Animal &
Plant Health
Agency

The need for harmonised data to ensure comparability of generic risk assessment

Rachel Taylor, Robin Simons

G-RAID Workshop, The Netherlands, 15/05/2019

Why do we want to harmonise data?

- To compare results between risk assessments we want to start from the same place
  - Differences in results due to methodology NOT input data.

- Risk Assessments have many input parameters
  - E.g. Prevalence of pathogen, Volume of trade, presence of animals

- Data for each parameter could come from many sources
  - Prevalence data from OIE WAHIS, Empres-I, Healthmap

Why do we want to harmonise data?

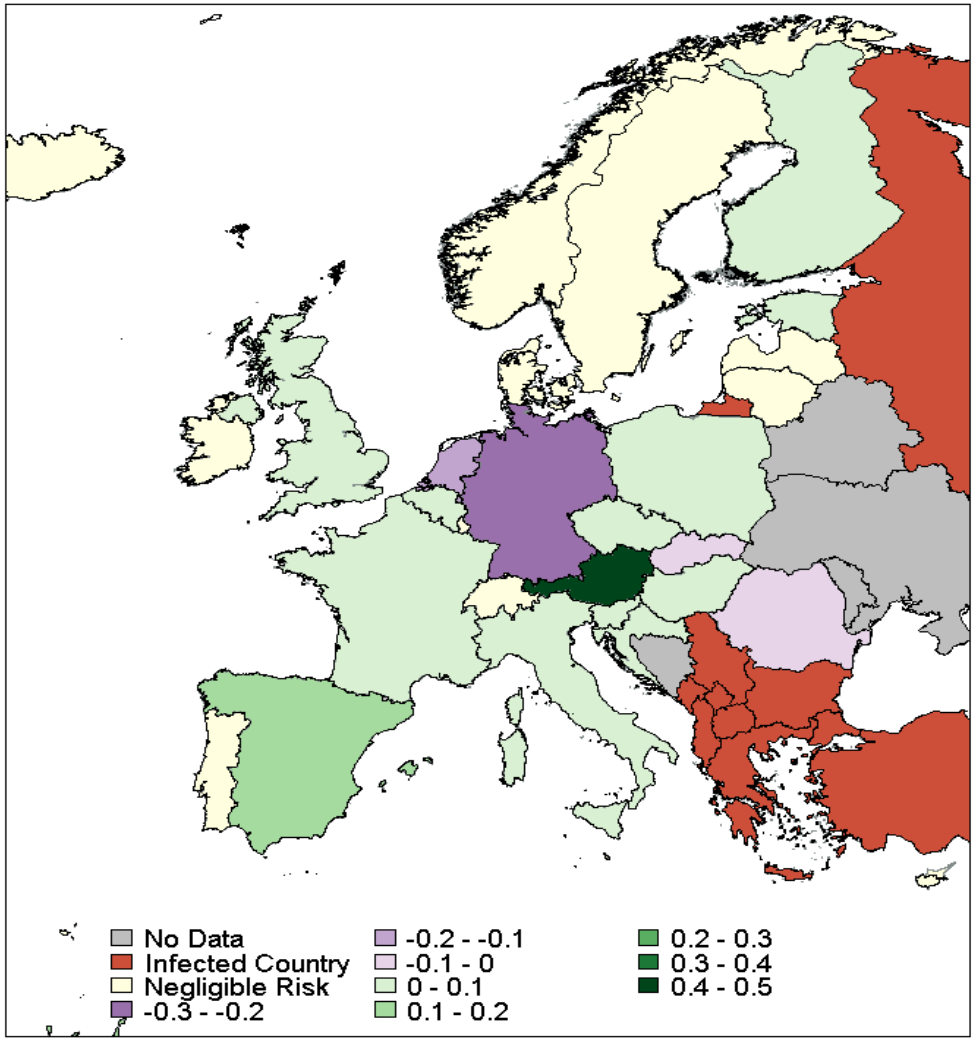
- Different sources may have different methodology
  - OIE based on reported cases only,
  - Healthmap includes public observations
  - Different estimates for input data

- Save time and resources retrieving and formatting data

Differences in data sources

Probability of infection for Lumpy Skin Disease in 2016 using COMPARE model

- Positive values (in green):
  - TRACES higher risk
- Negative values (purple):
  - COMEXT higher risk.



