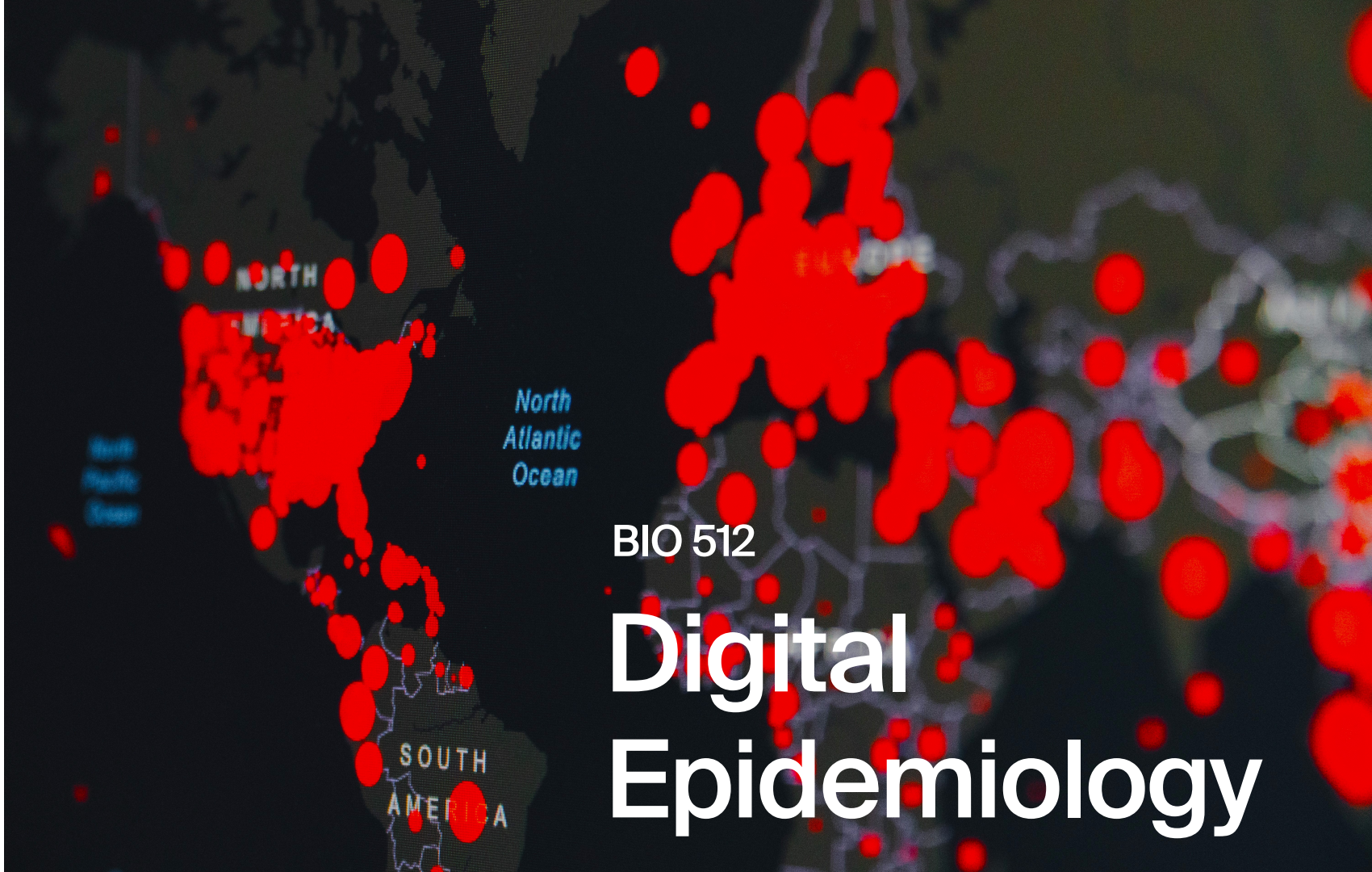


BIO 512

Digital Epidemiology

Introduction
Welcome!

- Marcel Salathé

- I’m a professor in SV with a joint appointment in IC.
- I run a group called digital epidemiology lab, ~~based at Campus Biotech.~~
- I’m trained as a biologist, with a PhD in evolutionary biology.

Introduction
About This Class

- This is a class about digital epidemiology.
- You are the third cohort to go through this class.
- For the second time, the class will be lectures + project.
- Future generations of EPFL students will be grateful for your feedback. #PayItForward
- Today, we will first go through some organizational details of the class.

Introduction
Class Structure

- Lecture Phase + Project Phase
- Lecture phase ~ in the middle of the semester
- Project phased begins right after

- **Mid-term (and only) exam: 1. April 2024 (no joke)**

Introduction
Lecture Phase

- 9 lectures - these correspond to 10 chapters in the book.
- Book: www.digitalepibook.com
- Slides will be made available on Moodle before the lecture.
- ~2 chapters / week

Course Schedule

PHASE 1: Lectures

18.02.25	Course Introduction Chapter 1: Epidemiology
20.02.25	Chapter 2: Testing & Diagnostics
25.02.25	Chapter 3: Epidemiological Studies
27.02.25	Chapter 4: Infectious Disease Epidemiology
04.03.25	Chapter 5: Modeling Infectious Diseases
06.03.25	Chapter 6: Spatial Models & Network Models
11.03.25	Chapter 7: Digital Contact Tracing
13.03.25	no class
18.03.25	no class
20.03.25	Chapter 8: Digital Public Health Surveillance
25.03.25	Chapter 9: Digital Cohorts & Trials
27.03.25	Chapter 10: Ethics of Digital Epidemiology
01.04.25	**Midterm Exam**

Introduction
Exercises

- **No exercise sessions** - if you need help, ask us
- In this class, you are strongly encourage to use LLMs.
- Exercise due date: Sunday night of week after exercise has been handed out.
- Exercises will not be graded for correctness, but for **reasonable submission.** 1% of grade per exercise, except 2% for first exercise.

Introduction
Project Phase

- Goal: hands-on experience with building a digital epidemiology prototype
- Find your project in an ideation session (10. April)
- Pitch it to class on 15. April
- Work in groups of 3
- 2 graded assessments (6. May, 15. May)
- 1 progress report to class
- Demo Day - Presentation (27. May) 🥳

03.04.25	Project Prototyping Workshop 1
08.04.25	Project Prototyping Workshop 2
10.04.25	Project ideation
15.04.25	Opening pitch to class
17.04.24	no class

PHASE 2: Project

29.04.25	No class - work on project
01.05.25	No class - work on project
06.05.25	**1st project assessment (graded, 5%)**
08.05.25	No class - work on project
13.05.25	1st progress report to class
15.05.25	**2nd project assessment (graded, 10%)**
20.05.25	No class - submit report, not graded
22.05.25	No class - get feedback on written report
27.05.25	**Demo (graded 10%)** **Due date final written project report (graded, 20%)**

Introduction
Midterm Exam

- Mid-term exam 45% of grade.
- Exam material is **everything discussed in class**, and **everything in the book**.
- You can take one sheet of A4 paper (hand written notes only) with you to the exam.

Introduction
Final grade

- Exercises: 10%
- Mid-term exam: 45%
- Project: 5% + 10% + 10% (demo) + 20% (report) = 45%

- You will always know where you stand, no surprises / gotchas

Introduction
Help & Communication

- TAs: **Rohan** Singh

Marouane Toumi

- There is no office hour - if you have questions, we are always available to meet with you (contact us via Moodle).
- All communication **MUST GO** through Moodle (unless agreed otherwise)

Introduction
Lecture Structure

- 1/3 Basic concepts in epidemiology
- 1/3 Infectious disease dynamics, computational approaches
- 1/3 Digital approaches to epidemiology
- Discussion about ethical issues

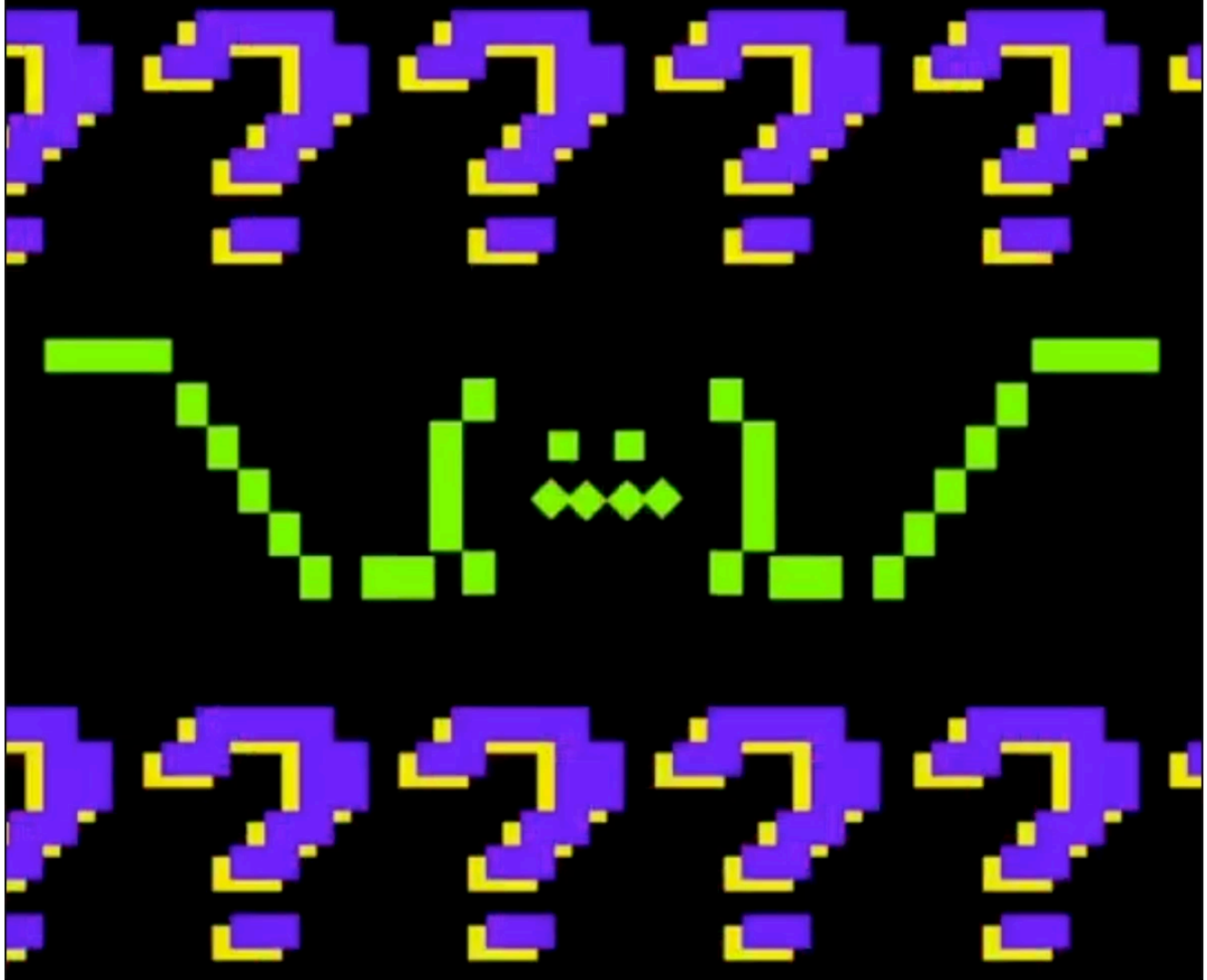
Introduction
General Remarks

- This class mixes a lot fields and approaches. Life sciences, CS, math, reading, programming, quantitive and qualitative reasoning - we'll be all over the map. Welcome to novel approaches in a multidisciplinary world.
- This is a breadth-first class. You will learn many things from many different fields, and you'll likely find it useful in unexpected ways.
- This is not a class about COVID-19, but you'll hear it often.


RESPECT

Introduction
RESPECT

- Limit use of laptops and phones in class (ideally zero).
- We will do our best to help you succeed. If you find anything suboptimal to your learning process, please let us know.



Epidemiology
Learning Goals

- Understand what epidemiology is about
- Understand surveillance, and different types of it
- Understand the concepts of incidence and prevalence
- Understand how a case is defined
- Understand measures of morbidity
- Understand measures of mortality

Epidemiology
What is epidemiology?

- Epidemiology (epi + demos + logos)
- The study of the distribution and determinants of health and disease in the population.
- *Who* is sick (*what*) or healthy, *when*, *where*, and *why*

Epidemiology
Who?

- Epidemiology is about groups, not individuals
- Always ask: what group are we talking about?

Original article

EDITOR'S CHOICE

From measures to models: an evaluation of air pollution exposure assessment for epidemiological studies of pregnant women

E Nethery,[1] S E Leckie,[1] K Teschke,[1,2] M Brauer[1]

► Additional appendices are published online only at http://oem.bmj.com/content/vol65/issue9

[1] School of Environmental Health, The University of British Columbia, Vancouver, BC

ABSTRACT

Objectives: To evaluate exposure estimation methods such as spatially resolved land-use regression models and ambient monitoring data in the context of epidemiological studies of the impact of air pollution on pregnancy outcomes.

importance of capturing within-city spatial variability in air pollution exposure.[13 14] Specifically, studies of traffic-related air pollution have used proximity (ie, living near a busy road),[15] traffic volume or density measures[5 16] or land-use regression (LUR)[17] models as exposure indicators. LUR

Epidemiology
Who?

