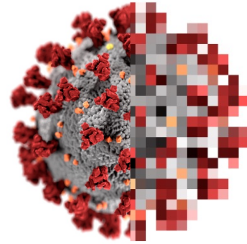


The **Epiverse** initiative



Adam Kucharski

Co-director, Centre for Epidemic Preparedness and Response

Epiverse/Harmonize workshop

July 2023

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Epidemic
Preparedness
& Response

What are we dealing with?

And what can we do about it?

Need to understand disease dynamics: COVID-19

Daily chart

How deadly is the new coronavirus?

Economist, 12th March 2020
Russell et al, Eurosurveillance, 2020

Coronavirus: What is the k number and can superspreading be stopped?

BBC News, 6th June 2020
Endo et al. Wellcome Open Res, 2020

Alpha

Coronavirus Variant Is Indeed More Transmissible, New Study Suggests

New York Times, 23rd Dec 2020
Davies et al. Science, 2021

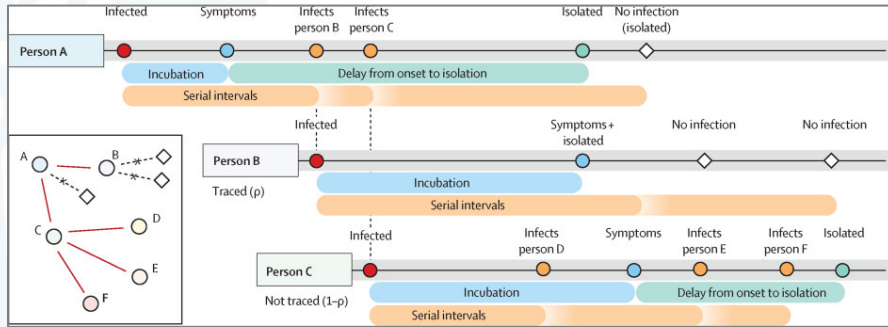
Delta

Is The Variant From India The Most Contagious Coronavirus Mutant On The Planet?

NPR, 14th May 2021
Kucharski et al. SPI-M/SAGE report, 12th May 2021

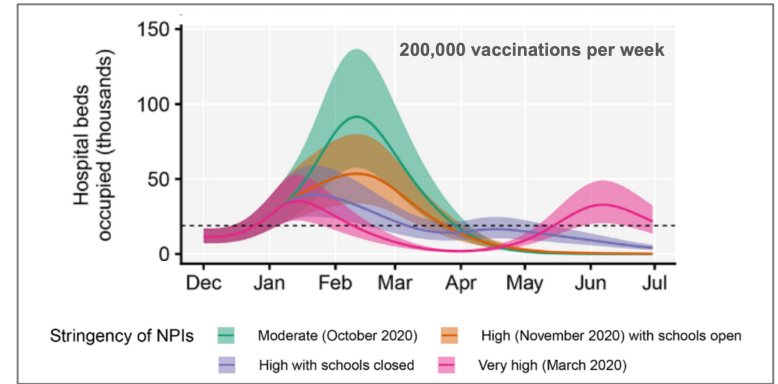
Need to understand control options: COVID-19

Early isolation/contact tracing modelling



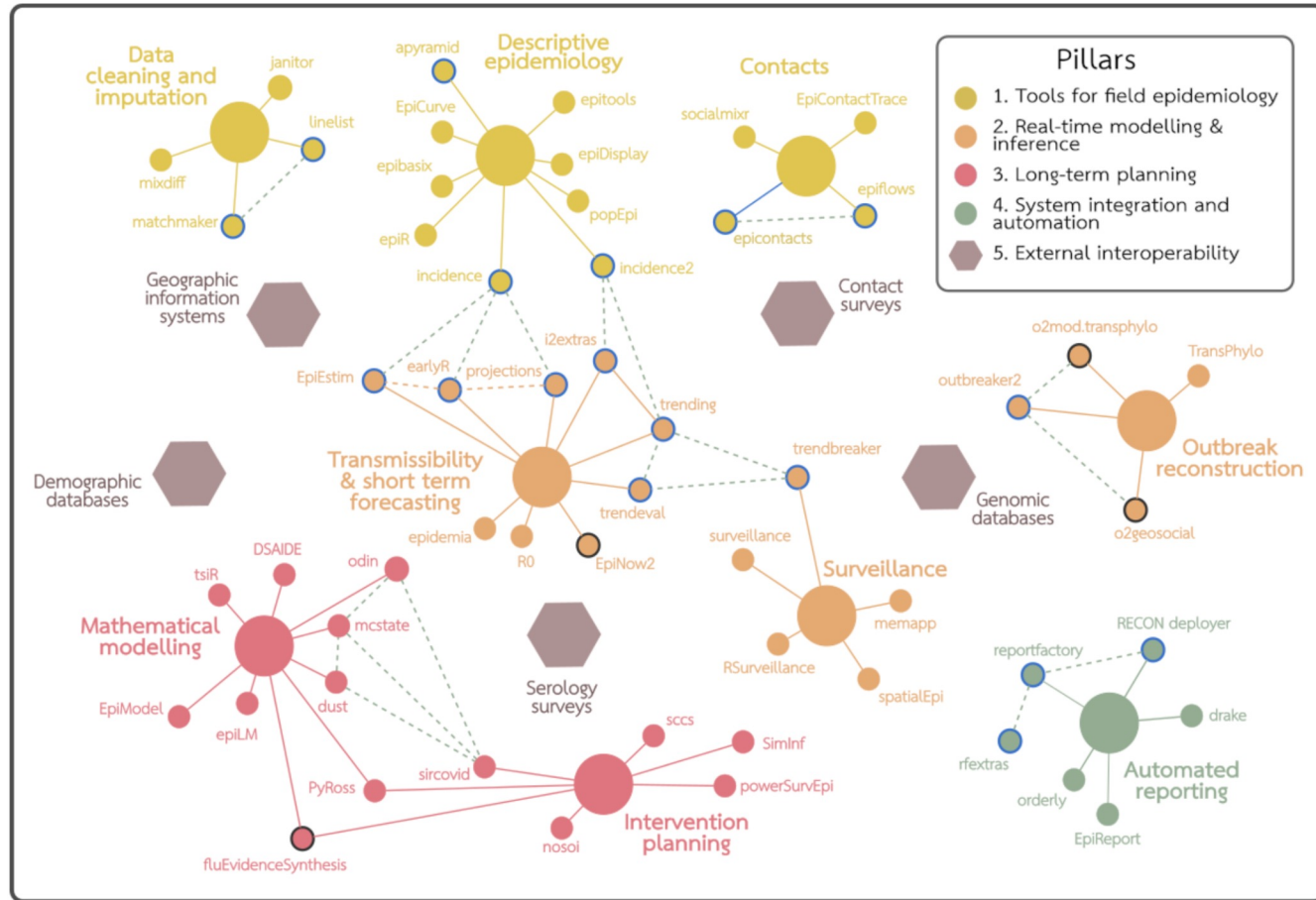
Hellewell et al, *Lancet Global Health*, 2020

