# Common analytics tasks and available R tools/packages

Dr. James Azam
Research Software Engineer
Epiverse-TRACE Initiative, London School of Hygiene and Tropical Medicine, UK

Presented at WHO Mpox Analytics Collaboratory Seminar
05/09/2024

Context: Mpox as of 01 September

Mpox cases in 2024, Africa
from 01 Jan 2024, as of 01 Sep 2024

World Health Organization

Confidence cases
- 0
- 1-9
- 10-99
- 100-999
- 1000+

Not applicable

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: World Health Organization
Map Production: WHO Health Emergencies Programme
© WHO 2024. All rights reserved.

Summary of Laboratory confirmed mpox cases

Country	Total cases	Total deaths	Case fatality ratio (%)	Cases in 2024[1]	Deaths in 2024[1]	Cases in the past four weeks[2]	Deaths in the past four weeks[2]	Clades detected in country	Date of last report
Democratic Republic of the Congo	4,489	27	1	3,244	25	47	0	Clades Ia and Ib	1 September 2024
Burundi	313	0	0	313	0	273	0	Clade Ib	1 September 2024
Nigeria	909	9	1	48	0	10	0	Clade II (a+b)	25 August 2024
Central African Republic	92	2	2	45	1	2	0	Clade Ia	18 August 2024
Côte d'Ivoire	28	1	4	28	1	22	1	Clade II (a+b)	18 August 2024
South Africa	29	3	10	24	3	0	0	Clade II (a+b)	18 August 2024

As of 01 Sep 2024

Context (cont'd)

Epiverse
powered by data.org

News ⌄   Use ⌄   Learn ⌄   Contribute ⌄   About

2024 mpox outbreak: common analytics tasks and available R tools

EPIVERSE-TRACE   MPOX   OUTBREAK   OUTBREAK-ANALYTICS   DOI

AUTHORS
James Azam
Hugo Gruson
Adam Kucharski

PUBLISHED
July 4, 2024

DOI
10.59350/z78kb-qrz59

Source: https://epiverse-trace.github.io/posts/mpox-preparedness/

The epidemiological modelling R tools ecosystem

CRAN Task View: Epidemiology

|	
Maintainer:	Thibaut Jombart, Matthieu Rolland, Hugo Gruson
Contact:	hugo.gruson+ctv at normalesup.org
Version:	2024-06-20
URL:	https://CRAN.R-project.org/view=Epidemiology
Source:	https://github.com/cran-task-views/Epidemiology/
Contributions:	Suggestions and improvements for this task view are very welcome and can be made through issues or pull requests on GitHub or via e-mail to the maintainer address. For further details see the Contributing guide.
Citation:	Thibaut Jombart, Matthieu Rolland, Hugo Gruson (2024). CRAN Task View: Epidemiology. Version 2024-06-20. URL https://CRAN.R-project.org/view=Epidemiology.
Installation:	The packages from this task view can be installed automatically using the ctv package. For example, ctv::install.views("Epidemiology", coreOnly = TRUE) installs all the core packages or ctv::update.views("Epidemiology") installs all packages that are not yet installed and up-to-date. See the CRAN Task View Initiative for more details.
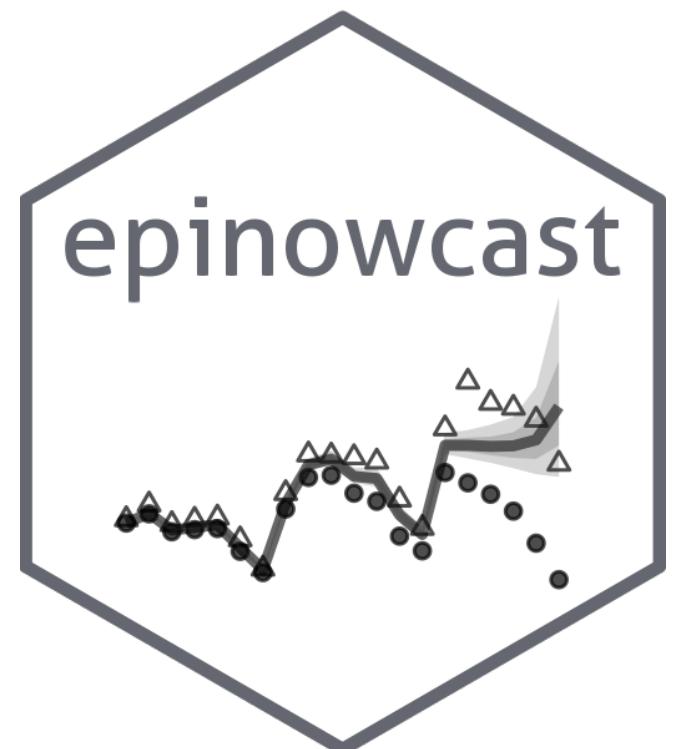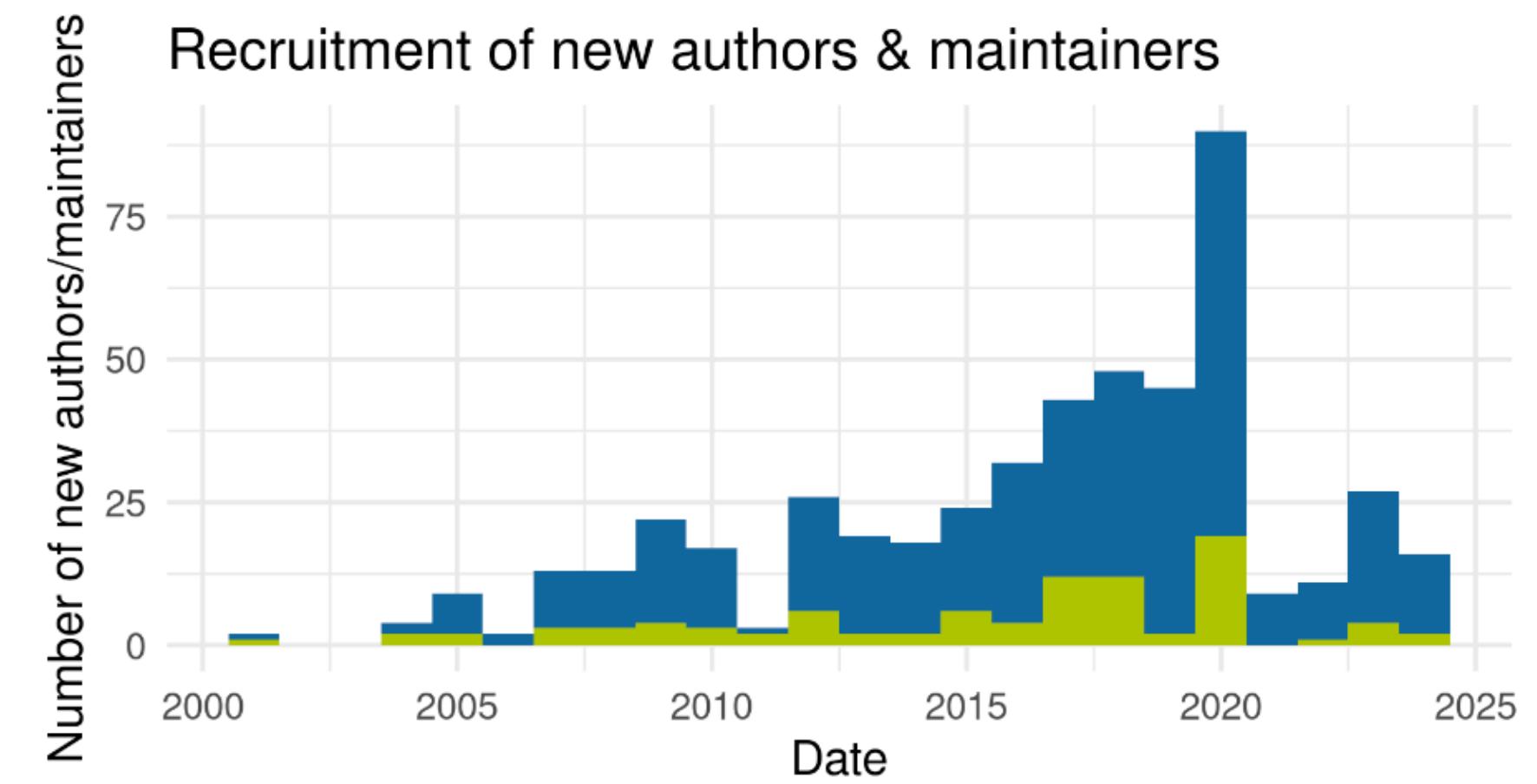
Contributors (in alphabetic order): Neale Batra, Solène Cadiou, Dylan Dijk, Christopher Endres, Rich FitzJohn, Hugo Gruson, Andreas Handel, Michael Höhle, Thibaut Jombart, Joseph Larmarange, Sebastian Lequime, Alex Spina, Tim Taylor, Sean Wu, Achim Zeileis.

