

RECON

A graph-based evidence synthesis approach for outbreak detection

Thibaut Jombart

20th September 2017

Imperial College London
MRC Centre for Outbreak Analysis and Modelling

Topics of the day

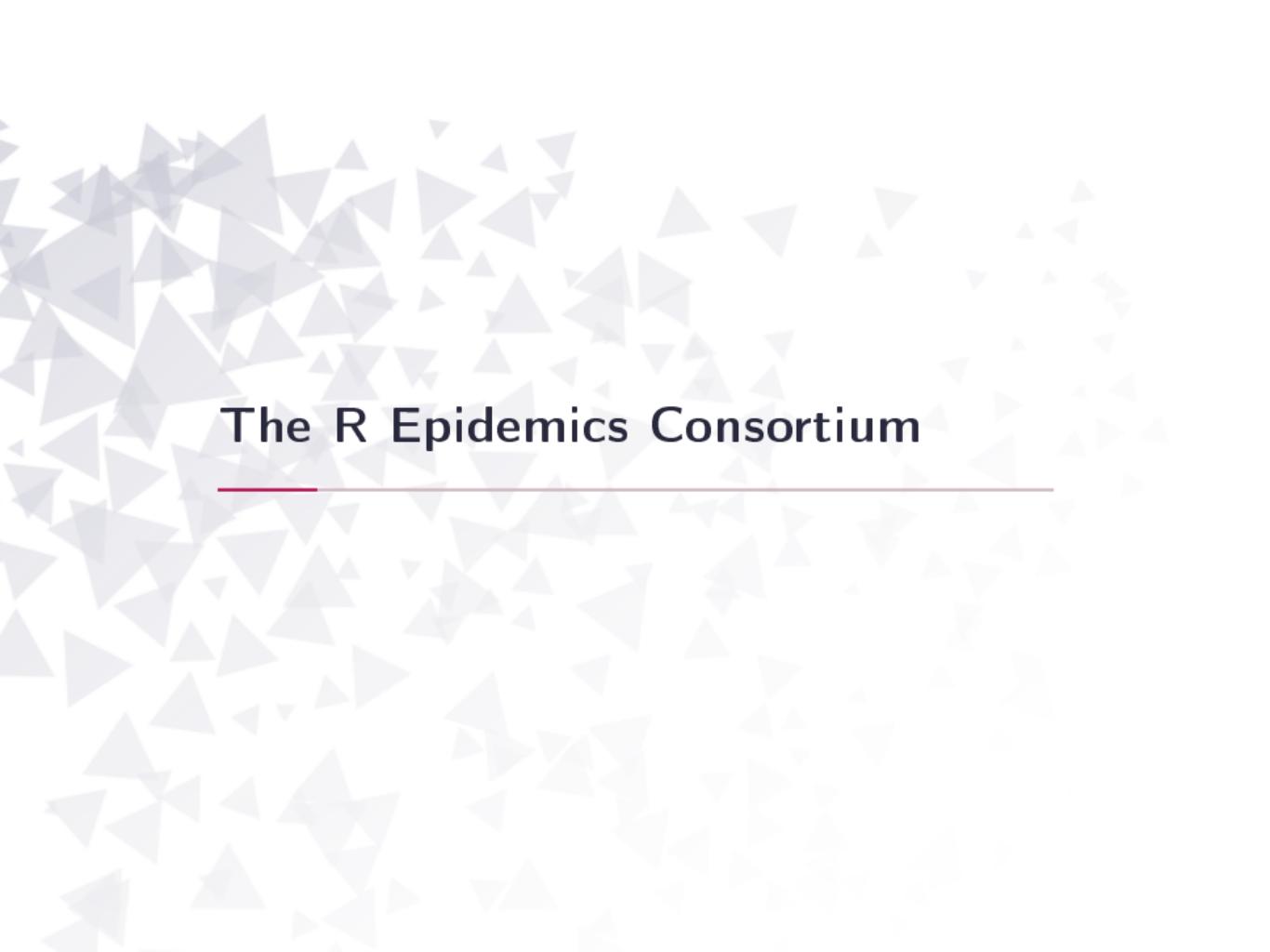


Topics of the day



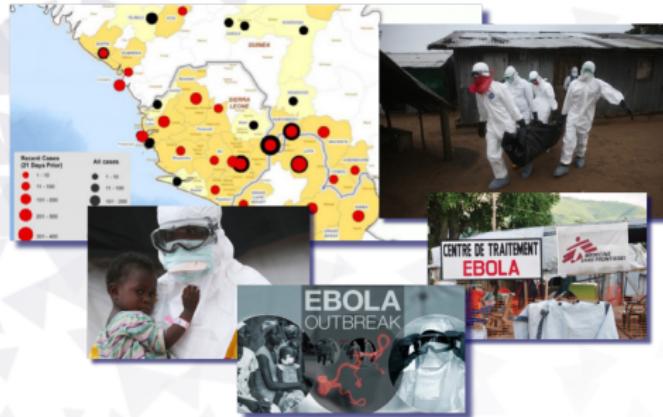
Topics of the day





The R Epidemics Consortium

Lessons learnt from the Ebola response



Lessons learnt from the Ebola response



Lessons learnt from the Ebola response

WHO Ebola response team

Help improving situation awareness

Imperial College Ebola team

Timeline:

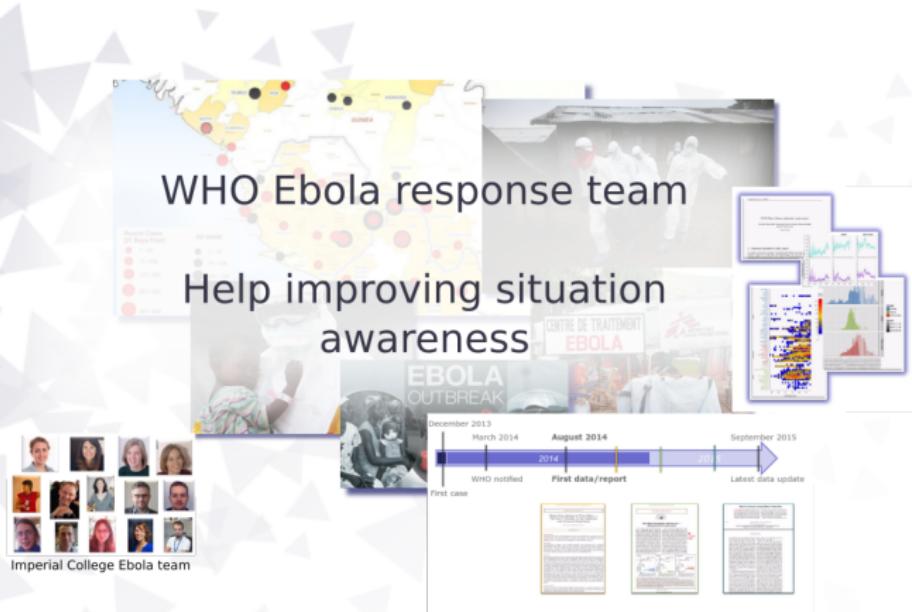
- December 2013: First case
- March 2014: WHO notified
- August 2014: First data/report
- September 2015: Latest data update

Centre de Traitement EBOLA

EBOLA OUTBREAK

WHO Ebola response team dashboard

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

functional
incubation
userfriendly secure dictionary
systems testing automated continuous
collection series repository
rpp efficiency number fast
secured bias outbreaks
parsing code integration
reporting gui
unit data delay
epidemiology security peak
situation anonymised
opensource contact
epiinfo clean time
compiled outbreaker interface tree
symptoms lineelist follow
tracing shiny
automation cdc
epicontacts edc
ggplot clusters rates
dashboard reliable
parallel contacttracing
parameters epidemics
genomics distribution
incidence bayesian site report
cleaning estimation
security transmission
peak curation model
contacts package
report reproduction
functional incubation
userfriendly secure dictionary
systems testing automated continuous
collection series repository
rpp efficiency number fast
secured bias outbreaks
parsing code integration
reporting gui
unit data delay
epidemiology security peak
situation anonymised
opensource contact
epiinfo clean time
compiled outbreaker interface tree
symptoms lineelist follow
tracing shiny
automation cdc
epicontacts edc
ggplot clusters rates
dashboard reliable
parallel contacttracing
parameters epidemics
genomics distribution
incidence bayesian site report
cleaning estimation
security transmission
peak curation model
contacts package
report reproduction

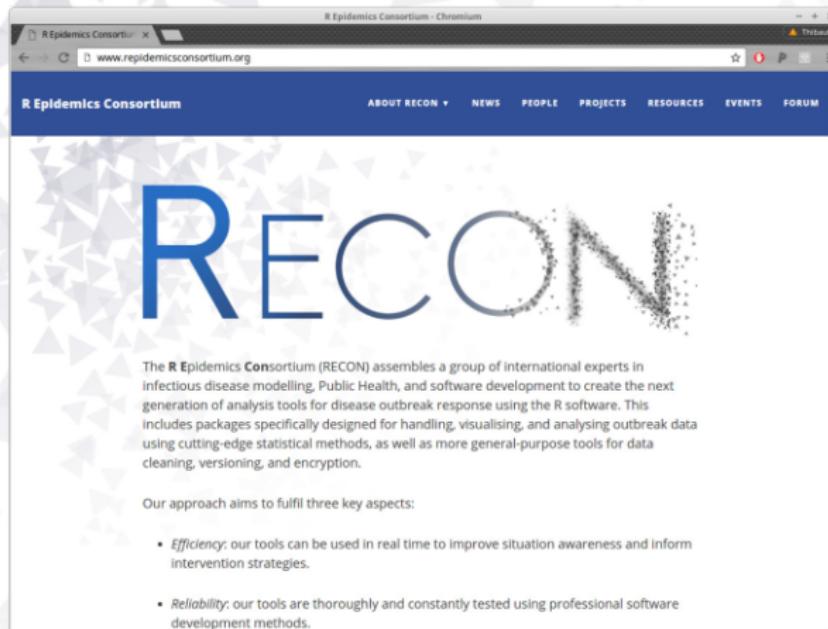
RECON
[The R Epidemics Consortium]

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" website. The title bar says "R Epidemics Consortium - Chromium". The address bar shows "www.repidemicsconsortium.org". The main content area features a large blue "RECON" logo where the letters are composed of small dots. Below the logo is a paragraph of text about the consortium's purpose. At the bottom, there is a section titled "Our approach aims to fulfil three key aspects:" followed by a bulleted list.

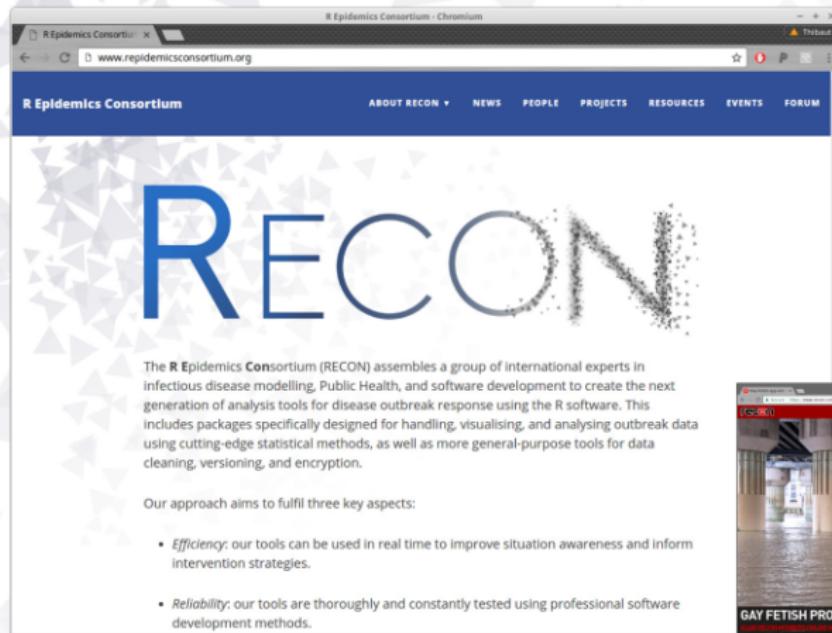
The R Epidemics Consortium (RECON) assembles a group of international experts in infectious disease modelling, Public Health, and software development to create the next generation of analysis tools for disease outbreak response using the R software. This includes packages specifically designed for handling, 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:

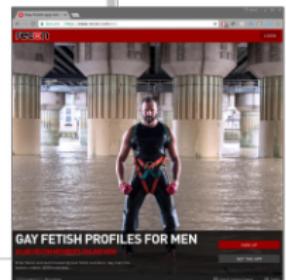
- *Efficiency*: our tools can be used in real time to improve situation awareness and inform intervention strategies.
- *Reliability*: our tools are thoroughly and constantly tested using professional software development methods.

RECON: the R Epidemics Consortium

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

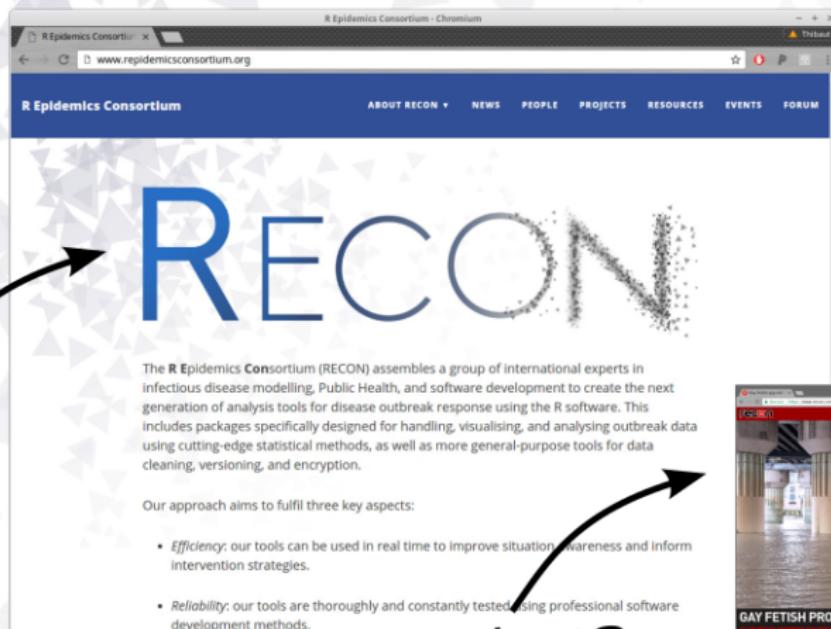


The screenshot shows the homepage of the RECON website. The header features the text "R Epidemics Consortium" and "RECON". Below the header is a large graphic where the letters of "RECON" are composed of small dots or particles. A descriptive paragraph explains the consortium's mission: "The R Epidemics 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." Below this text, a section titled "Our approach aims to fulfil three key aspects:" lists three bullet points: "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.", and "Scalability: our tools are designed to handle large amounts of data and can be scaled up as needed."



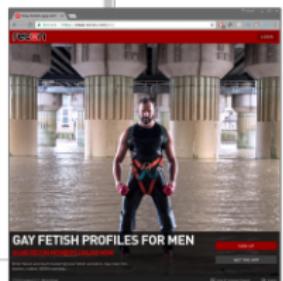
RECON: the R Epidemics Consortium

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



The screenshot shows the homepage of the R Epidemics Consortium (RECON) website. The title 'RECON' is prominently displayed in large blue letters, with the 'O' composed of small dots. Below the title is a paragraph describing the consortium's mission: "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." At the bottom of the page, there is a section titled "Our approach aims to fulfil three key aspects:" with two bullet points: "Efficiency: our tools can be used in real time to improve situation awareness and inform intervention strategies." and "Reliability: our tools are thoroughly and constantly tested using professional software development methods."

Not us



RECON

www.repidemicsconsortium.org

- started 6th September 2016
- ~70 members
- 17 countries, > 40 institutions
- ~ 3 packages released, 20 under development
- public forum, blog, online resources



Statistical software development

- **efficiency**: useful for improving situation awareness in real time
- **reliability**: outputs can be trusted
- **accessibility**: widely available, easy learning curve



[nicer
'strong and
stable']

Statistical software development

- **efficiency**: useful for improving situation awareness in real time
- **reliability**: outputs can be trusted
- **accessibility**: widely available, easy learning curve



Statistical software development

- **efficiency**: useful for improving situation awareness in real time
- **reliability**: outputs can be trusted
- **accessibility**: widely available, easy learning curve

Translation

- **disseminating knowledge**: free online training material, involvement with FETPs, workshops
- **outbreak response**: deployment to the field
- **RECON deployer**: portable data analysis environment

RECON: projects

The screenshot shows the 'R Epidemics Consortium' website at www.repidemicsconsortium.org/projects/. The page displays 14 projects, each represented by a circular icon containing gears and a brief description:

- epiflows**: Visualisation and analysis of passenger flows.
- epimaps**: Helpers and wrappers for mapping diseases.
- epimatch**: Finding matching patient records across tabular data sets.
- epitrix**: Small utility functions for epidemiology.
- gisfirstaid**: Tutorials and code gists for mapping infectious diseases.
- incidence.ul**: Graphical user interface for incidence.
- nomad**: Pack up R to take away.
- outbreaker2**: Inferring transmission chains by integrating epidemiological and genetic data.
- projections**: Projections of future incidence.
- recon.ul**: Template shiny GUI for RECON packages.
- recontools**: Tools to develop RECON packages.
- vimes**: Visualisation and Monitoring of Epidemics, including some outbreak detection algorithms.