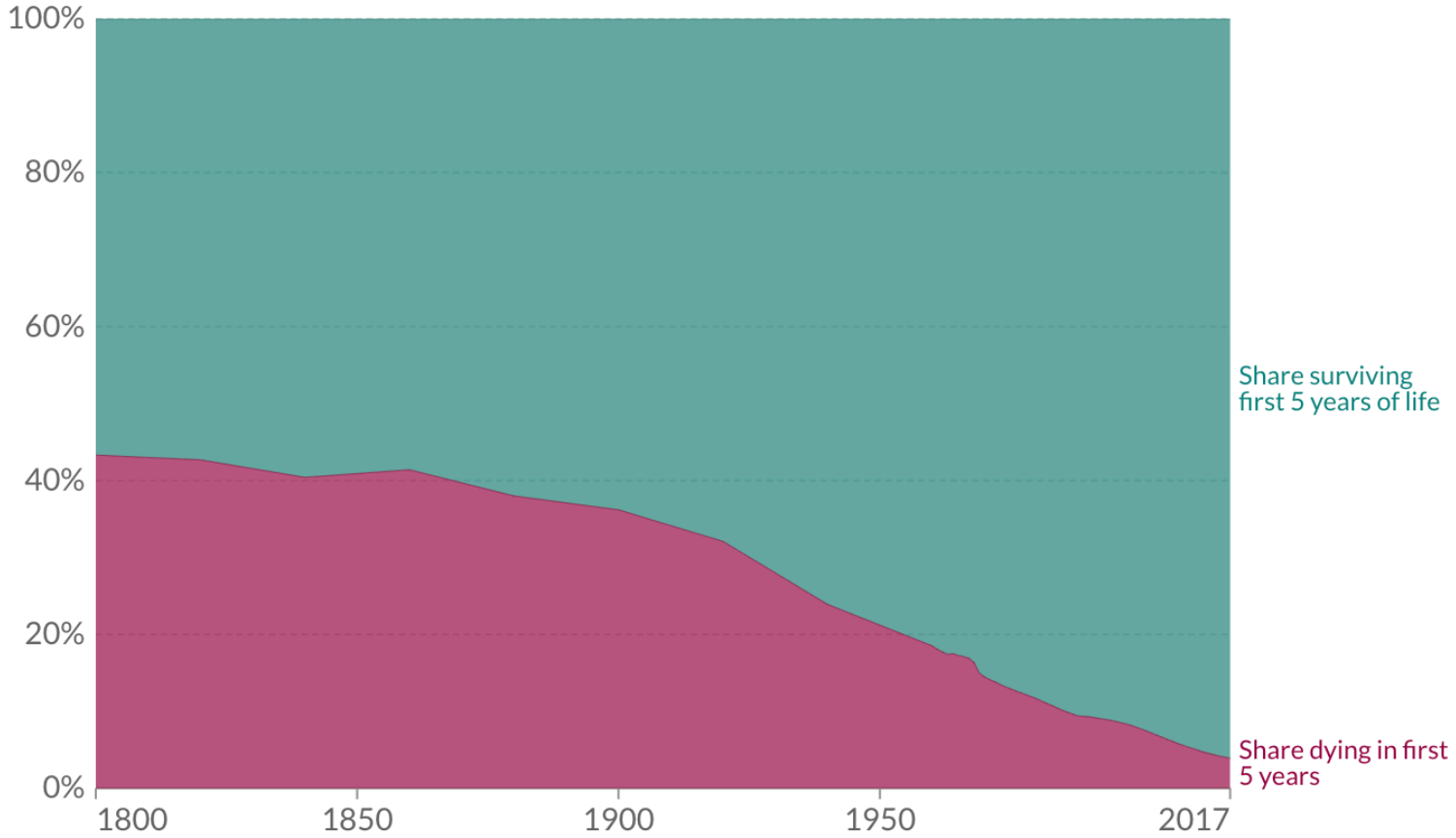
- How to study large groups: take a representative sample
- Descriptive statistics: helps us to *describe* the sample.
- Inferential statistics: helps us to *make inferences* about the larger group from the sample.
- Careful about biases

Epidemiology
What?

- Morbidity
- Mortality
- Why this distinction?
  1. Death was ubiquitous already early in life

Global child mortality

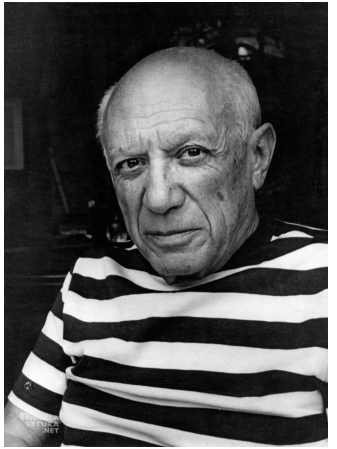
Share of the world population dying and surviving the first 5 years of life.



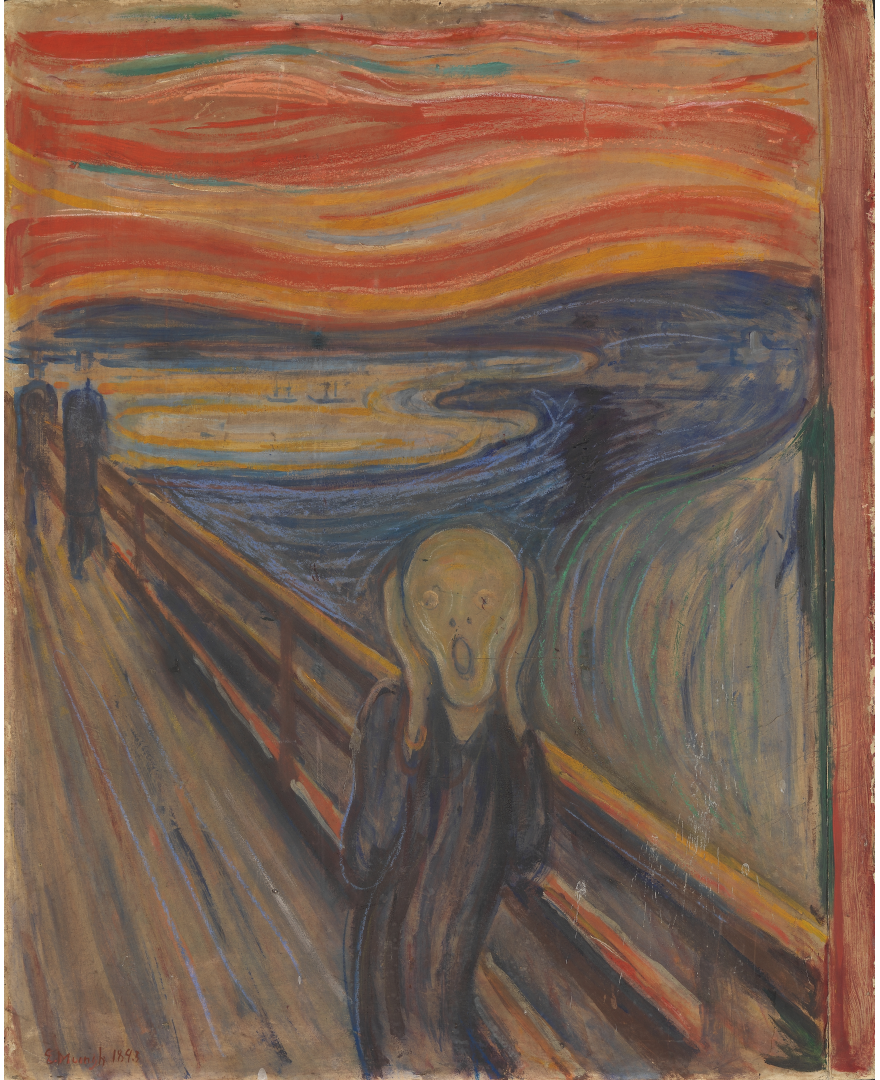
Source: Gapminder and the World Bank

OurWorldInData.org/child-mortality • CC BY

- Salvador Dalí: older brother (also Salvador), died of gastroenteritis before his second birthday. Dalí, who was born nine months later and given the same first name, believed for the rest of his life to be his dead brother's reincarnation.



- Pablo Picasso: lost sister to diphtheria when she was barely seven years old. Pleaded with God to save her in exchange for his sacrifice to give up painting, and he was convinced that her death was a sign that God wanted him to be a painter.



- Edvard Munch’s older sister died at age fifteen of tuberculosis, the same disease that killed his mother when he was just a five year old boy. Munch processed this tragic event in several paintings throughout his career.

Epidemiology
What?

- Morbidity
- Mortality
- Why this distinction?
  1. Death was ubiquitous already early in life
  2. Death is universally understood (as opposed to disease and disease severity, which can have different definitions)
  3. Death is final

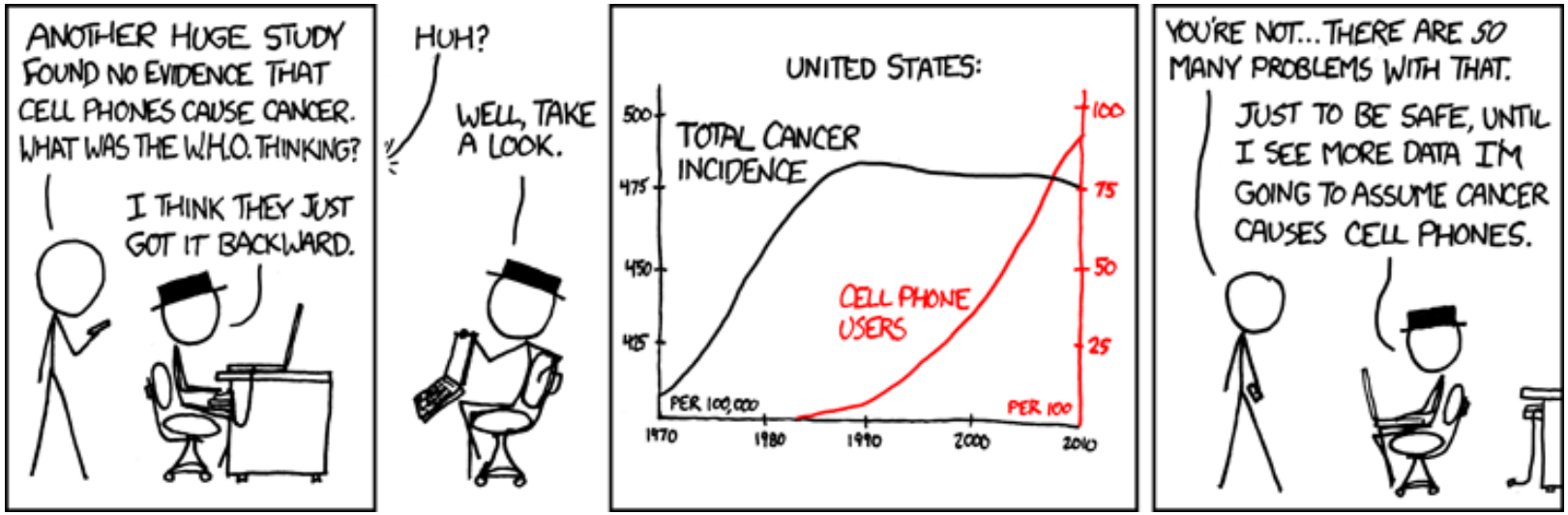
Epidemiology
When?

- When **during life** - childhood diseases and late-life diseases have different impacts
- When **in the calendar** - understanding timing of disease is critical to understand disease dynamics. E.g. many people getting sick at once, or a rapid growth of cases, has important implications.

Epidemiology
Why?

- Epidemiology often finds associations (correlations) - but this are not the same as causation.
- Association can already be enough for action, or a call for more studies - but understanding causes is ideal, as it maximizes our ability to control an outcome.

