Methods to estimate the epidemic reproduction number

Anne Cori

a.cori@imperial.ac.uk

MRC Centre for Global Infectious Disease Analysis
School of Public Health
Imperial College London

Modelling to Support Resilience for Pandemics
Cambridge
22 June 2022

MRC Centre for Global Infectious Disease Analysis

Imperial College London

With apologies in advance for the self sales pitch

The reproduction number R_t

MRC Centre for Global Infectious Disease Analysis

Imperial College London

R_t = average number of secondary cases infected by each case "at time t"

→ Easy to interpret

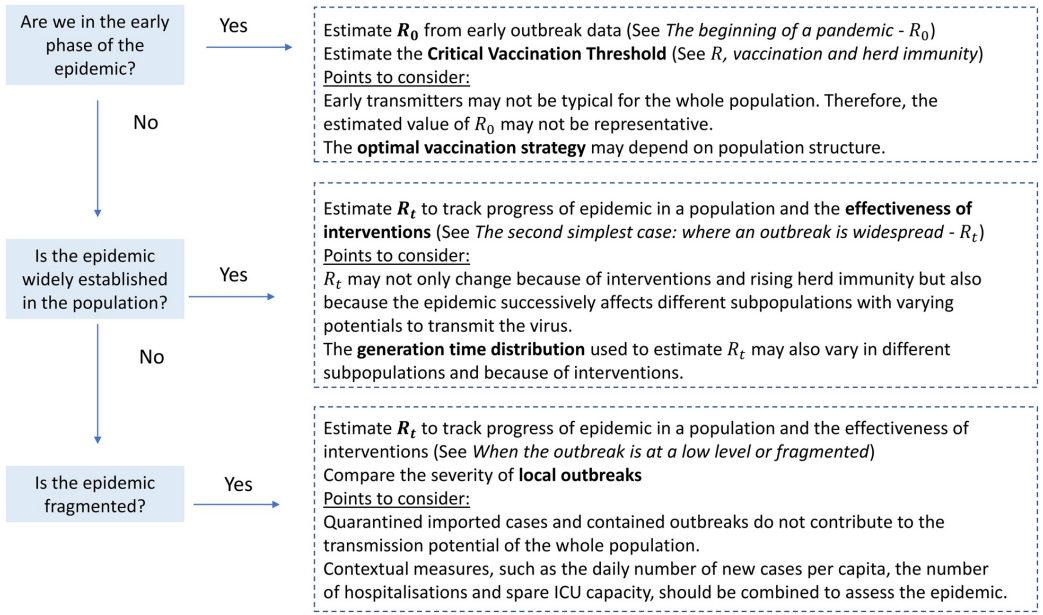
→ Useful at different stages of an outbreak, to:

1. **Predict** the potential impact of the outbreak
2. **Assess** the feasibility of control measures
3. **Track** potential changes in transmissibility over time
4. **Evaluate** the effectiveness of control measures

The reproduction number R_t

MRC Centre for Global Infectious Disease Analysis

Imperial College London



Vegvari et al. Statistical methods in Medical Research, 2021

Why should we care?

Why should we care?

MRC Centre for Global Infectious Disease Analysis

Imperial College London

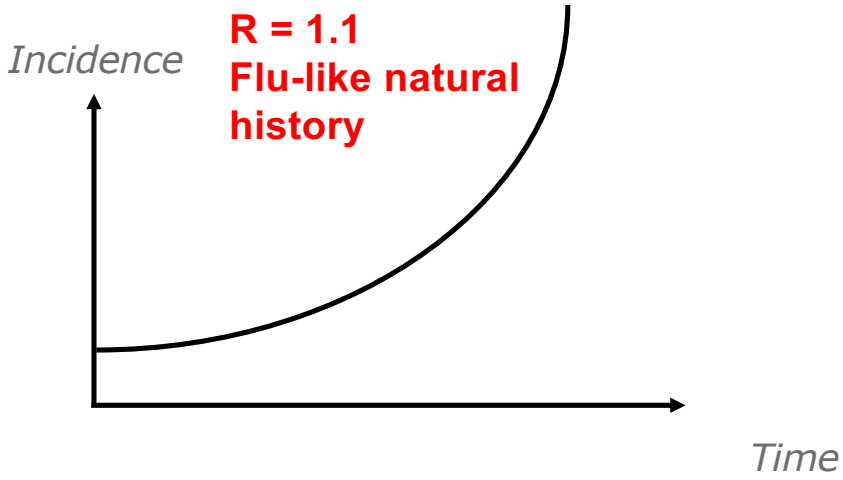
- "**We know when cases are going up and we are going to be overwhelmed**" (HCWs)
- "**The growth rate is more informative.**" (Many people)

Why should we care?

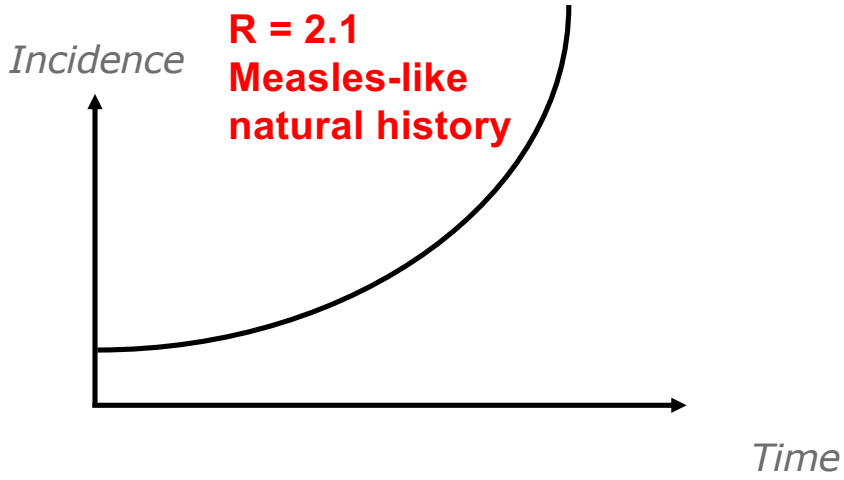
MRC Centre for Global Infectious Disease Analysis

Imperial College London

- "**We know when cases are going up and we are going to be overwhelmed**" (HCWs)
- "**The growth rate is more informative.**" (Many people)
  - **Yes**: the growth rate tells you how fast cases are growing / declining
  - **No**: the growth rate doesn't tell you how much effort is needed for control



Need to prevent ~9% of onwards infections.

