

# Deliverable

---

## D1.3 Generic Risk Assessment Framework

**Version: 1**

**Due: Month 46**

**Completed: Month 46**

**Authors: Rachel Taylor, Robin Simons, Roberto Condoleo, Paul Gale, Louise Kelly and Emma Snary  
Animal and Plant Health Agency (APHA) – Weybridge, UK**



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.



## Contents

Deliverable Description .....	3
Introduction .....	4
Generic Risk Assessment Framework .....	4
Risk of Infection .....	7
Risk of Spread .....	9
Overall Risk .....	11
Model Code .....	11
Pathways of Disease Incursion .....	12
Legal Live Animal Trade .....	12
Implementation of the Framework .....	12
Data .....	13
Case Study .....	13
Human Transportation .....	15
Implementation of the Framework .....	16
Data .....	17
Case Study (Draft Results) .....	18
Trade in Food Products .....	20
Implementation of the Framework .....	20
Data .....	22
Case Study (Draft Results) .....	23
Terrestrial Movement of Wild Animals .....	25
Implementation of the Framework .....	26
Data .....	27
Case Study (Draft Results) .....	28
Bird Migration .....	30
Implementation of the Framework .....	30
Data .....	31
Vector flight .....	32



COllaborative Management Platform for detection and Analyses  
of (Re-) emerging and foodborne outbreaks in Europe

Implementation of the Framework .....	33
Data.....	33
Summary.....	34
References.....	35



## Deliverable Description

Exotic diseases pose significant risks to human health, animal health and the livestock industry, and thus prevention methods such as horizon scanning, risk assessment and surveillance are crucial to stop such diseases from entering new areas and spreading. Risk assessments are useful tools to predict which diseases have the highest probability of entering new areas and therefore which pathogens and routes should be prioritized for surveillance. We present a generic spatial risk assessment framework, suitable for any pathogen, which aims to determine hotspots of potential infection of new diseases and the risk of spread should the pathogen enter via that hotspot. The fine spatial resolution of the model can aid risk-based surveillance schemes by reducing the surveillance area necessary to pinpoint disease incursions, potentially catching a new incursion before it has spread too far, thus saving time, money and resources. This document outlines this generic framework and the main pathways of disease incursion that have been included in the model. In addition, in order to demonstrate the type of outputs the risk assessment can produce, some results for the case studies we have used to test the framework are presented.

## Introduction

Risk assessment and surveillance are crucial to stopping the incursion and spread of diseases, thus they can provide a significant benefit for human health, animal health and the livestock industry. By assessing risk on a fine spatial scale, surveillance can be directed to areas identified as highest risk, which enables new incursions of the disease to be found faster. Similarly, management and control procedures, which depend on the risk in different locations, can be prepared in advance to prevent infections occurring and to reduce disease transmission should infection occur. Thus, risk assessment can prioritise surveillance and control management plans, which can save time, money and resources. We now have the ability to target prevention and control of diseases in this manner due to the increased availability of global datasets, more detailed data and improved computational power.

Whilst there are clear guidelines and recommendations on how to perform risk assessments for disease incursion and spread of infectious diseases generally, including how to conduct quantitative entry, exposure and consequence assessments (Murray 2004), there is no generic framework for performing quantitative risk assessments in a spatial setting. A framework allows for standardisation across different countries and organisations to facilitate policy and decision-making. We outline a proposed generic framework for completing spatial quantitative risk assessments for risk of infection and risk of spread. The defining feature of the framework is its emphasis on exotic disease introductions from one area to another. However, the aim is for the framework to be generic across any type of pathogen, method of transmission, species of host(s) and spatial resolutions.

We have created model code alongside the framework that, using appropriate data as input, will produce spatial results for each pathway of disease entry. For these incursion pathways, we outline what is included in the code and what is expected as an input data set. For some of the pathways of disease entry, we have provided example results from the case studies we have used to test our model. However, the case studies are not the focus of this deliverable report, and therefore we have not provided all results for each case study, but instead we present sample results to indicate the capabilities of the framework.

## Generic Risk Assessment Framework

The risk question to outline what our framework aims to achieve is: *“What is the risk of infection and spread of a pathogen in Area B given the presence of that pathogen in Area A?”* It is focused on generic areas, A and B, which will be defined when the framework is utilised for a specific disease situation. The spatial resolution within each area is also variable, depending on the aim of the risk assessor and the data available. The risk of infection outlines the probability that an infection would occur in Area B while the risk of spread outlines the probability that a location in Area B, if it were to be infected, would lead to infection in new locations in Area B. Therefore, the overall risk of infection and spread highlights which locations have the highest risk of both becoming infected and of spreading the disease further.

The risk pathway outlining the steps involved in the risk of infection and the risk of spread for the generic risk question is set out in Figure 1. Our framework fits within the OIE risk assessment framework involving entry, exposure and consequence (Murray 2004), but includes specific details for a quantitative spatial assessment.

Infection can only occur if there is incursion of infected species, non-detection of that species, the survival of that species, and subsequent exposure of native susceptible hosts resulting in transmission. We use the term “species” but it could be even more generic than this, such as infected products or feed, provided that they could come in contact with native susceptible hosts. The term “contemporaneous survival” in Figure 1 indicates that some species are only active part of the year (e.g. only summer months in Northern EU countries in the case of vector-borne diseases) and the two species need to coincide both in time and space for infection in the susceptible species to be possible. It also includes the period over which the infected species is infectious. We combine these five steps in the pathway to produce the probability of one or more infections occurring in Area B for each pathway. Different routes of transmission, different infected species entering Area B or even different pathogen strains require different risk pathways that are then combined together to create the total risk of infection for each location within Area B. We outline later how six of these different incursion pathways fit within the generic framework. Spread from a location in Area B to another location in Area B can only occur if there is an infected host, infectious material or infected secondary hosts at the original location; movement of that host, material or secondary host to a new location in Area B; contact with a new host; and transmission to this new host. Thus, the method incorporates spread from the infected location by any means. It does not have to be spread via direct movement of the original infected host, it could be via fomites, contaminated trucks, vectors etc. Different methods for spread can be calculated separately and then combined in order to compute the total risk of spread for that location. From Figure 1 and its description, it is clear that the steps in the risk of spread calculation are very similar to those in the risk of infection calculation. In fact, estimating the spread of the disease can essentially be thought of as estimating the result of a new incursion of a disease, but with Areas A and B redefined to incorporate the new knowledge of where disease is present. However, in our framework, we simplify the calculations for risk of spread compared to risk of infection, as we are not interested in where the disease spreads to next, just the propensity for that location to spread the disease to a new location. We calculate the risk of spread independently of the risk of infection by estimating spread assuming one infected host at the starting location. Therefore, we are able to compare the risk of spread across locations, regardless of their differences in risk of infection. We then combine the risk of infection and the risk of spread in order to estimate an overall risk of infection and spread. Locations at highest overall risk are of most interest to policy makers.

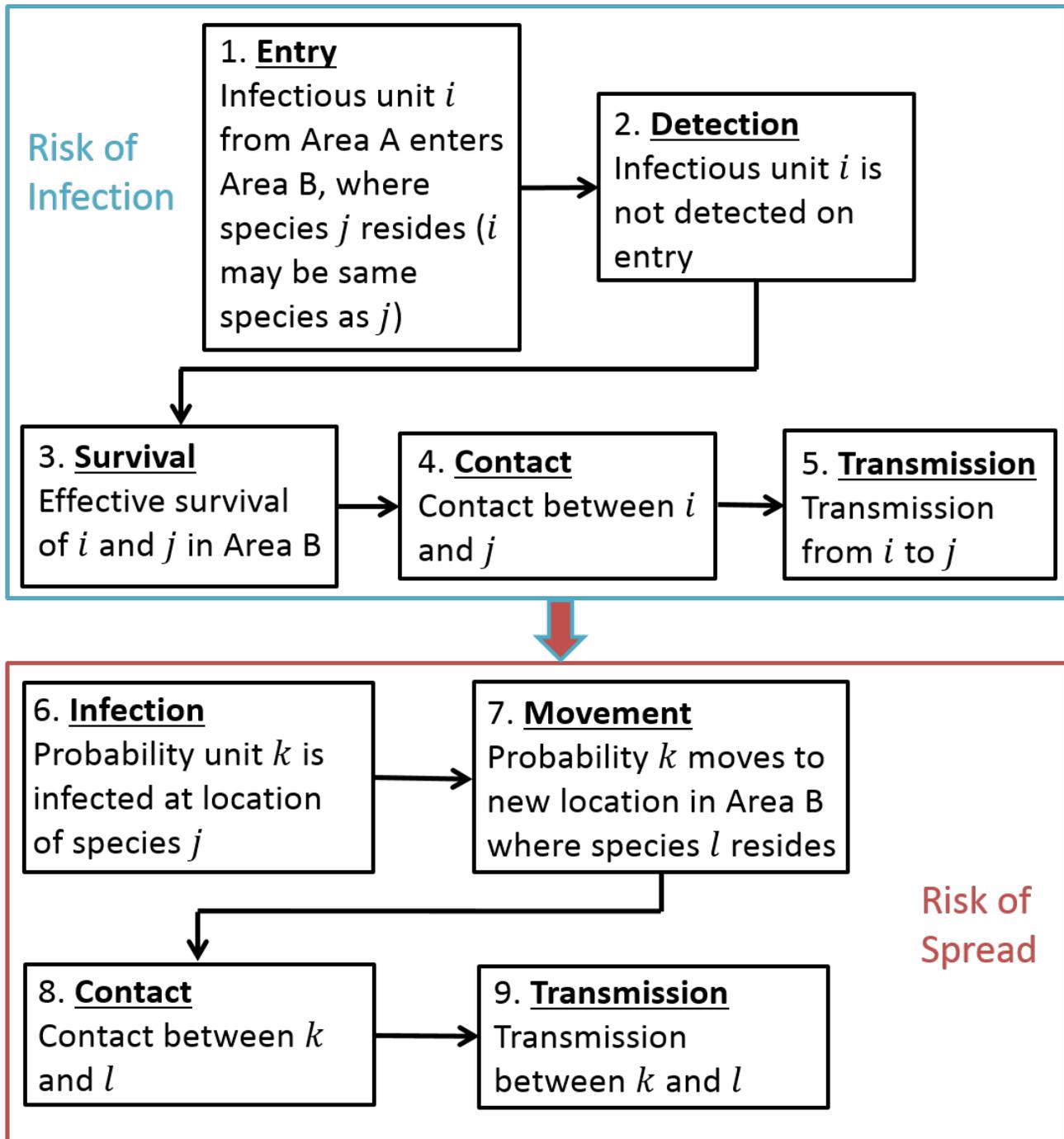


FIGURE 1 THE NINE STEPS OF THE RISK PATHWAY FOR THE GENERIC SPATIAL RISK QUESTION “WHAT IS THE RISK OF INFECTION AND SPREAD OF A PATHOGEN IN AREA B DUE TO THE PRESENCE OF THAT PATHOGEN IN AREA A?” THE TERM “UNIT” IN STEPS 1 AND 6 REFERS TO A GENERIC SOURCE OF INFECTION SUCH AS SPECIES, PRODUCTS OR FEED, THAT COULD HAVE DIRECT OR INDIRECT CONTACT WITH NATIVE SUSCEPTIBLE HOSTS,  $J$  AND  $L$ .

## Risk of Infection

We define the risk of infection as the probability of one or more initial infections in the native susceptible population in Area B. We can mathematically describe and combine the first five steps of the risk pathway (Figure 1) to compute a quantitative probability of initial infection at each location, which can be presented as a spatial risk map. Our mathematical description is adapted from a model that assesses the risk of species jumps in avian influenza (Hill et al. 2015). First, we describe a single disease pathway from species  $i$  to species  $j$ . How to combine other pathways, in which species  $i$  and  $j$  may be different, is outlined afterwards. Furthermore, in later sections of this report (see Pathways of Disease Incursion), we outline how each step of the risk of infection is calculated for different pathways. As the risk assessment may be calculated on different spatial scales, we use  $q$  to denote subregions of Area A and  $g$  for locations in Area B. Both of these will be determined by the resolution of the spatial data available.

