The **Epiverse** initiative



Adam Kucharski

Co-director, Centre for Epidemic Preparedness and Response

Epiverse/Harmonize workshop July 2023



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





github.com/epiverse-trace

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.



github.com/epiverse-trace

How does transmission vary over time? And what drives this?





Estimating reproduction number



Choose template:

New R Markdown				
Document	Template: (7) Using R Markdown Templates			
Document	Custom theming	{bslib} 🔺		
菜 Presentation	Legacy custom theming	{bslib}		
Chinu	Real-time theming	{bslib}		
Shiny	Transmissibility Report	{episoap}		
From Template	Flex Dashboard	{flexdashboard}		
	Flex Dashboard themed with {bslib}	{flexdashboard}		
	reprex (lots of features)	{reprex}		
	reprex (minimal)	{reprex} •		
	Name: Untitled			
	Location:			
	~	Browse		
Create Empty Docume	ent	OK Cancel		

Outbreak analytics pipelines

Work by Hugo Gruson et al

Select options:

An Integre or character indicating the (funct) size of the time interval used for computing the incidence. Passed as the "Interval" any incidence.2-incidence()'. week Number of days to exclude from the estimation of R1 since data is likely to still be incomplete. 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	integrate or character indicating the files (Files) indicates::::::::::::::::::::::::::::::::::::		
week. Kunser of days to acclude from the estimation of Rt since data is likely to still be incomplete. 7 7 7 8 Mouter of days to include to get the latest observed value of Rt. 21 9 9 9 9 9 9 9 9 9 9 9 9 9	Here if days is exclude from the estimation of Rt since data is likely to still be incomplete.	An integer o n`incidence	r character indicating the (fixed) size of the time interval used for computing the incidence. Passed as the 'interval' argument s2::incidence()'.
Automber of days to exclude from the estimation of Rf since data is likely to still be incomplete. 7 7 2 1 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2	mber of days to exclude from the estimation of R since data is likely to still be incomplete. 7 7 7 8 8 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9	week	
7 7 1 1 1 1 1 1 2 1 2 2 1 2 2 2 2 2 2 2	7 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	lumber of d	ays to exclude from the estimation of Rt since data is likely to still be incomplete.
Standard or of days to include to get the latest observed value of Rt. 21 30 Sould the serial interval distribution be extracted directly from the epiparameter package? term of the pathogen in the epiparameter database if 'use_parameter = TRUE'. SARS_CoV_2_velthype term of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'. 42 23 Decision 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'.	meter of days to include to get the latest observed value of Rt. T Should the senial interval distribution be extracted directly from the replarameter package? T Should the senial interval distribution for serial interval if you using value from epiparameter. Ignored If use, epiparameter = TRUE . T Should the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Ignored If use, epiparameter = TRUE . Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution for serial interval II not using value from epiparameter. Should deviation of the distribution epiparameter. Should deviation of the distribution of the distribution of the distribution epiparameter. Should deviation of t	7	0
21 23 Pould the serial interval distribution for serial interval If not using value from epiparameter. [spored If use_epiparameter a TRUE: 5456_CoV_2 with you Lean of the distribution for serial interval If not using value from epiparameter. [spored If use_epiparameter = TRUE: 42 43 44 45 45 45 45 45 45 45 45 45 45 45 45	11 Should the senial interval distribution be estructed directly from the epiparameter package? an of the pathogen in the epiparameter database if use_parameter = TRUE , AURS_CoV_2_withtype an of the distribution for serial interval if not using value from epiparameter. [gnored if use_epiparameter = TRUE , 12 underd deviation of the distribution for serial interval if not using value from epiparameter. [gnored if use_epiparameter = TRUE , 13 olice of probability distribution for serial interval if not using value from epiparameter. [gnored if use_epiparameter = TRUE , and mma me of the containing the court data over time (efeault: datacovid_heap_uk_20201024.stax) The bisected	umber of d	ays to include to get the latest observed value of Rt.