Alpha variant scenario modelling



Davies et al. *Science*, 2021

Some important tools and methods, but a fragmented ecosystem



The **Epiverse** initiative

Aim: change how analytics are used in the global infectious disease response, moving from inflexible analytical tools and ad-hoc collaboration to **integrated, generalisable and scalable community-driven software.**

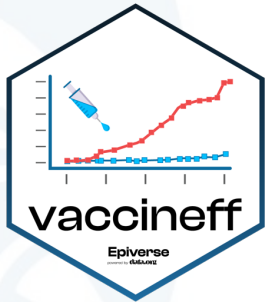


data.org



The **Epiverse** initiative

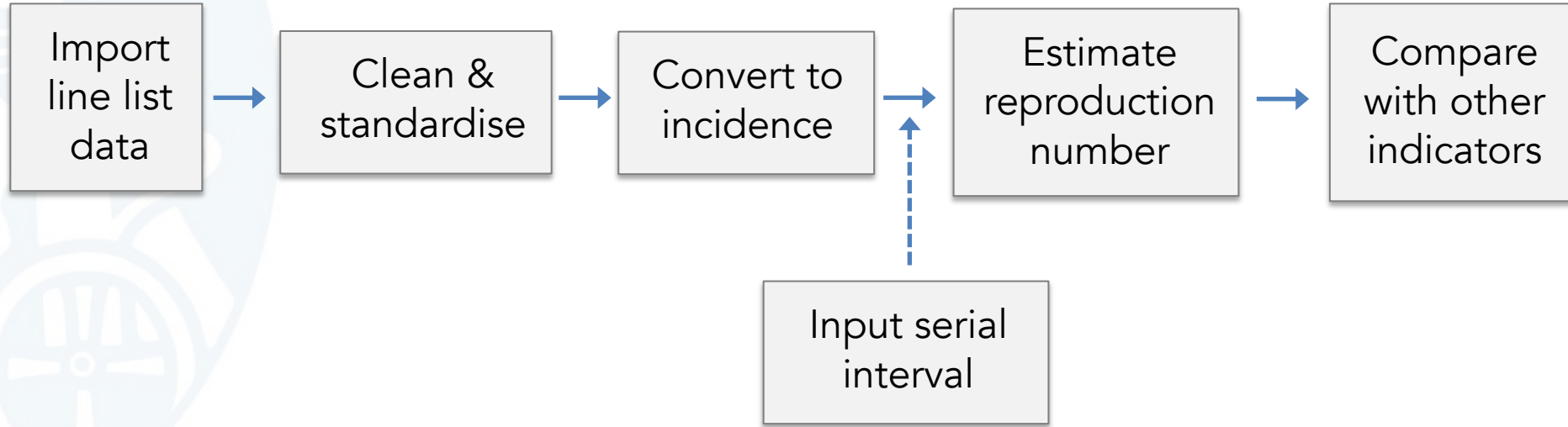
Aim: change how analytics are used in the global infectious disease response, moving from inflexible analytical tools and ad-hoc collaboration to **integrated, generalisable and scalable community-driven software.**



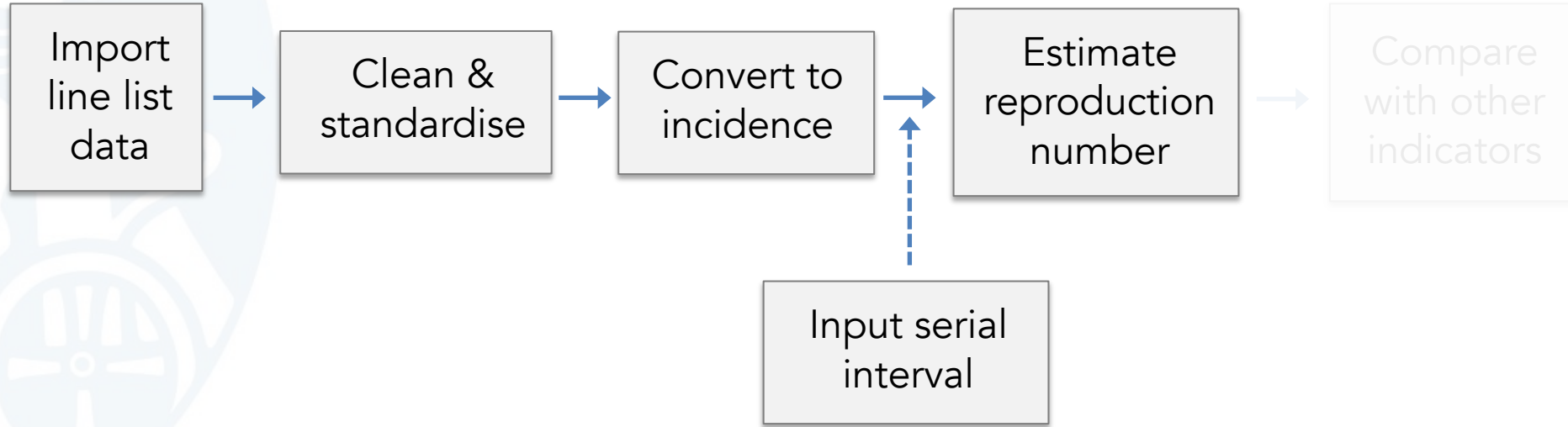
How does transmission vary over time?

And what drives this?