### Step 1: Entry

Step 1 of the calculation of the risk of initial infection is the estimation of the number of infected hosts entering Area B. Based on the prevalence in Area  $A_q$  and the total number of hosts exported from Area  $A_q$ , the number of infected hosts ( $I_q(g)$ ) entering location  $g$  of Area B from Area  $A_q$  during a set time interval follows a binomial distribution

$$I_q(g) \sim \text{Bin}(N_q(g), p_q).$$

Here  $N_q(g)$  is the number of hosts imported to or entering location  $g$  from Area  $A_q$  in a unit time interval and  $p_q$  is the prevalence of infected hosts in Area  $A_q$ . The number of infected hosts entering location  $g$  in Area B from Area A is derived by summing over all sub-regions in Area A, thus

$$I(g) = \sum_q I_q(g).$$

We use the stochastic representation above for the number of imported hosts that will be infected to better describe the potential variability. This requires an assumption of independence and therefore an assumption that infected and non-infected hosts are equally likely to be exported.

The prevalence is often calculated using datasets on the number of cases of the disease that have been reported over a relevant timespan but may be calculated differently dependent on data availability for the disease or the pathway of incursion. One simple way of calculating the prevalence,  $p_q$ , the estimated prevalence in Area  $A_q$ , is by the formula

$$p_q = U_q \frac{c_q}{N_p(i, q)} \frac{\tau}{365},$$

where  $c_q$  is the number of cases,  $N_p(i, q)$  is the total population of species  $i$  in Area  $A_q$  and  $U_q$  is an under-reporting factor to account for the fact that many cases go unreported due to lack of surveillance, no clinical symptoms or other factors. Lastly,  $\tau$  is the length (in days) of the infectious period or incubation period, depending on the disease or route. More complicated methods for calculating prevalence are also possible. Within our code, prevalence is a data input, therefore, we do not include in the code the above steps for calculating prevalence from the number of reported cases. A user of the code could calculate the prevalence using a more sophisticated

method than the formula above, if desired, and input that directly. However, for some pathways we have included additional considerations when they are specific to that pathway, outlined below.

## Step 2: Control

Some infected hosts may not make it through import control due to detection of symptoms or testing of hosts, Step 2 in Figure 1. The probability of detection and the sensitivity of the tests can also vary by location  $g$ . We assume, however, that the probability of detection is independent of which country is the exporter, although in practice this may vary. Therefore, we denote  $p_D(g)$  as the probability of successfully detecting and removing an infected host. The actual number of infected hosts  $J(g)$  entering location  $g$  in Area B follows a binomial distribution

$$J(g) \sim \text{Bin}(I(g), 1 - p_D(g)).$$

## Steps 3 – 5: Survival, Contact and Transmission

We next calculate Steps 3 – 5 (Figure 1), namely the survival rate of the species, the contact rate between hosts and the probability of transmission, given contact, leading to initial infection of susceptible hosts in each location  $g$  of Area B. We combine these components using the basic reproductive number  $R_0$ . The basic reproductive number gives the number of susceptible hosts likely to be infected by the introduction of one infected host at each location  $g$ .  $R_0$  is a fundamental metric of disease systems, but the equation to represent  $R_0$  depends on how the transmission of the disease is modelled. Therefore, using  $R_0$  facilitates adaptation to different methods of transmission due to different interactions between species  $i$  and  $j$ , e.g. vector-borne transmission, direct transmission or sexual transmission, as well as specific aspects that are only applicable in some cases, such as environmental factors, in determining survival of species. For example, direct transmission would be represented by the equation

$$R_0(g) = \frac{c\beta S(g)}{r}.$$

In this equation,  $c$  is the contact rate between hosts (Step 4),  $\beta$  is the probability that contact results in successful transmission (Step 5),  $S(g)$  is the population size of susceptible hosts of species  $j$  in location  $g$  and  $r$  is the recovery rate, based on the length of time infected hosts of species  $i$  remain present and infectious in location  $g$  (Step 3). Additionally,  $R_0$  can change depending on the location  $g$  to incorporate differences in transmission in different regions. And if the data are available,  $R_0$  could be a function of temperature, changing for each location based on average temperatures. For calculating the risk of initial infection occurring within location  $g$  we assume that there is homogeneous mixing between the newly introduced species and the native susceptible population, but this could be adapted to other scenarios in which homogeneous mixing is not a good assumption by changing the contact rate  $c$ .

## Risk of Infection Calculation

Based on our definition of risk of infection, we calculate the probability of one or more infections occurring in the susceptible population within location  $g$ , per unit time by combining the information from Steps 1 – 5. The probability of random events, such as infections, happening can be described by a Poisson process with parameter  $\lambda$ , where  $\lambda$  is the expected number of events occurring per unit time. For each observation of  $J(g)$  we can estimate the expected number of new infections occurring in a unit time in location  $g$  by  $R_0(g)J(g)$ . Hence, in our

framework the number of new infections per unit time is assumed to follow a Poisson process with parameter  $\lambda = R_0(g)J(g)$ . The probability of no events happening in a Poisson process is  $e^{-\lambda}$ . Hence, the risk of infection, alternatively the probability of one or more infections occurring in the susceptible population, in location  $g$  from introduction of infected hosts from Area A is given by:

$$R_I(g) = 1 - e^{-R_0(g)J(g)}.$$

If there are multiple routes of transmission (which could involve different species  $i$  and  $j$ ) then all parameters, including the contact rates, number of susceptible animals and number of imported infected hosts, may be different for each route. To incorporate these different routes, we denote the route with an additional subscript  $\omega$  and, hence, re-write the above equation as

$$R_{I,\omega}(g) = 1 - e^{-R_{0,\omega}(g)J_\omega(g)}$$

for each route  $\omega$ . Therefore, using the multiplication rule for independent probabilities (which is a reasonable assumption for the different pathways), we can compute the complete risk of infection over all routes of transmission as

$$R_I(g) = 1 - \prod_{\omega} [1 - R_{I,\omega}(g)].$$

This risk of infection calculation is capable of representing different methods of transmission, a wide range of pathogens with different environmental requirements, any spatial scale and any route of introduction.

## Risk of Spread

The risk of spread is focused on assessing which locations are most likely to be the source of further spread, not which locations the disease would spread to. This is for surveillance purposes, prior to disease entry, in order to prioritize which locations should have highest surveillance. We define the risk of spread at each location as the probability of one or more infections in the native susceptible population in a different location in Area B. We can mathematically describe and combine the last four steps of the risk pathway (Figure 1) to compute a quantitative probability of spread at each location, which can be presented as a spatial risk map.

The risk of spread calculations work on the assumption that there exists one infected host of species  $j$  at location  $g$  in Area B and that this is the only presence of the disease in Area B. This is then adjusted by the expected number of infected host species at location  $g$  to calculate the overall risk by using output from the risk of infection steps. We describe the framework for the risk of spread from the host species  $j$  to another host, species  $l$  at location  $\tilde{g}$ , in which  $j$  and  $l$  may be the same species. There are multiple methods of spread from one location to another, such as movement of domestic animals, contaminated clothes, trucks and cars, vectors and movement of wild animals. We describe first a single method of spread from location  $g$  to location  $\tilde{g}$ . How to combine other methods of spread from location  $g$  to location  $\tilde{g}$  is outlined afterwards.

We will not individually outline how each method of spread could be calculated using this framework. However, we do provide examples throughout the description of the framework. The framework outlines the basic steps in calculating the risk of spread, but it is up to the user to choose the level of detail to include for each method of

spread, which will be primarily determined by the specific risk question, data availability and time to perform the risk assessment.

### Step 6: Infection

The first step of the risk of spread calculation is to estimate the probability  $p_I(k, g)$  of infection at location  $g$  in the species/product of interest,  $k$ . The species/product of interest is the unit by which the spread will occur, such as fomites, soil, a wild animal species. If the species of interest is the same species as the original infection, i.e.  $j$  and  $k$  are the same species, then the probability of infection in species  $k$  will be the prevalence of the disease (based on the assumption of a single case) and will be calculated using the number of that species in location  $g$

$$p_I(k, g) = \frac{1}{N_p(k, g)},$$

where  $N_p(k, g)$  is the population size of species  $k$  in location  $g$ . However, if the species/product of interest  $k$  is different to the original infected species  $j$  then additional steps are required to compute  $p_I(k, g)$ . For example, if the method of spread from one location to another is contaminated trucks, then  $p_I(k, g)$  is the probability that if a truck was at location  $g$ , it would become contaminated. This could be calculated using data on the amount of shedding from the host species  $j$ , the probability that biosecurity measures would take place and the probability that the biosecurity procedures would successfully stop contamination of the truck.

### Step 7: Movement

This step involves calculating the amount of movement from location  $g$ , rather than where the movement is going to. We define it as the proportion of movement  $m(g)$  from location  $g$  relative to all other locations in Area B. Therefore, for the contaminated truck example from before, this would be the proportion of trucks that leave location  $g$  compared to all the trucks leaving every location in Area B. For movement of live animals, we assume that all wild animals perform some sort of movement that is independent of location. But the probability  $m(g)$  would be the proportion of wild animals that would perform movement to other locations (e.g. long-distance movement instead of home-range movement).

### Steps 8 – 9: Contact and Transmission

We next calculate Steps 8 – 9 (Figure 1), namely the contact rate between hosts and the probability of transmission leading to infection of susceptible hosts at location  $\tilde{g}$  of Area B. We combine these components using the basic reproductive number  $R_0$ , similar to the method in the risk of infection calculation. This equation for  $R_0$  can incorporate different transmission routes leading to infection, including disease-specific information and the effect of temperature on parameter values. It also includes the average number of susceptible hosts that we expect to be present in the new location  $\tilde{g}$ . The number of susceptible hosts could be determined by the (weighted) average across the locations which are reachable from location  $g$  by a specific method of spread. Therefore, we define our basic reproductive number for risk of spread by the notation  $R_0(\tilde{g})$ . Note that if  $k$  and  $l$  are the same species as  $i$  and  $j$ , it may be the case that the parameters for  $R_0$  remain unchanged.

### Risk of Spread Calculation

Based on our definition of the risk of spread, we calculate the probability of one or more infections occurring in the susceptible population in a new location  $\tilde{g}$  in Area B by combining the information from Steps 6 – 9. Similar to

the risk of infection calculation, we utilise the fact that random events occur as a Poisson process. Hence, in our framework the average rate at which new infections occur in other locations per unit time is given by  $R_0(\bar{g})m(g)p_l(k, g)$ . Therefore, the risk of spread by a single method of spread is

$$R_S(g) = 1 - e^{-R_0(\bar{g})m(g)p_l(k, g)}.$$

Different methods of spread, such as movement of live animals, contaminated trucks, vectors, involve different species/products  $k$  and  $l$  and therefore all the parameters may change. To incorporate these different methods of spread, we denote the method with an additional subscript  $v$  and, hence, re-write the above equation as

$$R_{S,v}(g) = 1 - e^{-R_{0,v}(\bar{g})m_v(g)p_{l,v}(k, g)}.$$

for each method of spread  $v$ . Therefore, using the multiplication rule for independent probabilities, we can compute the complete risk of spread over all routes of transmission as

$$R_S(g) = 1 - \prod_v [1 - R_{S,v}(g)].$$

## Overall Risk

The risk of infection and risk of spread calculated using the above framework can be combined together to produce an overall risk of infection and spread. However, it is not simply the case of multiplying the risk of infection with the risk of spread. This is because the risk of spread calculation is based upon the assumption that there is one infected host at the location. During our risk of infection calculations we estimated how many cases of initial infections were expected at location  $g$  by the formula  $R_0(g)J(g)$ . Therefore, combining this with the risk of spread scales the risk of spread appropriately to the actual expected number of cases. To do this, we multiply the expected number of infected cases by the risk of spread to get the overall risk of infection and spread,  $\Theta(g)$ ,

$$\Theta(g) = R_0(g)J(g)R_S(g).$$

We acknowledge that this method could ignore non-linear effects of multiple infected animals in the same location but as these are non-linear effects are difficult to quantify, we use this simpler assumption.

## Model Code

The framework has been programmed in R (R Core Team 2016) which is free software, thus allowing easy accessibility to all. However, the code is complex and is not designed with a front-end for users. Therefore, it is intended that the code would be used by someone with a sufficient knowledge of R to be able to understand the code, change data input for their disease and locations, and potentially adapt some of the code towards their specific situation. The first version of the code is available upon request from the authors. Potential plans for the future of the code are to include additional pathways and to create a front-end for easier use. The run-time of the code depends on the desired spatial resolution of the results and the size of Area B but if all the data are included in the correct format, the run-time can vary from <5 minutes to a day or two.

