Introduction to epidemiology and mathematical epidemiology

ICMS – Course 01

Julien Arino
FIGS-transitions/population-models-Gemini_Generated_Image_r55bcer55bcer55b.jpeg
University of Manitoba
julien.arino@umanitoba.ca

Fall 2024

December 2024

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis.
We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

Outline

FIGS-transitions/population-models-Gemini_Generated_Image_r55bccr55bccr55b.jpeg

Epidemiology

Mathematical Epidemiology

Compartmental models

FIGS-transitions/population-models-Gemini_Generated_Image_r55bccr55bccr55b.jpeg

The Kermack-McKendrick SIR model without demography

The endemic SIRS model with demography

Last remarks

Definition

Wiki **Epidemiology** is the study and analysis of the distribution (who, when, and where), patterns and determinants of health and disease conditions in defined populations

BMJ **Epidemiology** is the study of how often diseases occur in different groups of people and why. Epidemiological information is used to plan and evaluate strategies to prevent illness and as a guide to the management of patients in whom disease has already developed

Etymology: *the study of what is upon the people*, derived from the Greek *epi* (*upon*, *among*), *demos* (*people*, *district*) and *logos* (*study*, *word*, *discourse*)

What does ChatGPT say?

Epidemiology is the scientific discipline that focuses on the study of the distribution, determinants, and patterns of health-related events, conditions, and diseases within populations. It involves the systematic collection, analysis, and interpretation of data to understand how diseases and health-related issues spread and impact different groups of people. Epidemiologists use various research methods and statistical techniques to investigate factors such as the causes of diseases, risk factors, disease transmission, and the effectiveness of interventions and preventive measures. The findings from epidemiological studies are essential for public health decision-making, healthcare planning, and the development of strategies to control and prevent the spread of diseases.

Who, when and where

Recall part of the definition on Wikipedia

> **Epidemiology** *is the study and analysis of the distribution (who, when, and where)*

A terminologically heavy domain

A few pointers for modellers:

- Moghadas and Laskowski. Review of terms used in modelling influenza infection. NCCID 2014

- Milwid et al. Toward standardizing a lexicon of infectious disease modeling terms. Frontiers in Public Health 2016

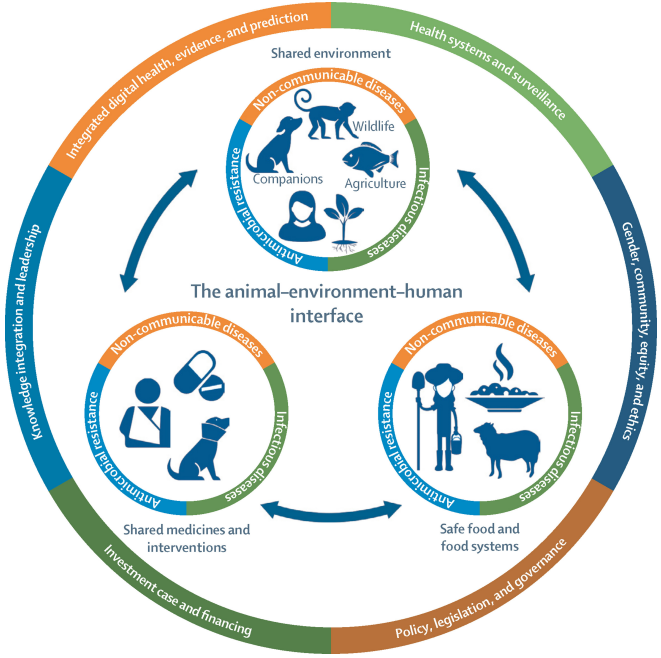
Who

- **Epidemiology** typically used when dealing with humans, but sometimes also generically when an easy description is sought; e.g., plant disease epidemiology

- **Epizootic**: denoting or relating to a disease that is temporarily prevalent and widespread in an animal population

- **Panzootic** is like a pandemic for animals

- **One Health**: considers health of humans, animals and their environment (including plants)


Integrated digital health, evidence, and prediction
Health systems and surveillance
Gender, community, equity, and ethics
Policy, legislation, and governance
Investment case and financing
Knowledge integration and leadership
Shared environment
Non-communicable diseases
Wildlife
Companions
Agriculture
Infectious diseases
Antimicrobial resistance
The animal–environment–human interface
Shared medicines and interventions
Safe food and food systems

Incidence & Prevalence (when?)

Incidence: number of new cases in a population generated within a certain time period

Prevalence: number of cases of a disease at a single time point in a population

$\Longrightarrow$ $I(t)$ in an epidemiological model is **prevalence**, not **incidence**

Exposition versus Exposed

- Some bright bulb (not sure who) in days of yore: let's call **exposed** someone who has contracted the disease but is not yet showing symptoms ($\Longrightarrow$ SEIR model)

- "Real" epidemiologist: let's trace people who were exposed to the virus, i.e., people having come into contact with the virus (whether they have contracted the disease or not)

- Interestingly, I have embarked on a quixotic quest to make people use L instead of E, only to be told by real epidemiologists that they don't care :)

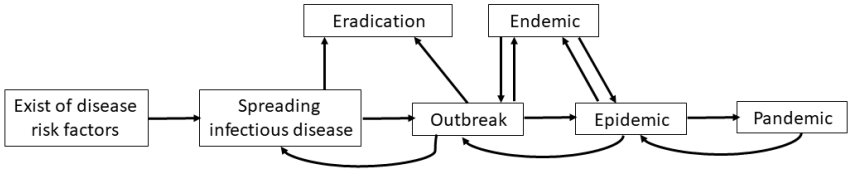
The different stages of propagation

Outbreak
sudden increase in occurrences of a disease in a particular time and place.

Endemic
constant maintained increase in occurrences of a disease in a geographic area.

Epidemic
rapid spread of disease to a large number of people in a given population within a short period of time.

Pandemic
spread across a large region, for instance multiple continents, or worldwide.

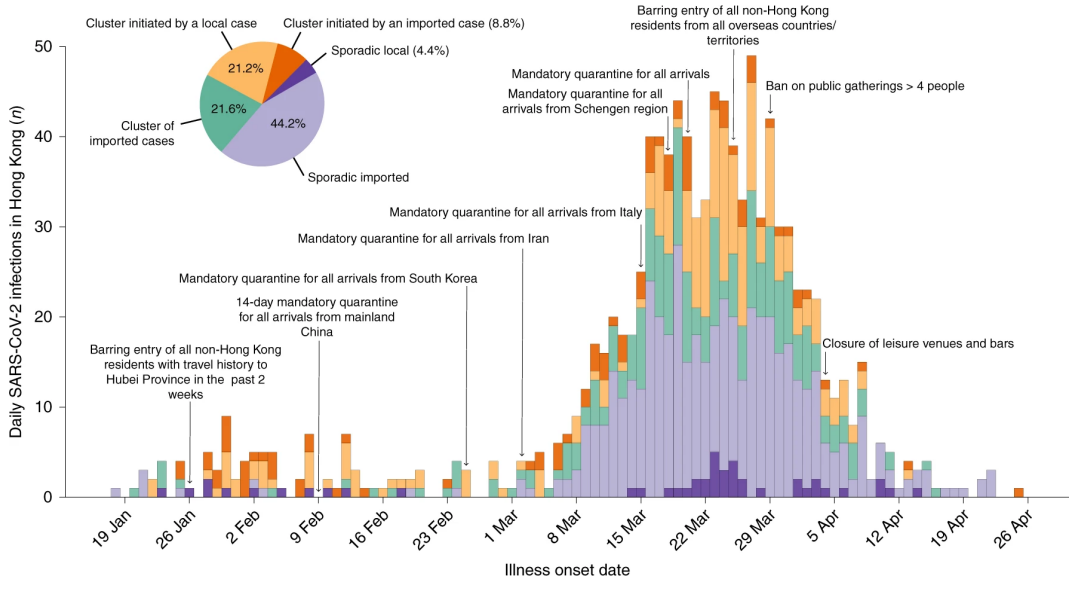


Epidemic curves

- Used to record the occurrence of new cases as a function of time

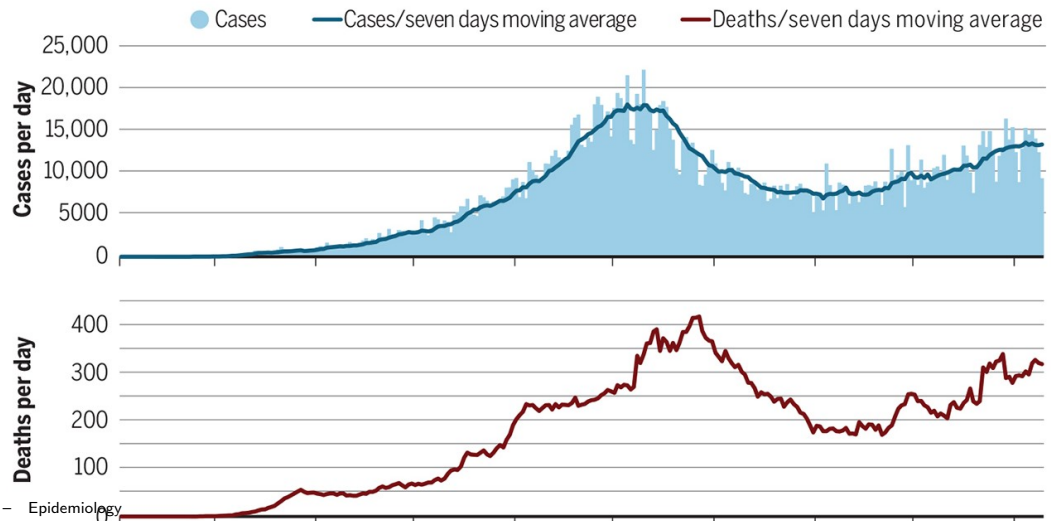
- When not too many cases, usually "individualised" (bar plots)

