NEOH Workshop on Evaluation of Data & Information Sharing in One Health Initiatives
Copenhagen, 20th & 21st April 2016

Prepare, Predict, Prevent: Creating Objectivity in Infectious Disease Risk Assessment using Big Data Approaches

K. Marie McIntyre

Institute of Infection & Global Health
UNIVERSITY OF LIVERPOOL

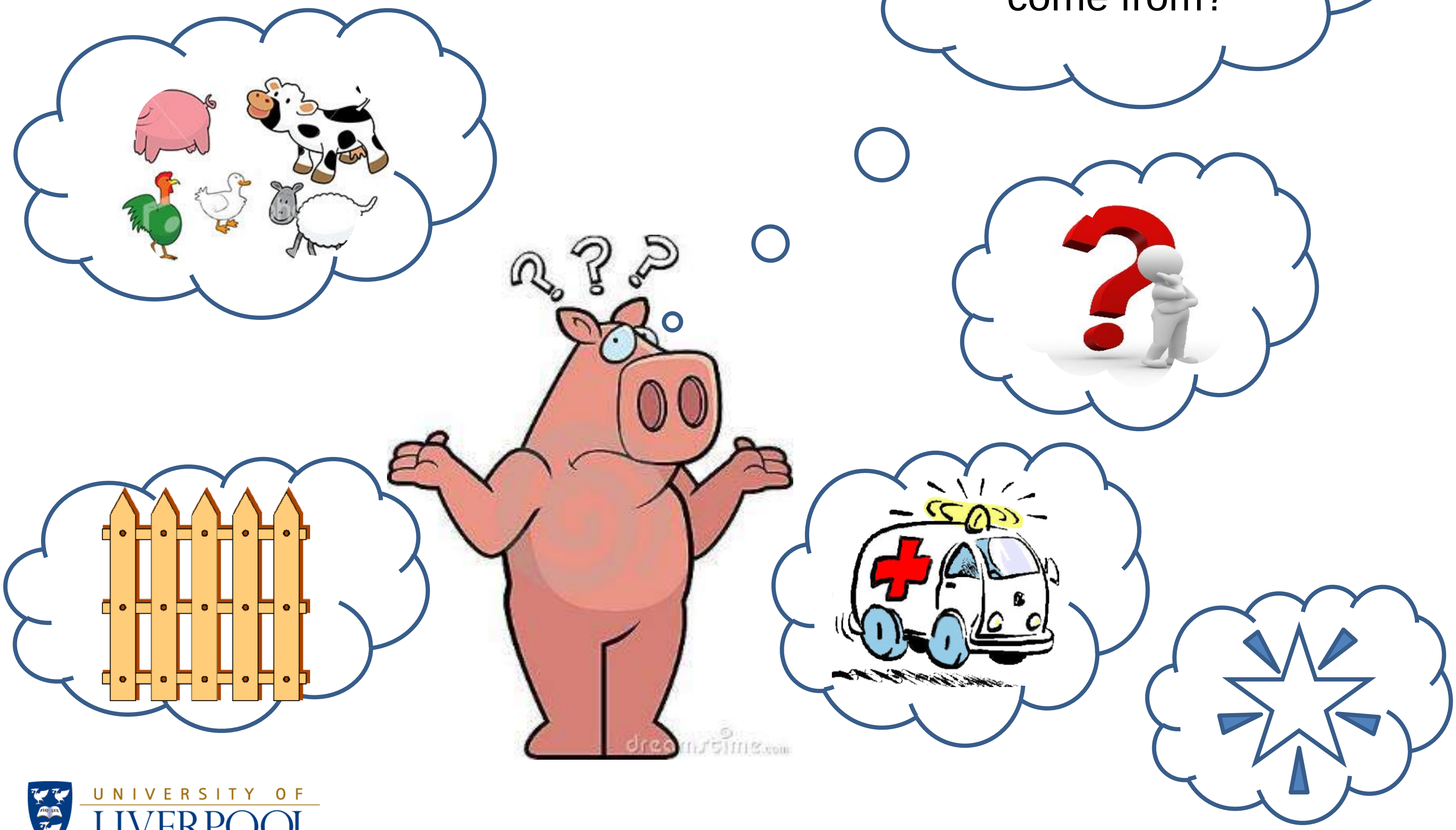
k.m.mcintyre@liv.ac.uk

Risk Assessments Provide…

UNIVERSITY OF LIVERPOOL

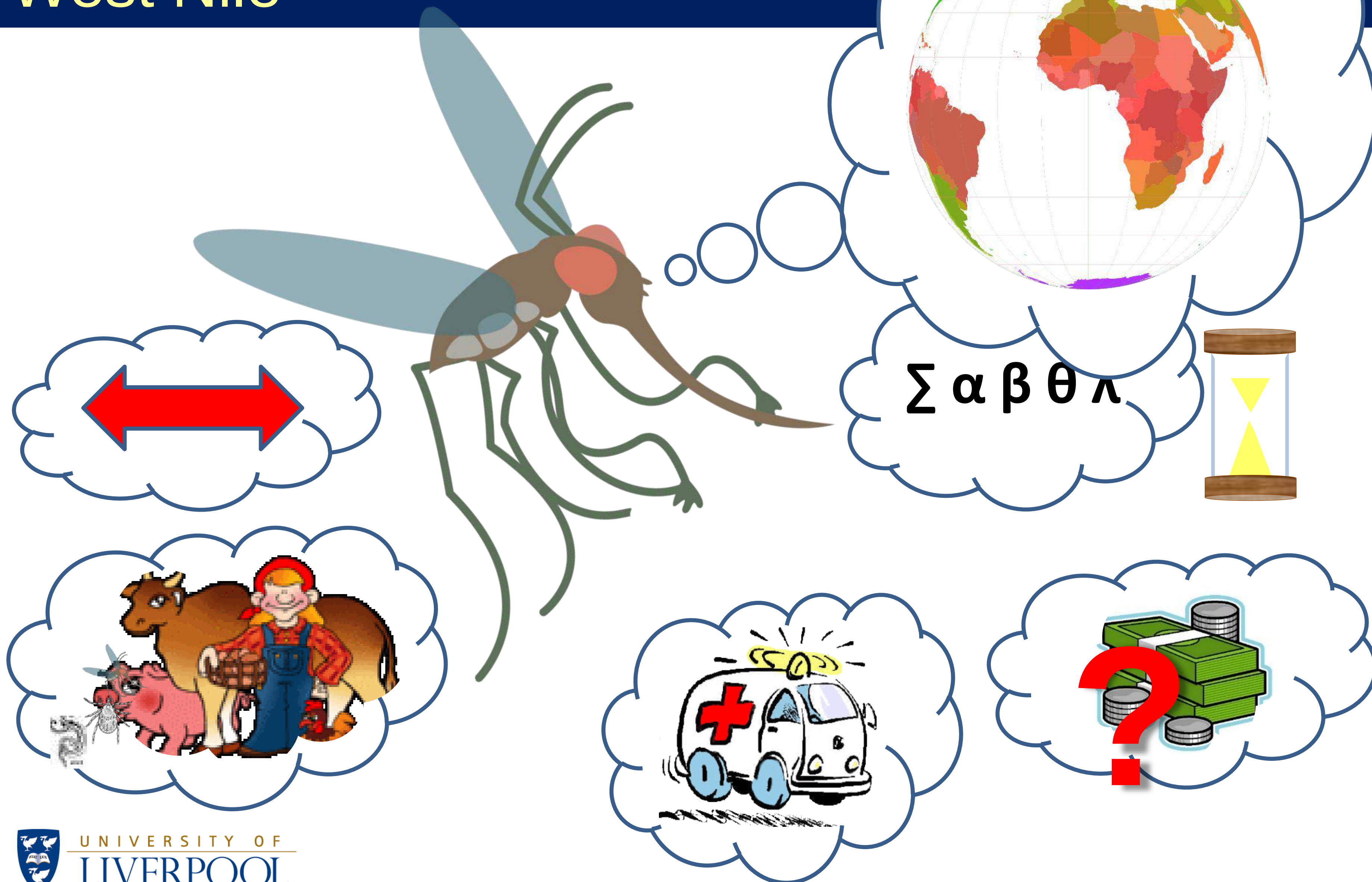
Prepare, Predict, Prevent: Schmallenberg Virus

Where has Schmallenberg come from?