## Pathways of Disease Incursion

The generic framework we have considered above for the risk of infection is applicable for all pathways of disease incursion that we might be interested in for a specific disease risk assessment. However, while all five steps of the risk of infection framework will be included for each pathway, the complexity of calculating each step may change depending on the disease and the pathway of incursion. Most importantly, for different pathways of incursion, the calculation of  $N_q(g)$ , the number of hosts entering location  $g$  from Area  $A_q$ , may change. This is because, for many pathways, there is no direct data available about entry of species and instead we create a sub-model to estimate this value based on other data. Similarly, depending on the disease, the calculation of  $R_0$  will change.

We consider six major pathways in this prototype version of the model code. These are a legal live animal trade route, human transportation, legal trade in food products, terrestrial movement of wild animals, bird migration and vector flight. It is not possible to include all potential pathways in the model and so we prioritized which pathways to develop within the prototype based on a review of the key routes and which were important for our chosen case studies (Horigan et al. 2018). Some pathways which we have not included but which may be important for other diseases are the movement of contaminated trucks, trade in genetic material, and illegal trade of animals and food products. These pathways do fit within our generic framework, but we have not formulated how to perform each step of the framework nor built code to compute this quantitatively.

For each of the six major pathways, we outline how to perform each step of the risk of infection framework, any specific aspects that are important for that pathway, and the type of data sources that are useful. For some of the pathways, we provide results from case studies that serve to elucidate how the method works, and the spatial resolution of results that are possible for each case study using our code for the framework.

### Legal Live Animal Trade

The trade in live animals pathway focusses on the largest component of live animal trade, namely livestock trade. It incorporates all legal trade in any livestock animal, traded for any reason, and considers whether those animals could be infected, mitigation measures on arrival at the new location, and how much contact with susceptible animals those infected animals could have in their new location. Within our code, we have not considered trade in horses, Balai animals (e.g. animals intended for zoos) or pets. This trade could be incorporated following the exact framework we outline, but the contact with other susceptible animals of the same or different species will need to be carefully considered and may involve changing the code. We have not considered transmission of the disease within a group of animals during transport to a new location and for some diseases, this may increase the probability of detection as more animals show signs of infection.

### Implementation of the Framework

This pathway is the simplest to implement as it follows the framework exactly. This is because there are global datasets available on the number of animals traded between countries and so we can input these directly into our model for the entry step. The one consideration required for the entry step is whether the live animal is being traded for slaughter purposes (i.e. slaughter in the country of import) and if this is indicated in the trade dataset. Trade for slaughter should be removed from the overall data set, if possible, as these animals will not arrive on to a farm and therefore it is assumed will not lead to infection in susceptible animals.

For all other steps of the framework, there are no specific adjustments required for the legal live animal trade pathway.

For this pathway, the code is set up to compute the risk of infection at a country level, a regional level, and at a farm level, as these were the different resolutions of trade data that we were able to source. However, the code has been implemented in such a way that provided the user has a Spatial Polygons Data Frame for the regions they wish to compute the risk for, computation on any regional scale (whether single counties, groups of counties, provinces etc.) should be feasible. Computation of the risk using raster cells of any size is possible within the framework but would require adapting the code.

## Data

We outline the possible data sources that can be used for some of the steps of the framework for this pathway in Table 1.

**TABLE 1 DATA REQUIRED TO COMPUTE THE RISK OF INFECTION WITHIN THE GENERIC FRAMEWORK FOR THE LIVE ANIMAL TRADE INCURSION PATHWAY.**

Parameter	Specific Data	Further Details	Potential Data Sources
Movement from Area A to Area B ( $N_q(g)$ )	Trade/Registered Movement	Trade in livestock or registered movement of hosts from regions in Area A to locations in Area B.	UN Comtrade data Eurostat COMEXT data Trade Control and Expert System (TRACES)
Prevalence ( $p_q$ )	Prevalence of the disease in Area A	Calculated using data on the number of cases that have been reported and the number of hosts in Area A. Preferably for the same regions in Area A as the movement data.	OIE Animal Disease Notification System FAO EMPRES-i
Susceptible Hosts ( $S(g)$ )	Size of farms in Area B	Depending on spatial scale, this could be the number of hosts on a specific farm, or the average number of hosts on farms in a region.	OIE Eurostat

## Case Study

For our case study we considered a disease of cattle in the Capripoxvirus genus: lumpy skin disease (LSD). This is an OIE-notifiable disease that causes nodules on the skin, mucus membranes and internal organs; reduction in milk production; fever; oedema; and sometimes death (Davies 1991, Tuppurainen and Oura 2012). Mortality is usually low (<10% Kumar (2011)), but it can cause significant economic losses. The disease had appeared to be in decline and restricted to sub-Saharan Africa, but a resurgence occurred in the 1980s and subsequently it has been steadily spreading northwards (Hunter and Wallace 2001). Although there had been infrequent incursions before, since 2006 LSD has become endemic throughout the Middle East (Tuppurainen and Oura 2012). Similarly, it has

been present in Turkey since 2013 and is now considered endemic. A few cases of LSD occurred in Greece for the first time during 2015, followed by a widespread outbreak in the Balkan regions in 2016, specifically in Greece, Bulgaria, The Former Yugoslav Republic of Macedonia, Serbia, Kosovo and Albania, as well as in Russia (Mercier et al. 2017). Therefore, with the recent resurgence of this disease within Europe, it was decided that LSD would be a good case study to explore the risk of this disease spreading to other countries in Europe due to legal trade.

We assessed the risk of infection within Europe due to the legal trade of live animals in 2016. Our Area B was defined as all countries in Europe excluding those countries that had notified cases in 2016, namely Greece, Bulgaria, The Former Yugoslav Republic of Macedonia, Serbia, Kosovo, Albania, Turkey and Russia. Our Area A was the whole world. We assessed the risk at different spatial scales to highlight the ability of the framework to cope with these different resolutions of data. We computed the risk of infection with LSD at a country level, a regional level and at a farm level. We used TRACES data for trade at a farm level (TRACES 2017), and Comext data for trade at a country level (Eurostat 2017). Our prevalence data for all countries in the world was taken from the EU-funded SPARE project (Simons et al. 2017), which estimates prevalence of disease around the world using OIE data on the number of outbreaks and the number of cases per outbreak of the disease in the past 10 years. For the number of susceptible animals that an imported animal would be in contact with, we used either country or regional averages of number of cattle on a farm, provided by Eurostat (Eurostat 2017). Disease related parameters were found from published literature.

In Figure 2, we present some of the results we obtained for this case study risk assessment, at a farm level. For further results, please see our published article on this case study (Taylor et al. 2018). In Figure 2, the mean annual risk of infection with LSD across Europe at an individual farm level is plotted. This indicates that the highest risk is focused in Croatia because a large number of farms in that country trade with countries that had non-zero prevalence of LSD, according to our prevalence data. Additionally, the two farms with the highest mean risk are located in Croatia. However, the next three highest-risk farms occur in Spain, even though only a small number of farms in Spain have non-zero risk. This provides much more detailed information than the regional- or country-level risk assessment, thereby allowing better surveillance plans to be in place for Spain. However, we found that the results were consistent across the spatial scales, and when time is short or data are not available, the country and regional risk assessments provide a useful and relevant measure of risk.

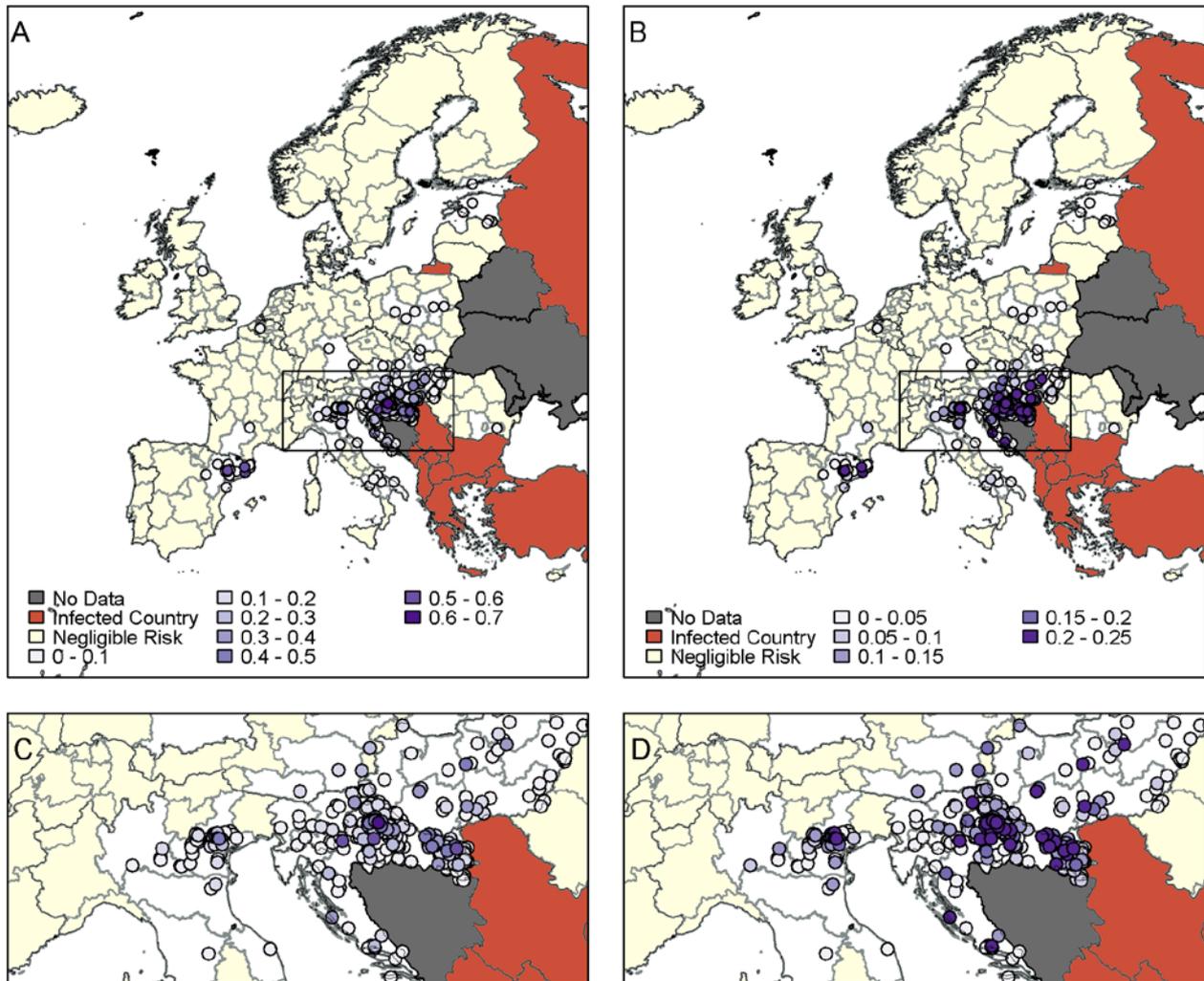


FIGURE 2 THE MEAN ANNUAL RISK OF INFECTION OF LSD (A) AND THE VARIANCE OF THIS RISK (B) IN 2016 DUE TO LEGAL TRADE ARE PLOTTED IN SHADES OF PURPLE, CALCULATED AT THE INDIVIDUAL FARM LEVEL. IN (C) AND (D), THE MEAN AND VARIANCE, RESPECTIVELY, OF THE ANNUAL RISK OF INFECTION ARE AGAIN PLOTTED, BUT ZOOMED IN TO THE AREAS OUTLINED IN RECTANGLES IN A AND B. REGIONS IN YELLOW HAVE NEGLIGIBLE RISK DUE TO FARMS WITHIN THOSE REGIONS ONLY TRADING WITH COUNTRIES THAT HAVE ZERO PREVALENCE, ACCORDING TO OUR PREVALENCE DATA. COUNTRIES THAT HAD NOTIFIED CASES IN 2016 ARE IN RED. COUNTRIES IN GREY HAVE INSUFFICIENT DATA FOR CALCULATING RISK.

