PANDHUB

Grant Agreement No:	607433
Project Acronym:	PANDHUB
Project Title:	Prevention and Management of High Threat Pathogen Incidents in Transport Hubs

D3.3 – Simulation of Passenger Flows

Lead participant:	Public Health England (Dept of Health)
Authors:	Caroline Walters, Emma Bennett and Ian Hall[1]
Reviewers:	Ilpo Kulmala and Nicolas Poirot
Dissemination level:	PU = Public
Date	06.02.2017

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 607433

[1] other staff in Public Health England have read and commented on the deliverable.

Executive Summary

1. Four case study infections were chosen as representative of the types of threat within scope of PANDHUB in D2.1. This report describes modelling frameworks for simulating passenger infection from transmission events involving these diseases, whether pandemic or deliberate release, and considers the role of transient contact and movement.

2. Mathematical models can be viewed as conceptual tools used for explaining real-world systems and predicting future outcomes and consequently they increase our understanding of a system. No single model is “correct” so it is important to assess what mathematical approach is most appropriate for each situation. A good model should be suited to its purpose and parametrised by available data.

3. We conducted a scoping review to identify techniques which have been employed to model the global spread of a pandemic potential disease and what data are available to parameterise these models. Validation is important because it indicates how well a model represents the real-world situation. We identified whether the models were validated and, if so, which data were used to do this.

4. The search terms for the scoping study included terminology for established modelling types. We searched explicitly for influenza as, from experience, many global disease spread models are for pandemic influenza. We also included a more general global disease spread term, along with terms ‘global’ and ‘pandemic’. The databases Embase, PubMed and Scopus, yielded 799 records after all duplicates were removed.

5. All 799 records were screened first on title and abstract. After full text screening and citation checking of included articles, 79 articles remained to be reviewed. The majority of these 79 articles did not have any validation data and population-level metapopulation models were the most common model type. Nine metapopulation model records reported validation data. This highlights that models currently exist and can be parameterised and validated: we do not consider it necessary to construct a new global disease spread model for the PANDHUB project.

6. Given that pandemic infections will spread globally, importation to European countries is likely, though the numbers may be low depending on international public health efforts and the disease transmission characteristics in question. This report presents a model for calculating the probability an infected individual becomes infectious or symptomatic during transit. This can be used to assess whether major transport hubs might have a role in mitigation by detecting imported cases.

7. For a deliberate release of biological material one needs to consider the mechanism of delivery, the airborne dissemination, and the effect of the material on humans in the vicinity. An eddy diffusion model can allow for partial reflection of material from surfaces. Critical to the model is the interaction between the eddy diffusion parameter, air exchange rates and dwelling time.

8. We recommend that for a detailed risk assessment of a bio-aerosol release in a transport hub, computational fluid dynamics simulations be conducted. However, given the mismatch in apparent precision between the fluid flow calculations and

individual location and movement patterns in a hub it is sufficient to use the coarser scale modelling with sensitivity analysis conducted around other parameters.

9. Assessing exposure to humans arising from a release of aerosol material may involve custom-built models in a full agent-based modelling system or a highly simplified estimate of passenger volumes and flows. From the number of exposed individuals, and the deposited dose for those individuals, characteristics of the disease outbreak may be calculated. Among these characteristics are the number of patients exhibiting different classes of symptoms, the time to onset of symptoms, and the expected number of deaths (perhaps with the inclusion of particular intervention strategies).

10. In general, the time to onset of symptoms will be of the order of days for the diseases considered here so it is likely that patients will seek healthcare away from the location of the transport hub. This dispersion of cases from a large transport hub, possibly across many countries, complicates the detection and intervention aspects of any response, which will be considered in D3.4 and D3.5.

11. This report presents a framework for modelling the passenger flows and interactions between users of the hub. In D3.4 we will build on this basis by parameterising the models and showing the spatial impact, both within the hub and external to the hub, of the pandemic and deliberate release infections. In D3.5 we then consider mitigation, such as screening, effects further. We also consider influences on behaviour, such as improved communication, and person to person contact reduction.

Abbreviations

CDC	Centers for Disease Control and Prevention
CDF	Cumulative distribution function
D N.N	PANDHUB deliverable N.N
GLEaM	Global Epidemic and Mobility
IATA	International Air Transport Association
PDF	Probability distribution function
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analysis
SEIR	Susceptible-Exposed-Infected-Recovered
SIR	Susceptible-Infected-Recovered
WHO	World Health Organization
WP N	PANDHUB work package N

Contents

Executive Summary....2
Abbreviations....4
Contents....5
1. Introduction....7
2. Pandemic Potential Diseases....8
2.1 Motivation....8
2.2 Disease Transmission Modelling....9
2.3 Scoping Study - Methods....13
2.3.1 Literature Searches....14
2.3.2 Record Screening....15
2.3.3 Information Extraction....16
2.4 Scoping Study - Results....16
2.4.1 Identified Modelling Approaches....16
2.4.2 Identified Datasets....18
2.5 Scoping Study - Discussion....20
2.6 Entry Screening....20
2.6.1 Motivation....20
2.6.2 Ebola....21
2.6.3 The Ebola Screening Process....21
2.6.4 Mathematical formulation....22
2.6.5 Summary....25
3. Bioterrorist Diseases....26
3.1 Introduction....26
3.2 Natural reservoir of infection....26
3.2.1 Anthrax....26
3.2.2 Plague....27
3.2.3 Summary....28
3.3 Aerosol dispersal....28
3.3.1 Mean field approach....28
3.3.2 Turbulent flow: the Gaussian Puff Model....29
3.3.3 Computational fluid dynamics....34
3.3.4 Summary of dispersion modelling....36
3.4 Individual Human Exposure....36
3.4.1 Breathing rate....36
3.4.2 Deposition....37
3.4.3 Competing risks framework....37
4. Passenger flows....38
4.1 Room occupancy and transit....38
4.2 Determination of passenger exposure profiles....38
4.2.1 Agent-based models for passenger movement....39