Harmonisation but…
the models are different

- All of the models reach a different endpoint
- All of the models perform different steps to reach endpoint
- Some models return a probability, others a number, others a risk score
- Some work on national level, others European
- Some incorporate multiple diseases, others work for one disease at a time

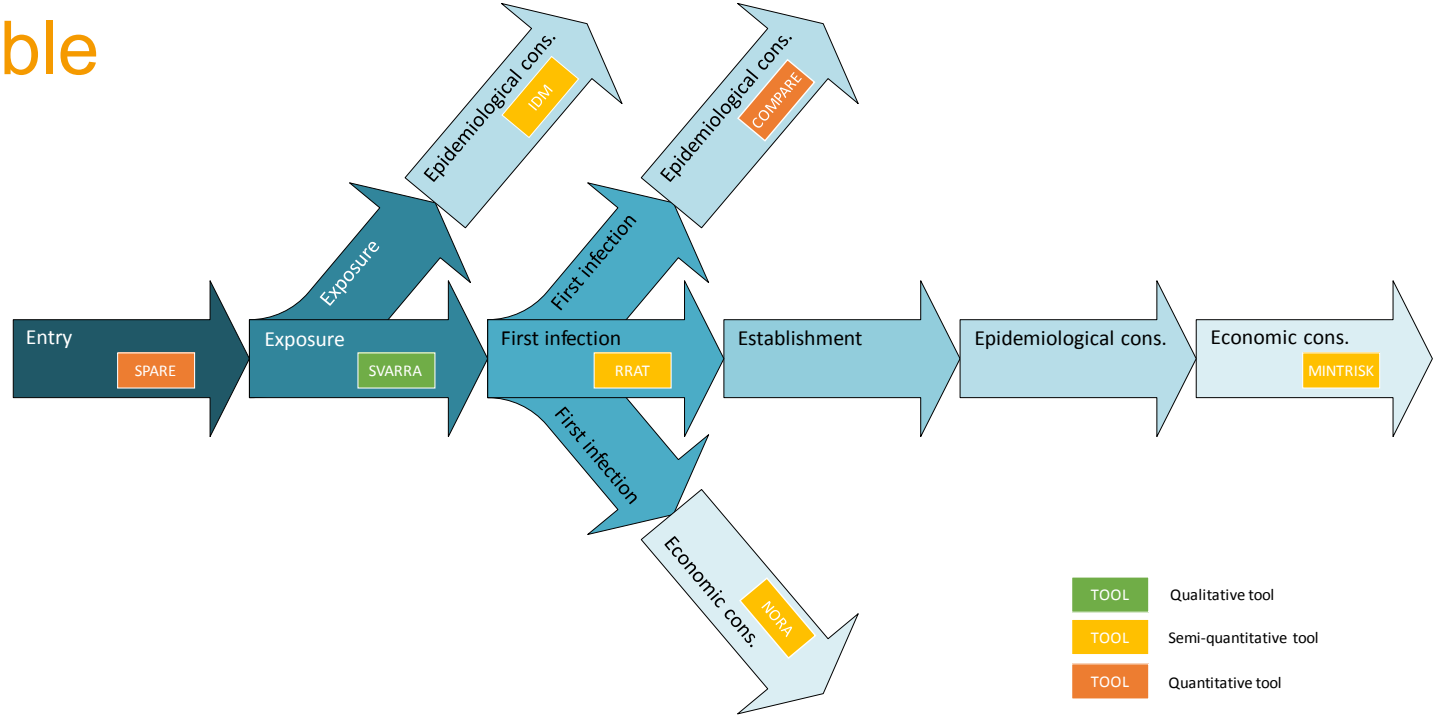
So how do we harmonise our data then?

Assess algorithms in the tools

→ find similarities

→ what input is needed when similarities occur?

→ choose same data if possible



What are the similarities?

- Principles of the Binomial model used for calculating entry for all
  - Data on volume along a pathway and probabilities of infection based on prevalence
  - NORA does not take into account prevalence
- Calculating the volume along a pathway
  - Models most similar for trade pathways
  - For other pathways different algorithms are used
  - Including using national data, proxy data or producing a sub-model (e.g. for wild animal movement)

Similarities continued

- All models other than SPARE estimate exposure or infection
  - Combine entry with exposure/infection to have a final risk estimate
  - This could be by multiplication, a risk matrix or a stochastic simulation
- Likelihood of disease reaching a farm/susceptible animal
  - SVARRA and RRAT split susceptible animals into classes e.g. dead-end hosts
  - COMPARE, NORA and MINTRISK use data sources directly or indirectly on abundance of susceptible animals
- COMPARE, IDM and MINTRISK use R_0 to estimate transmission
  - IDM uses published literature for value
  - MINTRISK uses R_0 for initial incursion, establishment and epidemic size
  - COMPARE has different formulas for R_0 depending on transmission pathways and disease