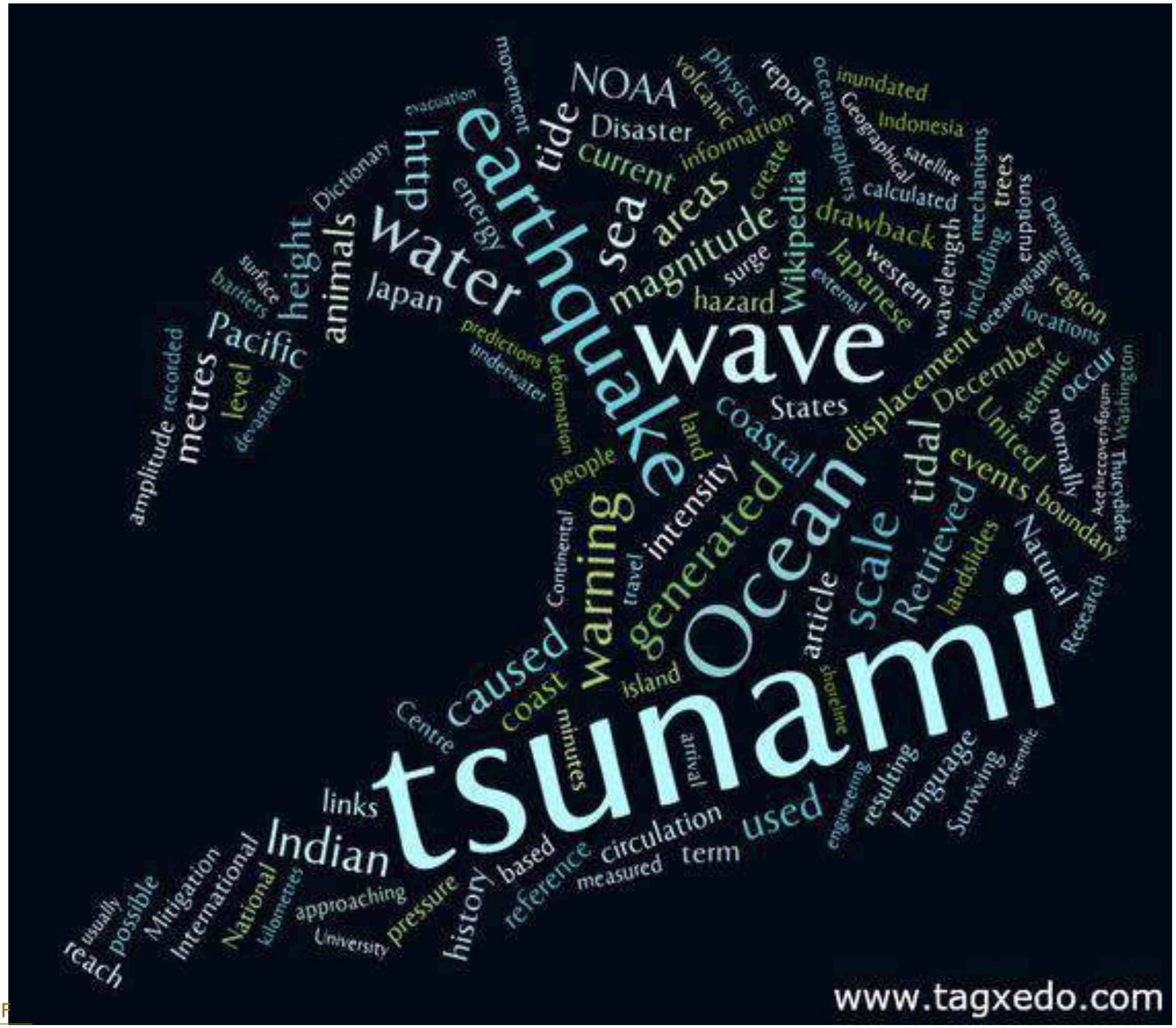
UNIVERSITY OF LIVERPOOL

Prepare, Predict, Prevent: West Nile

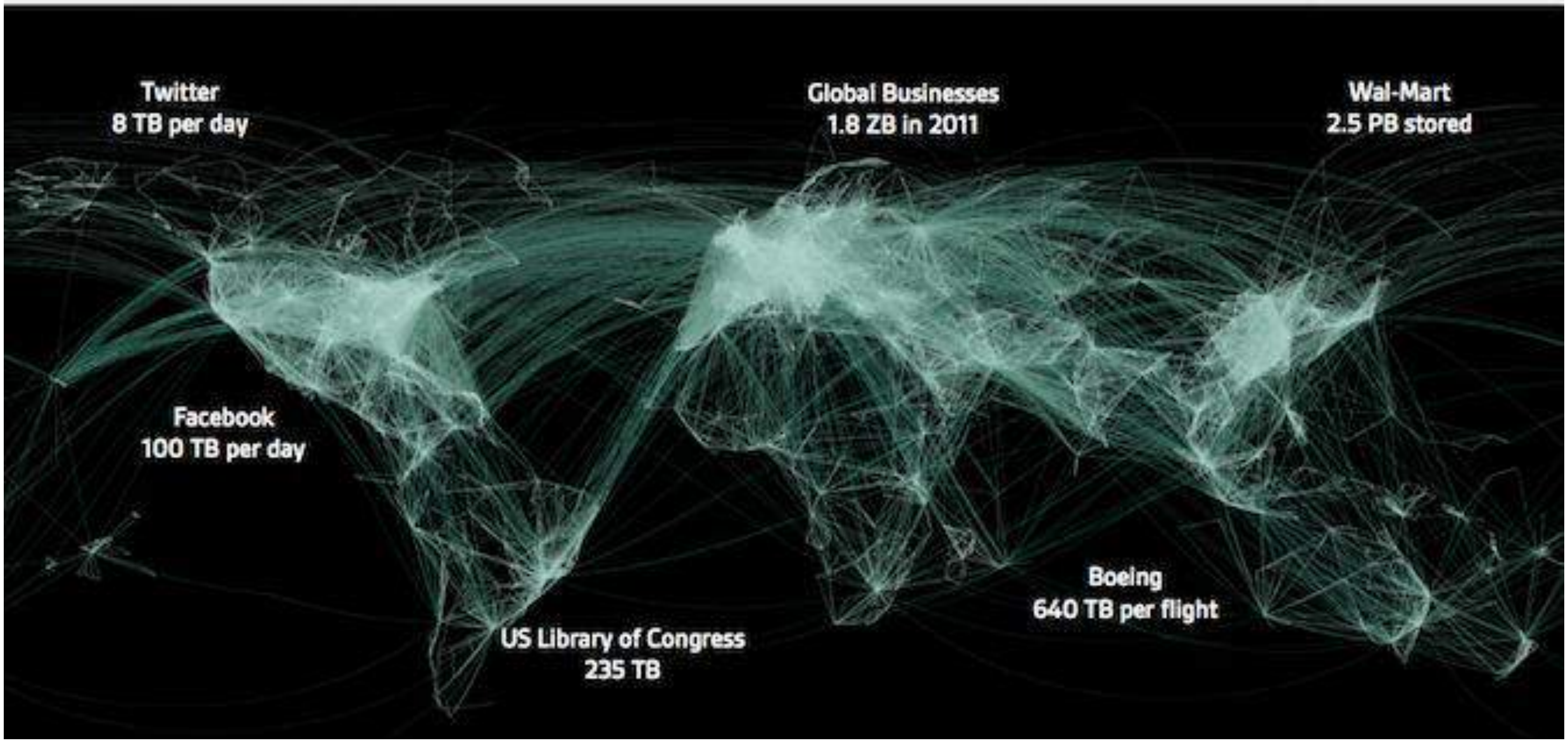


UNIVERSITY OF LIVERPOOL

Studying the Dynamics of Systems…



A Tidal Wave of Data



UNIVERSITY OF LIVERPOOL

ERA-NET Health & Climate in Europe (ENHanCE)