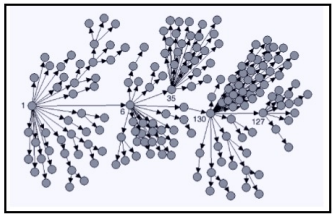


Need to prevent >50% of onwards infections.

A multitude of approaches

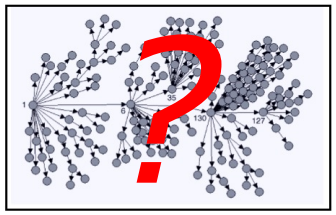
MRC Centre for Global Infectious Disease Analysis

Imperial College London



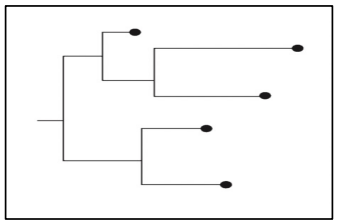
- **Contact tracing**

e.g. Tsang et al. NEJM 2003



- **Inferring who infected whom**

Wallinga & Teunis AJE 2004
Jombart et al. PLoS Comp Biol 2014

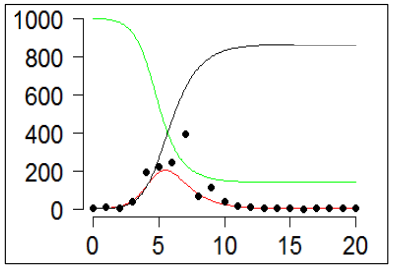


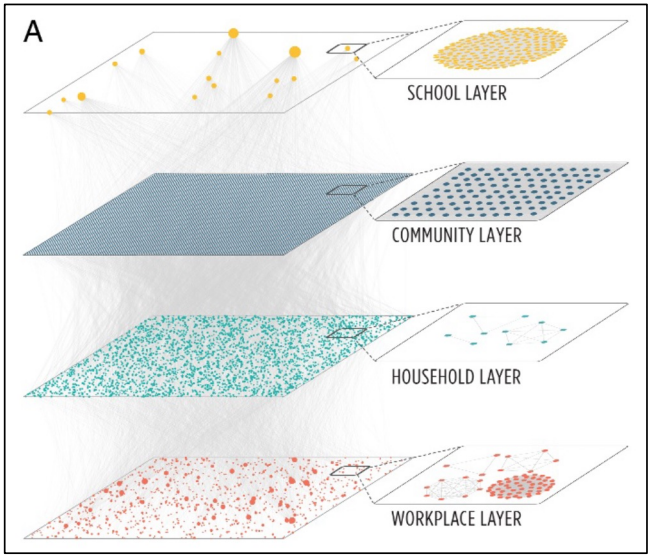
- **Fitting a model to genetic data**

Stadler et al. MBE 2011
Volz & Siveroni PLoS Comp Biol 2018

- **Mechanistic model fitting to epi data**

e.g. Kucharski et al. PNAS 2015,
Liu et al. PNAS 2018





NOT AN EXHAUSTIVE LIST!

Common issues

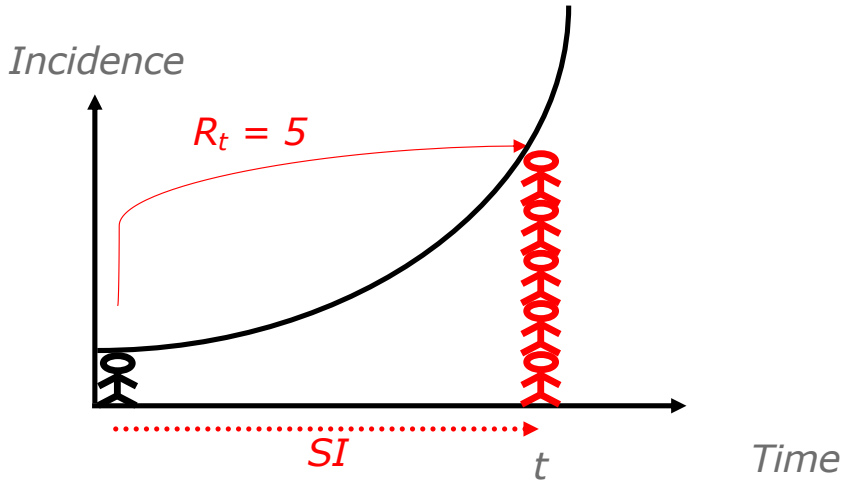
- Expensive, not easily scalable
- Not feasible in real-time (require future data AND/OR too long to develop)
- Require many assumptions (about the population, the pathogen, etc)