Data types to try to harmonise

- Movement from one area to another
- Prevalence in area of origin
- Susceptible animals in the target region
- Disease-related parameters

Prevalence

- What?
  - Global prevalence of disease
  - Number of reported outbreaks and cases of ASF
  - Split by species (domestic pigs and wild boar), country, year
- Where?
  - OIE WAHIS
    - Pros: whole world, officially reported data, multiple diseases, many years
    - Cons: bulk download hard (no API or access to database)
  - Empres-i
    - Pros: whole world, multiple sources (e.g. OIE, promed)
    - Cons: limited pathogens, no API for bulk download
  - Healthmap
    - Pros: whole world, multiple pathogens, multiple sources
    - Cons: quality control, no API for bulk download, limited years

Prevalence

World Organisation for Animal Health

- Harmonisation
  - Which?
    - OIE, WAHIS
  - Why?
    - Official data, multiple years
    - SPARE had collated most of the data already
  - How?
    - Historical data from annual reports
    - Data from the last year from the weekly disease information page
    - Data collated by SPARE and shared within consortium
    - Up to models how to transform raw data into prevalence

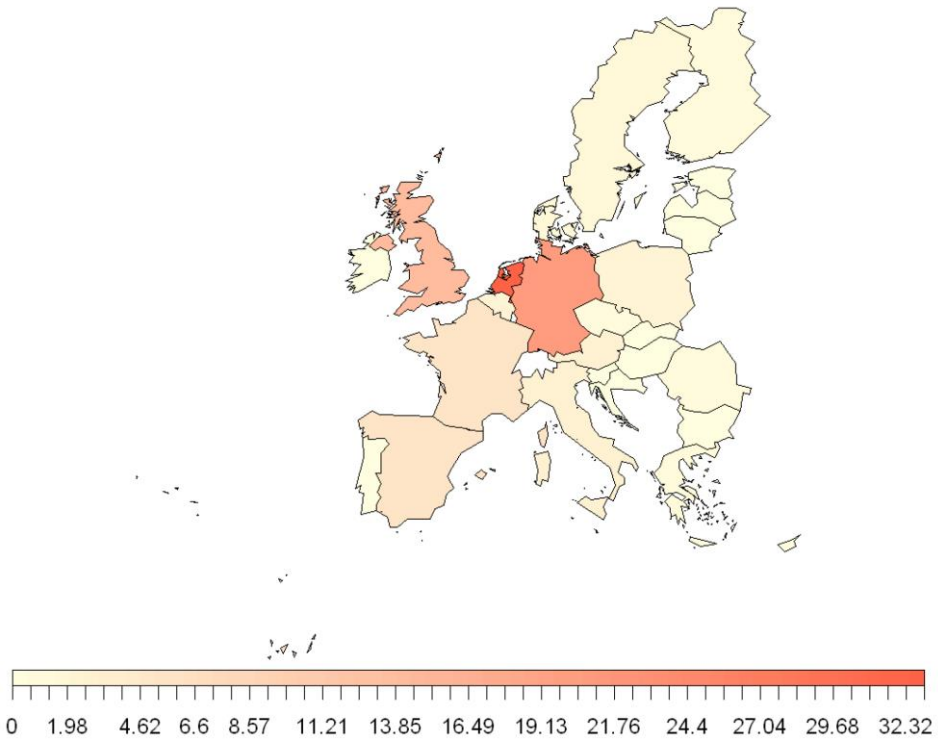
World Animal Health

Movement (Trade)

- What?
  - Trade of live animals, meat products
- Where?
  - Eurostat (comext)
    - Pros: Freely available, API to bulk download, updated frequently
    - Cons: EU MSs only, gaps for low numbers, not always up to date
  - Comtrade
    - Pros: Freely available, whole world, updated frequently
    - Cons: gaps for low numbers
  - FAOstat
    - Pros: Freely available, whole world, updated frequently
    - Cons: API not straightforward,
  - TRACES
    - Pros: Very detailed (can get down to postcode)
    - Cons: EU MSs only, require password, hard to bulk download, formatting required.

