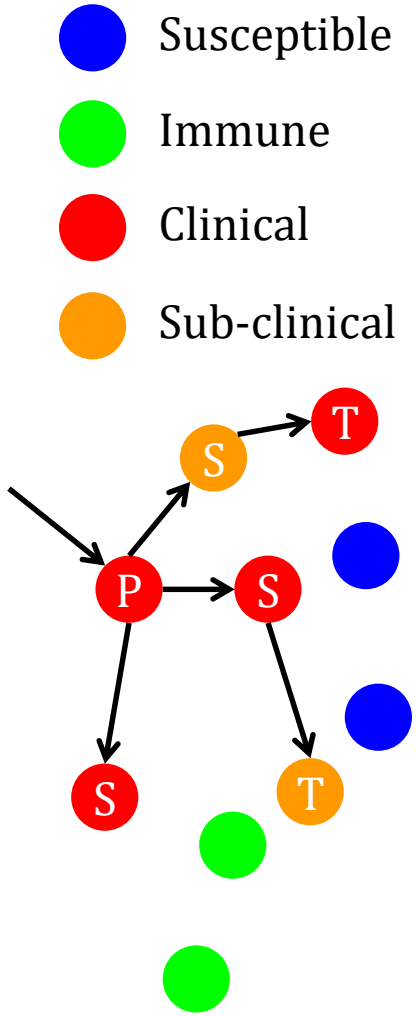



# Secondary attack rate

$$SAR = \frac{no.\,of\ cases}{no.\,of\ exposed\ (susceptibles)}$$

**Denominator**

| **Numerator** | All members | Susceptibles only |
|---|---|---|
| All infections | S S S T T / S S S T T ● ● ● ● | S S S T T / S S S T T ● ● |
| Clinical infections | S S T / S S S T T ● ● ● ● | S S T / S S S T T ● ● |

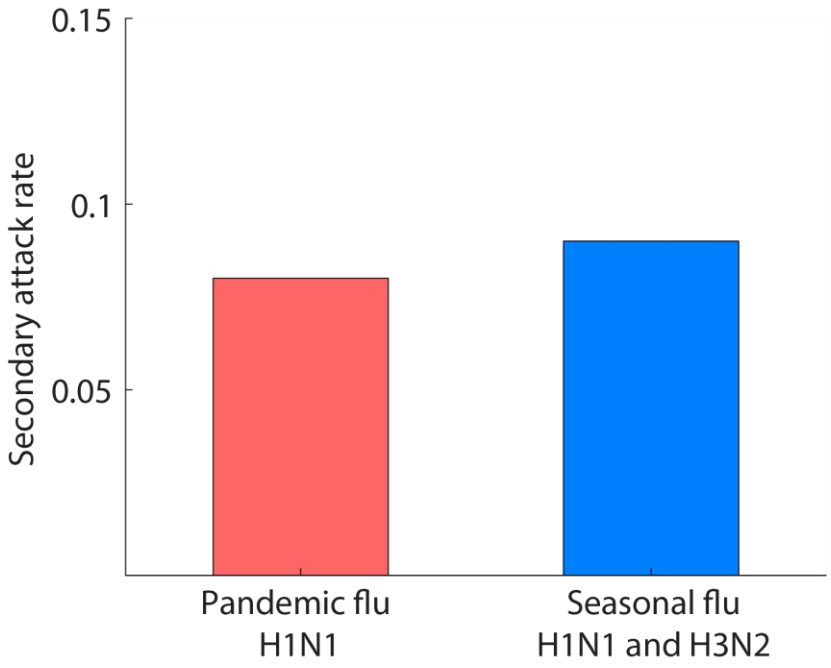


The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

# Comparative Epidemiology of Pandemic and Seasonal Influenza A in Households

Benjamin J. Cowling, Ph.D., Kwok Hung Chan, Ph.D., Vicky J. Fang, M.Phil., Lincoln L.H. Lau, B.Sc., Hau Chi So, B.N.S., Rita O.P. Fung, B.N.S., Edward S.K. Ma, M.Phil., Alfred S.K. Kwong, M.B., B.S., Chi-Wai Chan, M.B., B.S., Wendy W.S. Tsui, M.B., B.S., Ho-Yin Ngai, M.B., B.S., Daniel W.S. Chu, M.B., B.S., Paco W.Y. Lee, M.B., B.S., Ming-Chee Chiu, M.B., B.S., Gabriel M. Leung, M.D., and Joseph S.M. Peiris, D.Phil.




**BACKGROUND**

There are few data on the comparative epidemiology and virology of the pandemic 2009 influenza A (H1N1) virus and cocirculating seasonal influenza A viruses in community settings.

**METHODS**

We recruited 348 index patients with acute respiratory illness from 14 outpatient clinics in Hong Kong in July and August 2009. We then prospectively followed household members of 99 patients who tested positive for influenza A virus on rapid diagnostic testing. We collected nasal and throat swabs from all household members at three home visits within 7 days for testing by means of quantitative reverse-transcriptase–polymerase-chain-reaction (RT-PCR) assay and viral culture. Using hemagglutination-inhibition and viral-neutralization assays, we tested baseline and convalescent serum samples from a subgroup of patients for antibody responses to the pandemic and seasonal influenza A viruses.

**RESULTS**

Secondary attack rates (as confirmed on RT-PCR assay) among household contacts of index patients were similar for the pandemic influenza virus (8%; 95% confidence interval [CI], 3 to 14) and seasonal influenza viruses (9%; 95% CI, 5 to 15). The patterns of viral shedding and the course of illness among index patients were also similar for the pandemic and seasonal influenza viruses. In a subgroup of patients for whom baseline and convalescent serum samples were available, 36% of household contacts who had serologic evidence of pandemic influenza virus infection did not shed detectable virus or report illness.

**CONCLUSIONS**

Pandemic 2009 H1N1 virus has characteristics that are broadly similar to those of seasonal influenza A viruses in terms of rates of viral shedding, clinical illness, and transmissibility in the household setting.

Original Investigation | Global Health

# Factors Associated With Household Transmission of SARS-CoV-2
## An Updated Systematic Review and Meta-analysis

Zachary J. Madewell, PhD; Yang Yang, PhD; Ira M. Longini Jr, PhD; M. Elizabeth Halloran, MD, DSc; Natalie E. Dean, PhD


## Abstract

**IMPORTANCE** A previous systematic review and meta-analysis of household transmission of SARS-CoV-2 that summarized 54 published studies through October 19, 2020, found an overall secondary attack rate (SAR) of 16.6% (95% CI, 14.0%-19.3%). However, the understanding of household secondary attack rates for SARS-CoV-2 is still evolving, and updated analysis is needed.

**OBJECTIVE** To use newly published data to further the understanding of SARS-CoV-2 transmission in the household.

**DATA SOURCES** PubMed and reference lists of eligible articles were used to search for records published between October 20, 2020, and June 17, 2021. No restrictions on language, study design, time, or place of publication were applied. Studies published as preprints were included.

**STUDY SELECTION** Articles with original data that reported at least 2 of the following factors were included: number of household contacts with infection, total number of household contacts, and secondary attack rates among household contacts. Studies that reported household infection prevalence (which includes index cases), that tested contacts using antibody tests only, and that included populations overlapping with another included study were excluded. Search terms were *SARS-CoV-2* or *COVID-19* with *secondary attack rate*, *household*, *close contacts*, *contact transmission*, *contact attack rate*, or *family transmission*.

**DATA EXTRACTION AND SYNTHESIS** Meta-analyses were performed using generalized linear mixed models to obtain SAR estimates and 95% CIs. The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline was followed.

**MAIN OUTCOMES AND MEASURES** Overall household SAR for SARS-CoV-2, SAR by covariates (contact age, sex, ethnicity, comorbidities, and relationship; index case age, sex, symptom status, presence of fever, and presence of cough; number of contacts; study location; and variant), and SAR by index case identification period.

**RESULTS** A total of 2722 records (2710 records from database searches and 12 records from the reference lists of eligible articles) published between October 20, 2020, and June 17, 2021, were identified. Of those, 93 full-text articles reporting household transmission of SARS-CoV-2 were assessed for eligibility, and 37 studies were included. These 37 new studies were combined with 50 of the 54 studies (published through October 19, 2020) from our previous review (4 studies from Wuhan, China, were excluded because their study populations overlapped with another recent study), resulting in a total of 87 studies representing 1 249 163 household contacts from 30 countries. The estimated household SAR for all 87 studies was 18.9% (95% CI, 16.2%-22.0%). Compared with studies from January to February 2020, the SAR for studies from July 2020 to March 2021 was higher (13.4% [95% CI, 10.7%-16.7%] vs 31.1% [95% CI, 22.6%-41.1%], respectively). Results from


## Key Points

**Question** Are early estimates of household transmission of SARS-CoV-2 indicative of current household transmission?

**Findings** In this updated systematic review and meta-analysis of 87 studies representing 1 249 163 household contacts from 30 countries, the estimated household secondary attack rate was 19%. An increase in household transmission was observed over time, perhaps owing to improved diagnostic procedures and tools, longer follow-up, more contagious variants, and different study locations.

**Meaning** These findings suggest that the household remains an important site of SARS-CoV-2 transmission, and recent studies have generated higher household secondary attack rate estimates compared with the earliest reports; more transmissible variants and vaccines may be associated with additional changes in the future.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Madewell et al *JAMA Network Open 2021*

# Basic concepts for describing transmission

Basic reproductive number, $R_0$

- The average number of secondary cases generated by an index case when an epidemic begins in a completely susceptible population .

Mean generation time, $T_g$

- The average time it takes an index case to infect other individuals after he becomes infected.

Epidemic growth rate, $r$

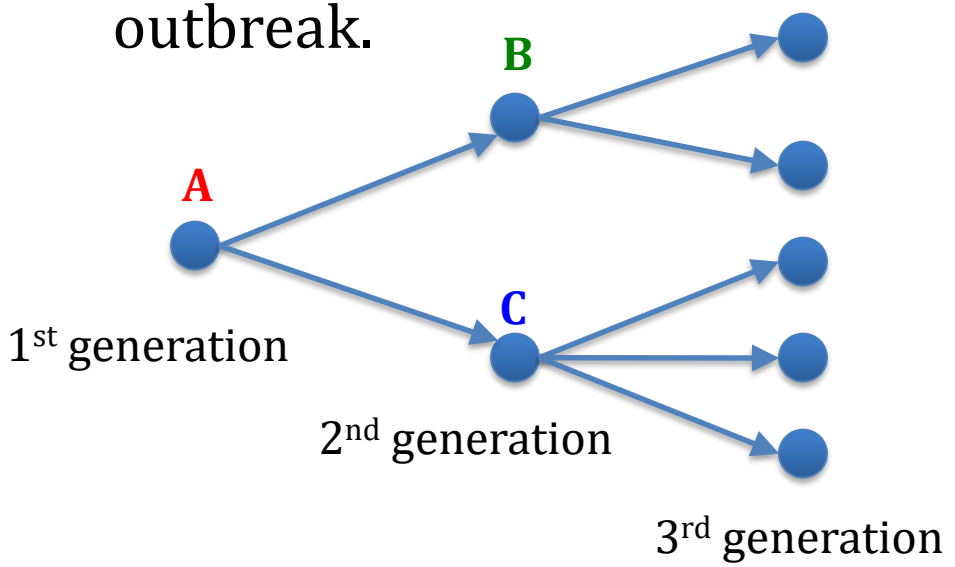
- The rate at which the number of infections is increasing exponentially
- During the early phase, the epidemic doubling time is $t_d = \ln 2/r$. Why? During this phase, the epidemic size is growing exponentially at rate $r$. This means

$$2 \times \text{current size} = \text{current size} \times \exp(rt_d)$$

# Basic reproductive number, $R_0$

The average number of secondary cases generated by an index case in a fully susceptible population:

- Reproductive number, $R$: Same as $R_0$ except that the population needs not be completely susceptible
  - Some time after the epidemic has started, in which case we denote the reproductive number by $R_t$.
  - If the population is partially vaccinated before the outbreak.

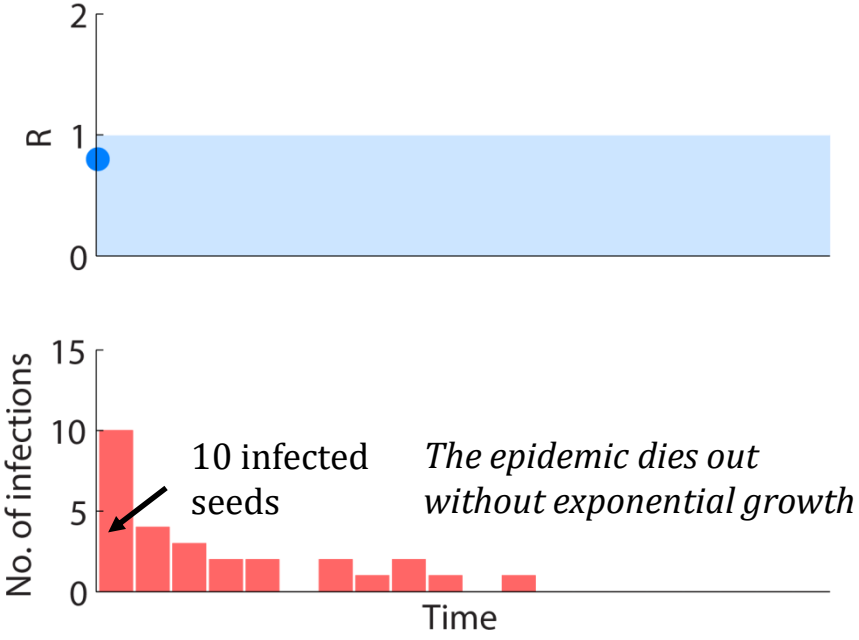



$$R_0 = \frac{2+2+3}{3} = 2.3$$

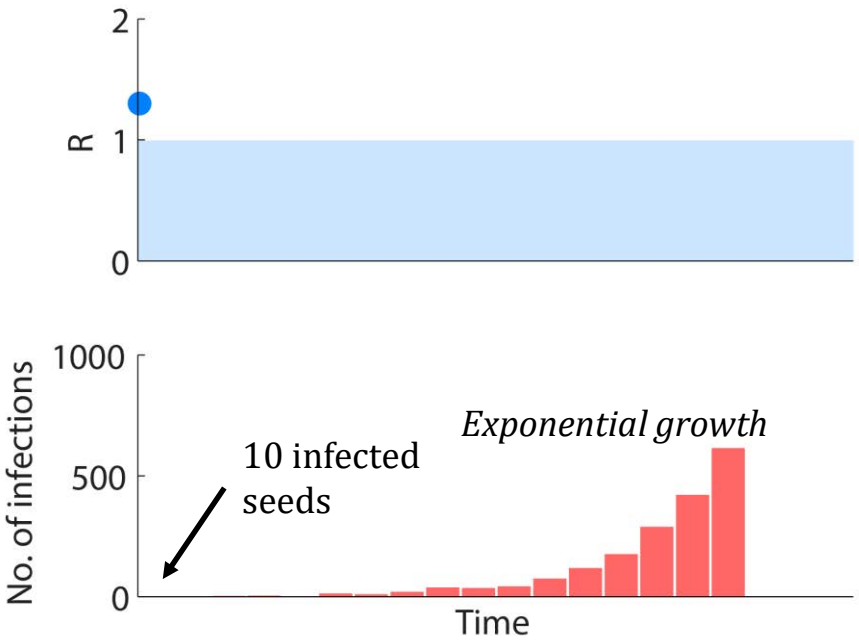
# Basic reproductive number, $R_0$

The average number of secondary cases generated by an index case in a fully susceptible population.

$R_0 < 1$: The epidemic will die out without exponential growth




$R_0 > 1$: The pathogen can lead to an exponentially growing epidemic




How to estimate $R_0$?