So let's go super simple

MRC Centre for Global Infectious Disease Analysis

Imperial College London

- R_t can be estimated using data on incidence and the serial interval SI

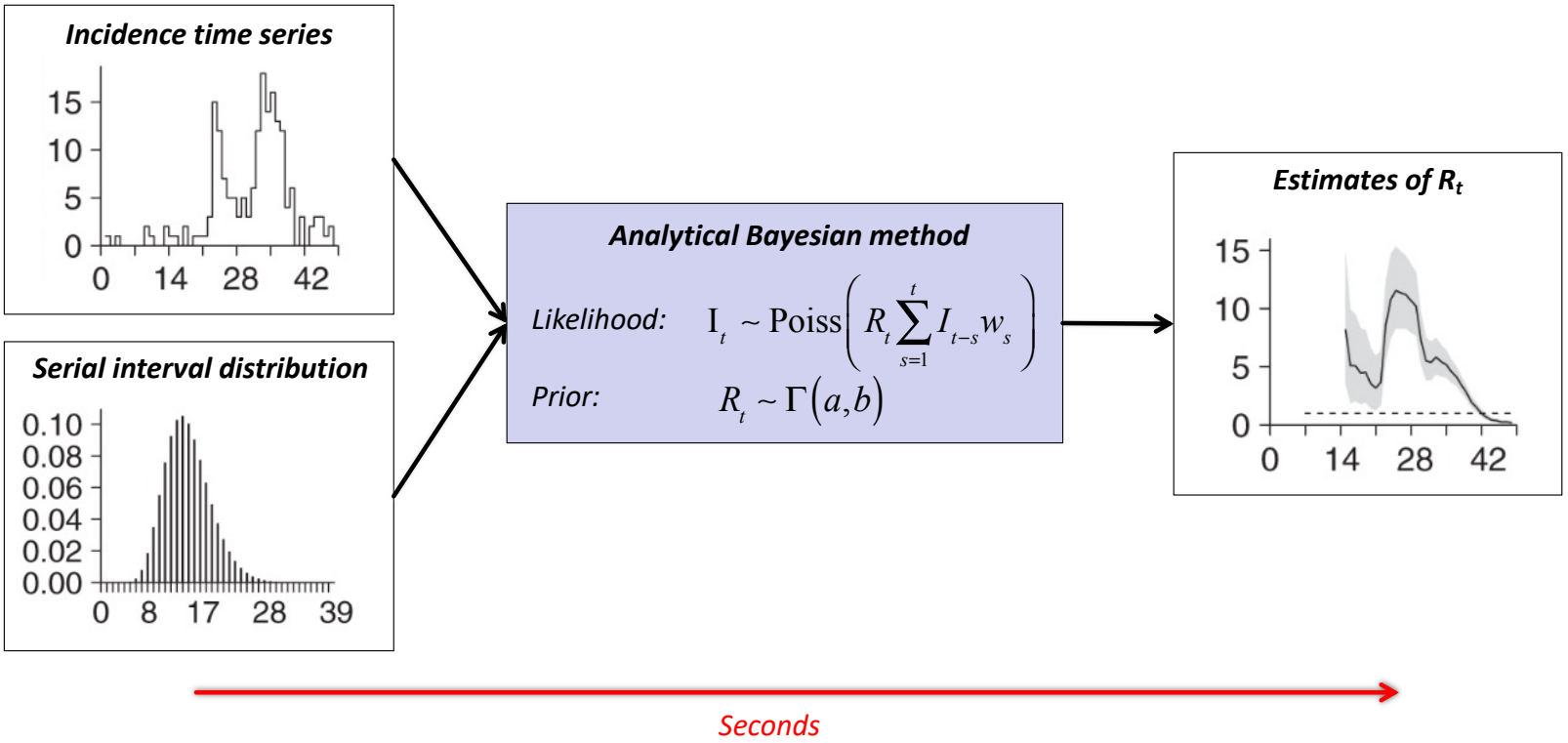


Cori et al. AJE 2013

The EpiEstim method

MRC Centre for Global Infectious Disease Analysis

Imperial College London



http://tools.epidemiology.net/EpiEstim.xls
http://cran.r-project.org/web/packages/EpiEstim/index.html
https://shiny.dide.imperial.ac.uk/epiestim/

Cori et al. AJE 2013

“But this is <u>really</u> simple”

(again, many people)

It's simple but it works well

MRC Centre for Global Infectious Disease Analysis

Imperial College London

"**For near real-time estimation of R_t, we recommend the approach of Cori and colleagues**, which uses data from before time *t* and empirical estimates of the distribution of time between infections."

Gostic et al. PLoS Comp Biol 2020

And people use it a lot

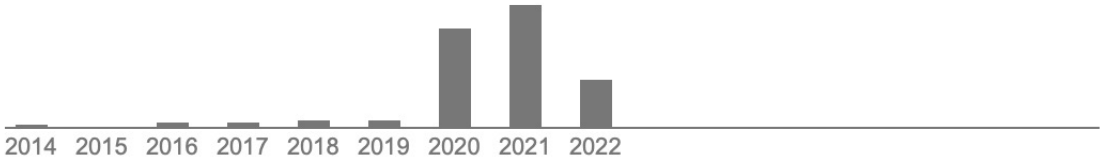
MRC Centre for Global Infectious Disease Analysis

Imperial College London

A new framework and software to estimate time-varying reproduction numbers during epidemics

[HTML] from oup.com
Full View

Authors	Anne Cori, Neil M Ferguson, Christophe Fraser, Simon Cauchemez
Publication date	2013/11/1
Journal	American journal of epidemiology
Total citations	Cited by 1063

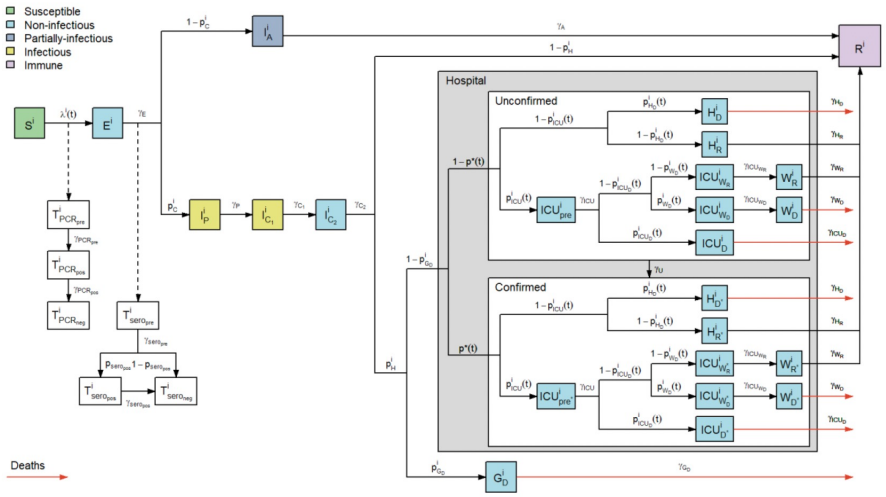


Is it too simple?