https://cran.r-project.org/web/views/Epidemiology.html

Recruitment of new authors & maintainers

https://epiverse-connect.github.io/ctv-analysis/

4

What can we anticipate in terms of analytics with emerging mpox data?
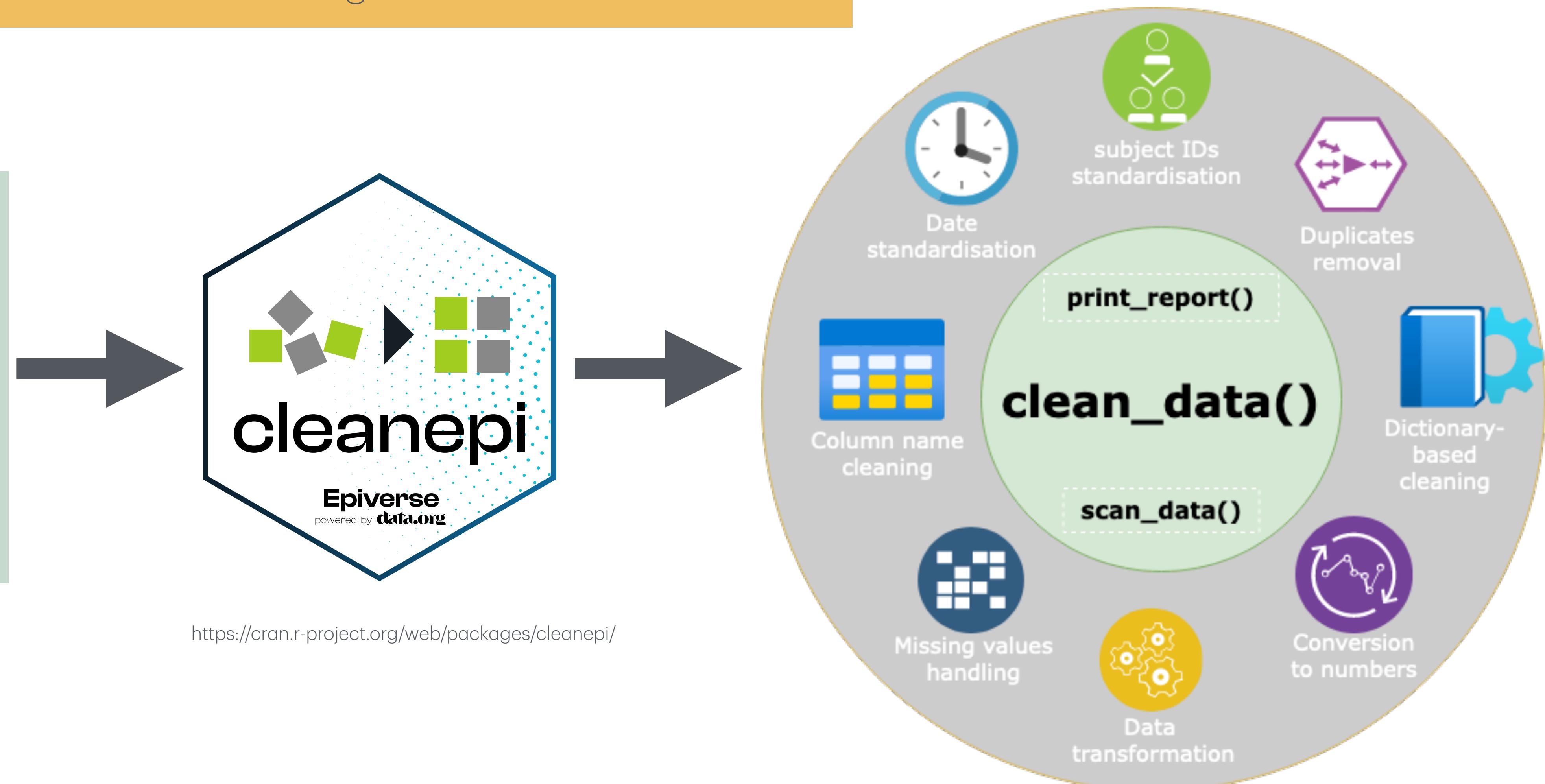
Summary of tasks