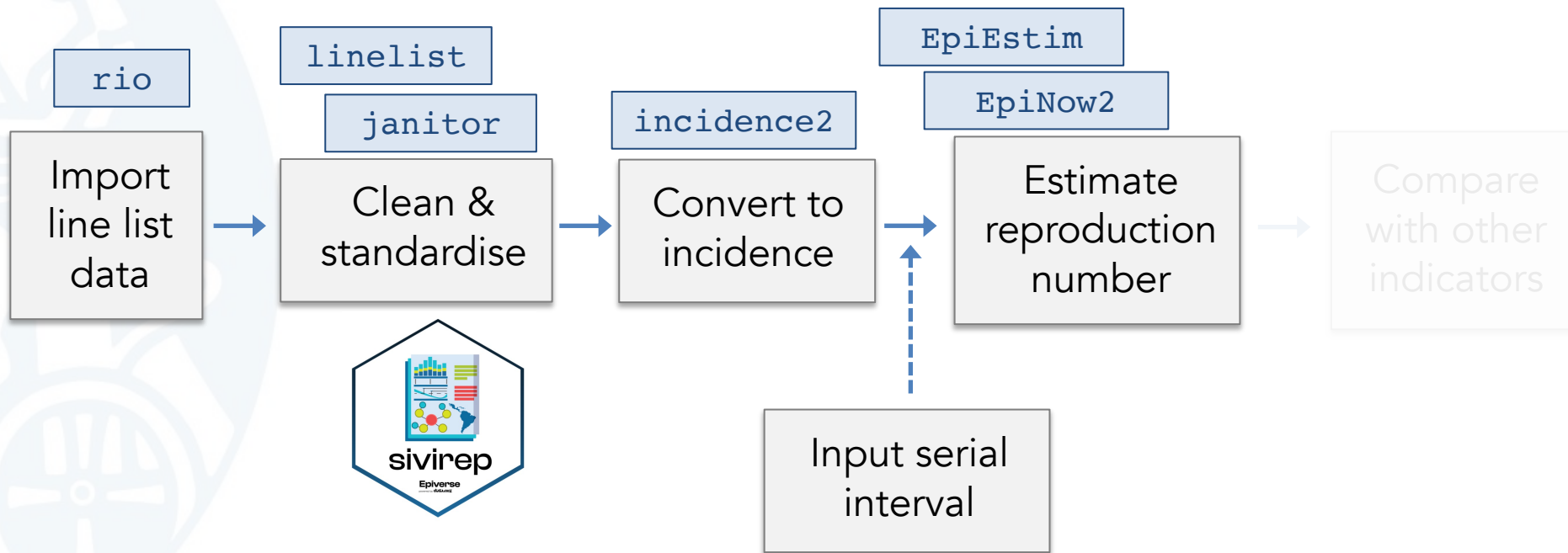
Steps to estimate transmission and compare to other indicators



Estimating reproduction number



Estimating reproduction number



Choose template:

New R Markdown

Document

Presentation

Shiny

From Template

Template: Using R Markdown Templates

- Custom theming {bslib}
- Legacy custom theming {bslib}
- Real-time theming {bslib}
- Transmissibility Report {episoap}
- Flex Dashboard {flexdashboard}
- Flex Dashboard themed with {bslib}
- reprex (lots of features) {reprex}
- reprex (minimal) {reprex}

This template contains multiple files. Create a new directory for these files:

Name: Untitled

Location: ~ Browse...

Create Empty Document

OK Cancel

Outbreak analytics pipelines

Work by Hugo Gruson et al



Select options:

An integer or character indicating the (fixed) size of the time interval used for computing the incidence. Passed as the 'interval' argument in 'incidence2:incidence()'.

week

Number of days to exclude from the estimation of Rt since data is likely to still be incomplete.

7

Number of days to include to get the latest observed value of Rt.

21

Should the serial interval distribution be extracted directly from the epiparameter package?

Name of the pathogen in the epiparameter database if 'use_parameter = TRUE'.

SARS_CoV_2_wildtype

Mean of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.

4.2

Standard deviation of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.

4.9

Choice of probability distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.

gamma

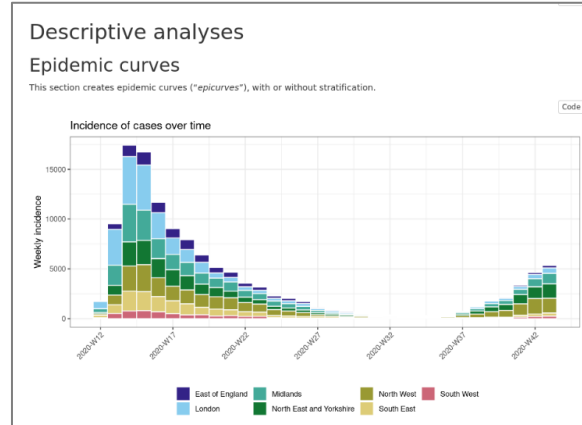
Name of file containing the count data over time (default: data/covid_hosp_uk_20201024.xlsx)

Browse... No file selected

Which R package to use for Rt estimation

- EpiEstim
- EpiNow2
- @centras
- Rt

Descriptive curves:



Choose template:

New R Markdown

Document

Presentation

Shiny

From Template

Template: Using R Markdown Templates

- Custom theming {bslib}
- Legacy custom theming {bslib}
- Real-time theming {bslib}
- Transmissibility Report {episoap}
- Flex Dashboard {flexdashboard}
- Flex Dashboard themed with {bslib}
- reprex (lots of features) {reprex}
- reprex (minimal) {reprex}

This template contains multiple files. Create a new directory for these files:

Name:

Location: Browse...

Create Empty Document

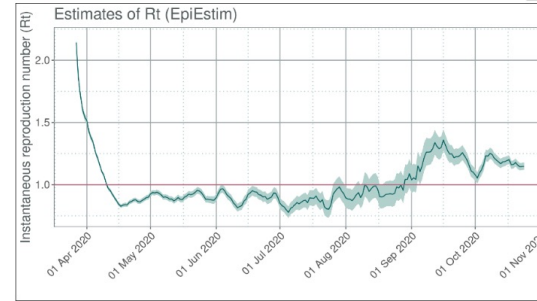
OK Cancel

Outbreak analytics pipelines

Work by Hugo Gruson et al



R estimation:



Select options:

An integer or character indicating the (fixed) size of the time interval used for computing the incidence. Passed as the 'interval' argument in 'incidence2:incidence()'.

week

Number of days to exclude from the estimation of Rt since data is likely to still be incomplete.