g Should the serial interval distribution be extracted directly from the apparameter package? tares of the gambagen in the epiparameter database if 'use_parameter = TRUE. SARS_CoV_2_withtype Kann of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_sepiparameter = TRUE. 4.2 Standard deviation of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_sepiparameter = TRUE. 4.9 Debice of possibility distribution for serial interval if not using value from epiparameter. Ignored if 'use_sepiparameter = TRUE. 4.9 Debice of possibility distribution for serial interval if not using value from epiparameter. Ignored if 'use_sepiparameter = TRUE.	Should the senial interval distribution be extracted directly from the explanameter package! there of the pathogen in the explanameter distabase if 'use_parameter a TRUE . ARRS_CVV_2_with/ge and the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 12 and and evidation of the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 13 and and evidation of the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 14 and devide the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 15 and the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 16 and the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 17 and the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 18 and the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 19 and the distribution for serial interval if not using value from explanameter. Ignored if 'use_explanameter = TRUE . 19 and the distribution for serial interval if (effectific distance) (hexp_uk_2001/1024.tlss) 19 and the distribution for serial interval if (effectific distance) (hexp_uk_2001/1024.tlss) 19 and the distribution for serial interval if (effectific distance) (hexp_uk_2001/1024.tlss) 19 and the distribution for serial interval if (effectific distance) (hexp_uk_2001/1024.tlss) 19 and the distribution (effectific distance) (hexp_uk_2001/1024.tlss) 19 and the distribution (effectific distance) (hexp_uk_2001/1024.tlss) 19 and the distribution (for serial interval interv	21	0
tame of the pathogen in the epiparameter database if use_parameter = TRUE . SARS_CoV_2_withsyse lean of the distribution for serial interval if not using value from epiparameter. Ignored If use_epiparameter = TRUE . 4.2 tandard deviation of the distribution for serial interval if not using value from epiparameter. Ignored If use_epiparameter = TRUE . 4.9 holice of probability distribution for serial interval If not using value from epiparameter. Ignored If use_epiparameter = TRUE .	me of the pathogen in the epiparameter database if 'use_parameter = TRUE . ARIS_COVwithing and the database if use_parameter = TRUE . and the database if use_parameter = TRUE . covwithing and the database if use parameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. Ignored if 'use_paparameter = TRUE . cov_within the distribution for serial interval if not using value from epiparameter. cov_within the distribution for serial interval if not using value from epiparameter. cov_within the distribution for serial interval if not using value from epiparameter. cov_within the distribution the	Should th	a serial interval distribution be extracted directly from the epiparameter package?
SARS_CoV_2, withype learn of the distribution for serial interval If not using value from epiparameter, Ignored IF use_spiparameter = TRUE : 42 43 45 45 46 47 48 49 49 49 40 40 40 40 40 40 40 40 40 40	ARRS_CoV_2 withtype an of the distribution for serial interval II not using value from epiparameter. Ignored II' use, epiparameter + TRUE . 12 and deviation of the distribution for serial interval II not using value from epiparameter. Ignored II' use, epiparameter + TRUE . 13 acics of probability distribution for serial interval II not using value from epiparameter. Ignored II' use, epiparameter + TRUE . 14 acics of probability distribution for serial interval II not using value from epiparameter. Ignored II' use, epiparameter + TRUE . 15 acies of probability distribution for serial interval II not using value from epiparameter. Ignored II' use, epiparameter + TRUE . 26 acies of the containing the court data over time (default: datacovid, heap_uk_20201024.stax) 20008. 2009. 2	ame of the	pathogen in the epiparameter database if 'use_parameter = TRUE'.