4.2.2 Data analysis for passenger movement 40
4.2.3 Using passenger tracks to calculate exposure 41
4.3 Evaluating the disease outbreak in wider community 41
4.4 Discussion 42
5. Next steps and Recommendations 44
Supplementary information - Global disease spread scoping review 45
Supplementary Information– Direct intervention at transport hub through screening 72

Adobe Acrobat Document 72
Supplementary Information– screening questionnaire 73
Bibliography 74

1. Introduction

The scenarios considered by the PANDHUB project were discussed in D2.1. Four case study infections were chosen as representative of the types of threat within the scope of PANDHUB. These were described in terms of different indicators of impact in a public health context. Changing the assumptions about these indicators allow for scenarios to be built with varying impacts. For example, an overt attack that gives prior warning to health authorities will enable preparedness and earlier intervention than a covert attack where the detection of an event is cases appearing in hospital. Early intervention may be facilitated by enhanced communication between transport hub operators and public health agencies following recognition of cases appearing in hospitals. Interventions such as delivery of effective countermeasures to exposed individuals allow the reduction of the number (or severity) of cases.

From the PANDHUB description of work we are ultimately interested in passenger flows around transport hubs. Following feedback from the consortium's stakeholder liaison group (SLG) and from subject matter experts (UK Department of Transport, Transport for London, Virgin Trains) the research team realised that passenger flows around hubs were a well studied problem, either in an empirical data-driven framework (with ticketing data, wireless network connections or video capture) or using simulation tools such as CAST[2]. For PANDHUB it is critical to understand the passenger flow in the context of the four agents. There are three aspects to the passenger flow: how passengers arrive at the hub (to understand the risk of importation of disease); how passengers interact in the hub (to understand how infection may occur within the hub); where passengers go after leaving the hub (to inform the challenge of case finding and contact tracing). This last aspect is the subject of D3.4 but we address the first two aspects in this deliverable. The first is relevant only in the context of infections of pandemic potential as we note that an entry to the hub of an infected individual attempting to cause deliberate infection is going to appear atypically.

The four diseases considered are inhalational anthrax, pneumonic plague, Ebola virus disease (later referred to as Ebola) and pandemic influenza. The latter two are considered illustrative of the challenges posed by pandemic disease and so it is natural to consider the importation risk of a case and the role of the transport hub in the transmission and mitigation process. This report contains a review of mathematical models for the global disease spread which is necessary for assessing importation potential. We also provide a framework for assessing the chance of cases developing symptoms in a hub which affords an assessment of the benefits of screening within the hub. Anthrax and pneumonic plague are illustrative of deliberate release and here we consider how case numbers may be modelled.

Mathematical models can be viewed as conceptual tools used for explaining real-world systems and predicting future outcomes to increase our understanding of a system (Keeling and Rohani 2008). Models come in many different forms, utilising different branches of mathematics, and often different techniques can be used to explain the same scenario. This is a result of no single model being "correct"; all models have a trade-off between complexity and accuracy so it is important to assess what approach is most appropriate for each individual situation (Keeling and Rohani 2008). Evaluation of a model comes down to a subjective measure of usefulness (Keeling and Rohani 2008) and deciding which mathematical methods to use to model a situation an important decision for the modeller

[2] http://www.airport-consultants.com/cast-simulation

(Fowkes and Mahony 1994). A good model should be suited to its purpose (as simple as possible but no simpler) and parameterisable by available data (Keeling and Rohani 2008).

In this report we describe a modelling framework for simulating passenger infection from events involving these diseases, both pandemic and deliberate release. We consider the role of movement of users of the hub environment, and the contacts between these users. Section 2 presents the literature review and simulation modelling developed for pandemic infections, whilst section 3 presents the framework for simulating deliberate releases and their effects. In section 4 we summarise the means of analysing passenger flows around the hub and then in section 5 summarise modelling choices and make recommendations for later deliverables.

Contingencies to mitigate the scenarios, such as deployment of medical countermeasures), will be analysed in D3.5, along with further analysis of the benefits of entry screening. Outputs of the modelling discussed below will appear in D3.4.

Due to the complexity of the modelling developed, the mathematical notation is reused for different parameters and variables throughout the document but is applied consistently within each section. Some sections are mathematically technical but it is important that this deliverable serves as a foundation for further work.

2. Pandemic Potential Diseases

2.1 Motivation

In order to assess mitigation strategies applied in a transport hub later in the project we must first understand how cases of a pandemic disease appear in the hub. For person-to-person transmissible diseases to spread widely, the movement of infected individuals is required. International movement is aided by the worldwide air and rail passenger networks and many infected individuals will enter a country through a transport hub. Whilst the PANDHUB project focuses on European countries, a pandemic will be global and thus disease dynamics outside of Europe cannot be ignored. Indeed, the increase in number of cases from the spread of disease outside Europe is a key contributor to an increasing likelihood of importation into Europe.

Severely reducing international travel will not prevent the spread of pandemic influenza (Ferguson, Cummings et al. 2006) so we assume that if a novel strain does start spreading internationally through person-to-person transmission then eventual importation to a particular well-connected country is inevitable. There are three key questions we address in this section of the report:

- What models exist to understand the international spread of disease?
- Given that a case travels, what is the chance that that case will develop symptoms, before, during or after transit through the transport?
- If a case is infectious in a hub what models may be used to predict the number of infections arising from that contact?

