Case and population geo-location data for infectious disease outbreak investigation and surveillance

Emmanuel Robesyn, European Centre for Disease Prevention and Control

Access for public bodies to privately-held data. European Commission Workshop, Brussels, 26 June 2017

European Centre for Disease Prevention and Control

ECDC : An EU agency dedicated to the prevention and control of infectious diseases

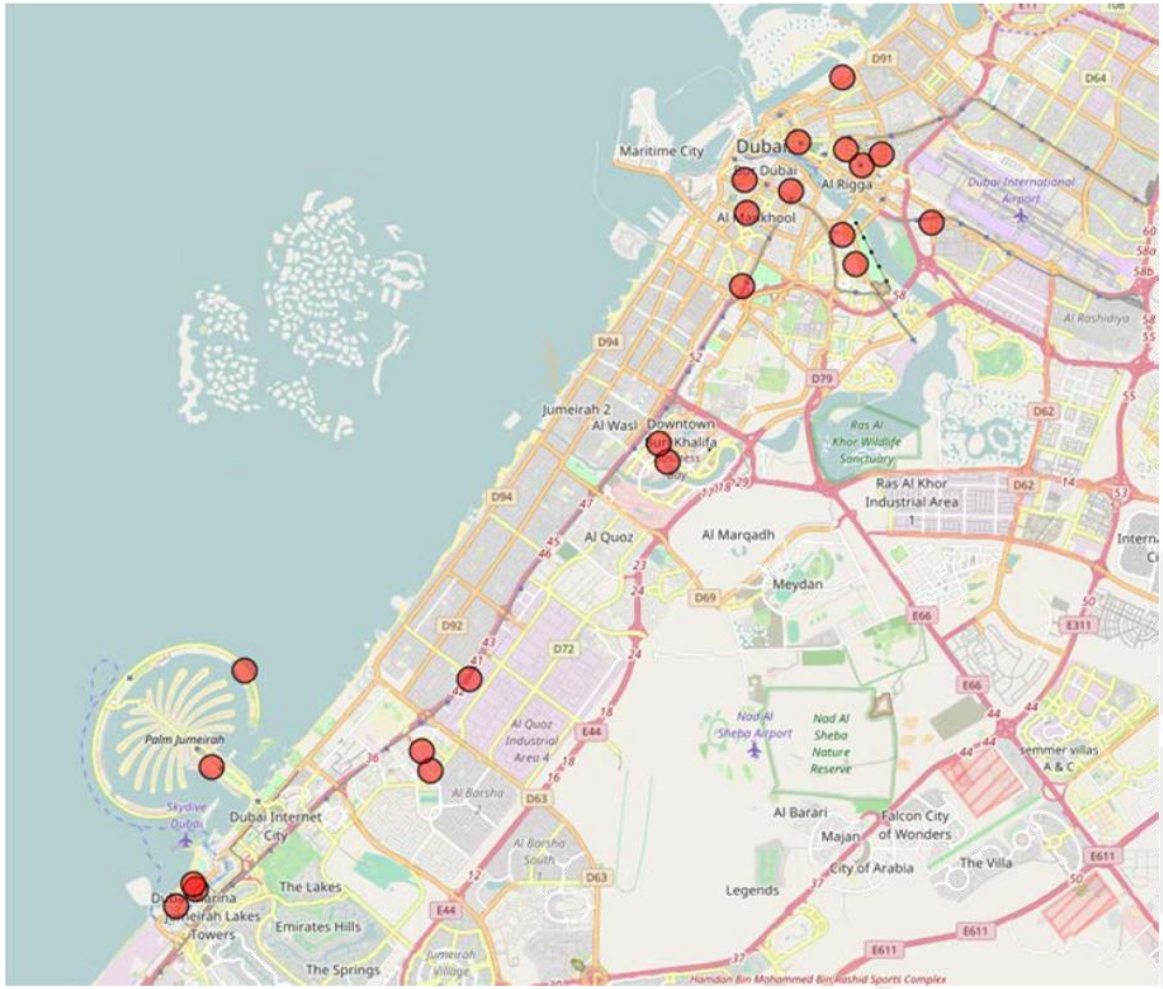
Identify, assess and communicate current and emerging health threats to human health from communicable diseases.