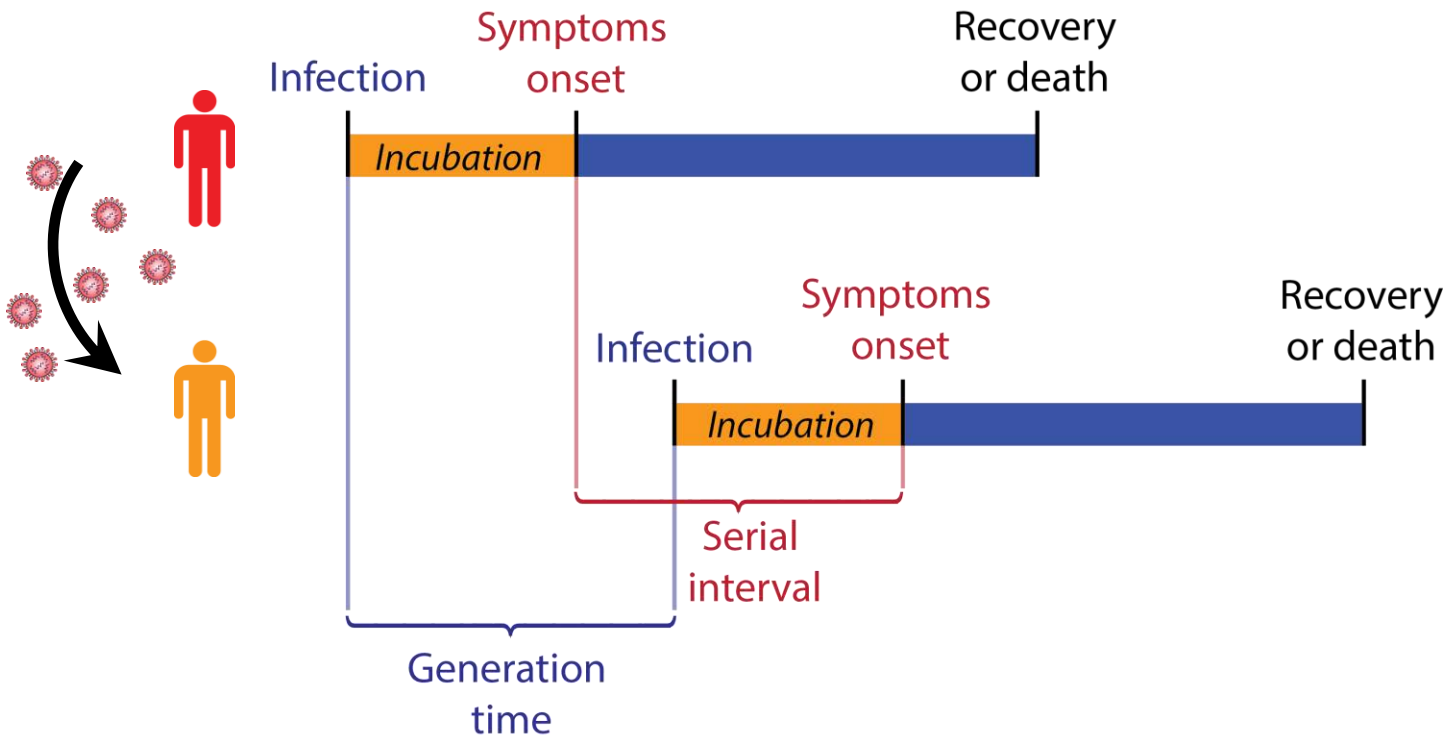
- Contract tracing and cohort follow-up.
- Fitting mathematical models to epidemic curves.

# Timescale of disease transmission

Reproductive number describes the number of secondary cases but not how long it takes for infections to occur.

Generation time: Time between successive infections.

Serial interval: Time between symptom onset of successive infections.

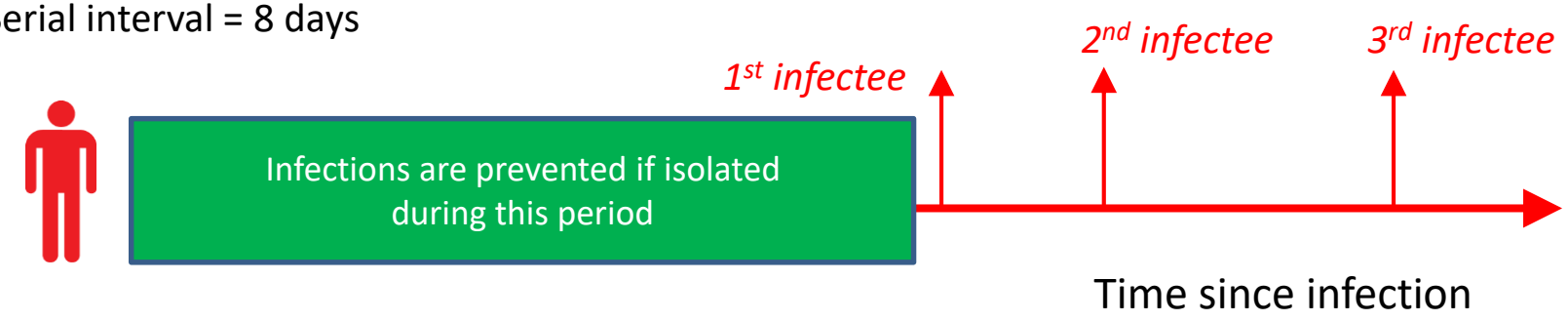


# Timescale of disease transmission

Shorter generation time or serial interval means that confirmed and probable cases would need to be identified and removed from the population sooner in order to prevent them from spreading the disease.

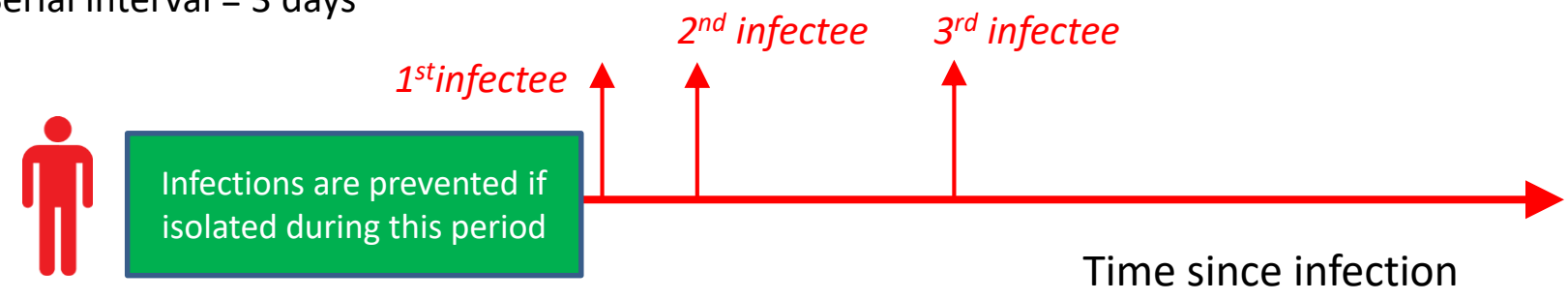
**SARS**

Serial interval = 8 days




**Influenza**

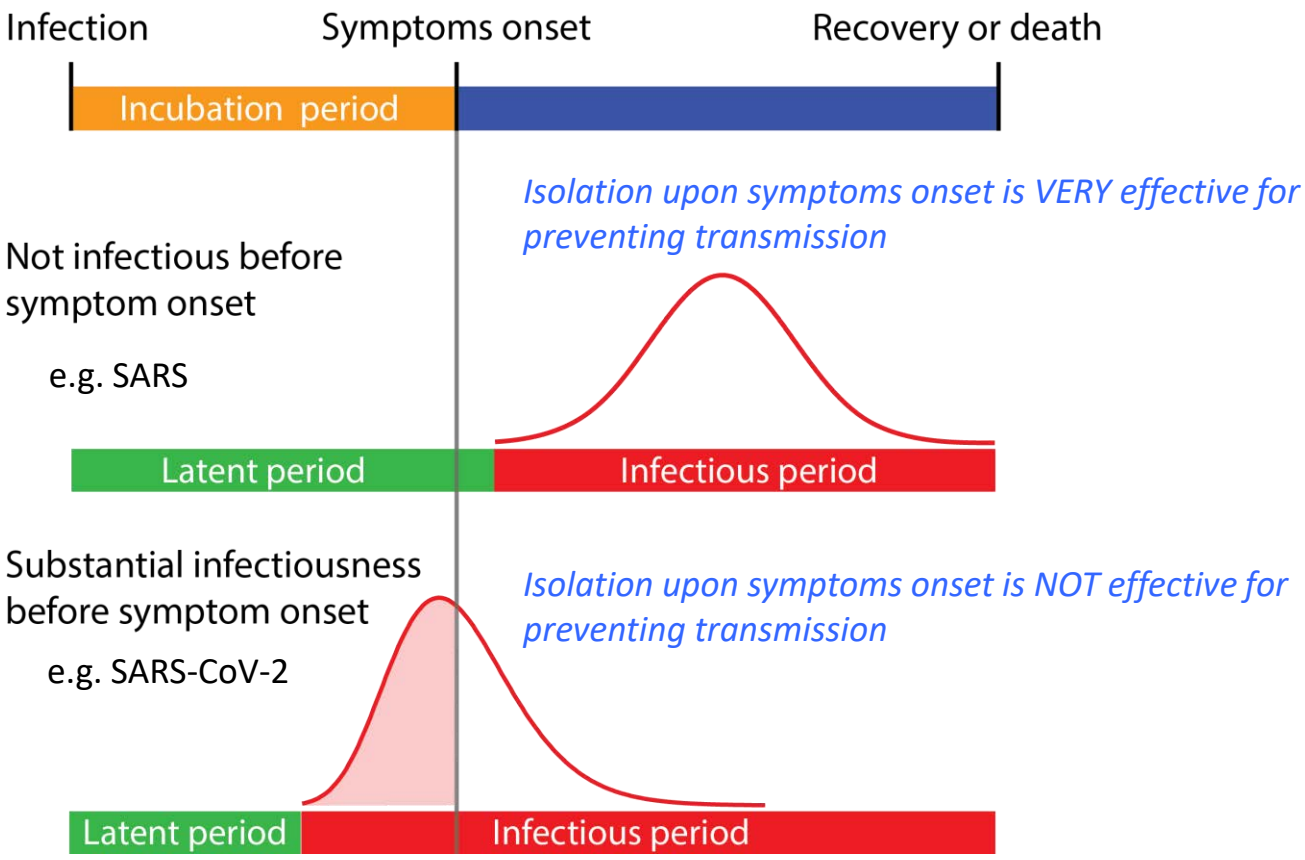
Serial interval = 3 days



# Timescale of disease transmission

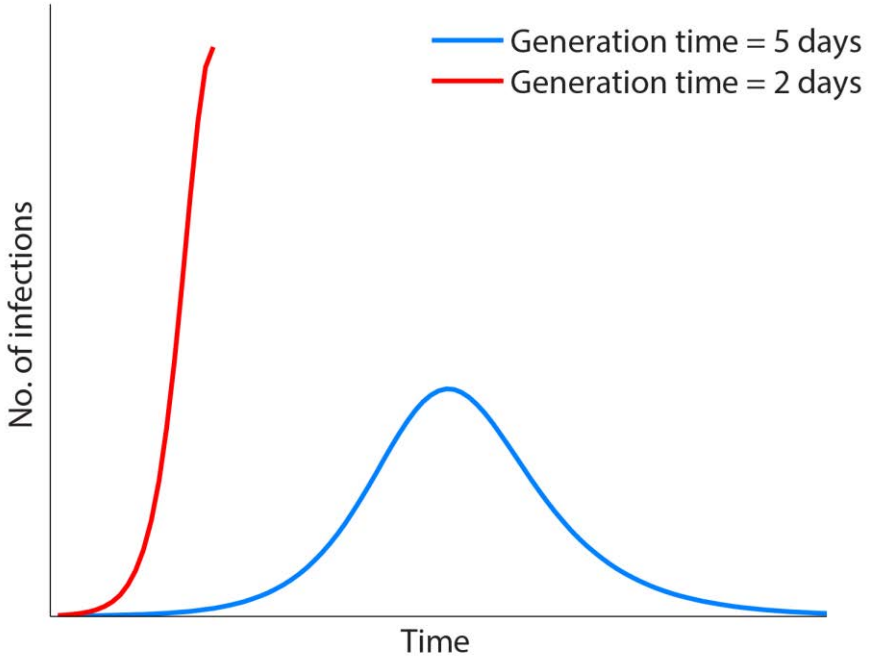
Latent period

- The time it takes to become infectious after infection.
- Can be longer or shorter than incubation period.