- Task 1: Cleaning data

- Task 2: Estimating transmissibility

- Task 3: Getting epidemiological parameter distributions

    - Task 4: Estimating epidemiological delays

    - Task 5: Nowcasting and forecasting infections

- Task 6: Estimating severity

- Task 7: Analysing transmission chains and contact networks

- Task 8: Analysing long term dynamics & scenarios
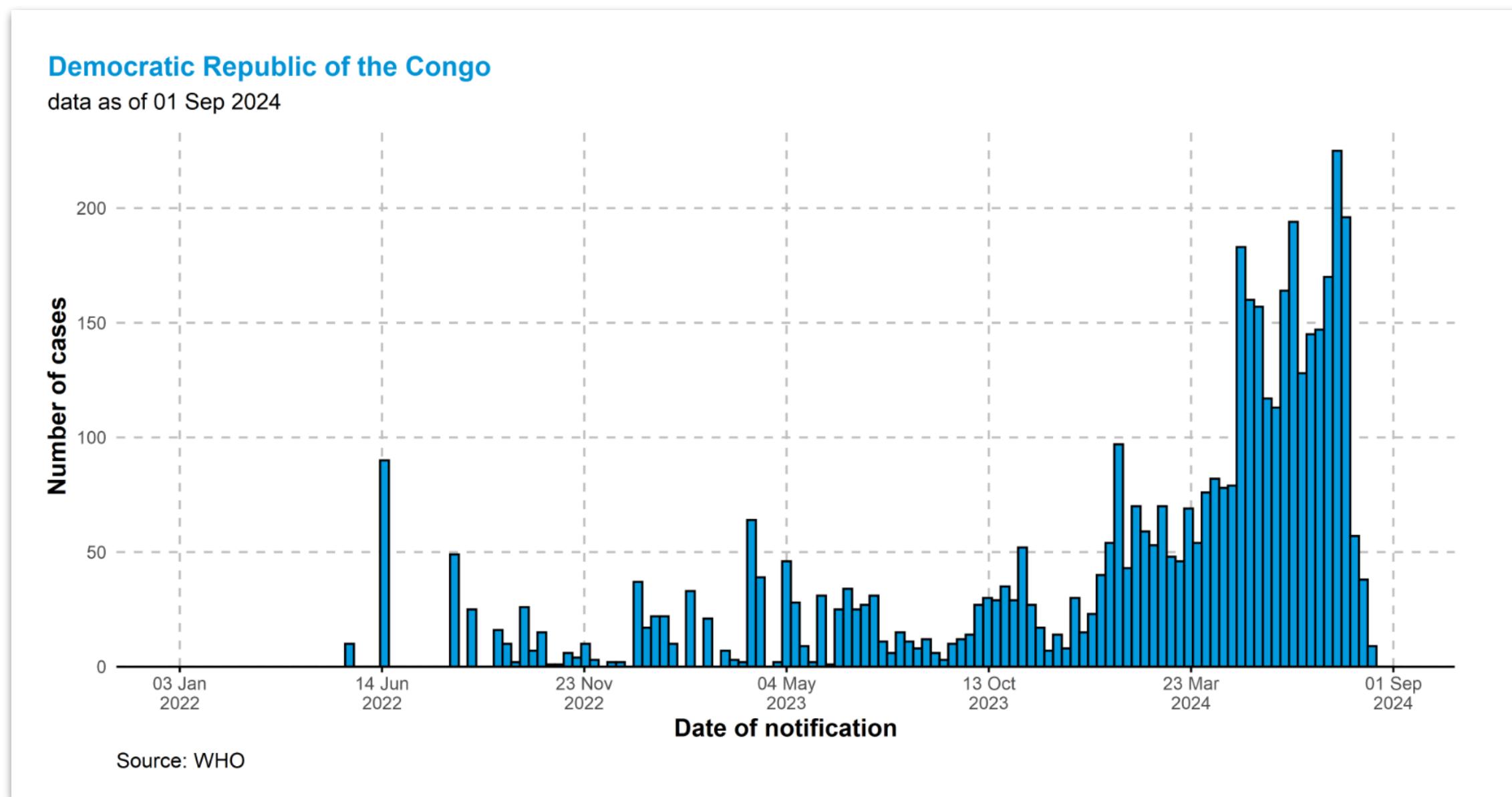
Task 1: data cleaning

An R package for modular data cleaning
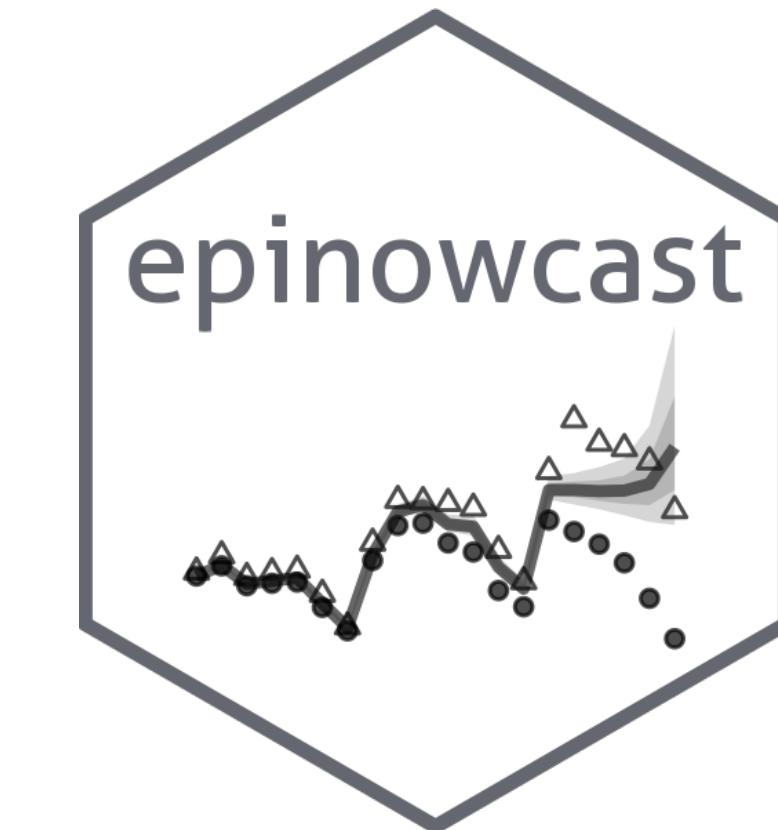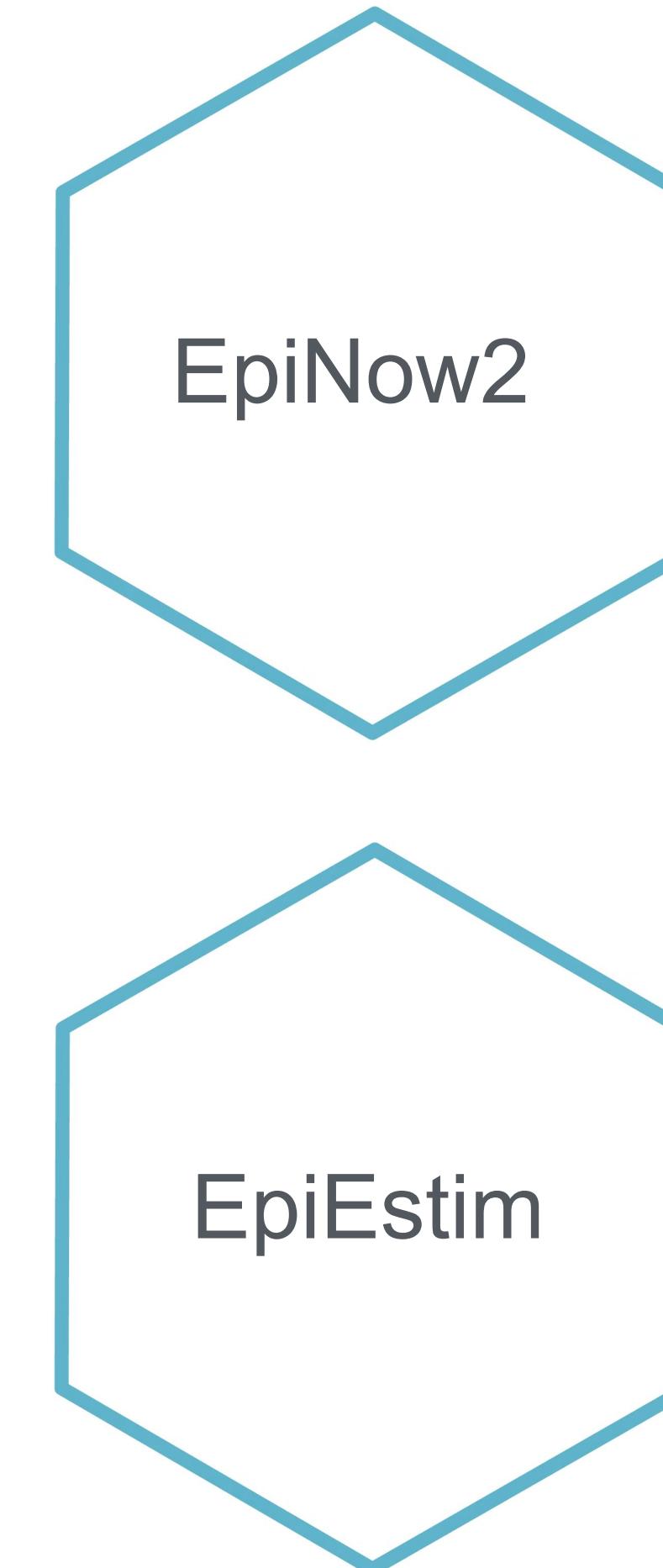
Engine

- janitor

- linelist

- lubridate

- matchmaker

- numberize

- snakecase

https://cran.r-project.org/web/packages/cleanepi/

Task 2: estimating transmissibility (Rt)

Source: https://worldhealthorg.shinyapps.io/mpx_global/

Task 3: Get epidemiological parameters

epiparameter

{epiparameter} is an R package that contains a library of epidemiological parameters for infectious diseases as well as classes and helper functions to work with the data. It also includes functions to extract and convert parameters from reported summary statistics.