## Human Transportation

The human transportation pathway includes all travel that is registered movement of humans from one location to another so it includes travel on planes, boats and rail. There exist statistics on the number of transport events move from one location to another, and how many passengers travelled for these methods of transport. This pathway is also adaptable to the incursion via transport in cars. However, there is a significant lack of data in this area and it was not required for our case study so we focused instead on transport by planes and boats. This

pathway includes travel by humans as well as (unintentional) hitchhikers on the transport such as vectors (e.g. mosquitoes).

### Implementation of the Framework

The implementation of the generic framework (Figure 1) for the entry step from Area A to Area B due to human transportation is relatively simple at a country level for planes and boats, as there exist global datasets on human travel to each country. For vectors travelling from Area A to Area B, instead of using data on the number of passengers, we instead use data on the number of transports. Using studies on the number of vectors found on international flights and maritime containers (Haseyama et al. 2007) and whether fumigation takes places in Area A on the plane/boat, we can then estimate how many vectors would travel from Area A to Area B by this pathway (Brown et al. 2012).

As stated above, for the estimation of prevalence, we calculate it based upon reported cases of the disease. We include one additional factor when modeling prevalence in humans in order to create a more realistic estimate of prevalence for travelers compared to the general population. We incorporate an average length of time that the human would have been in the country for, dependent on what proportion of passengers are tourists or travelling for business in Area A versus those that live in Area A (Simons et al. 2016). This reduces the prevalence in passengers to account for the fact that many on the transport would have had a short exposure time within Area A to have contracted the disease. Therefore, our prevalence  $p_q$  in Area  $A_q$  is calculated as

$$p_q = U_q \frac{c_q}{N_p(i, q)} \frac{\tau}{365} \sum_{m \in M} \vartheta_m \frac{T_m}{365},$$

where the first three terms are the same formula for prevalence as suggested in the generic framework above with  $c_q$  as the number of cases in Area  $A_q$ ,  $U_q$  an under-reporting factor for cases,  $N_p(i, q)$  the total population of species  $i$  in Area  $A_q$  and  $\tau$ , the length in days of the infectious period. Finally, in the last term we adjust for different passenger types by including different exposure times that are weighted by the proportion of each passenger type. Specifically,  $M$  is the set of all passenger types, e.g. resident of country in Area A, tourist, business traveler,  $\vartheta_m$  is the proportion of all passengers that are of that passenger type and  $T_m$  is the average duration in days spent in Area  $A_q$  for that passenger type. The prevalence within the vector population that have alighted the planes/boats is estimated to be the same as the prevalence in vectors in the origin region.

To calculate the incursion of infected species at a country level for both humans and vectors, we follow the framework above exactly, by inputting transport data and the prevalence into the Binomial formula to get the number infected  $I(C)$  entering each location  $g$ , which in this case we denote as a country,  $C$ . To calculate the incursion of infected species at a finer spatial scale, we make use of the raster format at a 100km<sup>2</sup> cell level, where we split geographical space into cells that are 10km wide and long at the equator. The calculations depend on whether humans or vectors are the species entering Area B. If humans, we calculate the incursion at a cell level by distributing the humans in Area B based upon human density in each raster cell and assuming they stay in the same country as they entered Area B. This assumes that human density, rather than distance to the airport/port, is the best predictor of where people are travelling to in each country. Per the framework, we need the number of infected entering each location  $g$ , which we calculate as

$$I(g) = \frac{N_p(i, g)}{N_p(i, C)} I(C),$$

where each location  $g$  is a raster cell in country  $C$ ,  $N_p(i, g)$  and  $N_p(i, C)$  give the total population of species  $i$ , in this case humans, in locations  $g$  and country  $C$  respectively, and  $I(C)$  is the number of infected humans entering country  $C$ . The risk of infection given infected human incursion is then given by the abundance of the susceptible species in each raster cell and the contact between this species and the infected humans, as well as the survival of both infected humans and the susceptible species. To do this, we follow the generic framework and use  $R_0(g)$  to represent the survival, contact and transmission for each specific disease.

For infected vector entry at a raster cell level, we assume the vectors are not able to travel far from the airport/port, unlike humans, and therefore assume that the vectors remain in the same raster cell as the airport/port they enter the country. For the calculation of the survival, contact and transmission, we have built in the ability in the code to include the effect of temperature in these parameters. Vectors are ectotherms and thus their vital rates, including survival and bite rate, are dependent on temperature. Therefore, within this pathway it is possible to use fine temporal scale travel data, and input average temperatures on the same temporal scale, in order to calculate these parameter values. Thus, the risk of infection given infected vector incursion will be estimated by combining the abundance of the susceptible species in the raster cell containing the relevant airport/port, and the temperature-dependent  $R_0$  equation that details the contact between the susceptible species and the infected vectors.

## Data

We outline the possible data sources that can be used for some of the steps of the framework for this pathway in Table 2.

TABLE 2 DATA REQUIRED TO COMPUTE THE RISK OF INFECTION WITHIN THE GENERIC FRAMEWORK FOR THE TRAVEL ON PLANES INCURSION PATHWAY.

Parameter	Specific Data	Further Details	Potential Data Sources
Movement from Area A to Area B ( $N_q(g)$ )	Travel on planes or boats	Registered movement of number of hosts or amount of transport from regions in Area A to locations in Area B.	Eurostat International Air Transport Association
	Vectors on the transport	An estimate of how many vectors of each species, on average, travel on each transport	Published literature
	Mitigation measures	Whether fumigation occurs on board transport to kill vectors, and its success rate	Published literature
Prevalence ( $p_q$ )	Prevalence of the disease in Area A	Calculated using data on the number of cases that have been reported. Preferably for the same regions in Area A as the movement data.	OIE Animal Disease Notification System FAO

			EMPRES-i
	Split of passenger types	Estimate of how many passengers are tourists/business travellers or from Area A to determine exposure to the disease in Area A.	Published literature
Susceptible Hosts ( $S(g)$ )	Abundance of susceptible species	Spatial distribution of the abundance of the susceptible species	Gridded Population of the World (SEDAC) Gridded Livestock Population of the World (FAO) Global Biodiversity Information Facility (GBIF)
Contact and Survival	Temperature data	Average temperatures in each location in Area B on a time scale the same as the travel data	UK Met Office European Environment Agency
	Vector life-history parameters	Effect of temperature on the vector life-history traits, such as mortality and bite rate	Published literature

### Case Study (Draft Results)

The case study that we are considering for the human transportation incursion pathway is Zika virus. Zika virus is an arbovirus of the family *Flaviviridae* that infects humans and is spread predominantly by *Aedes aegypti* mosquitoes, although *Aedes albopictus* has also been implicated as a potential vector (Chouin-Carneiro et al. 2016, Plourde and Bloch 2016). The disease originated in Africa and remained within a narrow region of the world, until 2007 when it started spreading eastwards through South-East Asia and across the Pacific until finally reaching the Americas. In 2015, an outbreak began in Brazil and then spread rapidly to other South and Central American countries and the Caribbean, as well as incursions within North America (Plourde and Bloch 2016). The significant number of cases reported in 2015 and 2016, alongside the insufficient knowledge of the long-term health effects, lead to the outbreak being declared a Public Health Emergency of International Concern, although the number of new cases fell dramatically in 2017. Although most infections of Zika virus involve only mild symptoms, a connection between Zika virus and microcephaly and Guillain–Barré syndrome cases led to the disease being a significant concern for pregnant women (Brasil et al. 2016). As well as transmission due to mosquito bites, infection can also occur due to sexual transmission (Musso et al. 2015).

Our human transportation incursion case study considers the risk of infection in Europe during 2016 when the cases of Zika were peaking. This case study highlights the ability of the generic framework to produce results for surveillance during an ongoing outbreak using limited data. For this case study, our Area A was those countries in the Americas that reported cases in 2016. Our Area B is the EU. For this case study, we only consider travel by plane. We include three different transmission processes by which an infection event could occur in Area B in

order to calculate the risk of infection although all involve the human transportation incursion pathway. These are the incursion of humans leading to infection in humans (via sexual transmission), incursion of humans leading to infection in EU indigenous mosquitoes, and incursion of mosquitoes leading to infection in humans in the EU. We use Eurostat data (Eurostat 2017) on the number of flights from each country in the Americas to each airport in the EU, and the number of passengers arriving at each airport from each country in Area A for 2016. For the number of mosquitoes hitching a ride on planes, we apply the same method as Brown et al. (2012) which determined the number of West Nile virus-infected mosquitoes entering the UK aboard aircraft. For the prevalence data we use reported cases of Zika in each country in the Americas in 2016 extracted from the Pan American Health Organization (Pan American Health Organization (PAHO)) and use the data on passenger types and length of stay in each country from Simons et al. (2016). As we are considering multiple incursion routes, the susceptible species may be either humans or indigenous EU mosquitoes. We use a human density map provided by SEDAC to distribute the infected humans that enter each country to each raster cell, as well as indicating how many susceptible humans are in each cell. We use presence/absence maps provided by the European Centre for Disease Control (ECDC) of *A. aegypti* and *A. albopictus* combined with information from published literature on mosquito life history traits to determine the spatial abundance of these mosquitoes. We combine these all together using our framework to calculate the risk of infection in humans and mosquitoes across the EU.

In Figure 3, we present a map of the entry of infected humans into Europe and in to the Netherlands specifically. We use a log scale in Figure 3A because it highlights differences between cells that all have values very close to zero. Within each country, the infected humans are distributed according to human density, with clear darker spots in locations such as Madrid, Paris and London. However, there are also clear differences between countries, with most of the risk of Zika-infected human entry occurring in Western Europe. We also highlight the Netherlands as a case study in Figure 3B. This indicates the risk is focused in Amsterdam and The Hague/Rotterdam areas as these have the highest density, with up to 1.2 estimated infected humans entering a single cell during 2016. In total, we estimated an average of 88 Zika-infected humans entering the Netherlands in 2016.

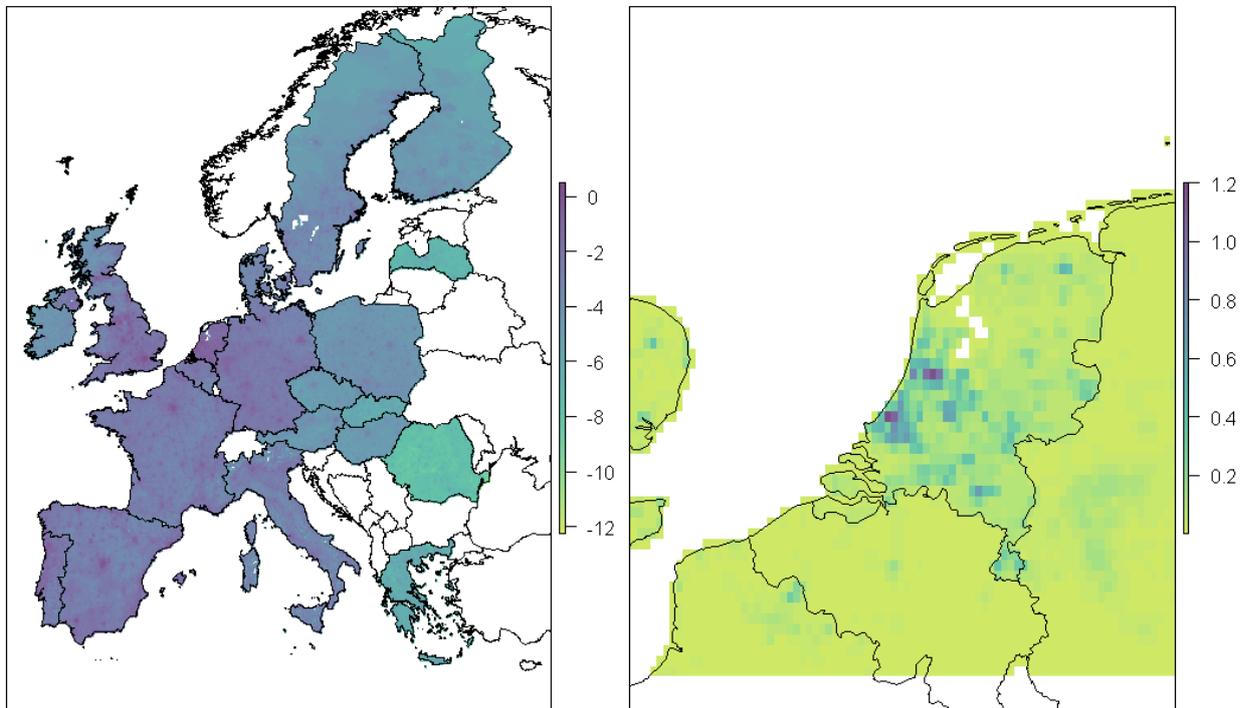


FIGURE 3 THE MEAN NUMBER OF ZIKA-INFECTED HUMANS PREDICTED TO HAVE ENTERED EACH 100KM<sup>2</sup> RASTER CELL ACROSS EUROPE (A) AND THE NETHERLANDS (B) IN 2016. THIS IS CALCULATED BASED ON TRAVEL BY PLANES TO EACH COUNTRY FROM THE AMERICAS ONLY. IN (A) THE NUMBER OF HUMANS ENTERING IS PLOTTED ON A LOG SCALE IN ORDER TO COMPARE THE DIFFERENCES ACROSS VERY LOW VALUES. IN (B) EXACT VALUES ARE USED.