A red circle highlights the **vimes** project icon.



An integrative approach for outbreak detection



Aims: develop a new method which..

- **detects outbreaks** i.e. groups of related cases (on the same transmission chain)



¹ well, really, I made that up because I was reading 'Snuff' at the time; at least this one is not a dodgy website (yet); incidentally, Terry Pratchett was a huge fan of using long footnotes, which were often quite entertaining to read; note that it does not apply here: if you are still reading this, you probably missed what I just said



Aims: develop a new method which..

- **detects outbreaks** i.e. groups of related cases (on the same transmission chain)
- **integrates different data:** temporal, spatial, genetic, etc.

¹ well, really, I made that up because I was reading 'Snuff' at the time; at least this one is not a dodgy website (yet); incidentally, Terry Pratchett was a huge fan of using long footnotes, which were often quite entertaining to read; note that it does not apply here: if you are still reading this, you probably missed what I just said

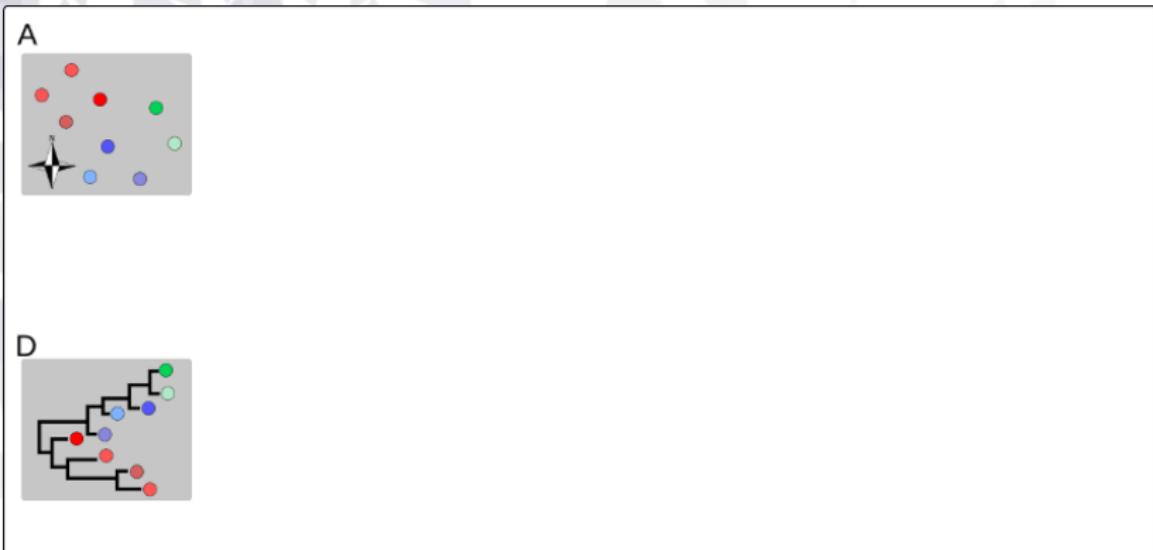


Aims: develop a new method which..

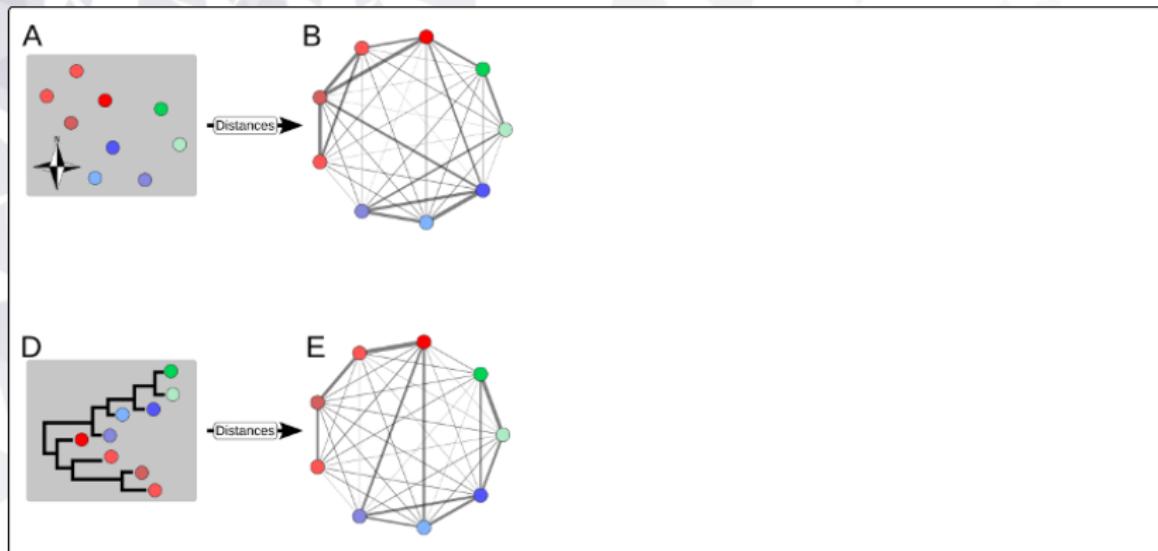
- **detects outbreaks** i.e. groups of related cases (on the same transmission chain)
- **integrates different data:** temporal, spatial, genetic, etc.
- **works fast, scales well:** so that it can be used for real-time outbreak detection