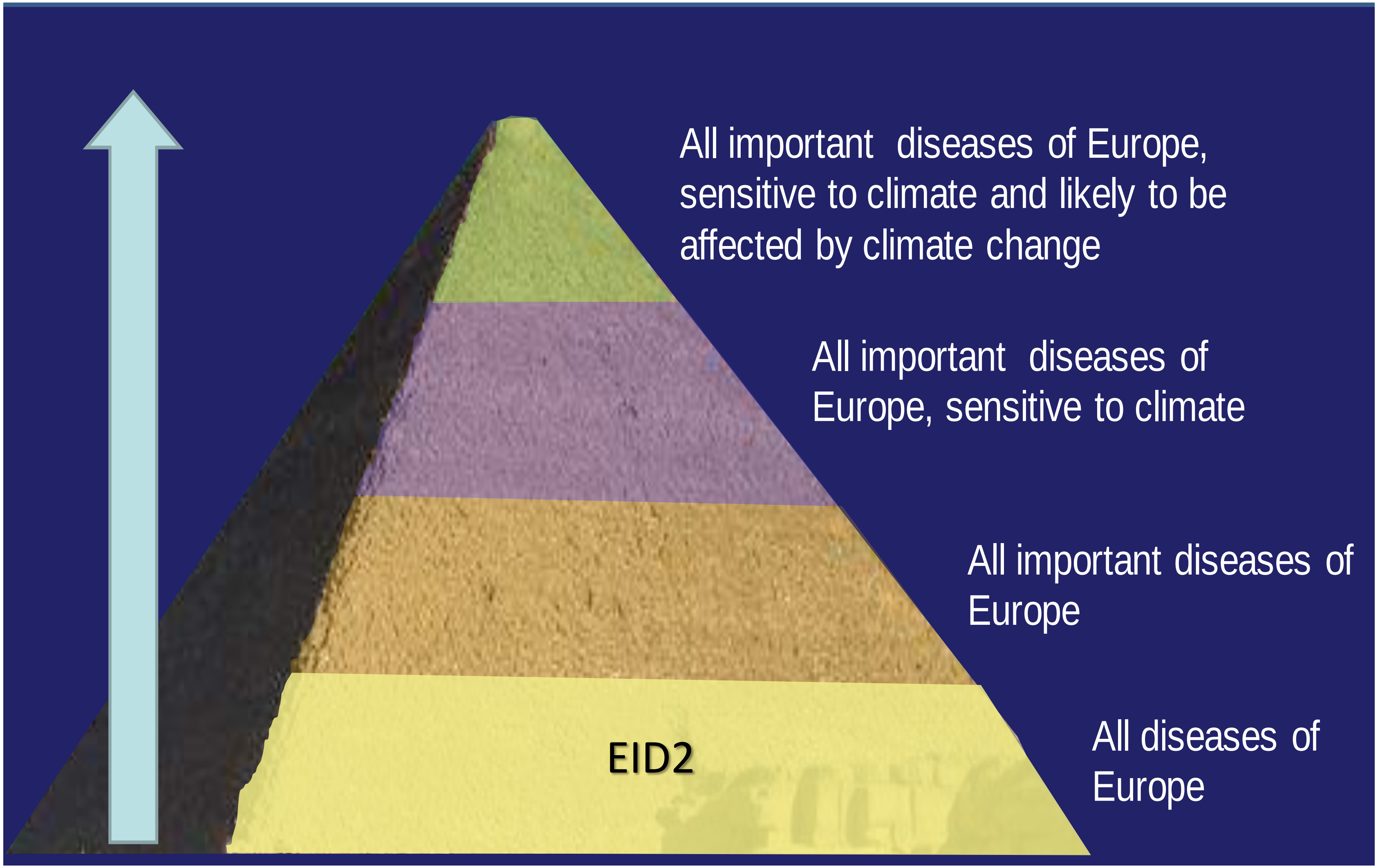
How realistic is the threat to human & animal health from climate change's effects on infectious diseases?

Will most diseases respond to climate change or just a few?

Will there be a net increase or decrease in disease burden?

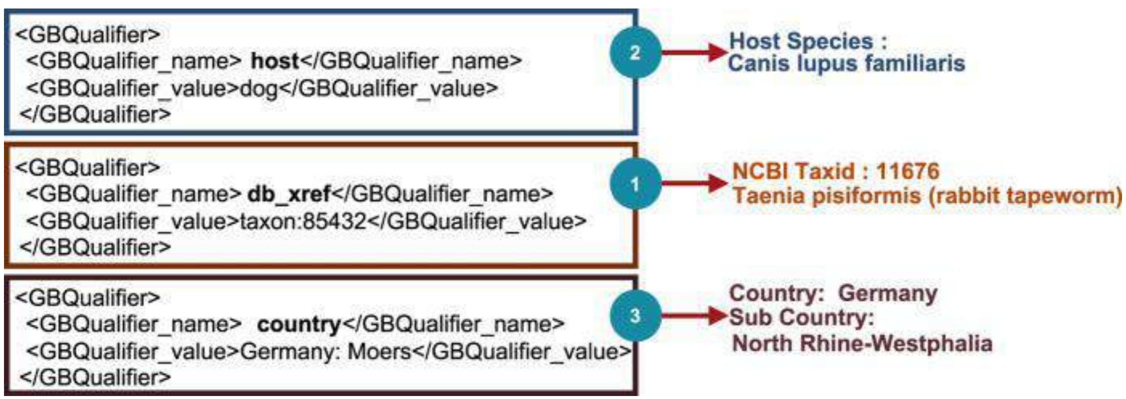
Is it possible that the diseases 'that matter most' are the least likely to respond to climate change?