Movement (Trade)

- Harmonisation
  - Which?
    - Eurostat (comext)
  - Why?
    - Everyone can access
    - Only looking at EU
  - How?
    - Full 2017 trade data downloaded from bulk download page
    - Selection of which products to include left up to individual models



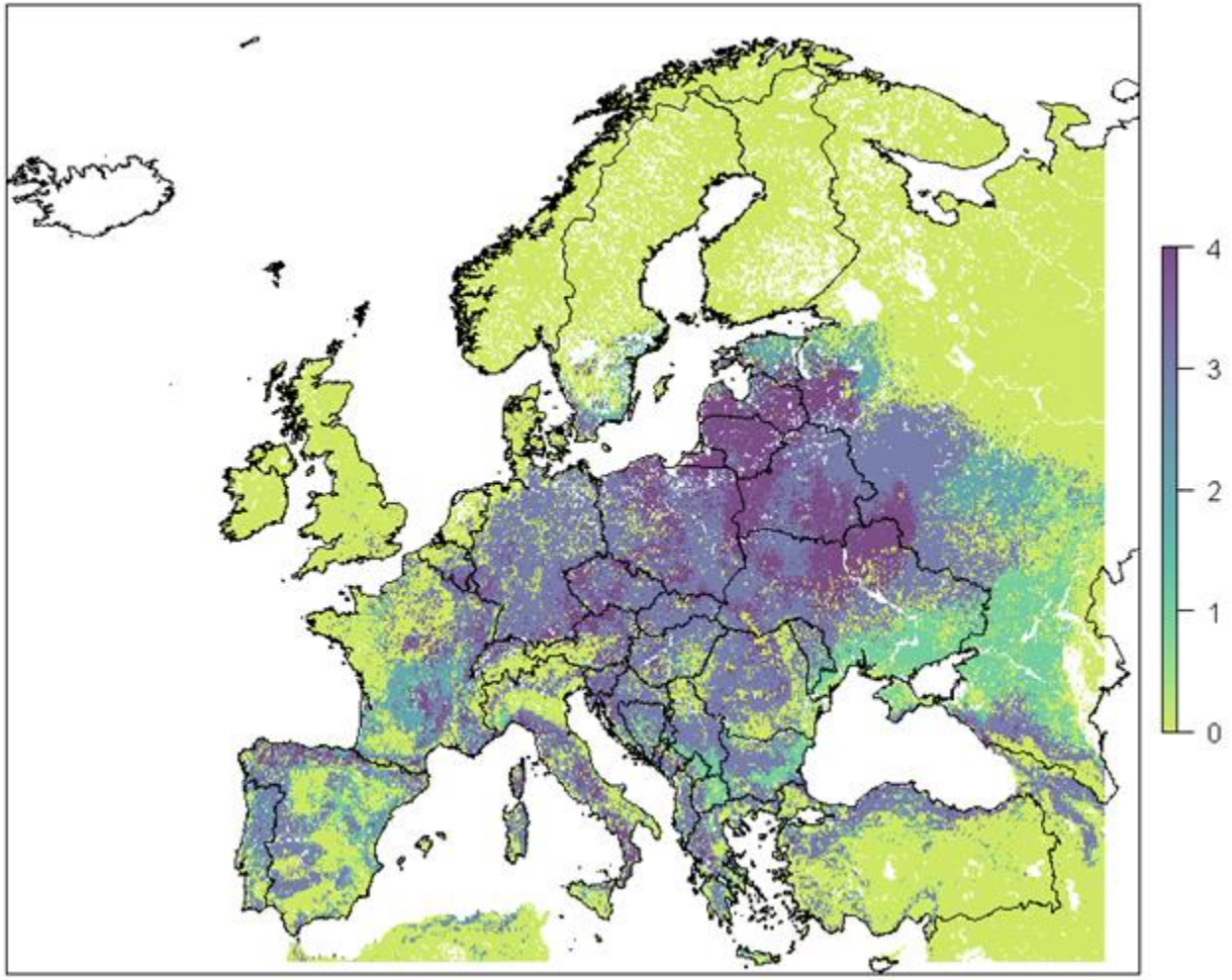
Percentage contribution of 2010 EU MS fruit and pig product trade from NiVregions. Simons *et al, viruses, 2014*

Movement (animal density):

- What?
  - Density, location of animal species
- Where?
  - FAO, Gridded livestock of the World
    - Pros: Freely available, peer reviewed, georeferenced
    - Cons: livestock only
  - Global Biodiversity Information Facility
    - Pros: Freely available, georeferenced
    - Cons: Presence points only, includes museums, onerous referencing for reports
  - Wild boar density map
    - Pros: peer reviewed publication,
    - Cons: wild boar only