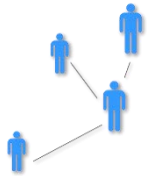
Generation time = 5 days

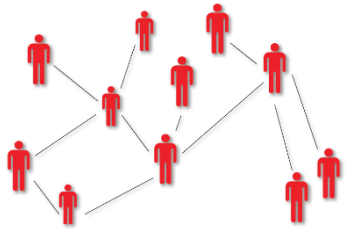
Generation time = 2 days

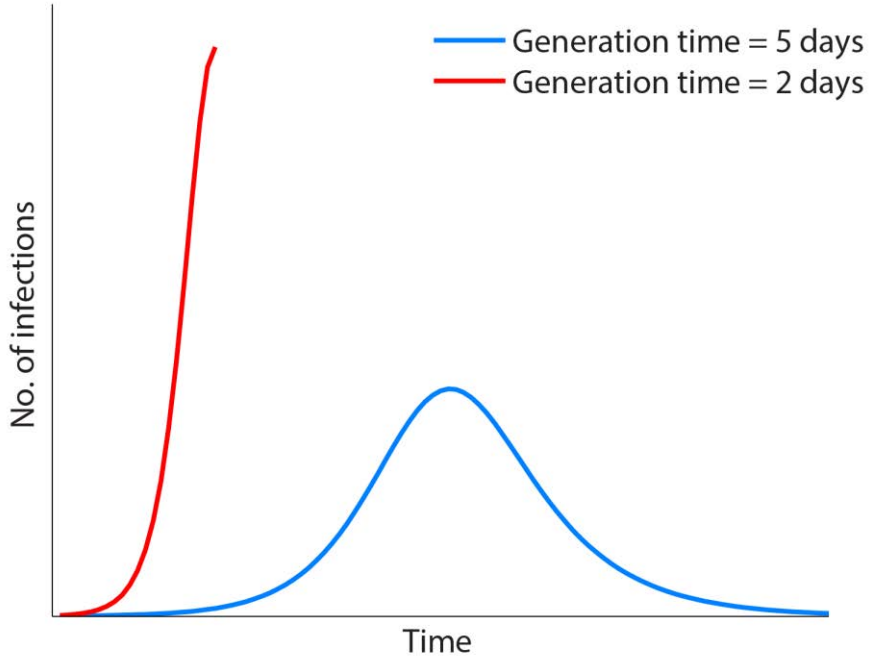


Generation time = 5 days

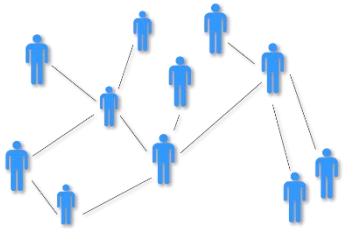
Generation time = 2 days



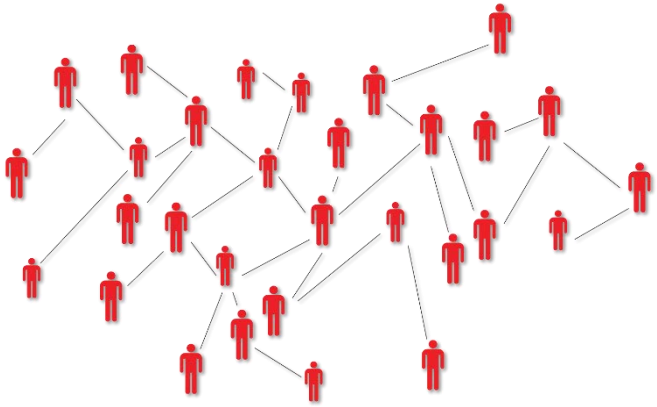


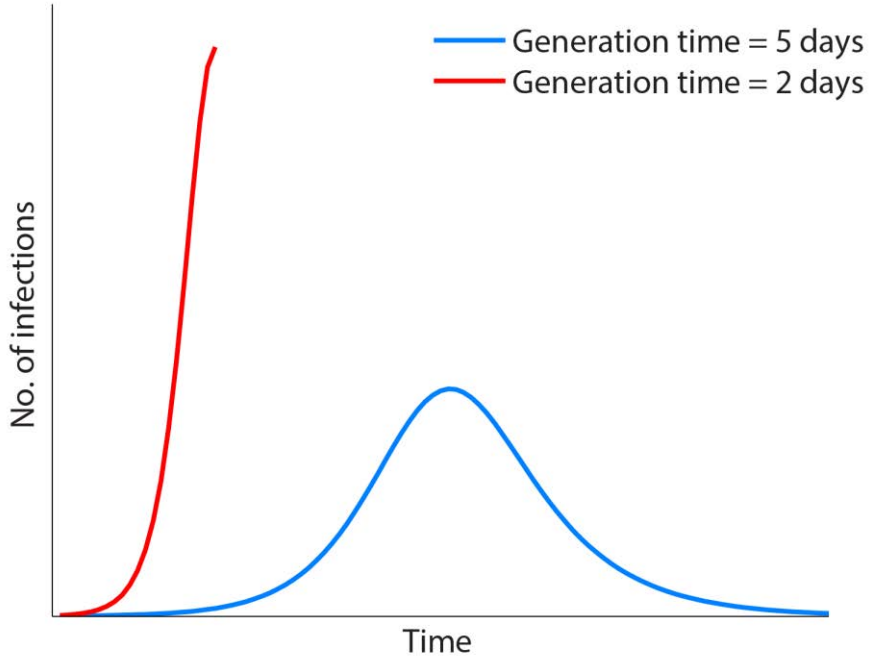


Generation time = 5 days

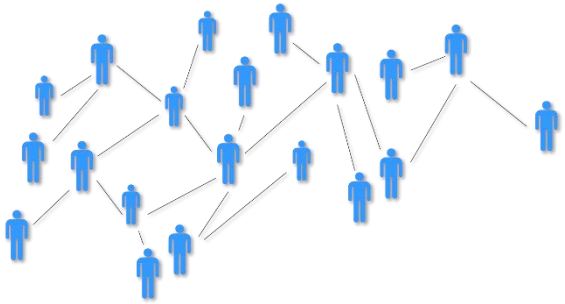


Generation time = 2 days

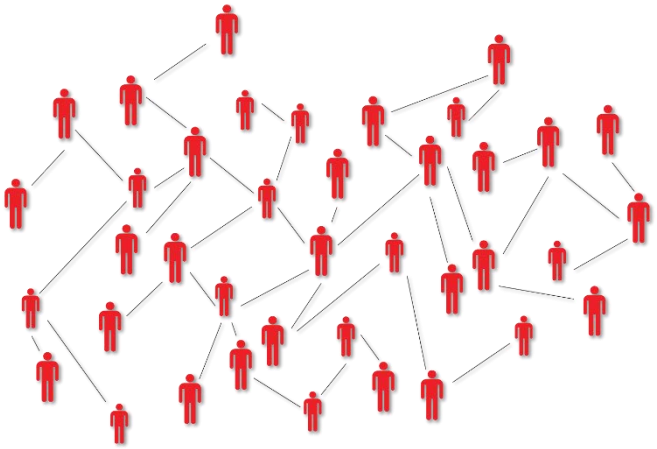


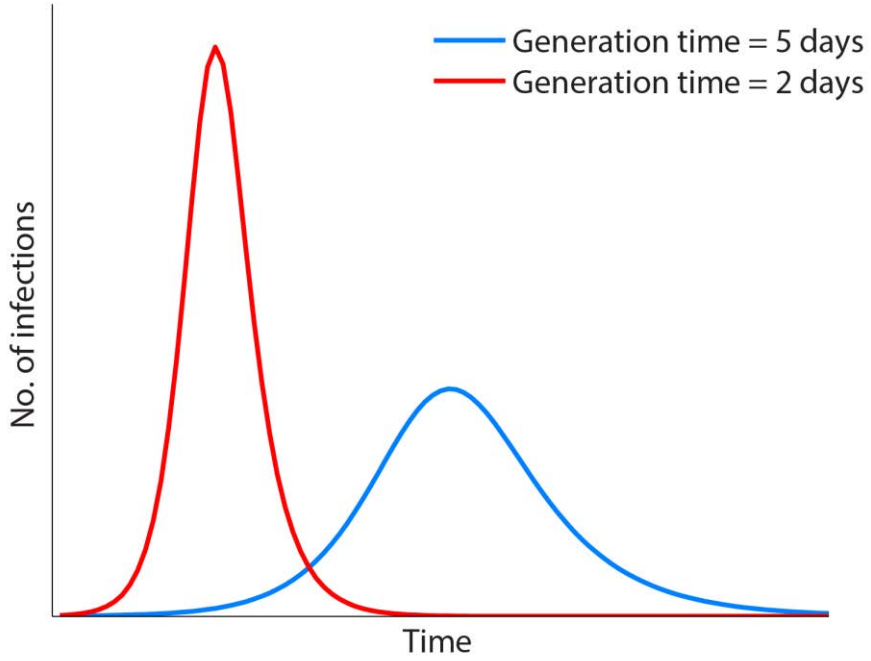


Generation time = 5 days

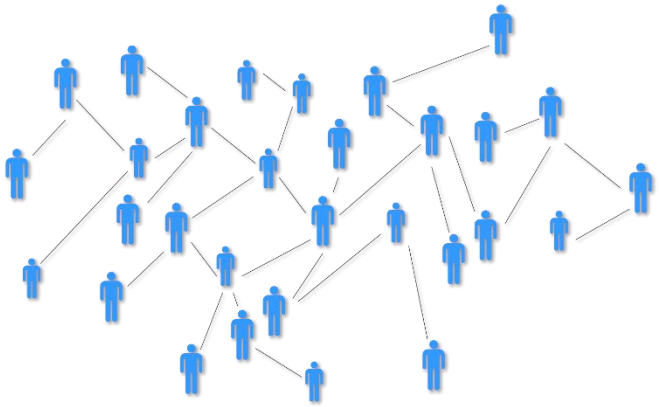


Generation time = 2 days

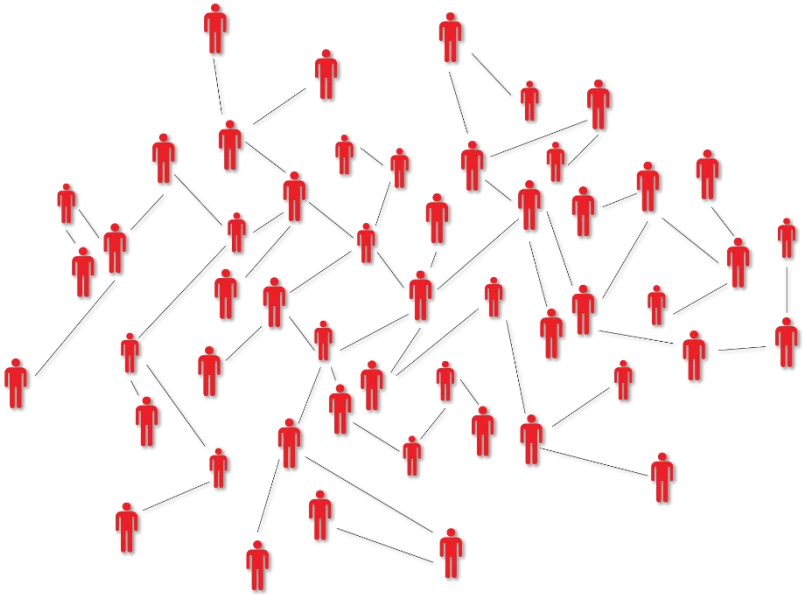


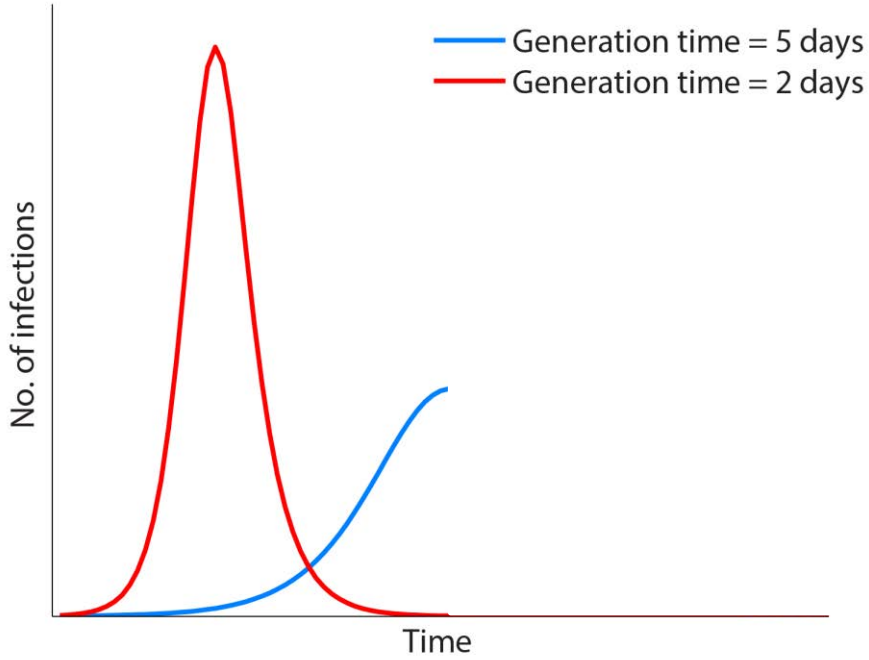


Generation time = 5 days

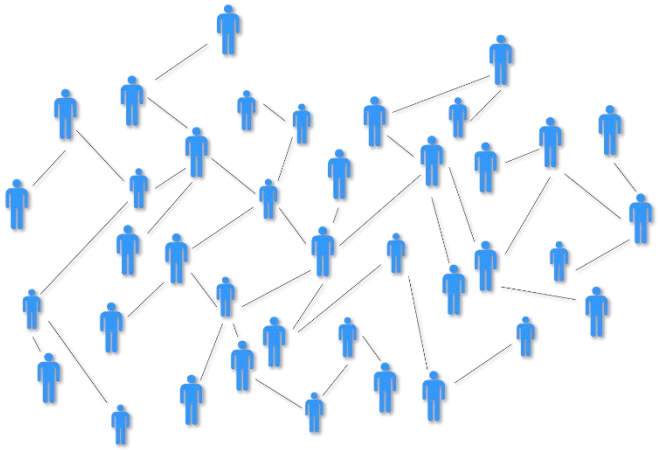


Generation time = 2 days

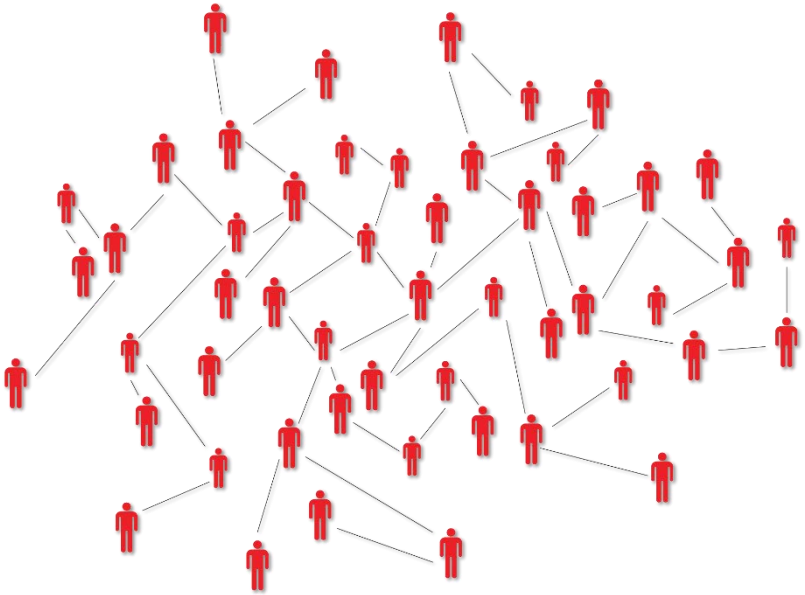


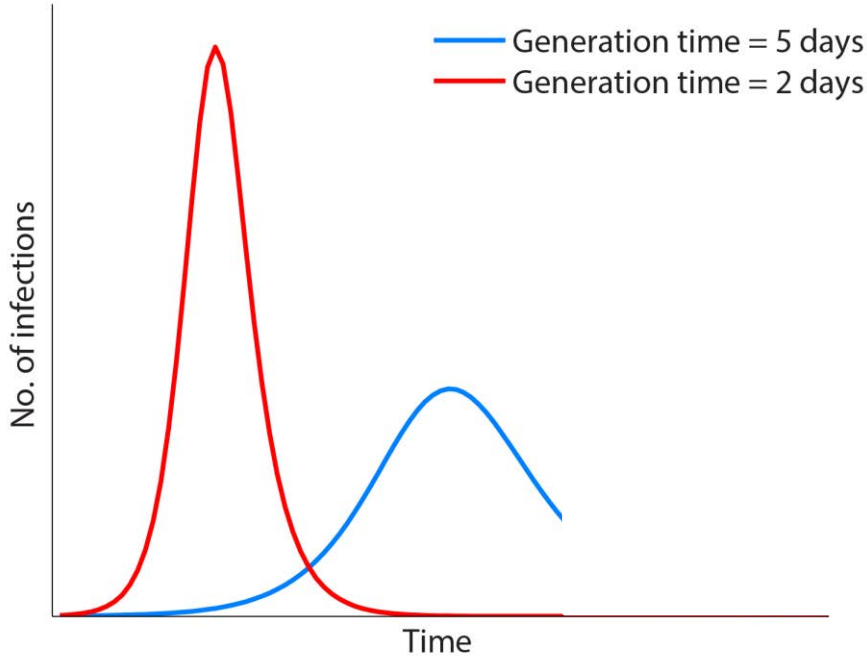


Generation time = 5 days

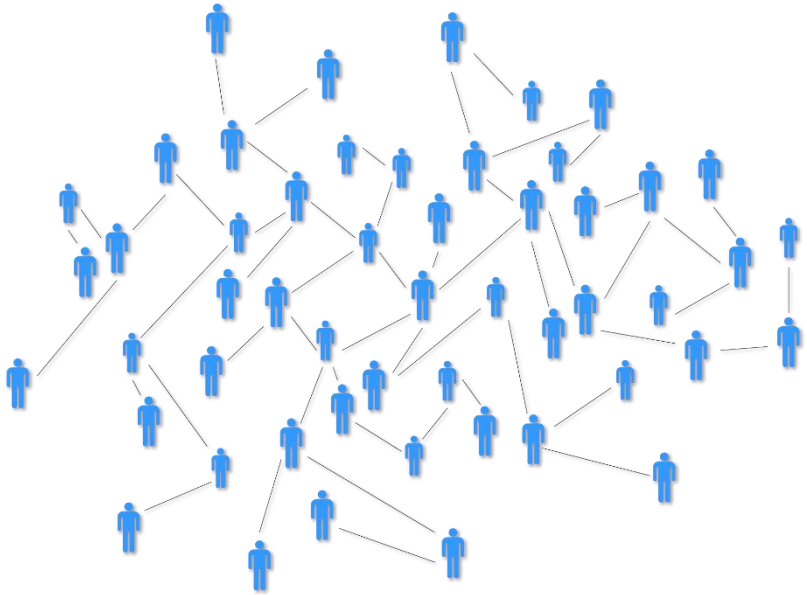


Generation time = 2 days

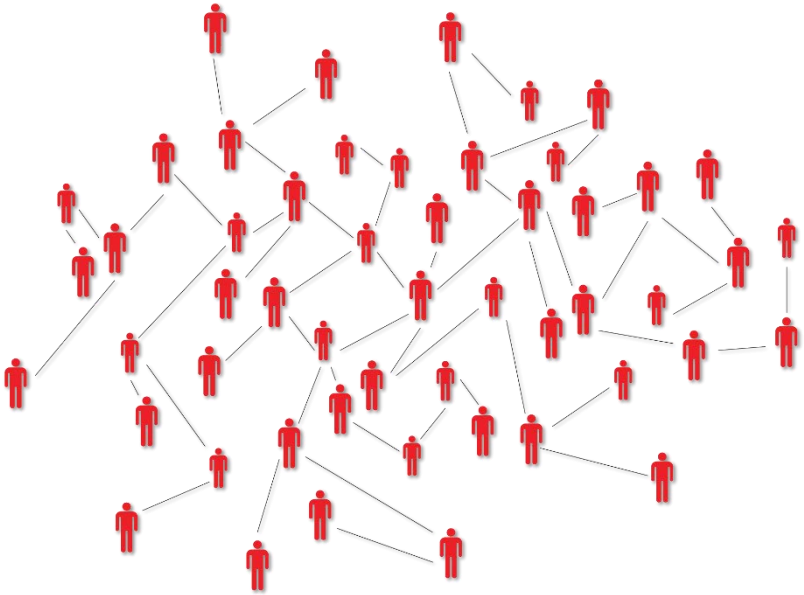


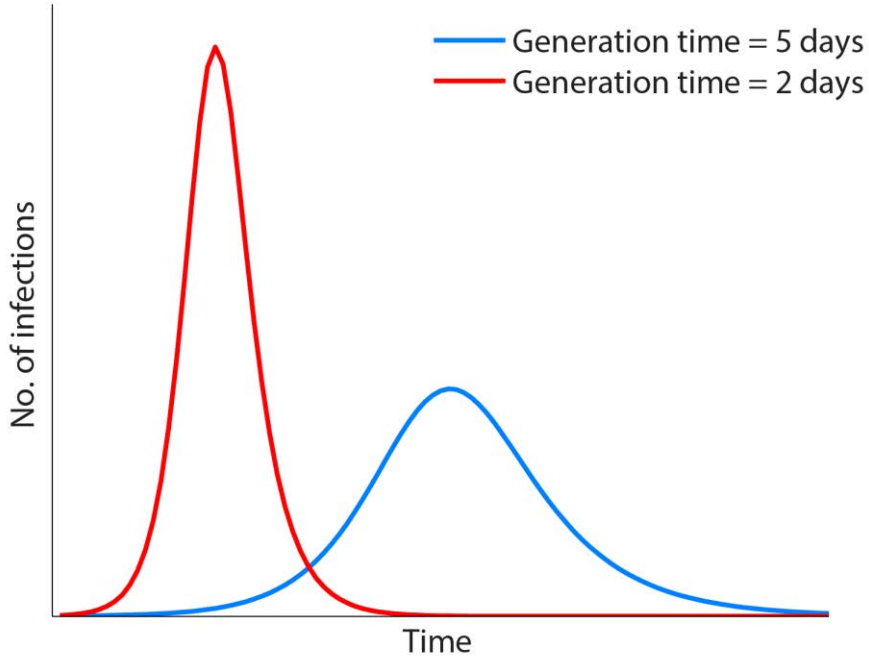


Generation time = 5 days



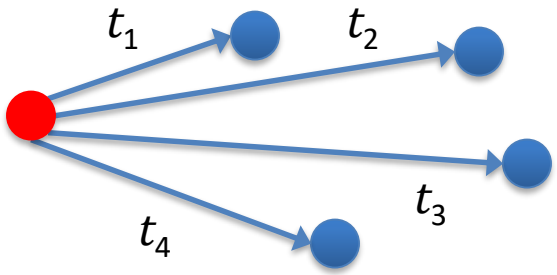
Generation time = 2 days





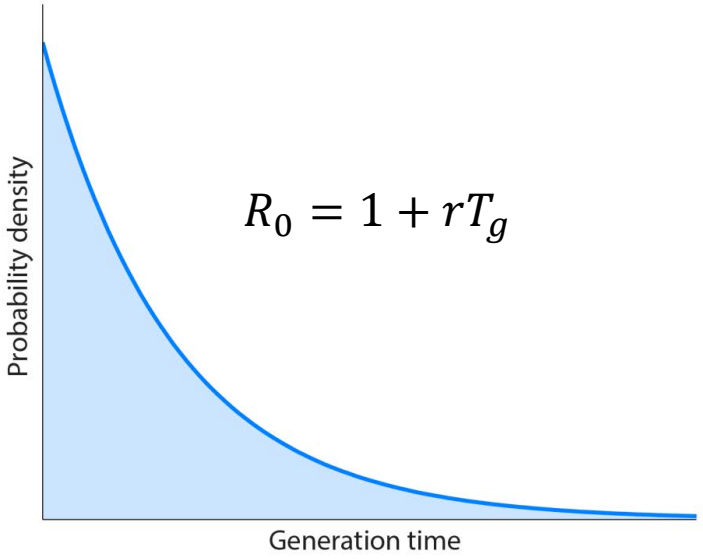
# Epidemic growth rate depends on $R_0$ and $T_g$

- Mean generation time $T_g$: the average time it takes an index case to infect other individuals after he becomes infected.