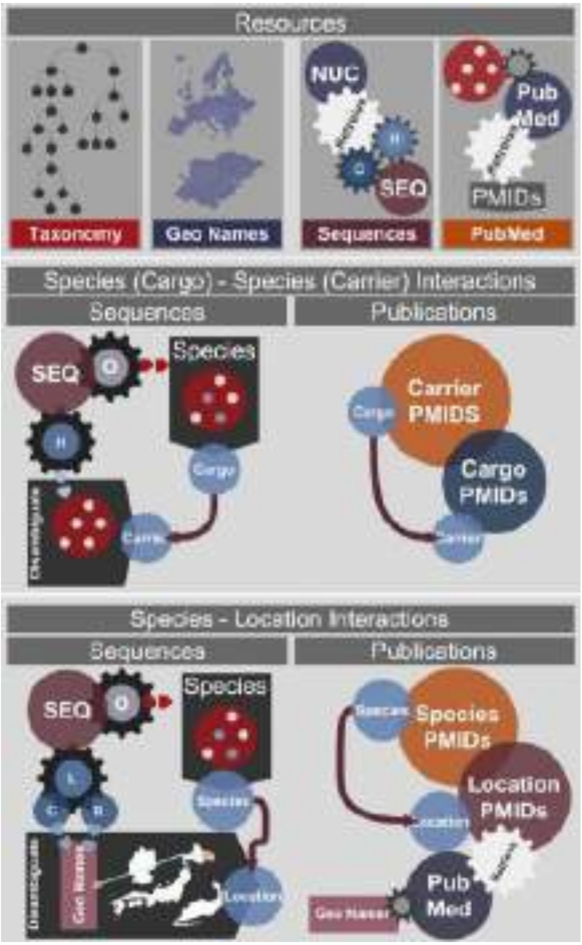
ENHanCE



UNIVERSITY OF LIVERPOOL

ENHanCEd Infectious Diseases (EID2) database
www.zoonosis.ac.uk/eid2

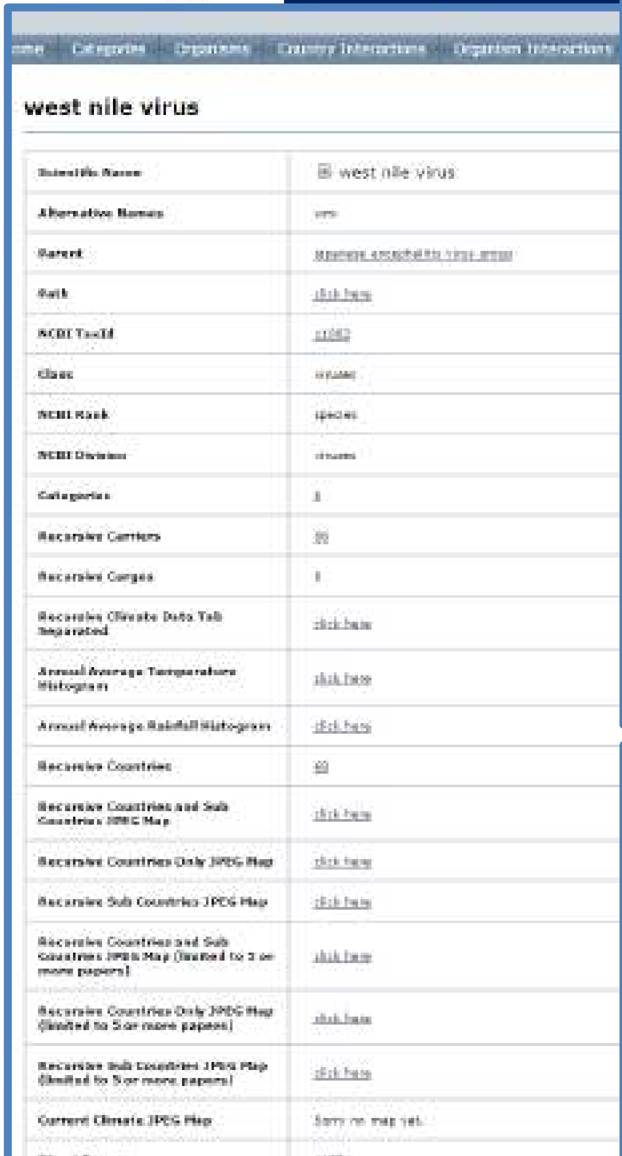


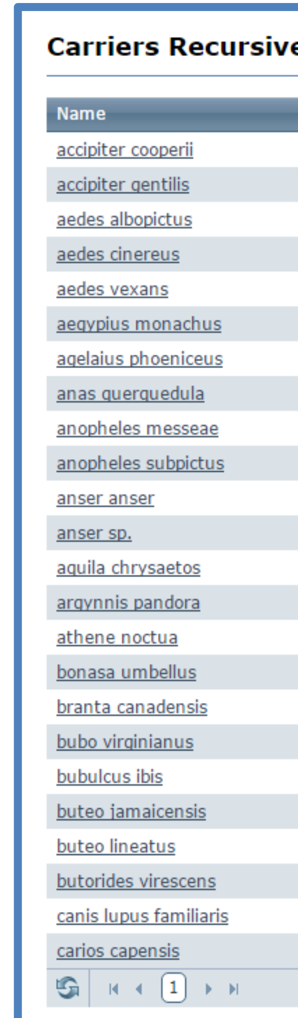


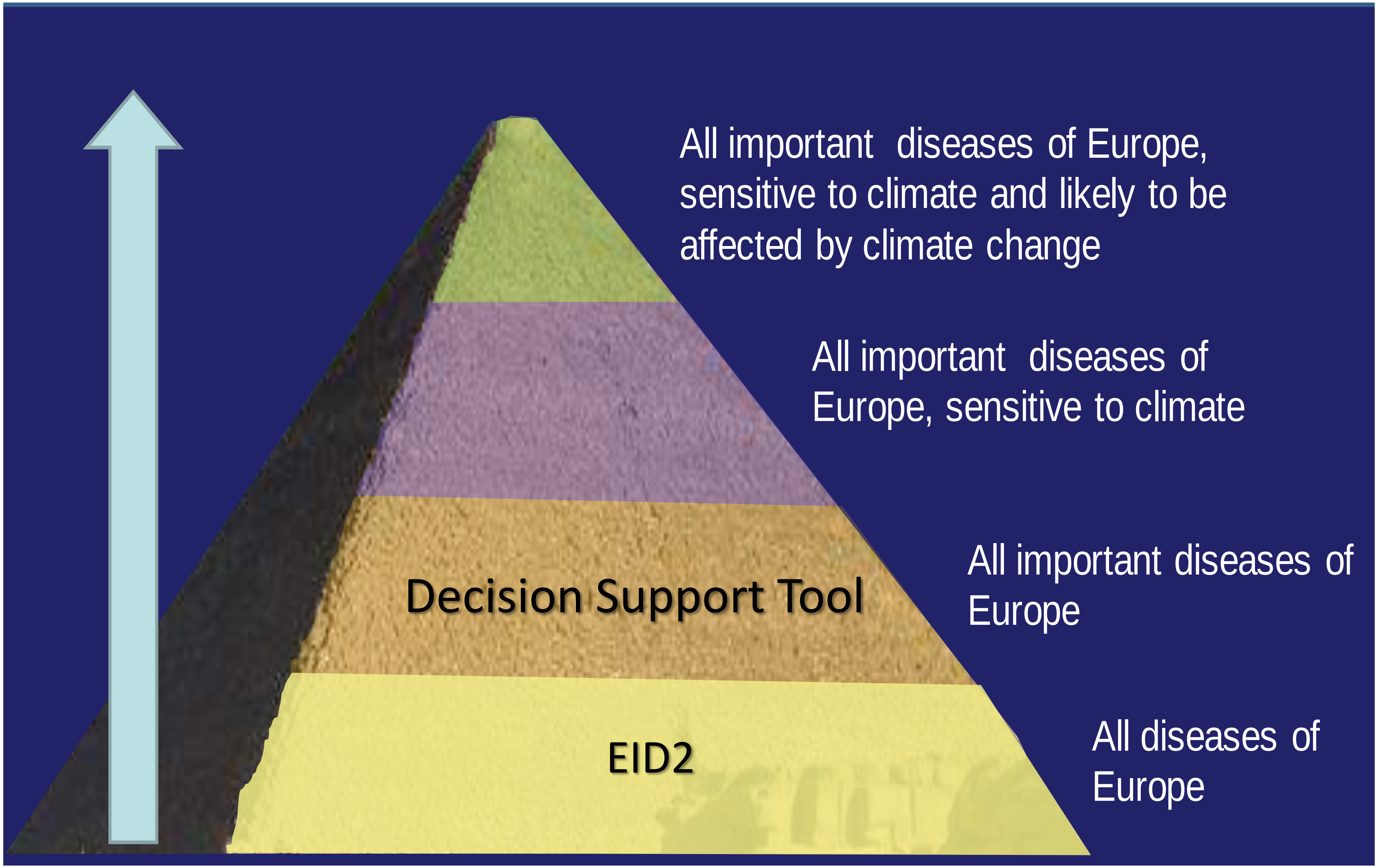
McIntyre et al. (2014) Prev Vet Med 116: 325–335

Wardeh et al. (2015) Scientific Data 2(150049)