Those cases which develop symptoms before or during time spent at the transport hubs may potentially be observed by authorities there.

One concept useful to address the first and last questions is the force of infection of a disease in each country. The force of infection if defined as "the *per capita* rate at which susceptible individuals contract the infection" (Keeling and Rohani 2008). The force of infection gives the probability that an individual will become infected within a certain time and consequently the growth rate of the infection. These are of interest as they may affect public health planning, such as whether to initiate a screening process. More broadly, an individual's perception of personal risk may vary with the force of infection. With a higher perception of risk, behaviour may change during a pandemic. For example, specific transport hubs or wider regions may be avoided by those viewing the visit as high risk. With an influenza pandemic, the expected number of sick employees may have serious implications for businesses and ensuring that services can keep running, including at transport hubs.

The force of infection informs of the value in investing in mitigation strategies within the hub. For instance, with pandemic influenza the rapidly increasing force of infection at the start of the pandemic means that local community transmission will quickly take over from the importation of cases as the predominant cause of new cases in that community and mitigation strategies focussed on incoming international travellers will be effective over a shorter timescale than for a disease with a more prolonged build-up.

To assess the need to create a bespoke model for PANDHUB we performed a scoping review of existing literature to assess whether a suitable model already exists. We sought to identify global disease spread models and for each determine: the model type; the input data used to parameterise the model; the validation data used to assess model accuracy and judge quality. A model could serve the needs of the PANDHUB project if it provides insight into the force of infection on European countries and the transport hubs within them and datasets exist with which to parameterise and validate the model.

For our second question, considering the probability that an international traveller develops symptoms in, or just after, transit, one needs to develop the concept of competing timeframes. A greater transit time increases the probability of symptom onset during the journey and this probability further depends on the disease natural history and when the journey occurs relative to the time of infection.

We continue in this section by describing the derivation of disease transmission models as might be applicable to a transport hub setting, then consider the scoping review of global spread of infection (and so consider the means of modelling the arrival of cases into a hub) and then consider the modelling needed to assess the probability of those cases being detected in a hub.

2.2 Disease Transmission Modelling

The most common and simple way of modelling disease transmission is with a compartmental model. In such models, the population of individuals can be partitioned into different disease or demographic states. Many variations exist but one of the most well-known is an SIR model where the compartments are: those susceptible to catching the disease (denoted S); those currently infected and infectious (I); those removed (R) (recovered and immunologically protected from reinfection or dead). Interactions between a susceptible and an infected individual can lead to the susceptible individual becoming infected. Infected individuals are removed at some assumed rate.

Such models are conceptually intuitive and appear mathematically simple, but as seen below, the mathematical equations are nonlinear and cannot be easily solved, if at all. For a detailed derivation of the SIR model the reader is referred to, for instance, (Keeling and Rohani 2008),

In essence the factors which make a contact sufficient for disease transmission will vary depending on the disease and the form and intensity of the contact, these can be combined as a single average rate parameterβ. In this model formulation, the chance that a susceptible individual comes into contact with an infected individual in such a way to become infected is proportional to $\beta I/N$ and this expression is the force of infection for the model.. Finally we assume that infected individuals recover after an average time length of $1/\gamma$ so the rate of recovery is γ.

This model can be formulated as a system of three differential equations which describes the size of the population of each state (S, I and R) over time. These equations contain the nonlinear term $\beta SI/N$. Analysis of the SIR model can show that the so called reproduction number, $R_0 = \beta/\gamma$, the number of secondary cases created following importation to an otherwise completely susceptible population, is a threshold parameter. If $R_0 > 1$ an importation of disease will lead to an epidemic whilst if $R_0 < 1$ the disease will eventually fade out which is useful measure when considering control options. For a more in-depth description of compartmental models see, for instance, (Hethcote 2000) (Keeling and Rohani 2008).

One assumption behind the SIR example model is homogenous mixing which means that all contacts appear with equal probability. This is a major simplification as in reality people will have more contact with, for example, those in the same household or workplace environment. In some contexts and scenarios this assumption may have little impact on the predicted outcomes, but in others it may significantly reduce the model accuracy.

Two common types of modelling approaches that add structure concerning disease spread between individuals are population-level models (including the SIR compartmental model discussed above) and agent-based models. Population-level models have been developed which do not assume homogeneous mixing within the total population. For instance, the population may be split by age groups where intra- and inter-group contact rates may differ. Another common population structure is to include a spatial component, where individuals located within the same region have a different contact rate with individuals within that region than with those individuals in a region further away.

In an agent-based model, each individual is modelled as an autonomous agent and the overall state of the system is determined by the combined actions of all of these individual agents. This contrasts with the population-level models where individuals within the population are not modelled explicitly, but only information regarding subsets of the population, such as the proportion of the population infected at any one time.

Agent-based network models consider each node to be an individual and edges connections between individuals. Disease transmission in such models is between a susceptible node and an infected node which are connected. Alternatively, a population-level network model could be created where nodes represent transport hubs and edges the transportation connections. In such a model a specific infected individual would not be modelled, but certain nodes could contain an infected class of individuals and thus disease would spread to other nodes via transportation connections.

Person-to-person transmission is possible for three of the infections considered by PANDHUB. Pandemic influenza is often modelled as a compartmental SEIR-type system (where the E compartment is the exposed class of individuals who are in their disease latent period) with specific age structured mixing (Diekmann and Heesterbeek 2000); pneumonic plague models have a similar mathematical structure but require closer contact between individuals for infection to occur (Gani and Leach 2004); Ebola has a more complicated natural history than the other infections and transmission is possible after death (Legrand, Grais et al. 2007).