It gives similar estimates to more complex models

MRC Centre for Global Infectious Disease Analysis

Imperial College London



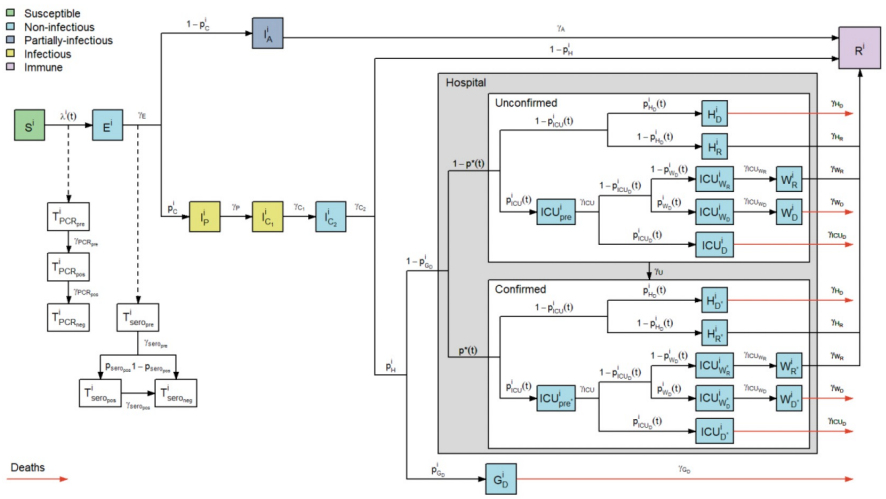
(722 compartments)

Knock et al. Science Translational Medicine 2021

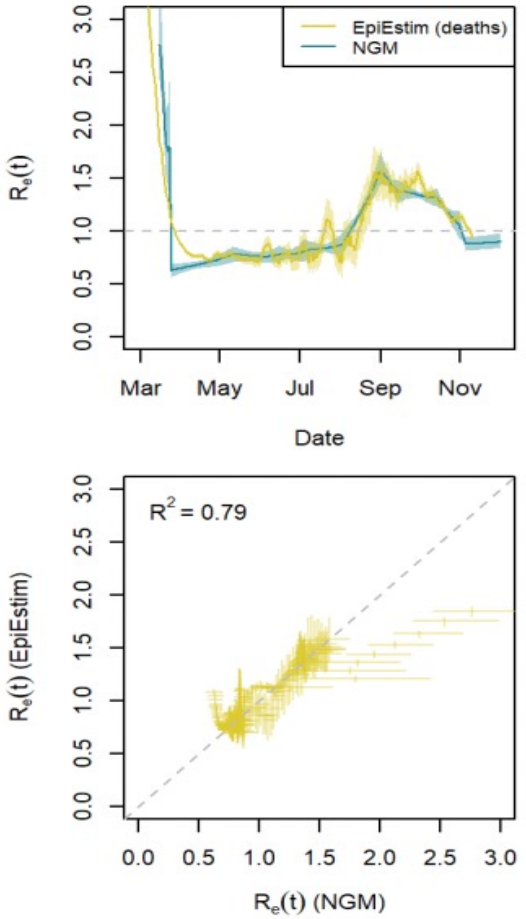
It gives similar estimates to more complex models

MRC Centre for Global Infectious Disease Analysis

Imperial College London



(722 compartments)



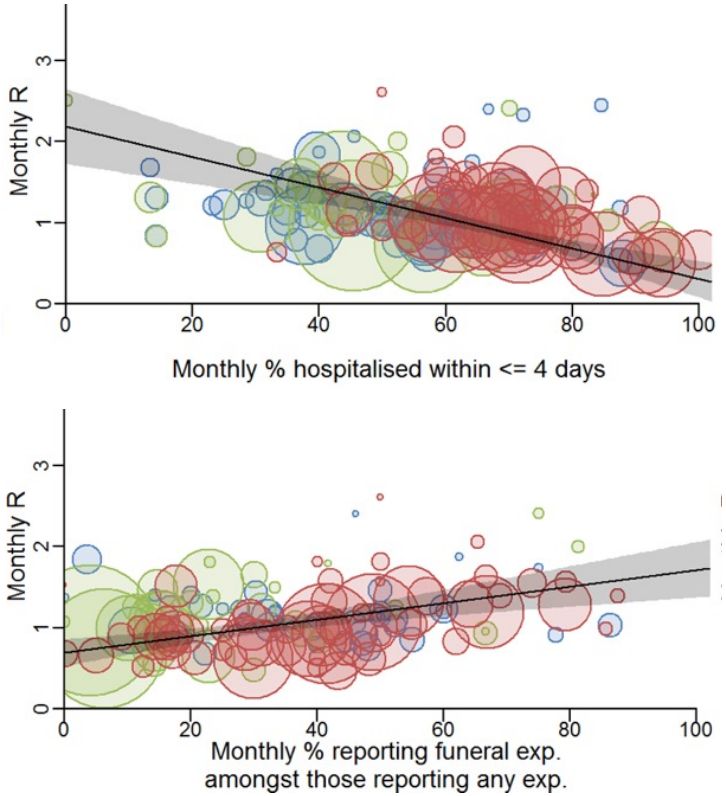
Knock et al. Science Translational Medicine 2021

