

Methods for Short Term Projections in epidemics (Projections Package)

Pierre Nouvellet, Anne Cori, Thibaut
Jombart, Sangeeta Bhatia

pierre.nouvellet@sussex.ac.uk

- Context
 - Basic principle: from model to inference to predictions?
 - Caveats

Structure

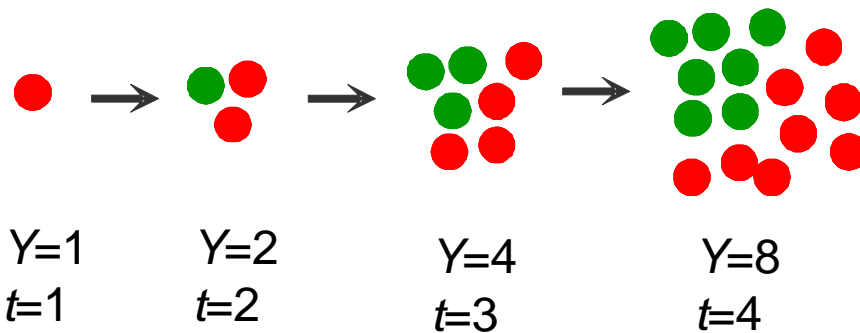
- What do I mean by projections/forecasts/predictions?
 - Projections: short term not mechanistic – taking current trend and continuing
 - Forecasts: relies on somehow more mechanistic model but typically assumes conditions in future remain stable
 - Predictions: relies on understanding the system and making hypothesis about future conditions – closer scenario modelling

Projection/Forecasting

- Importance, especially in context of public agencies and stakeholders:
 - Advocacy and planning
 - Monitoring the situation
 - Implementation/evaluation of control strategies
- Challenges:
 - Uncertainties surrounding the data
 - Uncertainties surrounding the dynamics of transmission
- In such context, we initially focussed on projecting case incidence:
 - Pro: Robust methodology
 - Con: weak mechanistic underlying model, so limited use for modelling the impact of interventions

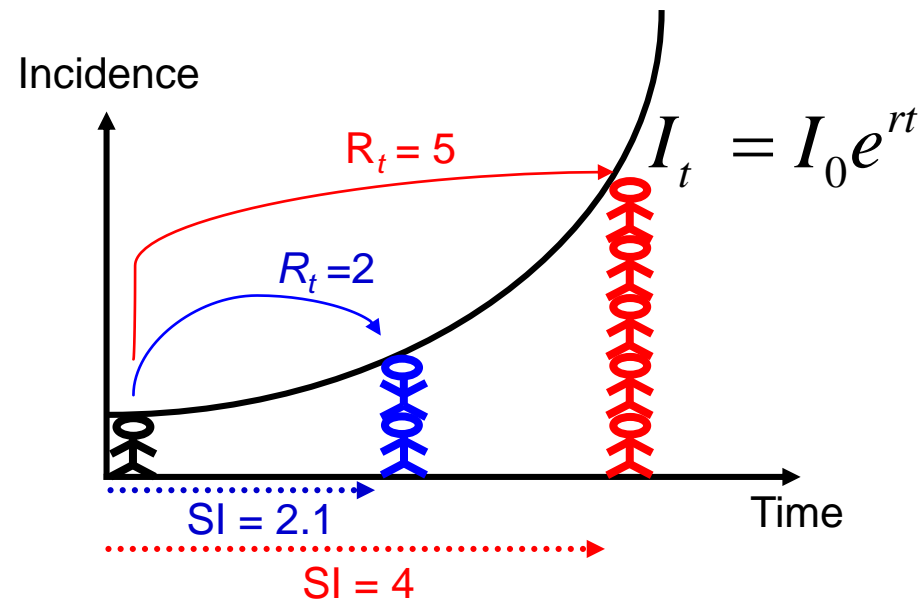
The reproduction number

- Basic reproduction number R_0 : average number of secondary cases generated by an index case in a large entirely susceptible population



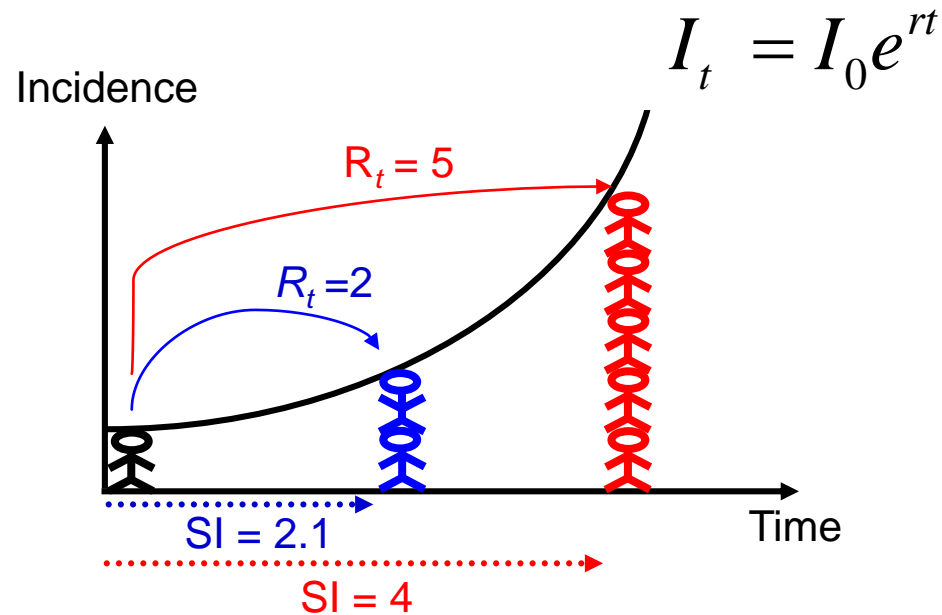
Contagion

- Effective reproduction number R_t
→ equivalent at time t



Estimation of R_0 and R_t :

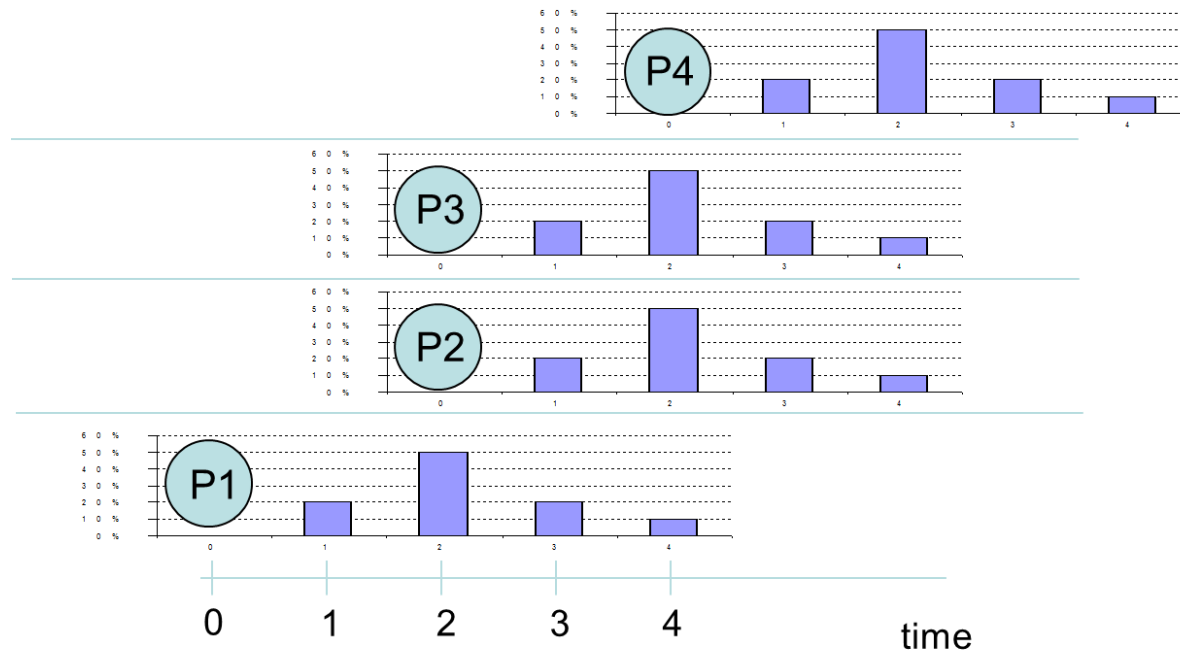
As long as there is a large proportion of susceptibles in the population, the epidemic will grow exponentially R_0 (later we define R_t)