- When number of cases is large, continuous curve


Cluster initiated by a local case
Cluster initiated by an imported case (8.8%)
Sporadic local (4.4%)
21.2%
21.6%
44.2%
Cluster of imported cases
Sporadic imported
Barring entry of all non-Hong Kong residents from all overseas countries/ territories
Mandatory quarantine for all arrivals
Mandatory quarantine for all arrivals from Schengen region
Ban on public gatherings > 4 people
Mandatory quarantine for all arrivals from Italy
Mandatory quarantine for all arrivals from Iran
Mandatory quarantine for all arrivals from South Korea
14-day mandatory quarantine for all arrivals from mainland China
Barring entry of all non-Hong Kong residents with travel history to Hubei Province in the past 2 weeks
Closure of leisure venues and bars
Daily SARS-CoV-2 infections in Hong Kong (n)
0
10
20
30
40
50
19 Jan
26 Jan
2 Feb
9 Feb
16 Feb
23 Feb
1 Mar
8 Mar
15 Mar
22 Mar
29 Mar
5 Apr
12 Apr
19 Apr
26 Apr
Illness onset date

COVID-19 cases and deaths in Africa

The trend of daily reported cases of COVID-19 for the African continent, February to November 2020, shows the first peak of cases occurred July to August (mostly attributed to the Southern African Region) followed by a second peak, which started in October (mostly attributed to the Northern Region).



Some terminology for "where"

- **Epidemic**: diseases that are *visited upon* a population
- **Pandemic**: (will revisit this later in the course) epidemic that has spread across a large region, e.g., multiple continents or worldwide
- **Endemic**: diseases that *reside within* a population
- We don't say "panendemic"

Where? 1854 cholera outbreak

Cholera outbreak near Broad Street, London (UK)
Studied by John Snow

> *I found that nearly all the deaths had taken place within a short distance of the [Broad Street] pump*

WHO pandemic (influenza) phases

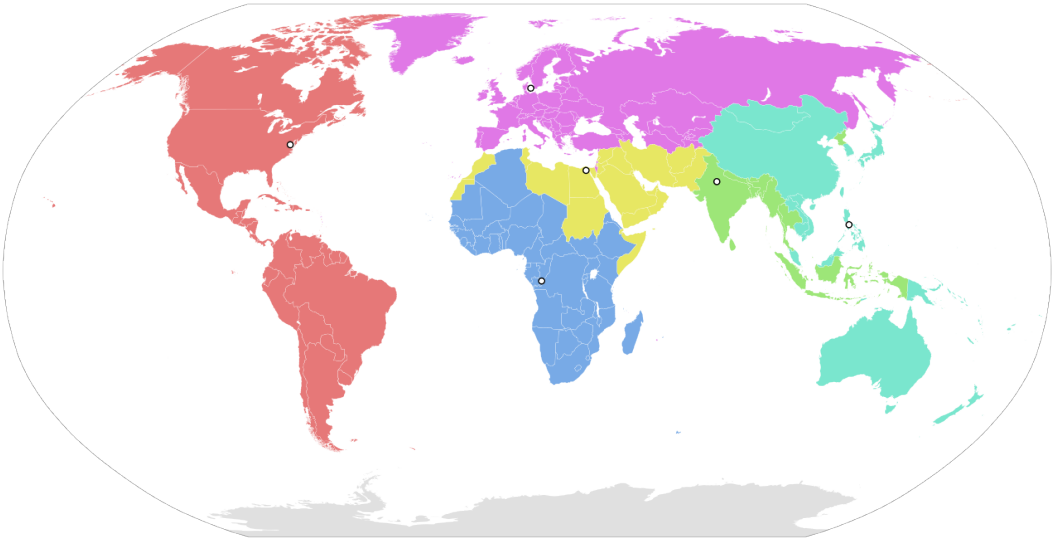
Period	Phase	Description
Interpandemic	1	No animal influenza virus circulating among animals has been reported to cause infection in humans
	2	Animal influenza virus circulating in domesticated or wild animals known to have caused infection in humans and therefore considered a specific potential pandemic threat

WHO pandemic (influenza) phases

Period	Phase	Description
Pandemic alert	3	Animal or human-animal influenza reassortant virus has caused sporadic cases or small clusters of disease in people, but has not resulted in H2H transmission sufficient to sustain community-level outbreaks
	4	Human-to-human transmission of an animal or human-animal influenza reassortant virus able to sustain community-level outbreaks has been verified

WHO pandemic (influenza) phases

Period	Phase	Description
Pandemic alert	5	Same identified virus has caused sustained community-level outbreaks in at least 2 countries in 1 WHO region
Pandemic	6	In addition to criteria in Phase 5, same virus has caused sustained community-level outbreaks in at least 1 other country in another WHO region



Fighting against infections

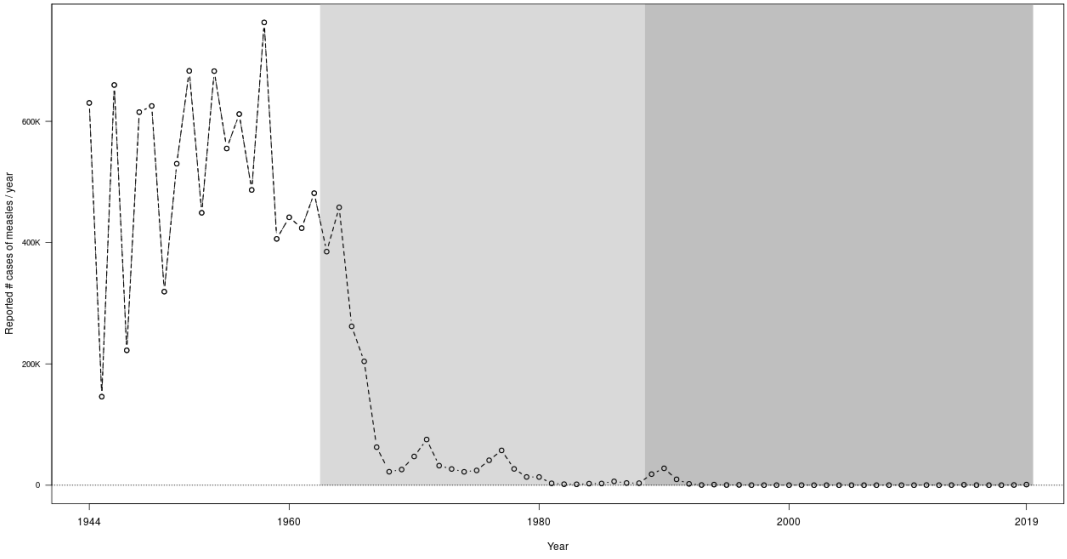
Epidemiological information is used to plan and evaluate **strategies to prevent illness** *and as a guide to the* **management of patients** *in whom disease has already developed*

- Preventing illness
  - Prophylactic measures
  - Vaccination
- Managing illness
  - Prevention of further spread (e.g., in hospital)
  - Treatment

Immunisation

- Smallpox first disease for which it was known
- Mentioned in a 1549 Chinese book
- China: powdered smallpox scabs blown up noses of the healthy; variolation-induced mortality not negligible (0.5-2%) but lower than normal (20%)
- 1798: Edward Jenner introduces safer inoculation with cowpox (vaccination)
- 1880s: Pasteur extends vaccination to chicken cholera and anthrax in animals and human rabies

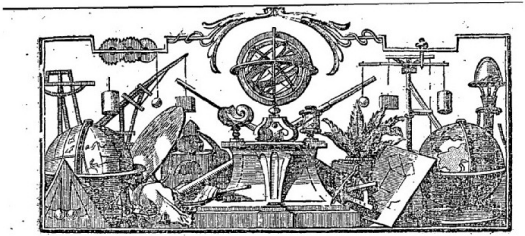
At the time, **herd immunity** was not understood so this was for personal protection


Reported # cases of measles / year
600K
400K
200K
0
1944
1960
1980
2000
2019
Year

The domain is quite old ..

.. but has only become a thing in recent years!

Daniel Bernoulli (1760)



MÉMOIRES
DE
MATHÉMATIQUE
ET
DE PHYSIQUE,
TIRÉS DES REGISTRES
de l'Académie Royale des Sciences;
De l'Année M. DCCLX.

ESSAI D'UNE NOUVELLE ANALYSE
De la mortalité causée par la petite Vérole, & des
avantages de l'Inoculation pour la prévenir.
Par M. DANIEL BERNOULLI.

- BNF scan or pdf
- Probably the first epidemic model
- About petite vérole (smallpox) inoculation

Ross (early 1900)

- On 20 August 1897, observed malaria parasites in the gut of a mosquito fed several days earlier on a malaria positive human
- Nobel Prize for Medicine 1902
- Started considering malaria eradication using mathematical models; for some history, read this 2012 paper

Kermack and McKendrick (1927+)

- We spend a lot more time on this later
- Groundbreaking set of papers starting in 1927
- We will see one particular case, the most well known, but this is just the tip of the iceberg of their work

Macdonald, Dietz and malaria

- Read for instance this paper, which presents a history of the development of the so-called Ross-Macdonald model

- Klaus Dietz also worked a lot on malaria

Some activity later, but not much until 1990s

- In recent years, explosion
- Since the beginning of COVID-19: just nuts..