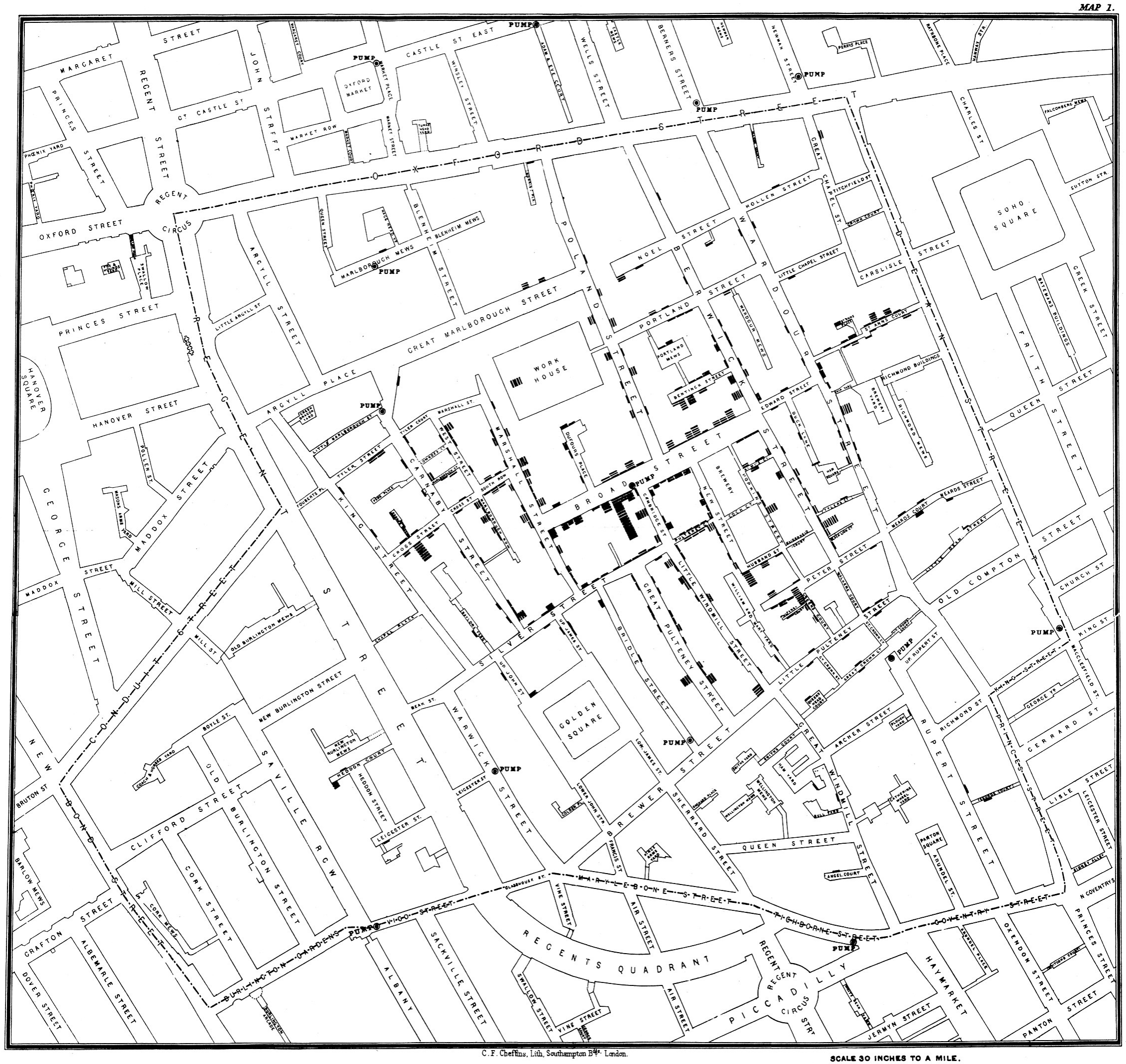
Epidemiology
Why?



Source: XKCD

Epidemiology
Where?

- Cholera outbreak in Soho district London, 1854
- John Snow: Hypothesized that cholera was caused by agents in polluted water (not miasma)
- Cholera was a massive problem at the time - a 1849 outbreak killed 14,137 residents in London.


MAP 1.
OXFORD STREET
REGENT CIRCUS
PRINCES STREET
HANOVER SQUARE
HANOVER STREET
GREAT MARLBOROUGH STREET
ARGYLL PLACE
ARGYLL STREET
MARLBOROUGH MEWS
POLAND STREET
BERWICK STREET
WARDOUR STREET
SOHO SQUARE
FRITH STREET
DEAN STREET
WORK HOUSE
BROAD STREET
PUMP
MARSHALL STREET
CARNABY STREET
KING STREET
NEW STREET
CROSS STREET
REGENT STREET
GREAT PULTENEY STREET
LITTLE WINDMILL STREET
GOLDEN SQUARE
WARWICK STREET
BREWER STREET
OLD COMPTON STREET
QUEEN STREET
RUPERT STREET
GERRARD ST
LISLE STREET
CONDUIT STREET
NEW BURLINGTON STREET
SAVILLE ROW
CLIFFORD STREET
OLD BURLINGTON STREET
CORK STREET
BURLINGTON GARDENS
VIGO STREET
MARYLEBONE STREET
REGENTS QUADRANT
PICCADILLY
REGENT CIRCUS
COVENTRY STREET
HAYMARKET
SACKVILLE STREET
ALBANY
ALBEMARLE STREET
DOVER STREET
GRAFTON STREET
BRUTON ST
NEW BOND STREET
GEORGE STREET
MADDOX STREET
JERMYN STREET
PANTON STREET
C.F. Cheffins, Lith, Southampton Bdgs. London.
SCALE 30 INCHES TO A MILE.

Epidemiology
Surveillance

- Public health surveillance - many different definitions of what counts as surveillance.
- WHO: “the continuous, systematic collection, analysis and interpretation of health-related data”

Epidemiology
Surveillance

- Passive vs active surveillance

CS324866-A

NATIONAL NOTIFIABLE DISEASES SURVEILLANCE SYSTEM

To protect Americans from serious disease, the National Notifiable Diseases Surveillance System (NNDSS) helps public health monitor, control, and prevent about 120 diseases. These national notifiable diseases are important to monitor nationwide and include infectious diseases such as Zika, foodborne outbreaks such as *E. coli*, and noninfectious conditions such as lead poisoning. About 3,000 public health departments gather and use data on these diseases to protect their local communities. Through NNDSS, CDC receives and uses these data to keep people healthy and defend America from health threats.

Accessible version: https://www.cdc.gov/nndss/about/index.html

NNDSS BY THE NUMBERS

120 — **120 diseases under surveillance**
- infectious diseases
- bioterrorism agents
- sexually transmitted diseases
- noninfectious conditions

2.7M — **Nearly 2.7 million** disease events reported through NNDSS each year

3K — **About 3,000 public health departments** send disease data to 60 state, territorial, and other public health departments, who then send the data to CDC

100% — **100% of the American population protected**

Epidemiology
Surveillance

- Passive vs active surveillance

Article

Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'

https://doi.org/10.1038/s41586-020-2488-1

Received: 2 April 2020

Accepted: 23 June 2020

Published online: 30 June 2020

Check for updates

Enrico Lavezzo[1,13], Elisa Franchin[1,13], Constanze Ciavarella[2], Gina Cuomo-Dannenburg[2], Luisa Barzon[1], Claudia Del Vecchio[1], Lucia Rossi[3], Riccardo Manganelli[1], Arianna Loregian[1], Nicolò Navarin[4,5], Davide Abate[1], Manuela Sciro[3], Stefano Merigliano[6], Ettore De Canale[3], Maria Cristina Vanuzzo[3], Valeria Besutti[3], Francesca Saluzzo[1], Francesco Onelia[1], Monia Pacenti[3], Saverio G. Parisi[1], Giovanni Carretta[3], Daniele Donato[3], Luciano Flor[3], Silvia Cocchio[7], Giulia Masi[1], Alessandro Sperduti[4,5], Lorenzo Cattarino[2], Renato Salvador[6], Michele Nicoletti[8], Federico Caldart[8], Gioele Castelli[8], Eleonora Nieddu[8], Beatrice Labella[8], Ludovico Fava[8], Matteo Drigo[8], Katy A. M. Gaythorpe[2], Imperial College COVID-19 Response Team*, Alessandra R. Brazzale[9], Stefano Toppo[1,5], Marta Trevisan[1], Vincenzo Baldo[7], Christl A. Donnelly[2,10], Neil M. Ferguson[2], Ilaria Dorigatti[2,14]✉ & Andrea Crisanti[1,11,14]✉

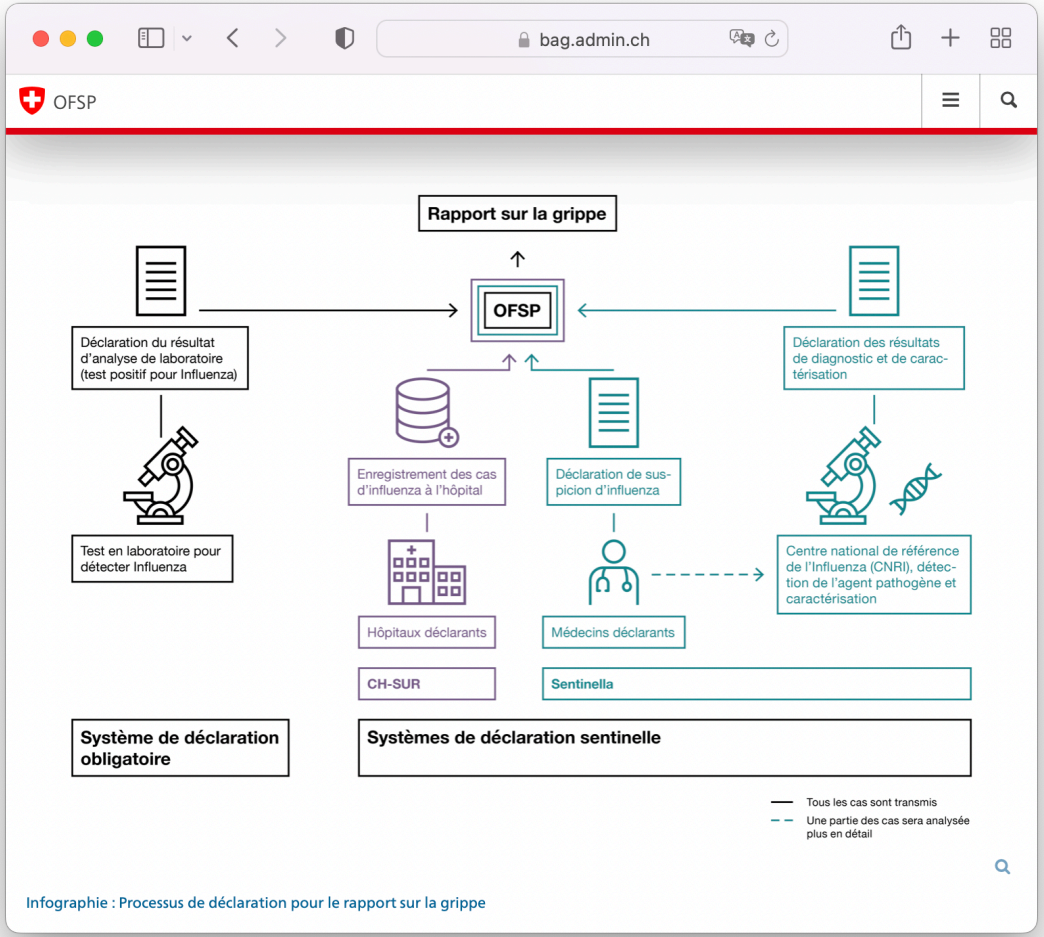
On 21 February 2020, a resident of the municipality of Vo', a small town near Padua (Italy), died of pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection[1]. This was the first coronavirus disease 19 (COVID-19)-related death detected in Italy since the detection of SARS-CoV-2 in the Chinese city of Wuhan, Hubei province[2]. In response, the regional authorities imposed the lockdown of the whole municipality for 14 days[3]. Here we collected information on the demography, clinical presentation, hospitalization, contact network and the presence of SARS-CoV-2 infection in nasopharyngeal swabs for 85.9% and 71.5% of the population of Vo' at two consecutive time points. From the first survey, which was conducted around the time the town lockdown started, we found a prevalence of infection of 2.6% (95% confidence interval (CI): 2.1–3.3%). From the second survey, which was conducted at the end of the lockdown, we found a prevalence of 1.2% (95% CI: 0.8–1.8%). Notably, 42.5% (95% CI: 31.5–54.6%) of the confirmed SARS-CoV-2 infections detected across the two surveys were asymptomatic (that is, did not have symptoms

Epidemiology
Surveillance

- Population-based vs sentinel surveillance

Epidemiology
Surveillance

- Influenza:
  1-5% of physicians voluntarily reporting



Infographie : Processus de déclaration pour le rapport sur la grippe

Epidemiology
Surveillance

- Case-based vs aggregate surveillance

  often a question of resources

Epidemiology
Surveillance

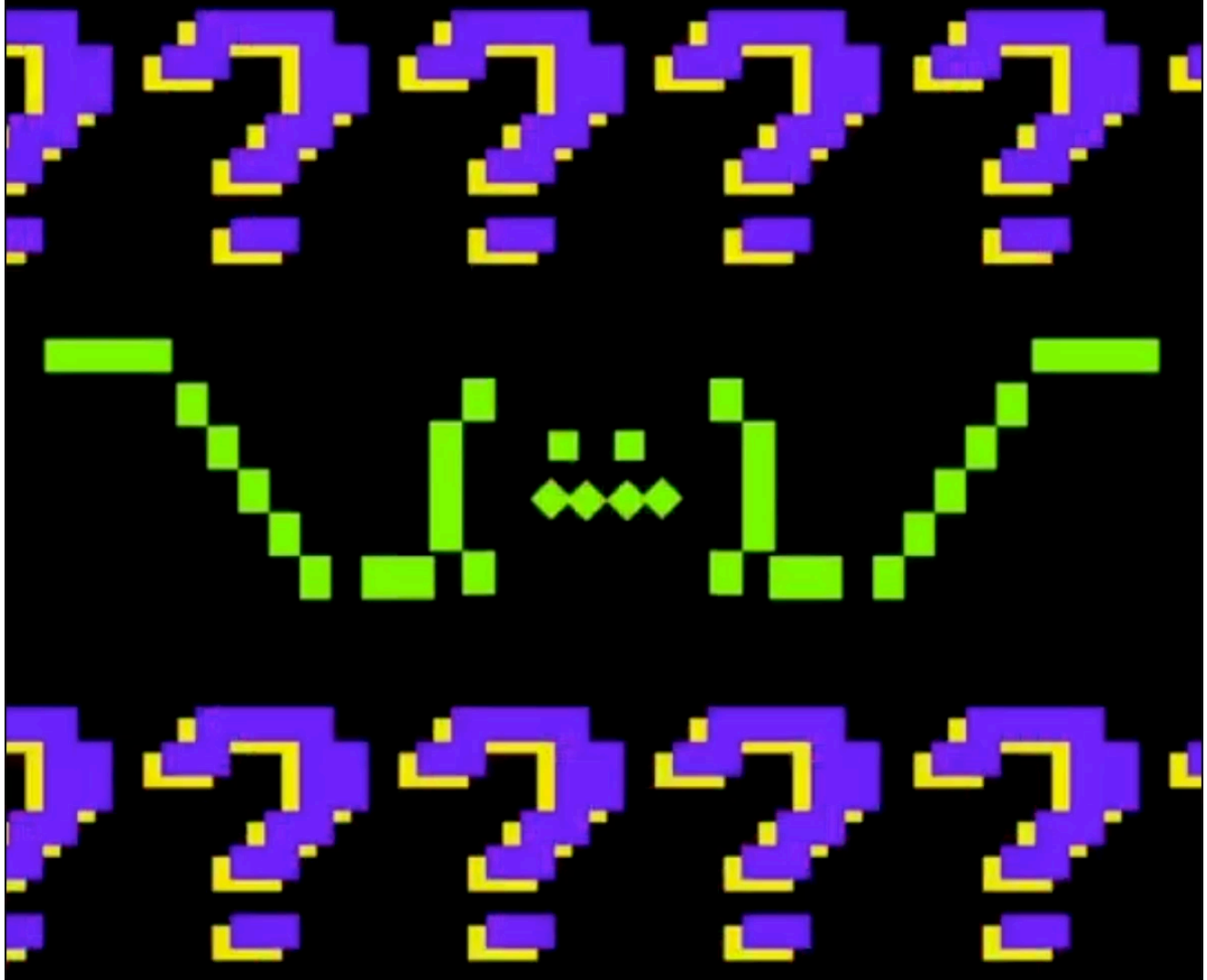
- Syndromic vs laboratory surveillance

Digital public health surveillance always syndromic, but may change soon.