7

Number of days to include to get the latest observed value of Rt.

21

Should the serial interval distribution be extracted directly from the epiparameter package?

Name of the pathogen in the epiparameter database if 'use_epiparameter = TRUE'.

SARS_CoV_2_wildtype

Mean of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.

4.2

Standard deviation of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.

4.9

Choice of probability distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.

gamma

Name of file containing the count data over time (default: data/covid_hosp_uk_20201024.xlsx)

Browse... No file selected

Which R package to use for Rt estimation

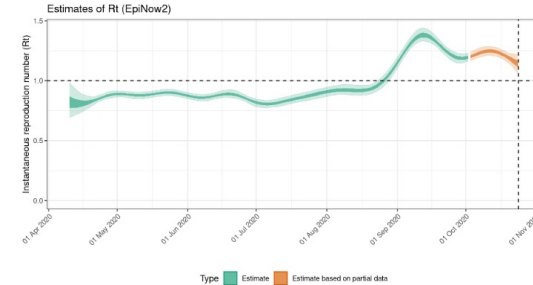
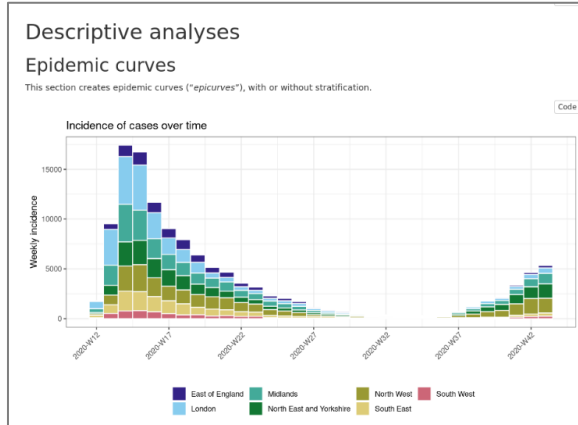
EpiEstim

EpiNow2

@estras

Ri

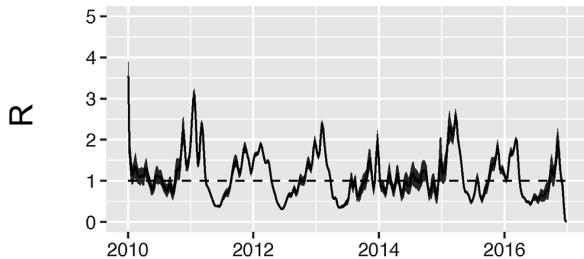
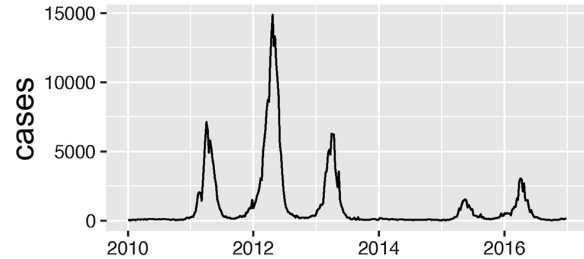
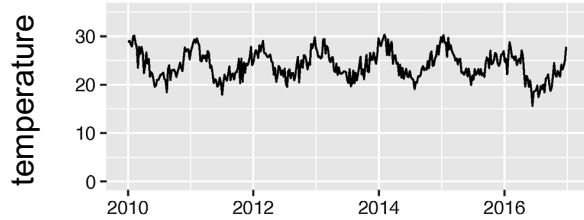
Descriptive curves:



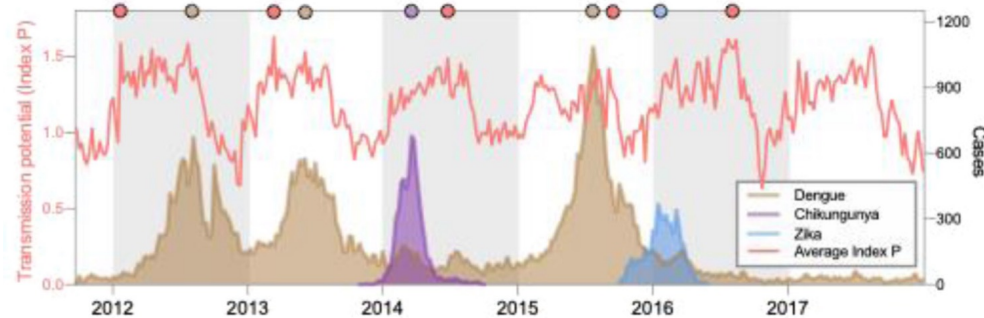
region	mean	median	95% ci
Midlands	1.3	1.3	[1.20 ; 1.40]
East of England	1.2	1.2	[1.10 ; 1.30]
London	1.2	1.2	[0.99 ; 1.30]
North East and Yorkshire	1.2	1.2	[1.10 ; 1.30]
South East	1.2	1.2	[1.10 ; 1.30]
South West	1.1	1.1	[0.95 ; 1.30]
North West	1.0	1.0	[0.90 ; 1.10]