Some landmarks in mathematical epidemiology (IMBO)

- Macdonald. The epidemiology and control of malaria. 1957
- Baroyan, Rvachev et al. Deterministic epidemic models for a territory with a transport network. Kibernetika, 1967
- Hethcote & Yorke. Gonorrhea Transmission Dynamics and Control. LNBM 56, 1984
- Anderson & May. Infectious diseases of humans: dynamics and control. 1991
- Capasso. Mathematical Structures of Epidemic Systems. LNBM 97, 1993
- Hethcote. The mathematics of infectious diseases. SIAM Review, 2000
- van den Driessche & Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. MBS, 2002

A more recent trend

- Some rare numerical work $\leq$ 1980s, mostly simulation of math models
  - Baroyan, Rvachev et al. Computer modelling of influenza epidemics for the whole country (USSR). *Advances in Applied Probability* (1971)
  - Rvachev & Longini. A mathematical model for the global spread of influenza. *Mathematical Biosciences* (1986)
  - Flahault, Letrait et al. Modelling the 1985 influenza epidemic in France. *Statistics in Medicine* (1988)

- More and more frequent now, to the point that some modelling studies are purely simulation-based

Agent-based models (ABM)

- Early in the life of these models, they were called IBM (individual-based models)

- Over the years, a "philosophical" distinction has emerged:
  - IBM are mathematical models that consider individuals as the units; e.g., DTMC, CTMC, branching processes, etc.
  - ABM are computational models whose study is, for the most part, only possible numerically

Network models

- Network models endow vertices with simple systems and couple them through graphs

- Can be ABM, but some networks can also be studied analytically

Has happened all along, undergoing a transformation

- Epidemiology has long relied on data
- Many developments in statistics originate there
- Data has traditionally been better for chronic diseases than for infectious ones
- Near-real-time surveillance of infectious diseases ongoing since the 1980s (e.g., Réseau Sentinelles)
- SARS-CoV-1 saw the beginning of a move towards real-time emerging infectious disease data
- With SARS-CoV-2, the system has really progressed a lot, both in terms of "citizen science" and governmental initiatives

Compartmental models

- Have become synonymous with epidemiological models

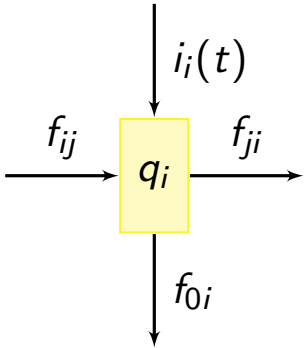
- Many epidemiological models are compartmental models, but the development of compartmental models in the 1970-1980s was not at all specific to epidemiology

- See in particular the works of John Jacquez, Carl Simon, GG Walter

- Unjustly fell into disuse: there are some very nice results in the area

Compartment (Jacquez 1979)

A **compartment** *is an amount of some material which acts kinetically like a distinct, homogeneous, well-mixed amount of material. A* **compartmental system** *consists of one or more compartments which interact by exchanging the material. There may be inputs into one or more compartments from outside the system and there may be excretions from the compartments of the system.*



- q_i size of the compartment, i.e., quantity of kinetically homogeneous material present in i; $q_i \geq 0$
- f_{ij} and f_{ji} transfer coefficients/functions
- f_{0i} excretion coefficient/function
- $i_i(t)$ entries from outside the system

Above is a **flow diagram**, which summarises the different flows acting on the compartment

Paper series worth reading

Model here is a particular case in

- Kermack & McKendrick. A contribution to the mathematical theory of epidemics (1927)

That paper was followed by a series of "Contributions to the mathematical theory of epidemics."

- II. The problem of endemicity (1932)
- III. Further studies of the problem of endemicity (1933)
- IV. Analysis of experimental epidemics of the virus disease mouse ectromelia (1937)
- V. Analysis of experimental epidemics of mouse-typhoid; a bacterial disease conferring incomplete immunity (1939)

What is the *size* of an epidemic?

- If we are interested in the possibility that an epidemic occurs
  - Does an epidemic peak always take place?
  - If it does take place, what is its size?

- If an epidemic traverses a population, is everyone affected/infected?

The Kermack-McKendrick SIR model without demography

- The period of time under consideration is sufficiently short that demography can be neglected (we also say the model has *no vital dynamics*)

- Individuals are either *susceptible* to the disease or *infected* by (and *infectious* with) the disease

- After recovering or dying from the disease, individuals are *removed* from the infectious compartment (R)

- Incidence is of **mass action** type and takes the form βSI

The state variables

We formulate the model as a system of **differential equations**

Differential equations: unknowns are *functions* (instead of scalars, like in algebraic equations)

At time $t \geq 0$ (we typically assume time starts at $t = 0$, but could also consider $t \geq t_0 > 0$), the **state variables**, in the current model, are the numbers of individuals who are

- susceptible to the disease: $S(t)$
- infected and infectious with the disease: $I(t)$
- removed from the infectious comparment: $R(t)$

Often, we drop the dependence on t if it is not explicitly required and write S, I, R

Important – Incidence functions

Incidence is the rate at which new cases arise, the incidence function then describes how contacts lead to new infections

If there are S susceptible individuals and I infectious individuals in the population, we use a function of the form

$$f(S, I)$$

The function can also explicitly depend on the total population N, i.e., $f(S, I, N)$

We return to incidence functions in Lecture 06

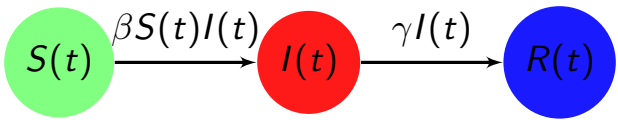
For now, just know the most common incidence functions are

- **mass action incidence** $f(S, I, N) = \beta SI$
- **standard** (or **proportional**) **incidence** $f(S, I, N) = \beta SI/N$

The Kermack-McKendrick model

This model is typically called the **Kermack-McKendrick** (KMK) **SIR model**

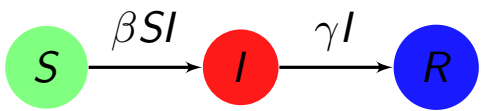
$$\frac{d}{dt}S(t) = -\beta S(t)I(t)$$
$$\frac{d}{dt}I(t) = \beta S(t)I(t) - \gamma I(t)$$
$$\frac{d}{dt}R(t) = \gamma I(t)$$



The Kermack-McKendrick model

As indicated, we often drop dependence on t of the state variables; we also write $X' := dX(t)/dt$. So the KMK model is usually written

$$S' = -\beta SI \tag{1a}$$
$$I' = \beta SI - \gamma I \tag{1b}$$
$$R' = \gamma I \tag{1c}$$



Reduction of the model

3 compartments, but when considered in detail, we notice that *removed* do not have a direct influence on the dynamics of S or I, in the sense that R does not appear in (1a) or (1b)

Furthermore, the total population (including deceased who are also in R)
$N = S + I + R$ satisfies

$$N' = (S + I + R)' = 0$$

Thus, N is constant and

$$S(t) + I(t) + R(t) = N_0, \quad t \geq 0. \tag{2}$$

so the dynamics of R can be deduced from $R = N - (S + I)$. So we can consider

$$S' = -\beta SI \tag{3a}$$
$$I' = \beta SI - \gamma I \tag{3b}$$

Equilibria

Let us consider the equilibria of

$$S' = -\beta SI \tag{3a}$$
$$I' = (\beta S - \gamma)I \tag{3b}$$

From (3b)

- either $S^\star = \gamma/\beta$
- or $I^\star = 0$

Substitute into (3a)

- in the first case, $(S^\star, I^\star) = (\gamma/\beta, 0)$
- in the second case, any $S^\star \geq 0$ is an EP

The second case is an *issue*: the usual linearisation does not work when there is a *continuum* of equilibria as the EP are not *isolated*

What is the problem with non-isolated EP?

Proposition 1

The Kermack-McKendrick model SIR model (1) *has the continuum of equilibria*

$$E_0^{KMK} := \{(S^\star, I^\star, R^\star) = (S_\infty, 0, N_0 - S_\infty), \quad S_\infty \in [0, N_0]\} \tag{5}$$

Proof

Let us consider (1) and start with $I = I^\star = 0$. Substitute this value into (1a) at equilibrium, giving $0 = -\gamma S^\star I^\star (= 0)$, meaning that any value of $S^\star$ satisfies this relation. From the conservation of the total population (2), the equilibrium E_0^{KMK} takes the form given by (5)

Now consider $S = S^\star = \gamma/\beta$. Substituting this value into (1a) at equilibrium gives $0 = -\gamma I^\star$, from which it follows that $I^\star = 0$, and, using the conservation of total population (2),

$$(S^\star, I^\star, R^\star) = \left(\frac{\gamma}{\beta}, 0, N_0 - \frac{\gamma}{\beta}\right) \tag{6}$$

is an equilibrium of (1). The equilibrium (6) is biologically relevant only when $N_0 - \gamma/\beta \geq 0$. Note that (5) includes (6) when the latter is biologically relevant

Adapting slightly the definitions in [Hirsch and Smale, 1974], consider the ordinary differential equation

$$x' = f(x) \tag{7}$$

where $x(t) \in W$ and $f : W \to E$ is a function such that solutions to (7) exist uniquely, e.g., a C^1 function, from an open set W of the vector space E into E