ENHanCEd Infectious Diseases (EID2) database
www.zoonosis.ac.uk/eid2





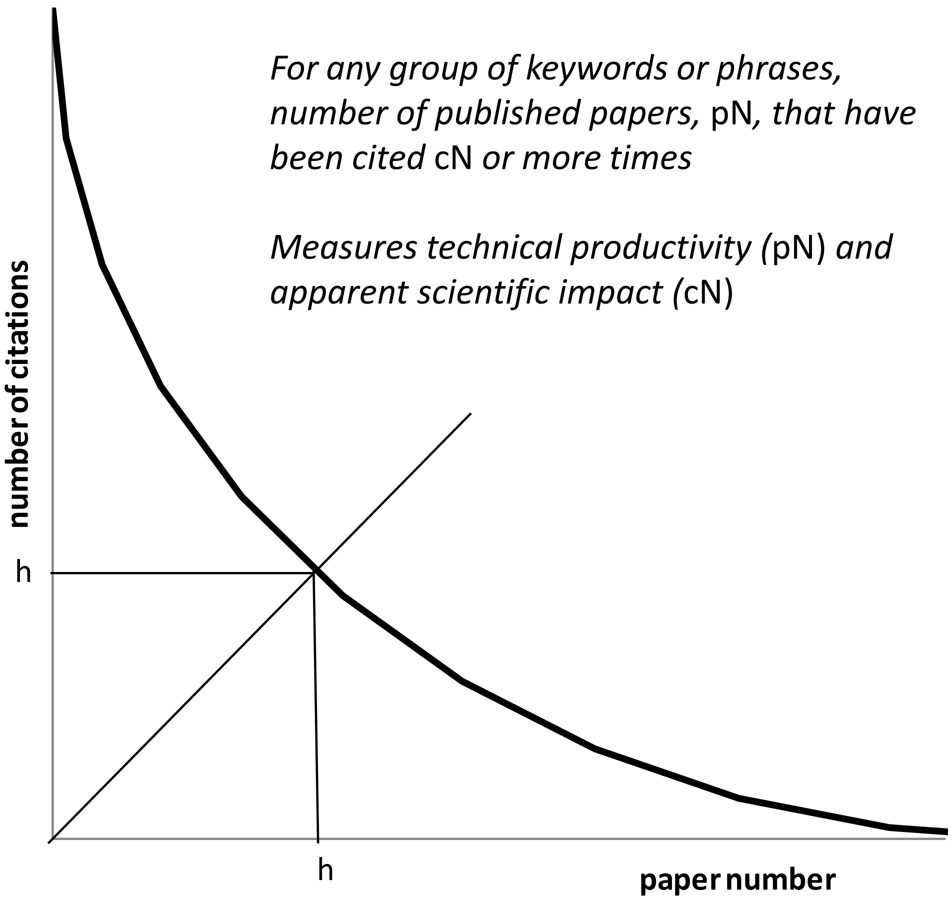


UNIVERSITY OF LIVERPOOL

Which pathogens/diseases have greatest impact?

For any group of keywords or phrases, number of published papers, pN*, that have been cited* cN *or more times*

Measures technical productivity (pN) *and apparent scientific impact* (cN)



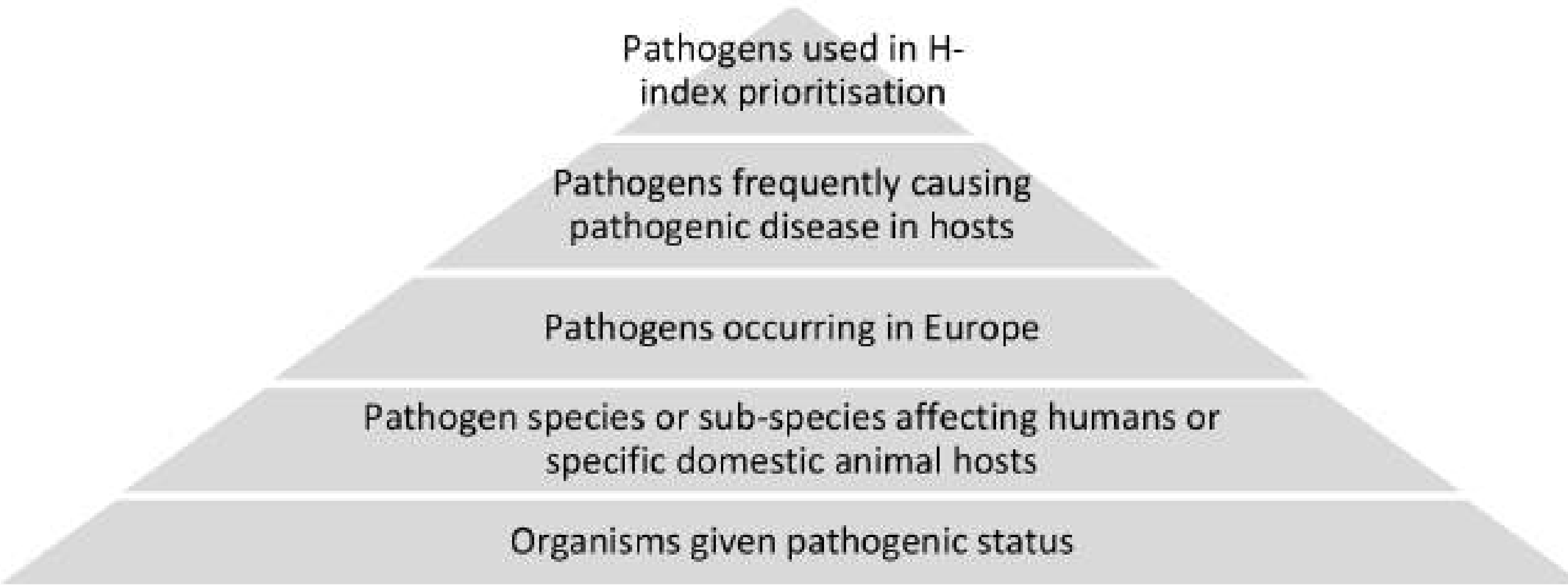
Use the (Hirsch-) H-index to assess

Searches run in WoS

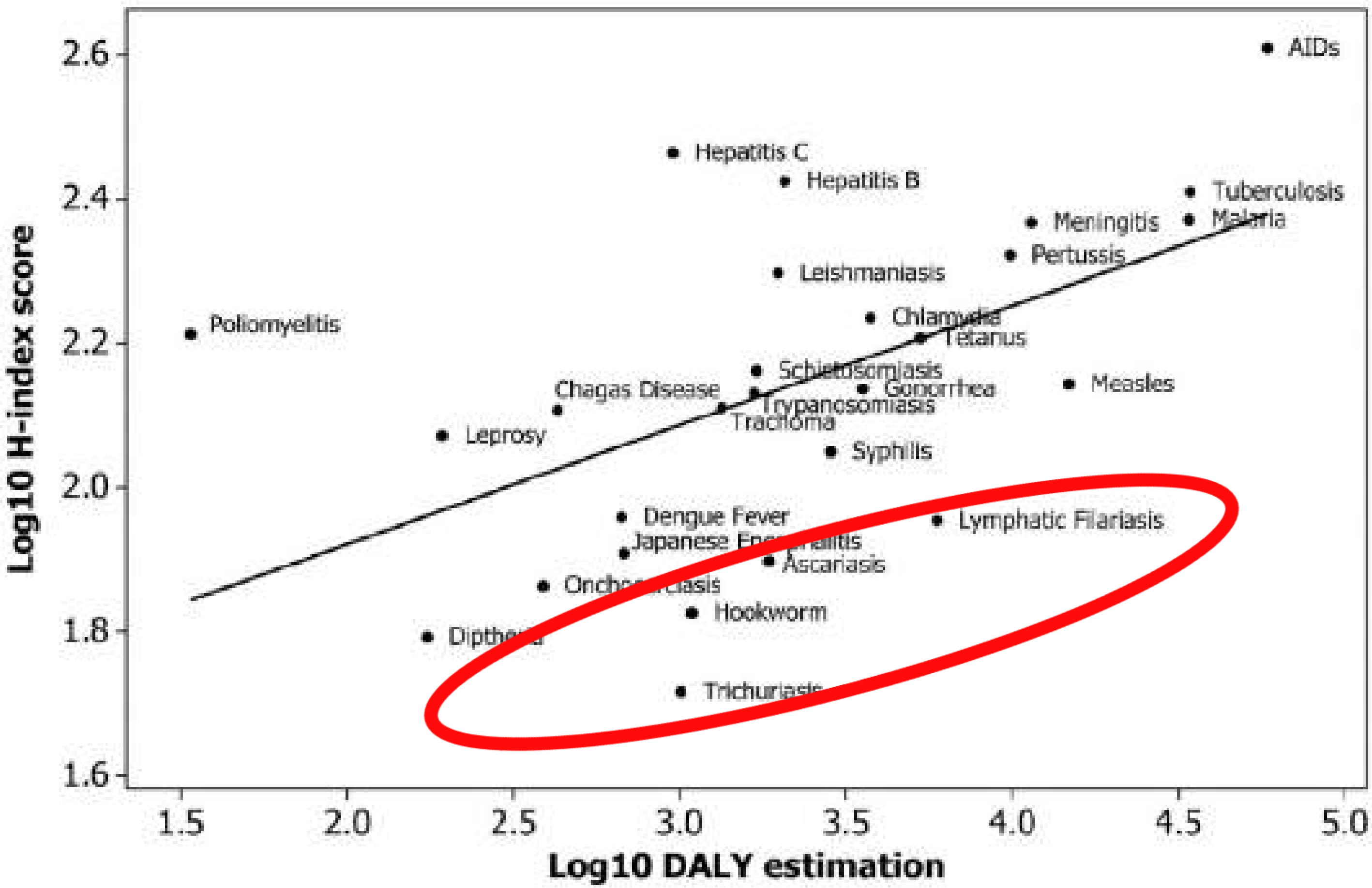
Correlation between H-index scores & other bibliographic indexes