Movement (animal density):

- Harmonisation
  - Which?
    - FAO & wild boar maps
  - Why?
    - Difference between models
  - How?
    - Manual download

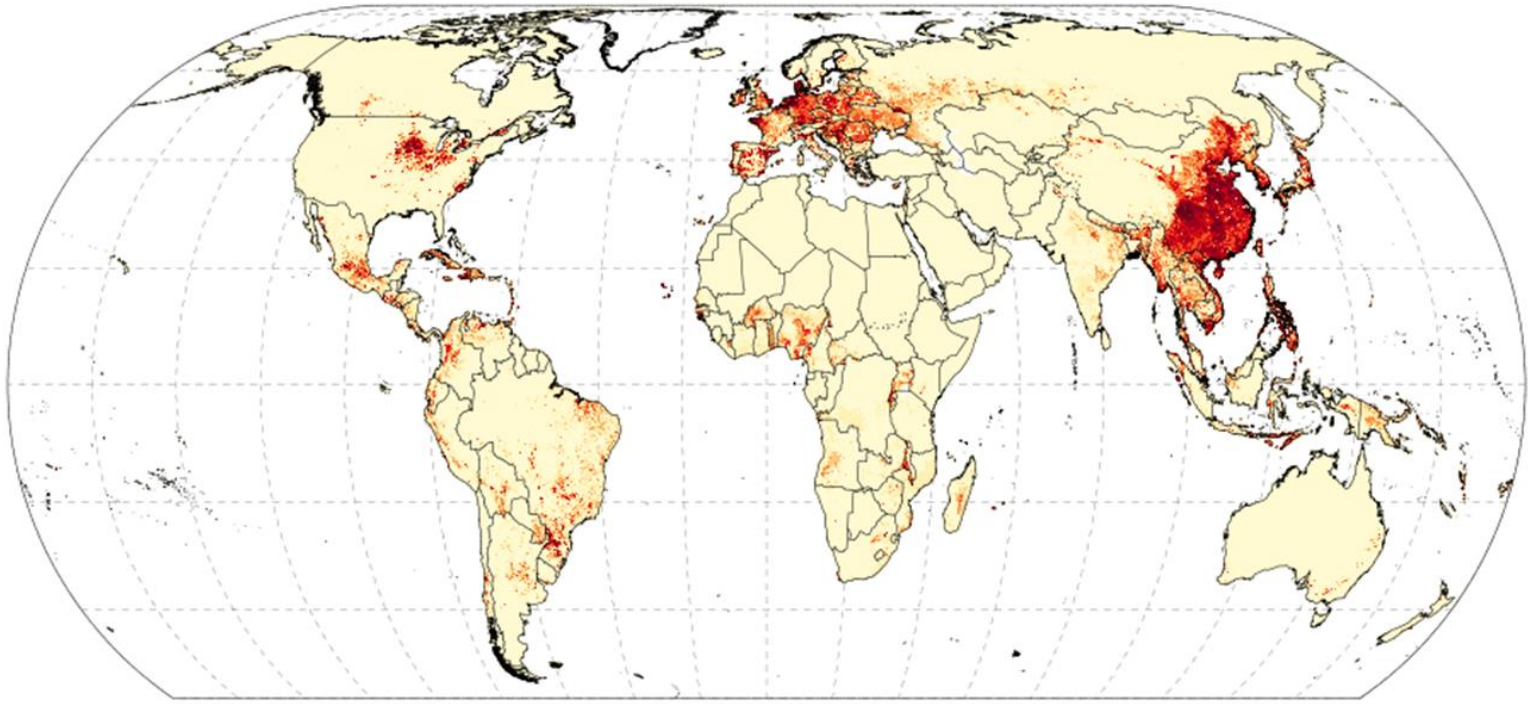


Susceptible Animals

- What?
  - Varies between models due to methodology
    - Susceptible animals in the target region
    - Whether contact/infection could occur
    - Population abundance maps at a fine scale
    - Regional data on average farm size
- Where?
  - National farm registries
  - FAO, Gridded livestock of the world
  - Published literature
  - National databases
  - Expert Opinion

Susceptible Animals

- Harmonisation
  - Which?
    - Combination of all
  - Why?
    - Differences between models
  - How?
    - Manual download
    - National data shared within the consortium



Pig Density, FAO Gridded livestock of the world map, FAO (2019), http://www.fao.org/livestock-systems/en/

Disease Specific

- What?
  - Varies between models, E.g.
    - Duration of clinical infection,
    - Time to clinical signs,
    - Survival time in different media,
    - Transmission rates between animals
- Where?
  - Peer reviewed journal articles
  - Official online sources such as WHO, CDC
  - National statistics
  - Expert opinion

Disease Specific

- Harmonisation
  - Which?
    - Combination of all
  - Why?
    - Difference between models
  - How?
    - Table of parameter estimates from all models were collated.
    - National data shared within the consortium
    - Models that used the same parameter agreed on the most appropriate estimate
    - Not always possible due to different algorithms

Online database key criteria (Horigan et. al.)