The serial interval (time between symptoms onset of infector and symptoms onset of infectee), informs on the value of R_t

Distribution of serial interval: w_t

proxy for infectiousness: when the R_0/t new infection will occur



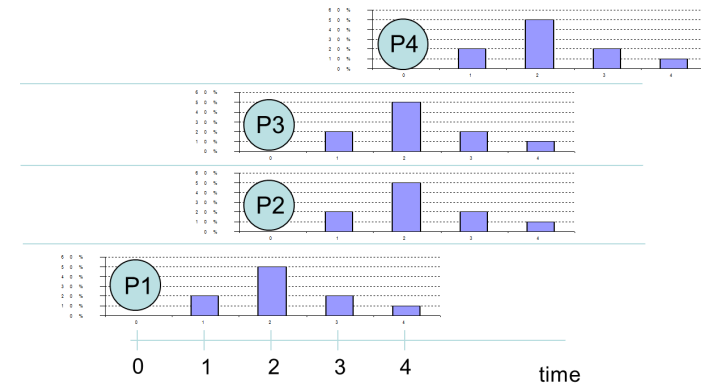
Distribution of serial interval: w_t

proxy for infectiousness: when the R_0/t new infection will occur

$$I_t = \mathcal{P} \left(R_t \sum_{s=1}^t I_{t-s} w_{t-s} \right)$$

Same equation used to:

- Infer R_t
- Project I_t in the future (typically assuming the last observed R_t remain constant)



Given knowledge of the serial interval distribution, we are able:

- Estimate R_t , doubling time

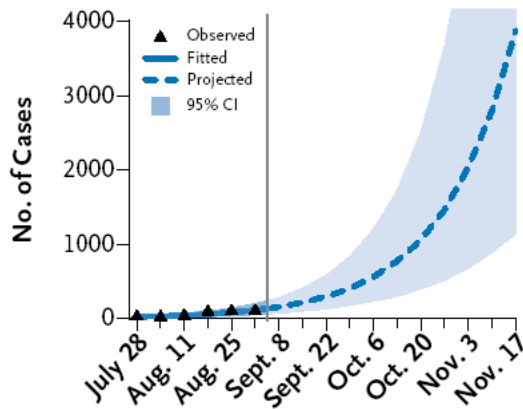
Given a time-series of incident cases and knowledge of R_t , we are able to:

- Predict the future number of cases (should the situation remains the same) - Projections

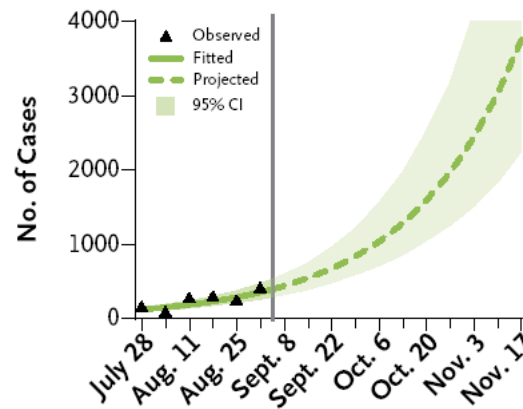
$$I_t = \mathcal{P} \left(R_t \sum_{s=1}^t I_{t-s} w_{t-s} \right)$$

	Guinea	Liberia	Sierra Leone
R_t	1.81 (1.60–2.03)	1.51 (1.41–1.60)	1.38 (1.27–1.51)
Initial doubling time (days)	15.7 (12.9–20.3)	23.6 (20.2–28.2)	30.2 (23.6–42.3)

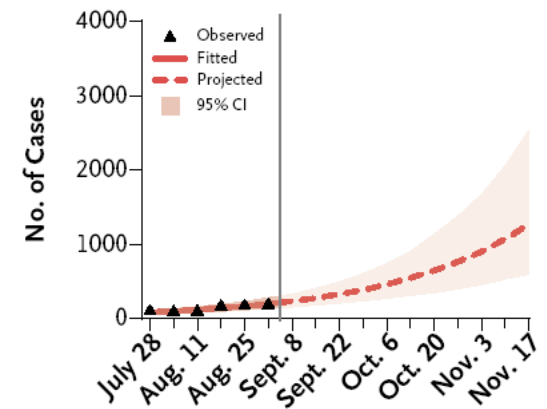
A Guinea



B Liberia



C Sierra Leone

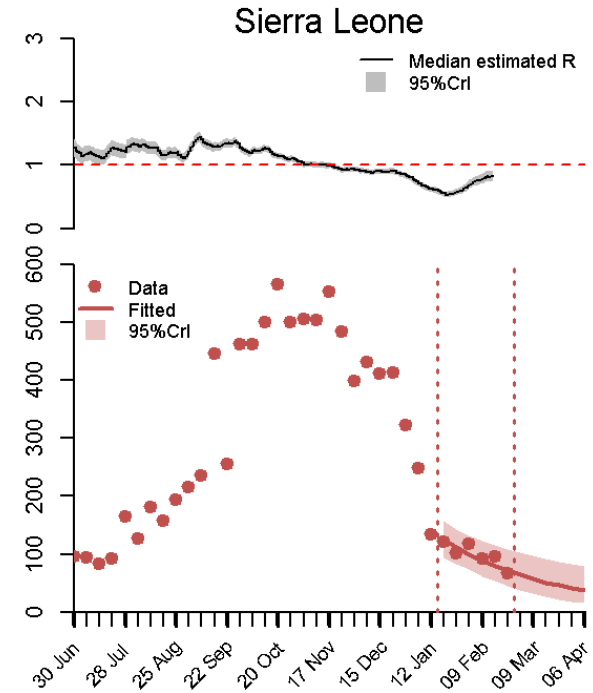
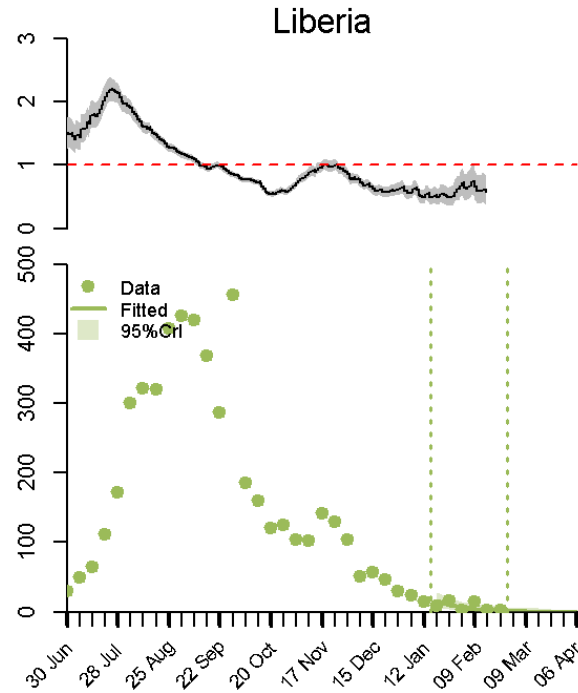
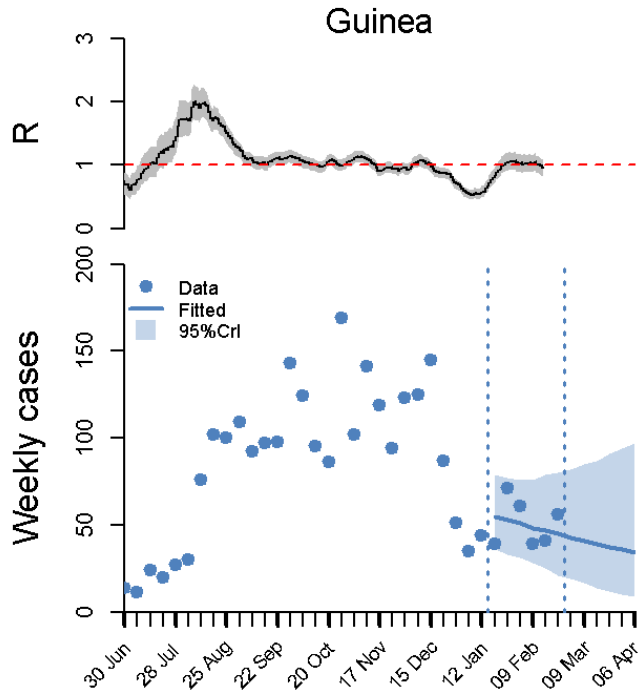