- SCOPUS - P<0.001
- Google Scholar - P<0.001

McIntyre et al. (2011) PLoS ONE 6(5): e19558

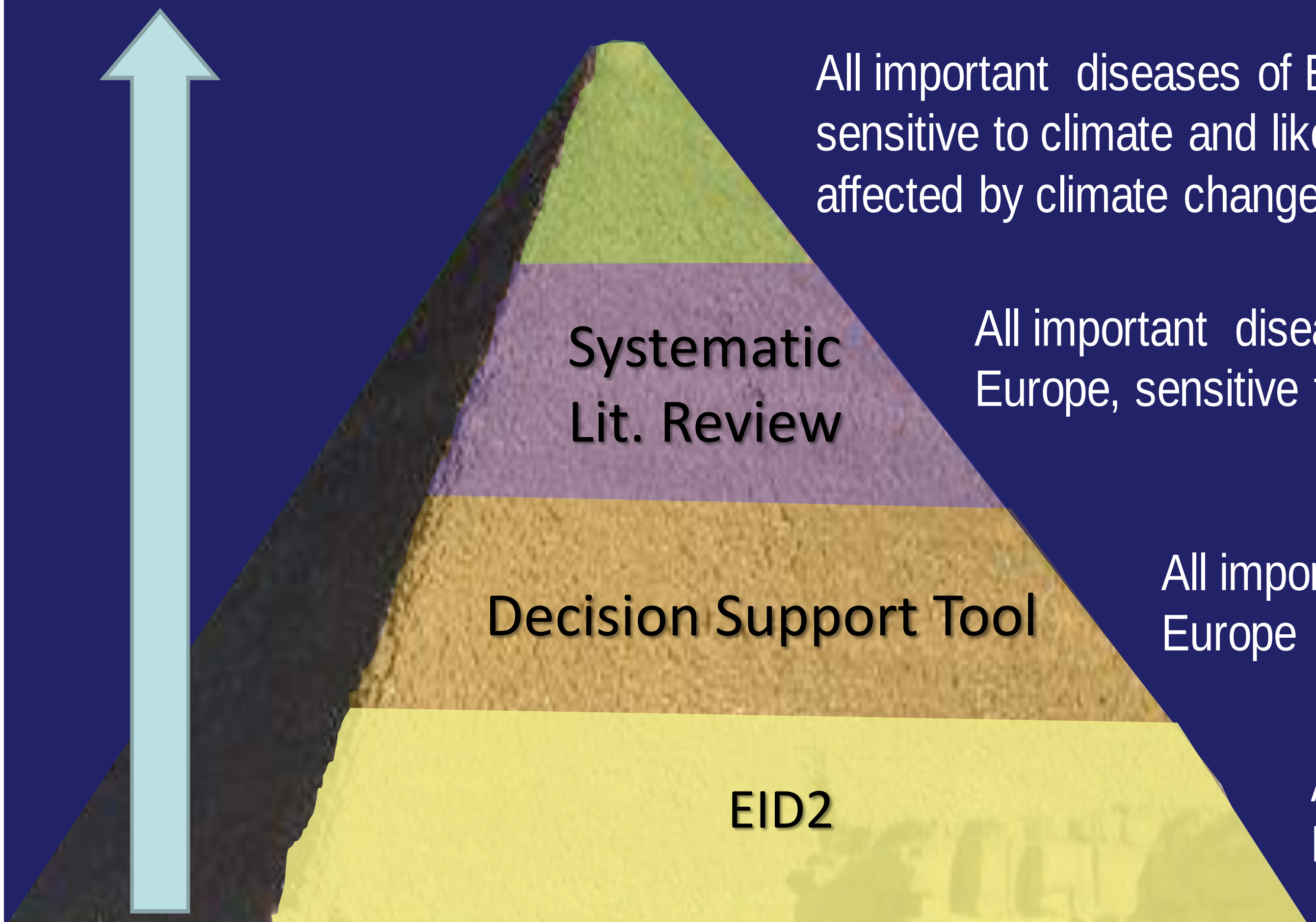

Pathogens used in H-index prioritisation
Pathogens frequently causing pathogenic disease in hosts
Pathogens occurring in Europe
Pathogen species or sub-species affecting humans or specific domestic animal hosts
Organisms given pathogenic status

Which pathogens/diseases have greatest impact?



McIntyre et al. (2011) PLoS ONE 6(5): e19558

ENHanCE



All important diseases of Europe, sensitive to climate and likely to be affected by climate change

All important diseases of Europe, sensitive to climate

All important diseases of Europe

All diseases of Europe

UNIVERSITY OF LIVERPOOL

ENHanCE

Automated searches

Climate driver terms n=190 x pathogens n=157

Combinations n=29830

~27 000 rows describing climate driver term x pathogen

Examined ~5000 rows for evidence

Stratified by:

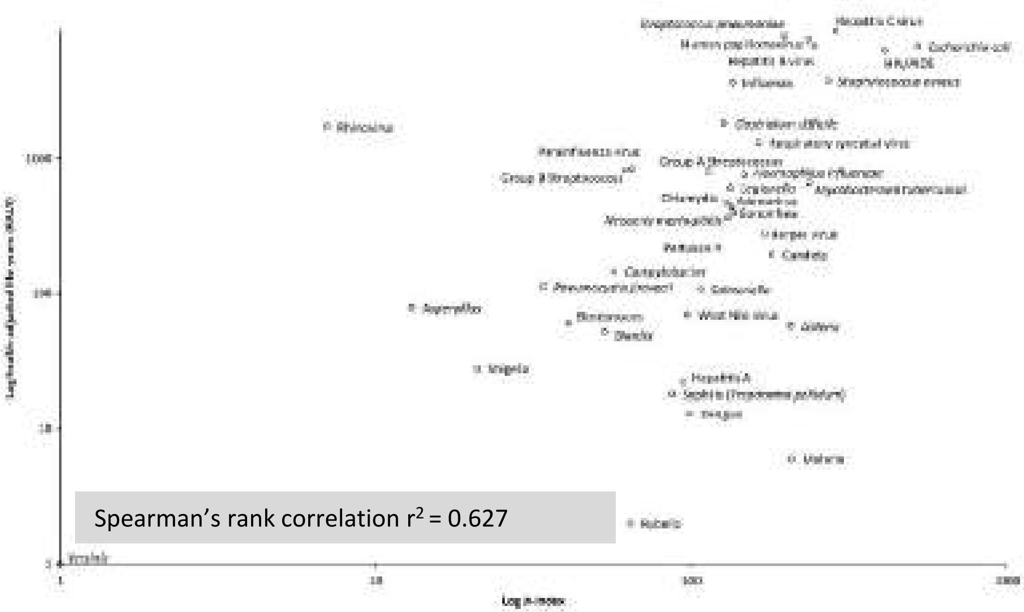
- pathogen
- climate driver

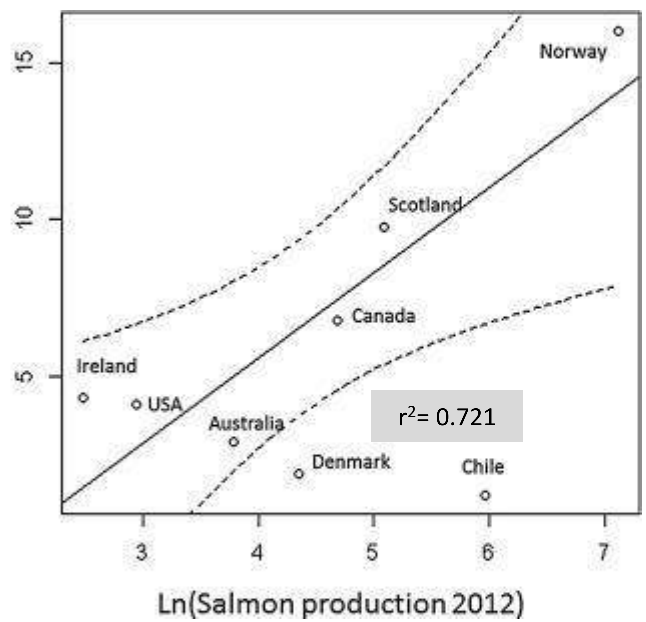
Prioritised according to:

- impact factor
- year of publication

McIntyre et al. (*submitted*)

Which Pathogens/Diseases Have Greatest Impact in N America & Salmon?!