It can still allow exploring the impact of interventions (albeit in a simple way)

MRC Centre for Global Infectious Disease Analysis

Imperial College London

(as long as you record information about interventions…)



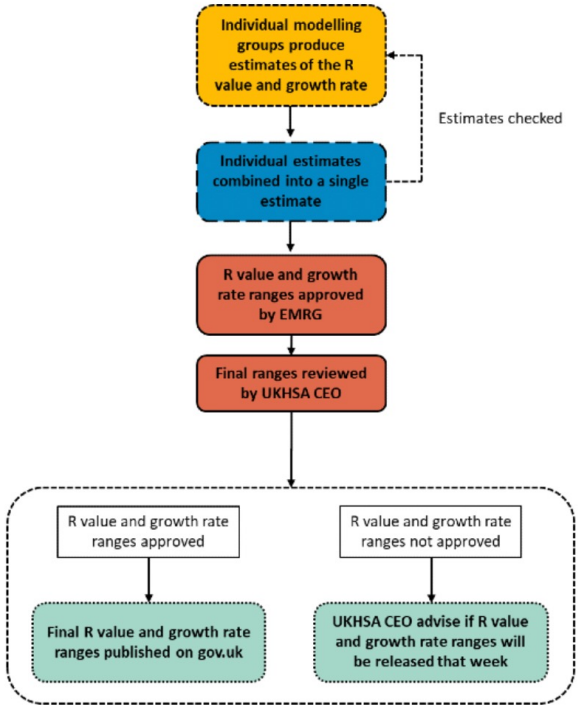
WHO Ebola Response Team PLoS Med 2016
Cori et al. Phil Transac Roy Soc B 2017

It is perhaps best used in combination with other more complex models?

MRC Centre for Global Infectious Disease Analysis

Imperial College London

Flowchart: estimating the R value and growth rate – from production to publication



More complex models take time to develop, and may be the first to be dropped as we get out of the most acute phase

→ useful to have simpler model informing pooled estimates

https://www.gov.uk/government/publications/reproduction-number-r-and-growth-rate-methodology/reproduction-number-r-and-growth-rate-methodology

MRC Centre for Global Infectious Disease Analysis

Imperial College London

Benefits of it being so simple:

We can add features pretty easily

whilst keeping the model and software relatively simple

Using the latest data informing the serial interval

MRC Centre for Global Infectious Disease Analysis

Imperial College London

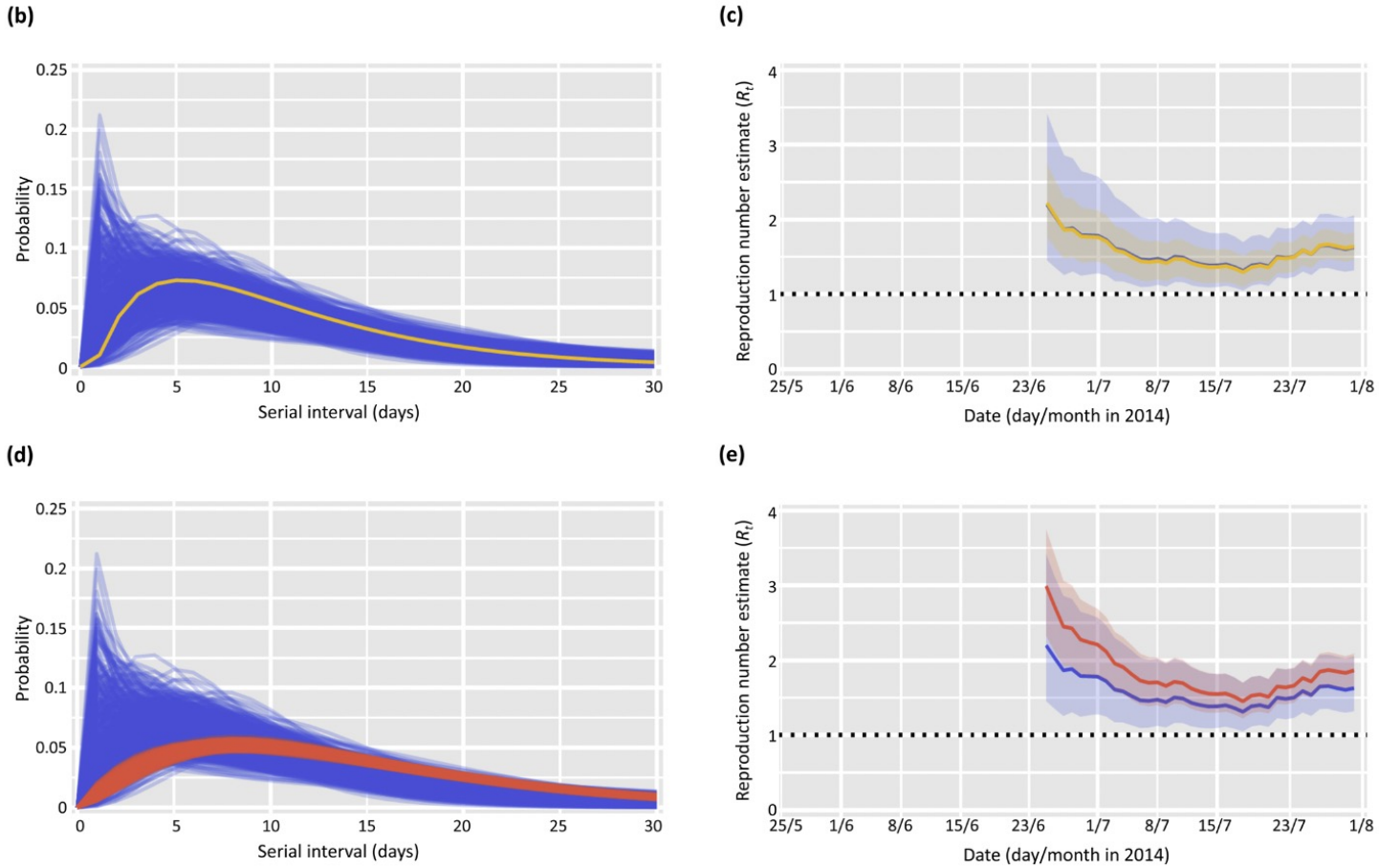


Fig. 4. Estimates of the time-dependent reproduction number, R_t, for the early part of the 2013–2016 Ebola outbreak in Liberia (International Ebola Response Team,