¹ well, really, I made that up because I was reading 'Snuff' at the time; at least this one is not a dodgy website (yet); incidentally, Terry Pratchett was a huge fan of using long footnotes, which were often quite entertaining to read; note that it does not apply here: if you are still reading this, you probably missed what I just said

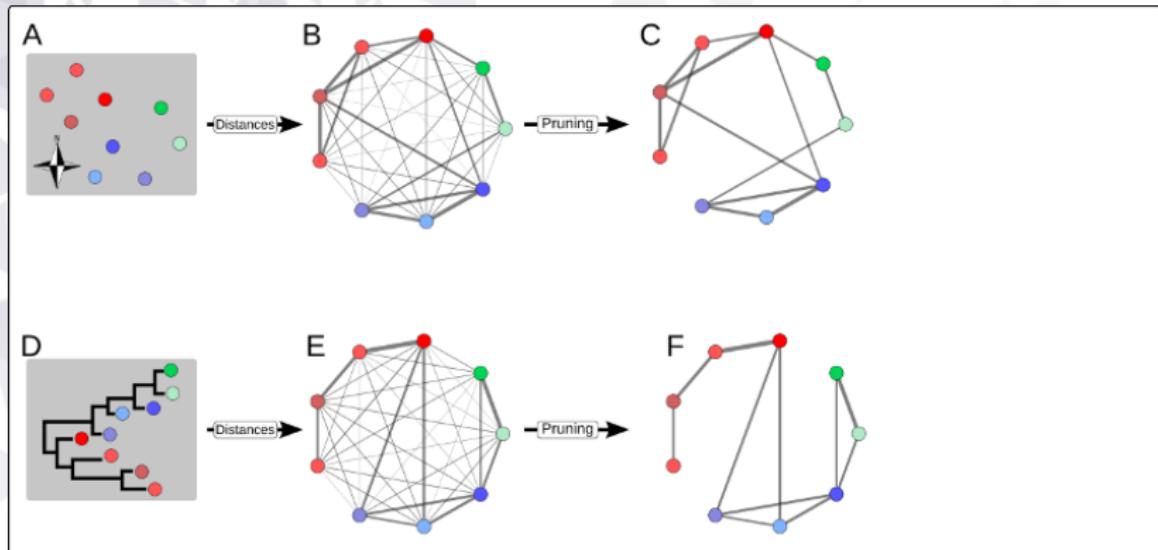
A graph-based evidence synthesis approach



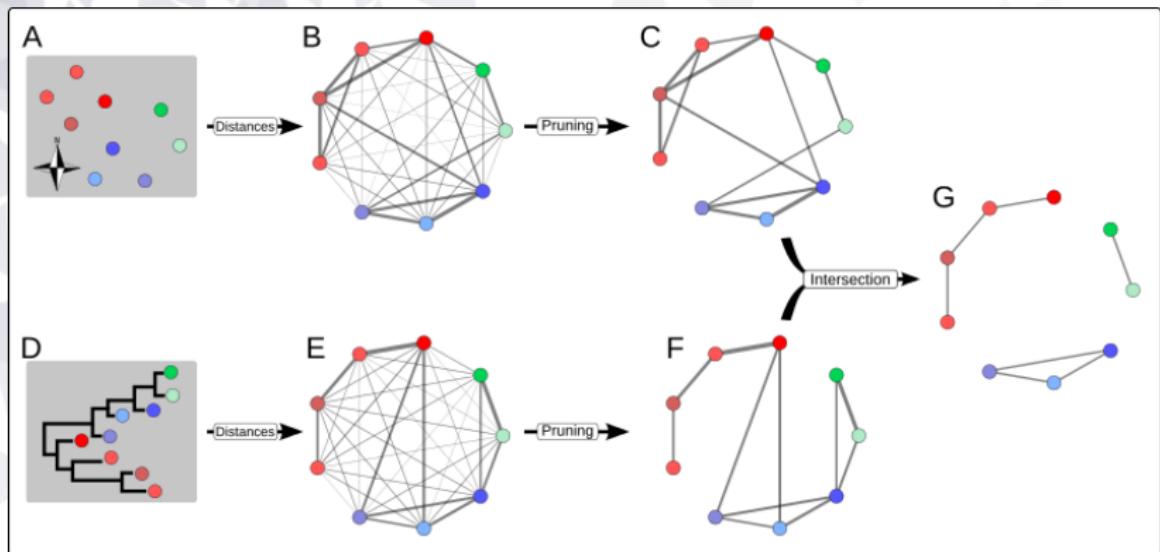
A graph-based evidence synthesis approach



A graph-based evidence synthesis approach

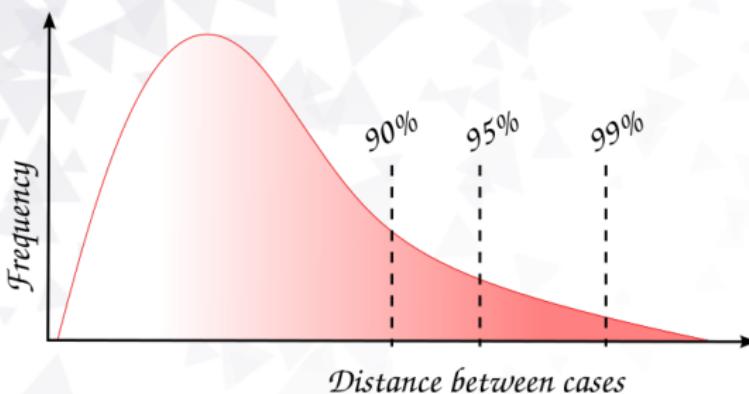


A graph-based evidence synthesis approach

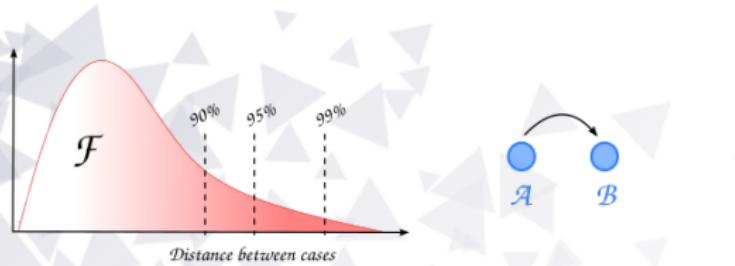


Pruning graphs: where to cut?