Denote $x(t, x_0)$ the solution to (7) through the initial value $x(t_0) = x_0$

A point $x^\star \in W$ is an **equilibrium** if $f(x^\star) = 0$

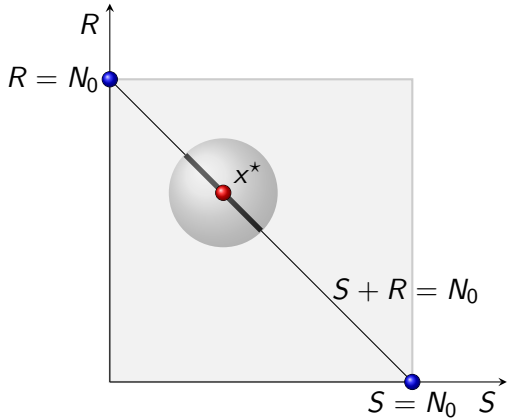
Definition 2 (Locally stable equilibrium)

An equilibrium point $x^\star$ of (7) is **locally stable** (LS) if for every neighbourhood $\mathcal{N}(x^\star)$ of $x^\star$ in W, there is a neighbourhood $\mathcal{N}_1 \subseteq \mathcal{N}(x^\star)$ of $x^\star$ such that every solution $x(t, x_0)$ with $x_0 \in \mathcal{N}_1$ is defined and in $\mathcal{N}(x^\star)$ for all $t > t_0$

Definition 3 (Locally asymptotically stable equilibrium)

If $\mathcal{N}_1$ can be chosen so that in addition to the properties in Definition 2, $\lim_{t\to\infty} x(t, x_0) = x^\star$ for all $x_0 \in \mathcal{N}_1$, then $x^\star$ is **locally asymptotically stable** (LAS)

DFE (5) of (1) are not **isolated**: any (open) neighbourhood of an equilibrium contains infinitely many other equilibria



Neighbourhood $\mathcal{N}(x^\star)$ of $x^\star \in E_0^{\mathrm{KMK}}$ lying on the $S-R$ plane (the neighbourhood extends above and below the $S-R$ plane in the I direction, not shown here). The thin line is E_0^{KMK}, the thick line is $E_0^{\mathrm{KMK}} \cap \mathcal{N}(x^\star)$

Proposition 4

Consider a disease-free equilibrium $x^\star \in E_0^{KMK}$ of (1). Then $x^\star$ is LS but not LAS

This means in particular that considering the Jacobian of (1) at the DFE **makes no sense**!

Proof

Let $x_1^\star \in E_0^{\text{KMK}}$ be an equilibrium of (1). Consider $\mathcal{S}_\mathcal{N}(x_1^\star) \subset E_0^{\text{KMK}}$, open subset of E_0^{KMK} containing $x_1^\star$. Now take some $x_2^\star \in \mathcal{S}_\mathcal{N}(x_1^\star)$. Since $x_2^\star \in \mathcal{S}_\mathcal{N}(x_1^\star) \subset E_0^{\text{KMK}}$, $x_2^\star$ is an equilibrium of (1) and thus $x(t, x_2^\star) = x_2^\star \in \mathcal{S}_\mathcal{N}(x_1^\star)$ for all $t \geq t_0$. As a consequence, $x_1^\star$ is locally stable

$\Rightarrow$ any open neighbourhood $\mathcal{N}(x_1^\star)$ contains $\mathcal{S}_\mathcal{N} = \mathcal{N}(x_1^\star) \cap E_0^{\text{KMK}}$

Consider, then, some $x_2^\star \in \mathcal{S}_\mathcal{N}$. Since $x_2^\star \in \mathcal{S}_\mathcal{N}$, $x_2^\star$ is an equilibrium and as a consequence, $\lim_{t\to\infty} x(t, x_2^\star) = x_2^\star$. Therefore, any open neighbourhood of $x_1^\star$ contains points x_0 not such that $\lim_{t\to\infty} x(t, x_0) = x_1^\star \implies x_1^\star$ is LS but not LAS

The next generation matrix method in this context

Consider the method in [van den Driessche and Watmough, 2002]

To construct $\mathcal{R}_0$, they require *local stability*

Theorem 2 in [van den Driessche and Watmough, 2002] pertaining to LAS, on the other hand, has one assumption (assumption A5) that the DFE be *locally asymptotically stable*, with the assumption that all eigenvalues of the linearisation near a disease-free equilibrium have negative real parts

Clearly, this cannot be true with (1)

Another approach – Study dI/dS

$$S' = -\beta SI \tag{3a}$$
$$I' = \beta SI - \gamma I \tag{3b}$$

What is the dynamics of dI/dS?

$$\frac{dI}{dS} = \frac{dI}{dt}\frac{dt}{dS} = \frac{I'}{S'} = \frac{\beta SI - \gamma I}{-\beta SI} = \frac{\gamma}{\beta S} - 1 \tag{8}$$

provided $S \neq 0$

Note – Recall that S and I are $S(t)$ and $I(t)$.. (8) thus describes the relation between S and I over solutions to the original ODE (3)

Integrate (8) and obtain trajectories in state space

$$I(S) = \frac{\gamma}{\beta} \ln S - S + C$$

with $C \in \mathbb{R}$

IC $I(S_0) = I_0 \Rightarrow C = S_0 + I_0 - \frac{\gamma}{\beta} \ln S_0$ and the solution to (1) is, as a function of S

$$I(S) = S_0 + I_0 - S + \frac{\gamma}{\beta} \ln \frac{S}{S_0}$$

$$R(S) = N - S - I(S) = R_0 - \frac{\gamma}{\beta} \ln \frac{S}{S_0}$$

(since $N_0 = S_0 + I_0 + R_0$)

Trajectories of (3) in (S, I)-space, normalised, with IC $(S_0, 1 - S_0)$ and $\beta/\gamma = 2.5$

Let us study

$$I(S) = S_0 + I_0 - S + \frac{\gamma}{\beta} \ln \frac{S}{S_0}$$

We have

$$\frac{d}{dS} I(S) = \frac{\gamma}{\beta S} - 1$$

So, in the previous curves, the max of $I(S)$ happens when $S = \gamma/\beta$ ($S = 0.4$ in the example)

At that point,

$$I(S) = I_0 + \left(1 - \frac{1}{\mathcal{R}_0} - \frac{\ln(\mathcal{R}_0)}{\mathcal{R}_0}\right) S_0$$

Theorem 5 (Epidemic or no epidemic?)

Let $(S(t), I(t))$ be a solution to (3) and $\mathcal{R}_0$ defined by

$$\mathcal{R}_0 = \frac{\beta}{\gamma} S_0 \tag{9}$$

- *If $\mathcal{R}_0 \leq 1$, then $I(t) \searrow 0$ when $t \to \infty$*
- *If $\mathcal{R}_0 > 1$, then $I(t)$ first reaches a maximum*

$$I_0 + \left(1 - \frac{1}{\mathcal{R}_0} - \frac{\ln(\mathcal{R}_0)}{\mathcal{R}_0}\right) S_0 \tag{10}$$

then goes to 0 as $t \to \infty$

```r
rhs_SIR_KMK <- function(t, x, p) {
  with(as.list(c(x, p)), {
    dS = - beta * S * I
    dI = beta * S * I - gamma * I
    dR = gamma * I
    return(list(c(dS, dI, dR)))
  })
}
# Initial condition for S (to compute R_0)
S0 = 1000
gamma = 1/14
# Set beta so that R_0 = 1.5
beta = 1.5 * gamma / S0
params = list(gamma = gamma, beta = beta)
IC = c(S = S0, I = 1, R = 0)
times = seq(0, 365, 1)
sol_KMK <- ode(IC, times, rhs_SIR_KMK, params)
```

```r
library(latex2exp)
plot(sol_KMK[, "time"], sol_KMK[, "I"],
     type = "l", lwd = 2,
     main = TeX("Kermack-McKendrick SIR, $R_0=1.5$"),
     xlab = "Time (days)", ylab = "Prevalence")
```

```
plot(sol_KMK[, "time"], sol_KMK[, "I"], type = "l",
main = TeX("Kermack-McKendrick␣SIR,␣$R_0=1.5$"),
xlab = "Time␣(days)", ylab = "Prevalence")
```

The basic reproduction number $\mathcal{R}_0$

- Indicator often used in epidemiology. Verbally
  average number of secondary cases of infection produced when a single infectious individual is introduced in a wholly susceptible population

- If $\mathcal{R}_0 < 1$, then each infectious individual infects on average less than 1 person and the epidemic is quite likely to go extinct

- If $R_0 > 1$, then each infectious individual infects on average more than 1 person and an epidemic is quite likely to occur

A few sample values of $\mathcal{R}_0$

$\mathcal{R}_0$ can be estimated from data

Infection	Location	Period	$\mathcal{R}_0$
Measles	Cirencester, England	1947-50	13-14
	England and Wales	1950-68	16-18
	Kansas, USA	1918-21	5-6
	Ontario, Canada	1912-3	11-12
	Willesden, England	1912-3	11-12
	Ghana	1960-8	14-15
	East Nigeria	1960-8	16-17

Final size of an epidemic

For a nonnegative valued integrable function $w(t)$, denote

$$w_0 = w(0), \qquad w_\infty = \lim_{t\to\infty} w(t), \qquad \hat{w} = \int_0^\infty w(t)\ dt$$