Current database

Source: vignettes/database.Rmd

```
## Returning 122 results that match the criteria (99 are parameteri
## Use subset to filter by entry variables or single_epiparameter t
## To retrieve the citation for each use the 'get_citation' functio
```

Show 50 entries   Search: mpox

| Disease	Pathogen	Distribution	Reference
Mpox	Monkeypox Virus	offspring distribution	Lloyd–Smith et al. (2005), DOI: 10.1038/nature04153
Mpox	Mpox Virus	incubation period	Miura et al. (2022), DOI: 10.2807/1560-7917.ES.2022.27.24.2200448
Mpox	Mpox Virus	incubation period	Charniga et al. (2022), DOI: 10.1101/2022.06.22.22276713
Mpox	Mpox Virus	incubation period	Guzzetta et al. (2022), DOI: 10.3201/ei-d2810.221126
Mpox	Mpox Virus	incubation period	Madewell et al. (2022), DOI: 10.1101/2022.10.26.22281516
Mpox	Mpox Virus	incubation period	Madewell et al. (2022), DOI: 10.1101/2022.10.26.22281516
Mpox	Mpox Virus	serial interval	Madewell et al. (2022), DOI: 10.1101/2022.10.26.22281516
Mpox	Mpox Virus	serial interval	Madewell et al. (2022), DOI: 10.1101/2022.10.26.22281516
Mpox	Mpox Virus	serial interval	Guo et al. (2022), DOI: 10.1002/jmv.28248
Mpox	Mpox Virus	serial interval	Wang et al. (2022), DOI: 10.1093/jtm/taac105

Collaboratory
Pandemic and Epidemic Intelligence

EpiParameter Community

About   Community   News   Resources   GitHub

About

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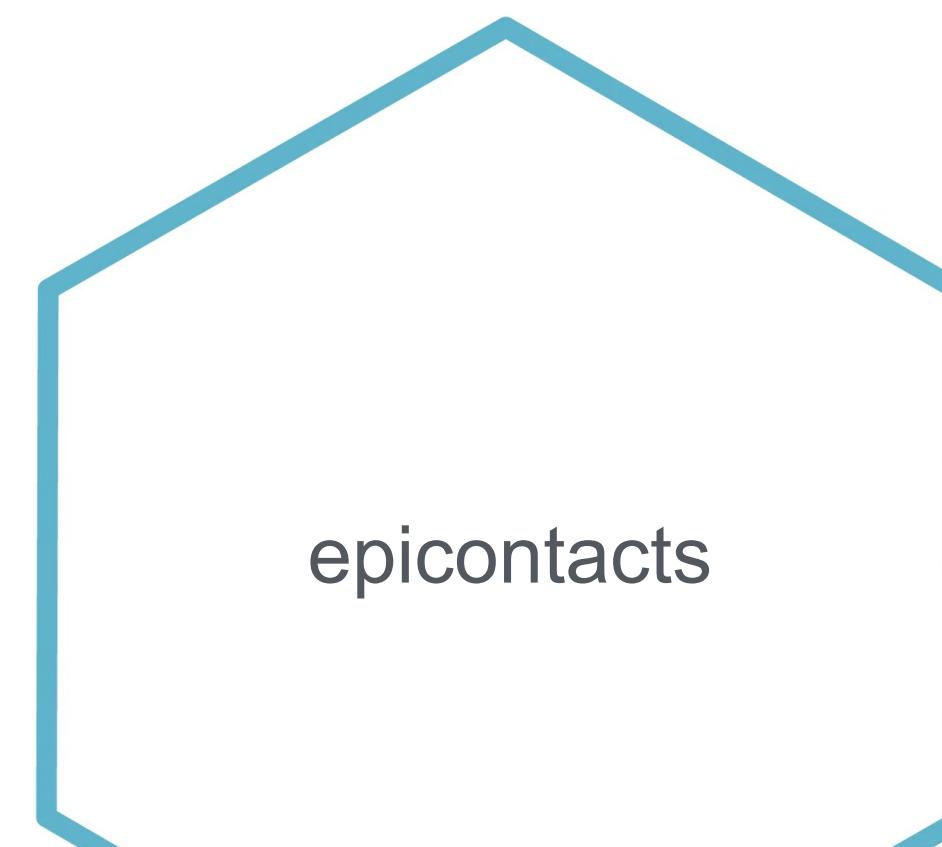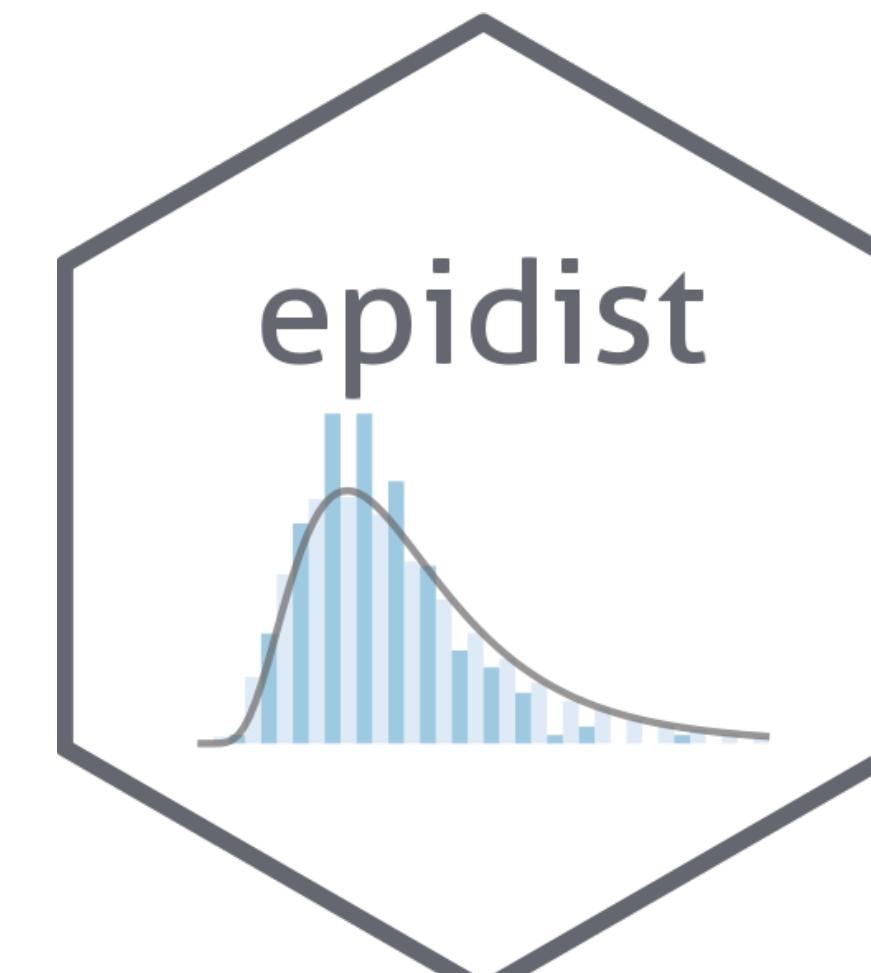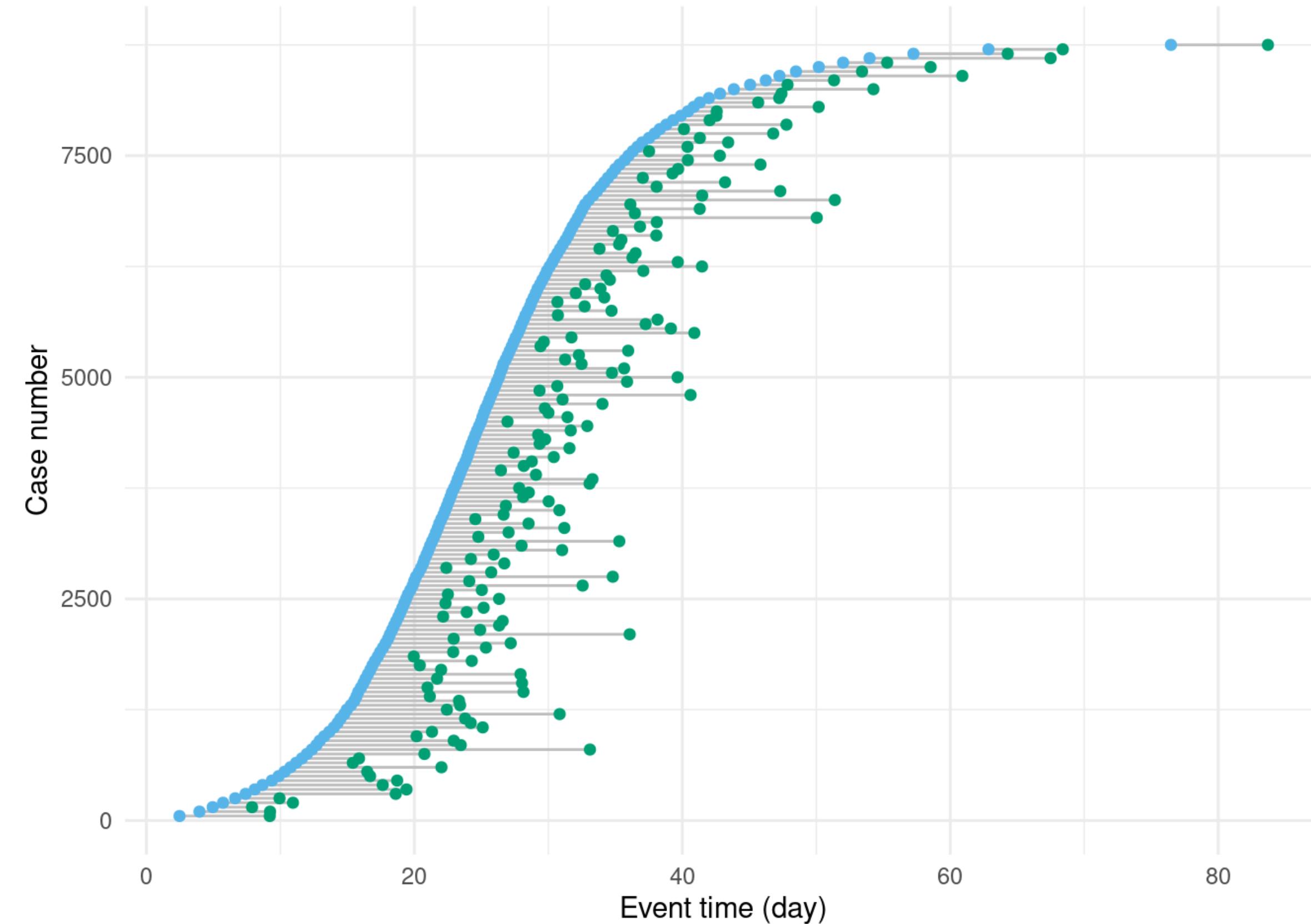
The epidemiological parameters community consist of a global collaborative working group coordinated by WHO, which aims to develop a global repository of epidemiological parameters. This repository will be publicly accessible by modellers, epidemiologists, subject matter experts and decision makers to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