Important for advocacy, planning

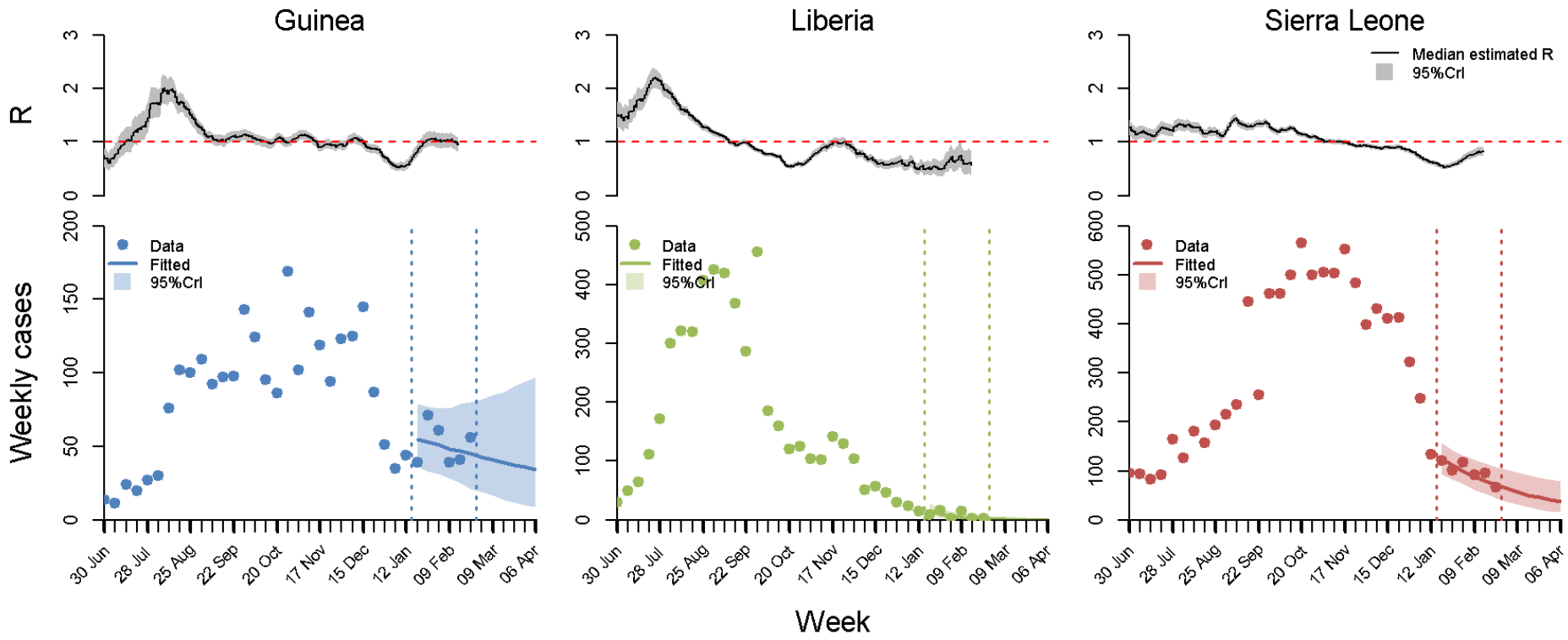
How quickly was the virus spreading?

Imperial College
London

March 2015



Week



	Guinea	Liberia	Sierra-Leone
	0.93 (0.77 ; 1.09)	0.43 (0.26 ; 0.68)	0.82 (0.74 ; 0.91)
Time to extinction	> 1 year (2015-07-16, > 1 year)	2015-03-22 (2015-02-18, 2015-06-12)	2015-11-22 (2015-07-13, > 1 year)

Implemented in a R package available in Recon website
(projection)

R Epidemics Consortium - Chromium

www.repidemicsconsortium.org

R Epidemics Consortium

ABOUT RECON NEWS PEOPLE PROJECTS RESOURCES EVENTS FORUM

RECON

The **R Epidemics Consortium (RECON)** assembles a group of international experts in infectious disease modelling, Public Health, and software development to create the next generation of analysis tools for disease outbreak response using the R software. This includes packages specifically designed for handling, visualising, and analysing outbreak data using cutting-edge statistical methods, as well as more general-purpose tools for data cleaning, versioning, and encryption.

Our approach aims to fulfil three key aspects:

- *Efficiency*: our tools can be used in real time to improve situation awareness and inform intervention strategies.
- *Reliability*: our tools are thoroughly and constantly tested using professional software development methods.

Released projects and packages

These projects are in a usable form. Packages have been developed following RECON's standards, are fully functional, documented and tested, and have been released on CRAN.

discrete
Discretized probability distributions.

earlyR
Estimation of infectiousness in the early stage of an outbreak.

epicontacts
Handling, visualisation and analysis of epidemiological contacts.

epitrix
Small helpers and tricks for epidemics analysis.

incidence
Computation, handling, visualisation and simple modelling of Incidence.

aweeek
Convert dates to arbitrary week definitions including epiweek, isoweek, and more.

outbreaker2
Modular framework for outbreak reconstruction.

outbreaks
Collection of outbreak data.

projections
Projections of future incidence.

RECON learn
Open training platform for epidemics analysis.

epiflows
Visualisation and analysis of passenger flows.

Implemented in a R package available in Recon website

projections 0.3.1 [Home](#) [Reference](#) [Changelog](#)

Welcome to the *projections* package!

This package uses data on *daily incidence*, the *serial interval* (tir *number* to simulate plausible epidemic trajectories and project fit follows a Poisson process determined by a daily infectiousness,

$$\lambda_t =$$

where $w()$ is the probability mass function (PMF) of the serial in

Installing the package

To install the current stable, CRAN version of the package, type:

```
install.packages("projections")
```

To benefit from the latest features and bug fixes, install the deve

```
devtools::install_github("reconhub/projections")
```

Note that this requires the package *devtools* installed.