Within a hub, contact is likely to be brief and different from that expected for the wider community. The force of infection quoted above is a mathematical simplification and the transmission rate parameter β can be recast in terms of the number of contacts made and probability of infection occurring at each contact. We consider an infectious individual making n contacts per time interval. Of these contacts some fraction S/N is with susceptible individuals and if we define p as the representative probability of successful disease transmission following contact (and so does not change at each contact event meaning that the inherent infectiousness of the case does not change and the strength of contact made does not change) then T, the number of transmission events per time interval, is modelled by a binomial distribution $T \sim B(\frac{pS}{N}, n)$.

If we assume that the pandemic is in its early stages and the chance of repeated contacts within the hub is small, so that $S = N$ (or define n as the number of *unique* contacts and subsume repeated contacts into p) and we assume that the probability of transmission changes at each contact (to reflect different types and duration of contact) then $T = \sum_{i=1}^{n} \hat{B}(p_i)$, where $\hat{B}$ is a Bernoulli distribution (and so returns either a 0 or 1 for each evaluation i).

If the change in transmission probability is random and can be modelled by a beta distribution then the number of transmission events per time interval can be modelled by a beta-binomial distribution. The probability of observing x transmission events given the n contacts is then given by:

$$P(x;\ \alpha_1, \alpha_2, n) = \frac{b(x + \alpha_1, n - x + \alpha_2)}{b(\alpha_1, \alpha_2)} \binom{n}{x} = \frac{\Gamma(x + \alpha_1)\Gamma(n - x + \alpha_2)\Gamma(\alpha_1 + \alpha_2)\Gamma(n + 1)}{\Gamma(n + \alpha_1 + \alpha_2)\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(x + 1)\Gamma(n - x + 1)}$$

where $b(\alpha_1, \alpha_2)$ is the beta function and $\Gamma(x)$ is the gamma function. Further we can consider the probability of no further infection arising from the case, whilst in the hub,

$$P_0 = P(0; \alpha_1, \alpha_2, n) = \frac{b(\alpha_1, n + \alpha_2)}{b(\alpha_1, \alpha_2)} = \frac{\Gamma(n + \alpha_2)\Gamma(\alpha_1 + \alpha_2)}{\Gamma(n + \alpha_1 + \alpha_2)\Gamma(\alpha_2)}$$

whilst the mean r (the expected number of cases arising from each infected case in the hub setting) of such a distribution is $r = \frac{\alpha_1}{\alpha_1 + \alpha_2} n$. Given that $\frac{\alpha_1}{\alpha_1 + \alpha_2}$ is the mean of the beta distribution affecting the probability infection this means that on average we would expect the same result as from a more traditional Binomial model, though critically with this structure the model allows for greater variability around that average value.

Such a model can give a simple illustrative assessment of the transmission risk for the first few generations of the disease with small numbers of importations. The model formulation implicitly incorporates the movement and interaction of infectious individuals in the hub

provided changes in contact type are random as predicted by the beta distribution. This is conceptually similar to the application of a Galton-Watson branching process to epidemic spread (House 2014). This is only valid though for the flow and interaction within the hub setting and not necessarily the wider community.

Defining the notation $\omega = \alpha_1 + \alpha_2$ gives $\alpha_1 = r\,\omega/n$ then substituting this into the equation above gives

$$P_0 = \frac{\Gamma(n + \left(1 - \frac{r}{n}\right)\omega)\Gamma(\omega)}{\Gamma(n + \omega)\Gamma(\left(1 - \frac{r}{n}\right)\omega)},$$

which is illustrated in figure 1 below for different values of ω and number of contacts. The parameter ω can be viewed as a controlling measure of the variance arising from the beta-binomial distribution given that the variance is defined to be:

$$\sigma^2 = \frac{\alpha_1\alpha_2 n}{(\alpha_1 + \alpha_2)^2}\left(\frac{\alpha_1 + \alpha_2 + n}{\alpha_1 + \alpha_2 + 1}\right) = r\left(1 - \frac{r}{n}\right)\left(\frac{\omega + n}{\omega + 1}\right)$$

The mean r may be thought of as a scaling of the basic reproduction number of the disease R_0 such that $r = T_D R_0/T_I$ where T_D is the time spent in the hub by the infectious case and $T_I$is the total duration of infectiousness for the case, critically, assuming for illustration at this stage that infectiousness remains constant throughout the infectious period. Consider then a 2 hour dwelling time in the hub and a 48 hour infectious period for a disease with a reproduction number of 2 (illustrative of pandemic influenza) then $r = 1/12$.

In Figure 1 we have shown P_0 for five different numbers of contacts between the infectious case and other individuals (one contact, black dot; two contacts, red line and points; five contacts, green; ten contacts, dark blue; and twenty contacts, pale blue) for this value of r. When $n = 1$ the $P_0 = 1 - r = \frac{11}{12}$ and so does not change with ω.

The coloured filled dot on each line is a limiting case when ω (and thus α_1 and α_2) is very small. The filled square is the point at which $\alpha_2 = 1$ (between dot and square the underlying distribution exhibits 'bimodal' behaviour, as might be the case if the infectious person is having many close contacts and many transient contacts). The empty dot then is the point at which $\alpha_1 = 1$ (so between the empty dot and filled square the underlying beta distribution is unimodal but with mode at zero), and then between empty dot and line ending the distribution is unimodal with $\alpha_2 > \alpha_1 > 1$ and mode = $(\alpha_1 - 1)/(\omega - 2)$. Notice also that for all numbers of contact that as $\omega \to \infty$ then $P_0 = \left(1 - \frac{r}{n}\right)^n$ which would be expected from the simpler binomial model.

