

RECON

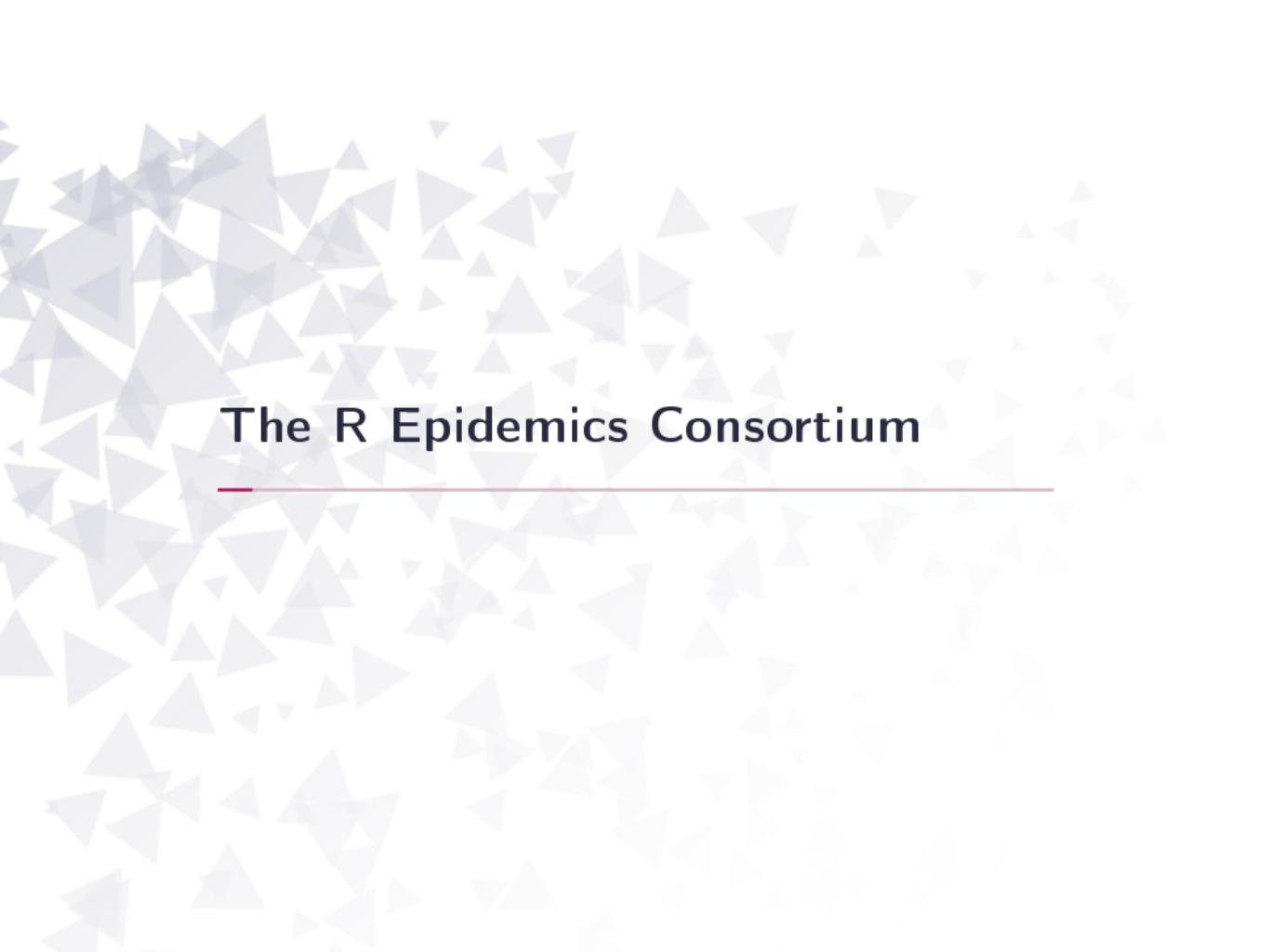


Operational modelling for outbreak response

Thibaut Jombart

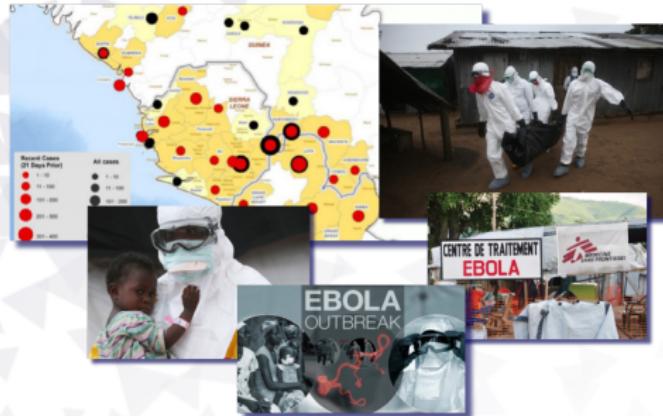
13th February 2018

Imperial College London
MRC Centre for Outbreak Analysis and Modelling



The R Epidemics Consortium

Lessons learnt from the Ebola response



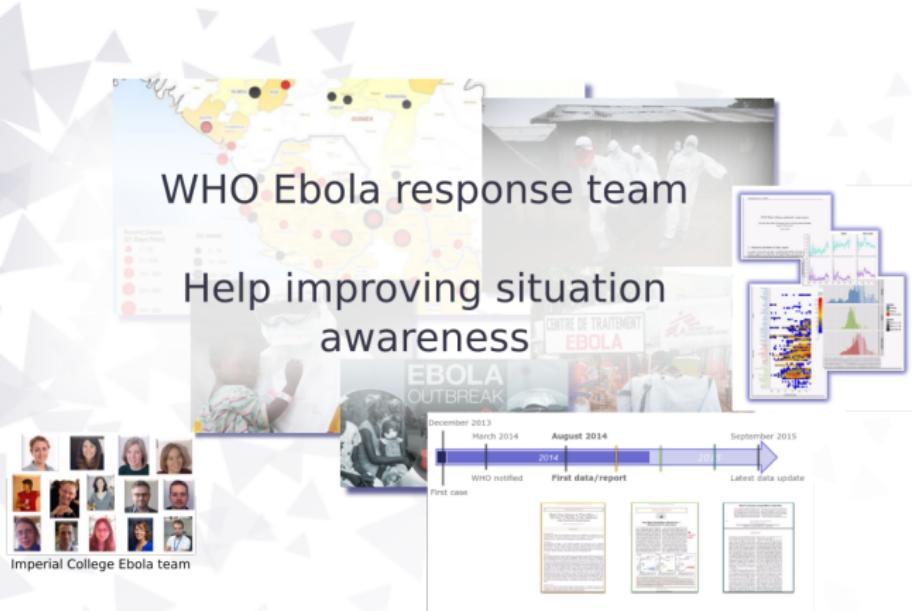
Lessons learnt from the Ebola response



Lessons learnt from the Ebola response



Lessons learnt from the Ebola response



Most statistical/modelling tools for situation awareness were missing.

Who do we need to develop these tools?



Who do we need to develop these tools?

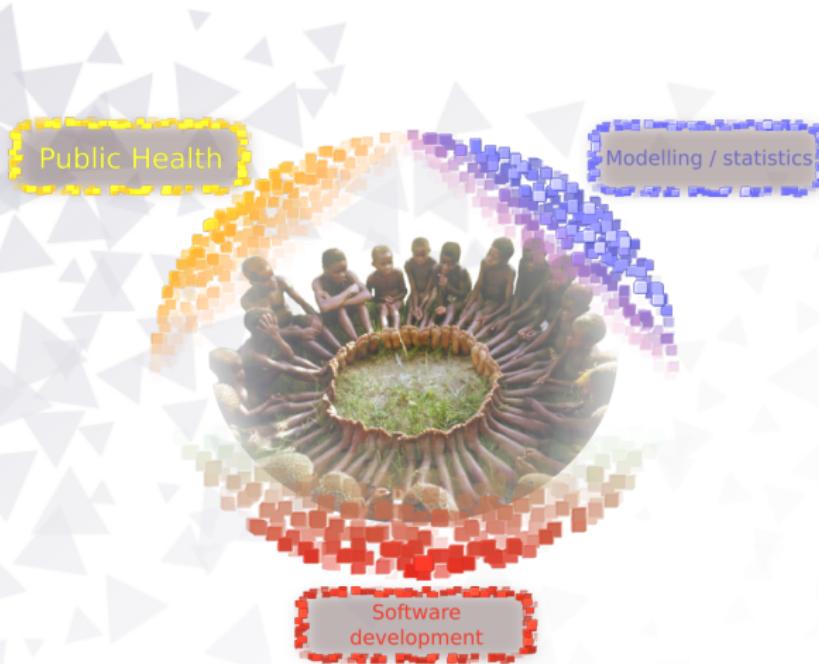
Public Health



Who do we need to develop these tools?