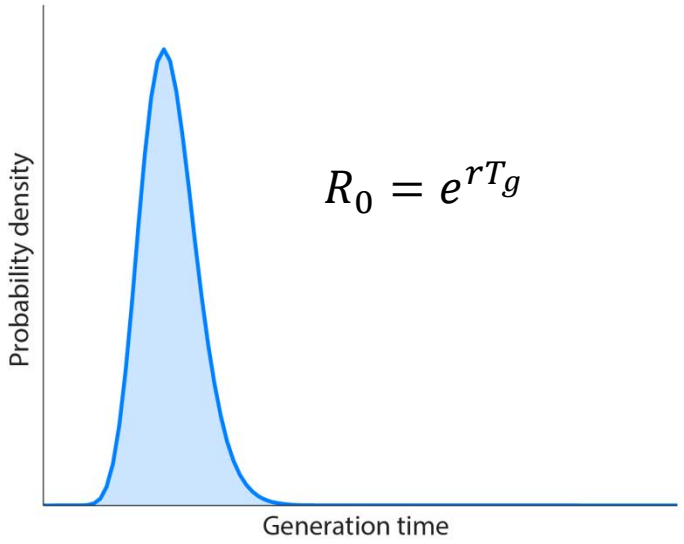



$T_g$ = the average of $t_1$, $t_2$, $t_3$, $t_4$

- $R_0$, $T_g$ and epidemic growth rate $r$ are interrelated. For a given $R_0$, a shorter $T_g$ means a higher growth rate.




$$R_0 = 1 + rT_g$$



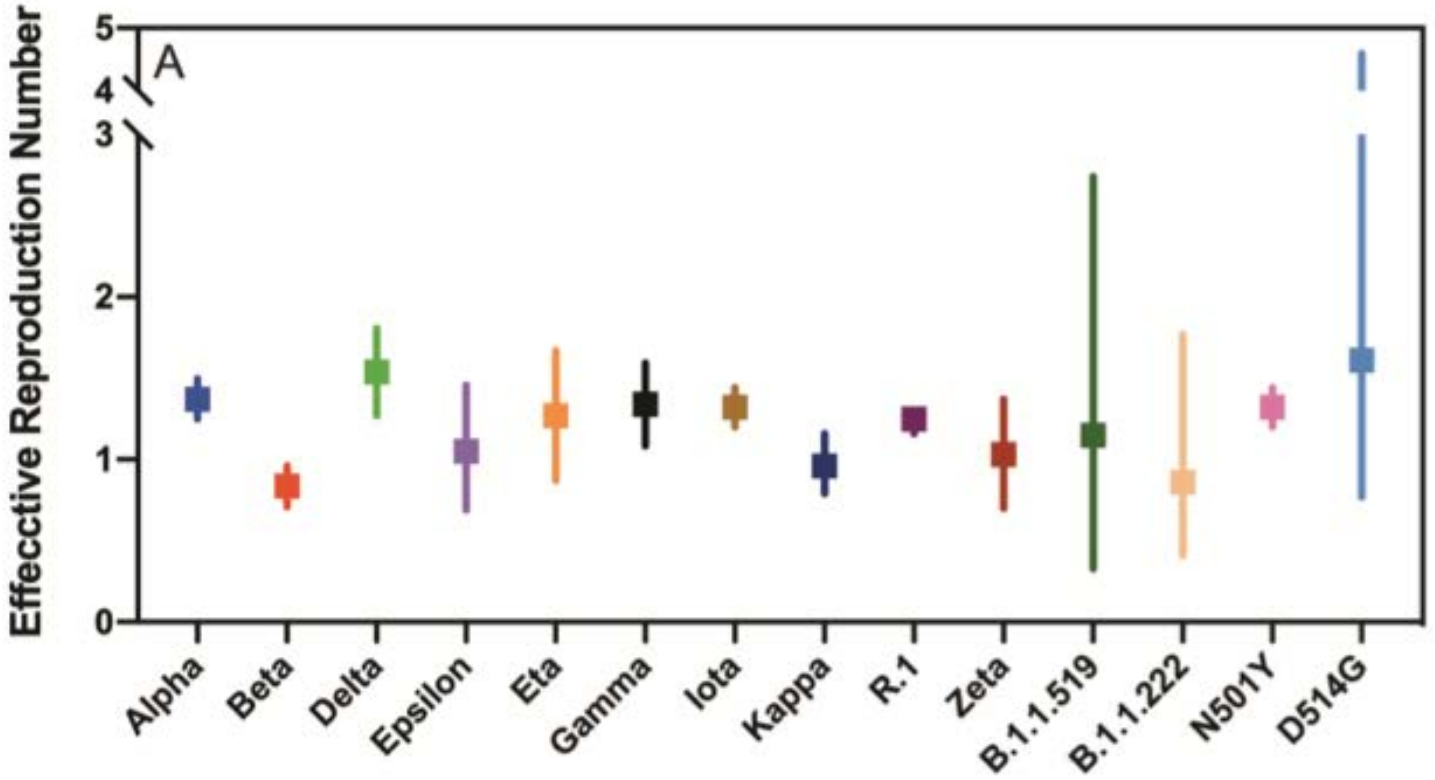

$$R_0 = e^{rT_g}$$

# Basic reproductive number and mean serial intervals of some pathogens

| Disease | $R_0$ | Mean serial interval (days) |
|---|---|---|
| Pandemic influenza | 1.5 – 3 | 2.5 |
| SARS | 2 – 3 | 8 |
| Measles | 12 – 19 | 12 |
| Smallpox | 4 – 7 | 15 |
| Chickenpox | 4 – 9 | 14 |

Heffernan et al *Royal Society Interface* 2005
Fine *Am J Epidemiol* 2003

# Basic reproductive number of variant of SARS-COV-2




Du *Clinic infect dis.* 2022

# Infectious disease modeling

- A systematic way of translating assumptions and data regarding disease transmission into quantitative description of how an epidemic evolves.

- Just like animal models, infectious disease modeling is a tool for helping you collect your thoughts about a complex problem to generate and test hypotheses.

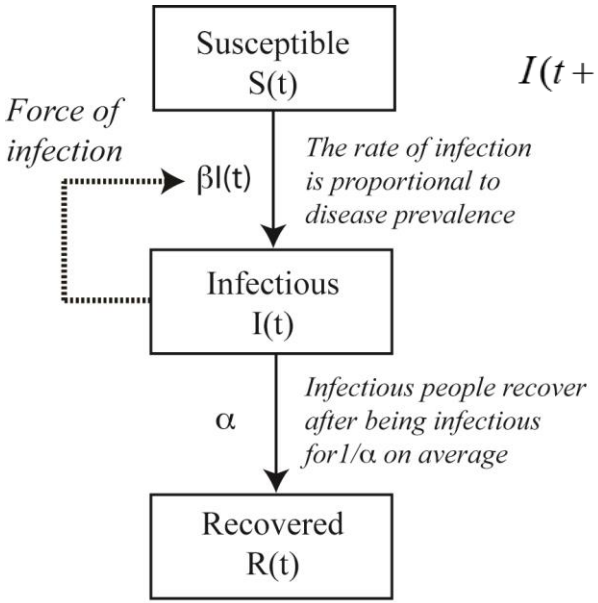
- Some uses of infectious disease modeling:
  - Estimating transmission parameters from epidemiologic data
  - Estimating effectiveness of interventions from epidemiologic data, e.g. vaccination or school closure
  - Predicting the cost and effectiveness of intervention strategies, e.g. vaccinating core transmission group vs high-risk group

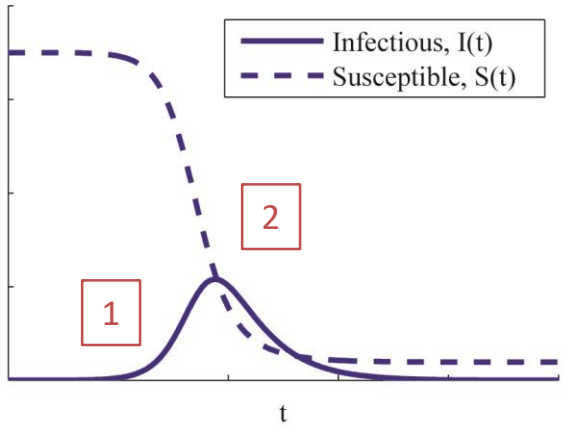
# The simplest epidemic model

- A closed population of size $N$ is partitioned into three compartments of individuals: "Susceptible", "Infectious", and "Recovered" (SIR model).
- All individuals are assumed to be identical in terms of their (i) susceptibility to infection, (ii) infectiousness if infected, and (iii) mixing behavior associated with disease transmission (the so-called homogenous mixing assumption).

$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$




**Hallmarks of an epidemic:**

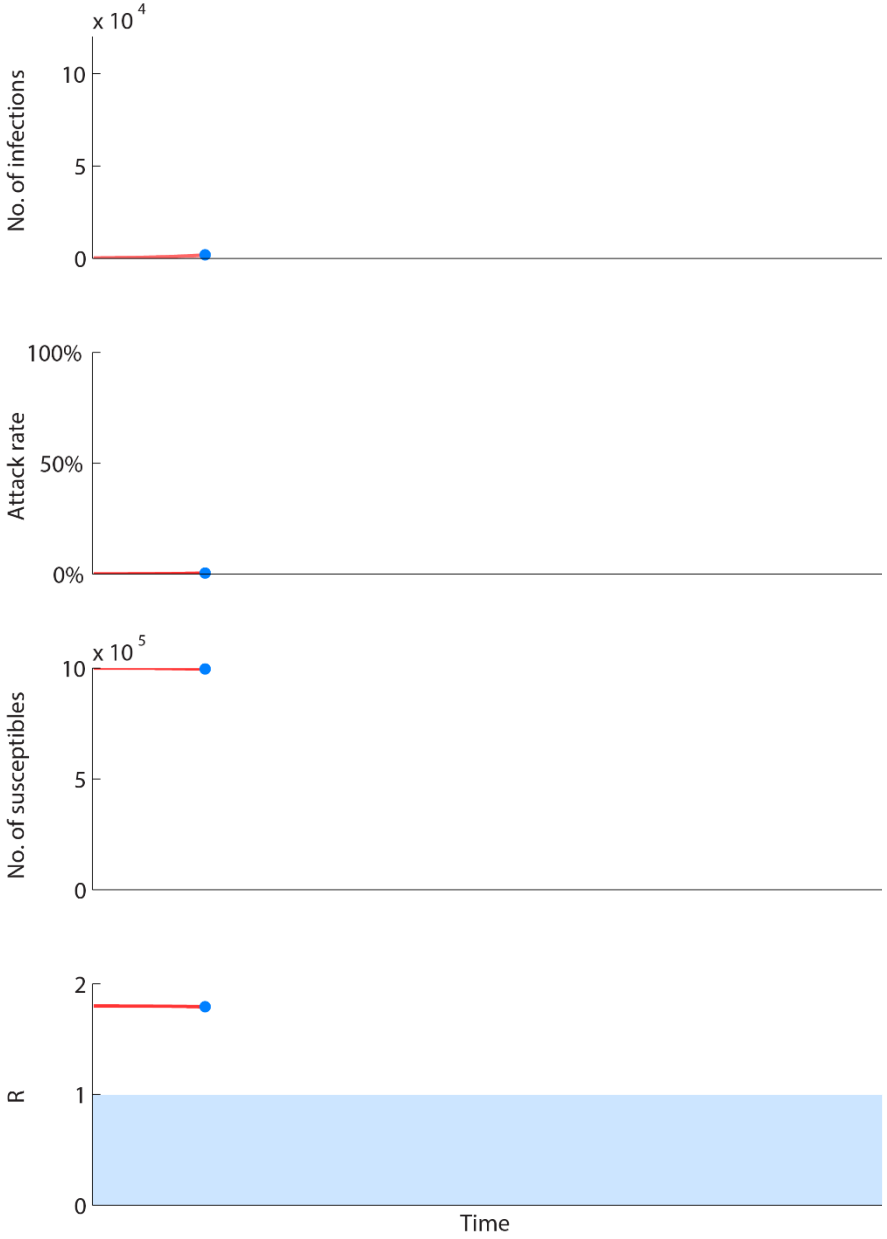
1. The number of infections increases exponentially during the early phase of a growing epidemic
2. The epidemic curve is unimodal and peaks when the susceptible pool has been sufficiently depleted (such that $R_t < 1$)