In the subsystem

$$S' = -\beta SI \tag{3a}$$
$$I' = \beta SI - \gamma I \tag{3b}$$

compute the sum of (3a) and (3b), making sure to show time dependence

$$\frac{d}{dt}(S(t) + I(t)) = -\gamma I(t)$$

Integrate from 0 to ∞:

$$\int_0^\infty \frac{d}{dt}(S(t) + I(t))\ dt = -\int_0^\infty \gamma I(t) dt$$

The left hand side gives

$$\int_0^\infty \frac{d}{dt}(S(t) + I(t))\ dt = S_\infty + I_\infty - S_0 - I_0 = S_\infty - S_0 - I_0$$

since $I_\infty = 0$

The right hand side takes the form

$$-\int_0^\infty \gamma I(t) dt = -\gamma \int_0^\infty I(t) dt = -\gamma \hat{I}$$

We thus have

$$S_\infty - S_0 - I_0 = -\gamma \hat{I} \tag{11}$$

Now consider (3a):

$$S' = -\beta SI$$

Divide both sides by S:

$$\frac{S'(t)}{S(t)} = -\beta I(t)$$

Integrate from 0 to ∞:

$$\ln S_\infty - \ln S_0 = -\beta \hat{I} \tag{12}$$

Express (11) and (12) in terms of $-\hat{I}$ and equate

$$\frac{\ln S_\infty - \ln S_0}{\beta} = \frac{S_\infty - S_0 - I_0}{\gamma}$$

Thus we have

$$(\ln S_0 - \ln S_\infty)S_0 = (S_0 - S_\infty)\mathcal{R}_0 + I_0\mathcal{R}_0 \tag{13}$$

Theorem 6 (Final size relation)

Let $(S(t), I(t))$ *be a solution to* (3) *and* $\mathcal{R}_0$ *defined by* (9)

The number $S(t)$ *of susceptible individuals is a nonincreasing function and its limit* S_∞ *is the only solution in* $(0, S_0)$ *of the transcendental equation*

$$(\ln S_0 - \ln S_\infty)S_0 = (S_0 - S_\infty)\mathcal{R}_0 + I_0\mathcal{R}_0 \tag{13}$$

The (transcendantal) final size equation

Rewrite the final size equation

$$(\ln S_0 - \ln S_\infty)S_0 = (S_0 - S_\infty)\mathcal{R}_0 + I_0\mathcal{R}_0 \tag{13}$$

as

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \tag{14}$$

Thus, we seek the zeros of the function $T(S_\infty)$

We seek S_∞ in $(0, S_0]$ s.t. $T(S_\infty) = 0$, with

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \tag{14}$$

Note to begin that

$$\lim_{S_\infty \to 0} T(S_\infty) = \lim_{S_\infty \to 0} -S_0 \ln(S_\infty) = \infty$$

Differentiating T with respect to S_∞, we get

$$T'(S_\infty) = \mathcal{R}_0 - S_0/S_\infty$$

When $S_\infty \to 0$, $\mathcal{R}_0 - S_0/S_\infty < 0$, so T decreases to $S_\infty = S_0/\mathcal{R}_0$

So if $\mathcal{R}_0 \leq 1$, the function T is decreasing on $(0, S_0)$, while it has a minimum if $\mathcal{R}_0 > 1$

Case $\mathcal{R}_0 \leq 1$

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \tag{14}$$

- We have seen that T decreases on $(0, S_0]$
- Also, $T(S_0) = -I_0\mathcal{R}_0 < 0$ ($I_0 = 0$ is trivial and not considered)
- T is continuous

$\Longrightarrow$ there exists a unique $S_\infty \in (0, S_0]$ s.t. $T(S_\infty) = 0$

Case $\mathcal{R}_0 > 1$

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \tag{14}$$

- We have seen that T decreases on $(0, S_0/\mathcal{R}_0]$
- For $S_\infty \in [S_0/\mathcal{R}_0]$, $T' > 0$
- As before, $T(S_\infty) = -I_0\mathcal{R}_0$
- T is continuous

$\implies$ there exists a unique $S_\infty \in (0, S_0]$ s.t. $T(S_\infty) = 0$. More precisely, in this case, $S_\infty \in (0, S_0/\mathcal{R}_0)$

We solve numerically. We need a function

```
final_size_eq = function(S_inf, S0 = 999, I0 = 1, R_0 = 2.5) {
  OUT = S0*(log(S0)-log(S_inf)) - (S0+I0-S_inf)*R_0
  return(OUT)
}
```

and solve easily using `uniroot`, here with the values by default that we have set for the function

```
uniroot(f = final_size_eq, interval = c(0.05, 999))
$root
[1] 106.8819
$f.root
[1] -2.649285e-07
$iter
[1] 10
$init.it
[1] NA
$estim.prec
[1] 6.103516e-05
```

To use something else than the default values, e.g.,

```
N0 = 1000
I0 = 1
S0 = N0-I0
R_0 = 2.4
uniroot(
  f = function(x)
    final_size_eq(S_inf = x,
                  S0 = S0, I0 = I0,
                  R_0 = R_0),
  interval = c(0.05, S0))
```

A function to use this..

```
final_size = function(L) {
  with(as.list(L), {
  S_inf = uniroot(f = function(x) final_size_eq(S_inf = x,
                                                S0 = S0, I0 = I0,
                                                R_0 = R_0),
                  interval = c(0.05, S0))
  return(S_inf$root)
  })
}
```

A figure with all the information

```
S = seq(0.1, S0, by = 0.1)
fs = final_size_eq(S, S0 = S0, I0 = I0, R_0 = R_0)
S_inf = uniroot(f = function(x) final_size_eq(S_inf = x,
                                              S0 = S0, I0 = I0,
                                              R_0 = R_0),
                interval = c(0.05, S0))
plot(S, fs, type = "l", ylab = "Value␣of␣equation␣(10)")
abline(h = 0)
points(x = S_inf$root, y = 0, pch = 19)
text(x = S_inf$root, y = 0, labels = "S_inf", adj = c(-0.25,-1))
```

```r
N0 = 1000
I0 = 1
S0 = N0-I0
R_0 = 0.8
S = seq(0.1, S0, by = 0.1)
fs = final_size_eq(S, S0 = S0, I0 = I0, R_0 = R_0)
S_inf = uniroot(f = function(x) final_size_eq(S_inf = x,
                                              S0 = S0, I0 = I0,
                                              R_0 = R_0),
                interval = c(0.05, S0))
plot(S, fs, type = "l", ylab = "Value of equation (10)")
abline(h = 0)
points(x = S_inf$root, y = 0, pch = 19)
text(x = S_inf$root, y = 0, labels = "S_inf", adj = c(-0.25,-1))
```

$\mathcal{R}_0 = 0.8$

```r
N0 = 1000
I0 = 1
S0 = N0-I0
R_0 = 2.5
S = seq(0.1, S0, by = 0.1)
fs = final_size_eq(S, S0 = S0, I0 = I0, R_0 = R_0)
S_inf = uniroot(f = function(x) final_size_eq(S_inf = x,
                                              S0 = S0, I0 = I0,
                                              R_0 = R_0),
                interval = c(0.05, S0))
plot(S, fs, type = "l", ylab = "Value of equation (10)")
abline(h = 0)
points(x = S_inf$root, y = 0, pch = 19)
text(x = S_inf$root, y = 0, labels = "S_inf", adj = c(-0.25,-1))
```

$\mathcal{R}_0 = 2.4$

```r
library(lattice)
library(viridis)
library(latex2exp)
values = expand.grid(
  R_0 = seq(0.01, 3, by = 0.01),
  I0 = seq(1, 100, 1)
)
values$S0 = N0-values$I0
L = split(values, 1:nrow(values))
values$S_inf = sapply(X = L, FUN = final_size)
values$final_size = values$S0-values$S_inf+values$I0
values$attack_rate = (values$final_size / N0)*100

p = levelplot(attack_rate ~ R_0*I0, data = values,
              xlab = TeX("$R_0$"), ylab = "I(0)",
              col.regions = viridis(100))
print(p)
```

A little nicer

```
values = expand.grid(
  R_0 = seq(0.01, 3, by = 0.01),
  I0 = seq(1, 100, 1)
)
values$S0 = N0-values$I0
L = split(values, 1:nrow(values))
values$S_inf = sapply(X = L, FUN = final_size)
values$final_size = values$S0-values$S_inf+values$I0
values$attack_rate = (values$final_size / N0)*100

p = levelplot(attack_rate ~ R_0*I0, data = values,
              xlab = TeX("$R_0$"), ylab = "I(0)",
              col.regions = viridis(100))
print(p)
```

(requires `lattice`, `viridis` and `latex2exp` librairies)

Attack rate (in %)

The simplest vaccination model

To implement vaccination in KMK, assume that vaccination reduces the number of susceptibles

Let total population be N with S_0 initially susceptible

Vaccinate a fraction $p \in [0, 1]$ of susceptible individuals

Original IC (for simplicity, $R(0) = 0$)

$$IC : (S(0), I(0), R(0)) = (S_0, I_0, 0) \tag{15}$$

Post-vaccination IC

$$IC : (S(0), I(0), R(0)) = ((1 - p)S_0, I_0, pS_0) \tag{16}$$

Vaccination reproduction number

Without vaccination

$$\mathcal{R}_0 = \frac{\beta}{\gamma} S_0 \tag{9}$$

With vaccination, denoting $\mathcal{R}_v$ the reproduction number,

$$\mathcal{R}_v = \frac{\beta}{\gamma}(1-p)S_0 \tag{17}$$

Since $p \in [0,1]$, $\mathcal{R}_v \leq \mathcal{R}_0$

Herd immunity

Therefore

- $\mathcal{R}_v < \mathcal{R}_0$ if $p > 0$
- To control the disease, $\mathcal{R}_v$ must take a value less than 1