— ECDC Founding Regulation (851/2004), Article 3

- Case reports (aggregated or not)
- Outbreak event reports
- Population data
- Other (e.g. flights, vectors)

Legionnaires' disease, a motivating example for use of location data

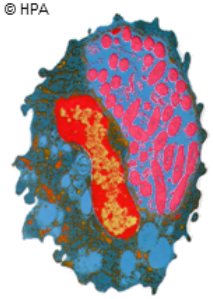
Figure 3. Accommodation sites in Dubai (N=22) where TALD cases* stayed, as of 21 December 2016



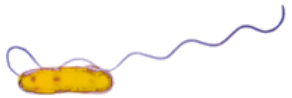
Welcome to the Legionnaires' disease outbreak investigation toolbox

Questionnaire:

Where have you been in 14 days before falling ill ?



Stain of Legionella pneumophila within a human lung macrophage



Individual Legionella pneumophila bacterium

REVIEW ARTICLES

The application of geographic information systems and spatial data during Legionnaires' disease outbreak responses

M Bull (matthew.bull@hpa.org.uk)[1], I M Hall[1], S Leach[1], E Robesyn[2]

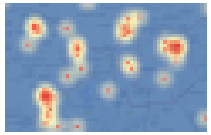
1. Microbial Risk Assessment, Emergency Response Department, Health Protection Agency, Porton Down, United Kingdom
2. Surveillance and Response Support Unit, European Centre for Disease Prevention and Control, Stockholm, Sweden

Case density map

- **Hotspot of case density suggest proximity of source**
- **Visualisation without disclosing case location data**

2. Select model

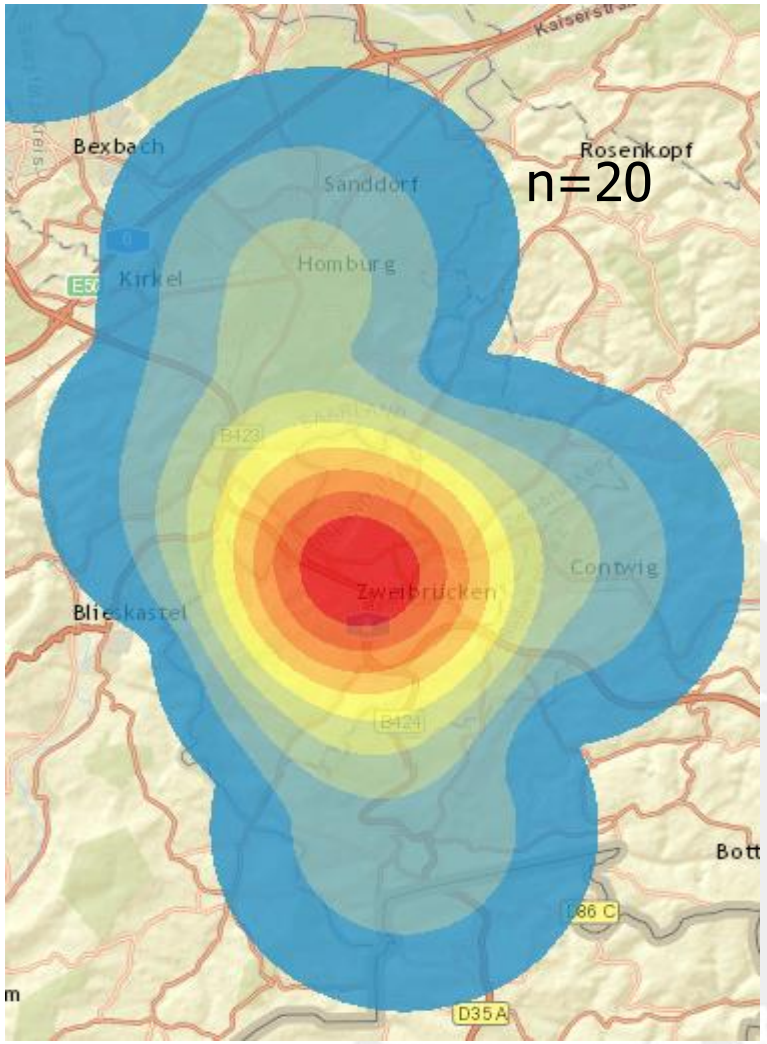
- Model 1: Case density ?
- Model 2: Disease risk ?
- Model 3: Buffer density ?
- Model 4: Buffer risk ?