Who do we need to develop these tools?



From a hack to a pack



Hackout 3, summer 2016, Berkeley

From a hack to a pack



Hackout 3, summer 2016, Berkeley



From a hack to a pack



Hackout 3, summer 2016, Berkeley



RECON

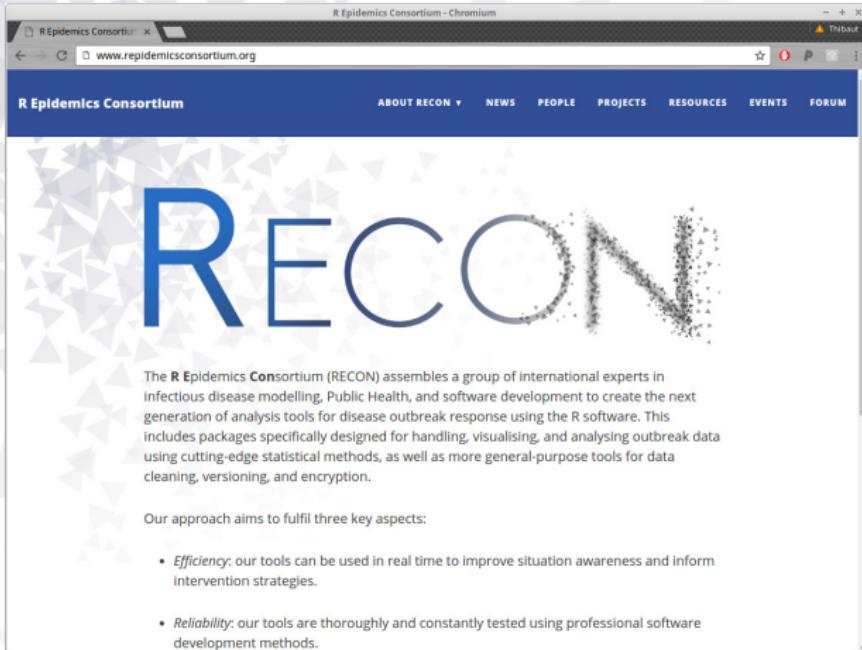
[The R Epidemics Consortium]

From a hack to a pack



RECON: the R Epidemics Consortium

A taskforce to build a new generation of outbreak response tools in .



The screenshot shows a web browser window for the "R Epidemics Consortium - Chromium" tab, displaying the URL "www.repidemicsconsortium.org". The page has a dark blue header with the "R Epidemics Consortium" logo and navigation links for "ABOUT RECON", "NEWS", "PEOPLE", "PROJECTS", "RESOURCES", "EVENTS", and "FORUM". Below the header is a large, stylized "RECON" logo where each letter is composed of small, dark grey dots. A descriptive paragraph follows:

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.

www.repidemicsconsortium.org

RECON

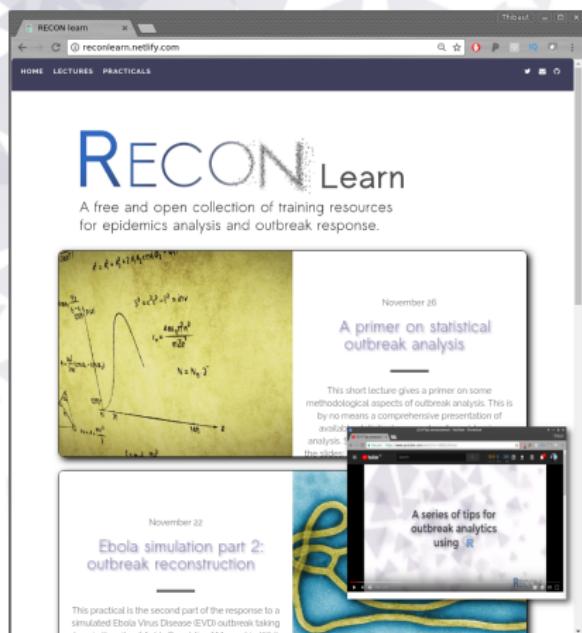
www.repidemicsconsortium.org

- started 6th September 2016
- ~70 members
- 20 countries, > 40 institutions
- ~ 9 packages released, 15 under development
- public forum, blog, online resources

RECON: activities

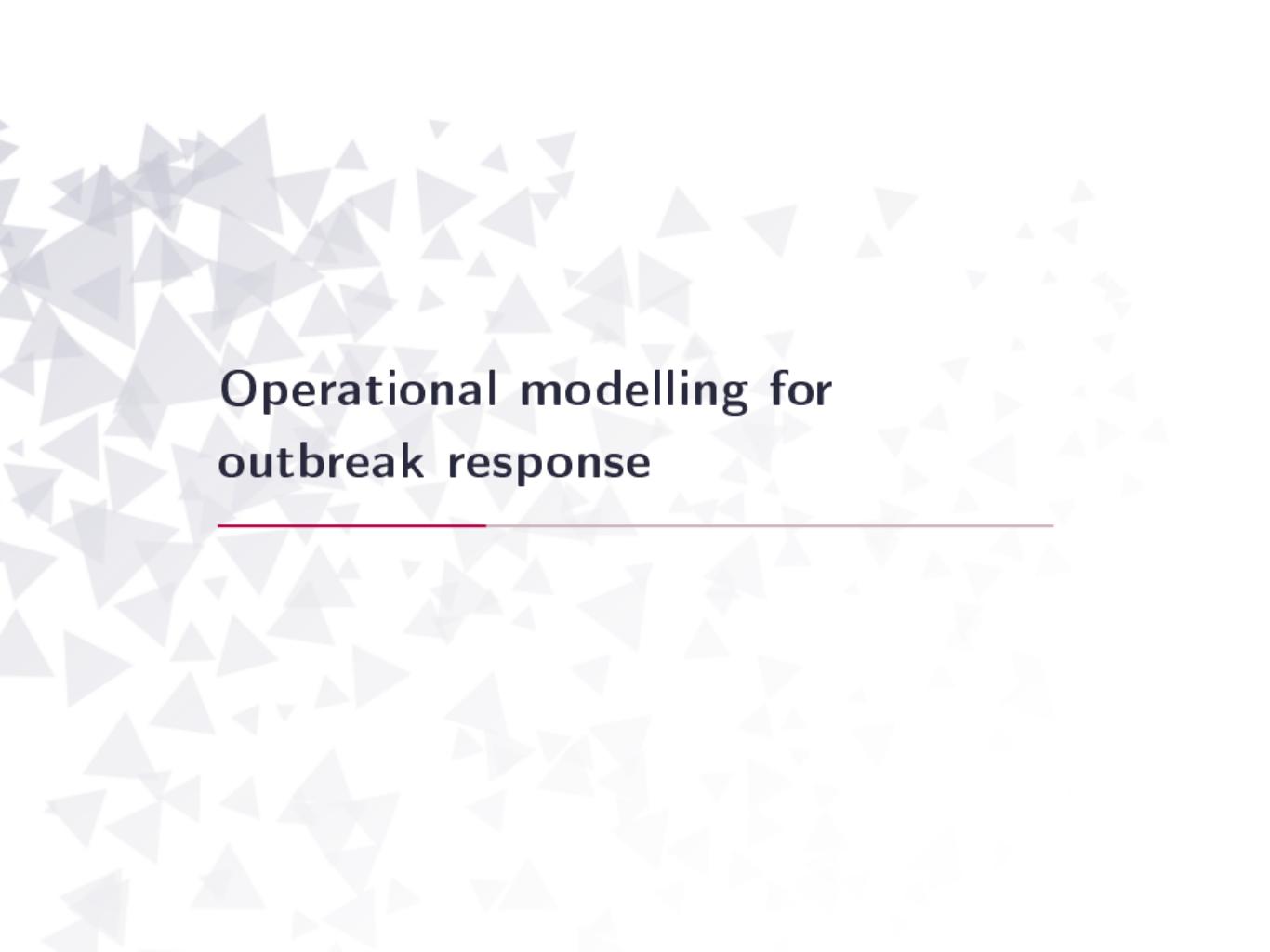
- **software development**: tools for real-time outbreak analytics (from data cleaning to modelling)
- **community support**: best practices, forums, hackathons, meetings
- **outbreak response**: deployment to the field
- **training**: RECON learn, podcast

RECON learn: training resources for epidemics analysis



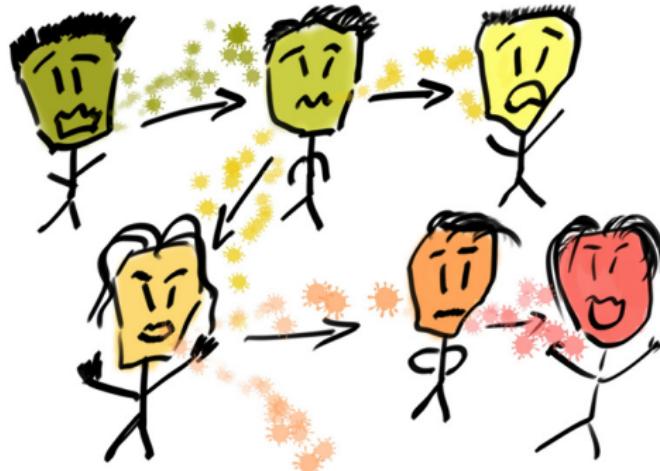
<https://reconlearn.netlify.com>

- repository for free, **open** training material
- lectures, practicals, case studies, code gists
- emphasis on **community contributions**
- podcast "**R Tips**" on Youtube
- **workshops** in 2018:
CDC, MSF, WHO, EPIET
Alumni Network, ...