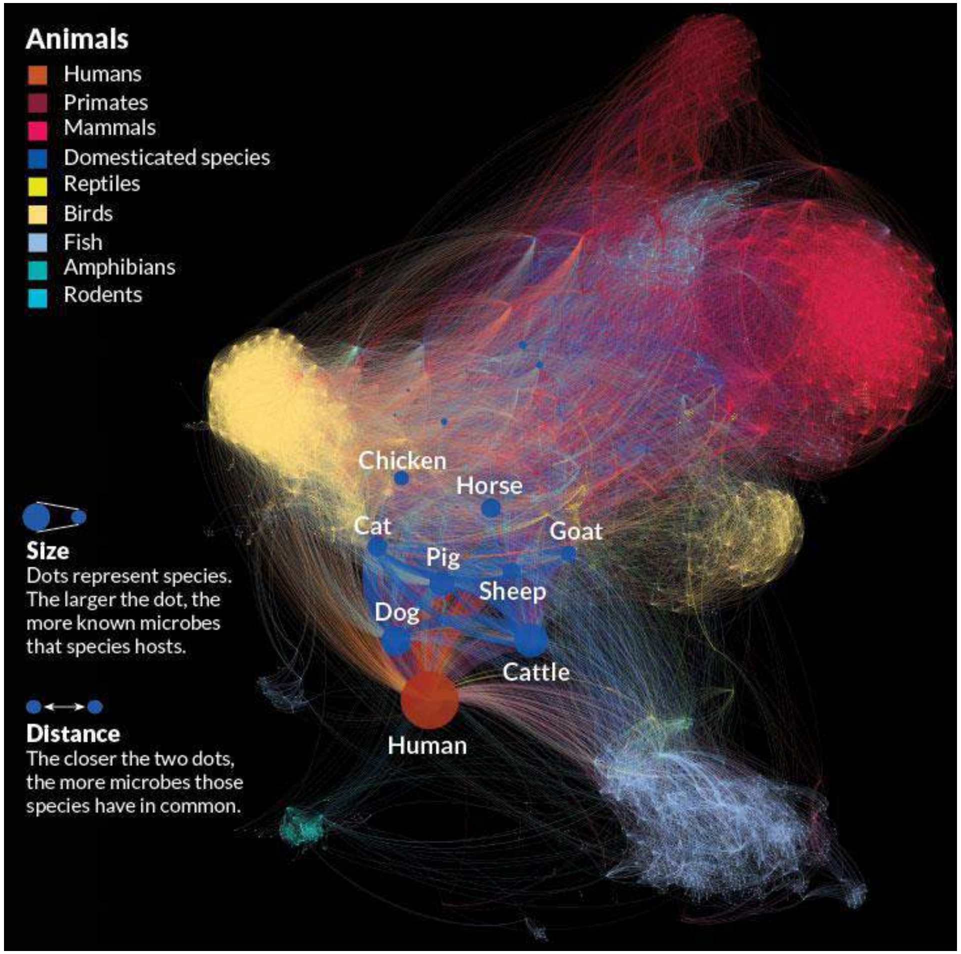


Cox et al. (2014) Trans Emerg Dis 63(1):79-91

Murray et al. (2016)
Prev Vet Med 126:199-207

UNIVERSITY OF LIVERPOOL

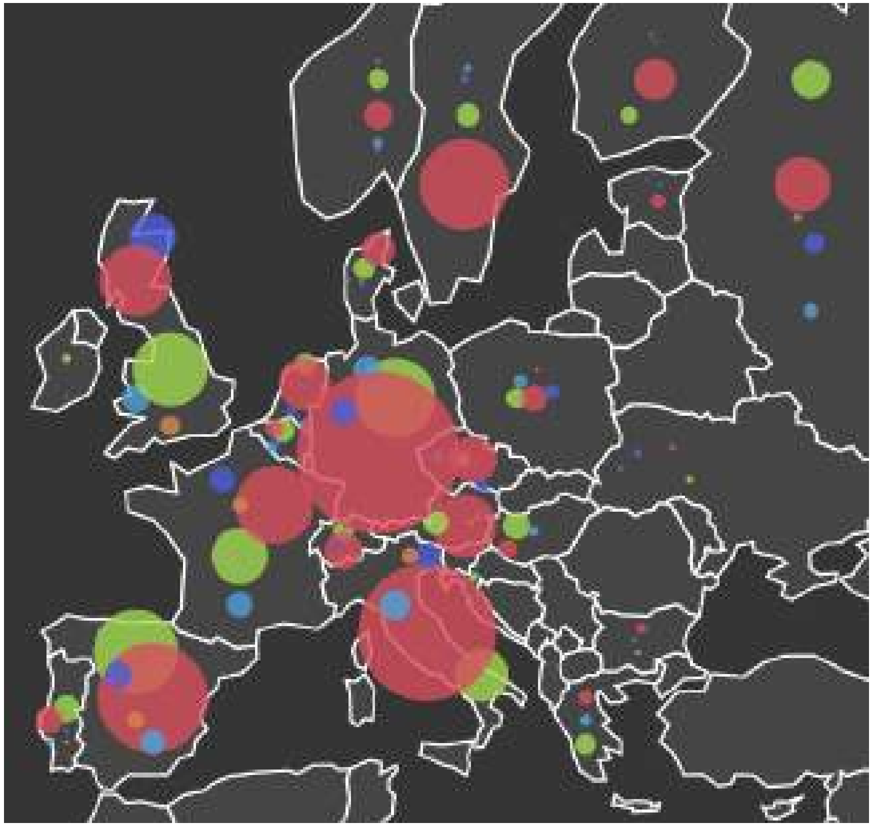
Which Host Groups are Most Important for Disease Transmission?



UNIVERSITY OF LIVERPOOL

Wardeh et al. (2015) Scientific Data 2(150049)
McIntyre et al. (2014) Prev Vet Med 116: 325–335

What Evidence do we Have for Pathogens in EU countries?