Alexe of the distribution for serial interval if not using value from opiparameter. Ignored if 'use_opiparameter = TRUE'. 4.2 Alexa deviation of the distribution for serial interval if not using value from opiparameter. Ignored if 'use_opiparameter = TRUE'. 4.9 Diolo of possibility distribution for serial interval if not using value from opiparameter. Ignored if 'use_opiparameter = TRUE'.	an of the distribution for serial interval If not using value from opiparameter. Ignored If 'use_spiparameter = TRUE , 12 13 14 15 15 15 15 16 16 17 16 17 17 17 17 17 17 17 17 17 17 17 17 17	SARS Col	/ 2 wildtype
4.2 Standard deviation of the distribution for serial interval if not using value from epiparameter, Ignored If use_epiparameter = TRUE 4.9 Debics of probability distribution for serial interval if not using value from epiparameter. Ignored If use_epiparameter = TRUE - memory	1.2 andard deviation of the distribution for serial interval If not using value from epiparameter, Ignored II' use_epiparameter = TRUE . 1.9 color of probability distribution for serial interval II' not using value from epiparameter, Ignored II' use_epiparameter = TRUE . 1.9 color of file containing the count data over time (default: data/covid_hosp_uk_20201094.stts) 2009	Mean of the	distribution for serial interval if not using value from epiparameter. Ignored if 'use epiparameter = TRUE',
Randard deviation of the distribution for serial interval if not using value from epiparameter. Ignored II' use_epiparameter = TRUE 4.9 4.9 Debies of possibility distribution for serial interval if not using value from epiparameter. Ignored II' use_epiparameter = TRUE .	andard deviation of the distribution for serial interval if not using value from epiparameter. [gnored If 'use_epiparameter = TRUE, 13 close of probability distribution for serial interval If not using value from epiparameter. [gnored If 'use_epiparameter = TRUE, and the of the containing the count data over time (default: data.covid_hosp_uk_20201024.atax) comma. No the selected	4.2	
4.9 thoice of probability distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.	13 acial of probability distribution for serial interval II not using value from opiparameter. Ignored I' use _epiparameter = TRUE . amma or file containing the court data over time (default: datacovid_heap_uk_20201024.stac) acouse. No the selected	Standard de	viation of the distribution for serial interval if not using value from epiparameter. Ignored if 'use epiparameter = TRUE'.
Choice of probability distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.	olos of probability distribution for serial interval II not using value from epiparameter.lpnored if 'use_epiparameter = TRUE . amma me of the containing the count data over time (default: data.covid_hosp_uk_20201024.stax) Do the selected	4.9	
namma	amma me of file containing the count data over time (default: data:covid_hosp_uk_20201024.xisx) Storesa. No file selected	hoice of pr	obability distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE'.
gamma	en of file containing the count data over time (default: data/covid_hosp_uk_20201024.xfsx) 3rome	gamma	•
-	Browse No file selected	Name of file	containing the count data over time (default: data/covid hosp uk 20201024.xisx)
Browse No file selected		Browse	No file selected
Ibiah D nackase to use for Dt actimation		rimon n pac	wage to use for the estimation
Which R package to use for Rt estimation ⊃ EciEstim	EpiEstim	C EpiEstim	
Which R package to use for Rt estimation ○ EpEtim ○ EpRive2	Inci n peckage to use for nt esumation EpiEstim EpiNox2) EpiEstim) EpiNow2	
Which R package to use for Rt estimation) EDEStim) EphPow2) Entrans	Hind Figure (and for the example) EpiNond Central) EpiEstim) EpiNow2) i2extras	