### Trade in Food Products

The trade in food products incursion pathway includes all legal trade in food that could be intended for human or animal consumption, and whether that food will go to waste and then have contact with an unintended species. It covers infected food products imported for human consumption and able to lead to infection in humans; those same products going to waste and then coming into contact with and leading to infection in livestock or wild animals; and imported feed products which are intended for livestock. The cases involving unintended recipients could also include illegal usage of this food such as swill feeding. These can all be incorporated into the same pathway as it only requires switching some parameters to indicate if it is human food or animal feed and whether the model needs to consider the waste product or the non-wasted product.

### Implementation of the Framework

Similar to the trade in live animals, implementation of the framework with regard to the entry step of the products is relatively simple as there are global databases of trade in food products. These are usually broken down into a huge number of different types of products but not to the individual items. Thus, if the pathogen is found in a particular fruit, it would be difficult to isolate the trade data solely for that fruit, but if the pathogen is found in a

certain type of product, for example pig meat, then all products that involve pig meat can be extracted. This is an issue with data availability rather than a drawback of our model.

We thus calculate the total amount of infected product entering each location  $g$ ,  $I_n(g)$ , as per the framework, split for each product type  $n \in N$ , and using the prevalence input by the user. We do not include specific adaptations to the prevalence calculation for this pathway. However, for the trade in food products pathway, we will perform calculations based upon the total viral load of the products, rather than the amount of product that is infected. This is because food products may be composed of different components (e.g. for meat products these could be bone, muscle, offal etc.) that may have different initial concentrations of viral load and decay rates. Therefore, to complete the first step in the framework for this pathway, we need to convert the amount of infected product entering each location  $g$  into the total amount of viral load  $V_n(g)$  entering, by the following formula

$$V_n(g) = I_n(g) \sum_{d \in D} \pi_d \rho_d,$$

where  $D$  is the set of all components that the product  $n$  is composed of,  $\pi_d$  is the initial viral load in that component  $d$  and  $\rho_d$  is the proportion of the product  $n$  that is composed of the component  $d$ .

For the survival step of the framework, we include many different food processes that may occur to each product type (e.g. smoking, salting, and drying) and reduce the viral load based on the effect of these processes on viral load. These processes occur in Area A and are determined by the product code of the data. We therefore update the viral load in each product going to location  $g$  by the following formula

$$\bar{V}_n(g) = V_n(g) (1 - \lambda_p)^{T_p}.$$

Here,  $\lambda_p$  is the decay rate of viral load in product  $n$  when it has undergone process  $p$  and  $T_p$  is the duration of time for process  $p$ . We also include the decay of viral load during transport from Area A to Area B, which is measured based on the average time to travel between different locations by different transport methods (Simons et al. 2016). To do this, we assume that transport of food products can occur through multiple methods (e.g. by air, sea or truck) which would have different speeds. Thus, the updated viral load is

$$\bar{\bar{V}}_n(g) = \bar{V}_n(g) \sum_{w \in W} \rho_w (1 - \lambda_p)^{T_w},$$

where  $W$  is the set of all transport methods,  $T_w$  is the time taken for transport method  $w$ , and  $\rho_w$  is the proportion of the product that will be imported via that transport method. Lastly, we include a further process within Area B, cooking. The product is cooked or not depending on the product type and the resulting viral load for cooked products is calculated by multiplying the above viral load with the probability that cooking will be sufficient to eliminate the virus. At this point, the total viral load entering location  $g$  is calculated by summing over the viral load in each product  $n$

$$V(g) = \sum_{n \in N} \bar{\bar{V}}_n(g).$$

It is possible to complete the risk of infection for this pathway on a raster cell, regional or country level. In order to perform the method on a raster level, we need an additional step that converts country level entry data to the amount going to each raster cell. We assume that the amount of product that is entering each country will be

distributed amongst the raster cells in that country based upon the population density of the intended recipient of the food. This is similar to the human transportation pathway. The viral load in cell  $c$ ,  $V(c)$ , is

$$V(c) = V(C) \frac{N_p(j, c)}{N_p(j, C)}$$

where  $V(C)$  is the viral load in the country the cell is located in, calculated from the above steps with  $g = C$ , and  $N_p(j, c)$  and  $N_p(j, C)$  are the populations of the intended recipients, species  $j$ , of the food in cell  $c$  and country  $C$ . For the contact step of the framework, it is necessary to think carefully about how to include the number of susceptible animals within the framework as there are many potential recipients of the food. When we wish to model contact and transmission with the intended recipients of the food, this is measured by splitting the viral load amongst the population in the raster cell and accounting for the percentage of recipients that eat the product (e.g. accounting for non-meat eaters). If the product goes to waste or is fed to unintended recipients, then the contact rate is adapted to include the population abundance of the unintended recipient in those raster cells, the probability that the food will go to waste and the probability that it then will be in contact with the unintended recipients. Thus, we can arrive at a risk of infection for the intended or unintended species at the raster cell level by combining all these terms into the final risk of infection calculation in the generic framework.

## Data

We outline the possible data sources that can be used for some of the steps of the framework for this pathway in Table 3.

TABLE 3 DATA REQUIRED TO COMPUTE THE RISK OF INFECTION WITHIN THE GENERIC FRAMEWORK FOR THE TRADE OF LEGAL FOOD INCURSION PATHWAY.

Parameter	Specific Data	Further Details	Potential Data Sources
Movement from Area A to Area B ( $N_q(g)$ )	Trade in food products	Trade in food products from regions in Area A to locations in Area B.	UN Comtrade data Eurostat COMEXT data
Prevalence ( $p_q$ )	Prevalence of the disease in Area A	Calculated using data on the number of cases that have been reported. Preferably for the same regions in Area A as the movement data.	OIE Animal Disease Notification System FAO EMPRES-i
	Initial viral load	Concentration of initial viral load in infected product	Published literature
Survival	Food processes	A list of which processes are performed on each product type	Published literature
	Transport times	Duration of transport and the split between different transport times	Published literature

			Distance calculation via maps
	Viral load decay rate	The rate at which viral load decays in each product type	Published literature
Susceptible Hosts ( $S(g)$ )	Abundance of intended recipient	Spatial distribution of the abundance of the intended recipient of the food product	Gridded Population of the World (SEDAC) Gridded Livestock Population of the World (FAO)
	Abundance of unintended recipient	Spatial distribution of the abundance of the unintended recipient of the food product, if measuring contact with wasted food product	Gridded Livestock Population of the World (FAO) Global Biodiversity Information Facility (GBIF)
	Proportion of food that goes to waste	Amount of food that is not eaten by the intended recipient	Published literature

### Case Study (Draft Results)

The case study that we are considering for the trade in food products incursion pathway is African Swine Fever (ASF). This is a disease of pigs that is of significant concern for the pig industry as it can spread rapidly, leading to a high death toll in pigs and boar, and there is no vaccine currently available (Costard et al. 2013). The disease is an OIE-notifiable disease and would lead to restriction of trade in pigs and pork products if found in a region, as well as significant culling procedures (Arias et al. 2018). It therefore has the potential to be a very costly disease. Furthermore, the disease was found in Georgia in 2007 after previously being predominantly associated with sub-Saharan Africa, and subsequently has been spreading westwards into Europe and eastwards across Russia (Costard et al. 2013). In 2014, Latvia, Lithuania and Estonia were the first European countries to report and since then there has been steady spread to Poland, Hungary, Romania, Bulgaria, the Czech Republic as well as a recent jump to Belgium (Arias et al. 2018). China, with the world’s largest pig industry, reported its first case in 2018 followed swiftly by 5 other cases in different regions (World Organisation for Animal Health 2018). Thus, this disease clearly poses a significant risk for animal health and the pig industry in many countries.

As ASF is a disease of pigs and wild boar, we considered the trade in pig meat. This is a product that is intended for human consumption but we were interested in the risk of infection in unintended recipients, domestic pigs and wild boar. Our Area A was the whole world and our Area B was the EU. We used Comext data (Eurostat 2017) for trade in food products and extracted all products that are related to pig meat. As there are numerous, very

specific product codes, we combined these into groups of product types determined by the processes that would occur to the different products. We included within the prevalence calculation the likelihood that an animal would be sent to slaughter if infected in Area A. Specifically, we reduced the prevalence in Area A to only account for animals that are currently in their incubation period of the disease, as we assumed that infected animals with clinical signs of a disease would not be sent to slaughter. We used data on the percentage of each product type that is formed of meat, bone, skin, fat, offal and other in order to calculate the initial viral load in each product, based on the prevalence within the product (Simons et al. 2017) and data on initial concentrations in meat, bone etc. We used published literature to determine the decay rate of ASF in meat products. We used data from Simons et al. (2016) on the average speed of different transport methods and the proportion of product that would be transported using that method and calculated distances between Area A and B to determine the viral load decay during transport. As the intended recipient of the pig meat is humans, we distributed the viral load across each country in the EU to each raster cell by the abundance of humans in each cell. Whether the pig meat was cooked or not was dependent on the processing type. If uncooked, then the amount of viral load that ends up in waste meat in that raster cell was determined solely by the amount of wastage of food within that country. If the food was cooked, then an additional probability that the food was cooked to a high enough temperature to kill the virus was included when calculating the total viral load in the waste meat. There are two unintended recipients in this case study that we are interested in, domestic pigs and wild boar. For the pigs, we assumed that contact with the waste meat could only occur through the illegal practice of swill feeding, which we assumed would only occur on backyard pig farms and not commercial farms. We estimated the number of backyard pig farms in each raster cell and the average size of backyard pig farms in each country using Eurostat data (Eurostat 2017). Combining the amount of viral load within each household with the number of backyard pig farms, the average number of pigs per backyard farm, the probability that swill feeding occurs, and a dose-response in pigs determines the risk of infection in backyard pig farms in each raster cell. For infection of wild boar, we assumed that all waste meat that is not swill fed to backyard pig farms is thrown away in bins and sent to landfill. We did not model this process exactly due to a lack of data on locations of landfills, but we assumed that the waste food remained in the same raster cell as the individuals who threw the waste out. We used a contact rate of wild boar with landfills and other waste, and the dose-response of wild boar to ASF, to determine the risk of infection in wild boar. Currently, the code for this case study pathway is deterministic, unlike the other pathways, but it will be made stochastic in the future.

In Figure 4, we show the total viral load that is present across the EU in each raster cell from trade in pork products, after we have accounted for processing, transport, cooking and the likelihood the product goes to waste. As expected, since the pork products are intended for human consumption, the viral load is correlated to human density within a country. However, between countries there are differences due to where each country imports their pig meat from. For example, Latvia and Lithuania have relatively high viral load across the countries, which is likely due to them trading mostly amongst Eastern European countries, which have a high estimated prevalence for ASF as there has been many reported cases in that region. Using this map of viral load, we then estimate the risk of infection in domestic pigs and wild boar. Our full results for this pathway will be included in a paper on the risk of ASF through multiple pathways.