Climate data link: comparison of transmission & other indicators

Reproduction number in Rio de Janeiro



Vector transmission potential vs cases in Dominican Republic



Petrone et al. (2021) *Nature Comms*

Codeço et al. (2018) *Epidemics*

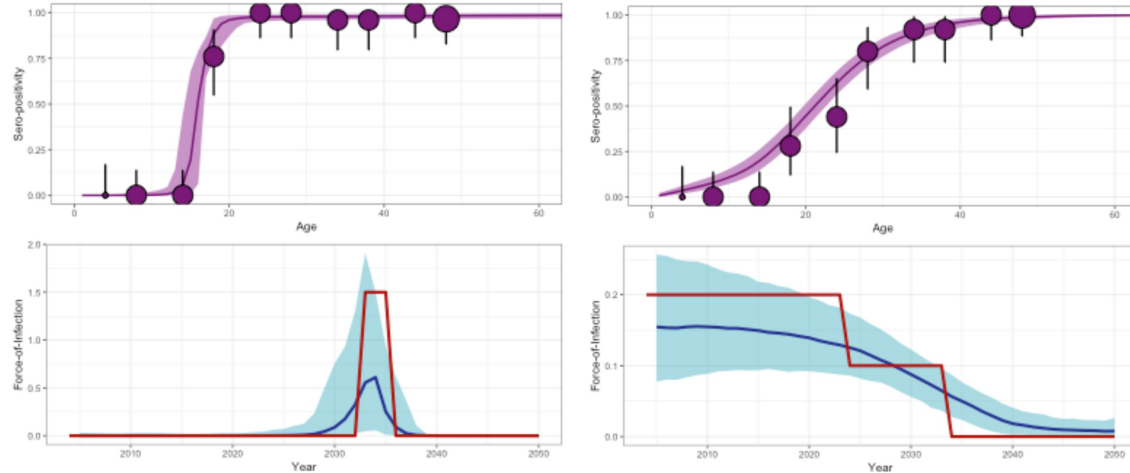
Climate data link: compare with estimates from serological data

- Can also use age-stratified serological data to estimate infection risk over longer period
- Complementary method to using case surveillance data

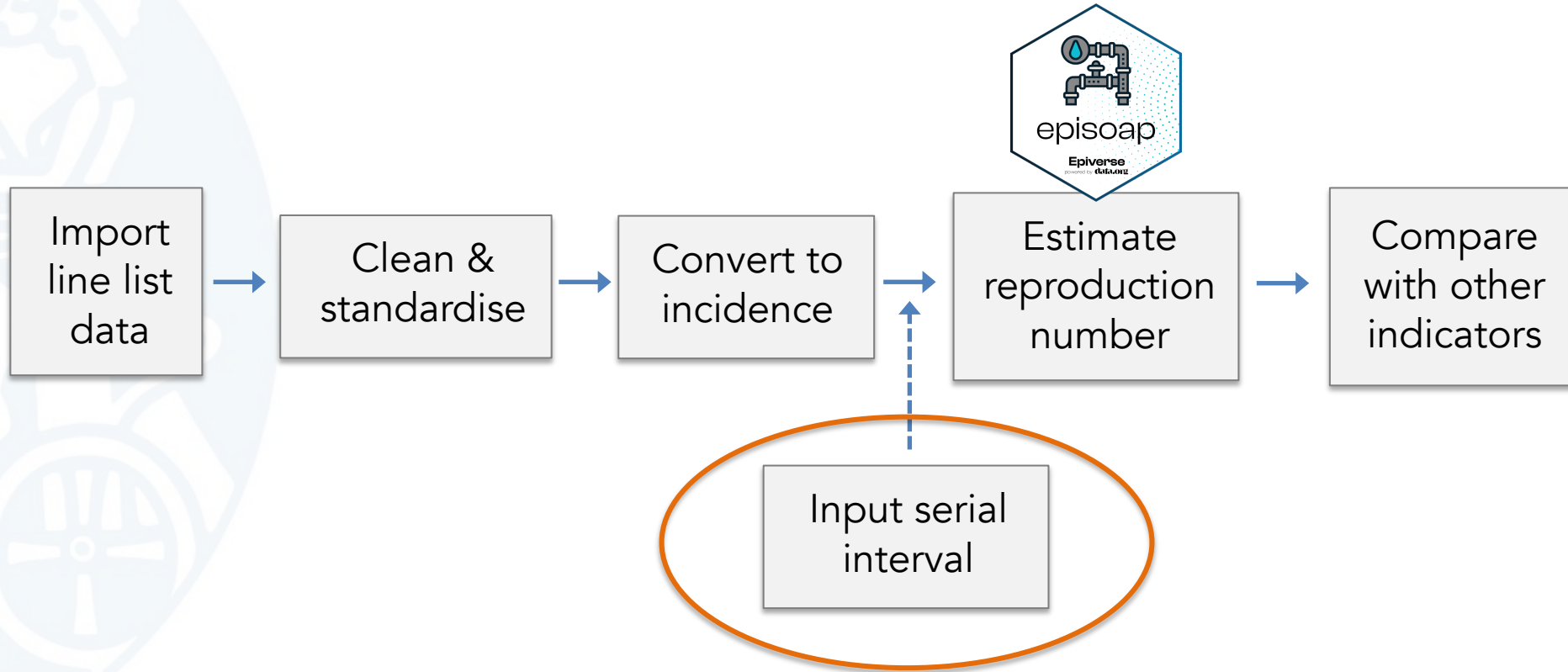



Work by Zulma Cucunubá,
Nicolás Domínguez et al

Simulated antibody positivity by age vs estimated infection risk over time:



Importing parameters





“Fitting a log-normal distribution to the data, we estimated the mean serial interval of COVID-19 to be 4.9 days (95% CI: 4.4–5.7 days)”

Extract, store and reuse epidemiological parameters



Extract distributions from summary statistics:

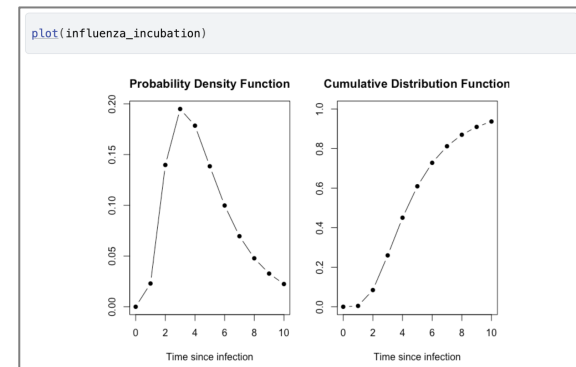
```
convert_gamma_summary_stats(mean = 2, sd = 2)
#> $shape
#> [1] 1
#>
#> $scale
#> [1] 2
```

```
extract_param(
  type = "range",
  values = c(10, 5, 15),
  distribution = "lnorm",
  samples = 25
)
```

Load parameters from library:

```
epidist_db(
  disease = "COVID-19",
  epi_dist = "incubation_period",
  author = "Bui_et_al"
)
#> Using Bui et al. (2020) <10.1371/journal.pone.0243889> PMID: 33362233.
#> To retrieve the short citation use the 'get_citation' function
#> Numerical approximation used, results may be unreliable.
#> Disease: COVID-19
#> Pathogen: SARS-CoV-2
#> Epi Distribution: incubation period
#> Study: Bui et al. (2020) <10.1371/journal.pone.0243889> PMID: 33362233
#> Distribution: weibull
#> Parameters:
#> shape: 2.217
#> scale: 7.226
```