To make $\mathcal{R}_v$ less than 1

$$\mathcal{R}_v < 1 \iff p > 1 - \frac{1}{\mathcal{R}_0} \tag{18}$$

By vaccinating a fraction $p > 1 - 1/\mathcal{R}_0$ of the susceptible population, we thus are in a situation where an epidemic peak is precluded (or, at the very least, the final size is reduced)

This is **herd immunity**

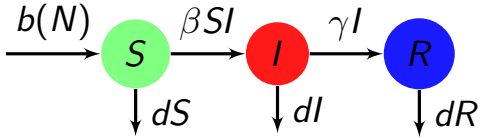
Two potential variations on the Kermack-McKendrick model

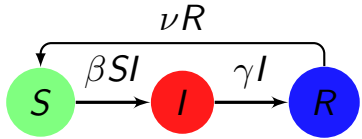
- Add *vital dynamics*, i.e., consider demographic processes

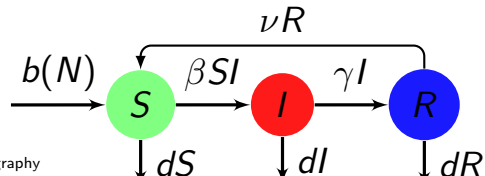
- Individuals do not die from the disease; after recovering, individuals are *immune* from infection for some time

- We can of course combine both!

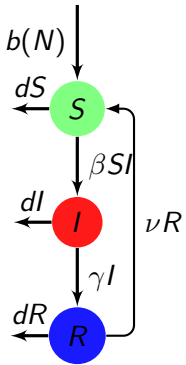
Potential variations







The model



$$S' = b(N) + \nu R - dS - \beta SI \tag{19a}$$
$$I' = \beta SI - (d+\gamma)I \tag{19b}$$
$$R' = \gamma I - (d+\nu)R \tag{19c}$$

Consider the initial value problem consisting in (19) to which we adjoin initial conditions $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ and $R(0) = R_0 \geq 0$

Typically, we assume $N_0 = S_0 + I_0 + R_0 > 0$ to avoid a trivial case

Birth and death are *relative*

Remark that the notions of *birth* and *death* are relative to the population under consideration

E.g., consider a model for human immunodeficiency virus (HIV) in an at-risk population of intravenous drug users. Then

- birth is the moment the at-risk behaviour starts
- death is the moment the at-risk behaviour stops, whether from "real death" or because the individual stops using drugs

Choosing a form for demography

Before we proceed with the analysis proper, we must discuss the nature of the assumptions on demography

To do this, we consider the behaviour of the total population

$$N(t) = S(t) + I(t) + R(t)$$

Behaviour of the total population

Summing the equations in (19)

$$N' = b(N) - dN \tag{20}$$

There are three common ways to define $b(N)$ in (20)

1. $b(N) = b$
2. $b(N) = bN$
3. $b(N) = bN - cN^2$

Case 3 leads to logistic dynamics of the total population and is not discussed here

Case of a birth rate constant *per capita*

If $b(N) = bN$, then birth in (20) satisfies $N'/N = b$; we say that birth is **constant per capita**

In this case, (20) takes the form

$$N' = bN - dN = (b - d)N$$

with initial condition $N(0) = N_0$

The solution to this scalar autonomous ODE is easy

$$N(t) = N_0 e^{(b-d)t}, \quad t \geq 0$$

Thus there are 3 possibilities:

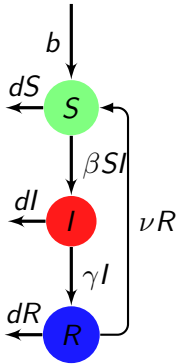
- if $b > d$, $N(t) \to \infty$, the total population explodes
- if $b = d$, $N(t) \equiv N_0$, the total population remains constant
- if $b < d$, $N(t) \to 0$, the total population collapses

From now on, assume $b(N) = b$

- We want a reasonable case, we could therefore suppose that $b(N) = d$, which would lead to a constant total population

- However, this is a little reductive, so we choose instead $b(N) = b$, which, we will see, works as well even though it can initially be thought of as not being very realistic

The model (for good this time)



$$S' = b + \nu R - dS - \beta SI \tag{21a}$$
$$I' = \beta SI - (d + \gamma)I \tag{21b}$$
$$R' = \gamma I - (d + \nu)R \tag{21c}$$

Consider the initial value problem consisting in (21) to which we adjoin initial conditions $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ and $R(0) = R_0 \geq 0$

Typically, we assume $N_0 = S_0 + I_0 + R_0 > 0$ to avoid a trivial case

Is the system well-posed?

For an ODE epidemiological model

- Do solutions to (21) exist and are they unique?
- Is the positive cone invariant under the flow of (21)?
- Are solutions to (21) bounded? Some models have unbounded solutions but they are rare and will need to be considered specifically

Solutions exist and are unique

- The vector field is always C^1, implying that solutions exist and are unique

If we had instead considered an incidence of the form $f(S, I, N) = \beta SI/N$ and, say, demography with $b(N) = bN$, then some discussion might have been needed if $b < d$

Invariance of $\mathbb{R}^3_+$ under the flow (1)

Let us start by assuming that $I(0) = I_0 = 0$. Then (21b) remains $I' = 0$, meaning that the SR-plane (i.e., the set $\{I = 0\}$) is positively invariant under the flow of (21)

On that plane, (21) reduce to

$$S' = b + \nu R - dS \tag{22a}$$
$$R' = -(d + \nu)R \tag{22b}$$

$\implies$ a solution with $I_0 > 0$ cannot enter the plane $\{I = 0\}$. Indeed, suppose that $I_0 > 0$ but $\exists t_\star > 0$ such that $I(t_\star) = 0$. Then at $(S(t_\star), I(t_\star) = 0, R(t_\star))$, there are two solutions to (21): the one we just generated as well as the one governed by (22)

This contradicts uniqueness of solutions to (21)

Invariance of $\mathbb{R}^3_+$ under the flow (2)

We saw that $I(t) > 0$ if $I(0) > 0$

Suppose now that $S = 0$. Equation (21a) is then

$$S' = b + \nu R > 0$$

So if $S(0) = S_0 > 0$, then $S(t) > 0$ for all t. If, on the other hand, $S_0 = 0$, then $S(t) > 0$ for $t > 0$ small; from what we just saw, this is then also true for all $t > 0$

We say the vector field points *inward*

$\implies$ S cannot become zero

Do the same for R

To summarise, for invariance

For simplicity, denote $\mathbb{R}^\star = \mathbb{R} \setminus \{0\}$

- If $(S(0), I(0), R(0)) \in \mathbb{R}_+ \times \mathbb{R}_+^\star \times \mathbb{R}_+$, then $\forall t > 0$,

$$(S(t), I(t), R(t)) \in (\mathbb{R}_+^\star)^3$$

- If $(S(0), I(0), R(0)) \in \mathbb{R}_+ \times \{0\} \times \mathbb{R}_+$, then $\forall t \geq 0$,

$$(S(t), I(t), R(t)) \in \mathbb{R}_+^\star \times \{0\} \times \mathbb{R}_+$$

The model is therefore satisfactory in that it does not allow solutions to become negative

Remark – Know your audience

This reasoning has its place in an MSc of PhD manuscript: you need to demonstrate that you know what to do and how to do it

In a research paper, this is not really necessary and actually often superfluous; the statement *it is easy to show that solutions exist uniquely and that the positive orthant is invariant under the flow of the system* is typically sufficient

(However, be sure to cover your bases: don't show the proof in the paper but have it in your notes.. *it is easy to show* can be a dangerous statement if it is not easy...)

The total population is asymptotically constant

Since $b(N) = b$, the total population equation (20) takes the form

$$N' = b - dN$$

This equation has a unique equilbrium $N^\star = b/d$ and it is very easy to check that this equilibrium is GAS: this is a scalar autonomous equation, so solutions are monotone; they increase to $N^\star$ if $N_0 < N^\star$ and decrease to $N^\star$ if $N_0 > N^\star$

So we can work at the limit $N^\star$ where $R = N^\star - (S + I)$ and thus drop the equation for R

Boundedness

It follows from what we just saw that the positive cone $\mathbb{R}^3_+$ is (positively) invariant under the flow of (21)

Since $N(t) \to N^\star$, we deduce that solutions of (21) are bounded

Seeking equilibria

We seek $S = S^\star, I = I^\star, R = R^\star$ such that

$$0 = b + \nu R - dS - \beta SI \tag{23a}$$
$$0 = \beta SI - (d + \gamma)I \tag{23b}$$
$$0 = \gamma I - (d + \nu)R \tag{23c}$$

From (23b), either $I^\star = 0$ or $\beta S - (d + \gamma) = 0$, i.e., $S^\star = (d + \gamma)/\beta$

When $I^\star = 0$, substituting $I^\star = 0$ into (23c) implies that $R^\star = 0$ and, in turn, substituting $I^\star = R^\star = 0$ into (23c) gives $S^\star = b/d$. This gives the disease-free equilibrium (DFE)

$$\boldsymbol{E}_0 := (S^\star, I^\star, R^\star) = \left(\frac{b}{d}, 0, 0\right) \tag{24}$$