A more exhaustive discussion of parameterisation of this model will form part of D3.4 but this provides a framework for assessing the number of infection arising within a transport hub infections from each infectious case within it.

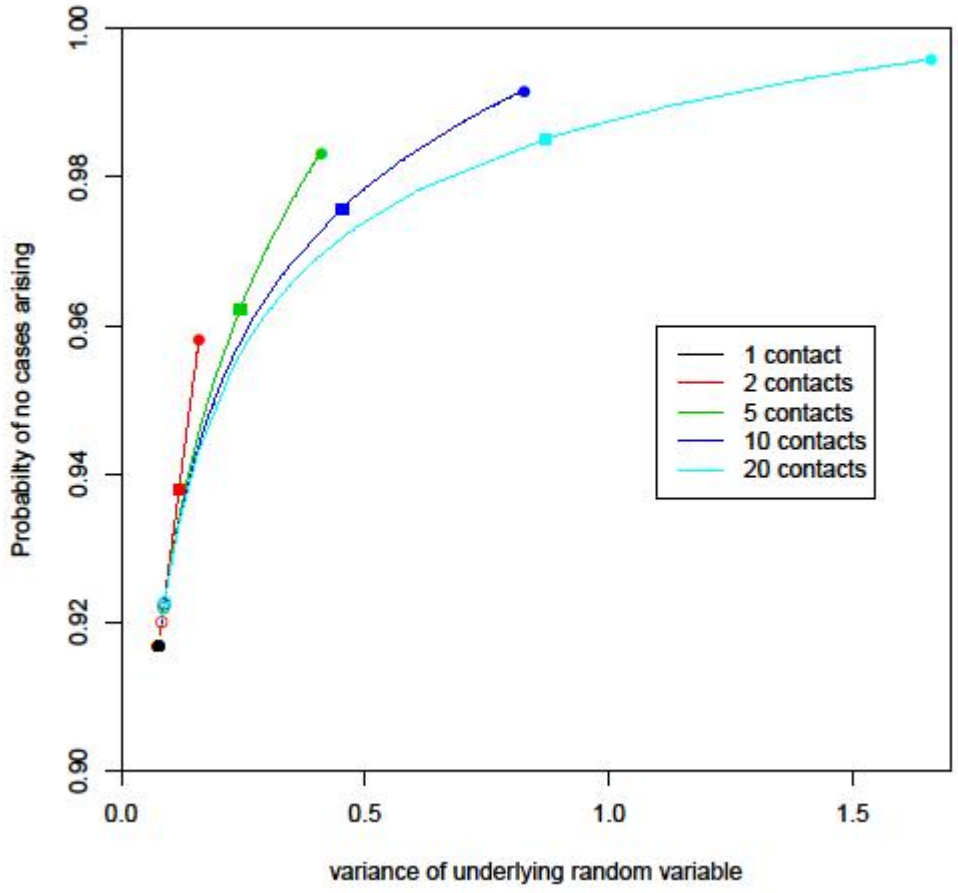


Figure 1: Figure showing the probability of no cases arising from an ill case (P_0, assuming 2 hour dwelling in the hub, 48 hour infectious period and a reproduction number of 2) for different numbers of contacts during visit.

2.3 Scoping Study - Methods

Models should be created with a specific purpose in mind, that is, there should be a specific question or set of questions that the model seeks to answer. Thus validity of a model is determined with respect to the specific purpose (Sargent 2011). Model validation is done to assure developers that the model does represent the real-world situation to a sufficient level of accuracy (Carson 2002) (Sargent 2011). What constitutes a sufficient level of accuracy must be a judgement call made by the modeller, or the commissioner of the modelling; no validation can be 100% because all models are approximations to the real world (Carson 2002). Model validation should not be confused with model verification, which is the exercising of an apparently correct model with the purpose of identifying and fixing modelling errors. Validation is checking that a model, combined with its assumptions, provide a sufficiently accurate real-world representation.

As many global disease spread models already exist, it was posited by the authors that for the PANDHUB project an entirely new approach may not be necessary and instead there is

greater benefit in identifying a range of existing models and assessing their suitability in considering the pandemic disease scenarios. We performed a scoping study to assess the need for the development of a novel global disease spread model. This study involved conducting a systematic literature search and identifying good practice modelling methods and finding data sources with which to parameterise and validate them.

2.3.1 Literature Searches

We are interested in pandemic potential diseases, particularly pandemic influenza, and the role of the global transportation network in the spread of such diseases. The search terms (which were developed by two individuals) have been chosen to reflect this and are listed in Table 1. Within the model terms, we included the known model types of: metapopulation, agent-based, and network, but also the general terms 'mathematical' and 'simulation' to capture any other modelling methods. Similarly we searched explicitly for influenza which is the classical example of a pandemic disease. However we also included a more general global disease spread term "disease spread", along with terms 'global' and 'pandemic' to pick up any other models. Finally we included terms (with wildcards) to capture the movement of individuals.[3]

The literature searches yielded 1453 papers across three databases: 332 from Embase (1974-present), 439 from PubMed (1946-present), 682 from Scopus (1970-present). Record titles and abstracts from the three databases were imported into EndNote X7.3.1 and duplicates were identified and removed with the aid of the software's `remove duplicate' function. This left 799 records after all duplicates were removed.

Search number	Search terms
Model terms	
1	Mathematical
2	Metapopulation OR meta-population
3	Agent-based
4	Simulation
5	Network
6	#1 OR #2 OR #3 OR #4 OR #5
Disease terms	
7	Disease spread
8	Influenza OR flu
9	#7 OR #8
Disease spread terms	
10	Global
11	Pandemic
12	#10 OR #11
Movement terms	
13	Travel*
14	Import*
15	Transport*

[3] We wish to acknowledge the PHE Knowledge and Library Service for their help and advice regarding scoping study methods.