**Epidemiologic parameters:**

$R_0 = \beta/\alpha$, $T_g = 1/\alpha$

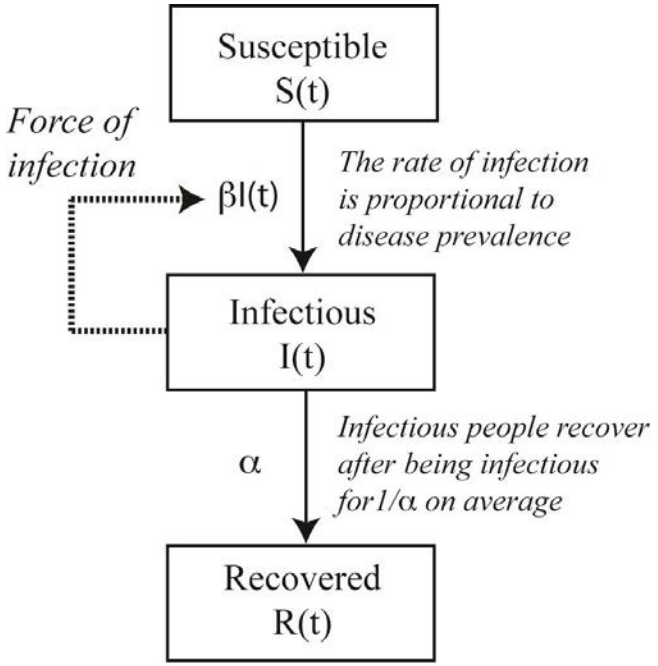
$r = (R_0 - 1)/T_g = \beta - \alpha$

# SIR epidemic dynamics

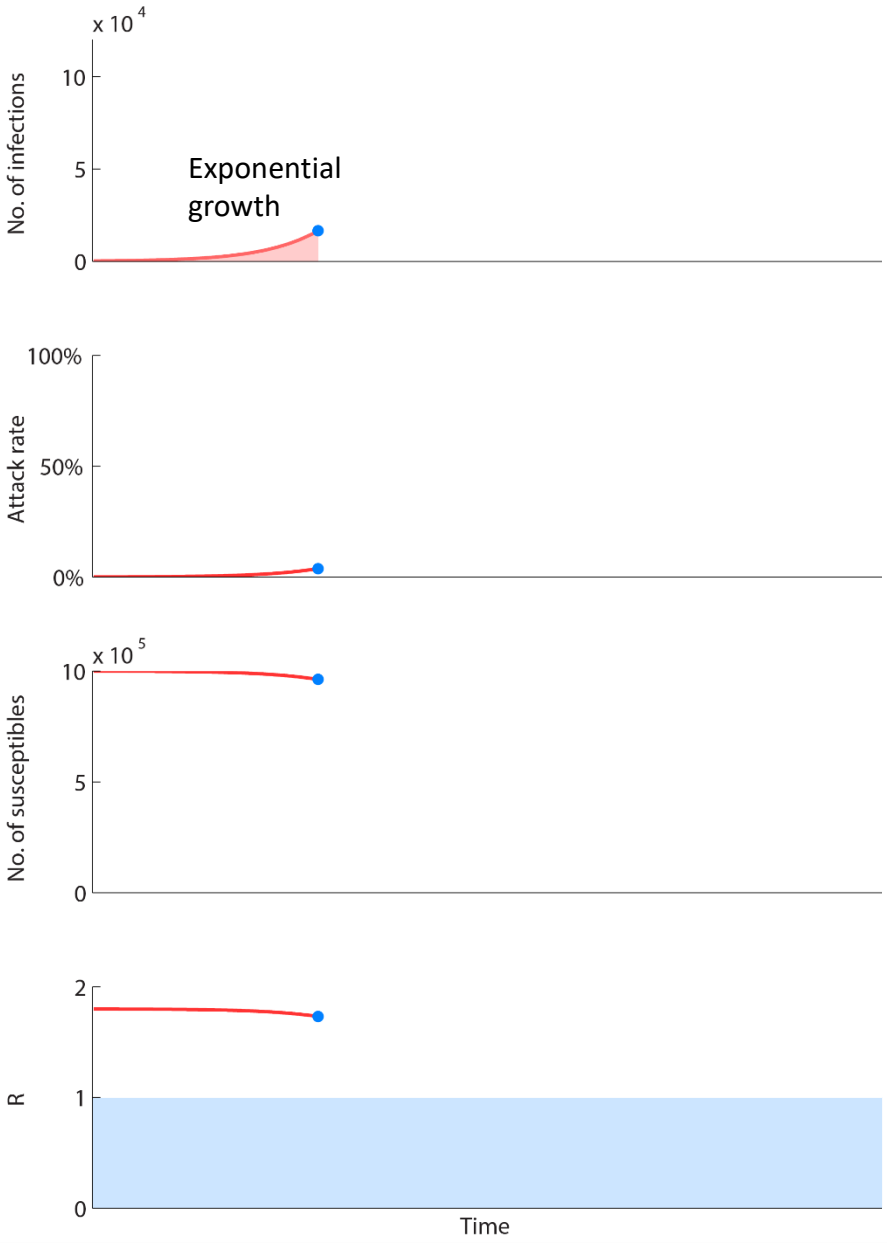



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

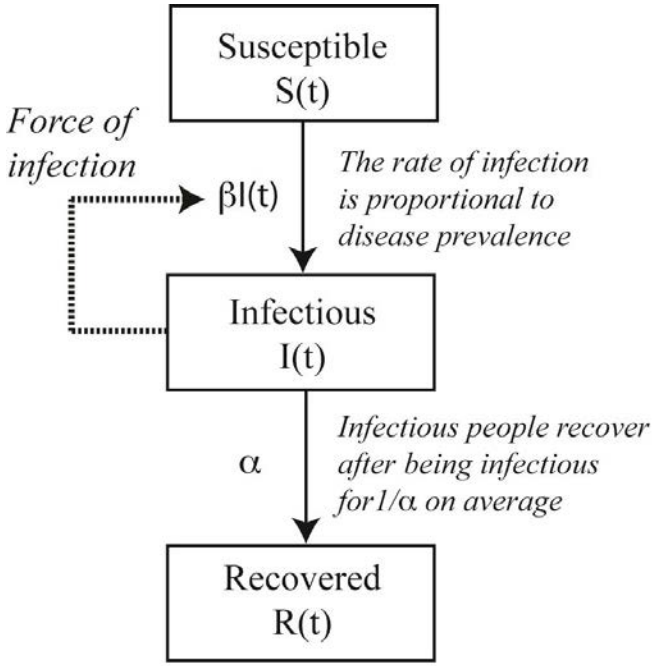


# SIR epidemic dynamics

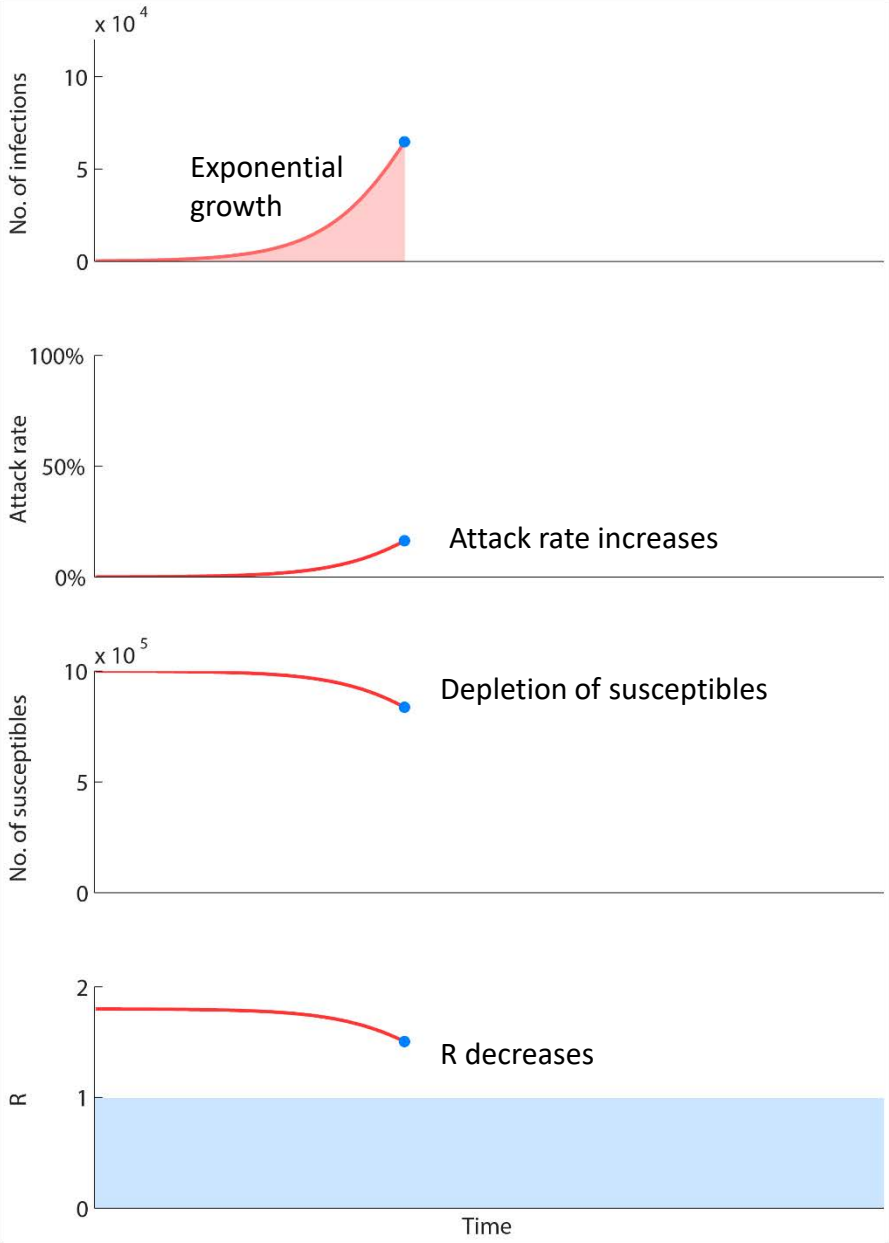



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

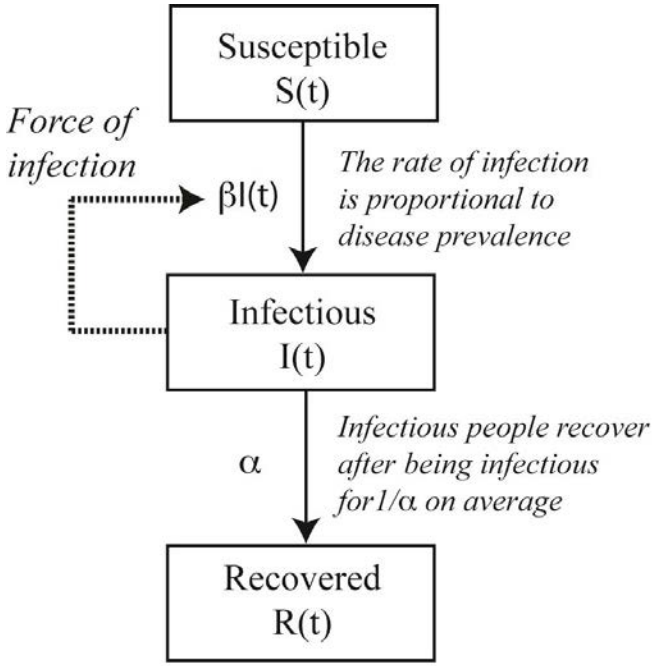


# SIR epidemic dynamics

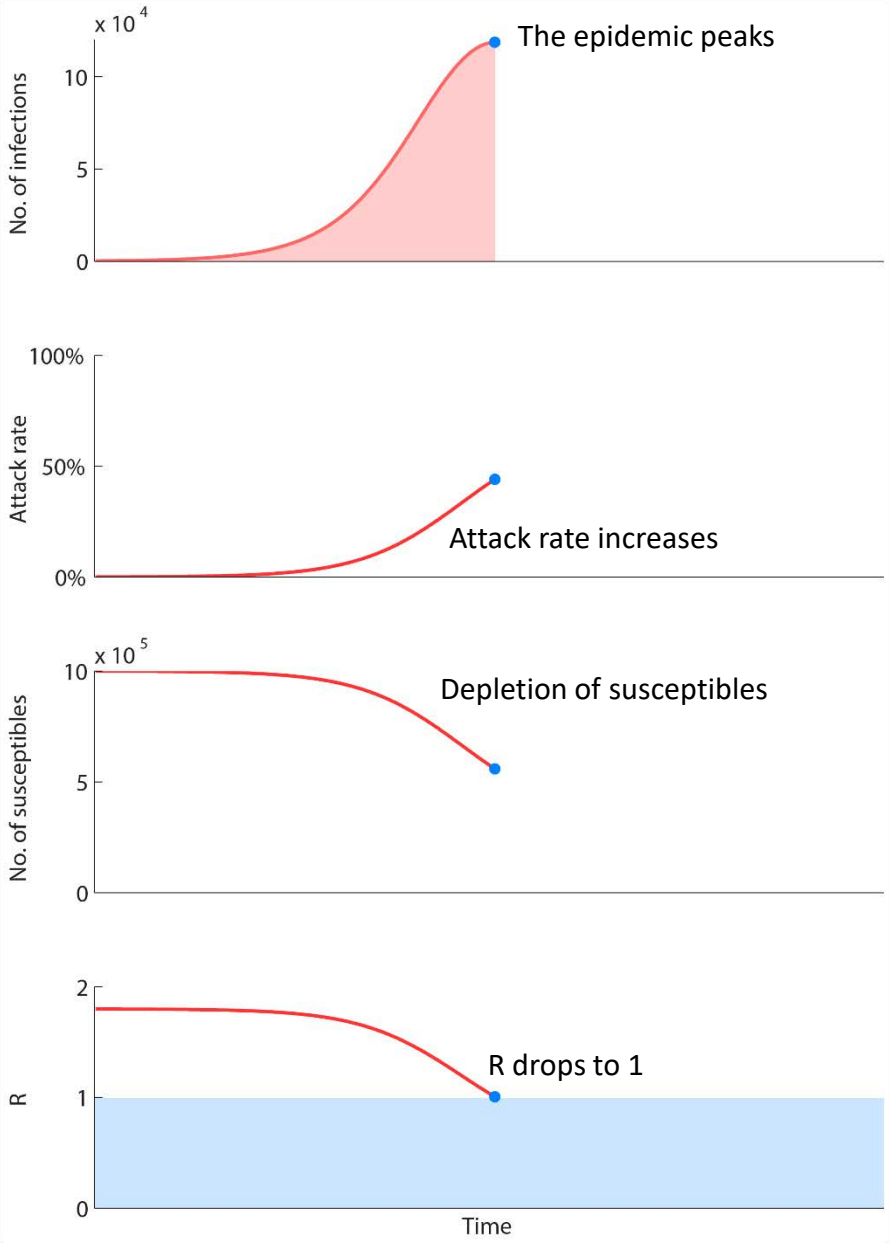



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

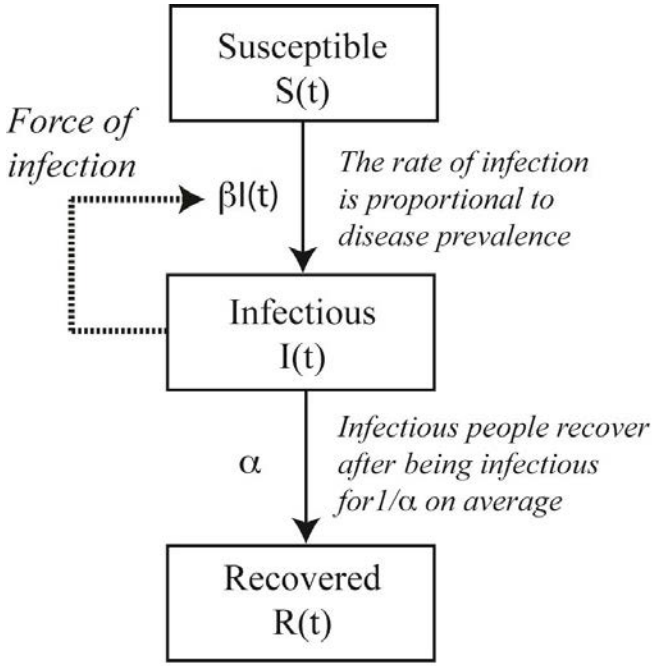


# SIR epidemic dynamics

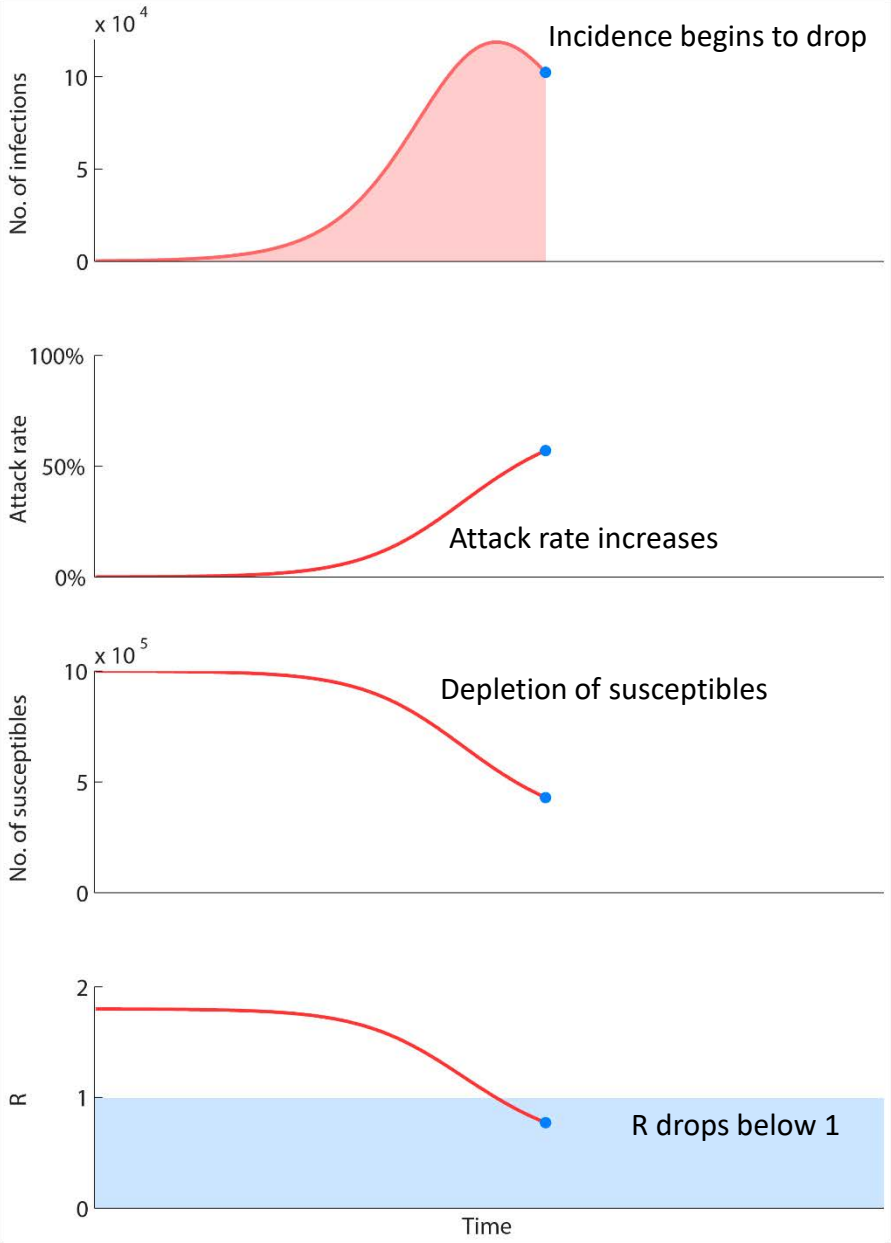



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

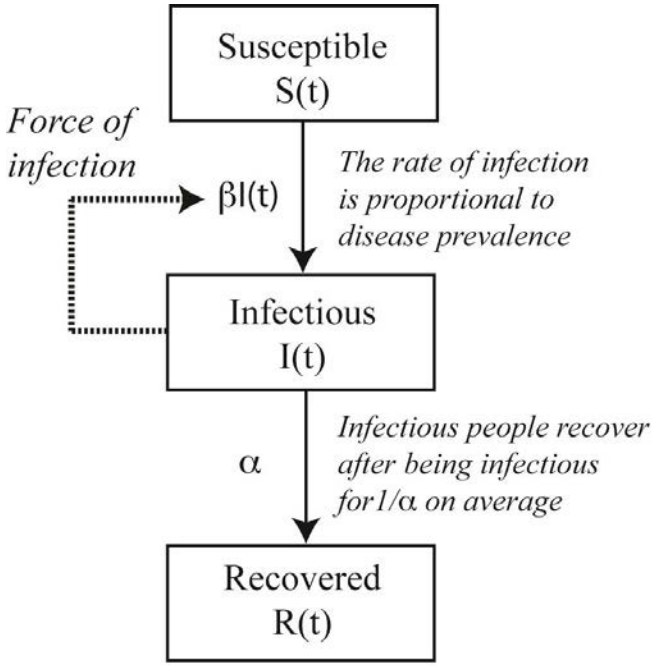


# SIR epidemic dynamics

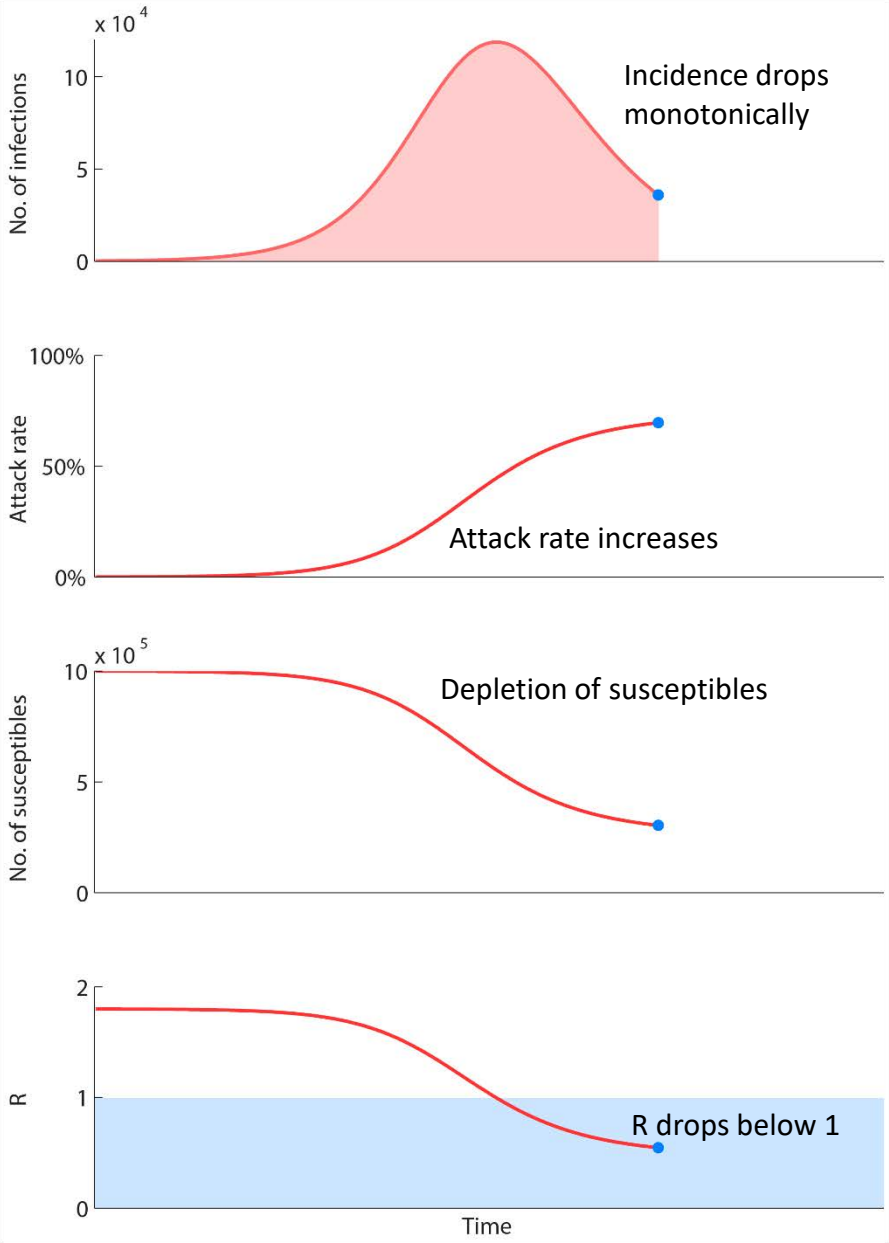



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