Descriptive curves:





github.com/epiverse-trace/episoap

Outbreak analytics pipelines Choose template: New R Markdown Template: ② Using R Markdown Templates Document R estimation: Custom theming {bslib} 4 Diresentation {bslib} Legacy custom theming Real-time theming {bslib} R Shiny Estimates of Rt (EpiEstim) episoap Transmissibility Report {episoap} Ξ Flex Dashboard From Template Work by Hugo Flex Dashboard themed with {flexdashboard} {bslib} Epiverse reprex (lots of features) {reprex} data.org reprex (minimal) {reprex} Gruson et al This template contains multiple files. Create a new directory for these files: Name: Untitled S Location Browse... Create Empty Document ок Cancel Estimates of Rt (EpiNow2) Select options: Descriptive curves: 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()' Descriptive analyses week **Epidemic curves** Number of days to exclude from the estimation of Rt since data is likely to still be incomplet This section creates epidemic curves ("epicurves"), with or without stratification Number of days to include to get the latest observed value of Rt. Code Incidence of cases over time 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 Type Estimate Estimate based on partial data Mean of the distribution for serial interval if not using value from epiparameter. Ignored if 'use_epiparameter = TRUE 4.2 region mean median 95% ci Standard deviation of the distribution for serial interval if not using value from epiparameter. Ignored if use epiparameter = TRUE. Þ, 1.3 [1.20; 1.40] Midlands 1.3 4.9 East of England 1.2 1.2 [1.10 : 1.30] Choice of probability distribution for serial interval if not using value from epiparameter, Ignored if use epiparameter = TRUE London 1.2 1.2 [0.99; 1.30] aamma North East and Yorkshire 1.2 1.2 [1.10 : 1.30] Name of file containing the count data over time (default: data/covid_hosp_uk_20201024.xlsx) South East 1.2 1.2 [1.10; 1.30] Browse... No file selected

North West South West

Midlands

North East and Yorkshire South East

South West

North West

github.com/epiverse-trace/episoap

1.1

1.0

1.1 [0.95 ; 1.30]

1.0 [0.90:1.10]

Which R package to use for Rt estimation O EpiEstim

O EpiNow2

i2extras
 R0

Climate data link: comparison of transmission & other indicators

Reproduction number in Rio de Janeiro



Vector transmission potential vs cases in **Dominican Republic** ion potential (Index P) 1200 -600 Case Dengue 0.5 Chikungunya 300 Zika Transn Average Inde 2013 2014 2015 2016 2012 2017

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
- Complemenary method to using case surveillance data

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





Work by Zulma Cucunubá, Nicolás Domínguez et al

github.com/epiverse-trace/serofoi

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:

<pre>convert_gamma_summary_stats(mean = 2, sd = 2)</pre>
#> \$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_etal"

#> 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

github.com/epiverse-trace/epiparameter

WHO collaboratory community to generate standardized library

Collaboratory Pandemic and Epidemic Intelligence

EpiParameter Community

About Community Resources GitHub News 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



Mordecai et al. (2017) PLOS NTD



Importing, cleaning and analysing field data



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

github.com/WorldHealthOrganization/godata R



Simulate epidemic scenarios

Define model options:

Prepare epidemiological parameters as an infection object
pandemic <- infection(
 r0 = 1.5.</pre>

```
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))</pre>

Run model

run an epidemic model using `epidemic`
output <- epidemic(</pre>

- population = uk_population,
- infection = pandemic, intervention = close_schools,

```
time end = 600, increment = 1.0
```

)

Work with Pratik Gupte, Roz Eggo et al



Show outputs



github.com/epiverse-trace/epidemics

Climate data link: inform transmission rate in models

Vector-borne transmission model



Climate data link: inform transmission rate in models



Wider activites: community blueprints for best practice



epiverse-trace.github.io/blueprints

Wider activites: 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? Convert Your R Fund		Ensuring & Sho Statistical Cor Package ction to an S3	owcasing the rectness of your R Episode 119 of the #rstats @rweek podverse.fm/episode/bvGTUi	kly_org Highlights podcast is out! arto @jimjam_slam
	Generic: Benefits, P Considerations	Pitfalls & Design	 Extending data frames @joshu Solar system plots with {ggsol 	ıa_lambert ar} @hrbrmstr
Improving the an Rcpp Packag	C++ Code Quality of ge		_	
			epiverse-trace.	github.io/blog.html

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