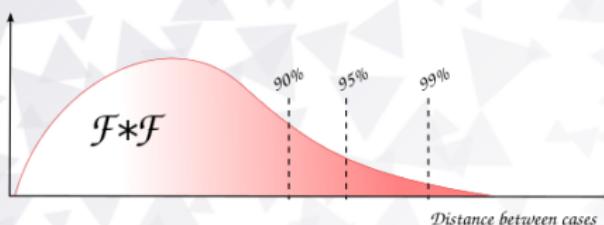
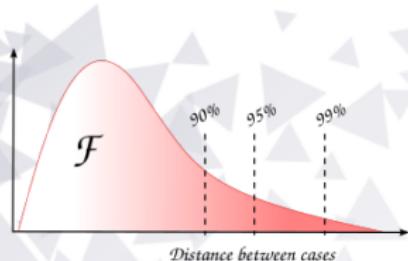
Assuming a known expected distribution between pairs of cases (e.g. serial interval, spatial kernel, molecular clock), different quantiles can be used:



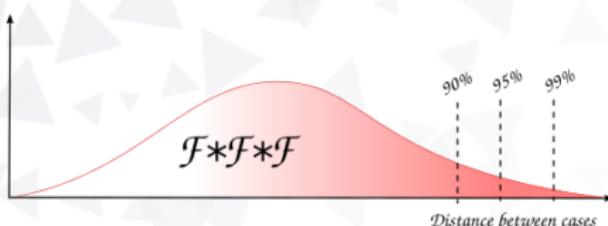
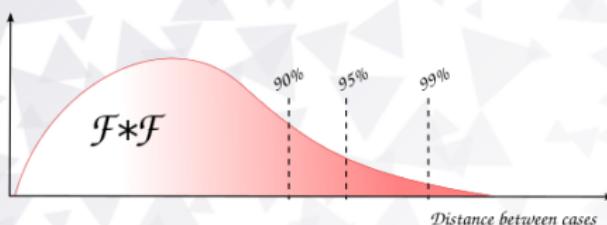
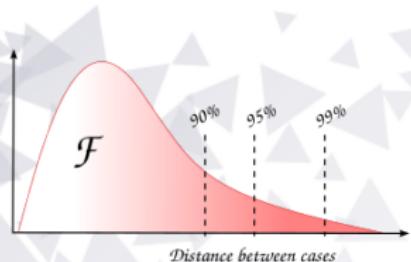
Pruning graphs: where to cut?



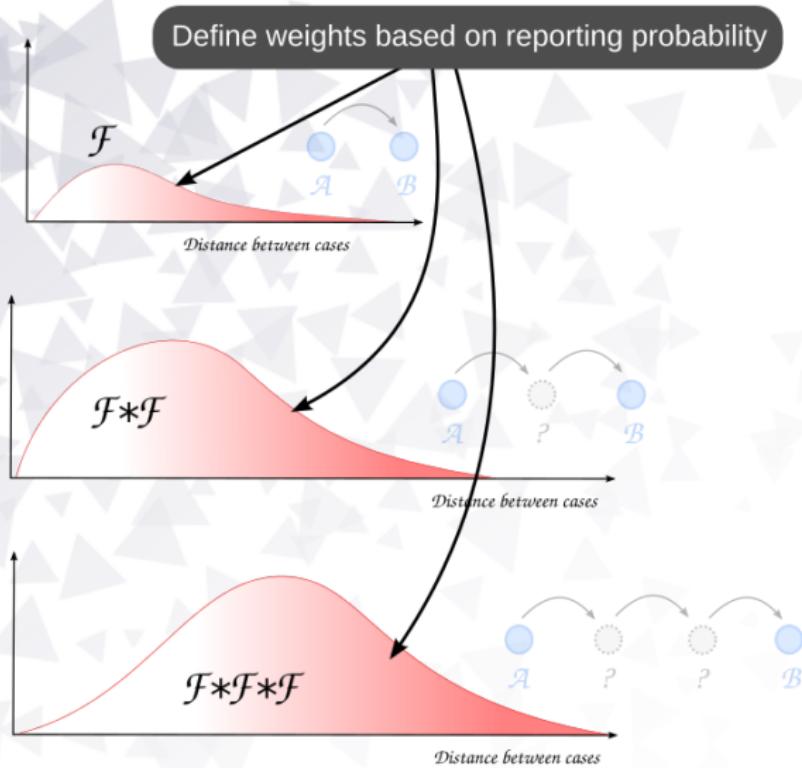
Pruning graphs: where to cut?