Therefore, by enabling faster and more transparent insight generation at the beginning and during an outbreak, the global epidemiological parameter repository will serve as an essential global public good to inform and guide public health interventions designed to mitigate the spread of diseases and reduce their impact on affected populations.

https://who-collaboratory.github.io/collaboratory-epiparameter-community/#/

https://epiverse-trace.github.io/epiparameter/
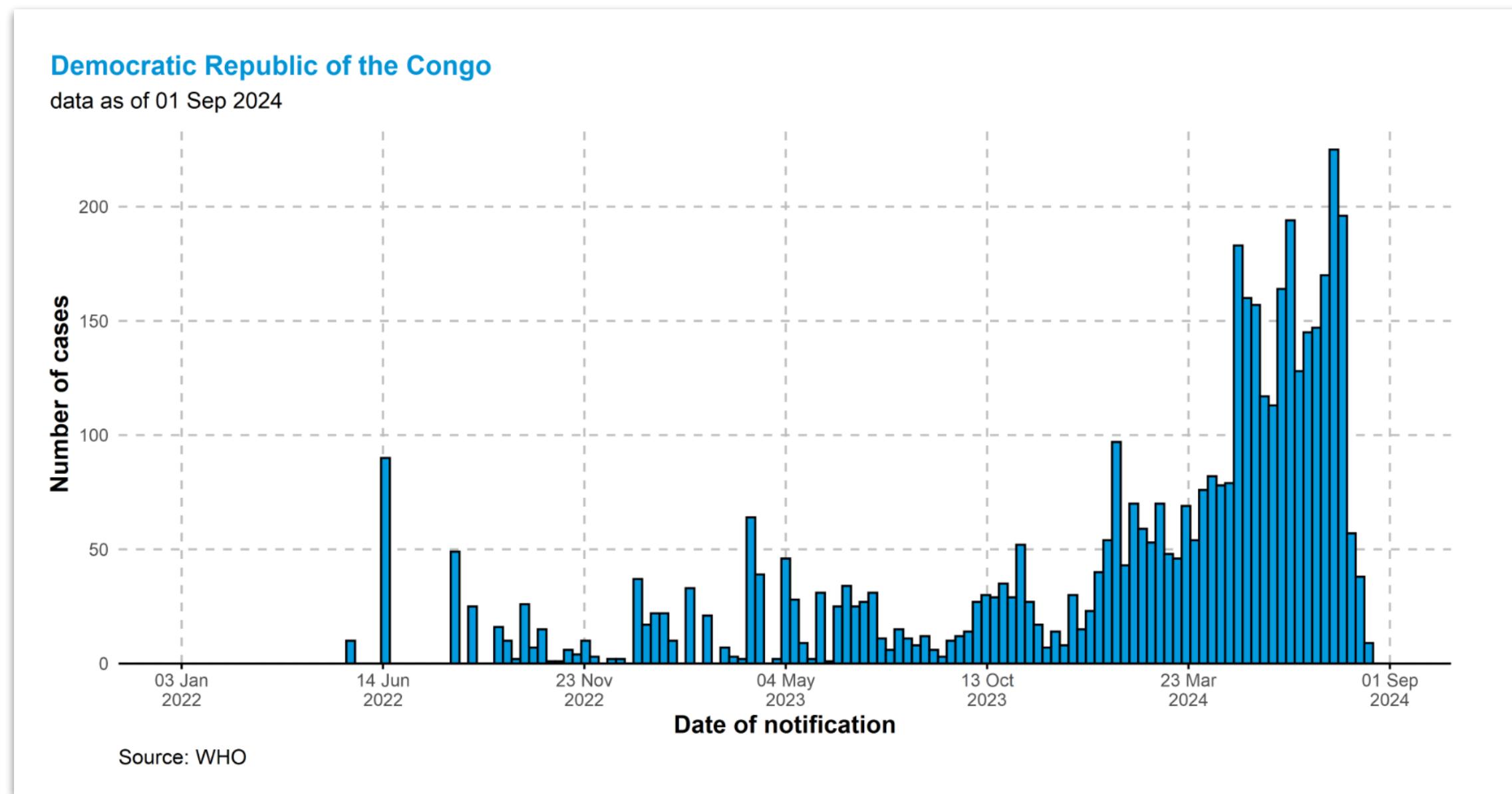
Task 4: estimate delays