We return to $S^\star = (d + \gamma)/\beta$ in a while

Classic method for computing $\mathcal{R}_0$

$\mathcal{R}_0$ is the surface in parameter space where the DFE loses its LAS

To find $\mathcal{R}_0$, we therefore study the LAS of the DFE

In an arbitrary (S, I, R), the Jacobian matrix of (21) takes the form

$$J_{(S,I,R)} = \begin{pmatrix} -d - \beta I & -\beta S & \nu \\ \beta I & \beta S - (d + \gamma) & 0 \\ 0 & \gamma & -(d + \nu) \end{pmatrix} \tag{25}$$

The LAS of the DFE depends on the sign of the real parts of the eigenvalues of (25) at that equilibrium point, so we evaluate

$$J_{\boldsymbol{E}_0} = \begin{pmatrix} -d & -\beta S^\star & \nu \\ 0 & \beta S^\star - (d+\gamma) & 0 \\ 0 & \gamma & -(d+\nu) \end{pmatrix} \tag{26}$$

Block upper triangular matrix $\implies$ eigenvalues are $-d < 0$, $-(d+\nu) < 0$ and $\beta S^\star - (d+\gamma)$

$\implies$ LAS of the DFE determined by sign of $\beta S^\star - (d+\gamma)$

Sign of $\beta S^\star - (d+\gamma)$

Recall that at the DFE (24), $S^\star = b/d$, so

$$\text{sign}(\beta S^\star - (d+\gamma)) = \text{sign}\left(\beta\frac{b}{d} - (d+\gamma)\right)$$

So the DFE is LAS if

$$\beta\frac{b}{d} < d+\gamma \iff \frac{\beta}{d+\gamma}\,\frac{b}{d} < 1$$

Denote

$$\mathcal{R}_0 = \frac{\beta}{d+\gamma}\,\frac{b}{d} \tag{27}$$

(We sometimes emphasise that $b/d = N^\star$, the total population, and thus write $\mathcal{R}_0 = \beta N^\star/(d+\gamma)$)

Seeking equilibria (2)

Now consider the second EP where $S^\star = (d+\gamma)/\beta = N^\star/\mathcal{R}_0$

Write (23c) as $R^\star = \gamma I^\star/(d+\nu)$

Since $S^\star + I^\star + R^\star = N^\star$, this means that

$$N^\star - S^\star - I^\star = \gamma I^\star/(d+\nu)$$

so substituting $S^\star = N^\star/\mathcal{R}_0$,

$$\left(1 + \frac{\gamma}{d+\nu}\right) I^\star = \left(1 - \frac{1}{\mathcal{R}_0}\right) N^\star$$

So finally

$$I^\star = \left(1 - \frac{1}{\mathcal{R}_0}\right) \frac{d+\nu}{d+\nu+\gamma} N^\star$$

The EEP

The **endemic equilibrium** (EEP) of (21) is

$$\boldsymbol{E}_\star := (S^\star, I^\star, R^\star) = \left(\frac{1}{\mathcal{R}_0} N^\star, \left(1 - \frac{1}{\mathcal{R}_0}\right) \frac{d+\nu}{d+\nu+\gamma} N^\star, N^\star - (S^\star + I^\star) \right) \quad (28)$$

Remark that $\boldsymbol{E}_\star$ is **not biologically relevant** when $\mathcal{R}_0 \leq 1$

Theorem 7

Let the basic reproduction number be

$$\mathcal{R}_0 = \frac{\beta}{d+\gamma}\, N^\star \tag{27}$$

and consider the EP of (21)*: the DFE*

$$\boldsymbol{E}_0 = \left(\frac{b}{d}, 0, 0\right) \tag{24}$$

and the EEP

$$\boldsymbol{E}_\star = \left(\frac{1}{\mathcal{R}_0}\, N^\star, \left(1 - \frac{1}{\mathcal{R}_0}\right)\, \frac{d+\nu}{d+\nu+\gamma}\, N^\star, N^\star - (S^\star + I^\star)\right) \tag{28}$$

- *If* $\mathcal{R}_0 < 1$*, then* $\boldsymbol{E}_0$ *is LAS and* $\boldsymbol{E}_\star$ *is not biologically relevant*
- *If* $\mathcal{R}_0 > 1$*, then* $\boldsymbol{E}_0$ *is unstable and* $\boldsymbol{E}_\star$ *is biologically relevant*

As you can probably guess, if $\mathcal{R}_0 > 1$, then $\boldsymbol{E}_\star$ is not only biologically relevant but actually also LAS

Recall the Jacobian

$$J_{(S,I,R)} = \begin{pmatrix} -d - \beta I & -\beta S & \nu \\ \beta I & \beta S - (d+\gamma) & 0 \\ 0 & \gamma & -(d+\nu) \end{pmatrix} \tag{25}$$

$$= \begin{pmatrix} -\beta I & -\beta S & \nu \\ \beta I & \beta S - \gamma & 0 \\ 0 & \gamma & -\nu \end{pmatrix} - d\mathbb{I}$$

From this, we get that $-d$ is an eigenvalue of J

- there is a theorem that tells us that if $\lambda \in \sigma(M)$, then $\lambda + k \in \sigma(M + k\mathbb{I})$ ($\sigma(M)$ is the spectrum of M, the set of eigenvalues of M)
- the first matrix on the second line has all column sums zero so has a zero eigenvalue

We could continue and after some blood, sweat and tears, get that $J_{\boldsymbol{E}_\star}$ has its eigenvalues with negative real parts when $\boldsymbol{E}_\star$ is biologically relevant, i.e., when $\mathcal{R}_0 > 1$

With even more blood, sweat and tears, we can actually show that the result is *global*

We express that on the next slide

Theorem 8

Let the basic reproduction number be defined by (27) *and consider the DFE* (24) *and the EEP* (28)

- *If* $\mathcal{R}_0 < 1$, *then* $\boldsymbol{E}_0$ *is globally asymptotically stable (GAS) and* $\boldsymbol{E}_\star$ *is not biologically relevant*
- *If* $\mathcal{R}_0 > 1$, *then* $\boldsymbol{E}_0$ *is unstable and* $\boldsymbol{E}_\star$ *is GAS*

In other words

- when $\mathcal{R}_0 < 1$, then all solutions go to the DFE, the disease goes **extinct**
- when $\mathcal{R}_0 > 1$, then all solutions go to the EEP, the disease becomes **endemic**

```r
library(deSolve)
rhs_SIRS <- function(t, x, p) {
  with(as.list(c(x, p)), {
    dS = b + nu * R - d * S - beta * S * I
    dI = beta * S * I - (d + gamma) * I
    dR = gamma * I - (d + nu) * R
    return(list(c(dS, dI, dR)))
  })
}
# Initial conditions
N0 = 1000
I0 = 1
R0 = 0
IC = c(S = N0-(I0+R0), I = I0, R = R0)
# "Known" parametres
d = 1/(80*365.25)
b = N0 * d
gamma = 1/14
nu = 1/365.25
```

```r
# Set beta s.t. R_0 = 1.5
R_0 = 1.5
beta = R_0 * (d + gamma) / (N0-I0-R0)
params = list(b = b, d = d, gamma = gamma, beta = beta, nu = nu)
times = seq(0, 365, 1)
# Call the numerical integrator
sol_SIRS <- ode(y = IC, times = times, func = rhs_SIRS,
                parms = params, method = "ode45")
# Plot the result
plot(sol_SIRS[,"time"], sol_SIRS[,"I"],
     type = "l", lwd = 2,
     xlab = "Time␣(days)", ylab = "Prevalence")
```

```r
library(deSolve)
rhs_SIRS <- function(t, x, p) {
  with(as.list(c(x, p)), {
    dS = b + nu * R - d * S - beta * S * I
    dI = beta * S * I - (d + gamma) * I
    dR = gamma * I - (d + nu) * R
    return(list(c(dS, dI, dR)))
  })
}
# Initial conditions
N0 = 1000
I0 = 1
R0 = 0
IC = c(S = N0-(I0+R0), I = I0, R = R0)
# "Known" parametres
d = 1/(80*365.25)
b = N0 * d
```

```r
gamma = 1/14
nu = 1/365.25
# Set beta s.t. R_0 = 1.5
R_0 = 1.5
beta = R_0 * (d + gamma) / (N0-I0-R0)
params = list(b = b, d = d, gamma = gamma, beta = beta, nu = nu)
times = seq(0, 500, 1)
# Call the numerical integrator
sol_SIRS <- ode(y = IC, times = times, func = rhs_SIRS,
                parms = params, method = "ode45")
# Plot the result
plot(sol_SIRS[,"time"], sol_SIRS[,"I"],
     type = "l", lwd = 2,
     xlab = "Time (days)", ylab = "Prevalence")
crop_figure("FIGS/lecture-01-SIRS_one_sim_prevalence")
```

```
## Error in crop_figure("FIGS/lecture-01-SIRS_one_sim_prevalence"):
could not find function "crop_figure"
```

I just did ...