- Accessibility
  - Does Data exist?
  - Is access limited? e.g. cost/passwords
- Availability
  - Can data be easily extracted
- Completeness
  - Are all data needed included
- Consistency
  - Are data similar to those in other datasets
- Quality
  - Are data fit for (our) purpose

ELSEVIER

Microbial Risk Analysis
Available online 1 March 2019
In Press, Corrected Proof

Full length article

Maximising data to optimise animal disease early warning systems and risk assessment tools within Europe

Verity Horigan [a, 1], Marco de Nardi [c, 1], Maria I. Crescio [b, 1], Agustin Estrada-Peña [d, 1], Amie Adkin [a], Cristiana Maurella [b], Silvia Bertolini [b], Anais Léger [c], Giuseppe Ru [b], Charlotte Cook [a], Katharina Stark [c], Robin R.L. Simons [a]

Show more

https://doi.org/10.1016/j.mran.2019.02.003

Summary of datasets (Horigan et al. 2019)

Data	Source	Accessibility	Availability	Completeness	Consistency	Quality
Pathogen Prevalence	OIE	High	Medium	Medium	Medium	Medium
	EMPRES-i	High	Medium	Medium	Medium	Medium
	Promed	High	Medium	Medium	Medium	Medium
	Healthmap	High	Medium	Medium	Medium	Medium
Movement	Comext (Eurostat): Trade	High	High	High	Medium	High
	Comtrade: Trade	High	High	High	Medium	High
	Faostat: Trade	High	High	High	Medium	High
	Eurostat: Transport	High	High	Medium	N/A	Medium
	International air transport association: Transport	Low	N/A	High	N/A	High
	Traces: Animal movement	Medium	Medium	High	High	High
Susceptible Animals	OIE: Livestock	High	Medium	Medium	Medium	Medium
	FAO Gridded livestock of the world: Livestock	High	Medium	Medium	N/A	Low
	Proxy based on habitat suitability	Low	Medium	Low	N/A	Low
	Global Biodiversity Information Facility	Medium	Medium	Low	Low	Low
	Walter Reed Biosystematics Unit: Vectors	Medium	Medium	Low	Low	Low
	ECDC: Vectors	Medium	Medium	Medium	Low	Low

Online database user wish list

- Standardisation of nomenclature
  - Use of ISO3 codes for countries
  - Latin names for animal species

- User-Friendly interface
  - Use of English language – increase accessibility for international users
  - Access to underlying databases
  - Advanced search functionality
  - Ability to bulk download
  - Official API to facilitate automatic downloading

Online database user wish list

- Metadata should always be available
  - Contact information for database managers
  - Dates of historical and future updates
  - Georeferenced data where relevant (e.g. presence maps)

- Simplified registration/permission process

- Development of an international platform
  - Sustained, reliable and rapid exchange of data and associated metadata between EU MSs

Summary

- Aim: perform a cross-validation of the G-RAID tools
- If differences in risk occur are they due to data input or different methodologies?
- Tried to harmonise as much as possible
  - Easiest for movement and prevalence data
  - Difference in algorithms led to need to use different data
- Data which are measuring the same thing (e.g. trade) may still produce different results

Acknowledgements

G-RAID consortium:

- Clazien De Vos (WBVR)
- Aline De Koeijer (WBVR)
- Manon Swanenburg (WBVR)
- Ronald Petie (WBVR)
- Cecilia Hultén (SVA)
- Kaisa Sörén (SVA)
- Arianna Comin (SVA)
- Anette Boklund (DTU)
- Sebastian Napp (IRTA)
- Tapani Lyytikäinen (Ruokavirasto)
- Leena Seppä-Lassila (Ruokavirasto)
- Helen Roberts (Defra)
- Emma Snary (APHA)

Others:

- Verity Horigan (APHA)
- Amie Adkin (APHA)