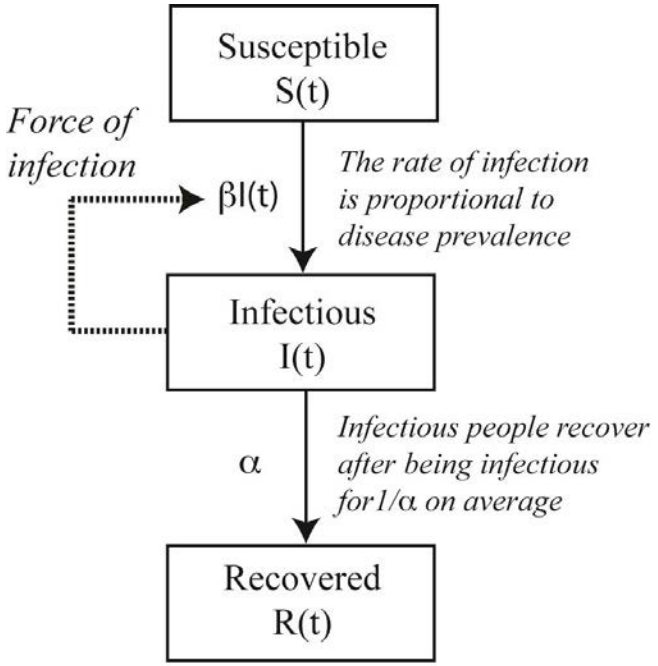


# SIR epidemic dynamics

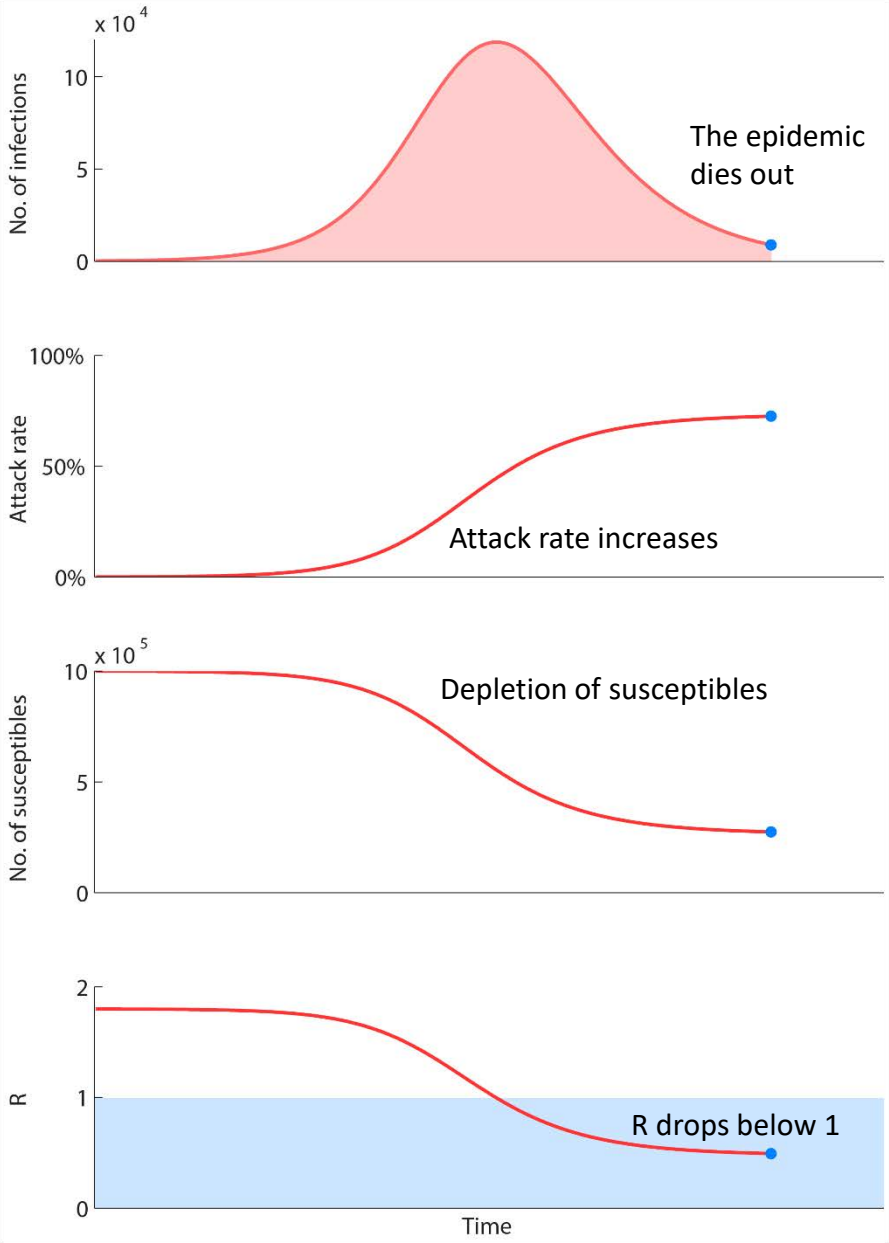



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

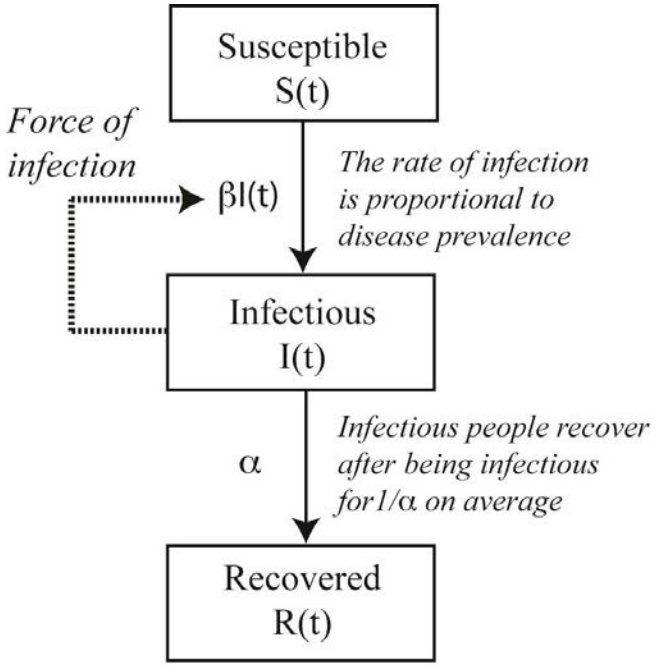


# SIR epidemic dynamics

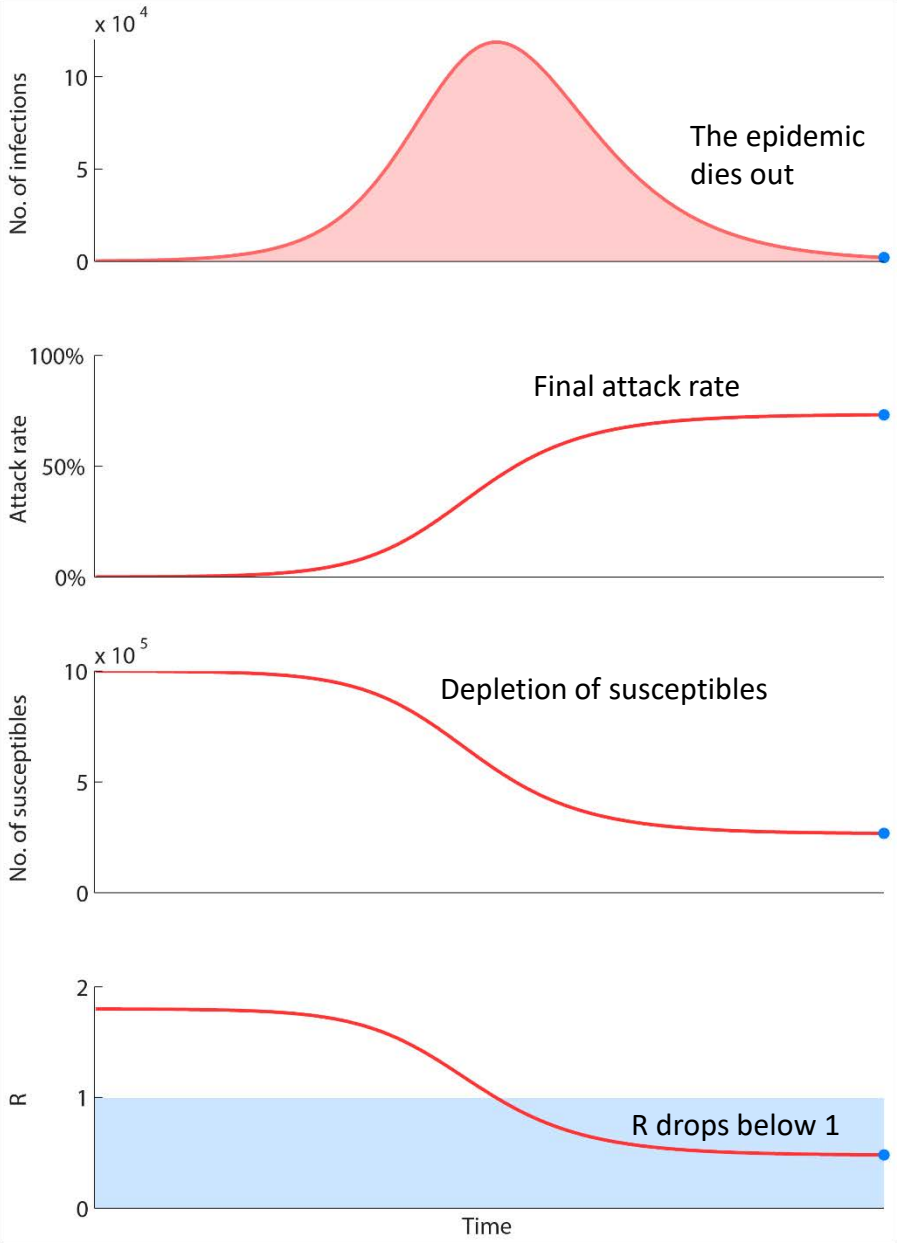



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

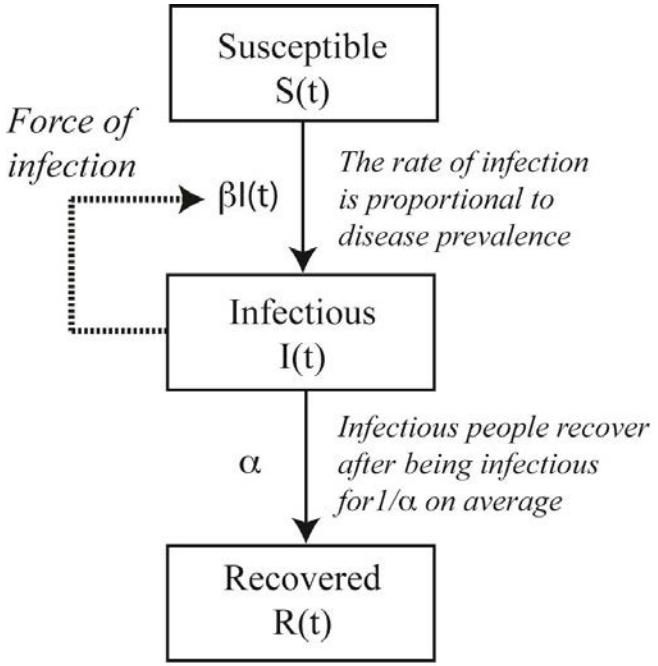


# SIR epidemic dynamics

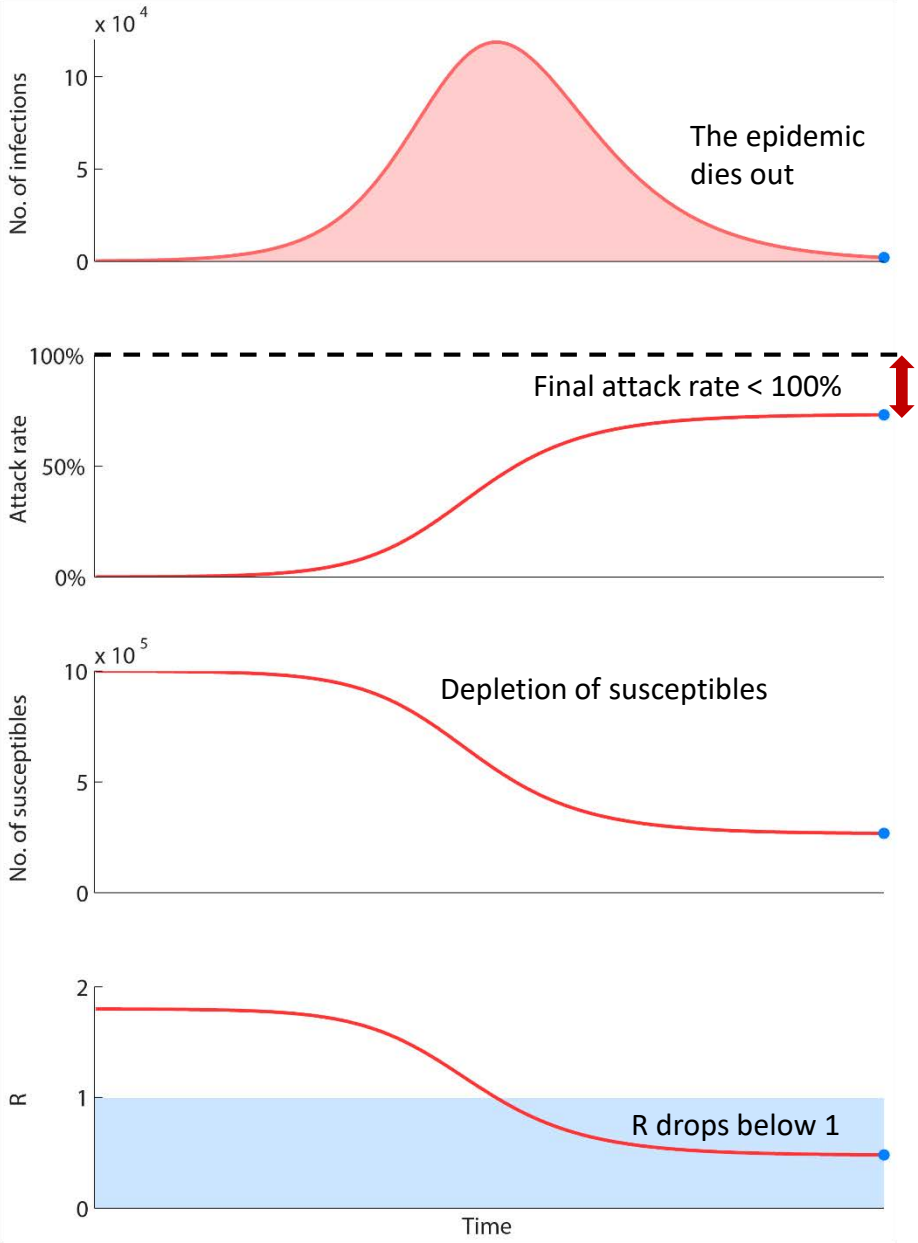



$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$

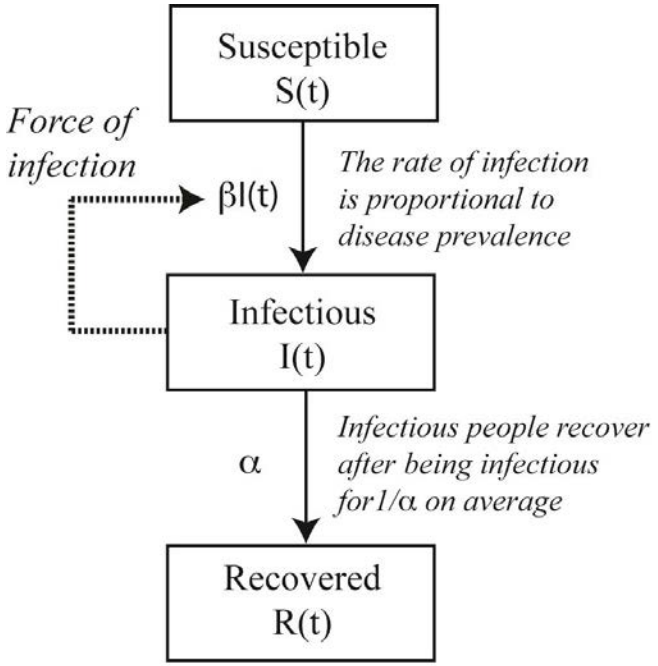


# SIR epidemic dynamics



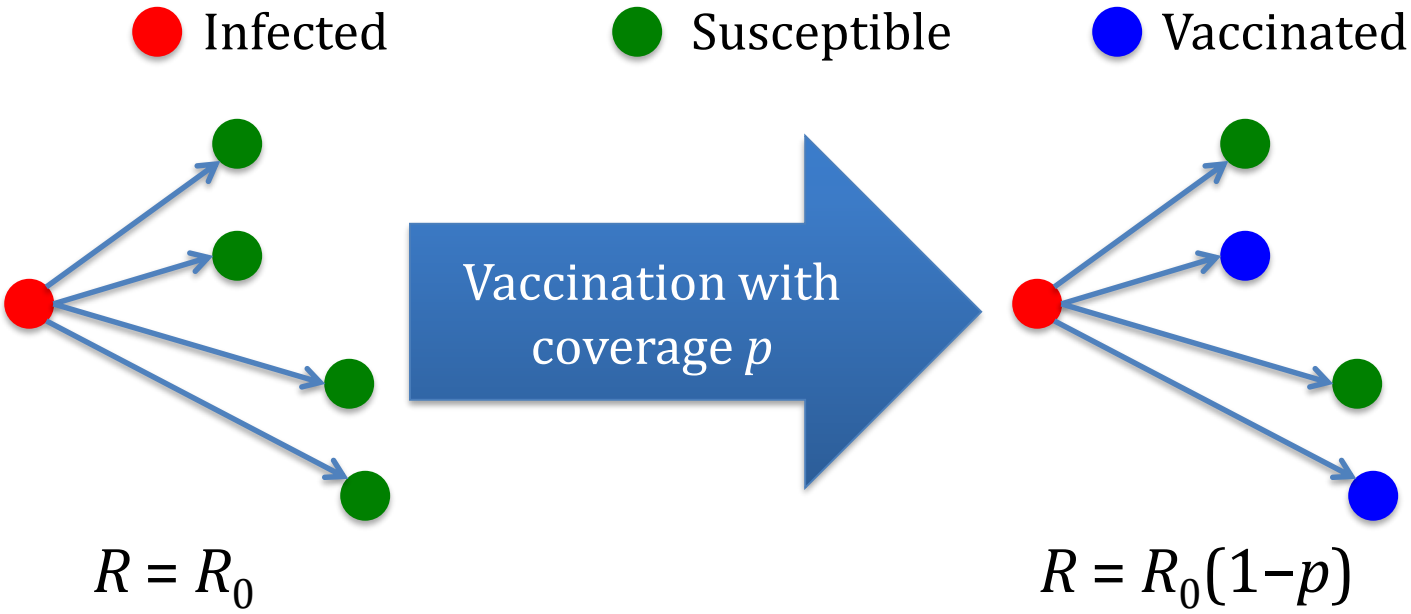

$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$



# Vaccination

- Suppose the vaccine is 100% efficacious.
- It is not necessary to vaccinate the whole population to halt a growing epidemic. Why?
  - Vaccinating an individual indirectly reduces the risk of infection of this individual's contacts (herd immunity).

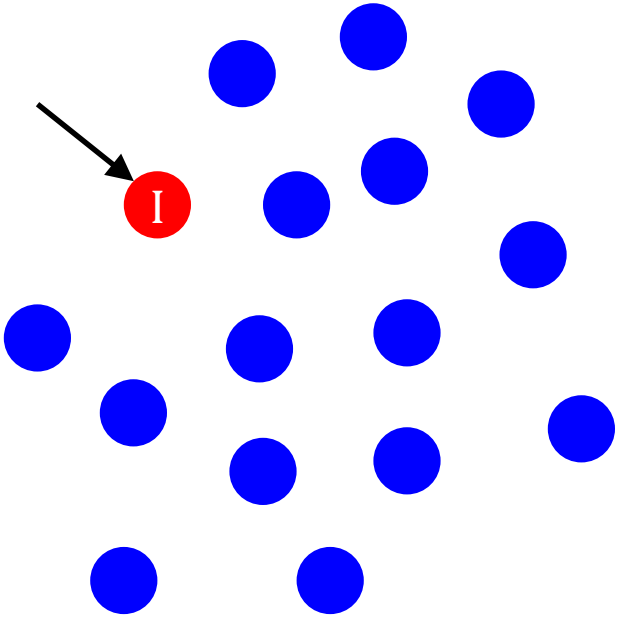