A hypothetical linelist showing the time between primary and secondary event.
Source: epidist Getting Started vignette

epidist

EpiNow2

epicontacts

Task 5: nowcast & forecast mpox infections

Democratic Republic of the Congo
data as of 01 Sep 2024

Source: WHO

Source: https://worldhealthorg.shinyapps.io/mpx_global/

Model → Evaluate

EpiNow2   epinowcast   scoringutils
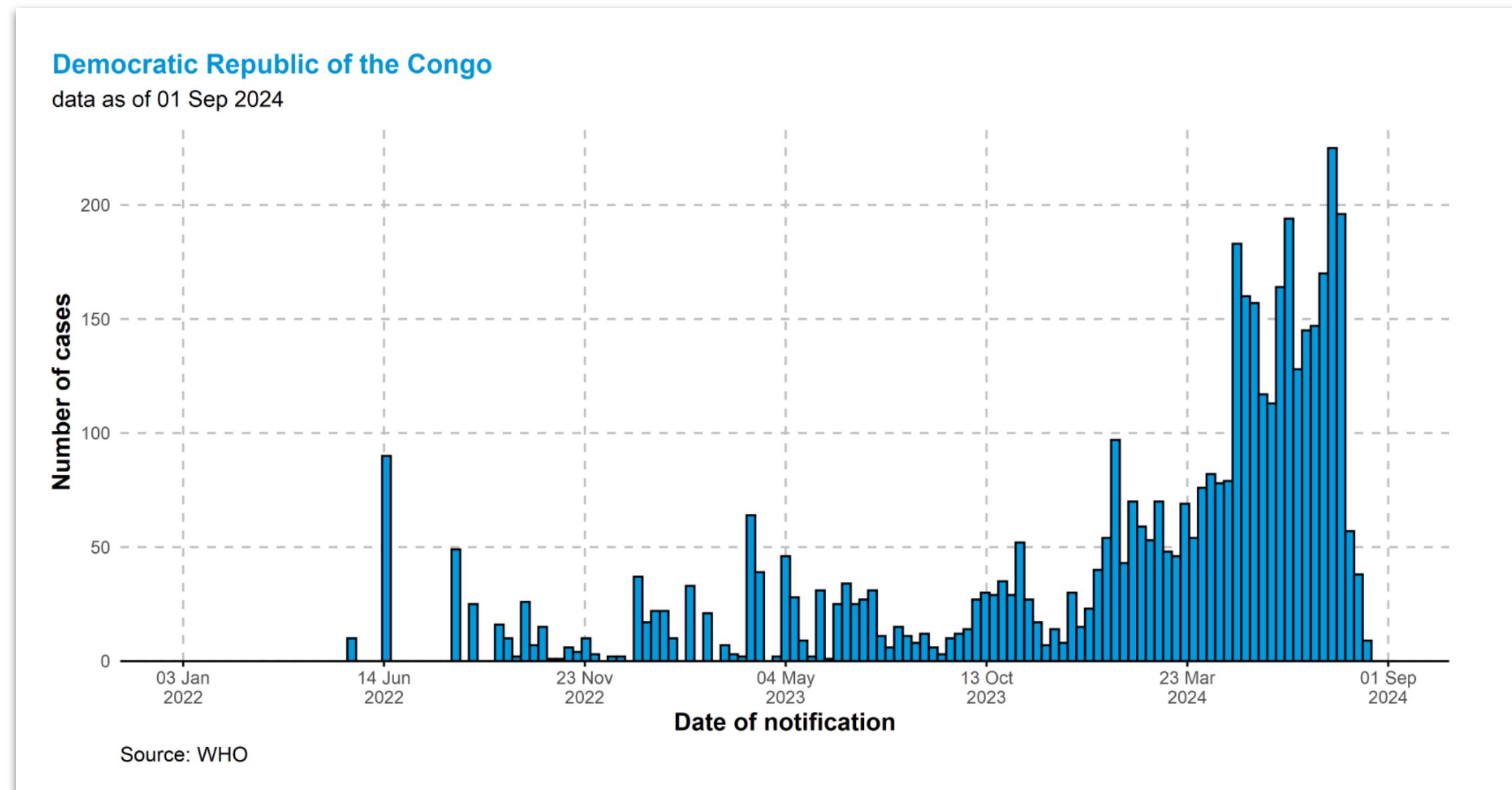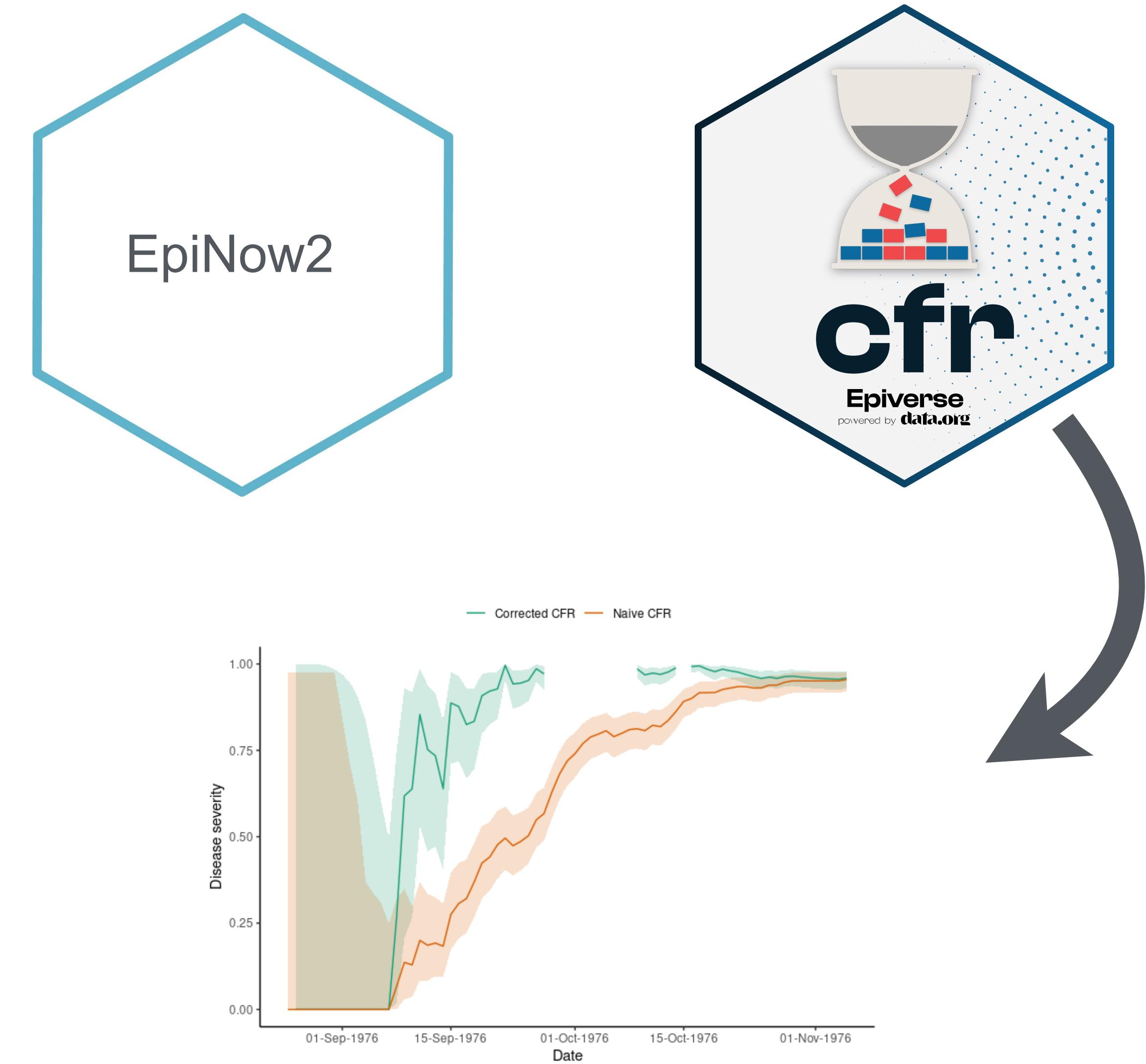
See end-to-end nowcasting example:

- EpiNow2

- epinowcast
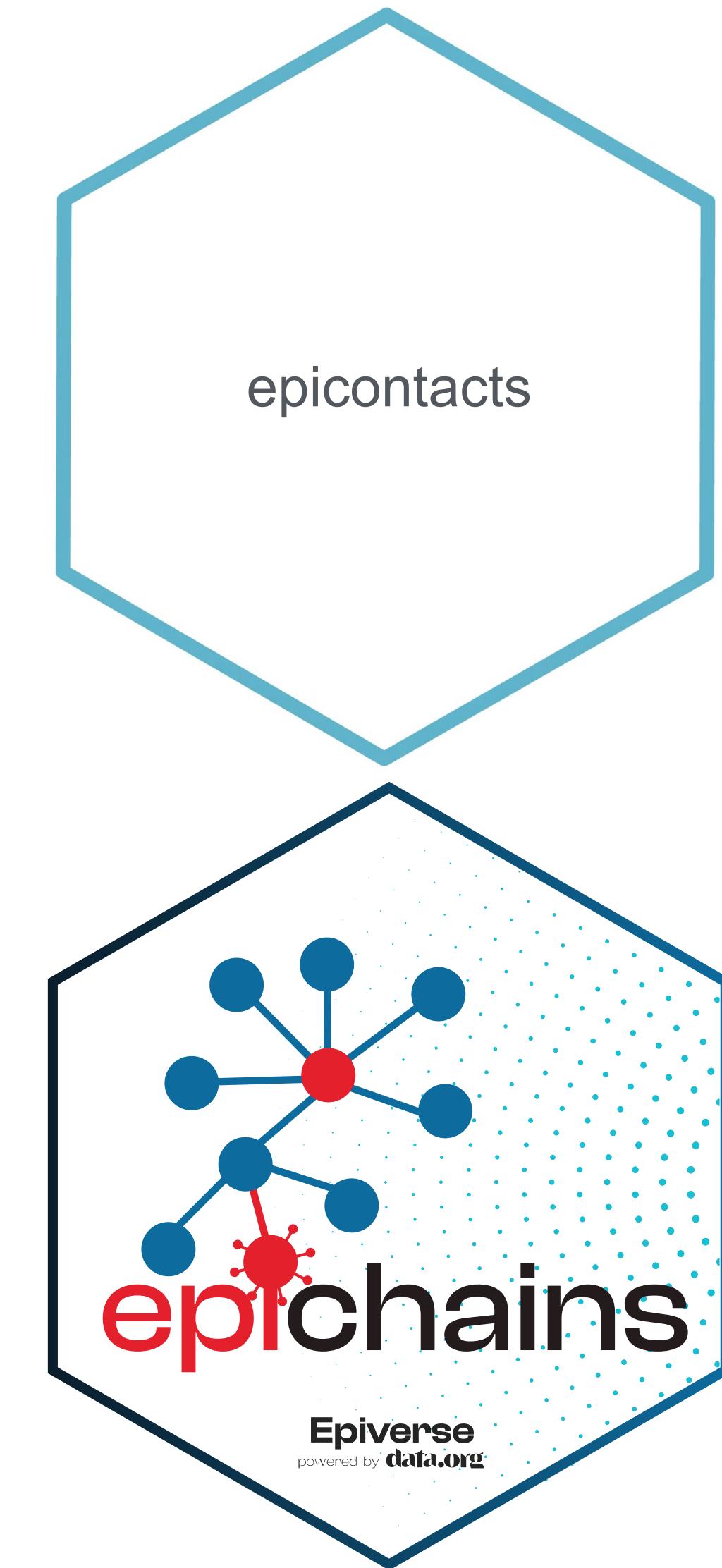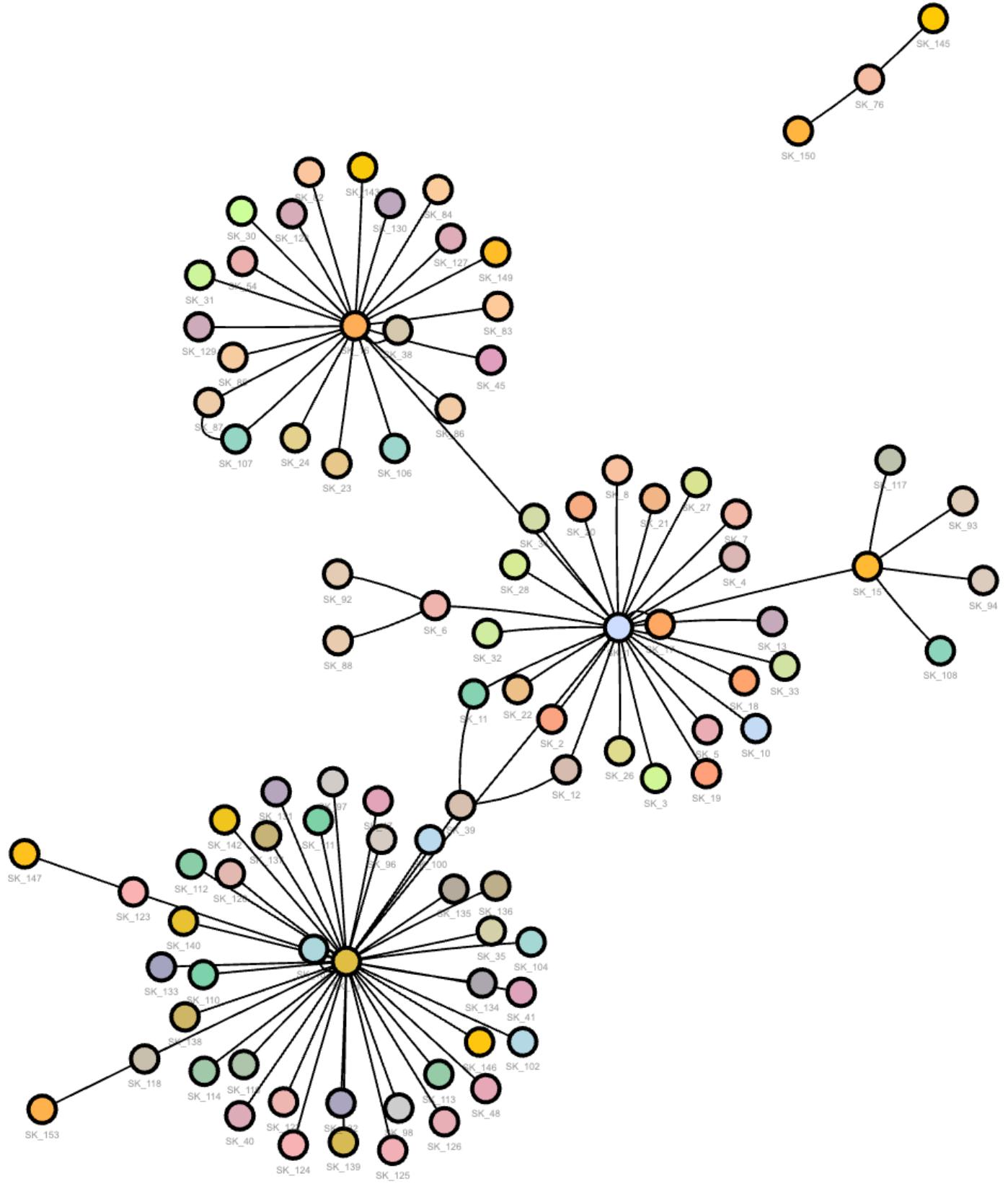
Task 6: estimate severity (cfr, etc)