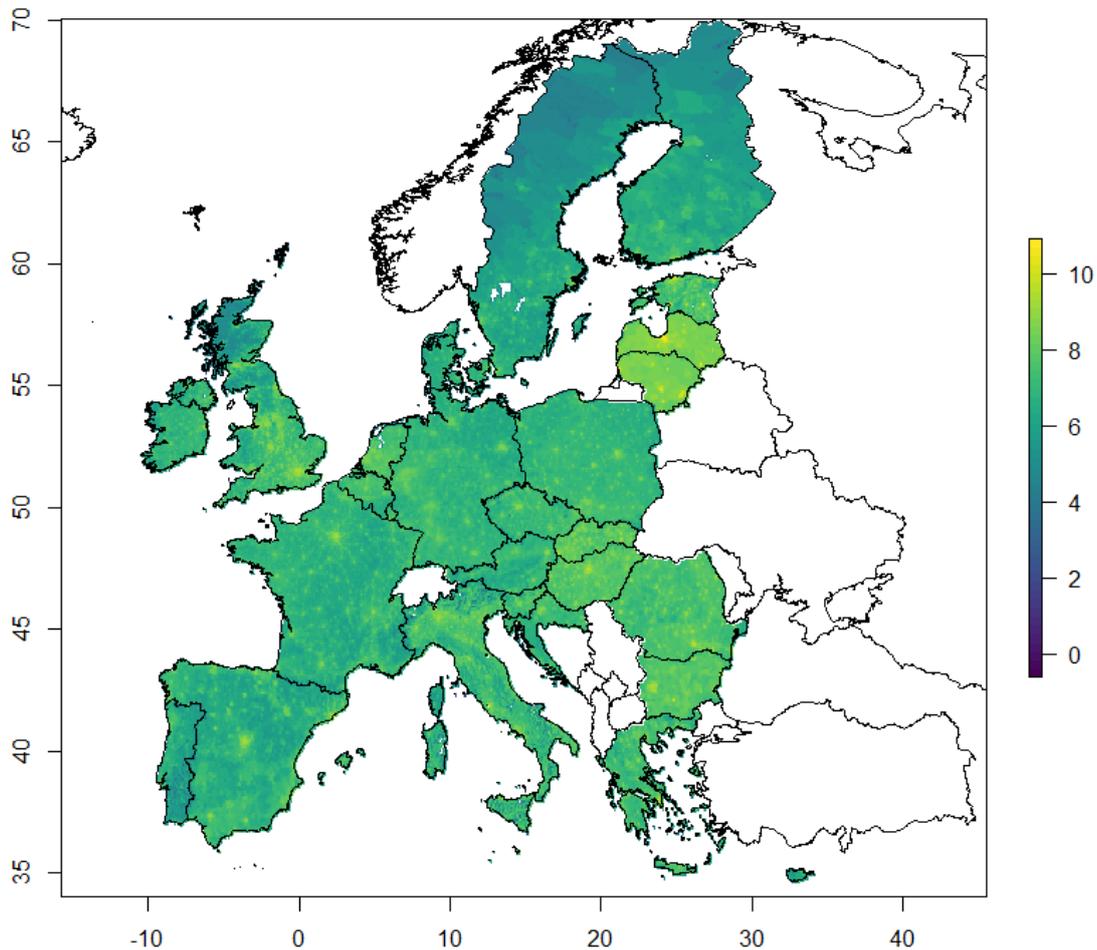


FIGURE 4 THE TOTAL AMOUNT OF VIRAL LOAD OF ASF PRESENT IN EACH CELL (ON A LOG SCALE) ACROSS THE EU FROM TRADE IN PORK PRODUCTS. THE VIRAL LOAD INCLUDES THE REDUCTION DUE TO PROCESSING AND TRANSPORT AS WELL AS ONLY INCLUDING THE PROPORTION THAT WOULD OCCUR IN MEAT THAT GOES TO WASTE.

### Terrestrial Movement of Wild Animals

Terrestrial movement of wild animals includes all those movements by wild animals that are, for the most part, restricted to land and are affected by land type and habitat, such as wild boar and red foxes. This route does not include the migration of birds or the flight (natural or driven by wind) of vectors. These have specific issues and have been included in other pathways within our code. Similarly, this pathway is restricted to natural movement of those wild animals, so human movement of those animals is not included. Movement of wild animals that is initiated due to large quantities of hunting in the area can be included by increasing the probability of animals likely to move long distances. We have not included human barriers to movement – for example, if fences were put up to ensure the wild animals did not move to specific locations, but this can be included in the model with some adaptation of the code.

## Implementation of the Framework

Terrestrial movement of wild animals is not as simple as the legal trade of live animals to implement within our generic framework (Figure 1), since there are no large-scale data sources for the number of wild animals moving from one location to another to input for our entry step. These do not exist even at a country level, and we wish to be able to perform our risk assessment on a finer scale than country level if possible. Therefore, we developed a module for the entry step to estimate the number of wild animals moving from region  $q$  in Area A to location  $g$  in Area B,  $N_q(g)$ , which allows us to compute our risk assessment on a finer spatial scale. The module also makes use of the raster format at a 100km<sup>2</sup> cell level.

Our module for entry incorporates data on the location of wild animals and determines where these animals will move to in a set amount of time. It does this by assuming two different types of movement – a home-range movement and a long-range dispersal. We assume that all animals perform the home-range movement but include a parameter that determines what percentage of the wild animals will undergo the long-range dispersal. This may be based on age or gender of the wild animals (or hunting pressure). The home range of the animal represents the whole area that they are able to cover over a set length of time. On the other hand, the long-range movement is based on the habitat suitability of each cell, where habitat suitability is a score between 0 and 1 determining how suitable each cell is for that wild animal to live. This method assumes that long-range movement is performed to find good territory and the wild animal will only stop in this new area if the habitat is suitable. Specifically, long-range movement is modelled by assuming the wild animal performs a biased random walk from cell to cell, where the probability of moving from one cell to any of its neighbours is determined by the habitat suitability. This ensures that the wild animal will choose to move to better territory, but allows for the animal to stay in the current cell if the habitat suitability in surrounding cells is much lower. Given the maximum long-range distance that the wild animal could travel, we calculate the total number of steps (defined as a movement from one cell to a neighbouring cell) the animal can travel,  $n$ , to be the total distance in km divided by the width of a cell, in our case 10km. We use  $N(c)$  to indicate the neighbourhood of cell  $c$ , specifically the 8 cells to the north, east, south and west, as well as the cell itself. The probability that a wild animal will be in cell  $c$  after  $n$  steps,  $p(c, q, n)$ , given that the animal started from cell  $q$  in Area A, is then given by the recurrence relation:

$$p(c, q, n) = \sum_{\tilde{c} \in N(c)} B(c, \tilde{c}) p(\tilde{c}, q, n - 1),$$

where  $B(c, \tilde{c})$  is a probability of moving from cell  $\tilde{c}$  to cell  $c$  determined by the benefit an animal will receive by moving, and is multiplied by the probability of being in cell  $\tilde{c}$  after  $n - 1$  steps. The recurrence relation has initial conditions at step 0 of:

$$\begin{cases} p(c, q, 0) = 1 & \text{if } c = q \\ p(c, q, 0) = 0 & \text{if } c \neq q \end{cases}$$

Note that for cells neighbouring  $q$ , the origin cell, we do not allow movement back to  $q$  in order to ensure dispersal from the home range does actually take place. The benefit probability  $B(c, \tilde{c})$  is based on the habitat suitability score,  $h(c)$ , in each cell  $c$ . It is calculated by comparing the difference in the suitability between the two cells  $c$  and  $\tilde{c}$  and normalising this by the benefit that could be gained by moving to any of the neighbouring cells  $c^*$  of cell  $\tilde{c}$ :

$$B(c, \tilde{c}) = \frac{1 + h(c) - h(\tilde{c})}{\sum_{c^* \in N(\tilde{c})} (1 + h(c^*) - h(\tilde{c}))}$$

We model one movement event within the time period chosen for the model but do not specifically model when this event occurs relative to the time period. Once the animal arrives at their final destination, we assume they perform home-range movement in this new cell for the rest of the time period.

As we are calculating the entry step on a cell level, rather than from some region in Area A to a cell in Area B, we provide the ability to calculate prevalence in this pathway at a cell level as well (although it is possible to use regional estimates of prevalence instead of cell estimates). Therefore, the code includes user input of exact locations of reported cases of the disease in the wild animal in Area A. We calculate prevalence at a cell level using the suggested method in the outline of the generic framework above, which we denote here as  $\tilde{p}_q$  in cell  $q$ , namely,

$$\tilde{p}_q = U_q \frac{c_q}{N_{P(i,q)}} \frac{\tau}{365}$$

However, since we are calculating prevalence on a fine spatial scale, it is possible that many cases could be close to the borders of other cells, and therefore neighbouring cells could also have had cases that were unreported. We therefore include a smoothing effect to the prevalence across neighbouring cells. The final prevalence,  $p_q$ , is therefore calculated as

$$p_q = 0.5\tilde{p}_q + \sum_{q^* \in N(q), q^* \neq q} \frac{0.5}{8} \tilde{p}_{q^*},$$

where we split the weight the new prevalence predominantly by the old prevalence in cell  $q$  but include the effect of each of the 8 neighbouring cells,  $N(q)$ , of cell  $q$ .

For the contact and transmission steps of the framework, we have to estimate the number of susceptible animals that could be in contact with the infected species  $i$ . For this, we use population abundance maps of the susceptible species on the same raster cell level as the model calculations.

## Data

We outline the possible data sources that can be used for some of the steps of the framework for this pathway in Table 4.

TABLE 4 DATA REQUIRED TO COMPUTE THE RISK OF INFECTION WITHIN THE GENERIC FRAMEWORK FOR THE TERRESTRIAL MOVEMENT OF WILD ANIMALS INCURSION PATHWAY.

Parameter	Specific Data	Further Details	Potential Data Sources
Movement from Area A to Area B	Movement of wild animals	Average home range size or distance travelled by wild animals.	Published literature
$(N_q(g))$	Location and abundance of wild animals in Area A	Spatial distribution and numbers of wild animals in Area A.	Published literature Global Biodiversity Information Facility (GBIF)

	Habitat suitability of wild animals in Area A and B	A map with the same spatial extent as the abundance map which gives a score for how suitable every cell is for the wild animal	Published literature
Prevalence ( $p_q$ )	Prevalence of the disease in Area A	Calculated using data on the number of cases that have been reported and the number of hosts in Area A. Preferably for the same regions in Area A as the movement data.	OIE Animal Disease Notification System FAO EMPRES-i
Susceptible Hosts ( $S(g)$ )	Spatial distribution of wild animals	To determine where and how many susceptible hosts the imported hosts will have contact with	Published literature Global Biodiversity Information Facility (GBIF)

### Case Study (Draft Results)

For the case study of the terrestrial movement of wild animal incursion pathway, we once again considered ASF as the disease of interest. See the case study description in the previous pathway, trade in food products, for a description of the disease. As stated, ASF is a disease that has led to cases within both wild boar and domestic pigs. Within Europe, most countries reported wild boar cases before any cases in pigs. Thus, wild boar have been implicated as a reservoir for ASF and the main mode of transmission within Europe (Arias et al. 2018). We decided to use ASF as a case study to investigate, in particular, the role that wild boar may play in spreading the disease. We model the risk of infection in both pigs and boar in new cells across Europe due to wild boar moving from cells that had reported cases, using our terrestrial movement of wild animal incursion pathway.

For this case study, our Area A is all cells across Europe that are estimated to have a non-zero prevalence according to the method previously mentioned, namely calculating a smoothed prevalence based on an under-reporting factor and the locations of reported cases of ASF in wild boar. The reported cases of wild boar were extracted from Empres-i. Our Area B is all cells in Europe. The movement model is calculated based on abundance and habitat suitability maps that we adapted from Alexander et al. (2016). We estimated the length of long-distance movement and the size of the home range of boar from published literature. We assumed that detection of wild boar moving from cell to cell does not occur and therefore we set this parameter in the detection step to zero. We also used abundance maps of the susceptible populations. For the susceptible boar population, we used the same boar abundance maps as above, but we reduced the number of animals in each cell by the number of estimated cases of ASF in boar in that cell in order to get a total number of susceptible boar. For the pig density maps, we used data from the FAO gridded livestock of the world (FAO 2014). We modelled the contact between live infected boar and susceptible boar and pigs using direct contact transmission, as outlined in the generic framework, with the relevant parameters found in published literature. However, as ASF is a highly virulent disease and can remain in the environment and in meat for a long time, we chose to include the potential situation that the boar dies from infection but the resulting carcass is still able to transmit the disease. We thus adapted the contact rates and transmission probabilities to include the probability that live boar contact boar carcasses and

the probability of transmission from such an event (Probst et al. 2017). However, we assumed that pigs will not have any contact with boar carcasses, as we assumed that biosecurity of farms is high enough to prevent such an event.