# Herd immunity

Susceptible

Infected

Vaccinated(100% directly protected)

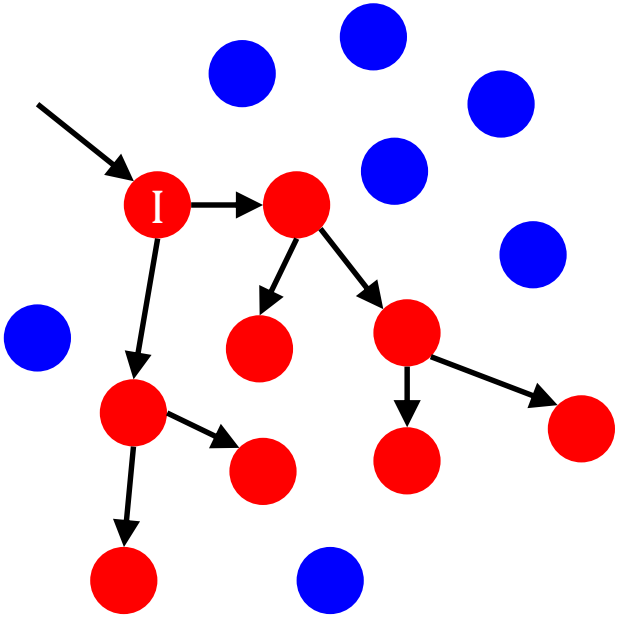


# Herd immunity

Susceptible

Infected

Vaccinated(100% directly protected)

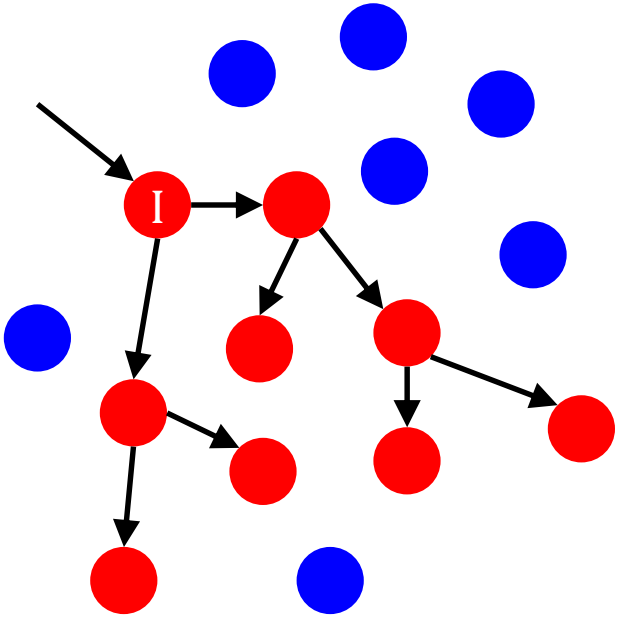


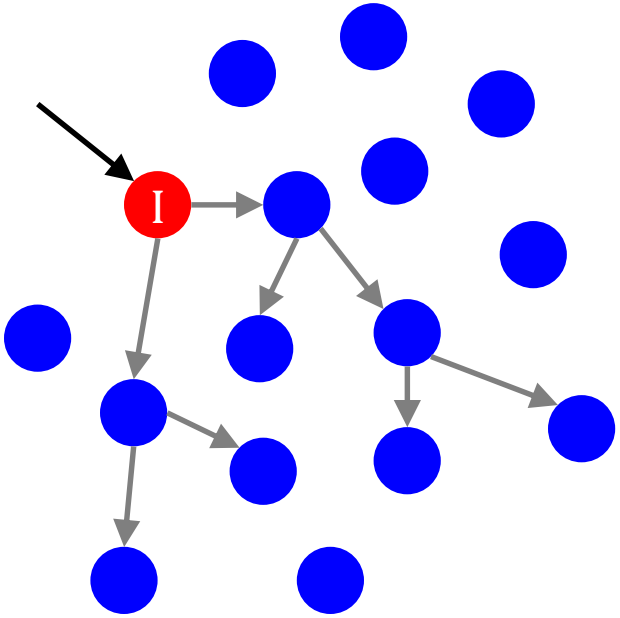
# Herd immunity

Susceptible

Infected

Vaccinated(100% directly protected)



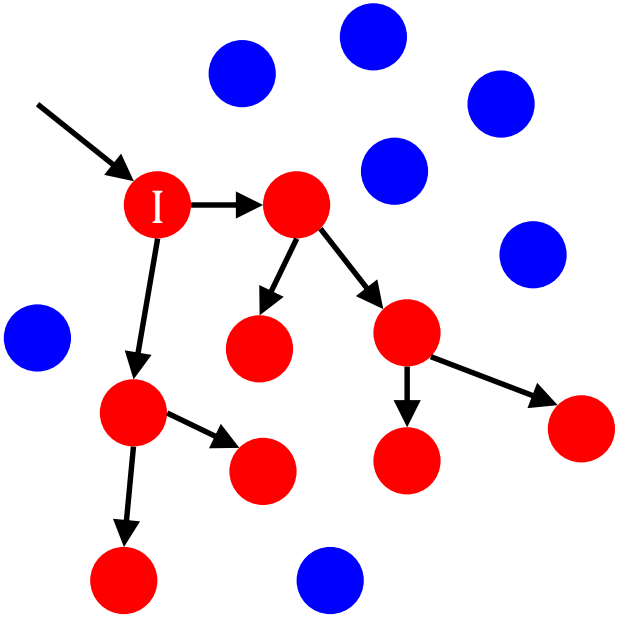



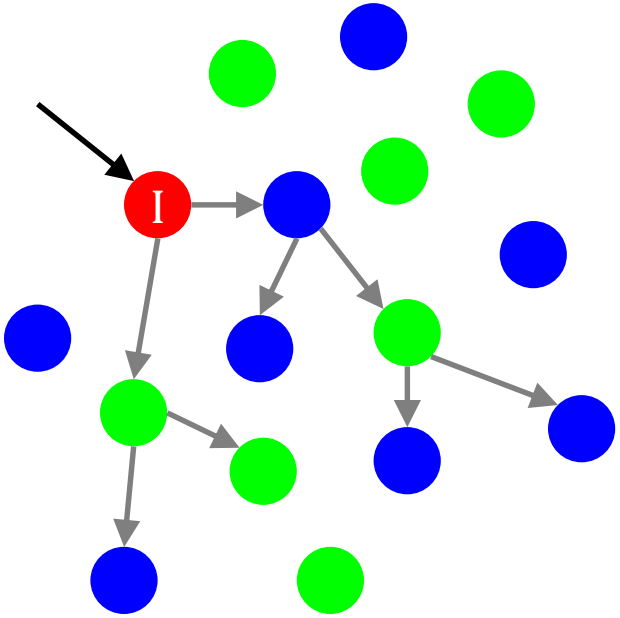
# Herd immunity

Susceptible

Infected

Vaccinated(100% directly protected)



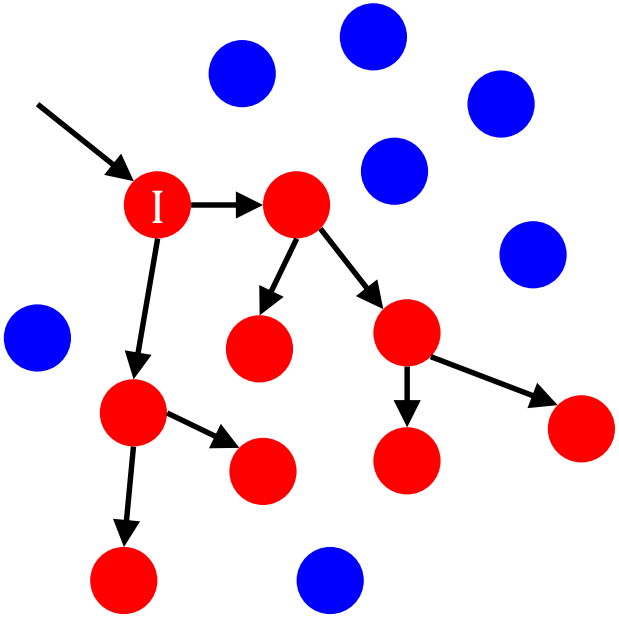



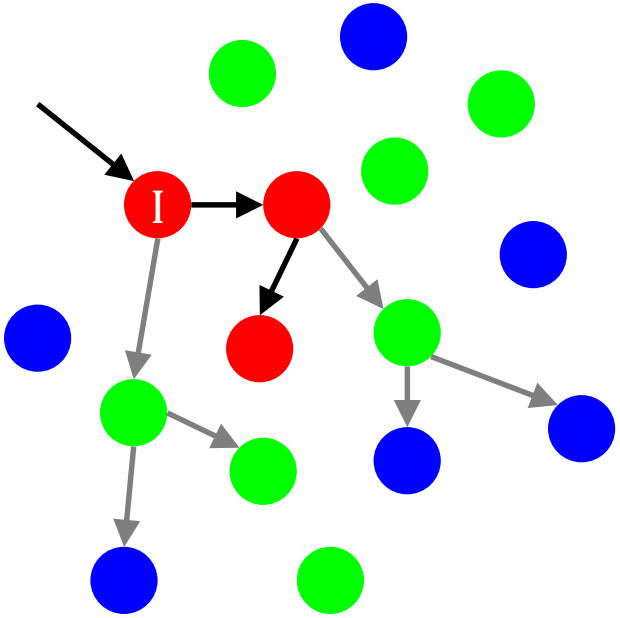
# Herd immunity

Susceptible

Infected

Vaccinated(100% directly protected)



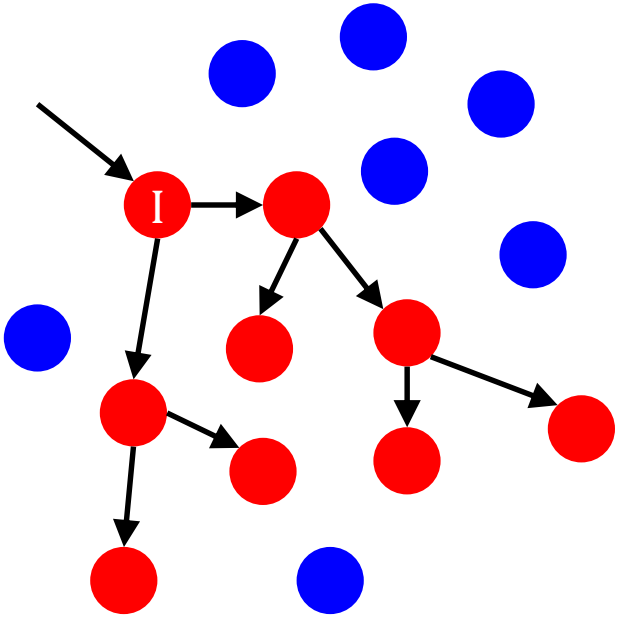
# Herd immunity

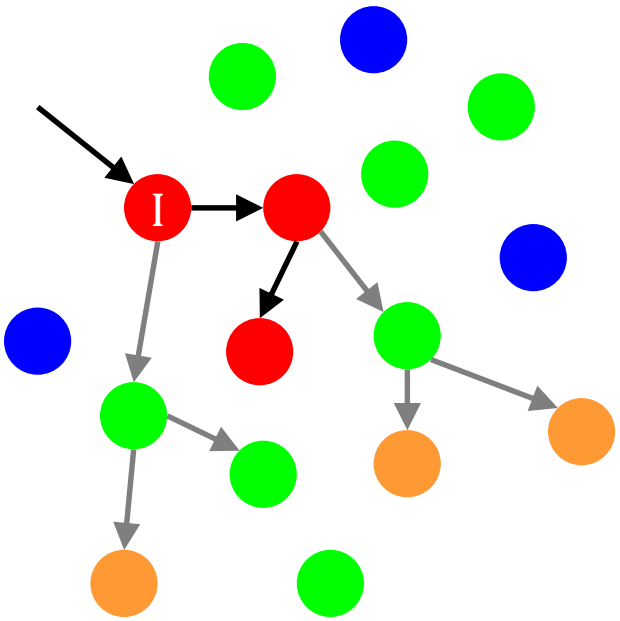
Susceptible

Infected

Vaccinated(100% directly protected)

Indirectly protected



# Critical vaccination coverage

- Suppose the vaccine is 100% efficacious.