16	#13 OR #14 OR #15
Combining terms for final search	
17	#6 AND #9 AND #12 AND #16

Table 1: Database search terms used for the literature review.

2.3.2 Record Screening

All 799 records were independently screened by two individuals on title and abstract and any discrepancies were resolved. This process yielded 75 articles to be screened on full text, a task which was split between these two individuals. The inclusion/exclusion criteria for the initial sift and the full text screening are listed in Table 2. After this process 57 records remained. The reference lists of all articles were screened (title and abstract) by a single individual, yielding a further 24 articles to be screened on full text. Of these, 22 articles were included in the review. The reference lists of these 22 were also screened. In total, 79 articles remained to be reviewed. This process is represented in the flow diagram in the supplementary information.

Inclusion Criteria
Global spread of a human to human infectious disease
Specific models for far-reaching outbreaks of ILIs
Use of appropriate datasets
Exclusion Criteria
Abstract or full text not available in English or French
Articles not containing mathematical models: review papers, empirical studies, emergency response, microbiological studies, or disease surveillance.
Disease type (not influenza or other potential pandemic disease)
Vector or non-human host
Focus on treatment or vaccination
Not relevant models: in host, not disease-specific, not involving population-wide spread, computer science, social media/ internet/ phone modelling.

Table 2:Scoping review inclusion and exclusion criteria.

One article identified during title and abstract screening was found to be a letter to the editor detailing a correction to a paper which also flagged up in our initial searches. This is reference [67] in the supplementary information and we have included reference details for the original 2006 article and the 2009 correction. We considered the correction to be an extension to the original article thus analysed the two records as one throughout this review. When screening the references of [53] an unpublished manuscript was flagged from the title to be reviewed on full text. It was not possible to find any further records on, or a copy of, this manuscript[4] so we were unable to continue the process with full text screening.

At full text screening we rejected non-human diseases. Despite this, we have included record [72] which considers influenza transmission between birds and humans and seeks to describe the seeding of a novel pandemic-potential strain where the virus jumps species. On

[4] L A Belova, J W Donovan, P E M Fine, D W Fraser, M B Gregg, L A Rvachev, V A Shashkov, and V I Vasilyeva Experiment on pandemic process modeling (Part 1 – influenza), unpublished manuscript (in Russian and English), 1983.

reading the full article it was decided that it more strongly fit the inclusion criteria as: it considers a pandemic potential disease; it models person-to-person transmission; and it models the dynamics between different spatial regions.

2.3.3 Information Extraction

The information extraction was shared between two individuals. The themes of interest are: the study goal (noted for context); model type; model input data; model validation data. Records included mathematical models for the spread of ILIs (influenza like illnesses); mathematical models for the spread of human infectious diseases across the globe; mathematical models for the spread of human infectious diseases across localised regions, but which could be reparameterised to become global models; human infectious disease spread papers which refer to population movement data. Records were classified as one of four broad categories: metapopulation model; individual-based model; data analysis; or probabilistic model.

Model input data could be broken down thematically into: epidemiological data; population data; and travel data. Epidemiological data concerns model parameters which describe the disease, such as the average length of infection. Population data relates to difference within the total modelled population, for example splitting the population into different age brackets or determining how many individuals live in a particular region. Travel data consists of information on travel patterns of individuals, either commuting or long-distance travel. Validation data should be from a source independent to all input data sources so that model outputs can be compared against it.

2.4 Scoping Study - Results

In this section we consolidate the information extracted from the 79 records. We are interested in identifying modelling techniques and datasets, including how many records have used particular methods, rather than conducting a critical appraisal of each record. Thus we have not included a discussion of each record here. However detailed descriptions of the model type and datasets used in each record can be found in the supplementary information. In some instances we were not satisfied with the claims of some articles regarding model type or data sources. Where we were not satisfied with regard to model type, we used our expert judgement to define the model as one of our four categories. In terms of data, we found that some records would state that input data had come from a source and yet provided no detailed reference. As we could not identify these data sources we have classified them as “none” along with records where no mention of any input data was mentioned. Similarly, regarding validation data, we classified some records as “none” if the source of the validation data was either unclear or not clearly independent of the model input data. Instances where the model type or data sources were not obvious are listed as “unclear” in the supplementary information.

2.4.1 Identified Modelling Approaches

The 79 records have been uniquely categorised in terms of model type (see Table 3). Both agent-based and population-wide models cover a wide range of different modelling methods, therefore a breakdown detailing the more commonly known mathematical approaches has been included. In some records there is no explicit mention of population modelling so we have grouped these as ‘other’. Those records where the modelling technique involves analysing datasets, rather than using them to parameterise a model, have been grouped

together as ‘data analysis’. Probabilistic models are those which incorporate probability distributions into the model or seek to provide a probability as the solution.

Model Classification		
Model Type	Number of articles	Breakdown
Agent-based	8	*FluTE = 1* *Network = 2* *Other = 5*
Population-wide	57	*Cellular Automata = 1* *Gravity = 5* *Metapopulation = 45* *Metapopulation GLEaM = 5* *Network = 1*
Other	14	*Data analysis = 9* *Probabilistic = 5*

Table 3: Model classification

The majority of records (57 out of 77) are population-level models and of these 57, 50 are metapopulation models. Not all articles explicitly stated that the model type was metapopulation, however we deemed them to be so based on the characteristics of the modelling approach taken. One explicitly mentioned metapopulation model is GLEaM. This is a computational model which can be used to represent global infectious disease spread, which is based upon a metapopulation approach (MOBS Lab). The model world is split into spatial regions which are connected in a network. This represents population mobility patterns, both local commuting and global airline connections (MOBS Lab). A software tool is publically available at http://www.gleamviz.org/, as described in [73]. Similarly, the agent-based model FluTE (see [15]), has open source code, however the model considers travel within the US only rather than worldwide. However it may be possible to adapt the model to fit a global scale.