Sample model result: 

1. Add data

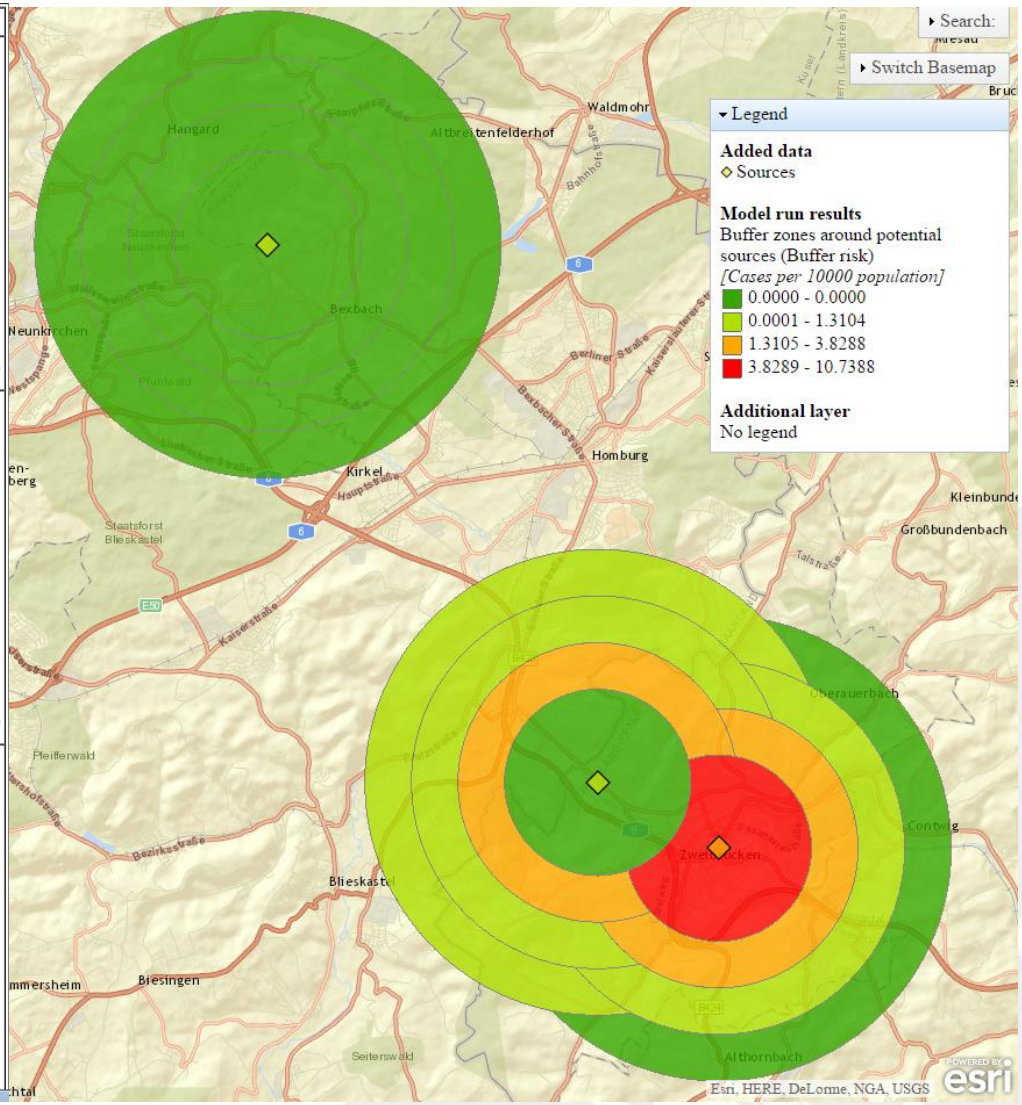
Cases: [0027] Import CSV Export CSV ?