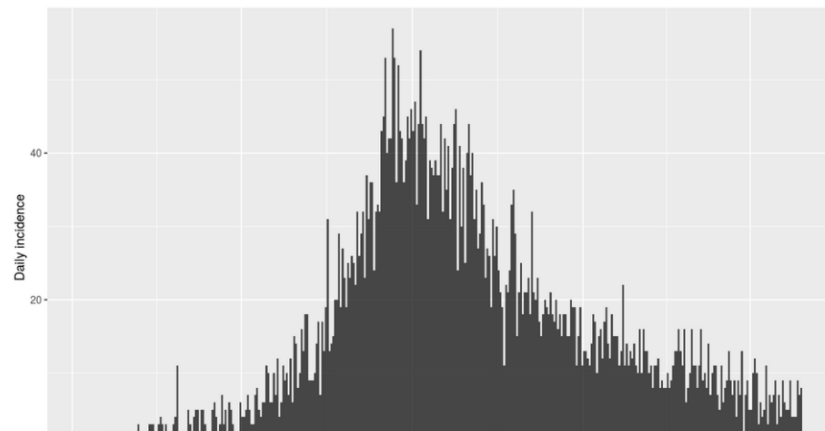
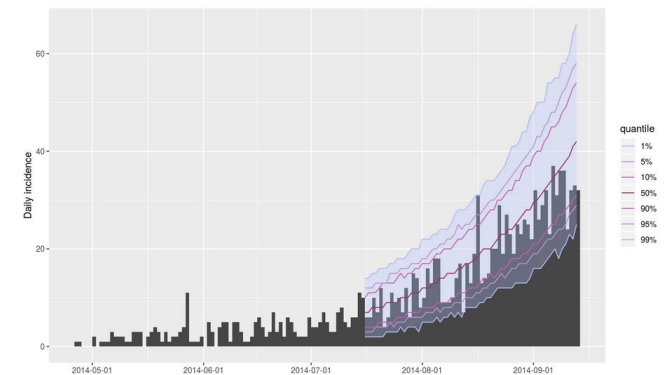
projections 0.3.1 [Home](#) [Reference](#) [Changelog](#)

Worked example

In the following, we project future incidence using a simulate Ebola ou

```
library(outbreaks)
library(incidence)

onset <- ebola_sim$linelist$date_of_onset
i <- incidence(onset)
plot(i) # full outbreak
```



From projections to forecasting?

Can we say more about the determinants of Ebola dynamics?

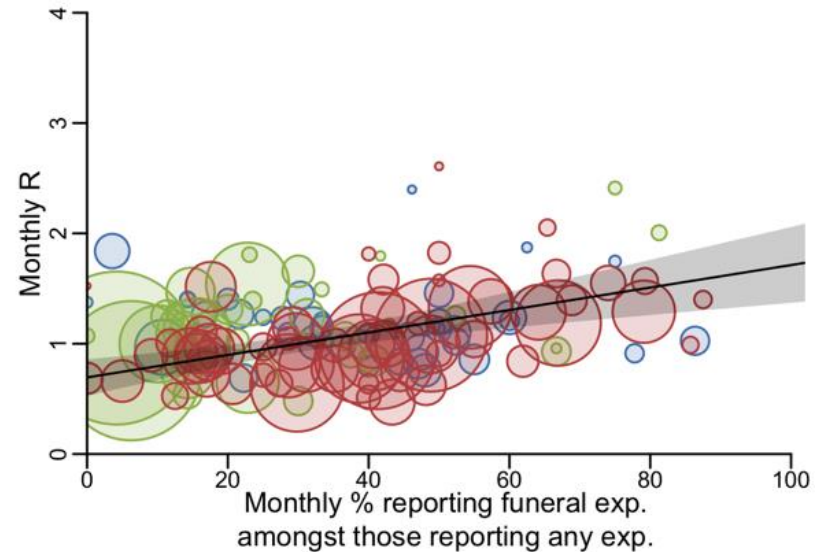
Exposure patterns driving Ebola transmission in West Africa
International Ebola Response Team (2016), *PLoS Medicine*

From projections to forecasting?

Can we say more about the determinants of Ebola dynamics?

Reproduction number for a given month was correlated with:

- % of individuals reporting funeral exposure (positive correlation)

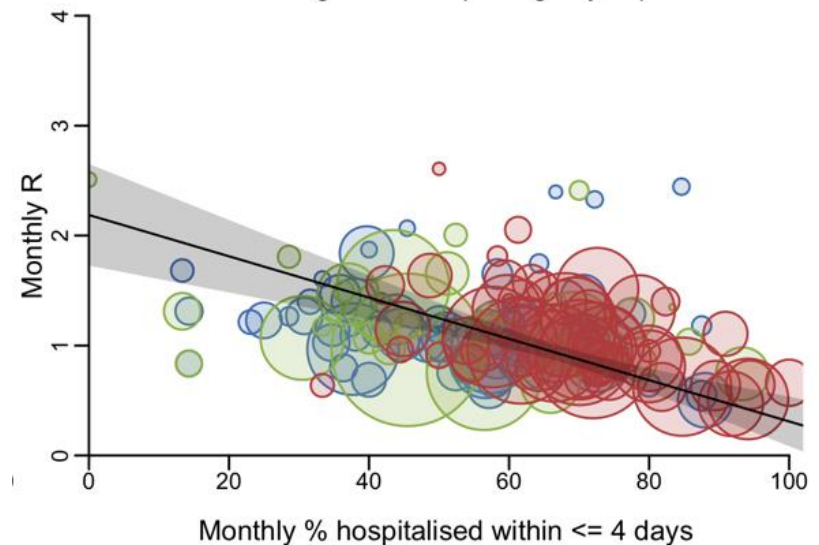
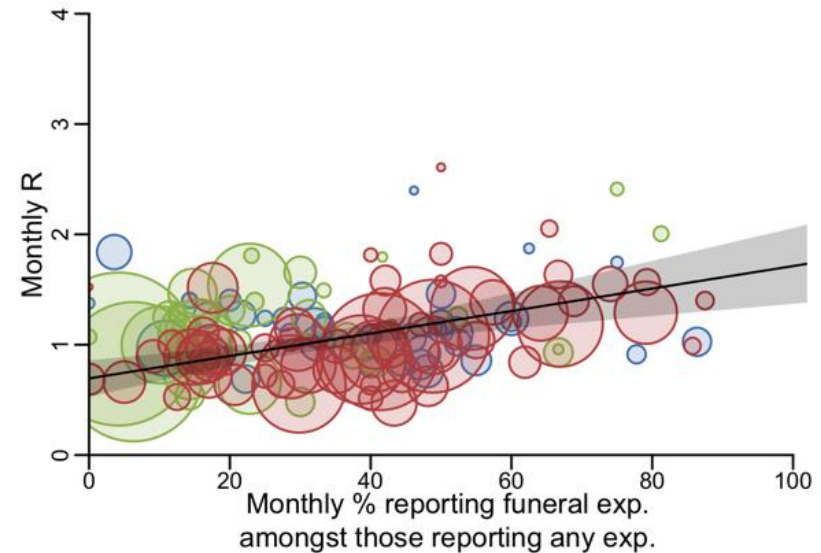


From projections to forecasting?

Can we say more about the determinants of Ebola dynamics?

Reproduction number for a given month was correlated with:

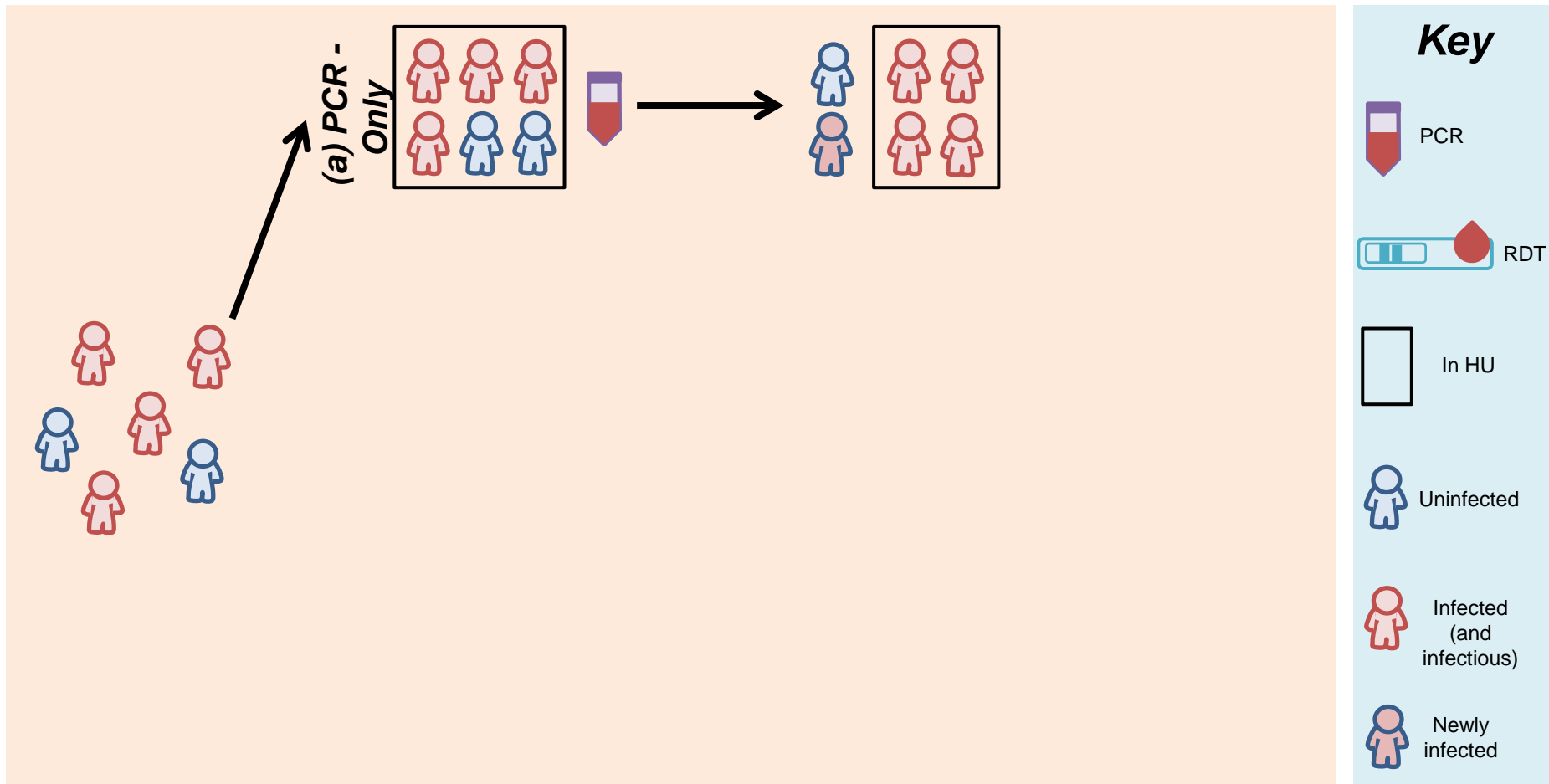
- % of individuals reporting funeral exposure (positive correlation)
- % of individuals hospitalised within 4 days (negative correlation)



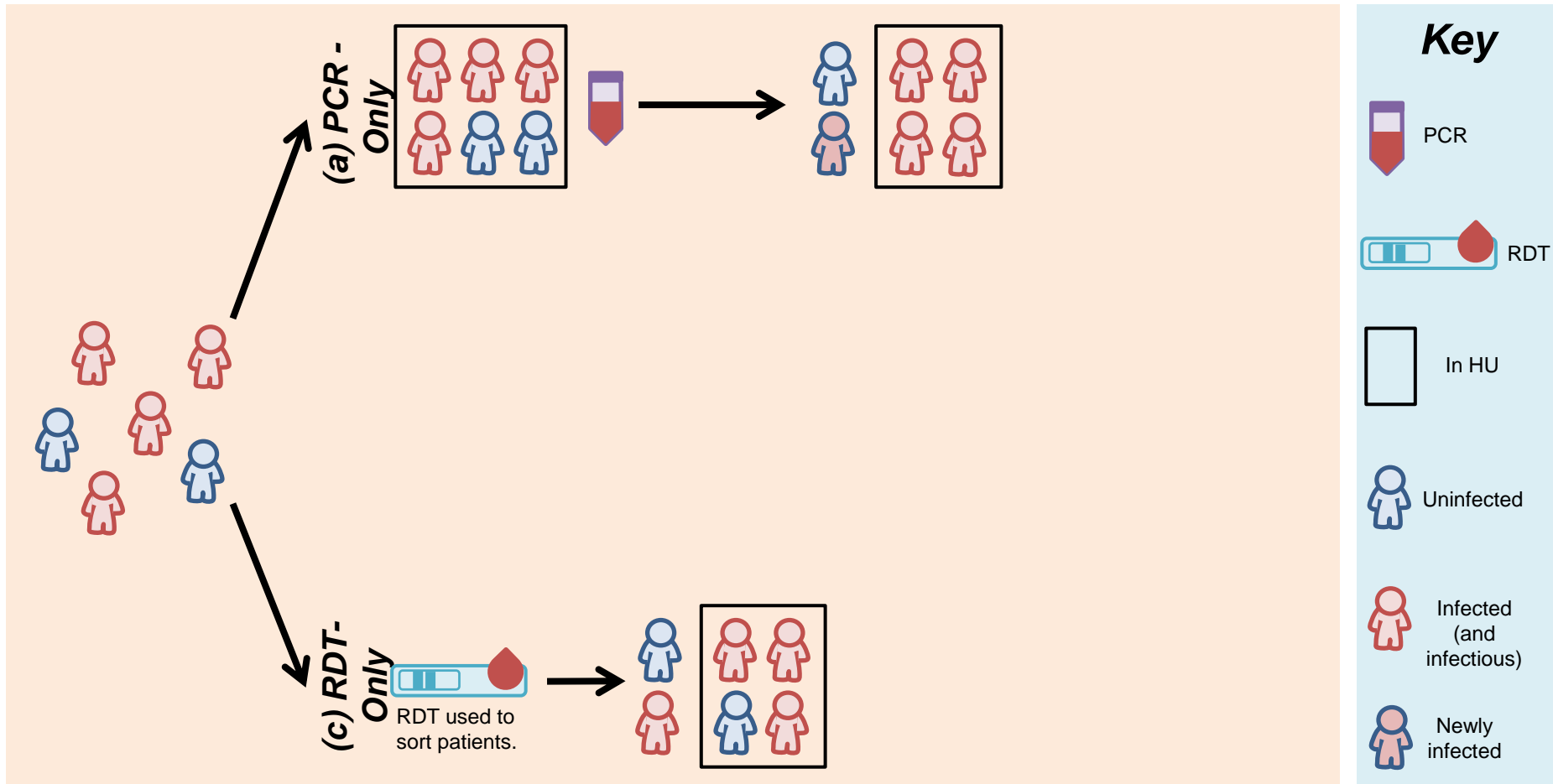
From projections to predictions?

Can we make predictions if conditions were different?