Cellular automata models involve a grid lattice made up of cells. At each discrete time step, the state of an individual cell is affected by the states of its neighbours according to a predefined mathematical rule (Wolfram Math World). Infectious disease cellular automata models could involve each cell representing an individual in a population. At each time step a cell will be either susceptible, infected, or recovered; see, for instance, (Keeling and Rohani 2008) for a more detailed description. Alternatively, in the case of [79], each cell represents one of a number of discrete spatial regions each with corresponding. At each time step, the population for a discrete region is split into susceptible, infected, and recovered individuals and the sizes of these subpopulations vary dependent upon disease dynamics both within and between spatial regions.

Gravity models can be used to represent the flow of commodities, people, or information from one region to another, allowing them to be adapted to model infectious disease spread from a source location to a destination location (Rodrigue, Comtois et al. 2013). The measure of disease spread from one region to another is proportional to the sizes of the populations and the distance between them.

2.4.2 Identified Datasets

Sources of input data (epidemiological, population, and travel) and validation data have been identified from all 79 records, with some records using multiple sources for the same data input type. The results of the record screening are presented in Tables 4-7.

The majority of epidemiological data come from existing literature, by which we mean published journal articles as opposed to distinctly published datasets, as shown in Table 4 below. However, some models are parameterised with data from disease surveillance, or information from CDC or WHO. FluNet is an influenza surveillance system, maintained by WHO, which is briefly described in D3.1, section 6.2.3.1.

Population data come from a wide range of sources. What sources are appropriate is model-dependent: the data format must match up with the model parameters. Information is commonly taken from a census for the modelled region, or (inter-)national databases or statistics as shown in Table 5 below.

The main sources of travel data, as shown in Table 6, aside from existing literature, are from national statistics or surveys, from IATA (International Air Transport Association) or from OAG. IATA is a trade association for the world's airlines and has a variety of available datasets for purchase (IATA). These include: passenger forecasts; air traffic statistics; and customisable datasets, which may provide more flexibility when designing a model to be parameterised by particular datasets. OAG is an air travel intelligence company which has a large network of air travel data, also available for purchase (OAG).

Most models (64 out of 79) do not have any validation data, Table 7. This may be because it can be difficult to find a separate dataset for pandemic or pandemic potential outbreaks than the one used to parameterise the model. This is because pandemics do not happen frequently and records and systems for recording outbreaks may have changed so much over time that those datasets which do exist may not be appropriate for comparison. When listing articles with validation data we sought evidence of a dataset independent from the input dataset. In some instances authors claimed to have validated their model from a dataset which was used to parameterise the model. This may actually be verification rather than validation and so, without more detailed information, we were not able to confirm that such models had indeed been validated. If we were unsure then we listed the model as having no validation data.

Epidemiological Data	
Source	Number of citing articles
Centers for Disease Control and Prevention (CDC)	4
Census	1
Existing literature	48
National reports/ statistics	8
Personal communications	1
Surveillance data	3
World Health organization (WHO) *including FluNet*	6
None	20

Table 4: Epidemiological data classification

Population Data	
Source	Number of citing articles
Census	22
Center for International Earth Science Information	1
Central intelligence Agency	1
Eurostat	3
Existing literature	16
International population Database	1
Landscan	1
National Geographic Information Services	1
National Statistics	3
Organisation datasets	3
Polymod	1
Population database for specific country	5
Socioeconomic Data and Applications Center (SEDAC)	2
Surveys	2
United Nations database	3
World Bank population estimates	2
World Gazetteer	1
None	38

Table 5: Population data classification

Travel Data	
Source	Number of citing articles
Airport/carrier-specific statistics/surveys	7
Census	4
Company statistics/data	2
Data In. Information out. (DIIO)	1
Eurostat	4
Existing literature	18
IATA (International Air Transport Association)	16
International Civil Aviation Organisation	1
National Statistics/surveys	21
OAG	15
University of Manitoba Transport Information Group	1
World Tourism Organisation	1
None	17

Table 6: Travel data classification

Validation Data	
Source	Number of citing articles
CDC data	1
Outbreak data	11
Surveillance data	1
WHO morbidity data	2
None	64

Table 7: Validation data classification

2.5 Scoping Study - Discussion

We found that the majority of records did not have any validation data and that population-level metapopulation models were the most common model type. Nine metapopulation records have validation data: [5], [19], [27], [42], [43], [45], [53], [54], and [71]. Of these, [27], [53] and [54] are all studies from prior to 1993 so, whilst mathematical techniques are unlikely to have dated, the computational implementation and available datasets are likely to have evolved to such an extent that these studies are of less interest. Two records are concerned with travel within the US only ([42], [43]) and one ([45]) has only existing literature for input data sources. This leaves three records which use a metapopulation approach to model global disease spread and also have a validation data set. This highlights that models do currently exist and can be parameterised and validated, that can simulate the international spread of infection which was the one of our three questions for pandemic infections thus we do not consider it necessary to construct a new global spread model for the PANDHUB project.