For this case study, we considered both the current risk of ASF across the whole of Europe based on the recent cases of ASF in wild boar in Europe, as well as the estimated risk of ASF compared against historic cases. In particular, we considered the specific case of Poland from 2014-2016. We used reported cases of ASF in wild boar in Poland in 2015 to predict the risk of infection in boar and pigs in 2016. In Figure 5, we plot the estimated mean risk of ASF in boar and pigs in 2016, using reported cases in 2015 to estimate the prevalence at a cell level. Our current model predicts the risk to be mainly focused in the area that had reported cases in 2015, but with a wide circle of potential transmission. This circle of potential transmission does include the clusters to the west and south of the plot, although the risk in these areas is very low. We also provide plots of reported cases of wild boar and pigs in Poland in 2016 separately, colour-coded according to the month they were reported (Figure 5B, D). In these plots, it is evident that in the cluster of cases to the west of the plots, all of the pig cases occurred before there were any reported cases in boar. Similarly, in the cluster in the south of the plots, most of the pig cases occurred in August with a couple of cases in September. Two wild boar cases were reported in this area in August but otherwise all the wild boar cases in this region occurred between October to December 2016. Specifically, the first pig case was reported on the 10<sup>th</sup> August and the first wild boar case was reported on the 13<sup>th</sup> August. This could be due to cases in boar being undetected rather than no cases but, overall, our model suggests that jumps between disease clusters, even on a medium range as seen within Poland in Figure 5, may not be solely caused by wild boar. Instead, these clusters may have been unintentionally caused by humans or other transmission pathways. We are currently preparing a manuscript on this case study including further results.

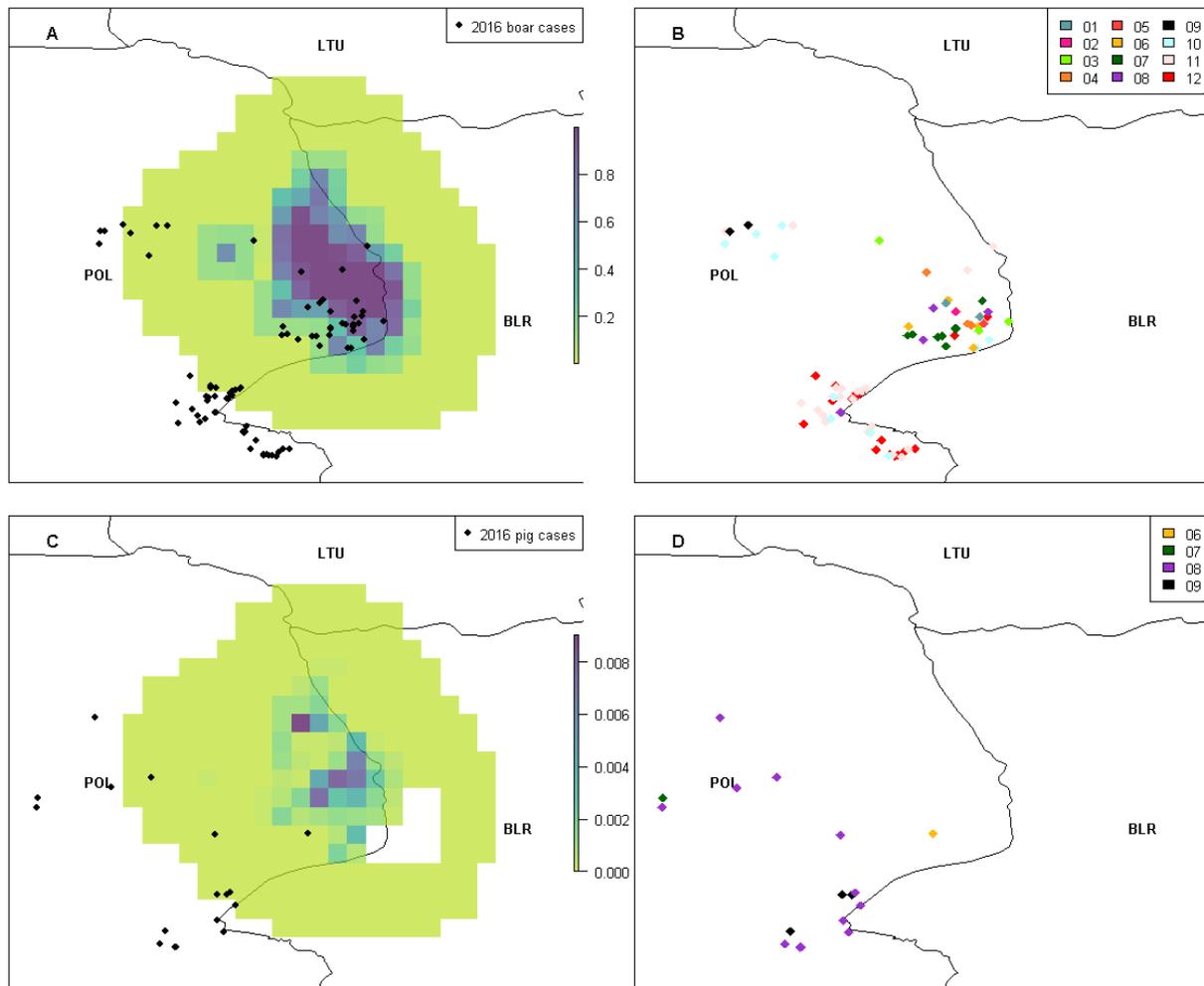


FIGURE 5 THE MEAN RISK OF INFECTION OF ASF IN BOAR AND PIGS IN 2016 DUE TO WILD BOAR MOVEMENT USING REPORTED CASES FROM 2015 TO ESTIMATE PREVALENCE IN BOAR. IN (A) AND (C) THE TOTAL ESTIMATED MEAN RISK IN BOAR AND IN PIGS, RESPECTIVELY, IS PLOTTED, DUE TO TRANSMISSION BY LIVE BOAR OR BOAR CARCASSES. REPORTED CASES ARE PLOTTED AS BLACK CIRCLES. IN (B) AND (D) THE REPORTED CASES OF WILD BOAR AND PIGS, RESPECTIVELY, ARE PLOTTED, COLOUR-CODED ACCORDING TO THE MONTH OF THE OBSERVATION DATE OF THE CASE. COUNTRIES ARE INDICATED BY THEIR ISO3 CODE.

### Bird Migration

The focus of the bird migration pathway is the incursion of disease due to birds migrating from one area to another. Thus, as well as the direct introduction of infected wild birds with disease such as Avian Influenza, it also accommodates incursion due to healthy migrating birds that carry infected ticks or other species as “hitchhikers”, such as the *hyalomma marginatum* tick which has been shown to carry Crimean-Congo haemorrhagic fever.

### Implementation of the Framework

The entry step of the framework, in which we estimate how many infected birds (or ticks on birds) are entering the new Area B, is calculated in a similar manner to the human transportation pathway. We consider the bird

migration routes that occur across the world and use these to determine which countries should be included in Area A and where in Area B the birds will enter. There exist datasets of these migration routes and which species of wild birds will use each flyway. We can therefore split the entry into the different species, similar to the different products entering in the trade of food products pathway. If data exists on the differences between the wild bird species, such as in contact or transmission rates with poultry or in prevalence, then it is best if the framework is used separately for each bird species. The risks of infection for each bird species can then be combined together to produce one risk of infection by any bird species using the description in the framework. To estimate how many wild birds of each species are using each migration route, it is possible to use data on the number of the migratory species that have reached the new Area B in the past, as well as species distribution methods to determine the abundance in areas with missing data. To calculate the entry at a finer spatial scale, we once again utilize the raster cell format as first outlined in the human transportation pathway. We assume that the migratory birds will travel to the same locations as they have in the past, and therefore use data on where the migratory birds usually go to in order to distribute the birds within Area B. If the data for number of birds entering is based upon the number of birds that usually reach Area B, then we use this data as is. However, if the data is based upon how many birds leave Area A, then we include an additional probability that the bird will survive the migration journey.

When calculating the prevalence of disease in wild birds that enter Area B, we adapt the prevalence calculation to include the fact that infected birds are less likely to survive the migration journey than non-infected birds. For the prevalence calculations when the pathway is considering hitchhikers on the birds, such as ticks, we include a similar additional parameter, namely the probability that the tick would survive the migration journey.

For contact with other species, we consider which hosts are in the cell that the migratory birds go to in Area B. If data exist on how many farms are outdoor versus indoor poultry farms, which would indicate different contact rates with wild birds, this can be included. There may be multiple locations in Area B that the same bird would stop at during its migration journey. The ability to model multiple stops is included within the pathway, as the risk is calculated at each location rather than for each bird. Similarly to multiple stops, it may be important to include a low contact rate between migratory birds and other hosts for the locations that the birds fly over but do not stop. This can also be included within the framework by adjusting the contact rate by a parameter that indicates it is a flyover cell.

For infected ticks hitchhiking on birds, we assume that, provided they survive the journey, they would enter Area B at the first location that the birds enter (if the bird has multiple stops in Area B). For the contact, survival and transmission of infection from ticks to a new host, we include the probability that the ticks were juvenile and therefore will bite another host, the mortality rate of ticks, and the abundance of host species in the location  $g$  where the tick enters Area B. Similar to the vectors on human transportation pathway, we include the effects of temperature and seasonality on the survival, contact and transmission rates by using temperature-dependent life history traits of vectors and average temperatures in each raster cell.

## Data

We outline the possible data sources that can be used for some of the steps of the framework for this pathway in Table 5.

TABLE 5 DATA REQUIRED TO COMPUTE THE RISK OF INFECTION WITHIN THE GENERIC FRAMEWORK FOR THE BIRD MIGRATION INCURSION PATHWAY.

Parameter	Specific Data	Further Details	Potential Data Sources
Movement from Area A to Area B ( $N_q(g)$ )	Movement of birds	Global migration pathways of birds and which bird species travel which routes.	Published literature
	Location and abundance of birds in Area A	Spatial distribution and numbers of birds in Area A.	Published literature Global Biodiversity Information Facility (GBIF)
	Location and abundance of birds in Area B	Spatial distribution and numbers of birds in Area B to determine where the birds will travel to.	Published literature Global Biodiversity Information Facility (GBIF)
	Survival	Probability that an infected bird or the attached tick would survive the migration journey.	Published literature
Prevalence ( $p_q$ )	Prevalence of the disease in Area A	Calculated using data on the number of cases that have been reported and the number of hosts in Area A. Preferably for the same regions in Area A as the movement data.	OIE Animal Disease Notification System FAO EMPRES-i
Susceptible Hosts ( $S(g)$ )	Spatial distribution of hosts	To determine where and how many susceptible hosts the imported hosts or ticks will have contact with.	Published literature Global Biodiversity Information Facility (GBIF)

### Vector flight

The focus of the vector flight incursion pathway is on any movement of vectors that is not by hitchhiking on another mode of transport (e.g. human transportation or birds). Therefore, it incorporates both the short-distance natural flight of vectors performed on a daily basis, and windborne longer-distance travel of vectors. For the longer-distance travel of vectors, we assume this is driven solely by the wind and the vector does not have a choice in where it lands afterwards.