Source: https://worldhealthorg.shinyapps.io/mpx_global/
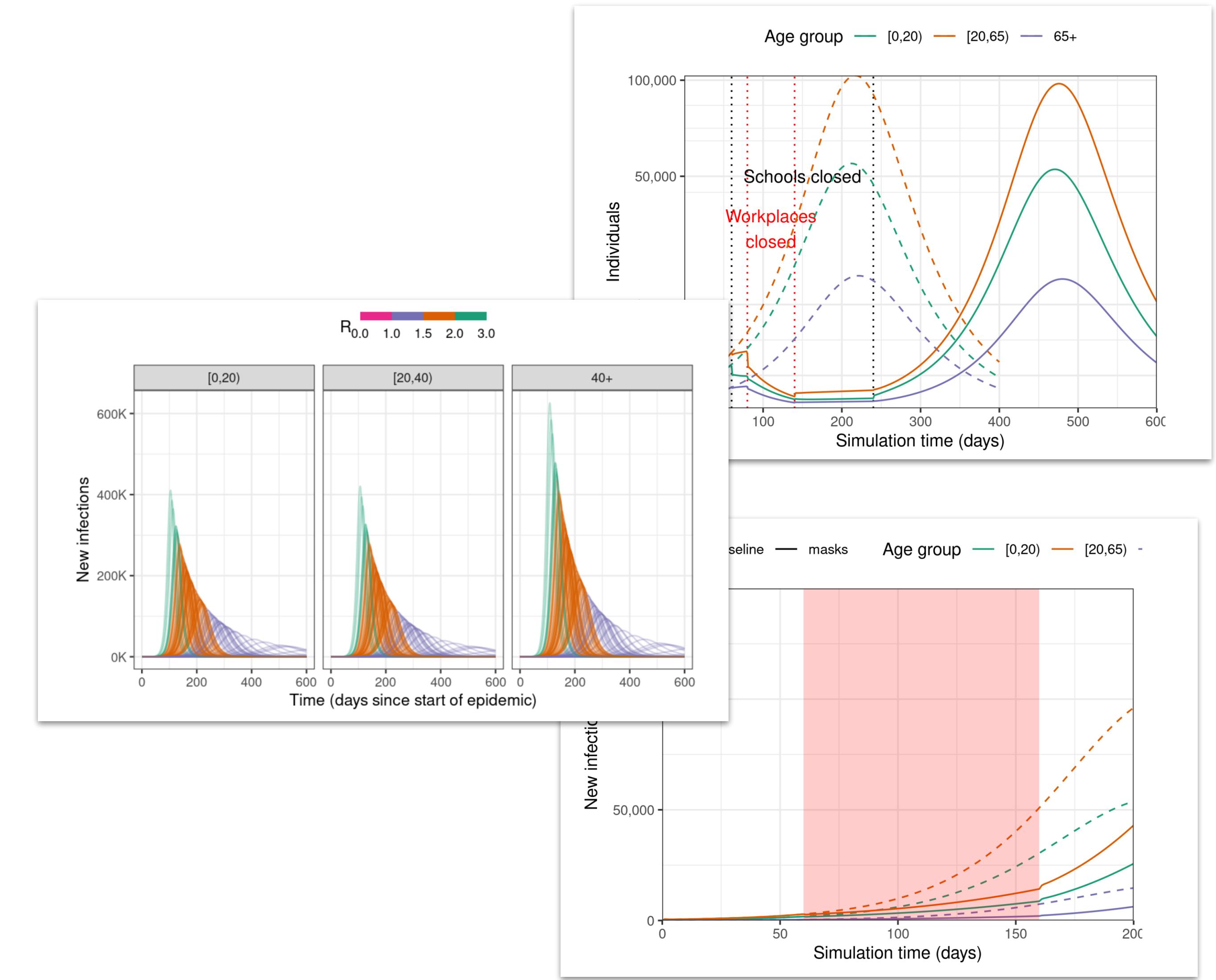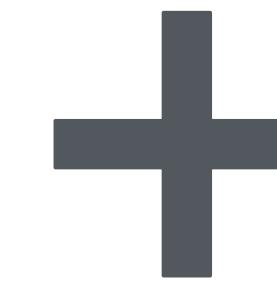
Task 7: analyse transmission chains + contact data

Transmission chains for Middle East respiratory syndrome in South Korea, 2015
Source: {epicontacts} R package Getting Started vignette

epicontacts

- Visualise and explore

- Estimate serial interval

- Cluster size distributions and contact degrees

epichains
Epiverse
powered by data.org

- Estimate R0 and k
- Fit to data using branching processes

Task 8: Analyse long term dynamics and scenarios

Source: epidemics R package vignettes

Summary of tasks

- Task 1: data cleaning

- Task 2: estimating transmissibility

- Task 3: Getting epidemiological parameter distributions

  - Task 4: Estimating epidemiological delays

  - Task 5: Nowcasting and forecasting infections

- Task 6: Estimating severity

- Task 7: Analysing transmission chains and contact networks

- Task 8: Analysing long term dynamics & scenarios

Which R package should I use?

Ten simple rules for finding and selecting R packages

Caroline J. Wendt, G. Brooke Anderson ✉

=

Rule 1: Consider your purpose

Rule 2: Find and collect options

Rule 3: Check how it's shared

Rule 4: Explore the availability and quality of help

Rule 5: Quantify how established it is

Rule 6: Seek evidence of peer acceptance and review

Rule 7: Find out who developed it

Rule 8: See how it's developed

Rule 9: Put it to the test

Rule 10: Develop your own package

Summary

- Mpox 2024 outbreak is growing

- Here, provided overview of some analytics tasks that can be anticipated

- Showcased existing and emerging R tools to tackle common tasks

- Try out with emerging data and share ideas, findings, frustrations, etc with the developer community for enhancements.

Acknowledgements

- Epiverse-TRACE Initiative

- Epinowcast community

- Epiforecasts community

- RECON

- Everyone, really!

    - R package developers & contributors

    - Methods developers

    - User community

Questions

- Are there any packages for incorporating phylogenetic data into reconstructing transmission chains?

  - Answer: outbreaker2, TransPhylo, phybreak