Plot, process and reuse parameters:



Work with Josh Lambert,
Carmen Tamayo Cuartero et al

WHO collaboratory community to generate standardized library

Collaboratory

Pandemic and Epidemic Intelligence

EpiParameter Community

[About](#) [Community](#) [News](#) [Resources](#) [GitHub](#)

About

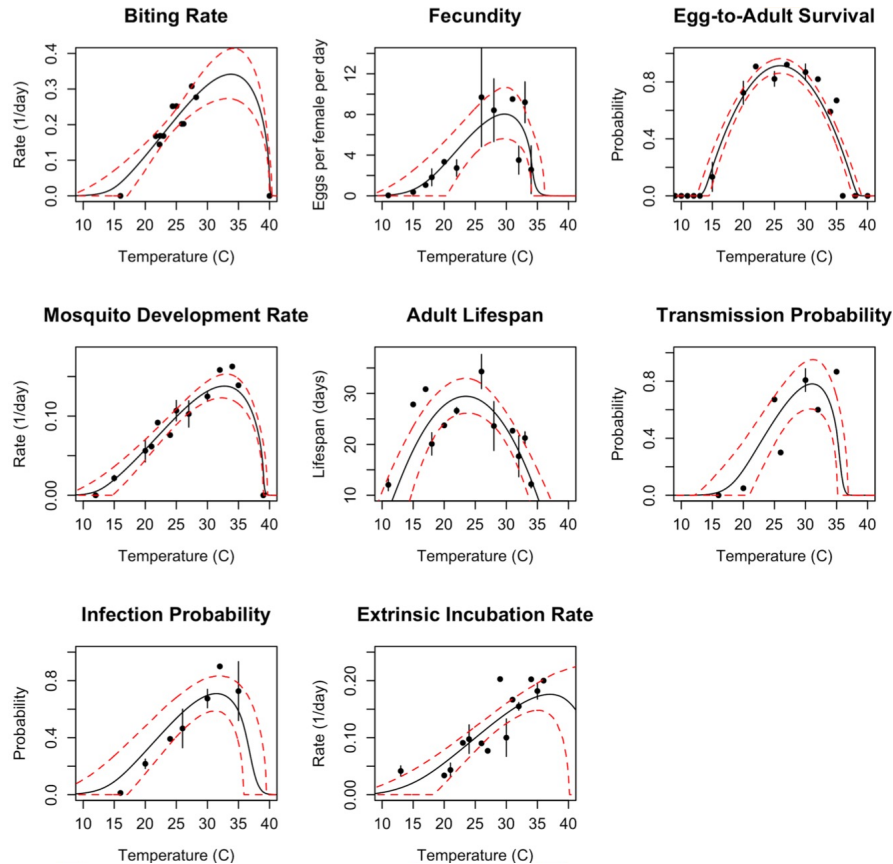
Overview

The epidemiological parameters community consist of a global collaborative working group coordinated by WHO, which aims to develop a global repository of epidemiological parameters. This repository will be publicly accessible by modellers, epidemiologists, subject matter experts and decision makers to inform mathematical models and public health response.

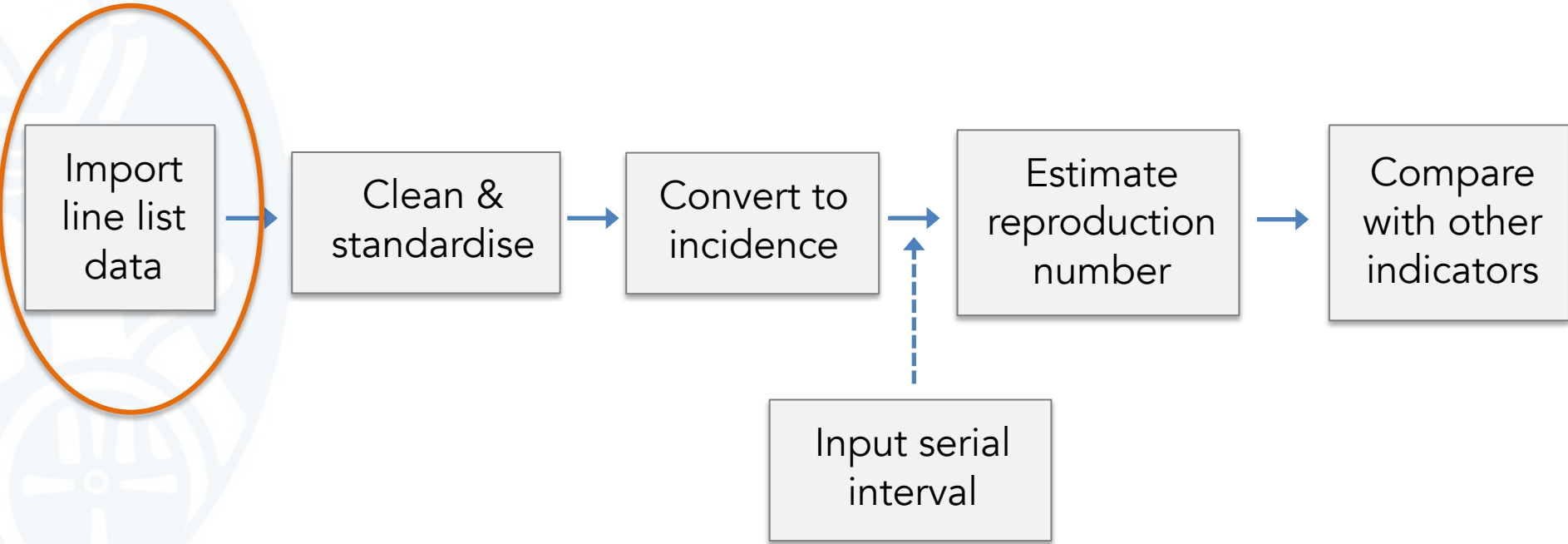
Epidemiological parameters are used by mathematical models that are critical to understand the transmission dynamics of pathogens and to determine the potential impact of outbreaks in terms of morbidity, mortality, and geographical spread over time.