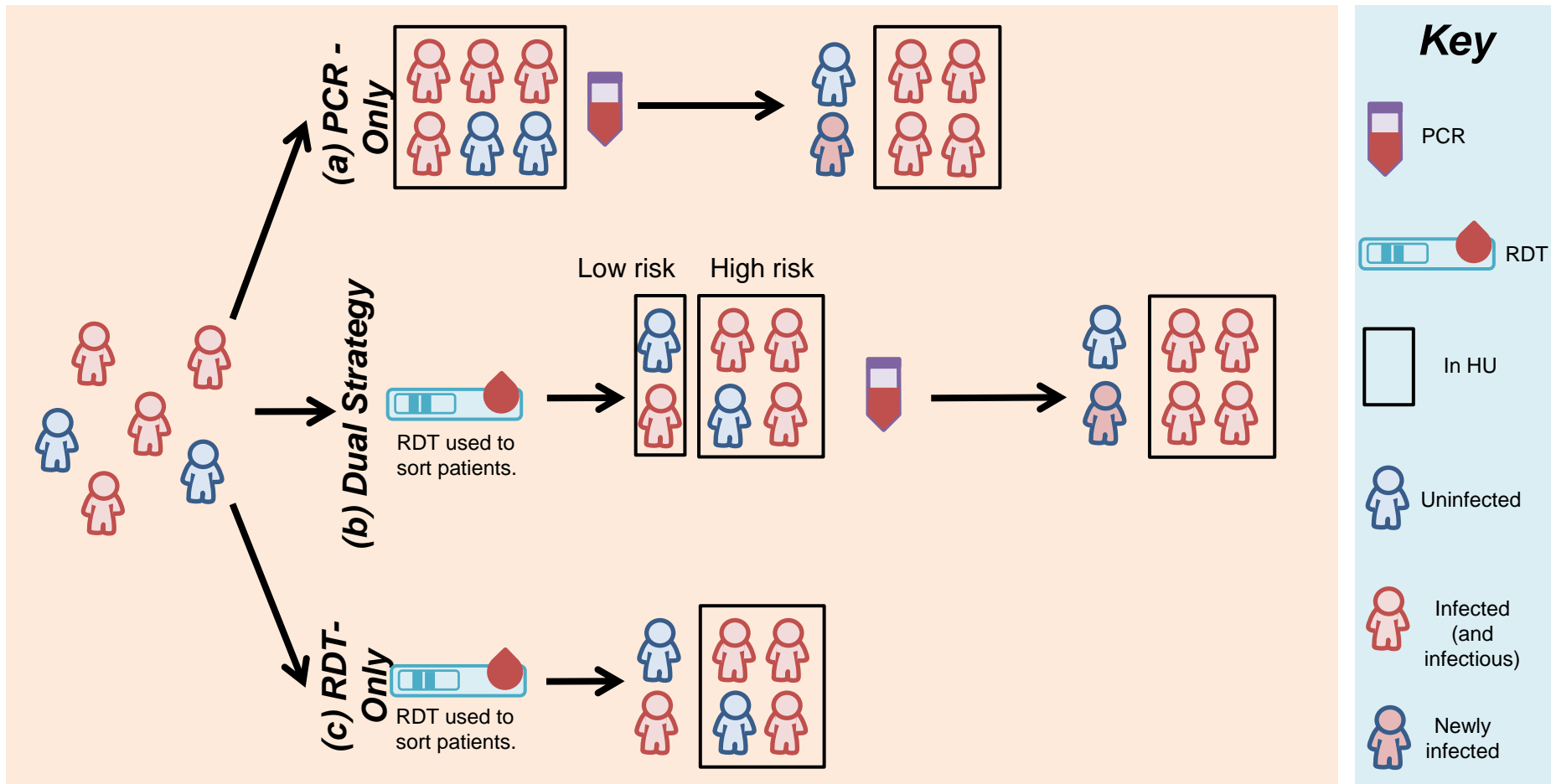
From projections to predictions?



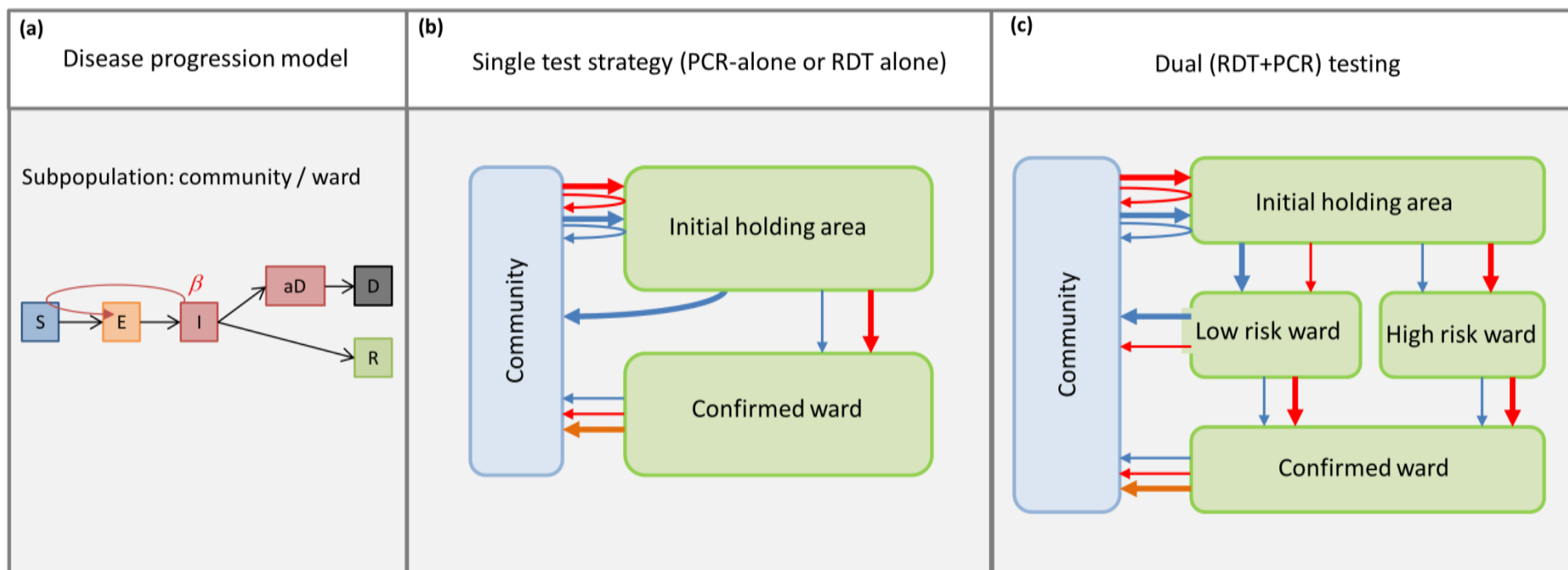
From projections to predictions?



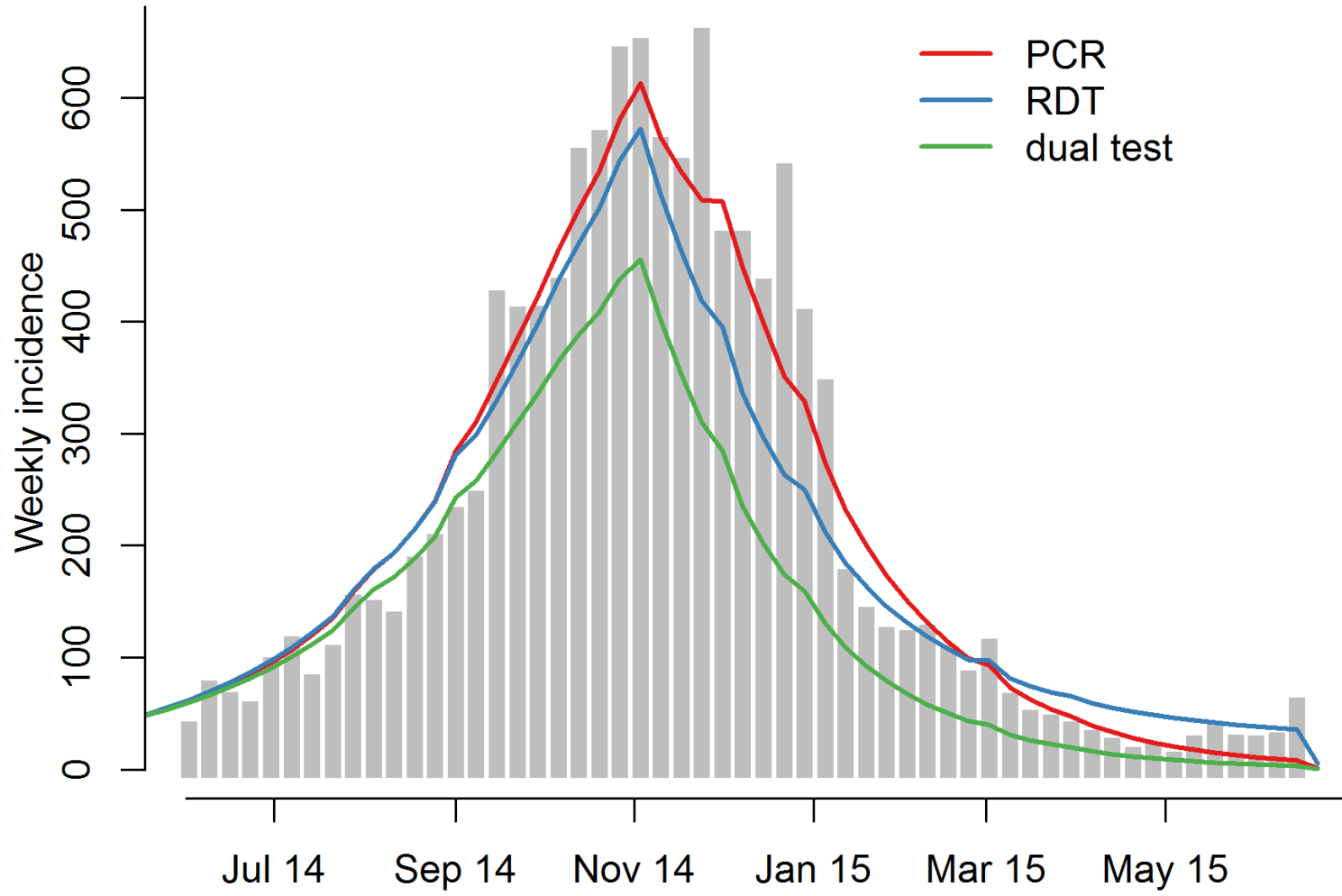
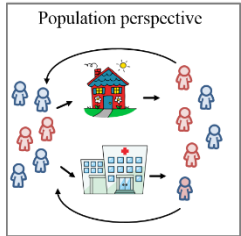
From projections to predictions?