→ Can make a large difference in the R estimates

Thompson et al. Epidemics 2019

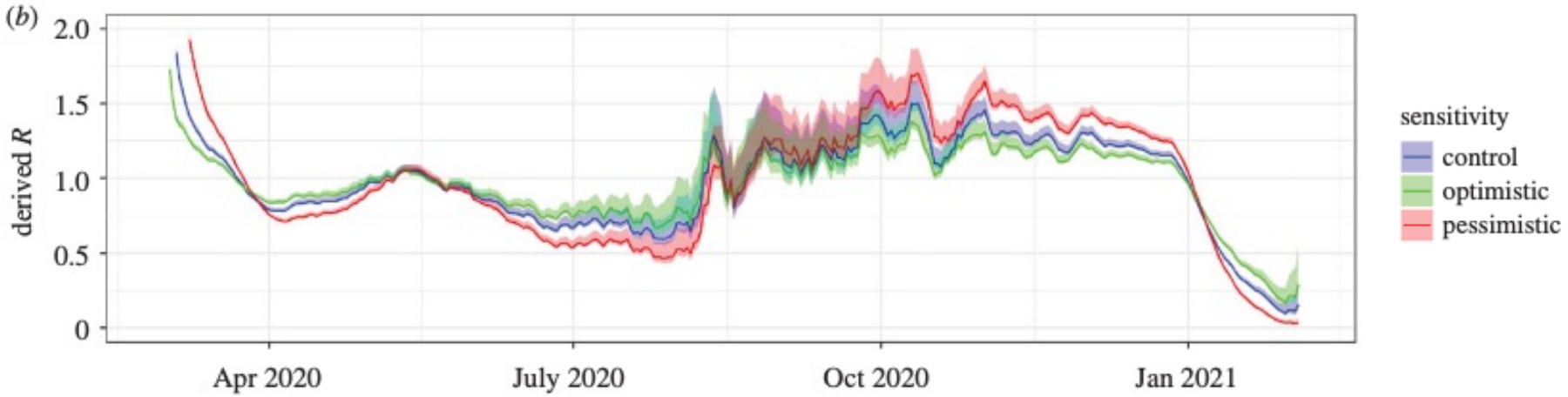
Accounting for heterogeneities in the serial interval

MRC Centre for Global Infectious Disease Analysis

Imperial College London

→ Using a single type branching process can lead to biases in R estimates

→ Bias is small for isolating / non isolating, but can be large for symptomatic / asymptomatic, vaccinated / unvaccinated if those have different serial intervals



→ Importance of characterising the serial interval of an “average” transmission pair, not just from symptomatic unvaccinated cases.

Green et al. J. R. Soc. Interface 2022

Estimating the transmission advantage of novel variants / strains

MRC Centre for Global Infectious Disease Analysis

Imperial College London

- Extending EpiEstim (analytically tractable marginal posteriors)

→ Fast, so can

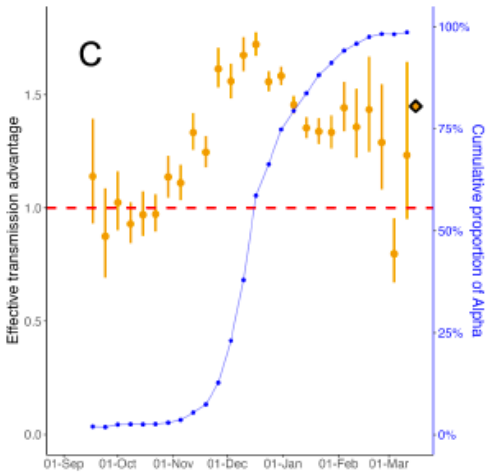
explore spatiotemporal heterogeneity

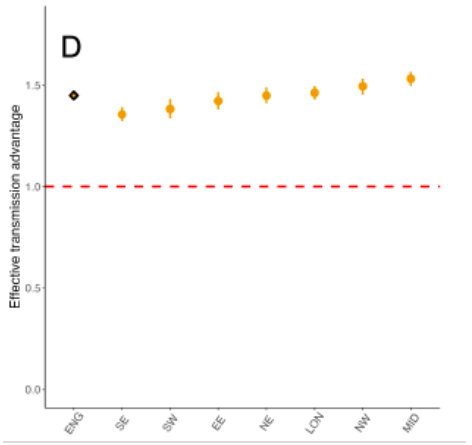
pool information where relevant

- Retrospective analysis of the Alpha variant,
- prospective analysis for Delta & Omicron

- Measures the “effective transmission advantage”