github.com/WorldHealthOrganization/collaboratory-epiparameter-community

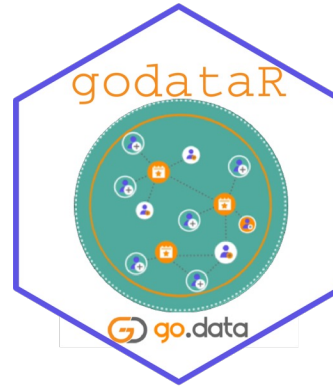
Climate data link: vector-specific parameters



Processing epidemiological data

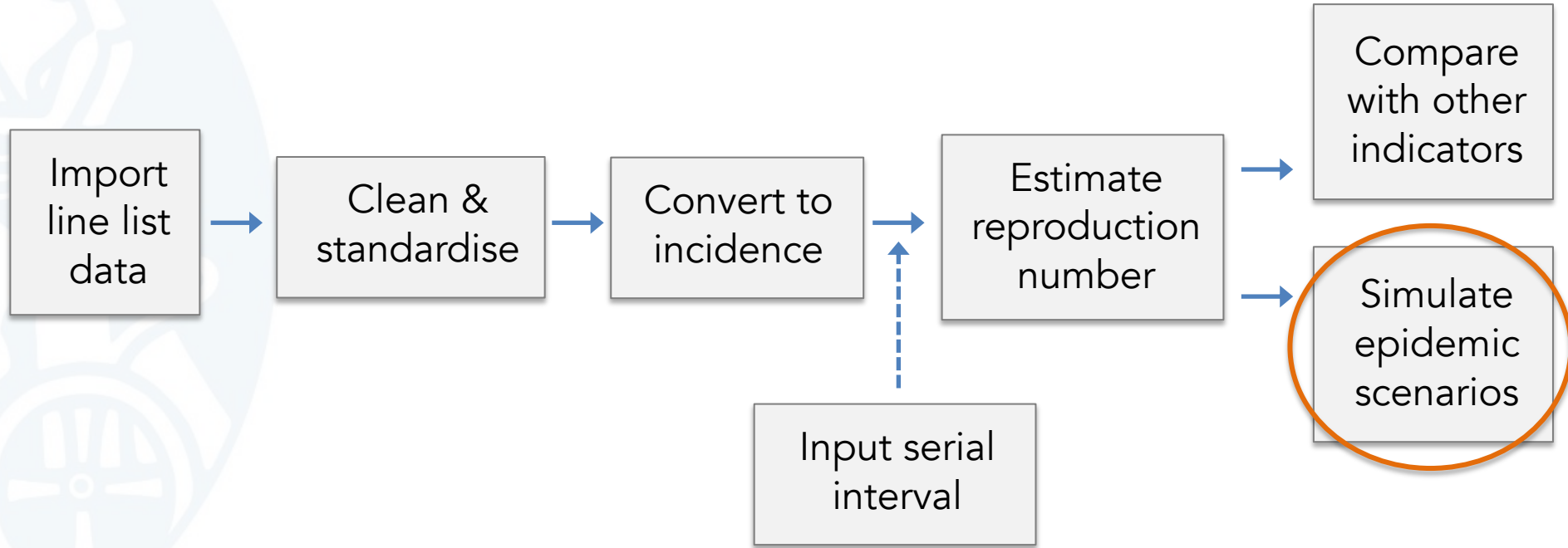


Importing, cleaning and analysing field data



Work with Sara Hollis, James Fuller,
Josh Lambert et al

Processing epidemiological data



Simulate epidemic scenarios



Work with Pratik
Gupte, Roz Eggo et al

Define model options:

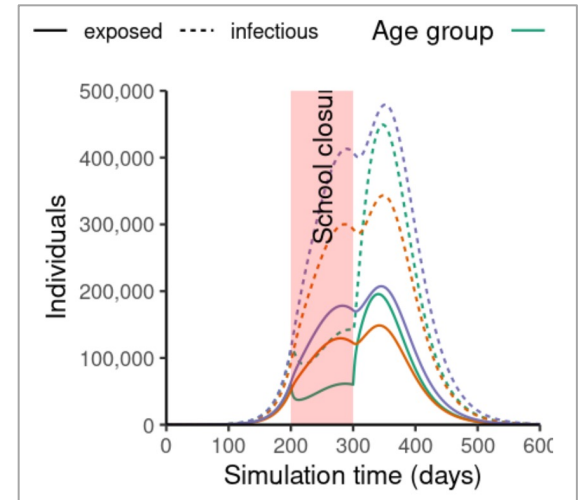
```
# Prepare epidemiological parameters as an infection object
pandemic <- infection(
  r0 = 1.5,
  preinfectious_period = 3,
  infectious_period = 7
)
```

```
# prepare an intervention with a differential effect on age groups
close_schools <- intervention(
  name = "School closure",
  time_begin = 200,
  time_end = 300,
  contact_reduction = matrix(c(0.5, 0.001, 0.001))
)
```

Run model

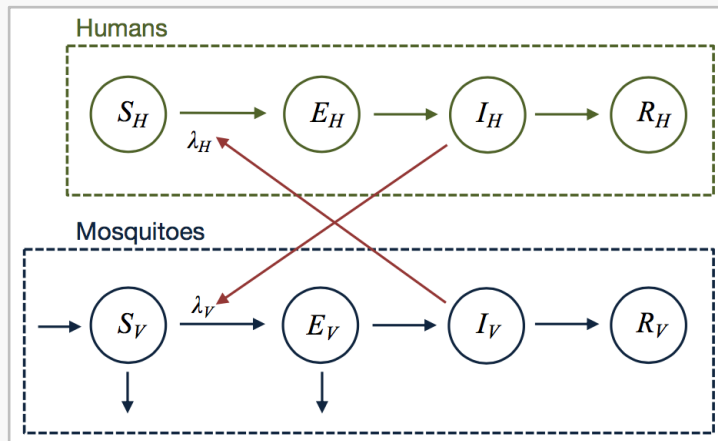
```
# run an epidemic model using `epidemic`
output <- epidemic(
  population = uk_population,
  infection = pandemic,
  intervention = close_schools,
  time_end = 600, increment = 1.0
)
```