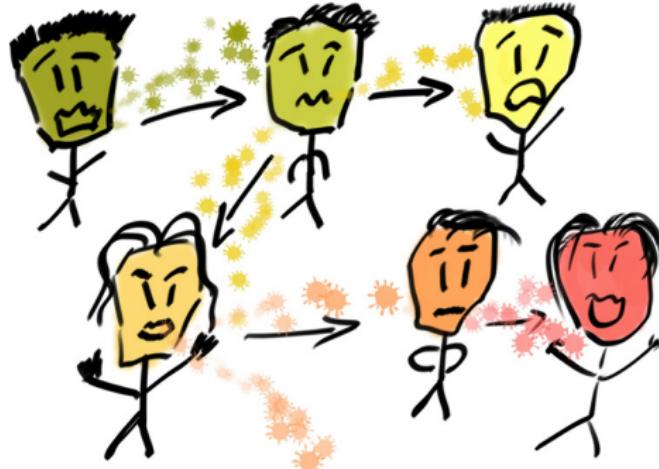
Operational modelling for outbreak response

Using WGS to infer who infected whom



Mutations accumulate along transmission chains.

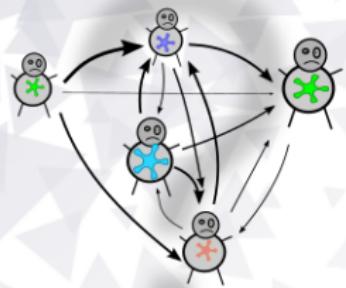
Using WGS to infer who infected whom



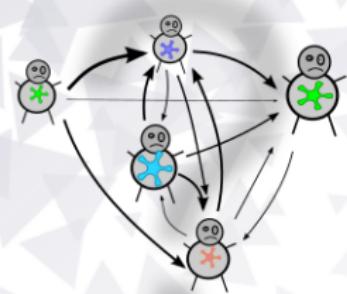
Mutations accumulate along transmission chains.

Can be used to reconstruct transmission trees.

Who infects whom? Many answers for a single question



Who infects whom? Many answers for a single question

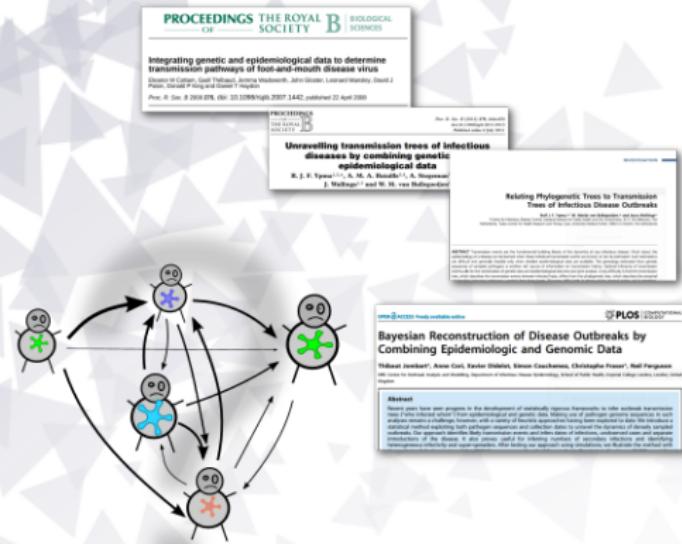


Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data

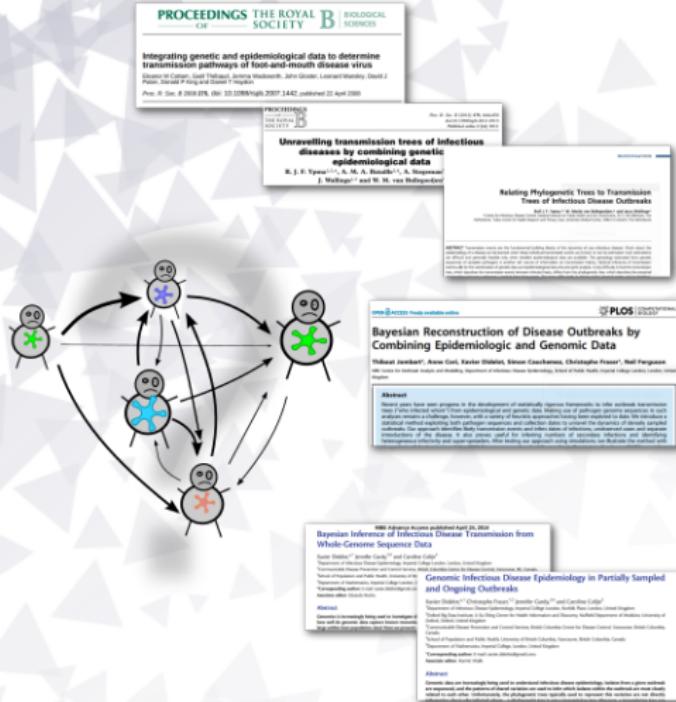
R. J. E. Verwoerd¹, S. M. A. Bouwink¹, J. A. Nijhuis², J. Hulstegen² & W. M. van Harten¹



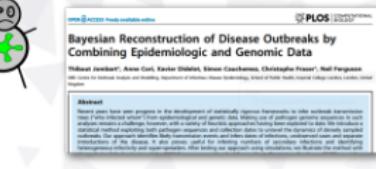
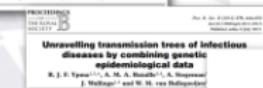
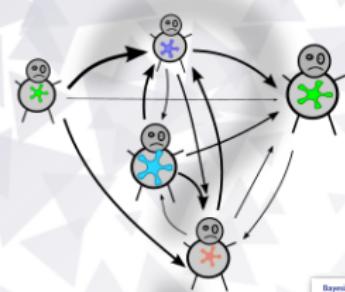
Who infects whom? Many answers for a single question



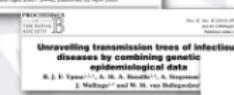
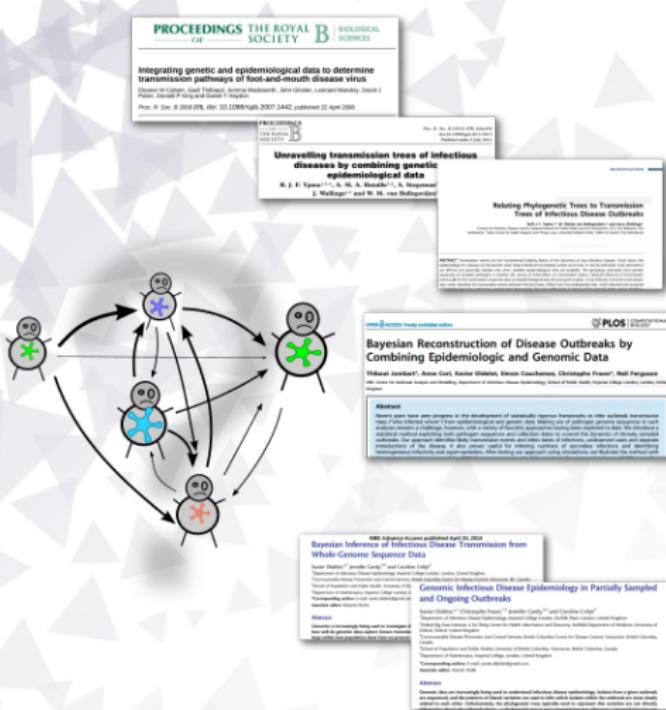
Who infects whom? Many answers for a single question



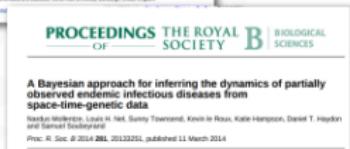
Who infects whom? Many answers for a single question