Epidemiology
Surveillance

- Event-based vs indicator-based surveillance

Most of digital public health surveillance is event-based



Epidemiology
Incidence and Prevalence

- Incidence: number of new cases during a specified period of time.
- Prevalence: number of cases present during a specified period of time.

Epidemiology
Incidence and Prevalence

- Incidence: number of new cases during a specified period of time.
- Prevalence: number of cases present during a specified period of time.

- Attack rate: incidence / population at risk (also called incidence rate)

Epidemiology
Incidence and Prevalence

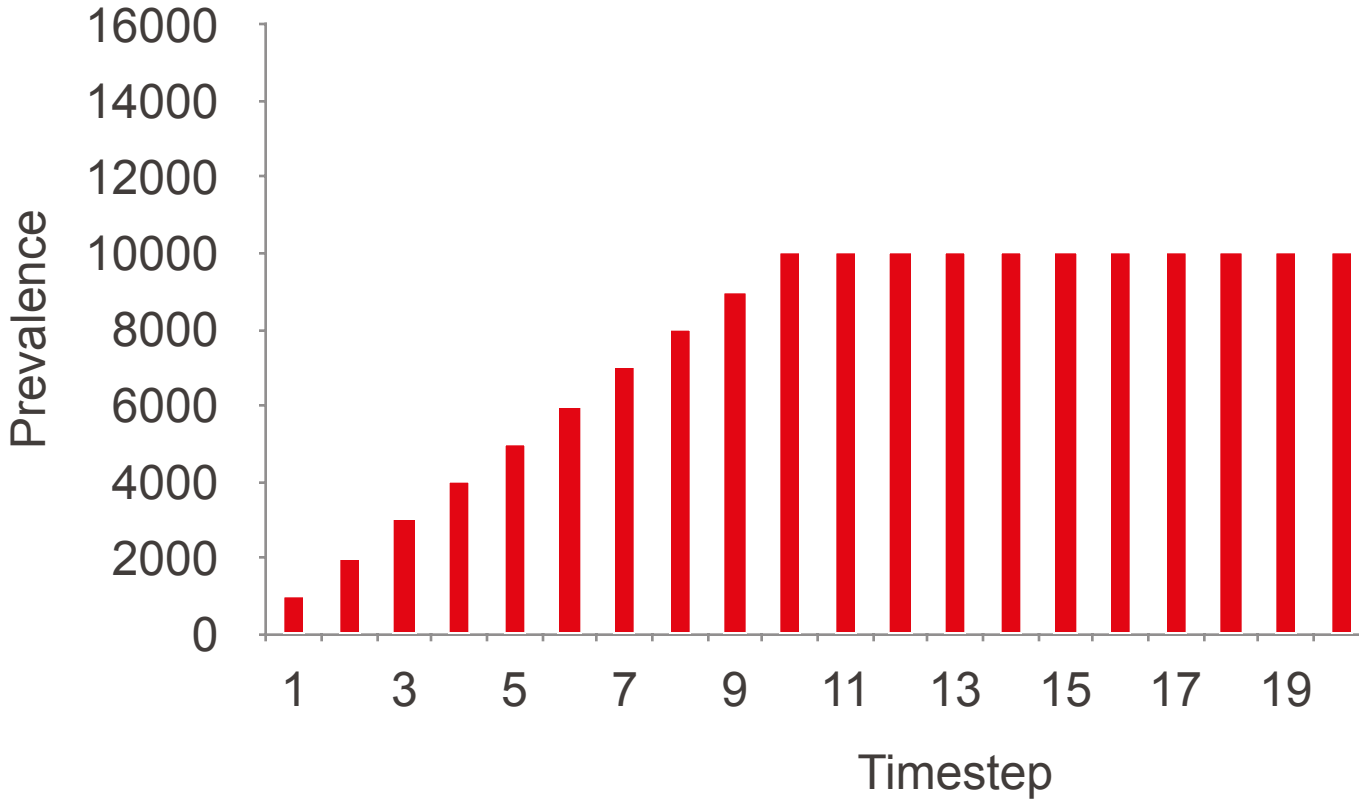
- Incidence: a measure of risk
- Prevalence: a measure of “diseases pressure / burden” in the system

Epidemiology
Incidence and Prevalence

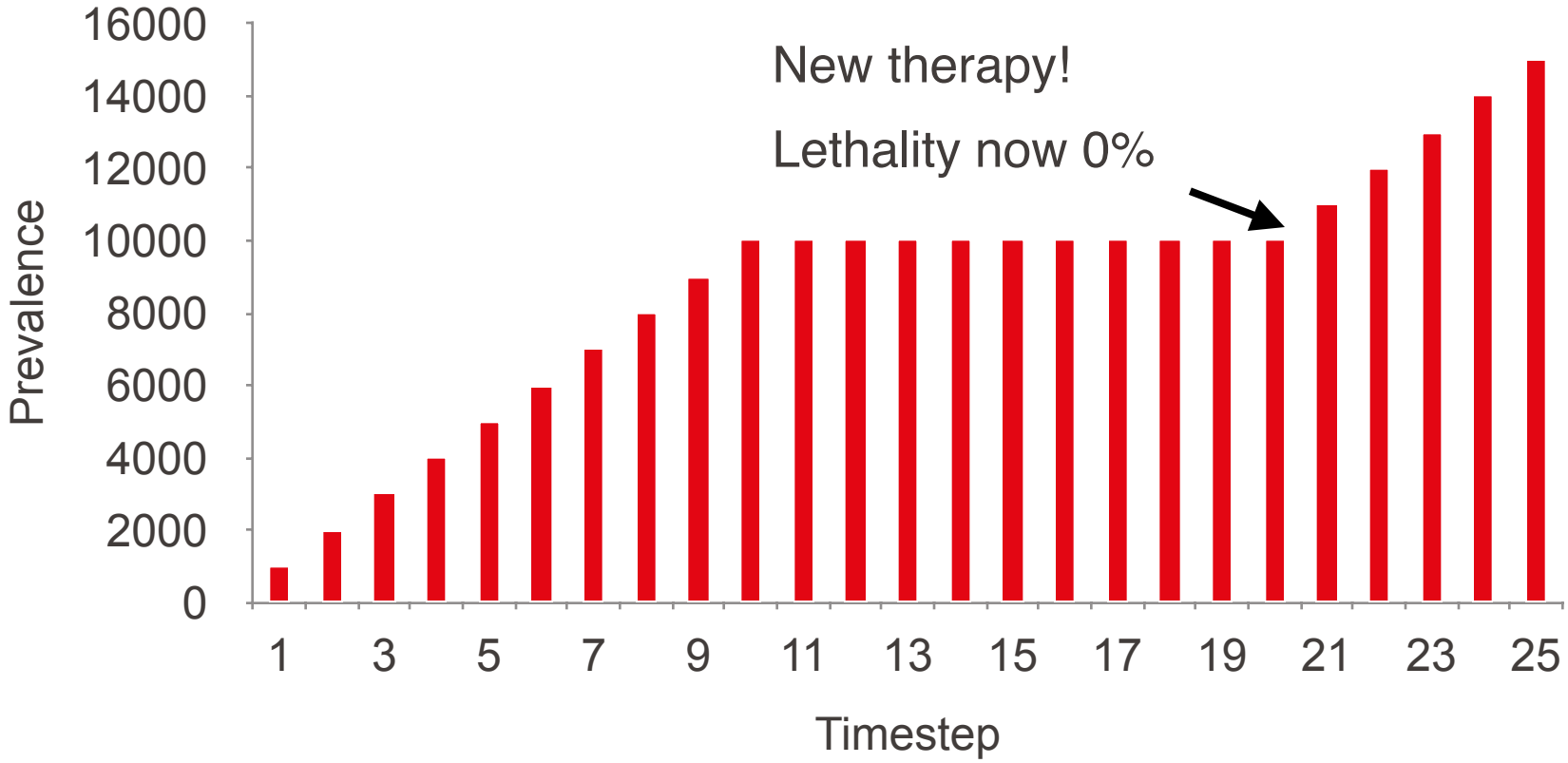
- Case study: A public health *improvement* leads to an increase in prevalence. How is this possible?

  Assume:
  1. Incidence of 1000 new cases per time step
  2. Disease incurable
  3. Disease 100% lethal after 10 time steps