What I advise not to do: illustrate a mathematical result without adding anything to the result itself

Let us make things a bit better. See the code

```r
# Compute the EPs
valeur_PE = function(params) {
  with(as.list(c(params)), {
    OUT = list()
    if (R_0<1) {
      OUT$S_EP = Pop
      OUT$I_EP = 0
      OUT$col = "dodgerblue4"
    } else {
      OUT$S_EP = 1/R_0*Pop
      OUT$I_EP = (1-1/R_0)*(d+nu)/(d+nu+gamma)*Pop
      OUT$col = "darkorange4"
    }
    return(OUT)
  })
}
# RHS function set in previous chunk
```

```r
# Put the parameters in a list
# "Known" parametres
params = list()
params$Pop = N0
params$d = 1/(80 * 365.25)
params$b = params$Pop * params$d
params$gamma = 1/14
params$nu = 1/365.25
params$t_f = 1200
params$I_0 = I0
# Note that we did not set R_0 or beta. This is done in a loop

# IC. "Static" part (N0, I0, R0) of IC are set in previous chunk
IC = c(S = N0-(I0+R0), I = I0, R = R0)

# Times at which the solution will be returned.
tspan = seq(from = 0, to = params$t_f, by = 0.1)
```

```r
# Now simulate the ODE. Loop on several values of R_0
R_0 = c(0.8, 1.5, 2.5)
# Save results in a list together with EP values
sol_ODE = list()
EP = list()
# Now loop on R_0
for (r_0 in R_0) {
  # Name for list entry
  entry_name = sprintf("$R_0$=%1.1f",r_0)
  # Keep the current value of R_0 to compute EPs
  params$R_0 = r_0
  # R0=(beta/(d+gamma)) => beta=R0*(d+gamma)
  params$beta = r_0 * (params$d+params$gamma) / (N0-I0-R0)
  # Call numerical integrator
  sol_ODE[[entry_name]] = ode(y = IC,
                              func = rhs_SIRS,
                              times = tspan,
```

```
                                parms = params)
  EP[[entry_name]] = valeur_PE(params)
  EP[[entry_name]]$lty = which(r_0 == R_0)
}

# Get maximum value of I across all simulations for plot. Note the use of
max_I = max(unlist(lapply(sol_ODE, function(x) max(x[,"I"]))))

# Plot
y_axis = plot_hr_yaxis(sol_ODE[[1]][,"time"], sol_ODE[[1]][,"I"],
                       y_range = c(0, max_I),
                       type = "l", lwd = 5, col = EP[[1]]$col, lty = EP[[1]
                       xlab = "Time (days)", ylab = "Prevalence")

## Error in plot_hr_yaxis(sol_ODE[[1]][, "time"], sol_ODE[[1]][, "I"],
y_range = c(0, :  could not find function "plot_hr_yaxis"

points(x = params$t_f, y = EP[[1]]$I_EP*y_axis$factor,
```

```
       col = EP[[1]]$col, pch = 19, cex = 2)

## Error:  object 'y_axis' not found

for (i in 2:length(sol_ODE)) {
  lines(sol_ODE[[i]][,"time"], sol_ODE[[i]][,"I"]*y_axis$factor,
        type = "l", lwd = 5, col = EP[[i]]$col, lty = EP[[i]]$lty)
  points(x = params$t_f, y = EP[[i]]$I_EP*y_axis$factor,
         col = EP[[i]]$col, pch = 19, cex = 2)
}

## Error:  object 'y_axis' not found

legend("topright", legend = TeX(names(EP)), cex = 0.8,
       col = unlist(lapply(EP, function(x) x$col)),
       lty = unlist(lapply(EP, function(x) x$lty)),
       lwd = c(3,3,3))

## Error in (function (s, units = "user", cex = NULL, font = NULL,
vfont = NULL, :  plot.new has not been called yet
```

We could continue, but with a model this simple, there is little more to do: the 3 parameters of the system are combined within $\mathcal{R}_0$ and the latter summarises the dynamics well

We are going to show something important: the bifurcation diagram

We saw that when $\mathcal{R}_0 < 1$, $I \to 0$, whereas when $\mathcal{R}_0 > 1$, $I \to (1 - 1/\mathcal{R}_0)N$. Let us represent this (code)

```r
# Values of the EPs
value_EPs = function(R_0, N) {
  EP_I = ifelse(R_0 < 1, 0, (1-1/R_0)*N)
  return(EP_I)
}

R_0 = seq(0.5, 5, by = 0.01)
EP_I = value_EPs(R_0, N = 1000)
# We also show the DFE when R_0>1, so prepare this
R_0_geq_1 = R_0[which(R_0>=1)]
DFE = rep(0, length(R_0_geq_1))

plot(R_0, EP_I,
     type = "l", lwd = 3,
     xlab = TeX("$R_0$"),
     las = 1,
     ylab = "Prevalence at equilibrium")
```

```
lines(R_0_geq_1, DFE,
      type = "l", lwd = 3,
      lty = 2)
legend("topleft", legend = c("LAS EP", "Unstable EP"),
       lty = c(1, 2), lwd = c(2,2),
       bty = "n")
```

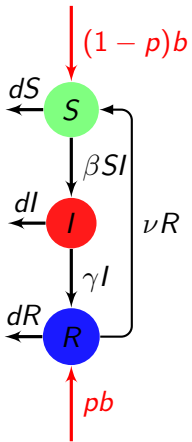
An SIRS model with vaccination

Take SIRS model (21) and assume the following

- Vaccination takes newborn individuals and moves them directly into the removed compartment, without them becoming infected/infectious

- A fraction p is vaccinated at birth

The model



$$S' = (1-p)b + \nu R - dS - \beta SI \tag{29a}$$
$$I' = \beta SI - (d+\gamma)I \tag{29b}$$
$$R' = bp + \gamma I - (d+\nu)R \tag{29c}$$

Consider the initial value problem consisting in (29) to which we adjoin initial conditions $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ and $R(0) = R_0 \geq 0$

Typically, we assume $N_0 = S_0 + I_0 + R_0 > 0$ to avoid a trivial case

This modification doesn't change much

Equation (20) for the total population is unchanged

The Jacobian (25) at arbitrary point is also unchanged

The DFE is affected, though; as a consequence, so is the reproduction number

The DFE for the SIRS vaccination model

Considering (29) at equilibrium and substituting $I^\star = 0$ into this system gives

$$\begin{aligned} 0 &= (1-p)b + \nu R^\star - dS^\star \\ 0 &= bp - (d+\nu)R^\star \end{aligned}$$

which we rewrite as the linear system

$$\begin{pmatrix} d & -\nu \\ 0 & d+\nu \end{pmatrix} \begin{pmatrix} S^\star \\ R^\star \end{pmatrix} = \begin{pmatrix} (1-p)b \\ bp \end{pmatrix}$$

Thus

$$\begin{aligned} \begin{pmatrix} S^\star \\ R^\star \end{pmatrix} &= \frac{1}{d(d+\nu)} \begin{pmatrix} d+\nu & \nu \\ 0 & d \end{pmatrix} \begin{pmatrix} (1-p)b \\ pb \end{pmatrix} \\ &= \frac{1}{d(d+\nu)} \begin{pmatrix} (d+\nu)(1-p)b + pb\nu \\ pbd \end{pmatrix} \end{aligned}$$

As a consequence, the DFE takes the form

$$\boldsymbol{E}_0^v := (S^\star, I^\star, R^\star) = \left(\left(1 - p + \frac{p\nu}{d+\nu} \right) N^\star, 0, \frac{pd}{d+\nu} N^\star \right) \tag{30}$$

Substituting (30) into the eigenvalue that determines stability of the DFE, $\beta S^\star - (d+\gamma)$, we get

$$\begin{aligned} \beta S^\star - (d+\gamma) < 0 &\iff \frac{\beta}{d+\gamma} S^\star < 1 \\ &\iff \frac{\beta}{d+\gamma} \left(1 - p + \frac{p\nu}{d+\nu} \right) N^\star < 1 \end{aligned}$$

So we define

$$\mathcal{R}_0^v = \frac{\beta}{d+\gamma} \left(1 - p + \frac{p\nu}{d+\nu} \right) N^\star \tag{31}$$

Herd immunity

Therefore

- $\mathcal{R}_0^{\mathrm{v}} < \mathcal{R}_0$ if $p > 0$
- To control the disease, $\mathcal{R}_{\mathrm{v}}$ must take a value less than 1, i.e.,

$$\mathcal{R}_{\mathrm{v}} < 1 \iff p > 1 - \frac{1}{\mathcal{R}_0} \tag{32}$$

By vaccinating a fraction $p > 1 - 1/\mathcal{R}_0$ of newborns, we thus are in a situation where the disease is eventually eradicated

This is **herd immunity** (*bis repetita*)

To simplify or not to simplify?

- In the KMK epidemic model (1) and the SIRS endemic model (21), since the total population is constant or asymptotically constant, it is possible to omit one of the state variables since $N^\star = S + I + R$
- We often use $R = N^\star - S - I$
- This can greatly simplify some computations
- Whether to do it or not is a matter of preference

To normalise or not to normalise?

- In the KMK epidemic model (1) and the SIRS endemic model (21), since the total population is constant or asymptotically constant, it is possible to normalise to $N = 1$

- This can greatly simplify some computations

- However, I am not a big fan: it is important to always have the "sizes" of objects in mind

- If you do normalise, at least for a paper destined to mathematical biology, always do a "return to biology", i.e., interpret your results in a biological light, which often implies to return to original values

Where we are

- An *epidemic* SIR model (the KMK SIR) in which the presence or absence of an epidemic wave is characterised by the value of $\mathcal{R}_0$

- The KMK SIR has explicit solutions (in some sense). **This is an exception!**

- An *endemic* SIRS model in which the threshold $\mathcal{R}_0 = 1$ is such that, when $\mathcal{R}_0 < 1$, the disease goes extinct, whereas when $\mathcal{R}_0 > 1$, the disease becomes established in the population

Bibliography


Hirsch, M. and Smale, S. (1974).
Differential Equations, Dynamical Systems, and Linear Algebra.
Academic Press.


van den Driessche, P. and Watmough, J. (2002).
Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission.
Mathematical Biosciences, 180:29–48.