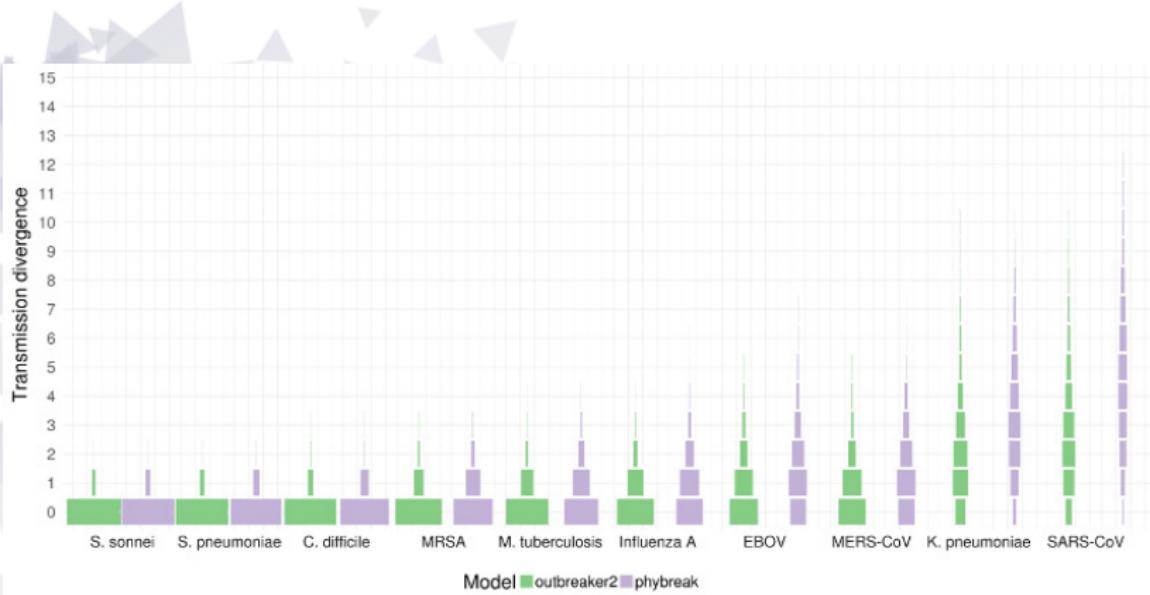
Who infects whom? Many answers for a single question



Methods heavily
rely on whole genome
sequence data



How informative are whole genome sequences?



Campbell *et al.* 2018 PLoS Pathogens

Insufficient diversity for most diseases.

Evidence synthesis approach to outbreak reconstruction



Combine different data to shrink the set of plausible trees.

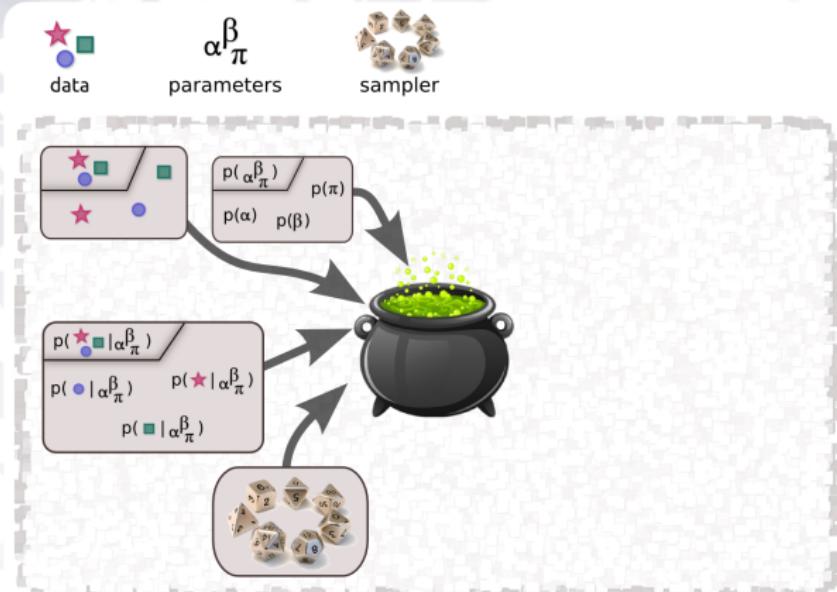
outbreaker2: evidence synthesis framework for outbreak reconstruction

Modularity: customise data, prior, likelihood, MCMC.



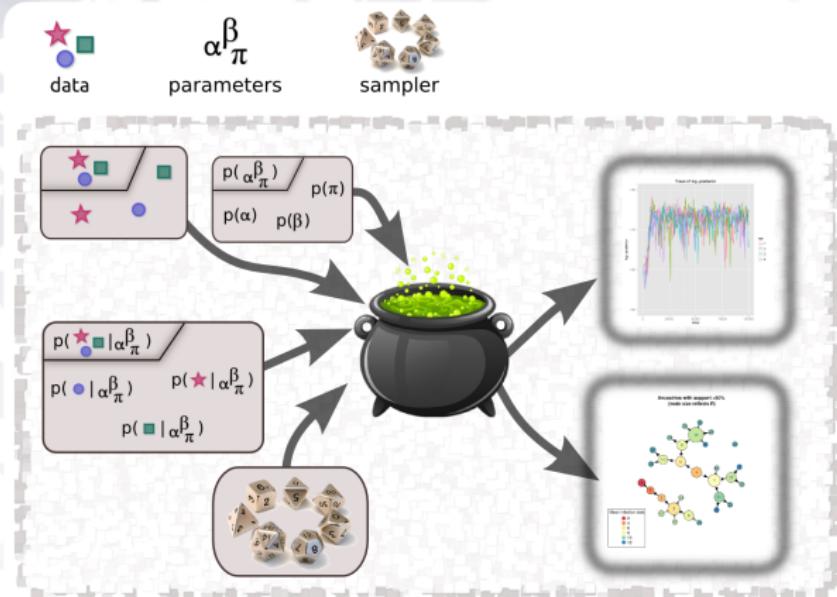
outbreaker2: evidence synthesis framework for outbreak reconstruction

Modularity: customise data, prior, likelihood, MCMC.

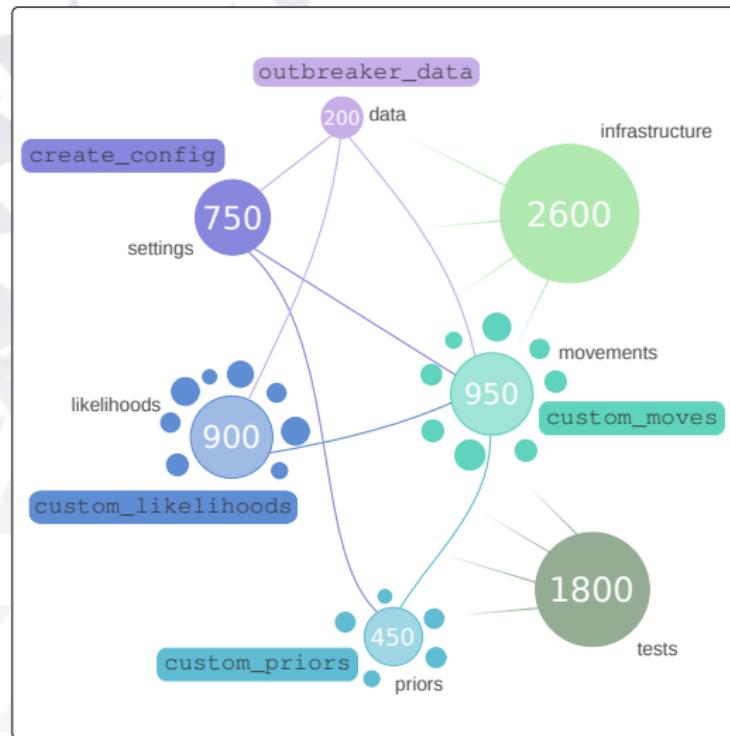


outbreaker2: evidence synthesis framework for outbreak reconstruction

Modularity: customise data, prior, likelihood, MCMC.



outbreaker2: a modular implementation



Example: implementing *TransPhylo* in *outbreaker2*

outbreaker likelihood

- $p(t, s|\alpha, T^{inf}, \kappa, \mu, \pi) = p(t|T^{inf})p(T^{inf}|\alpha, \kappa)p(s|\alpha, \kappa, \mu)p(\kappa|\pi)$
- i.e. *incubation* \times *generation time* \times *genetic (simple)* \times *missing cases*

Example: implementing *TransPhylo* in *outbreaker2*

outbreaker likelihood

- $p(t, s|\alpha, T^{inf}, \kappa, \mu, \pi) = p(t|T^{inf})p(T^{inf}|\alpha, \kappa)p(s|\alpha, \kappa, \mu)p(\kappa|\pi)$
- i.e. *incubation* \times *generation time* \times *genetic (simple)* \times *missing cases*