- Critical vaccination coverage, $c^*$
  - The minimal proportion needed to be vaccinated in order to prevent an epidemic.
  - Also called the herd immunity threshold

- In the SIR model, $c^* = 1–1/R_0$, i.e. vaccinating a proportion $p > 1–1/R_0$ of the population is sufficient to prevent an epidemic. Why?
  - After vaccination, the reproductive number is

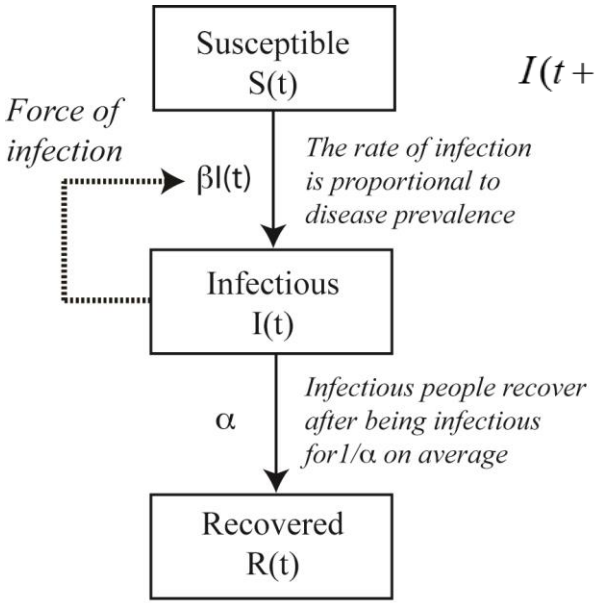
$$R = (1-p)R_0 < 1 \qquad \Leftrightarrow \qquad p > 1 - 1/R_0$$

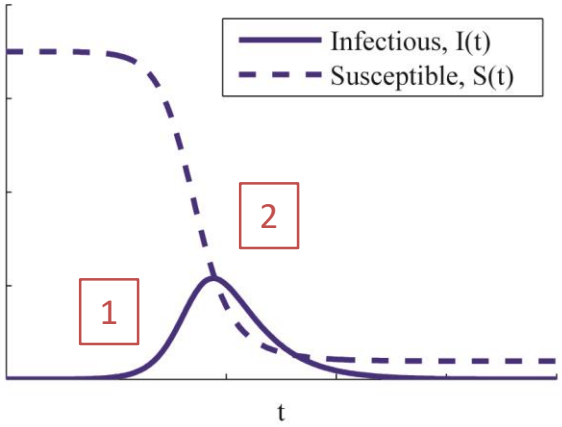
# Recap: The simplest epidemic model

- A closed population of size $N$ is partitioned into three compartments of individuals: "Susceptible", "Infectious", and "Recovered" (SIR model).
- All individuals are assumed to be identical in terms of their (i) susceptibility to infection, (ii) infectiousness if infected, and (iii) mixing behavior associated with disease transmission (the so-called homogenous mixing assumption).

$$S(t+\Delta t) = S(t) - \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t}$$

$$I(t+\Delta t) = I(t) + \underbrace{\beta S(t) I(t) \Delta t}_{\text{number of infections between time } t \text{ and } t+\Delta t} - \underbrace{\alpha I(t) \Delta t}_{\text{number of recovery between time } t \text{ and } t+\Delta t}$$




**Hallmarks of an epidemic:**

1. The number of infections increases exponentially during the early phase of a growing epidemic
2. The epidemic curve is unimodal and peaks when the susceptible pool has been sufficiently depleted (such that $R_t < 1$)

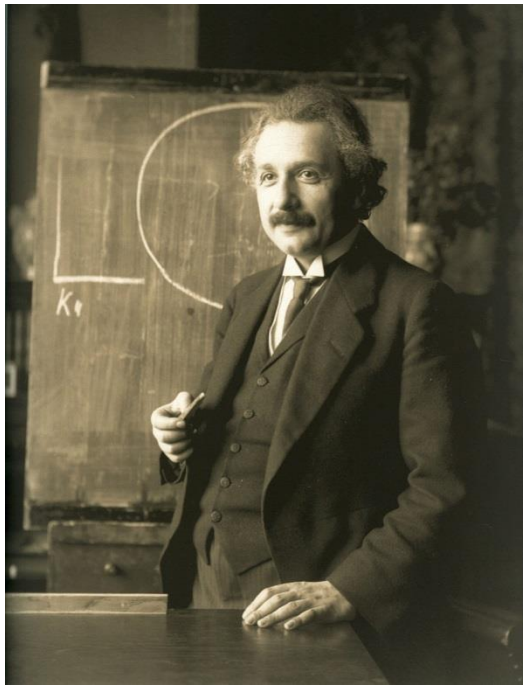
**Epidemiologic parameters:**

$R_0 = \beta/\alpha$, $T_g = 1/\alpha$

$r = (R_0 - 1)/T_g = \beta - \alpha$

# Building more complexities into the model

- Stochasticity ("chance effects")
- More detailed natural history , e.g. asymptomatic infections
- Age structure and multiple populations
- Aging, birth, deaths, immigration, emigration
- Seasonality and other time-dependent forcing
- Individual-based model instead of compartmental model



*"Everything should be made as simple as possible, but not simpler."*

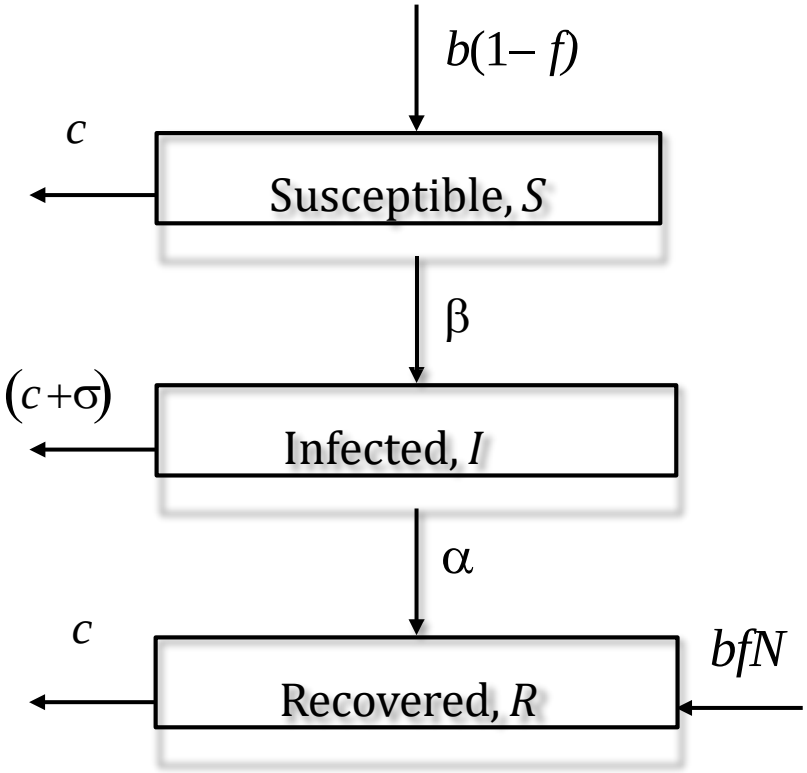
\- Albert Einstein

Nobel Prize in Physics 1921

Models should be:

- Complex enough to provide robust answers
- Simple enough to avoid unnecessary details

# SIR model with vaccination




$b$ = birth rate
$c$ = death rate
$f$ = vaccine coverage

$$\frac{dS}{dt} = b(1-f)N - \beta SI - cS$$

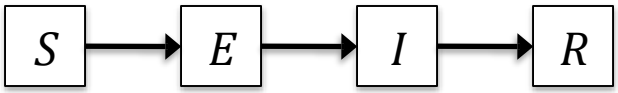
$$\frac{dI}{dt} = \beta SI - \alpha I - (c+\sigma)I$$

$$\frac{dR}{dt} = \alpha I - cR + bfN$$

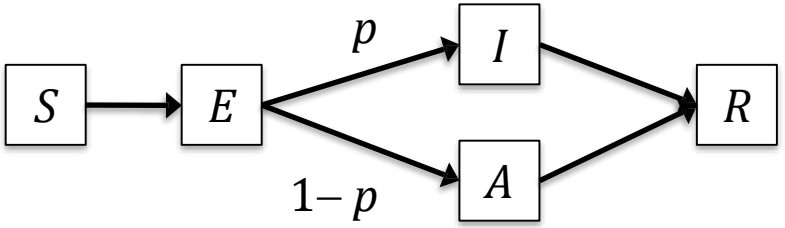
$$N = S + I + R$$

# More detailed natural history

- Latent period – the period of time during which the infected individual is not yet infectious (e.g. the latent period is 14 days on average for smallpox)
  - 4 compartments: “Susceptible”, “Exposed”, “Infectious”, and “Recovered”

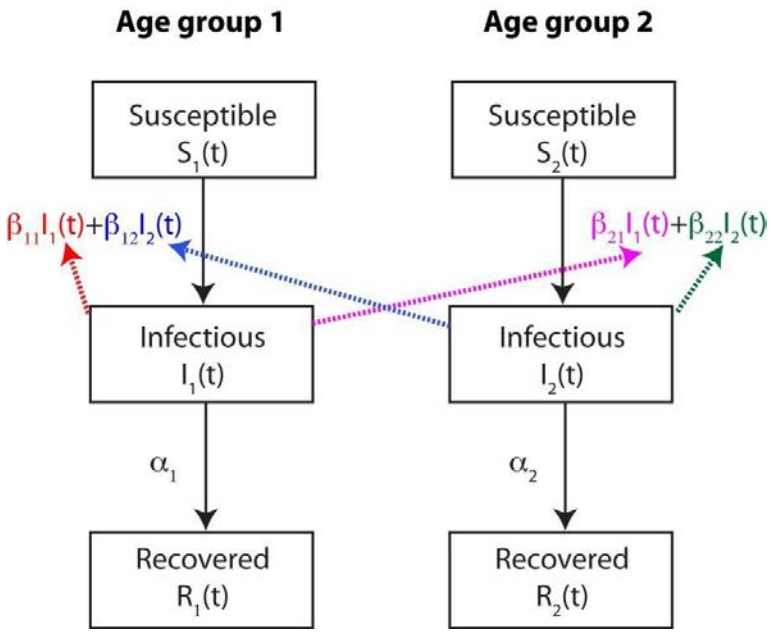



- More than one possible pathway, e.g. an infected individual develops symptoms only with probability *p*



# Age structure

- Age-specific transmissibility and natural history.
- $\beta_{ij}$: transmission rate from age group *j* to age group *i*. If there are *n* age groups, then there are $n^2$ $\beta_{ij}$'s.




$$S_1(t+\Delta t) = S_1(t) - \underbrace{\beta_{11}S_1(t)I_1(t)\Delta t}_{\text{number of infections in age group 1 generated by age group 1 between time } t \text{ and } t+\Delta t} - \underbrace{\beta_{12}S_1(t)I_2(t)\Delta t}_{\text{number of infections in age group 1 generated by age group 2 between time } t \text{ and } t+\Delta t}$$

$$I_1(t+\Delta t) = I_1(t) + \underbrace{\beta_{11}S_1(t)I_1(t)\Delta t}_{\text{number of infections in age group 1 generated by age group 1 between time } t \text{ and } t+\Delta t} + \underbrace{\beta_{12}S_1(t)I_2(t)\Delta t}_{\text{number of infections in age group 1 generated by age group 2 between time } t \text{ and } t+\Delta t} - \underbrace{\alpha_1 I_1(t)\Delta t}_{\text{number of recovery in age group 1 between time } t \text{ and } t+\Delta t}$$

$$S_2(t+\Delta t) = S_2(t) - \underbrace{\beta_{21}S_2(t)I_1(t)\Delta t}_{\text{number of infections in age group 2 generated by age group 1 between time } t \text{ and } t+\Delta t} - \underbrace{\beta_{22}S_2(t)I_2(t)\Delta t}_{\text{number of infections in age group 2 generated by age group 2 between time } t \text{ and } t+\Delta t}$$

$$I_2(t+\Delta t) = I_2(t) + \underbrace{\beta_{21}S_2(t)I_1(t)\Delta t}_{\text{number of infections in age group 2 generated by age group 1 between time } t \text{ and } t+\Delta t} + \underbrace{\beta_{22}S_2(t)I_2(t)\Delta t}_{\text{number of infections in age group 2 generated by age group 2 between time } t \text{ and } t+\Delta t} - \underbrace{\alpha_2 I_2(t)\Delta t}_{\text{number of recovery in age group 2 between time } t \text{ and } t+\Delta t}$$

$$\text{age group 1: } \frac{dS_1}{dt} = -\underbrace{\left(\beta_{11}I_1(t) + \beta_{12}I_2(t)\right)}_{\lambda_1(t)}S_1(t), \qquad \frac{dI_1}{dt} = \underbrace{\left(\beta_{11}I_1(t) + \beta_{12}I_2(t)\right)}_{\lambda_1(t)}S_1(t) - \alpha_1 I_1(t)$$

$$\text{age group 2: } \frac{dS_2}{dt} = -\underbrace{\left(\beta_{21}I_1(t) + \beta_{22}I_2(t)\right)}_{\lambda_1(t)}S_2(t), \qquad \frac{dI_2}{dt} = \underbrace{\left(\beta_{21}I_1(t) + \beta_{22}I_2(t)\right)}_{\lambda_2(t)}S_2(t) - \alpha_2 I_2(t)$$

# Multiple populations

- While the assumption of homogenous mixing may be reasonable for within-population epidemic dynamics, it is obviously not valid when considering epidemic dynamics among different populations that are weakly interacting (e.g. Hong Kong and Beijing)
- Common models to simulate epidemic dynamics among populations (linked by human travel): meta-population models, gravity model


Population 1: $\frac{dS_1}{dt} = -\underbrace{(\beta_{11}I_1(t) + \beta_{12}I_2(t))}_{\lambda_1(t)}S_1(t) - \underbrace{k_1S_1(t)}_{\text{traveling from pop 1 to pop 2}} + \underbrace{k_2S_2(t)}_{\text{traveling from pop 2 to pop 1}}$, $\frac{dI_1}{dt} = \underbrace{(\beta_{11}I_1(t) + \beta_{12}I_2(t))}_{\lambda_1(t)}S_1(t) - \alpha_1 I_1(t) - \underbrace{k_1I_1(t)}_{\text{traveling from pop 1 to pop 2}} + \underbrace{k_2I_2(t)}_{\text{traveling from pop 2 to pop 1}}$

Population 2: $\frac{dS_2}{dt} = -\underbrace{(\beta_{21}I_1(t) + \beta_{22}I_2(t))}_{\lambda_1(t)}S_2(t) + \underbrace{k_1S_1(t)}_{\text{traveling from pop 1 to pop 2}} - \underbrace{k_2S_2(t)}_{\text{traveling from pop 2 to pop 1}}$, $\frac{dI_2}{dt} = \underbrace{(\beta_{21}I_1(t) + \beta_{22}I_2(t))}_{\lambda_2(t)}S_2(t) - \alpha_2 I_2(t) + \underbrace{k_1I_1(t)}_{\text{traveling from pop 1 to pop 2}} - \underbrace{k_2I_2(t)}_{\text{traveling from pop 2 to pop 1}}$