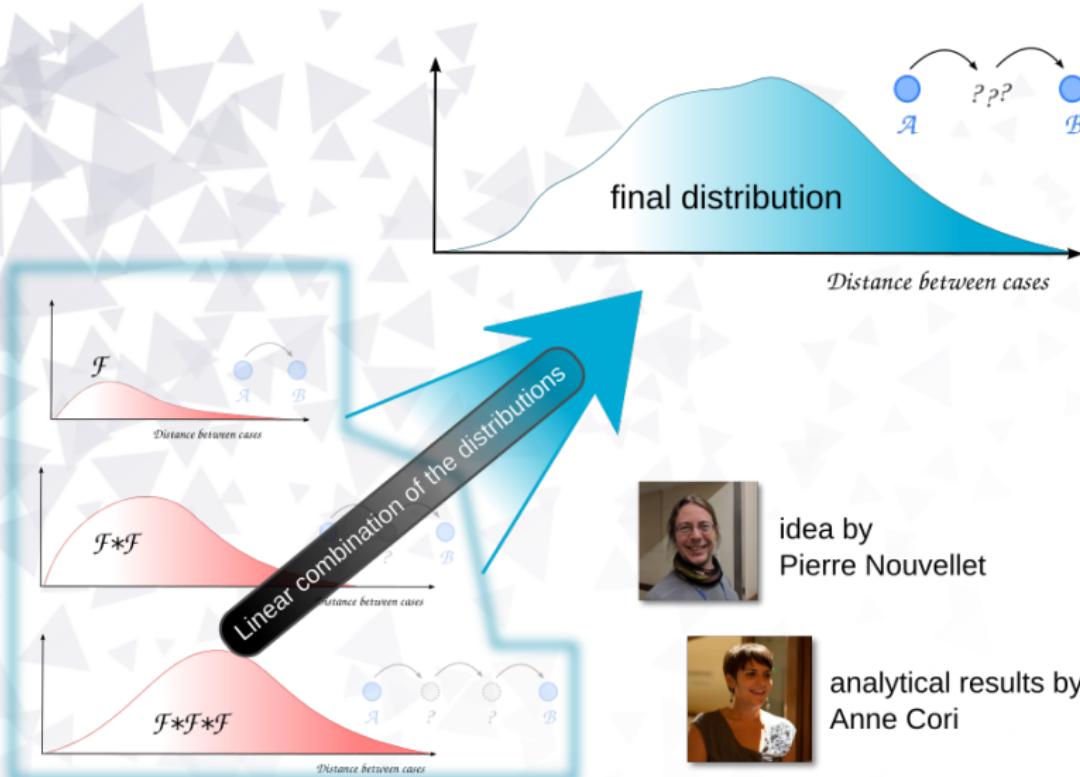
Pruning graphs: where to cut?



Pruning graphs: where to cut?



Pruning graphs: where to cut?



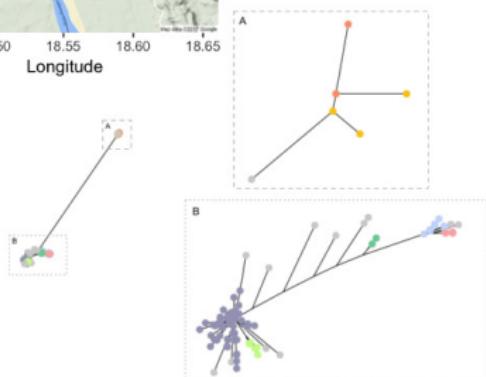
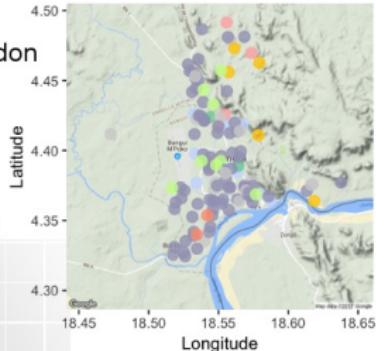
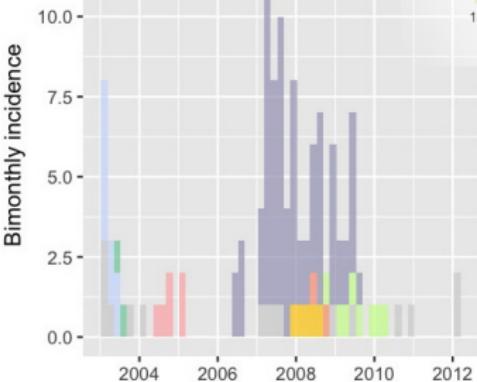
Application: dog rabies epidemics, Central African Republic