TransPhylo likelihood

- $p(G|\beta, \gamma, N_{eg}, \alpha) = p(G|N_{eg}, \alpha) \times p(\alpha|\beta, \gamma)$
- i.e. *phylogeny (coalescent)* \times *SIR*

Example: implementing *TransPhylo* in *outbreaker2*

outbreaker likelihood

- $p(t, s|\alpha, T^{inf}, \kappa, \mu, \pi) = p(t|T^{inf})p(T^{inf}|\alpha, \kappa)p(s|\alpha, \kappa, \mu)p(\kappa|\pi)$
- i.e. *incubation* x *generation time* x *genetic (simple)* x *missing cases*

TransPhylo likelihood

- $p(G|\beta, \gamma, N_{eg}, \alpha) = p(G|N_{eg}, \alpha) \times p(\alpha|\beta, \gamma)$
- i.e. *phylogeny (coalescent)* x *SIR*

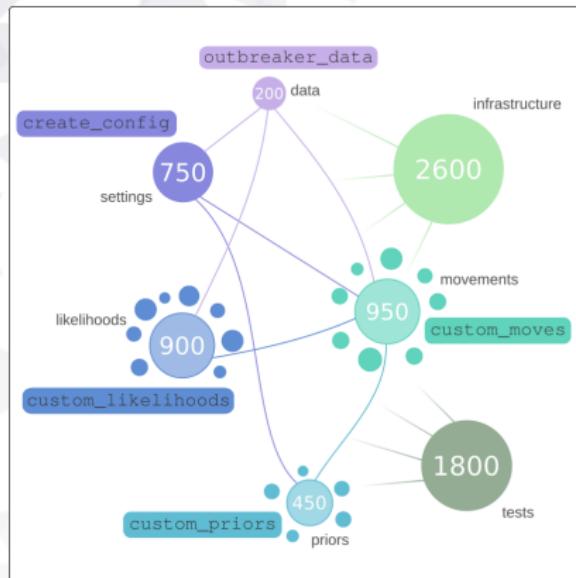
Can we combine the two models?

o2mod.transphylo: a new transmission model

$$p(t, G | \alpha, T^{inf}, \kappa, \pi, N_{eg}) =$$

$$p(t | T^{inf}) p(T^{inf} | \alpha, \kappa) p(G | N_{eg}, \alpha) p(\kappa | \pi)$$

i.e. *incubation x generation time x coalescent x missing cases*

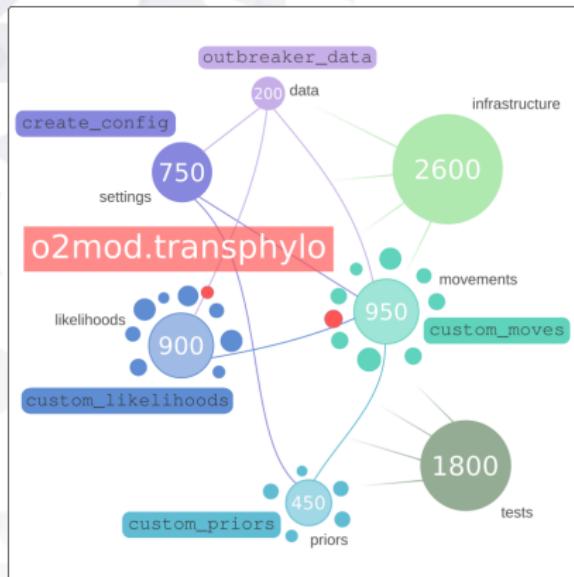


o2mod.transphylo: a new transmission model

$$p(t, G | \alpha, T^{inf}, \kappa, \pi, N_{eg}) =$$

$$p(t | T^{inf}) p(T^{inf} | \alpha, \kappa) p(G | N_{eg}, \alpha) p(\kappa | \pi)$$

i.e. *incubation x generation time x coalescent x missing cases*



o2mod.TransPhylo: TransPhylo module for outbreaker2

[implementation by Xavier Didelot and Finlay Campbell]

3 Custom likelihood

In order to calculate the likelihood under the TransPhylo model, we need to (i) extract the transmission tree from the outbreaker2 parameter, (ii) combine this transmission tree with the phylogenetic tree to form a colored tree, and (iii) calculate the likelihood of this colored tree. Step (i) is easy since transmission tree are encoded almost in the same way in TransPhylo and outbreaker2. For step (ii) we have to write the `combine` function which is tedious but not especially interesting (this function is included in this Rnw file but its code is not shown in the pdf). For step (iii) we only need to call the appropriate function of the TransPhylo package which is `probTreeGivenTree`. During step (ii) messages can arise indicating that the transmission tree and phylogenetic tree are in fact incompatible, in which case the likelihood is returned as -Inf.

```
lik_TransPhylo <- function(data, param) {
  ttree <- list(ttree = chind(param$tt_inf, data$dates, param$alpha),
    nam = data$pmtree$nam)
  ttree$ttree[which(is.na(ttree$ttree[,3])),3] <- 0
  txt <- capture.output(ctree <- combine(ttree, data$pmtree))
  if (length(txt)>0) {
    prob <- probTreeGivenTree(ctree, neg = 365 * 0.25)
  } else {
    prob <- -Inf
  }
  return(prob)
}

## Function (data, param, i = NULL, custom_functions = NULL)

new_move_tinf <- function(param, data, list_custom_ll = new_model) {
  for (i in 1:datas$N) {
    current_ll <- api$pp_ll_all(data, param, i = NULL, list_custom_ll)
    modif <- sample(c(-100:-1,1:100), 1)
    param$ll.inf[i] <- param$ll.inf[i] + modif
    new_ll <- api$pp_ll_all(data, param = NULL, list_custom_ll)
    if (log(modif) > 0 & new_ll > current_ll) {
      param$ll.inf[i] <- param$ll.inf[i] - modif
    }
  }
  return(param)
}

new_moves <- custom_moves(t.inf = new_move_tinf)
new_moves

## 
## //////////////////////////////////////////////////////////////////// outbreaker movement functions ////
## 
## classes: outbreaker_moves list
## number of items: 8
## 
## //////////////////////////////////////////////////////////////////// movement functions ////
## 
```

o2mod.TransPhylo: TransPhylo module for outbreaker2

[implementation by Xavier Didelot and Finlay Campbell]

3 Custom likelihood

In order to calculate the likelihood under the TransPhylo model, we need to (i) extract the transmission tree from the outbreaker2 parameter, (ii) combine this transmission tree with the phylogenetic tree to form a colored tree, and (iii) calculate the likelihood of this colored tree. Step (i) is easy since transmission tree are encoded almost in the same way in TransPhylo and outbreaker2. For step (ii) we have to write the `combine` function which is tedious but not especially interesting (this function is included in this Rnw file but its code is not shown in the pdf). For step (iii) we only need to call the appropriate function of the TransPhylo package which is `probTreeGivenTree`. During step (ii) messages can arise indicating that the transmission tree and phylogenetic tree are in fact incompatible, in which case the likelihood is returned as `-Inf`.

```
lik_TransPhylo <- function(data, param) {
  ttree <- list(ttree = chind(param$t_inf, data$dates, param$alpha),
               nan = data$ptree$nan)
  ttree$ttree[which(is.na(ttree$ttree[,3]),3] <- 0
  txt <- capture.output(ctree <- combine(ttree,data$ptree))
  if (length(txt)==0) {
    prob <- probTreeGivenTree(ctree, neg = 365 * 0.25)
  } else {
    prob <- -Inf
  }
  return(prob)
}
```

likelihood

movement function

```
args(api$cpp_ll_all)

## function (data, paran, i = NULL, custom_functions = NULL)
## NULL

new_move_tinf <- function(paran, data, list_custom_ll = new_model) {
  for (i in 1:datas$N) {
    current_ll <- api$cpp_ll_all(data,paran, i = NULL, list_custom_ll)
    modif <- sample(c(-100:-1,1:100), 1)
    paran$ll.inf[i] <- paran$ll.inf[i] + modif
    new_ll <- api$cpp_ll_all(data,paran, i = NULL, list_custom_ll)
    if (log(modif)>0) > (new_ll > current_ll) {
      paran$ll.inf[i] <- paran$ll.inf[i] - modif
    }
  }
  return(paran)
}

new_moves <- custom_moves(v_inf = new_move_tinf)
new_moves

##
## /////////////////////////////////////////////////////////////////// outbreaker movement functions ///
## 
## class: outbreaker_moves list
## number of items: 8
## 
## /// movement functions ///
## End
```

o2mod.TransPhylo: TransPhylo module for outbreaker2

[implementation by Xavier Didelot and Finlay Campbell]