→ First step in quantifying threat from new variants





Revised version of *Bhatia et al. MedRxiv 2021*

Estimating R from data reported over long time intervals

MRC Centre for Global Infectious Disease Analysis

Imperial College London

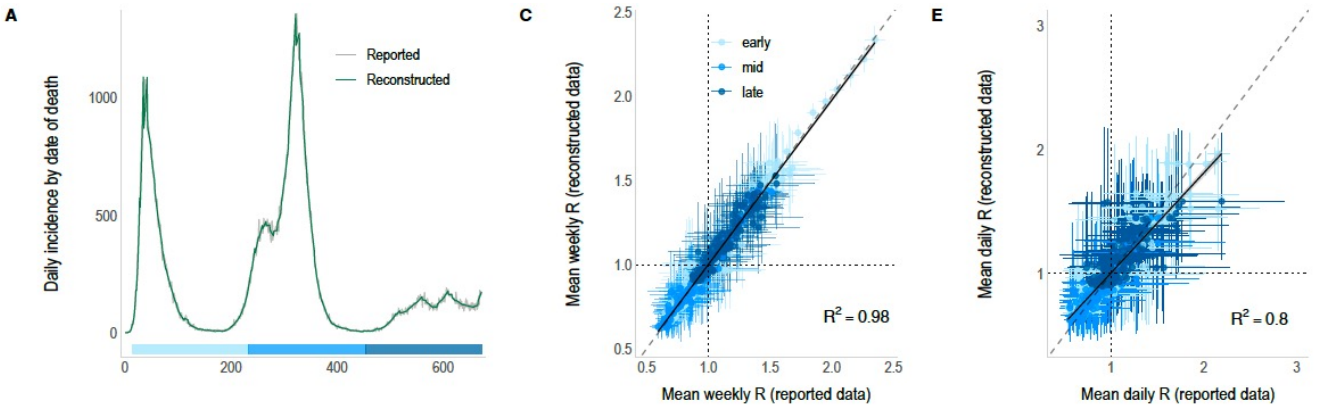
- Expectation-Maximisation algorithm combined with EpiEstim to reconstruct daily incidence and estimate R

- Applied to COVID and flu daily data which we aggregated

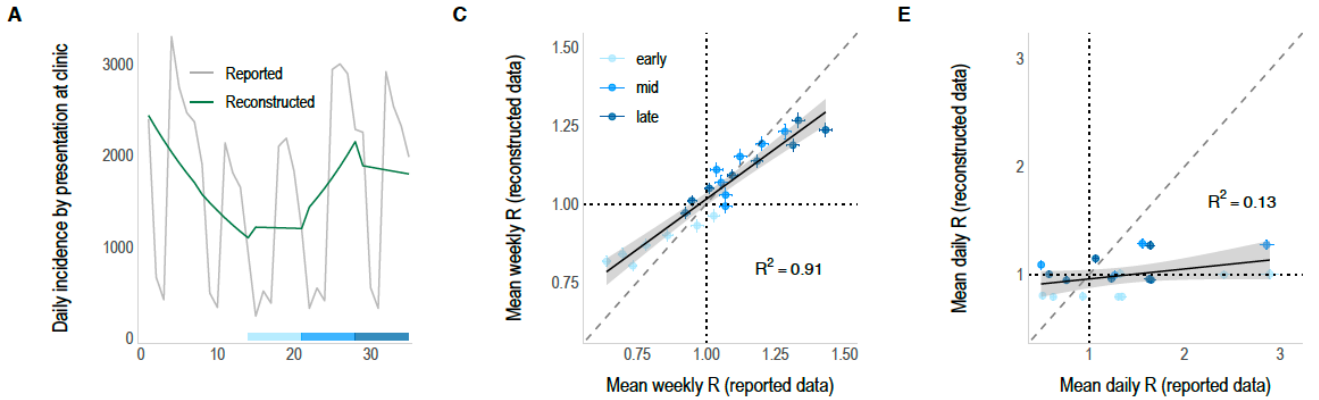
→ Expands applicability of EpiEstim (e.g. for less acute COVID phase)

→ May be useful even when data is reported daily but with administrative noise

COVID-19 deaths in the UK



ILI in US military (2009-10)



Nash et al. in prep

MRC Centre for Global Infectious Disease Analysis

Imperial College London

What context has EpiEstim been applied to, and what features have others been adding?

What are people using this for?

MRC Centre for Global Infectious Disease Analysis

Imperial College London

- Literature review + survey of EpiEstim users

Disease/disease agent investigated	*n*
COVID-19	198
Ebola Virus Disease (EVD)	7
Influenza	7
Measles	4
Multiple diseases	3
Cholera	3
Zika	2
Severe Acute Respiratory Syndrome (SARS)	2
Respiratory Syncytial Virus (RSV)	2
Middle East Respiratory Syndrome (MERS)	2
Other	12

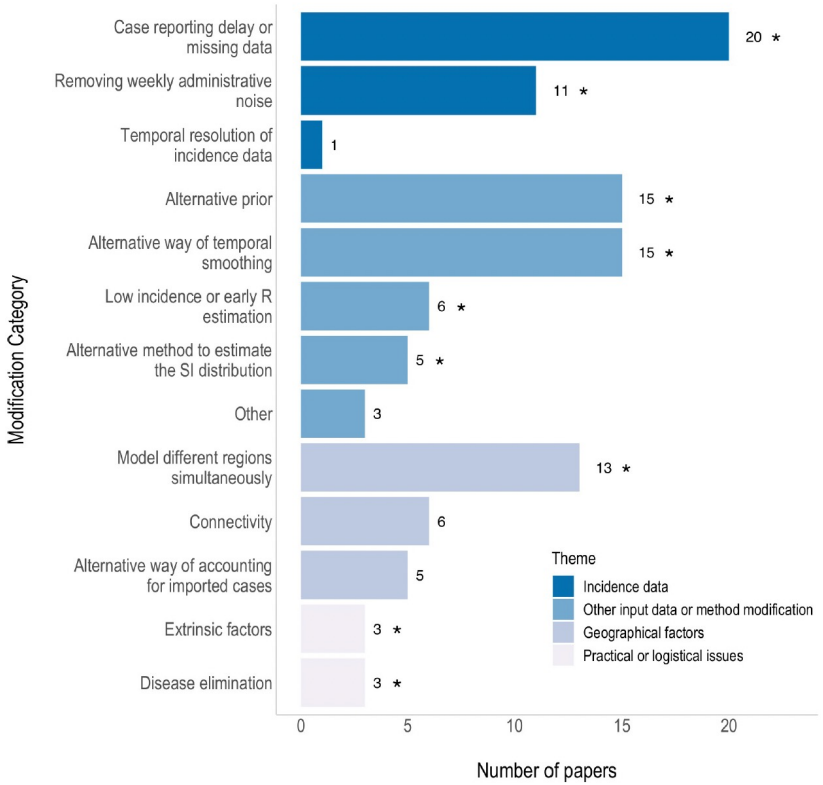


Fig 3. Summary of the modifications made to the EpiEstim method in 54 papers. Papers which made multiple modifications are counted more than once. Asterisks are present next to the modification category if an R package or opensource tool has been identified that addresses an issue within that category. See text for a description of all modifications and Table 2 for a summary of the R packages identified.