Show outputs



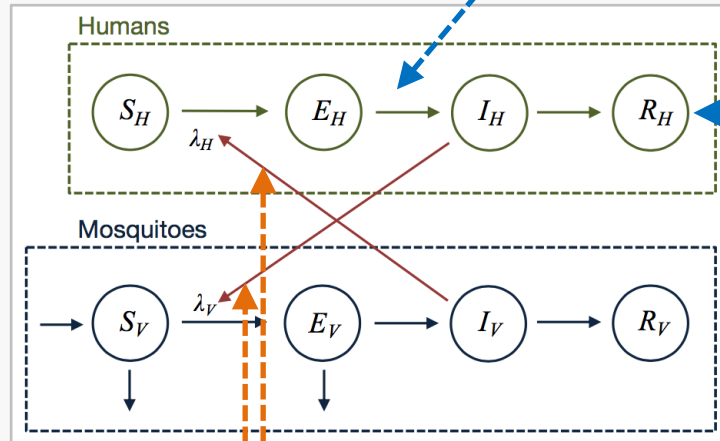
Climate data link: inform transmission rate in models

Vector-borne transmission model

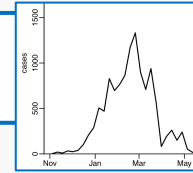


Climate data link: inform transmission rate in models

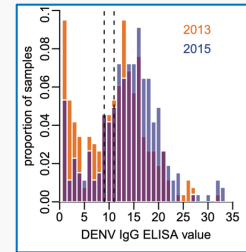
Vector-borne transmission model



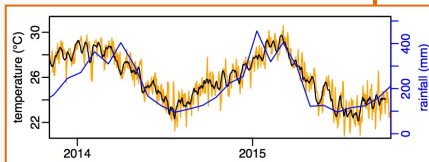
Surveillance data to compare with disease incidence



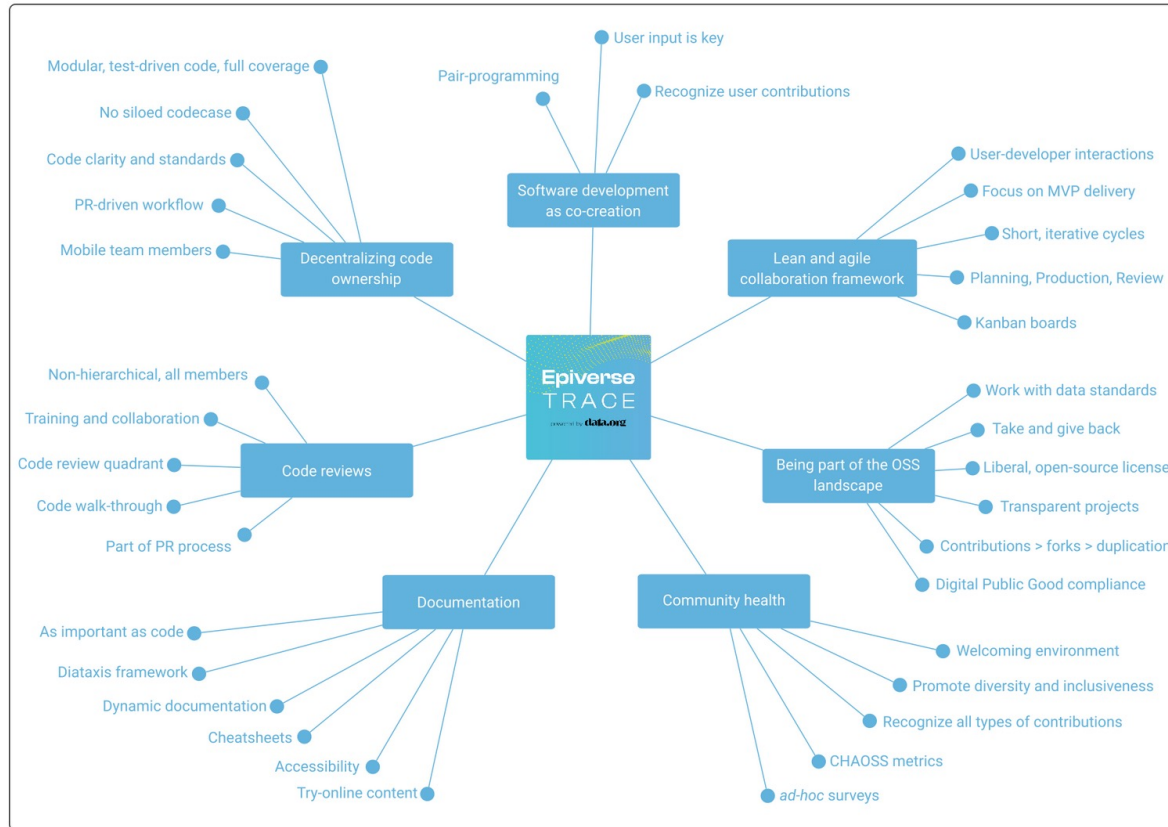
Serological data to compare with immunity



Climate data to inform temporal variation in transmission



Wider activities: community blueprints for best practice



Wider activities: tool showcases and developer blog

Epiverse TRACE: January 2023 showcase

Epiverse TRACE: Spring 2023 showcase




What Should the First 100 Lines of Code Written During an Epidemic Look Like?

Ensuring & Showcasing the Statistical Correctness of your R Package

Convert Your R Function to an S3 Generic: Benefits, Pitfalls & Design Considerations

Improving the C++ Code Quality of an Rcpp Package

Episode 119 of the [#rstats @rweekly_org](#) Highlights podcast is out! [podverse.fm/episode/bvGTUi...](#)

-  Dev containers with R and Quarto [@jimjam_slam](#)
-  Extending data frames [@joshua_lambert](#)
-  Solar system plots with {ggsolar} [@hrbmstr](#)

Summary

- Huge potential to improve the ecosystem for epidemic analytics, from **understanding dynamics** to **exploring control options** .
- Several areas where climate data could be integrated into this ecosystem, including **comparing transmission estimates to common indicators** and **incorporating climate-sensitive parameters** into analysis.

Acknowledgements

Colleagues and collaborators at: the London School of Hygiene & Tropical Medicine, MRC Gambia, data.org, GOARN, WHO Pandemic Hub, Uniandes, Pontificia Universidad Javeriana

data.org