3 Custom likelihood

In order to calculate the likelihood under the TransPhylo model, we need to (i) extract the transmission tree from the outbreaker2 parameter, (ii) combine this transmission tree with the phylogenetic tree to form a colored tree, and (iii) calculate the likelihood of this colored tree. Step (i) is easy since transmission tree are encoded almost in the same way in TransPhylo and outbreaker2. For step (ii) we have to write the `combine` function which is tedious but not especially interesting (this function is included in this Rnw file but its code is not shown in the pdf). For step (iii) we only need to call the appropriate function of the TransPhylo package which is `probTreeGivenTree`. During step (ii) messages can arise indicating that the transmission tree and phylogenetic tree are in fact incompatible, in which case the likelihood is returned as `-Inf`.

```
lik_TransPhylo <- function(data, param) {
  ttree <- list(ttree = chind(param$tt_inf, data$dates, param$alpha),
    nam = data$pmtree$nam)
  ttree$ttree[which(is.na(ttree$ttree[,3])),3] <- 0
  txt <- capture.output(ctree <- combine(ttree, data$pmtree))
  if (length(txt)==0) {
    prob <- probTreeGivenTree(ctree, neg = 365 * 0.25)
  } else {
    prob <- -Inf
  }
  return(prob)
}
```

likelihood

Total: 25 lines of R
outbreaker2: 7,500 lines of R/C++
Code difference: 0.3%

movement function

```
args(api$cpp_ll_all)

## function (data, param, i = NULL, custom.functions = NULL)
## NULL

new_move_tinf <- function(param, data, list_custom_ll = new_model) {
  for (i in 1:datasize) {
    current_ll <- api$cpp_ll_all(data, param, i = NULL, list_custom_ll)
    modif <- sample(c(-100:-1, 1:100), 1)
    param$ll.inf[i] <- param$ll.inf[i] + modif
    new_ll <- api$cpp_ll_all(data, param, i = NULL, list_custom_ll)
    if (log(modif[i]) > (new_ll - current_ll)) {
      param$ll.inf[i] <- param$ll.inf[i] + modif
    }
  }
  return(param)
}

new_moves <- custom_moves(v_inf = new_move_tinf)
new_moves

## 
## //////////////////////////////////////////////////////////////////
## classes: outbreaker_moves list
## number of items: 8
## 
## // movement functions //
## End
```

o2mod.TransPhylo: TransPhylo module for outbreaker2

[implementation by Xavier Didelot and Finlay Campbell]

3 Custom likelihood

In order to calculate the likelihood under the TransPhylo model, we need to (i) extract the transmission tree from the outbreaker2 parameter, (ii) combine this transmission tree with the phylogenetic tree to form a colored tree, and (iii) calculate the likelihood of this colored tree. Step (i) is easy since transmission tree are encoded as lists in outbreaker2 in a very similar way to phylogenetic trees. For step (ii), we have to write the `combine` function. This is where the new stuff is especially interesting (this function is included in the `outbreaker2` package). The main difference with the `combine` function in `outbreaker` is that it only needs to call the appropriate function for each parameter. In this case, `combine(ttree, data$ptree)`. Step (iii) is also interesting because it is based on the `likTransPhylo` function. The only difference is that if the likelihood is negative infinity, then the likelihood is returned as `-Inf`.

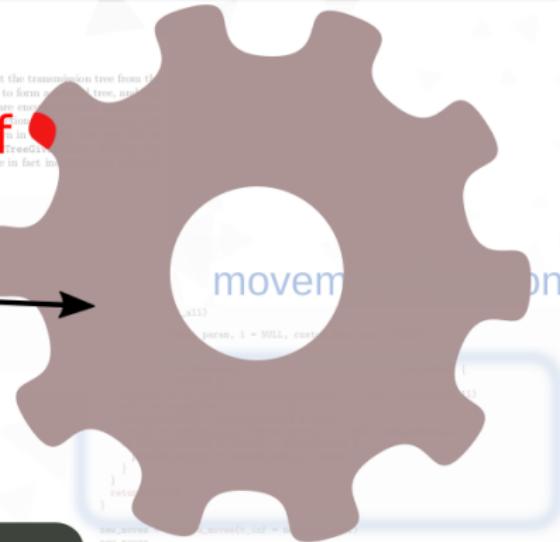
```
likTransPhylo <- function(data, param) {
  ttree <- list(ttree = chind(param$t.inf, data$dates, param$alpha),
               nan = data$ptree$nan)
  ttree$ttree[which(is.na(ttree$ttree[,3]))] <- 0
  txt <- capture.output(ptree <- combine(ttree, data$ptree))
  if (length(txt)==0) {
    prob <- 1
  } else {
    prob <- exp(-sum(log(ptree)))
    if (prob <= 0.25) {
      prob <- 0.25
    }
  }
  return(prob)
}
```

likelihood

Total: 25 lines of R
outbreaker2: 7,500 lines of R/C++
Code difference: 0.3%

New stuff

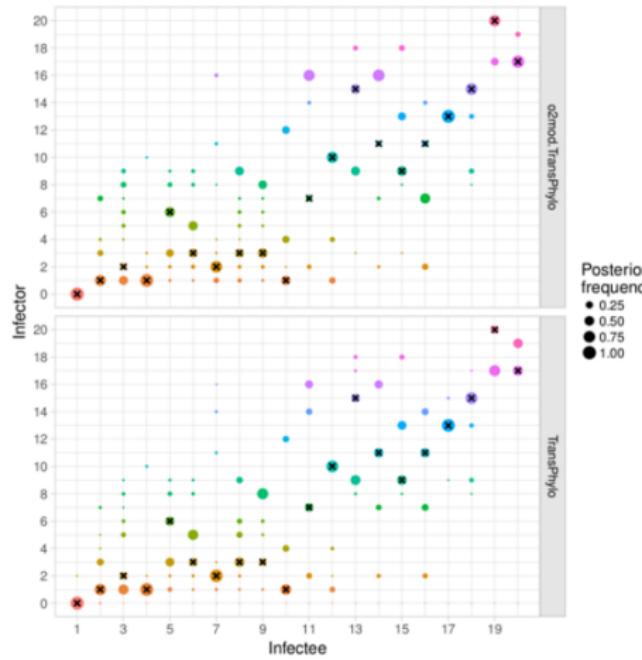
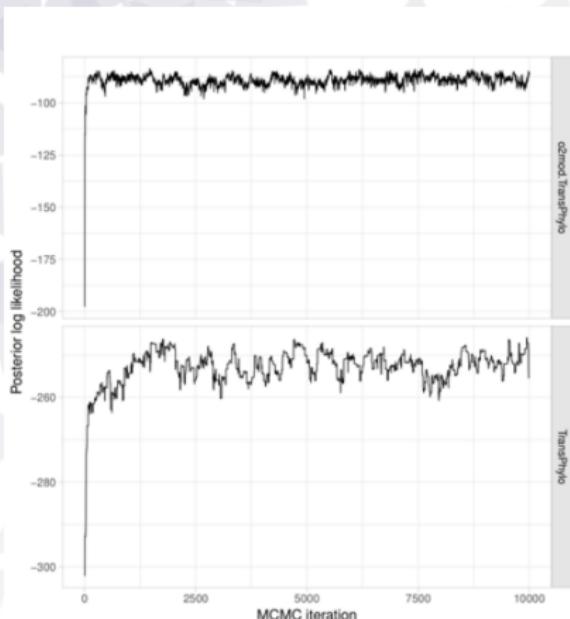
Old wheel



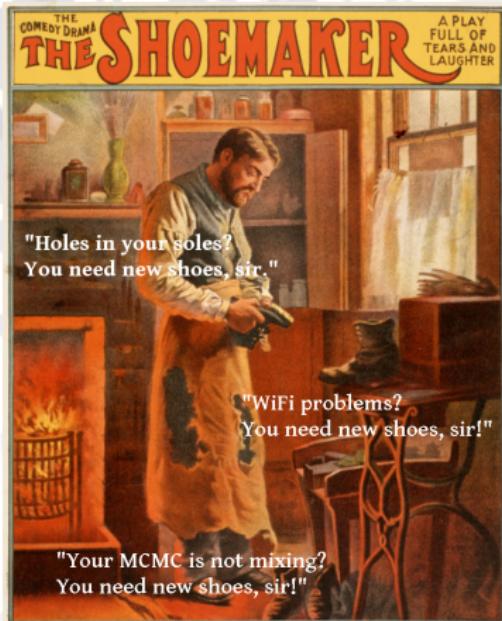
```
## movement functions //  
## class: outbreaker_movement list  
## number of items: 0  
##   
## movement functions //  
## len:
```

o2mod. TransPhylo: results

[material by Finlay Campbell]



Do we even need this to respond to outbreaks?



- in general, **not needed for forecasting or control**

Do we even need this to respond to outbreaks?