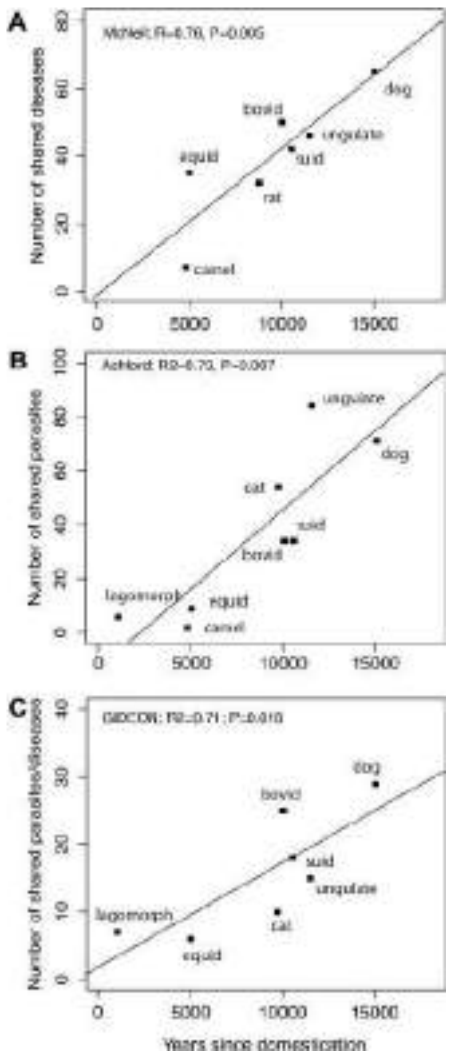


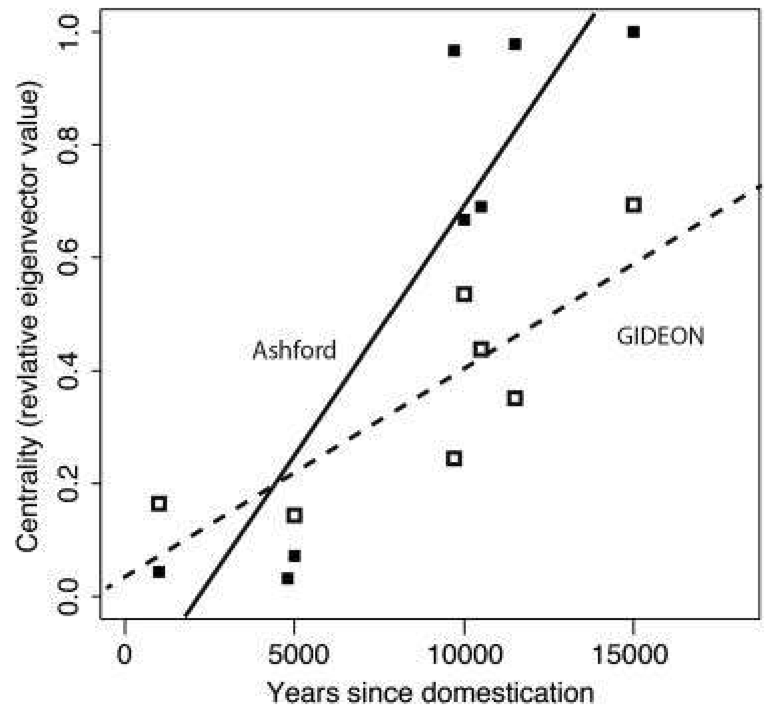
Colours represent taxonomic groups
Size represents number of pathogen species

UNIVERSITY OF LIVERPOOL

http://www.bbsrc.ac.uk/documents/1509-eid2-case-study/

Does Animal Domestication Play a Roll in the Development of Zoonoses?





Domestication time related to:

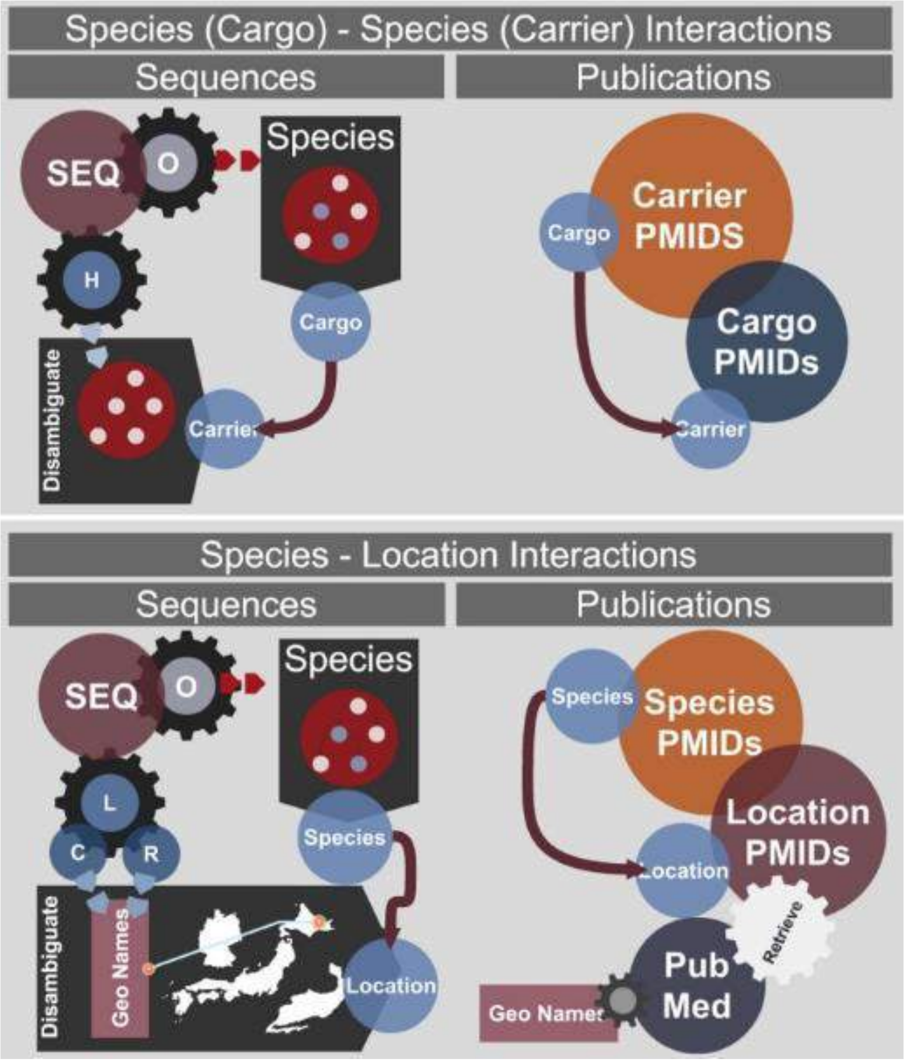
- number of shared pathogens
- host network centrality

Morand et al. (2014) Infect Gen Evol 24: 6–81

Challenges of Big Data Approaches

Accuracy of Incorporating Data

94.5% of carriers identified,
99.44% correctly



60% metadata with country contained additional location information which needed further geo-spatial assignment.

99.99% & 99.65% of countries & regions correctly identified, respectively

Wardeh et al. (2015) Scientific Data 2(150049)

95% of assumed pathogen-country interactions correct

McIntyre et al. (2014) Prev Vet Med 116: 325–335

How Complete is EID2 Host-Pathogen Data Compared to Other Sources?

THE ROYAL SOCIETY

doi 10.1098/rstb.2001.0888

Risk factors for human disease emergence

Louise H. Taylor[*], Sophia M. Latham† and Mark E. J. Woolhouse

Centre for Tropical Veterinary Medicine, University of Edinburgh, Easter Bush, Roslin, Midlothian, EH25 9RG, UK

A comprehensive literature review identifies 1415 species of infectious organism known to be pathogenic to humans, including 217 viruses and prions, 538 bacteria and rickettsia, 307 fungi, 66 protozoa and 287 helminths. Out of these, 868 (61%) are zoonotic, that is, they can be transmitted between humans and animals, and 175 pathogenic species are associated with diseases considered to be 'emerging'. We test the hypothesis that zoonotic pathogens are more likely to be associated with emerging diseases than non-emerging ones. Out of the emerging pathogens, 132 (75%) are zoonotic, and overall, zoonotic pathogens are twice as likely to be associated with emerging diseases than non-zoonotic pathogens. However, the result varies among taxa, with protozoa and viruses particularly likely to emerge, and helminths particularly unlikely to do so, irrespective of their zoonotic status. No association between transmission route and emergence was found. This study represents the first quantitative analysis identifying risk factors for human disease emergence.

THE ROYAL SOCIETY

Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence

S. Cleaveland[*], M. K. Laurenson and L. H. Taylor

[illegible]

[illegible]

McIntyre et al. (2014) Prev Vet Med 116: 325–335

Wardeh et al. (2015) Scientific Data 2(150049)

Data quality & sharing challenges

Need to consider:

- information is presence-only
- spatial resolution
- temporal resolution
- efficiency of data collection
- specificity & sensitivity of methods
- choice of keyword/phrases
- quality of & evidence available within secondary sources
- comparison (or combination) with manual (primary data) methods
- institutional inertia to systems-based approaches

Contributors

University of Liverpool

Prof Matthew Baylis
Dr Cyril Caminade
Dr Phil Hepworth
Prof Andy Morse
Dr Alan Radford
(Dr Claire Risley)
(Dr Christian Setzkorn)
Dr Maya Wardeh

Maastricht University

Dr Su-Mia Akin
Dr Maud Huynens
Prof Pim Martens

cirad

Dr. Serge Morand
(Dr. Agnès Waret-Szkuta)

marine scotland science — The Scottish Government

Dr Sandy Murray

University of Prince Edward Island

Dr Ruth Cox
Prof Crawford Revie
Dr Javier Sanchez

Natural Environment Research Council

afsset

Ministry of Housing, Spatial Planning and the Environment

BBSRC bioscience for the future