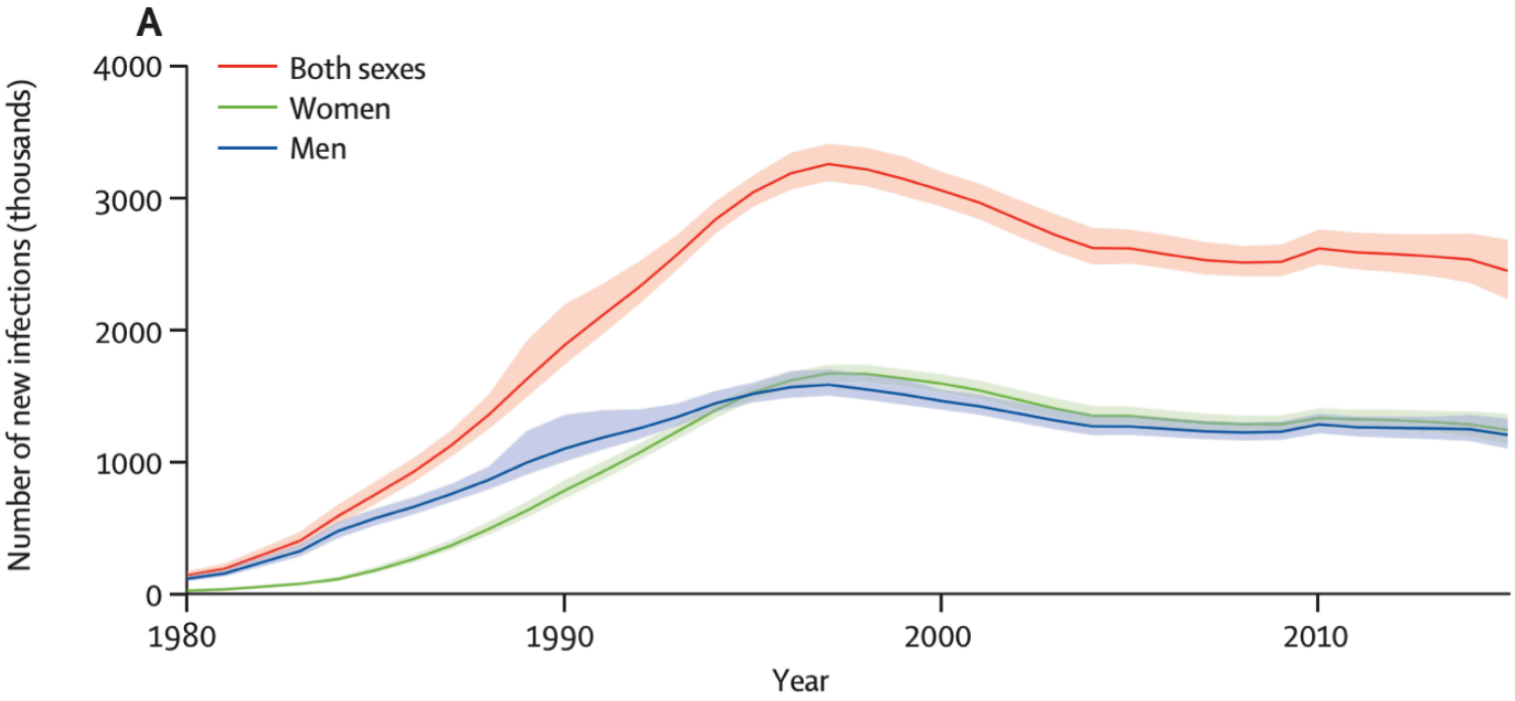
Epidemiology
Incidence and Prevalence



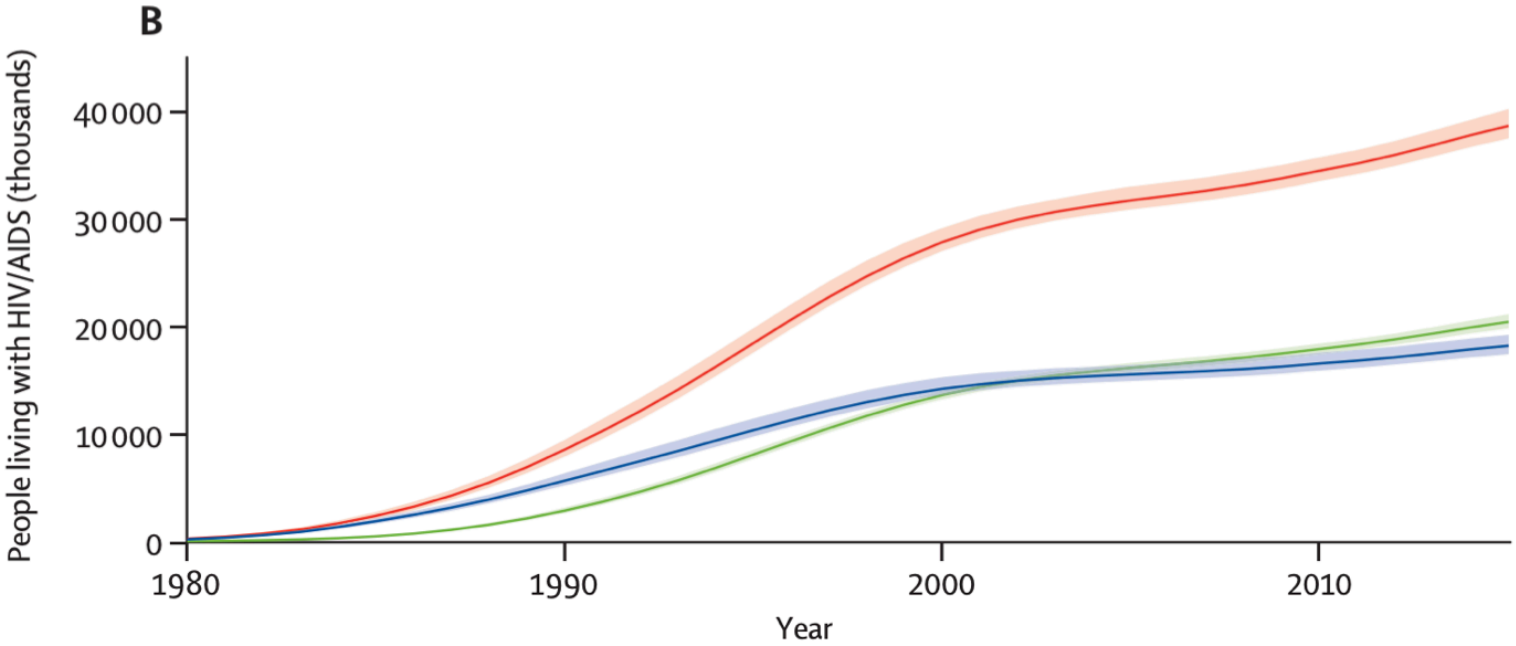
Epidemiology
Incidence and Prevalence



Epidemiology
HIV Incidence

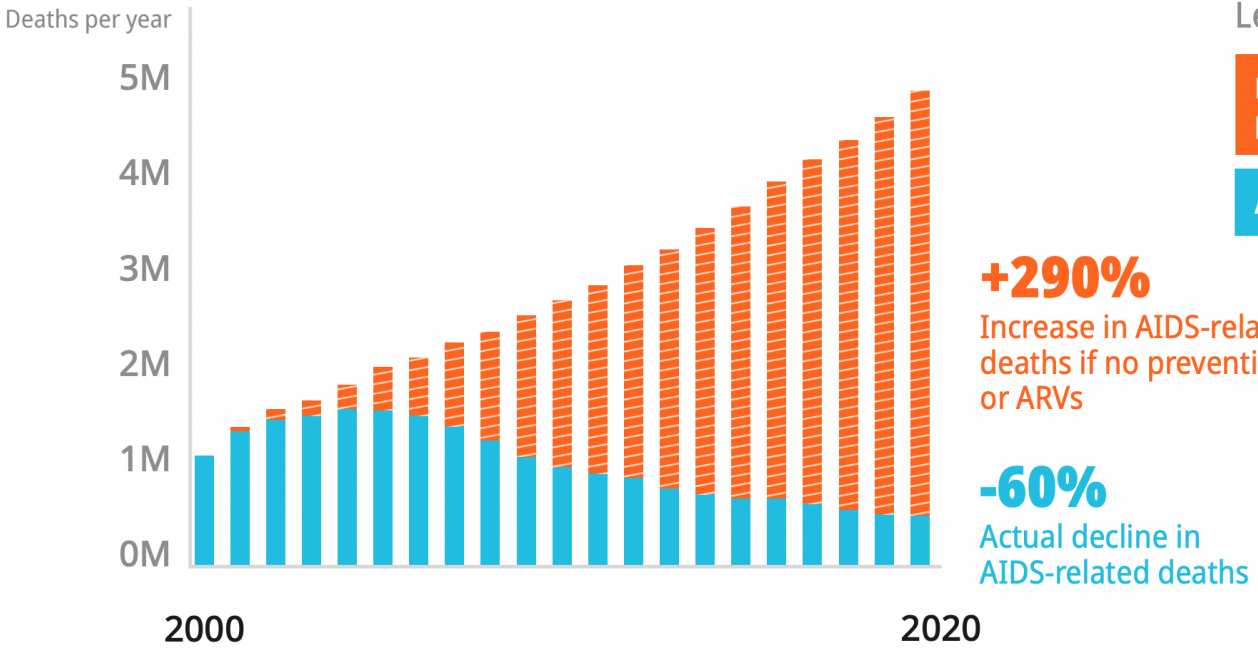


Epidemiology
HIV Prevalence



A dramatic shift toward progress in the HIV/AIDS epidemic

AIDS-related deaths



Legend

Estimated AIDS-related deaths without prevention or antiretroviral treatments (ARVs)

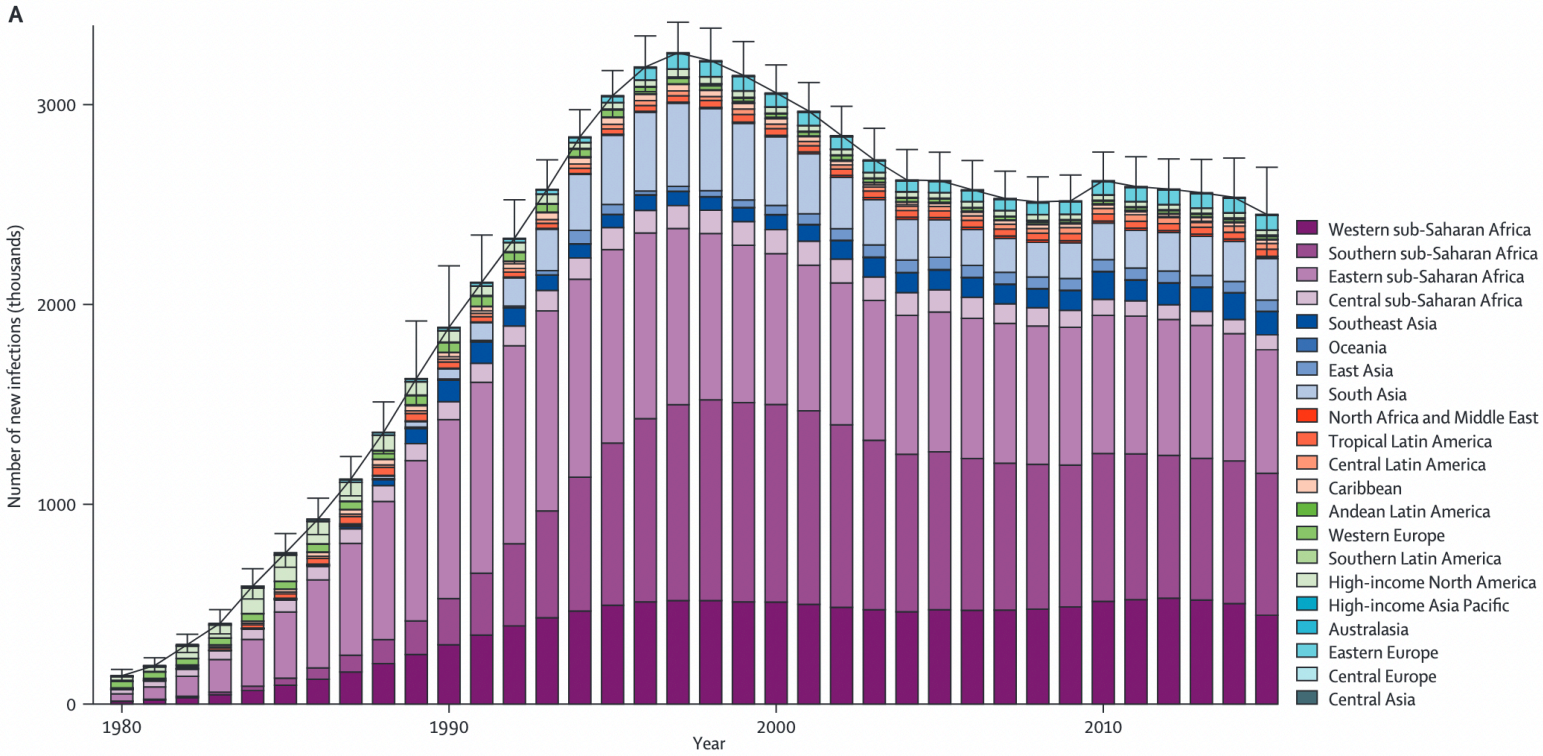
AIDS-related deaths

+290%
Increase in AIDS-related deaths if no prevention or ARVs

-60%
Actual decline in AIDS-related deaths

Source: Goalkeepers Report 2022, Gates Foundation

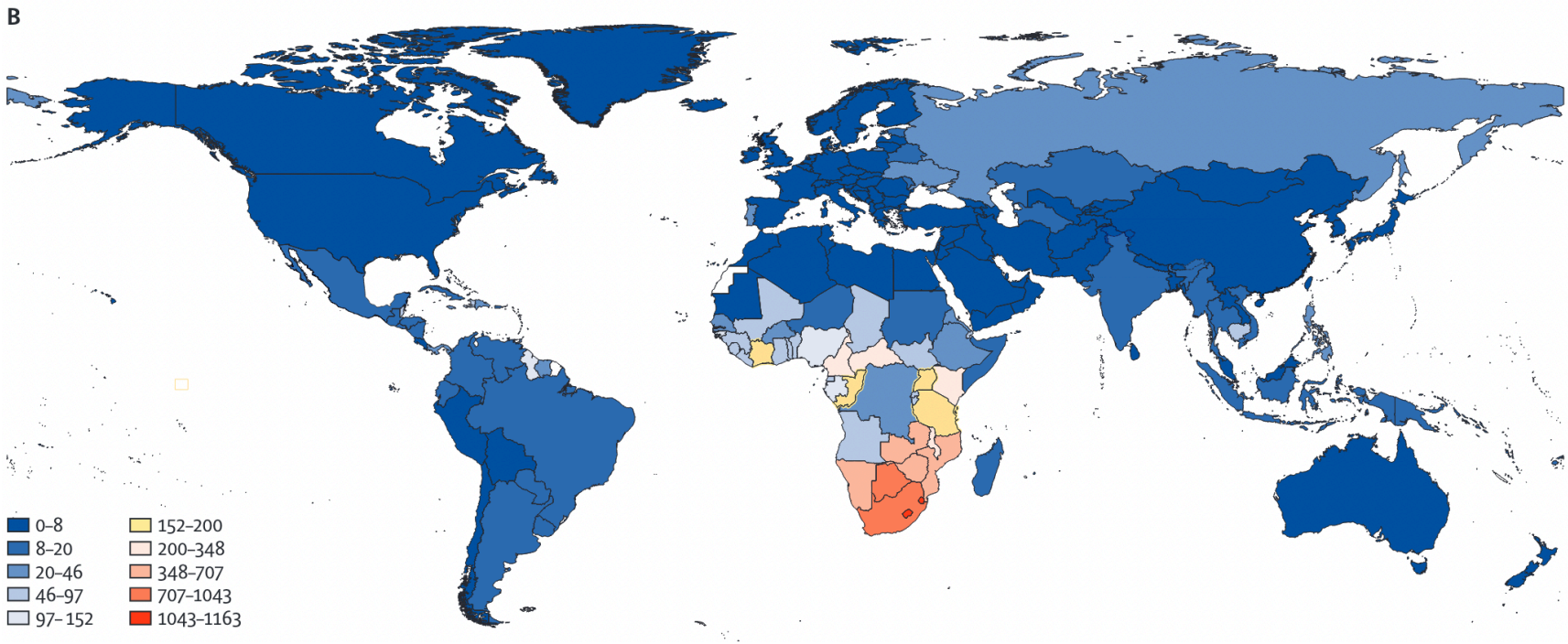
Epidemiology
HIV Incidence



Bars show the mean number of estimated new infections within a given year.