- in general, **not needed** for forecasting or control
- useful to detect **multiple introductions** or **superspreading**

Do we even need this to respond to outbreaks?



- in general, **not needed** for forecasting or control
- useful to detect **multiple introductions** or **superspreading**
- complement **contact tracing** data

Do we even need this to respond to outbreaks?

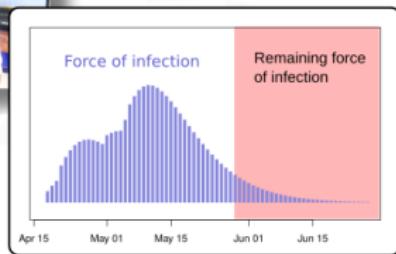


- in general, **not needed** for forecasting or control
- useful to detect **multiple introductions** or **superspreading**
- complement **contact tracing** data
- WGS are costly: **is it worth it?**

Ebola outbreak, Likati (DRC) 2017

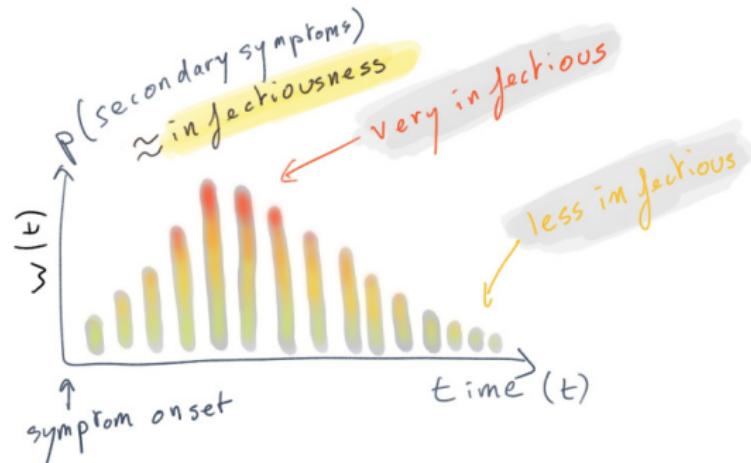


- EVD outbreak May 2017
- contact data visualisation tools used in contact tracing
- simple model informed response (scaling)
- end: 2nd July 2017



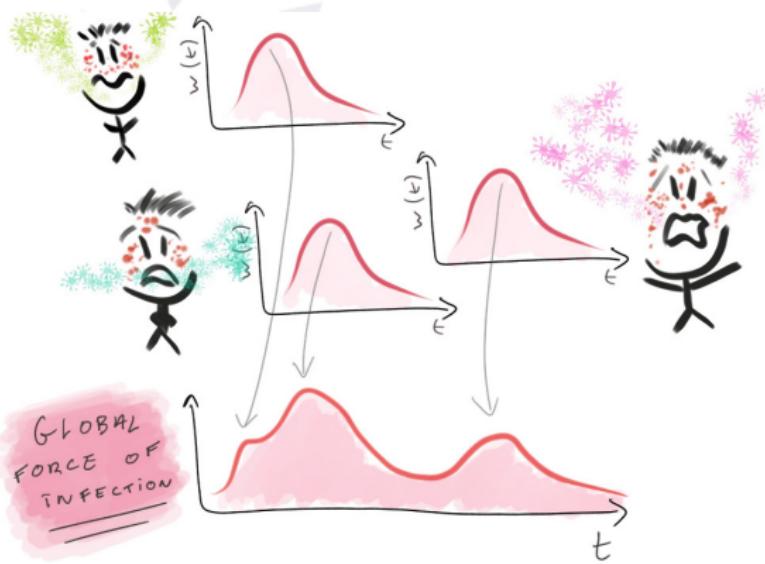
Individual infectiousness over time

Serial interval: delay between symptom onset in infector and infectees



Indicates when we expect new cases, if there are any.

A “simple” branching process model

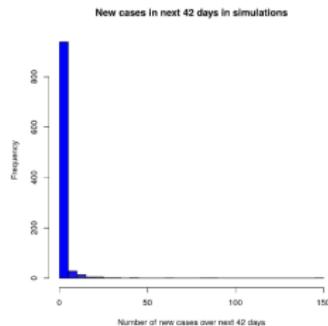
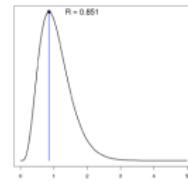
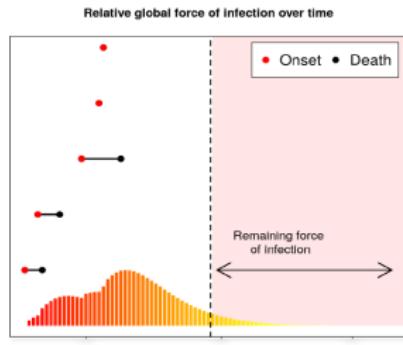


$$y_t \sim \mathcal{P}(\lambda_t) \quad ; \quad \lambda_t = R_0 \times \sum_i w(t - t_i)$$

y_t : incidence at time t ; $\mathcal{P}()$: Poisson distribution; λ_t : **global force of infection**; $w()$: serial interval distribution; t_i : date of symptom onset

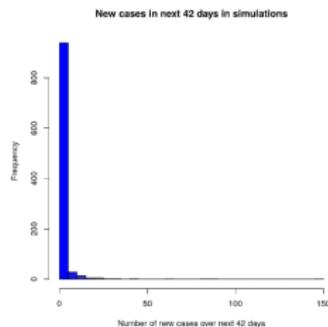
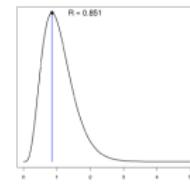
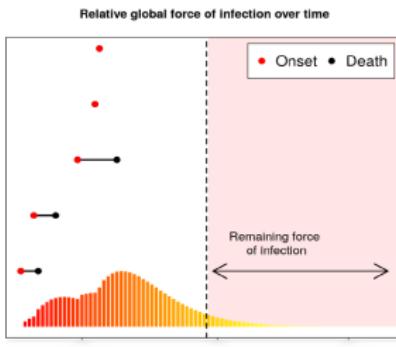
Scaling the response in real-time

Estimating remaining force of infection,
transmissibility (R), predicting new cases



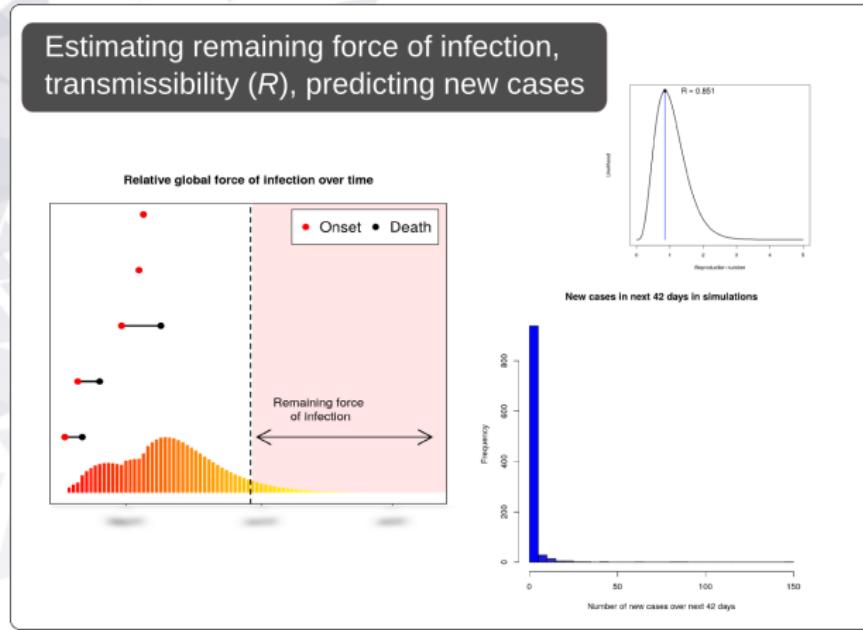
Scaling the response in real-time

Estimating remaining force of infection,
transmissibility (R), predicting new cases



Despite uncertainty in R_0 , new cases were unlikely.

Scaling the response in real-time



Despite uncertainty in R_0 , new cases were unlikely.

Discouraged scaling up in resource-limited context.