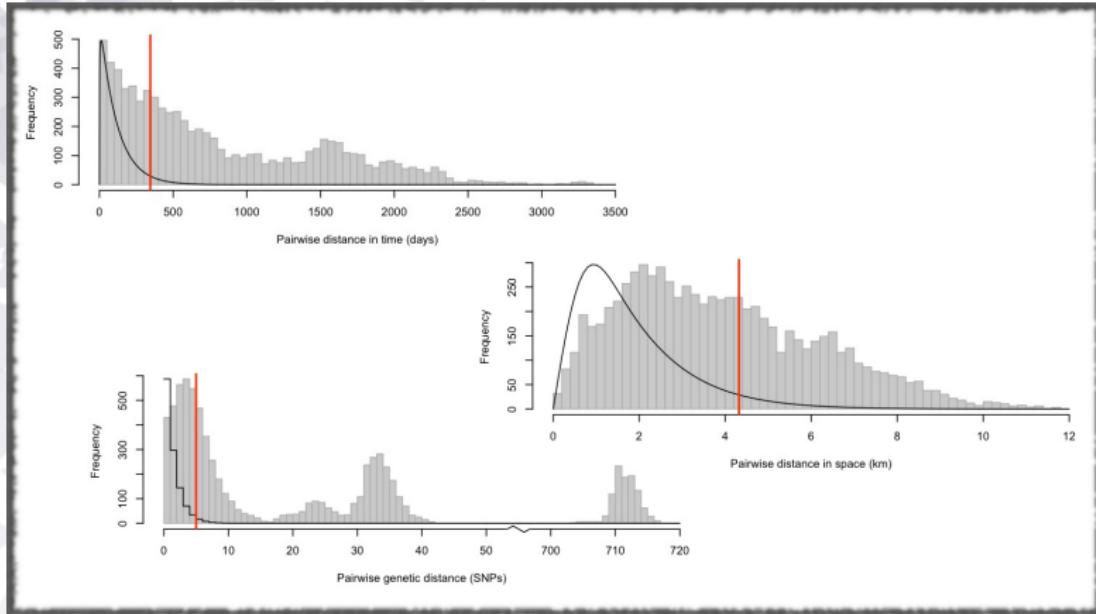
Anne Cori
Imperial College London



Pierre Nouvellet
University of Sussex

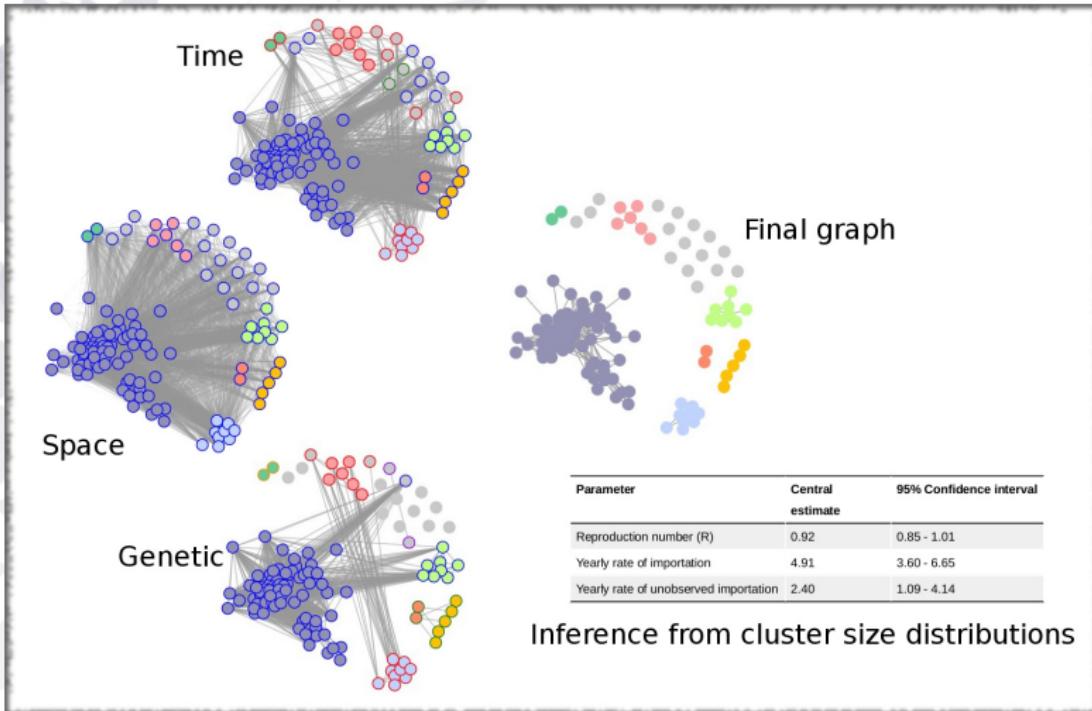


Distributions of distances between cases



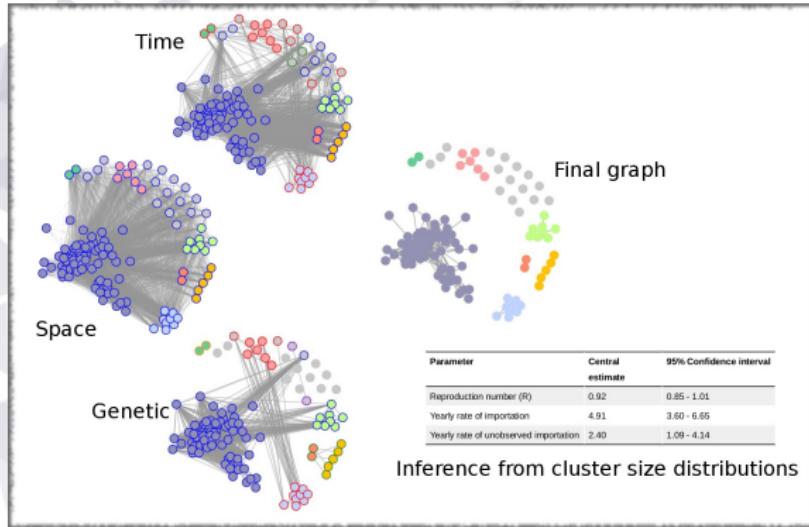
[material by Anne Cori]

Results



[material by Anne Cori and Pierre Nouvellet]

Results



- one large outbreak, low R_0 , $\sim 5 - 10$ introductions / year
- same results as more complex approaches (BEAST, epi model + particle filtering)
- much faster (<1s vs 1 week)

Perspectives

- **flexible** approach for detecting outbreaks using different data types

Perspectives

- **flexible** approach for detecting outbreaks using different data types
- **threshold**: unsatisfying, but sensitivity study easy

Perspectives

- **flexible** approach for detecting outbreaks using different data types
- **threshold**: unsatisfying, but sensitivity study easy
- **fast and scalable**: possible integration in routine surveillance

Perspectives

- **flexible** approach for detecting outbreaks using different data types
- **threshold**: unsatisfying, but sensitivity study easy
- **fast and scalable**: possible integration in routine surveillance
- can serve as **basis to other methods** for integrating different data sources

Thanks

- Conference organisers
- Colleagues: **Anne Cori, Pierre Nouvellet**, Tini Garske, Hervé Bourhy, Emmanuel Nakouné
- Groups: WHO Ebola Response Team, Hackout 1/2/3, RECON members, GOARN
- funding: HPRU-NIHR, MRC

RECON

www.repidemicsconsortium.org

vimes

[www.repidemicsconsortium.org/
vimes](http://www.repidemicsconsortium.org/vimes)

Questions?