### Implementation of the Framework

For the entry step of the framework, we have to estimate how many vectors are moving from one area to another. We implement this step of the vector flight pathway in a very similar manner to the terrestrial movement of wild animals entry step as they both contain two different lengths of movement – “home-range” movement and a longer-distance travel. We include the home-range movement for vectors, i.e. the movement they would perform naturally and unaided, in exactly the same way as for terrestrial movement of wild animals. For the longer-distance windborne flight of vectors, we use a similar method to the dispersal of terrestrial wild animals. Once again, we require a population abundance map of vectors at a cell level. Then, the total number of steps from cell to cell that the vectors can travel will be determined by average distances that vectors can be blown by the wind. The movement from cell to cell is driven solely by the wind. The probability that a vector will be in cell  $c$  after  $n$  steps,  $p(c, q, n)$ , given that the vector started from cell  $q$  in Area A, will be given by the adapted recurrence relation:

$$p(c, q, n) = \sum_{\tilde{c} \in N(c)} W(c, \tilde{c})p(\tilde{c}, q, n - 1),$$

where  $W(c, \tilde{c})$  is the proportion of time that the wind is blowing in the direction from cell  $\tilde{c}$  to cell  $c$ . In this first version of the code we assume only wind direction is important, but wind speed and altitude could be included. Unlike the terrestrial movement pathway in which movement over the sea is not possible, for the vector pathway the sea is not a barrier to movement as vectors can be blown over the sea by wind. It is also possible that vectors will travel further distances over sea than land which is a further aspect that could be included in a later version of this pathway.

The prevalence of the vectors can also be determined on a cell level, as in the terrestrial movement of wild animal pathway. However, the prevalence in vectors is usually calculated by using reported cases in the main disease host as it is rare for data to exist on the number of infections in vectors. We can use the same method as in the terrestrial movement of wild animal pathway, with the reported cases to estimate of the prevalence in the host species. Published literature is then used to estimate the conversion factor to calculate the equivalent prevalence in vectors, as prevalence in the vector population is usually lower than the prevalence in the host species (Anderson and May 1992).

For the survival of infected vectors and contact with susceptible species in Area B at a cell level, we use temperature-dependent life-history traits of the vectors, similar to the human transportation pathway. We include the effect of average temperatures at each cell level as well as differences in land type on the life-history traits; while vectors can travel over the sea, they are not able to survive for very long over the sea naturally. The abundance of susceptible species also has to be given by a population map on the same spatial scale as the vector abundance map.

### Data

We outline the possible data sources that can be used for some of the steps of the framework for this pathway in Table 5.

**TABLE 6 DATA REQUIRED TO COMPUTE THE RISK OF INFECTION WITHIN THE GENERIC FRAMEWORK FOR THE VECTOR FLIGHT INCURSION PATHWAY.**

Parameter	Specific Data	Further Details	Potential Data Sources
-----------	---------------	-----------------	------------------------

Movement from Area A to Area B ( $N_q(g)$ )	Movement of vectors	Average home range size or distance blown by the wind.	Published literature
	Location and abundance of vectors in Area A	Spatial distribution and numbers of vectors in Area A.	Published literature Global Biodiversity Information Facility (GBIF)
	Habitat suitability of vectors in Area A and B	A map with the same spatial extent as the abundance map which gives a score for how suitable every cell is for the vectors – primarily based on temperature.	Published literature UK Met Office European Environment Agency
	Wind direction	Average wind direction for each cell in Area A and Area B.	UK Met Office European Environment Agency
Prevalence ( $p_q$ )	Prevalence of the disease in Area A	Calculated using data on the number of cases that have been reported. Preferably for the same regions in Area A as the movement data.	OIE Animal Disease Notification System FAO EMPRES-i
	Conversion Factor	Prevalence of the disease in vectors relative to main hosts.	Published literature
Susceptible Hosts ( $S(g)$ )	Spatial distribution of susceptible hosts	To determine where and how many susceptible hosts the vectors will have contact with.	Published literature Global Biodiversity Information Facility (GBIF)
Contact and Survival	Vector life-history parameters	Effect of temperature on the vector life-history traits, such as mortality and bite rate.	Published literature

## Summary

We have presented a generic framework for the risk of infection and risk of spread of a disease in a new Area B given the presence of that disease in Area A. This is a powerful tool that can be utilized for rapid risk assessments when there are limited data or for in-depth risk assessment in which complicated methods and detailed data are

used in each step of the framework. The generic framework can be used on its own, by applying the nine steps of the framework to the user's risk question, without using the methods we have outlined for each pathway. Or it can be used more fully, by using the methods we have outlined for each pathway, downloading our code and inputting data relevant to the risk question. Currently, there are six major pathways of disease incursion included in the code, as outlined above. The framework is also flexible for various purposes because it is possible to use it to calculate the risk of infection or the risk of spread independently rather than both if desired.

Our framework for spatial quantitative risk assessments is ambitious in its scope, aiming to be generic across space, disease, pathway and transmission method, as well as performing the calculations at different spatial scales. However, in reality this may often not be possible. Data are the obvious limitation. Our case studies were performed with Area B as Europe, because there exist good European datasets curated by Eurostat with regulations in place to ensure EU partners submit data. However, across the world, the necessary data may not be collected at all, or not at a high enough standard, or may not be freely available. Therefore, this is likely to limit our ability to perform the risk assessment for any Area A and Area B, especially if we wish Area B to include multiple countries. Furthermore, many incursion pathways are complicated to add to the model because of a lack of data or the stochastic nature of the incursions. For example, it is believed that some of the cases of ASF in Europe have occurred due to illegal movements of wild boar carcasses or infected meat (Arias et al. 2018). However, finding quantitative data on how many people would transport meat, where exactly they would get the meat from, where and how they would travel, the likelihood that they drop some meat, and the probability that it would be found and eaten by a wild boar in another location, is nigh on impossible. There is no way to perform risk assessment quantitatively for those difficult cases when data just does not exist. However, our framework provides the tools to be able to perform the spatial quantitative risk assessment if the data becomes available or if suitable proxy data are found.

This deliverable report outlines our generic framework for spatial quantitative risk assessments, alongside information on the prototype of the code. Our prototype code includes six major pathways of disease incursion, which we detail above. We plan to produce a second version of the code, which will include more pathways of disease incursion. These will be chosen based upon the perceived importance of each pathway relative to others, as well as their relevance to any case studies we wish to complete. Furthermore, we will expand our current case studies, including performing a case study of the risk of spread framework. As it currently stands in this first version, we have produced a generic framework that is quantitative, able to consider multiple incursion pathways and intrinsically spatial, even at fine resolution. This allows for the relative risk to be compared across diseases and incursion pathways, disease hotspots identified and thus, surveillance and management plans to be put in place accordingly.

## References

- Alexander, N. S., G. Massei and W. Wint (2016). "The European Distribution of *Sus Scrofa*. Model Outputs from the Project Described within the Poster—Where are All the Boars? An Attempt to Gain a Continental Perspective." Open Health Data 4(1).
- Anderson, R. M. and R. M. May (1992). Infectious diseases of humans: dynamics and control, Oxford university press.
- Arias, M., C. Jurado, C. Gallardo, J. Fernández-Pinero and J. Sánchez-Vizcaíno (2018). "Gaps in African swine fever: Analysis and priorities." Transboundary and emerging diseases 65: 235-247.

- Brasil, P., J. P. Pereira Jr, M. E. Moreira, R. M. Ribeiro Nogueira, L. Damasceno, M. Wakimoto, R. S. Rabello, S. G. Valderramos, U.-A. Halai and T. S. Salles (2016). "Zika virus infection in pregnant women in Rio de Janeiro." New England Journal of Medicine **375**(24): 2321-2334.
- Brown, E. B., A. Adkin, A. R. Fooks, B. Stephenson, J. M. Medlock and E. L. Snary (2012). "Assessing the risks of West Nile virus–infected mosquitoes from transatlantic aircraft: implications for disease emergence in the United Kingdom." Vector-Borne and Zoonotic Diseases **12**(4): 310-320.
- Chouin-Carneiro, T., A. Vega-Rua, M. Vazeille, A. Yebakima, R. Girod, D. Goindin, M. Dupont-Rouzeyrol, R. Lourenço-de-Oliveira and A.-B. Failloux (2016). "Differential susceptibilities of *Aedes aegypti* and *Aedes albopictus* from the Americas to Zika virus." PLoS neglected tropical diseases **10**(3): e0004543.
- Costard, S., L. Mur, J. Lubroth, J. Sanchez-Vizcaino and D. Pfeiffer (2013). "Epidemiology of African swine fever virus." Virus research **173**(1): 191-197.
- Davies, F. G. (1991). "Lumpy skin disease of cattle: a growing problem in Africa and the Near East." World Animal Review **68**(3): 37-42.
- Eurostat. (2017). "Eurostat Bulk Download Listing." Retrieved 9th May, 2017, from <http://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing>.
- FAO. (2014). "Gridded Livestock of the World (GLW)." Retrieved May, 2017, from <http://www.fao.org/ag/againfo/resources/en/glw/home.html>.
- Haseyama, M., S. Iizuka, H. Omae and Y. Tsuda (2007). "Results of mosquito collection from international aircrafts arriving at Narita International Airport, Japan and mosquito surveillance at the airport." Medical Entomology and Zoology **58**(3): 191.
- Hill, A., T. Dewé, R. Kosmider, S. Von Dobschuetz, O. Munoz, A. Hanna, A. Fusaro, M. De Nardi, W. Howard and K. Stevens (2015). "Modelling the species jump: towards assessing the risk of human infection from novel avian influenzas." Royal Society open science **2**(9): 150173.
- Horigan, V., M. De Nardi, R. Simons, S. Bertolini, M. Crescio, A. Estrada-Peña, A. Léger, C. Maurella, G. Ru and M. Schuppers (2018). "Using multi-criteria risk ranking methodology to select case studies for a generic risk assessment framework for exotic disease incursion and spread through Europe." Preventive veterinary medicine **153**: 47-55.
- Hunter, P. and D. Wallace (2001). "Lumpy skin disease in southern Africa: a review of the disease and aspects of control." Journal of the South African Veterinary Association **72**(2): 68-71.
- Kumar, S. M. (2011). "An outbreak of lumpy skin disease in a Holstein dairy herd in Oman: a clinical report." Asian Journal of Animal and Veterinary Advances **6**(8): 851-859.
- Mercier, A., E. Arsevska, L. Bournez, A. Bronner, D. Calavas, J. Cauchard, S. Falala, P. Caufour, C. Tisseuil and T. Lefrançois (2017). "Spread rate of lumpy skin disease in the Balkans, 2015–2016." Transboundary and Emerging Diseases.
- Murray, N. (2004). Handbook on import risk analysis for animals and animal products: quantitative risk assessment, Office international des épizooties.
- Musso, D., C. Roche, E. Robin, T. Nhan, A. Teissier and V.-M. Cao-Lormeau (2015). "Potential sexual transmission of Zika virus." Emerging infectious diseases **21**(2): 359.
- Pan American Health Organization (PAHO). "Health Information Platform for the Americas." Retrieved January, 2018, from <http://www.paho.org/data/index.php/en/>.
- Plourde, A. R. and E. M. Bloch (2016). "A literature review of Zika virus." Emerging infectious diseases **22**(7): 1185.
- Probst, C., A. Globig, B. Knoll, F. J. Conraths and K. Depner (2017). "Behaviour of free ranging wild boar towards their dead fellows: potential implications for the transmission of African swine fever." Royal Society open science **4**(5): 170054.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria <https://www.R-project.org/>.
- Simons, R. R., V. Horigan, P. Gale, R. D. Kosmider, A. C. Breed and E. L. Snary (2016). "A Generic Quantitative Risk Assessment Framework for the Entry of Bat-Borne Zoonotic Viruses into the European Union." PLoS one **11**(10): e0165383.
- Simons, R. R. L., V. Horigan, M. De Nardi, G. Ru, A. E. Pena and A. Adkin (2017). Mighty models from little data grow: Estimating animal disease prevalence. Proceedings of the Society for veterinary epidemiology and preventative medicine. Inverness, Scotland: 166.
- Taylor, R. A., A. D. Berriman, P. Gale, L. A. Kelly and E. L. Snary (2018). "A generic framework for spatial quantitative risk assessments of infectious diseases: lumpy skin disease case study." Transboundary and emerging diseases.



TRACES. (2017). "TRACES - Trade Control and Expert System." Retrieved 12th July, 2017, from [https://ec.europa.eu/food/animals/traces\\_en](https://ec.europa.eu/food/animals/traces_en).

Tuppurainen, E. and C. A. Oura (2012). "lumpy skin disease: an emerging threat to Europe, the Middle East and Asia." *Transboundary and Emerging Diseases* **59**(1): 40-48.

World Organisation for Animal Health. (2018). "OIE - WAHIS Interface." Retrieved August, 2018, from <http://www.oie.int/en/animal-health-in-the-world/the-world-animal-health-information-system/the-world-animal-health-information-system/>.