Nash et al. PLoS Digital Health, in press

Review of software

MRC Centre for Global Infectious Disease Analysis

Imperial College London

Table 2. Summary table of the R packages and tool (*) that address some of the issues identified in the scoping review. The ticks (✓) indicate whether an R package includes that modification type and rows with hash lines show the modification types that have not been incorporated into a known R package or tool. The final three rows summarise additional exploration by the authors to assess how easily each package/tool can be installed and used. For full details of how each classification (very good = ✓✓, good = ✓, poor = ✗) was defined, see Table B in S1 Text and the associated text (section 3).

Theme	Modification type	APEestim *(v 0.0.1)*	bayEStim *(v 0.0.1)*	earlyR *(v 0.0.5)*	epicontacts *(v 1.1.2)*	Epidemia *(v 1.0.0)*	EpiFilter* *(*Tool)*	EpiNow2 *(v 1.3.2)*
Incidence	Case reporting delay or missing data		✓			✓		✓
	Removing weekly administrative noise							✓
	Temporal resolution of incidence data							
Other input data or method modification	Low incidence or early R estimation			✓			✓	
	Alternative prior					✓	✓	✓
	Alternative way of temporal smoothing	✓				✓	✓	✓
	Alternative method to estimate the SI distribution				✓			
Geographical factors	Connectivity							
	Model different regions simultaneously					✓		✓
	Alternative way of accounting for imported cases							
Practical/logistical	Extrinsic factors					✓		
	Disease elimination						✓	
Additional exploration (✓✓ = very good, ✓ = good, ✗ = poor)								
Usability	Ease of installation	✓	✗	✓✓	✓✓	✗	✗	✓
	Documentation and tutorials (e.g., vignette)	✗	✗	✓✓	✓✓	✓✓	✓✓	✓✓
	Speed of R estimation◇ (◇except for epicontacts, which estimates the SI)	✓✓	NA	✓✓	✓✓◇	✓	✓✓	✗

Nash et al. PLoS Digital Health, in press

An epidemic of methods / software?

MRC Centre for Global Infectious Disease Analysis

Imperial College London

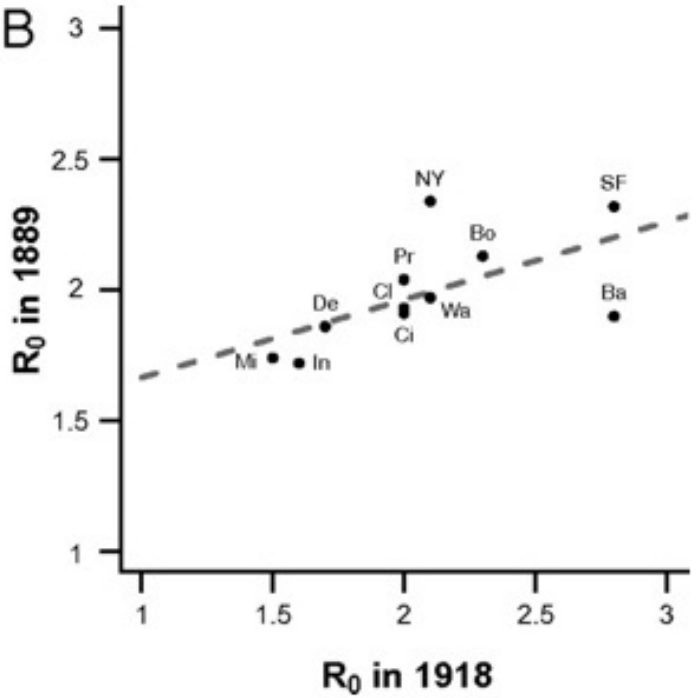
- With homoplasy?!
- Many methods are not implemented in easily useable ready to use software
- Little systematic comparison of performance / clarity on when one should use which software

Ongoing work and future priorities

MRC Centre for Global Infectious Disease Analysis

Imperial College London

- Strategy for method and software development? Discussion welcome!
- Comparison of methods + understanding what is “worth” the effort, in what context:
  - Changes in the serial interval over time
  - Super-spreading
  - Ideal temporal and spatial scale of analysis?
- Implications for forecasting
- Retrospective analyses can involve simple models to test hypotheses, e.g. locations at risk?



Valleron et al. PNAS 2010

Conclusion

MRC Centre for Global Infectious Disease Analysis

Imperial College London

- Simple ready-to-use tools are important for rapid outbreak assessment
  → estimation of R but many other epidemic characteristics
- Most countries do not have SPI-M
  with tens of groups contributing weekly situation updates based on complex modelling
- Even in countries who do, having simple ready-to-use tools is critical
  - in the early days
  - as a benchmark
- Tools are only useful if they are useable in practice
- R is important but doesn't tell you the whole picture

Acknowledgements

Imperial College London
- Sangeeta Bhatia, Neil Ferguson, Will Green, Rebecca Nash, Jack Wardle
- The Imperial College real time COVID-19 modelling team

University of Sussex
Pierre Nouvellet

University of Oxford
Christophe Fraser

Institut Pasteur
Simon Cauchemez

University of Warwick
Robin Thompson

EpiEstim contributors and users

Funders:
MRC, NIHR HPRU

Questions?