Source: Wang et al. 2016

Epidemiology
HIV Incidence



Rates are per 100 000 people.

Source: Wang et al. 2016

Epidemiology

HIV Incidence and Prevalence

	Country/Region	Adult prevalence of HIV/AIDS[1]	Number of people with HIV/AIDS	Annual deaths from HIV/AIDS[3]	Year of estimate
1	Eswatini	27.90%	200,000	2,300	2019
2	Lesotho	23.10%	340,000	4,800	2019
3	Botswana	22.20%	380,000	5,000	2019
4	Zimbabwe	21.40%	1,400,000	20,000	2019
5	South Africa	13.30%	7,500,000	72,000	2019
6	Namibia	12.70%	210,000	3,000	2019
7	Mozambique	12.10%	2,200,000	51,000	2019
8	Zambia	12.10%	1,200,000	17,000	2019
9	Malawi	9.50%	1,100,000	13,000	2016
10	Equatorial Guinea	7.00%	65,000	1,800	2019
79	United States	0.36%	1,189,700	-	2019[13]
102	Switzerland	0.20%	17,000	-	2019

Epidemiology
Prevalence

- Point prevalence vs period prevalence
- Seroprevalence

Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study

Silvia Stringhini, Ania Wisniak, Giovanni Piumatti*, Andrew S Azman*, Stephen A Lauer, Hélène Baysson, David De Ridder, Dusan Petrovic, Stephanie Schrempft, Kailing Marcus, Sabine Yerly, Isabelle Arm Vernez, Olivia Keiser, Samia Hurst, Klara M Posfay-Barbe, Didier Trono, Didier Pittet, Laurent Gétaz, François Chappuis, Isabella Eckerle, Nicolas Vuilleumier, Benjamin Meyer, Antoine Flahault, Laurent Kaiser, Idris Guessous*

Summary

Background Assessing the burden of COVID-19 on the basis of medically attended case numbers is suboptimal given its reliance on testing strategy, changing case definitions, and disease presentation. Population-based serosurveys measuring anti-severe acute respiratory syndrome coronavirus 2 (anti-SARS-CoV-2) antibodies provide one method for estimating infection rates and monitoring the progression of the epidemic. Here, we estimate weekly seroprevalence of anti-SARS-CoV-2 antibodies in the population of Geneva, Switzerland, during the epidemic.

Methods The SEROCoV-POP study is a population-based study of former participants of the Bus Santé study and their household members. We planned a series of 12 consecutive weekly serosurveys among randomly selected participants from a previous population-representative survey, and their household members aged 5 years and older. We tested each participant for anti-SARS-CoV-2-IgG antibodies using a commercially available ELISA. We estimated seroprevalence using a Bayesian logistic regression model taking into account test performance and adjusting for the age and sex of Geneva's population. Here we present results from the first 5 weeks of the study.

Findings Between April 6 and May 9, 2020, we enrolled 2766 participants from 1339 households, with a demographic distribution similar to that of the canton of Geneva. In the first week, we estimated a seroprevalence of 4·8% (95% CI 2·4–8·0, n=341). The estimate increased to 8·5% (5·9–11·4, n=469) in the second week, to 10·9% (7·9–14·4, n=577) in the third week, 6·6% (4·3–9·4, n=604) in the fourth week, and 10·8% (8·2–13·9, n=775) in the fifth week. Individuals aged 5–9 years (relative risk [RR] 0·32 [95% CI 0·11–0·63]) and those older than 65 years (RR 0·50 [0·28–0·78]) had a significantly lower risk of being seropositive than those aged 20–49 years. After accounting for the time to seroconversion, we estimated that for every reported confirmed case, there were 11·6 infections in the community.

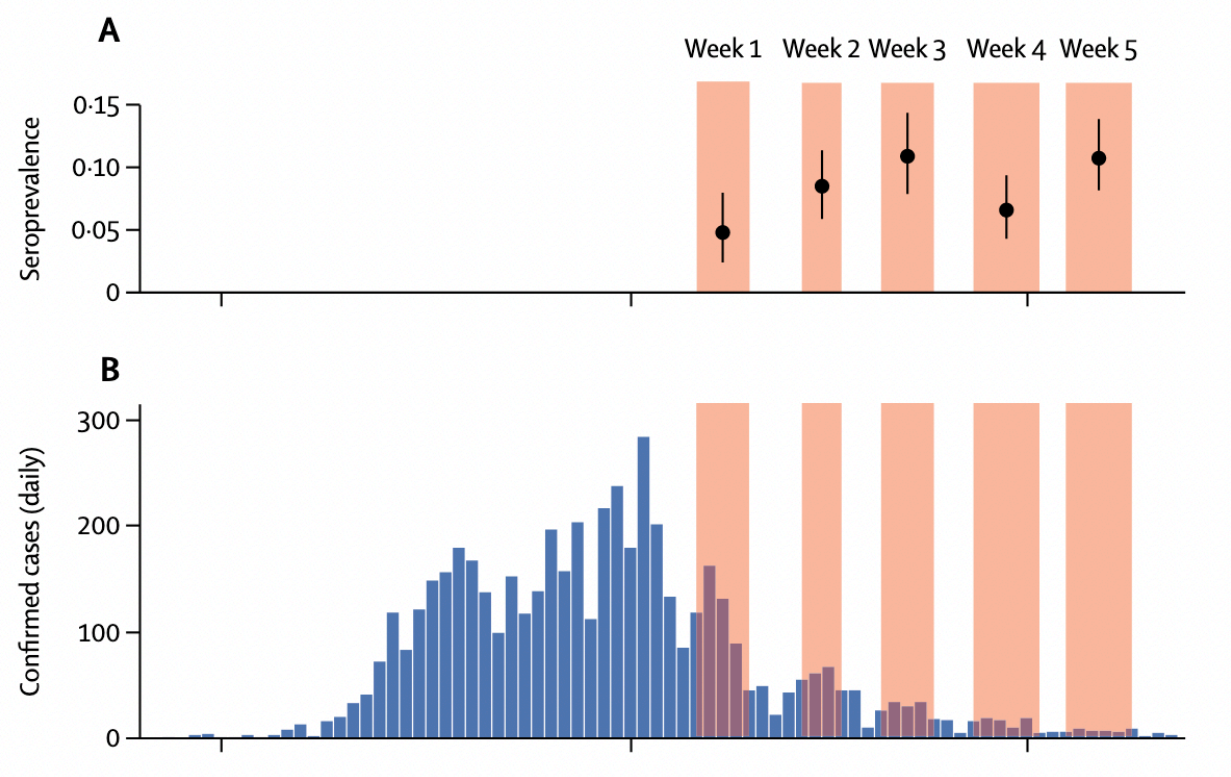
Lancet 2020; 396: 313–19

Published Online
June 11, 2020
https://doi.org/10.1016/S0140-6736(20)31304-0

*Contributed equally

Division of Primary Care (S Stringhini PhD, G Piumatti PhD, D Petrovic PhD, S Schrempft PhD, K Marcus MSc, Prof I Guessous PhD), **Division of Laboratory Medicine** (S Yerly MSc, Prof N Vuilleumier MD), **Geneva Center for Emerging Viral Diseases and Laboratory of Virology** (S Yerly, I Arm-Vernez MSc, Prof I Eckerle MD, Prof L Kaiser MD), **Division of General Pediatrics** (Prof K M Posfay-Barbe MD), **Infection Prevention and Control Program and World**

Epidemiology
Seroprevalence



Epidemiology
Seroprevalence

The preliminary results of this study provide an important benchmark to assess the state of the COVID-19 epidemic. At what appears to be the tail end of the first wave of the pandemic in Switzerland, about one in ten people have developed detectable antibodies against SARS-CoV-2, despite the fact that it was one of the more heavily affected areas in Europe.[15] Thus, assuming that the presence of the IgG antibodies measured in this study is, at least in the short term, associated with protection, these results highlight that the vast majority of the population is still immunologically naive to this new virus.

Over the course of the 5 study weeks, we observed an increase in seroprevalence from about 5% to about 11%, which is to be expected considering time to seroconversion after symptoms (median 10·4 days [IQR 8·1–13·4]) and that the peak of the epidemic was reached the week before the start of our survey (appendix p 8). As expected, our study also confirms that cases identified during the acute phase of disease provide little information on the state of the outbreak. Indeed, we observed that in the community, there were 11 infections for every COVID-19 confirmed case in Geneva, reflecting the variability in disease severity, testing access or practices, and care-seeking behaviours.

Epidemiology
Case Definition

- A set of criteria that defines a case for a given health condition.
- Can vary over time, and over different places
- Often divided in *suspected*, *probable*, and *confirmed* cases

WHO COVID-19: Case Definitions

Updated in Public health surveillance for COVID-19, 22 July 2022

Suspected case of SARS-CoV-2 infection (3 options)

A A person who meets the clinical **OR** epidemiological criteria:

Clinical criteria:

- acute onset of fever AND cough (ILI)

OR

- acute onset of **ANY THREE OR MORE** of the following signs or symptoms: fever, cough, general weakness/fatigue[1], headache, myalgia, sore throat, coryza, dyspnoea, nausea/diarrhoea/anorexia

OR

Epidemiological criteria [2] :

- contact of a probable or confirmed case, or linked to a **COVID-19 cluster.**[3]

B A patient with **severe acute respiratory illness**
(SARI: acute respiratory infection with history of fever or measured fever of ≥38 °C; and cough; with onset within the last 10 days; and requires hospitalization)

C A person with no clinical signs or symptoms **OR** meeting epidemiologic criteria with a **positive professional-use or self-test** SARS-CoV-2 Antigen-RDT.[4]

[1] Signs separated with slash (/) are to be counted as one sign.

[2] In light of the heightened transmissibility of emerging variants and the high likelihood that any close contact could be infected, epidemiological criteria alone are included in order to qualify asymptomatic contacts for testing, when possible, for the countries with the capacity to adapt more sensitive testing strategies; this is particularly relevant in high-risk populations and settings.

[3] A group of symptomatic individuals linked by time, geographic location and common exposures, containing at least **one NAAT-confirmed** case or at least **two** epidemiologically linked, symptomatic (meeting clinical criteria of Suspect case definition A or B) persons with **positive professional use OR self-test Ag-RDT** (based on ≥97% specificity of test and desired >99.9% probability of at least one positive result being a true positive)

Probable case of SARS-CoV-2 infection (2 options)

A A patient who meets **clinical criteria AND** is a **contact of a probable or confirmed case**, or linked to a **COVID-19 cluster**[3]

B **Death**, not otherwise explained, in an adult with **respiratory distress** preceding death AND who **was a contact of a probable or confirmed case** or linked to a **COVID-19 cluster**[3]

Confirmed case of SARS-CoV-2 infection (2 options)

A A person with a positive **Nucleic Acid Amplification Test (NAAT), regardless of** clinical criteria **OR** epidemiological criteria

B A person meeting clinical criteria **AND/OR** epidemiological criteria (suspect case A) with a **positive professional-use or self- test** SARS-CoV-2 Antigen-RDT.[4]

[4] **Ag RDT antigen-detection rapid diagnostic tests (Ag-RDT) are available for use by trained professionals or for self-testing by individuals:**

- **Professional-use SARS-CoV-2 antigen-RDT :** WHO EUL-approved Ag-RDT, in which sample collection, test performance and result interpretation are done by a trained operator

- **Self-test SARS-CoV-2 antigen-RDT :** WHO EUL-approved Ag-RDT in which sample collection, test performance and result interpretation are done by individuals by themselves.

Note: Clinical and public health judgment should be used to determine the need for further investigation in patients who do not strictly meet the clinical or epidemiological criteria. Surveillance case definitions should not be used as the sole basis for guiding clinical management.