From projections to predictions?



From projections to predictions?



Caveats for projections

- When using projections, things to consider:
 - Caveats linked to estimation of transmissibility (e.g. estimation issues if level reporting changes or delay in reporting)
 - Assume constant transmissibility in the future – to be used for short term projections (few serial intervals)
 - Be aware of the importance of accounting for
 - Delay in reporting
 - Uncertainty in current situation before projecting in the future (nowcasting)
 - Heterogeneity in transmission

Caveats for projections

- Heterogeneity in transmission

and heterogeneity in transmission



The cases of Amoy garden:

- over 300 cases
- Concentrated in 4 blocks
- Required quarantine
- Linked to drainage system

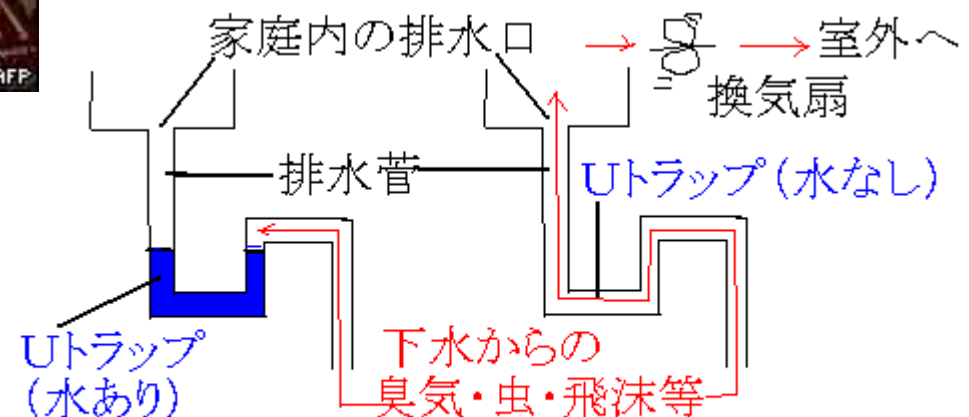


図3. Uトラップ (水なし) からの下水飛沫の侵入

and heterogeneity in transmission

SARS and heterogeneity in transmission

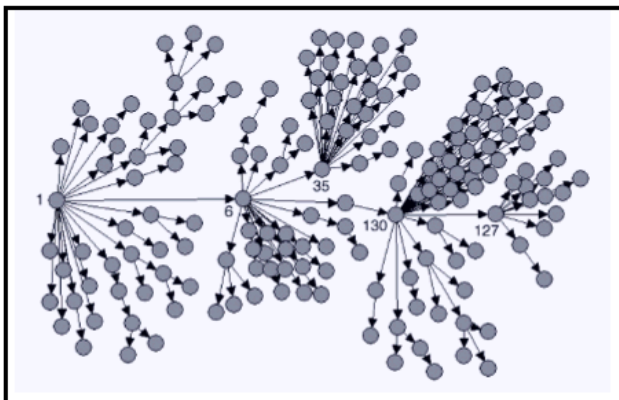
Reproduction number:

The number of cases one case generates on average over the course of its infectious period



Contagion

FIGURE 2. Probable cases of severe acute respiratory syndrome, by reported source of infection* — Singapore, February 25–April 30, 2003



Typically require detailed investigation

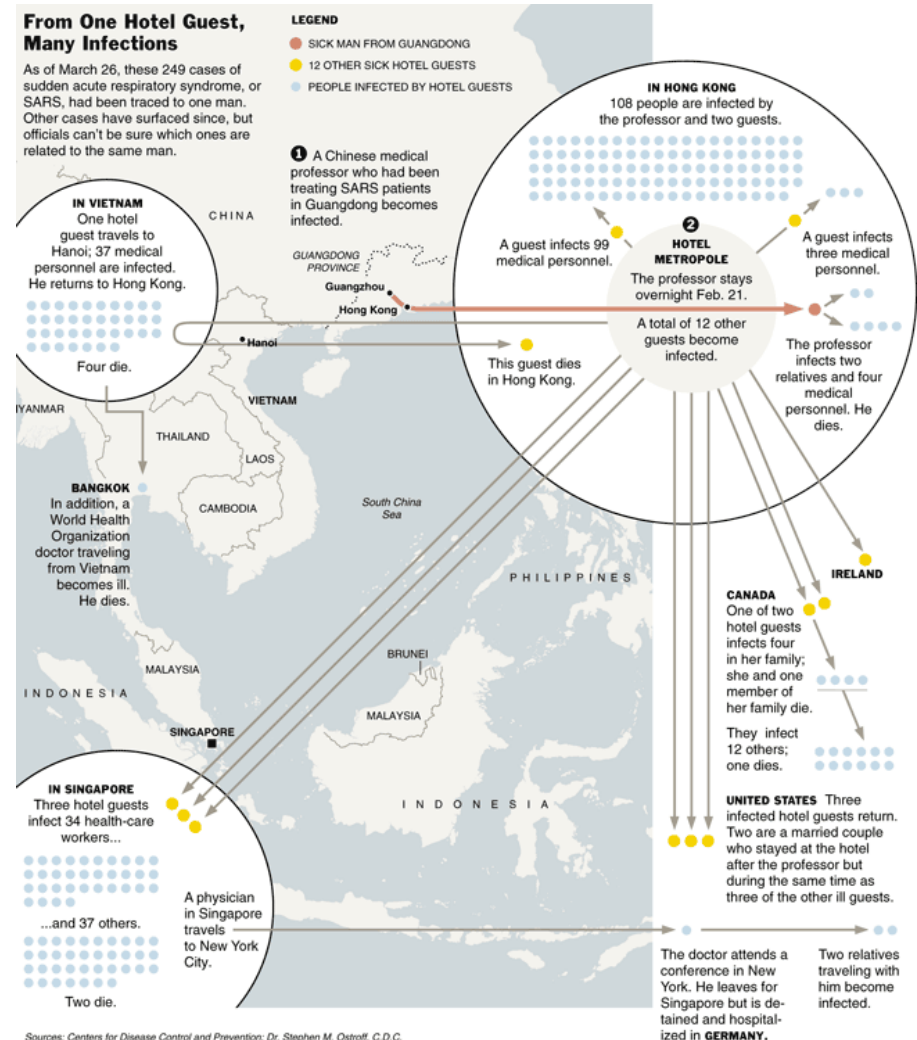
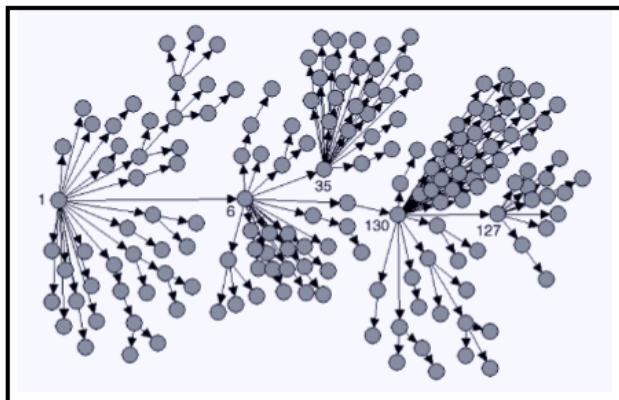
and heterogeneity in transmission

SARS and heterogeneity in transmission

Reproduction number:

The number of cases one case generates on average over the course of its infectious period,
BUT...