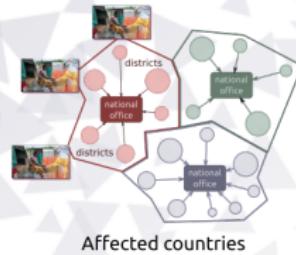
Challenges of a timely response



Challenges of a timely response

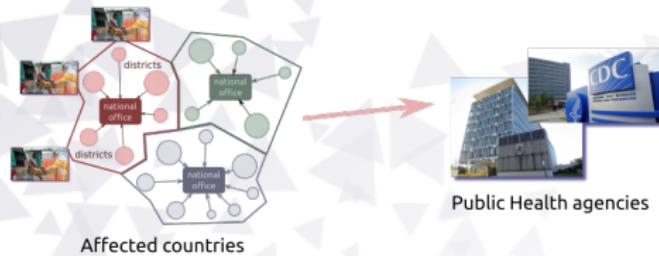


Challenges of a timely response

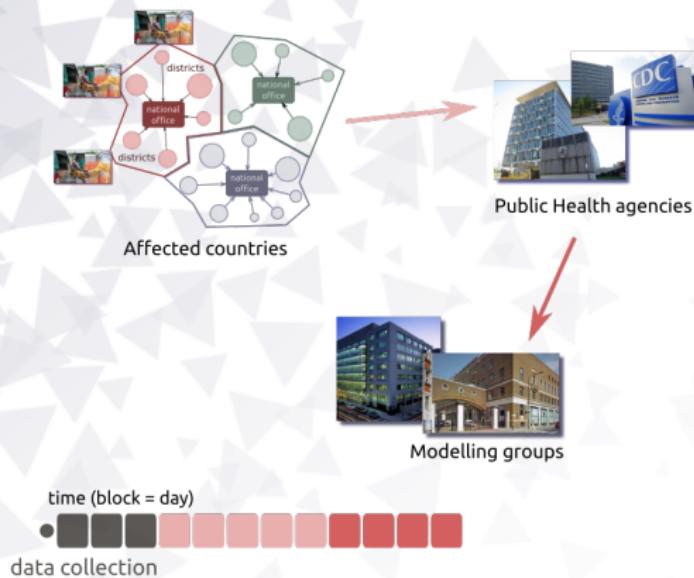


time (block = day)
• data collection

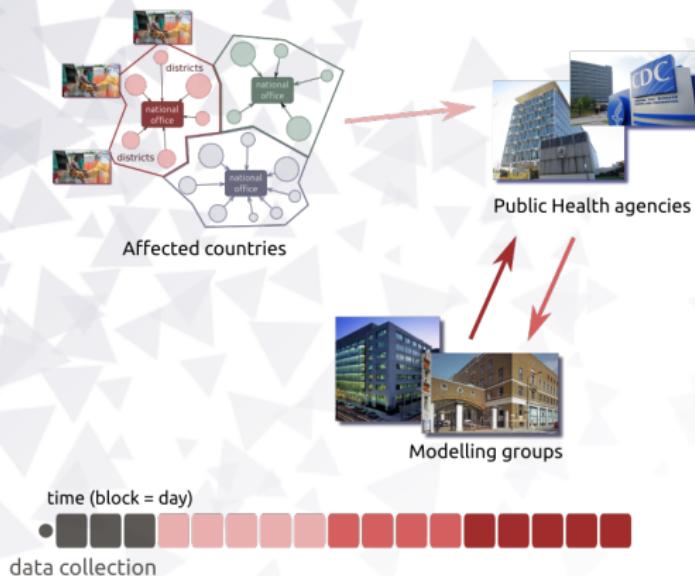
Challenges of a timely response



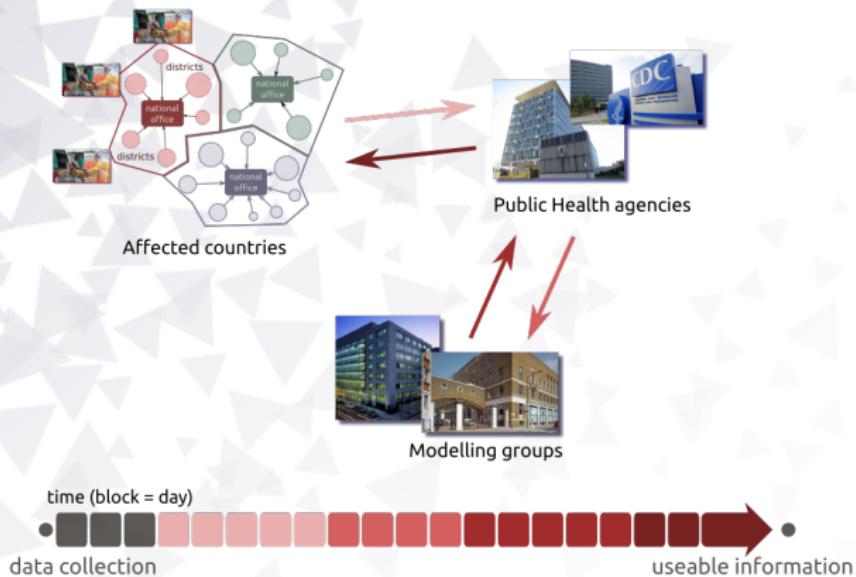
Challenges of a timely response



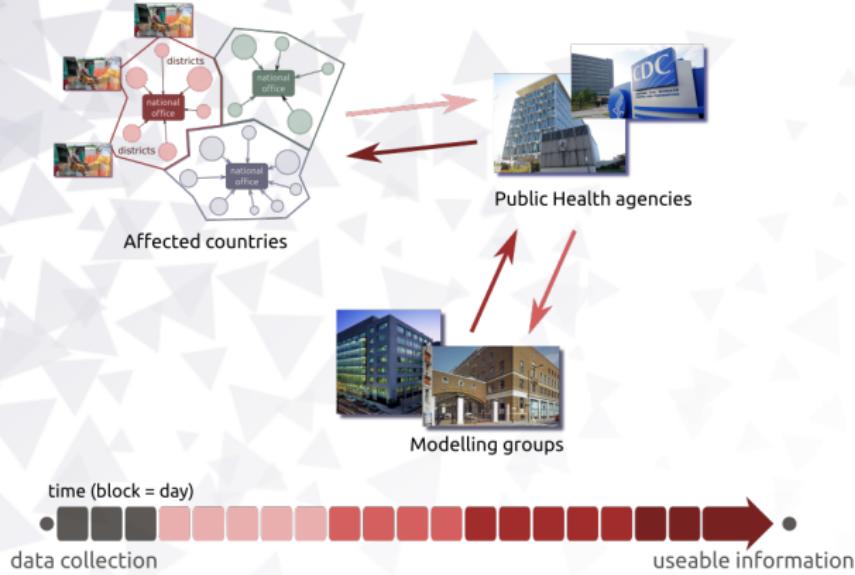
Challenges of a timely response



Challenges of a timely response

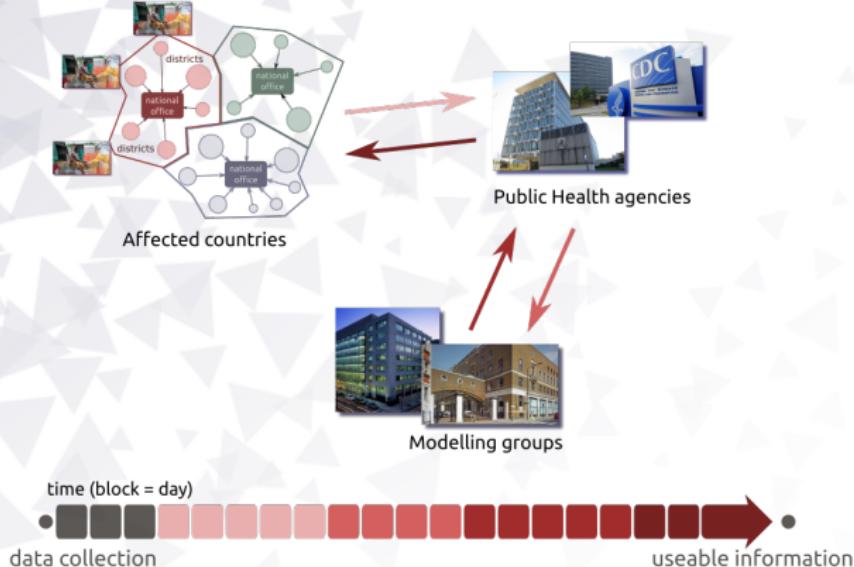


Challenges of a timely response



- efficient **tools** can shorten delays

Challenges of a timely response



- efficient **tools** can shorten delays
 - embed **methodologists** in outbreak response teams

Thanks to:

- **Eve Miguel**
- **Colleagues:** Finlay Campbell, Amrish Baidjoe, Anne Cori, James Hayward, Rich Fitzjohn, Neil Ferguson
- **Groups:** WHO Ebola Response Team, Ebola Likati team, RECON members, GOARN
- **funding:** HPRU-NIHR, MRC

RECON

www.repidemicsconsortium.org

RECON Learn

A free and open collection of training resources
for epidemics analysis and outbreak response.

reconlearn.netlify.com/

Upcoming event!

<https://recon-gathering-march2018.netlify.com/>