WHO reference number: **WHO/2019-nCoV/Surveillance_Case_Definition/2022.1**

Epidemiology
Case Definition

Suspected case of SARS-CoV-2 infection (3 options)

A A person who meets the clinical **OR** epidemiological criteria:

Clinical criteria:

- acute onset of fever AND cough (ILI)

OR

- acute onset of **ANY THREE OR MORE** of the following signs or symptoms: fever, cough, general weakness/fatigue[1], headache, myalgia, sore throat, coryza, dyspnoea, nausea/diarrhoea/anorexia

OR

Epidemiological criteria [2]:

- contact of a probable or confirmed case, or linked to a **COVID-19 cluster.**[3]

B A patient with **severe acute respiratory illness**
(SARI: acute respiratory infection with history of fever or measured fever of ≥38 °C; and cough; with onset within the last 10 days; and requires hospitalization)

C A person
with no clinical signs or symptoms **OR** meeting epidemiologic criteria
with a **positive professional-use or self-test** SARS-CoV-2 Antigen-RDT.[4]

Epidemiology
Case Definition

Probable case of SARS-CoV-2 infection (2 options)

A A patient who meets **clinical criteria AND** is a **contact of a probable or confirmed case**, or linked to a **COVID-19 cluster**[3]

B **Death**, not otherwise explained, in an adult with **respiratory distress** preceding death AND who **was a contact of a probable or confirmed case** or linked to a **COVID-19 cluster**[3]

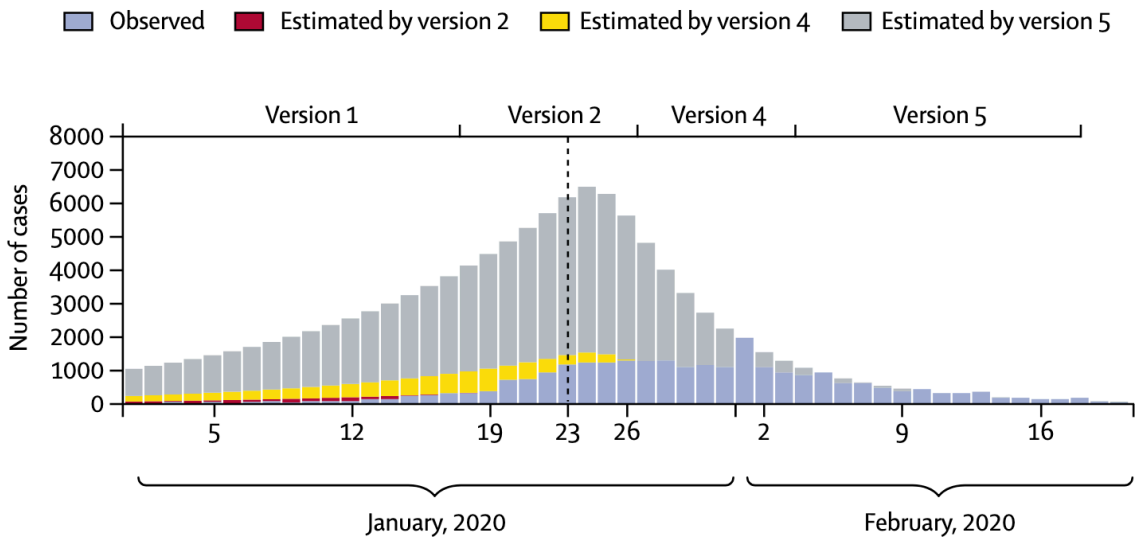
Confirmed case of SARS-CoV-2 infection (2 options)

A A person with a positive **Nucleic Acid Amplification Test (NAAT), regardless of** clinical criteria **OR** epidemiological criteria

B A person meeting clinical criteria **AND/OR** epidemiological criteria (suspect case A) with a **positive professional-use or self- test** SARS-CoV-2 Antigen-RDT.[4]

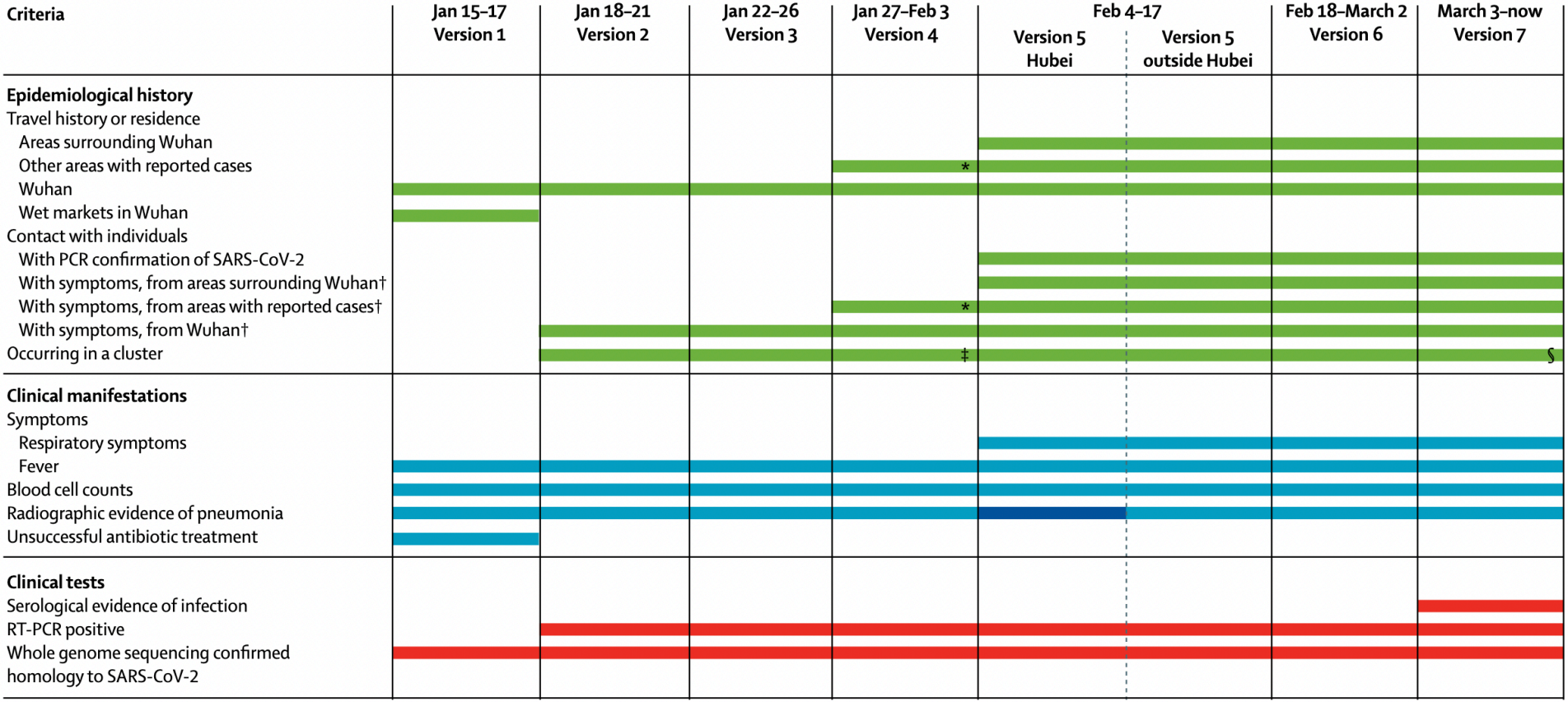
Epidemiology
Case Definition

- Case definitions can change over time, especially during the early phases of newly emerging diseases.

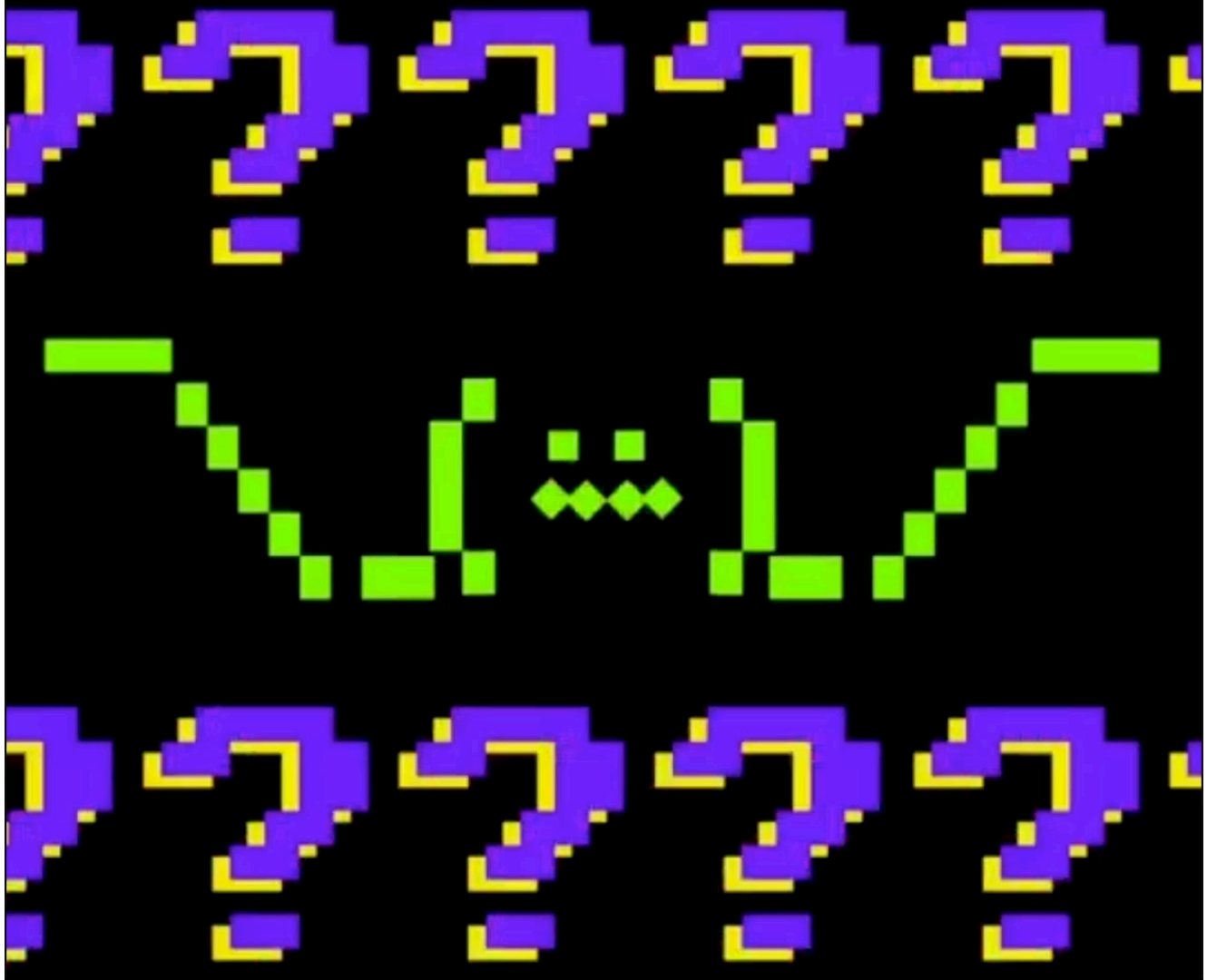


Source: Tsang et al. 2020

Epidemiology



Source: Tsang et al. 2020



Epidemiology
Morbidity

- Disease = reduction in health due to a biological factor (contrast with accidents / non-intentional injury)
- Factor can be external, i.e. pathogen -> infectious disease
- Factor can be internal, i.e cancer -> non-communicable disease (NCD)
- (Note that 10-15% of humans cancers are currently known to be caused by viruses)

Epidemiology
Disease Severity

- Many diseases have a range, from fully asymptotic to deadly.

Epidemiology
Disease Severity

Patient State	Descriptor	Score
Uninfected	Uninfected; no viral RNA detected	0
Ambulatory mild disease	Asymptomatic; viral RNA detected	1
	Symptomatic; independent	2
	Symptomatic; assistance needed	3
Hospitalised: moderate disease	Hospitalised; no oxygen therapy*	4
	Hospitalised; oxygen by mask or nasal prongs	5
Hospitalised: severe diseases	Hospitalised; oxygen by NIV or high flow	6
	Intubation and mechanical ventilation, pO_2/FiO_2 ≥150 or SpO_2/FiO_2 ≥200	7
	Mechanical ventilation pO_2/FIO_2 <150 (SpO_2/FiO_2 <200) or vasopressors	8
	Mechanical ventilation pO_2/FiO_2 <150 and vasopressors, dialysis, or ECMO	9
Dead	Dead	10

Epidemiology
Disease Severity

- Even when severe disease is relatively rare, in absolute numbers, they can overwhelm a healthcare system.

Epidemiology
Disease Severity

- Frequencies of different severities hard to measure - cases visible to healthcare system most likely more severe than average.

Summary of main findings

(1) For all 130 included studies, the IQR for the proportion of asymptomatic SARS-CoV-2 was 14% to 50%, prediction interval 2% to 90%) and for 84 studies based on screening of defined populations, IQR 20% to 65% (prediction interval 4% to 94%). In 46 studies that identified participants through contact tracing of index cases and outbreak investigations, the summary proportion from meta-analysis was 19% (95% CI 15% to 25%, prediction interval 2% to 70%). (2) The risk ratio for the secondary attack rate from asymptomatic compared with symptomatic infections was 0.32 (95% CI 0.16 to 0.64, prediction interval 0.11 to 0.95) and for presymptomatic infections compared with symptomatic infection was 1.00 (95% CI 0.37 to 2.71, prediction interval 0.11 to 9.12). (3) In mathematical modelling studies, estimated proportions of all SARS-CoV-2 infections that result from transmission from asymptomatic individuals were mostly below 15%, and from presymptomatic individuals mostly higher than 40%. Evidence

Source:
Buitrago-Garcia et al. 2022

LETTER

doi:10.1038/nature12060

The global distribution and burden of dengue

Samir Bhatt[1], Peter W. Gething[1], Oliver J. Brady[1,2], Jane P. Messina[1], Andrew W. Farlow[1], Catherine L. Moyes[1], John M. Drake[1,3], John S. Brownstein[4], Anne G. Hoen[5], Osman Sankoh[6,7,8], Monica F. Myers[1], Dylan B. George[9], Thomas Jaenisch[10], G. R. William Wint[1,11], Cameron P. Simmons[12,13], Thomas W. Scott[9,14], Jeremy J. Farrar[12,13,15] & Simon I. Hay[1,9]

Dengue is a systemic viral infection transmitted between humans by *Aedes* mosquitoes[1]. For some patients, dengue is a life-threatening illness[2]. There are currently no licensed vaccines or specific therapeutics, and substantial vector control efforts have not stopped its rapid emergence and global spread[3]. The contemporary worldwide distribution of the risk of dengue virus infection[4] and its public health burden are poorly known[2,5]. Here we undertake an exhaustive assembly of known records of dengue occurrence worldwide, and use a formal modelling framework to map the global distribution of dengue risk. We then pair the resulting risk map with detailed longitudinal information from dengue cohort studies and population surfaces to infer the public health burden of dengue in 2010. We predict dengue to be ubiquitous throughout the tropics, with local spatial variations in risk influenced strongly by rainfall, temperature and the degree of urbanization. Using cartographic approaches, we estimate there to be 390 million (95% credible interval 284–528) dengue infections per year, of which 96 million (67–136) manifest apparently (any level of disease severity). This infection total is more than three times the dengue burden estimate of the World Health Organization[2]. Stratification of our estimates by country allows comparison with national dengue re-

geographical range of endemic transmission[9]. Although the historical expansion of this disease is well documented, the potentially large burden of ill-health attributable to dengue across much of the tropical and subtropical world remains poorly enumerated.