FIGURE 2. Probable cases of severe acute respiratory syndrome, by reported source of infection* — Singapore, February 25–April 30, 2003



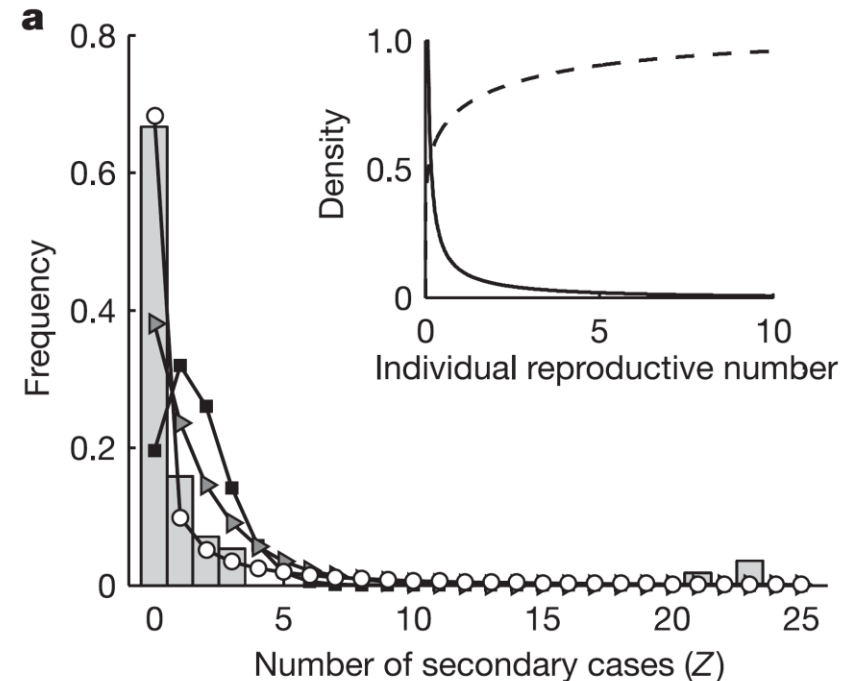
and heterogeneity in transmission

Simplest case, assumes:

- Number of secondary cases for each infectious individual follows a Poisson distribution (offspring distribution)
- Same mean for everyone (R)

Increased heterogeneity, assumes:

- Individual 'offspring distribution' is still Poisson
- Individual R is gamma distributed (not the same for everyone)
- **Negative binomial offspring distribution for the population**



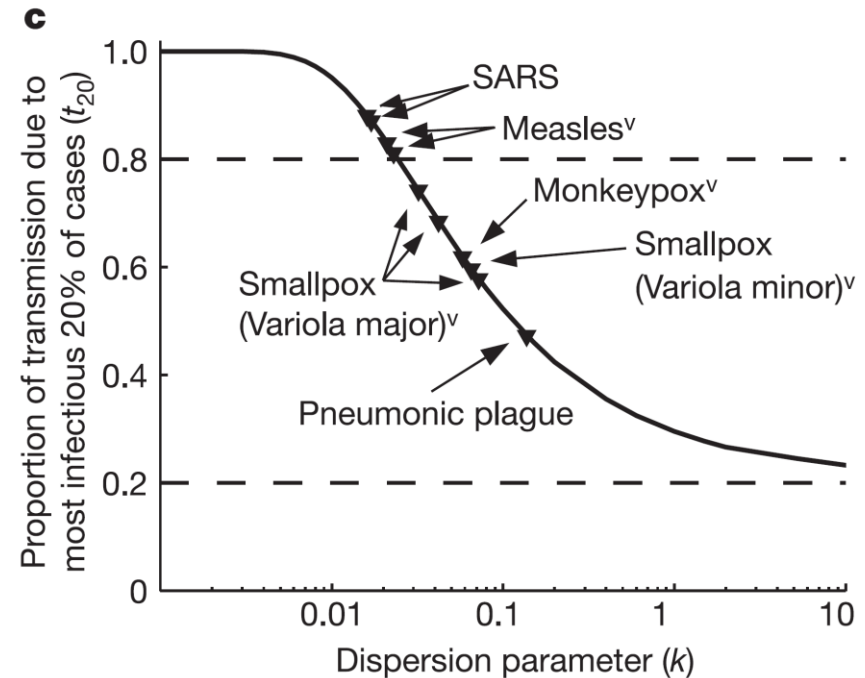
and heterogeneity in transmission

Simplest case, assumes:

- Number of secondary cases for each infectious individual follows a Poisson distribution (offspring distribution)
- Same mean for everyone (R)

Increased heterogeneity, assumes

- Individual 'offspring distribution' is still Poisson
- Individual R is gamma distributed (not the same for everyone)
- **Negative binomial offspring distribution for the population**



and heterogeneity in transmission

Simplest case, assumes:

- Number of secondary cases for each infectious individual follows a Poisson distribution (offspring distribution)
- Same mean for everyone (R)

Implications for Projections

$$I_t = \mathcal{P} \left(R_t \sum_{s=1}^t I_{t-s} w_{t-s} \right)$$

Increased heterogeneity, assumes:

- Individual 'offspring distribution' is still Poisson
- Individual R is gamma distributed (not the same for everyone)
- Negative binomial offspring distribution for the population

$$I_t = \text{NB} \left(R_t \sum_{s=1}^t I_{t-s} w_{t-s}, \delta \right)$$

and heterogeneity in transmission

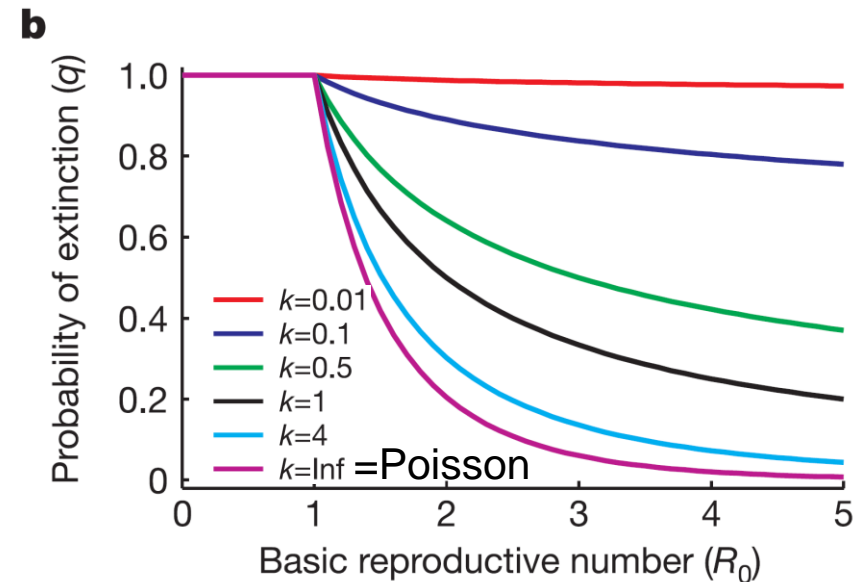
Simplest case, assumes:

- Number of secondary cases for each infectious individual follows a Poisson distribution (offspring distribution)
- Same mean for everyone (R)

Increased heterogeneity, assumes:

- Individual 'offspring distribution' is still Poisson
- Individual R is gamma distributed (not the same for everyone)
- **Negative binomial offspring distribution for the population**

Implications for
outbreak extinctions



Thank you!