Sources: [0009] Import CSV Export CSV ?



Case risk in rings around potential sources

ID	Name	Buffer, m	km²	Total pop.	Cases	Cases per 10000	Graph
2		4000	29.58	21628	0	0	
2		6000	36.99	24118	0	0	
2		8000	51.79	32834	0	0	
2		10000	66.58	46509	0	0	
3		4000	29.44	9312	10	10.739	
3		6000	36.82	13059	5	3.829	
3		8000	51.54	11338	1	0.882	
3		10000	66.27	12217	0	0	
4		4000	29.45	0	5	0	
4		6000	36.83	24025	7	2.914	
4		8000	51.57	22894	3	1.31	
4		10000	66.31	14102	1	0.709	



Use of several locations per person

- Assign weight to each location, proportional to time spent

  Example:
  - Case x, *residence location*, weight 0.5
  - Case x, *work place location*, weight 0.3
  - Case x, *hobby place location*, weight 0.2

Advantage of mobile phone location data

- More precise location, not limited to residence: usefulness dependent on tower coverage and spatial scale outbreak
- No recall bias
- Easier to collect
- Trajectory can still be verified and further detailed by case
- Includes time feature: can be related to most likely time of infection of each case

=> **mobile call data records for source detection**

- When multiple locations/movement data, risk calculation based on resident population not valid

=> **mobile call data records to estimate risk denominator**

Extending the use for other diseases

- Individual case location: source identification in outbreaks (environmentally linked infections)
  - Airborne from contaminated water (LD)
  - From infected animals (zoonosis e.g. Q fever, MERS, avian influenza)
  - Vector borne diseases (e.g. dengue, West-Nile, chikungunya, yellow fever)
- Population density (anonymous): population at risk
  - Outbreaks: attack rate, risk assessment, preventive action, transmission chain and characteristics
  - Surveillance: denominator real time (vs. static census data of resident population), more precise (e.g. commuting population, weekend vs. weekday, high vs. low season),
  - Mass gathering events (sports, religious, cultural, …)
  - Chemical/nuclear incident
- Population movement: migration (e.g. cholera Haiti*, ebola W-Africa**)

* Using Mobile Phone Data to Predict the Spatial Spread of Cholera, Bengtsson et al.

** Commentary: containing the ebola outbreak - the potential and challenge of mobile network data, Wesolowski et al.

Challenges

- Confidentiality – need for informed consent, data protection agreement
  - population (person identifiable data)
  - case (health data)
- Data storage and retrieval
  - Individual versus population density
  - Retrospective (how long back in time) versus prospective during outbreak (tracking to be turned on, on request of public health authorities)
  - Different mobile phone companies in outbreak area
- Accessibility: constraints for outbreak investigation purpose
  - timeliness (days - weeks)
  - established way to start tracking (if no default storage)
- Precision: constraints for outbreak investigation purpose
  - Legionnaires' disease outbreak - range 1 - 15 km
  - Vector borne or zoonosis - depending on range of vector, epizooty, …

=> NEED FOR FRAMEWORK TO FACILITATE USE

Thank you