Knowledge of the geographical distribution and burden of dengue is essential for understanding its contribution to global morbidity and mortality burdens, in determining how to allocate optimally the limited resources available for dengue control, and in evaluating the impact of such activities internationally. Additionally, estimates of both apparent and inapparent infection distributions form a key requirement for assessing clinical surveillance and for scoping reliably future vaccine demand and delivery strategies. Previous maps of dengue risk have used various approaches combining historical occurrence records and expert opinion to demarcate areas at endemic risk[10–12]. More sophisticated risk-mapping techniques have also been implemented[13,14], but the empirical evidence base has since been improved, alongside advances in disease modelling approaches. Furthermore, no studies have used a continuous global risk map as the foundation for dengue burden estimation.

The first global estimates of total dengue virus infections were based on an assumed constant annual infection rate among a crude approxi-

Epidemiology
Disease Severity

PLOS PATHOGENS

RESEARCH ARTICLE

Contributions from the silent majority dominate dengue virus transmission

Quirine A. ten Bosch[1¤a]*, Hannah E. Clapham[2¤b], Louis Lambrechts[3,4], Veasna Duong[5], Philippe Buchy[5,6], Benjamin M. Althouse[7,8,9], Alun L. Lloyd[10], Lance A. Waller[11], Amy C. Morrison[12], Uriel Kitron[13], Gonzalo M. Vazquez-Prokopec[13], Thomas W. Scott[12], T. Alex Perkins[1]*

1 Department of Biological Sciences, Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, United States, **2** Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD, United States, **3** Insect-Virus Interactions Group, Department of Genomes and Genetics, Institut Pasteur, Paris, France, **4** Centre National de la Recherche Scientifique, Unité Mixte de Recherche 2000, Paris, France, **5** Virology Unit, Institut Pasteur du Cambodge, Institut Pasteur International Network, Phnom Penh, Cambodia, **6** GlaxoSmithKline, Vaccines R&D, Singapore, **7** Institute for Disease Modeling, Bellevue, WA, United States, **8** Information School, University of Washington, Seattle, WA, United States, **9** Department of Biology, New Mexico State University, Las Cruces, NM, United States, **10** Department of Mathematics, Biomathematics Graduate Program and Center for Quantitative Sciences in Biomedicine, North Carolina State University, Raleigh, NC, United States, **11** Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, United States, **12** Department of Entomology and Nematology, University of California, Davis, CA, United States, **13** Department of Environmental Sciences,

Check for updates

Epidemiology
Disease Severity

of people with asymptomatic infections is 80% (median) that of people with apparent or inapparent symptomatic infections (95% credible interval (CI): 0–146%). Due to their numerical prominence in the infectious reservoir, clinically inapparent infections in total could account for 84% (CI: 82–86%) of DENV transmission. Of infections that ultimately result in any level of symptoms, we estimate that 24% (95% CI: 0–79%) of onward trans-

Epidemiology
Disease Severity

- How to quantify impact? We want to compare disease, and make comparison over time and space.

- DALY = YLD + YLL

Epidemiology
Disease Severity

- How to quantify impact? We want to compare disease, and make comparison over time and space.

DALY

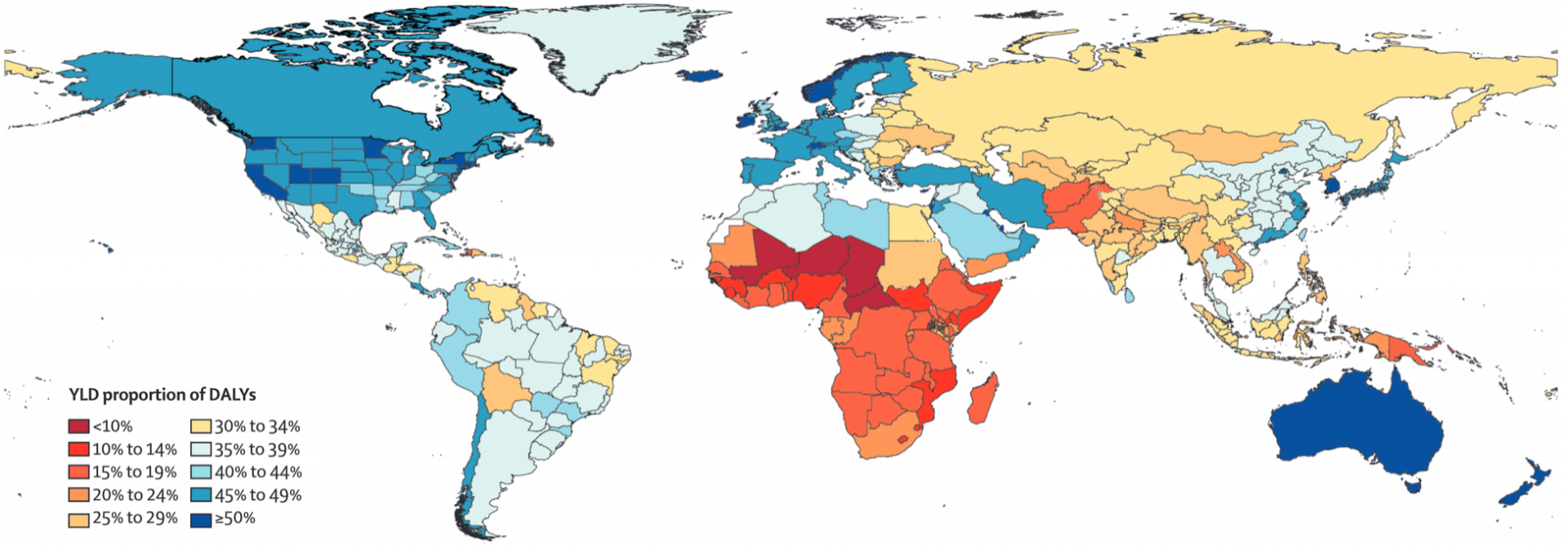
Disability Adjusted Life Year is a measure of overall disease burden, expressed as the cumulative number of years lost due to ill-health, disability or early death

= YLD
Years Lived with Disability

\+ YLL
Years of Life Lost

Source: Wikipedia

Epidemiology
Disease Severity



Source: GBD 2019

Epidemiology
Morbidity

- Comorbidities = presence of more than one health condition
- US: >27% of US adults have chronic comorbidities

Epidemiology
Mortality

- How deadly is a disease?

- How deadly is it in a population -> mortality
- How deadly is it for an individual -> lethality

Epidemiology
Mortality

- Mortality rate: people dying of disease / population of interest, in a specified period of time

Number of deaths for leading causes of death:

- Heart disease: 695,547
- Cancer: 605,213
- COVID-19: 416,893
- Accidents (unintentional injuries): 224,935
- Stroke (cerebrovascular diseases): 162,890
- Chronic lower respiratory diseases: 142,342
- Alzheimer's disease: 119,399
- Diabetes: 103,294
- Chronic liver disease and cirrhosis: 56,585
- Nephritis, nephrotic syndrome, and nephrosis: 54,358

- US Data 2021 (population size 331.9 million)

https://www.cdc.gov/nchs/fastats/deaths.htm

Epidemiology
Mortality

- Swiss Data 2019 (population size 8.575 million)

https://www.bfs.admin.ch/bfs/en/home/statistics/health/state-health/diseases/cancer.html

Cancer

2019	Men	Women
Cancer epidemiology		
Number of new cases[1]	25 302	21 107
Incidence rate[2]	428.5	334.7
Cumulative risk before age 70, as a %	23.9	20.4
Number of deaths from cancer	9 322	7 870

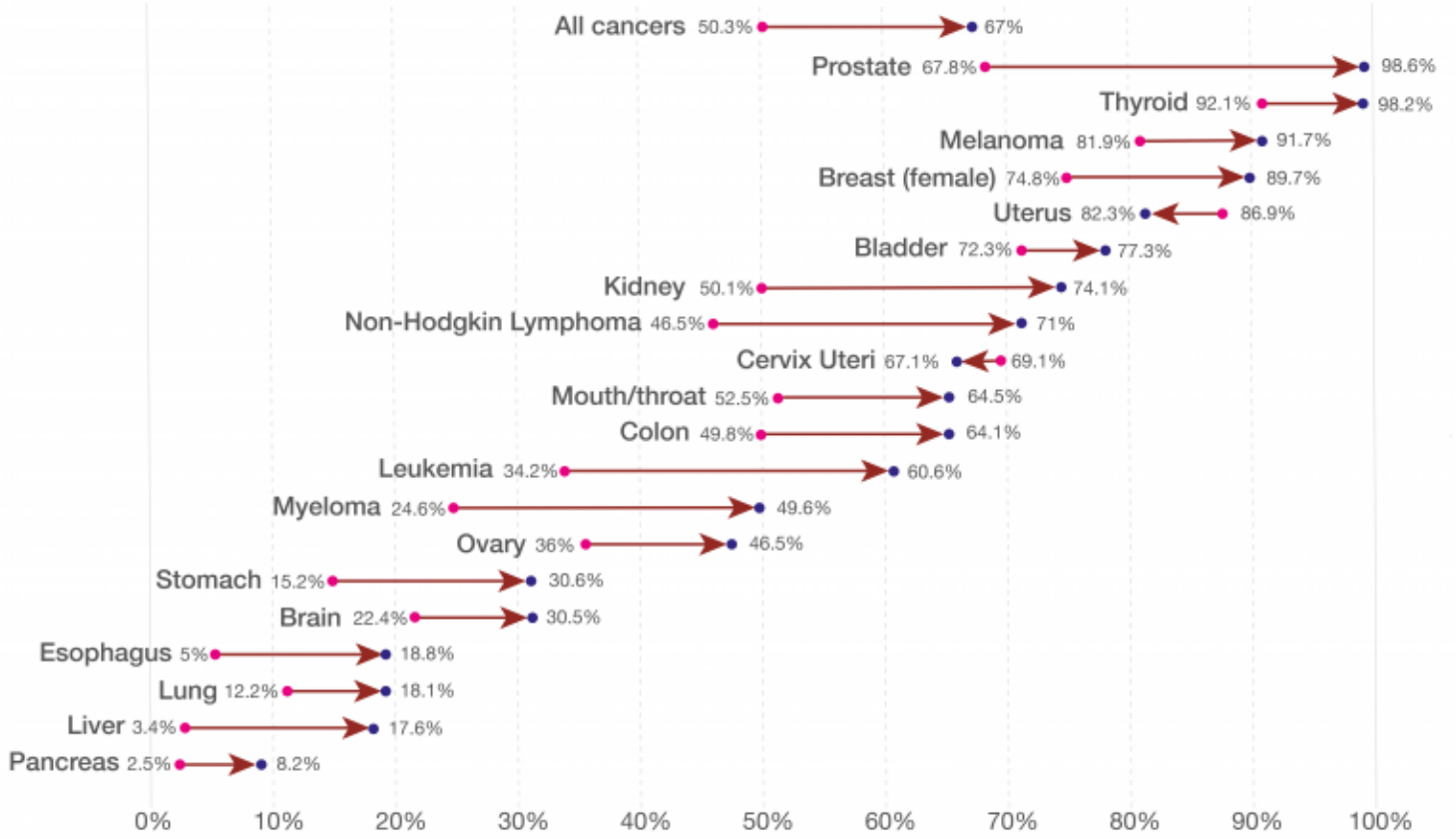
Epidemiology
Lethality

- How deadly is the disease once you have it? For *chronic* diseases, often expressed in survival rates per time period.

Five-year cancer survival rates in the USA

Average five-year survival rates from common cancer types in the United States, shown as the rate over the period 1970-77 [●] and over the period 2007-2013 [●]: 1970-77 → 2007-2013
This five-year interval indicates the percentage of people who live longer than five years following diagnosis.

Our World in Data



Based on data by Journal of the National Cancer Institute; Surveillance, Epidemiology and End Results Program.
The data visualization is available at OurWorldinData.org. There you find research and more visualizations on this topic.

Epidemiology
Lethality

- How deadly is the disease once you have it? For *acute* diseases, expressed in case fatality rate (CFR).
- CFR = number of deaths due to disease / cases

Epidemiology
Lethality

- IFR = number of deaths due to disease / infections
- Since infections often unknown, we need to estimate it from cases
- In early phase of outbreak, we haven’t seen all deaths yet

Epidemiology
Lethality

- IFR of COVID: now known to be 0.5% and 1% (pre-vaccine)
- IFR of influenza: 0.03% - 0.1%

Epidemiology
Lethality

- Malaria: CFR: 0.3%
- Untreated HIV: CFR 99+%
- Huge range of lethalities: from almost non-lethal (certain childhood diseases, CFR < 0.001%), to fully lethal (rabies after onset of symptoms, CFR 100%)