2.6 Entry Screening

2.6.1 Motivation

In 2014 many European governments opted to begin a screening programme for Ebola at international transportation hubs which had links with West Africa. The aim was to avoid the spread of Ebola into unaffected countries by preventing unwell individuals from travelling and by preparing for cases of Ebola which develop once an individual has already passed into the uninfected country. Part of this preparedness involved identifying passengers at risk of becoming symptomatic and providing them with appropriate information and advice regarding access to health care. Similar interventions have been discussed for pandemic influenza and modelling has shown little benefit from the adoption of airport screening *per se* (Ferguson, Cummings et al. 2005) though Ebola and influenza are very different diseases epidemiologically .

We are interested in how a transport hub could function as part of a mitigation strategy for diseases with pandemic potential, in particular influenza and Ebola. As Public Health England was responsible for the screening at UK ports of entry, much of our knowledge is based on the procedures adopted within the UK. Such processes, however, are likely to be similar to those in other European countries. One question posed by the UK government during the 2014 West African Ebola outbreak was “what is the chance of an infected individual entering the UK?" Addressing this question is within the scope of the PANDHUB project as this information would help governments accurately assess the risk of disease

importation and thus allow them to make appropriate cost-benefit analyses regarding mitigation strategies.

2.6.2 Ebola

As presented in D2.1, Ebola is a viral haemorrhagic fever for which, at the time of the 2014 outbreak, there was no vaccine or cure. The 2014 outbreak of *Zaire ebolavirus* is currently the largest recorded outbreak of Ebola (Baize, Pannetier et al. 2014) (Briand, Bertherat et al. 2014). Ebola is person to person transmissible through a susceptible individual coming into contact with body fluids of an infected individual, either directly or indirectly via fomites (Osterholm, Moore et al. 2015). The incubation period (that is, the time from exposure to the pathogen to symptom onset) is usually 2-21 days (Public Health England 2015). Initial symptoms include: fever; headache; joint and muscle pain; sore throat; and intense muscle weakness (Public Health England 2015). Subsequent symptoms include: diarrhoea; vomiting; a rash; stomach pain; impaired liver or kidney function; and bleeding either internally or from eyes, ears, nose or mouth (Public Health England 2015). An individual infected with Ebola virus is not infectious prior to becoming symptomatic (Public Health England 2015). This is illustrated in Figure 2. Findings from the 2014 West African outbreak indicate that Ebola can be sexually transmitted from recovered individuals to susceptible individuals as the virus has been found to remain in genital fluid after symptoms have subsided (Fischer and Wohl 2016).

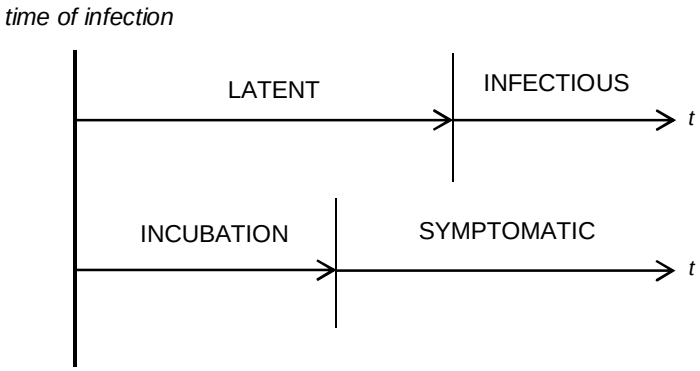


Figure 2: Schematic of the infection and symptom timelines. The latent period for Ebola is thought to be greater than or equal to the incubation period. Thus an infected individual must be symptomatic in order to be infectious so up until symptom onset they cannot pass on the disease.

2.6.3 The Ebola Screening Process

We describe the UK screening process from the 2014 Ebola outbreak as it was known to Public Health England at the time of the outbreak. Screening happens at the departure hub and the arrival hub. If a journey has a break in it, screening may also occur at the

changeover hub. For the Ebola outbreak, hubs were primarily airports, although screening did take place at St Pancras railway station in London, which is the UK Eurostar terminal.

Passengers arriving into the UK from three countries experiencing the Ebola epidemic (Sierra Leone, Liberia and Guinea) were screened by healthcare professionals for symptoms or epidemiological risk factors for Ebola infection. UK Border Force officials flagged these passengers who needed to be screened and then directed them to a designated screening area.

The screening process is informed by two parts: a questionnaire (known as a *Health Assessment form*) and a temperature measurement. Screeners would meet with passengers and follow an algorithm, similar to the one shown in the supplementary information. According to the algorithm, those passengers with a maintained raised temperature, or raised temperature and other symptoms from the list, should be immediately isolated. Passengers without a raised temperature and without other symptoms or epidemiological risk factors should be provided with information, reassured, and discharged. Those without significant signs or symptoms and without a raised temperature, but categorised as at elevated risk through hazardous activity, should be discharged, but with follow-up from a designated health protection team. A passenger without a raised temperature but with at least one of the listed significant signs or symptoms would have a discussion with the duty senior clinician (who may not be based at the screening area or terminal). Depending on the results of this consultation the patient would either be isolated, or discharged with follow-up arranged.

2.6.4 Mathematical formulation

Fundamentally we want to know the probability that an infectious individual arrives into a transport hub. To answer this question we first transform it into a mathematical problem which we can then analyse. Based on our knowledge of the natural history of Ebola (section 2.6.2) and the way in which the Ebola screening works (section 2.6.3), we can begin to tackle the problem.

First, we acknowledge that determining when an individual becomes infectious is not simple: infectiousness may vary as the disease progresses and also may vary from individual to individual. For Ebola, it is thought that an individual must be symptomatic before becoming infectious (Public Health England 2015) (see Figure 2) so we rephrase our question as: “what is the chance of a symptomatic individual arriving into the UK?” Ignoring the small risk of sexual transmission the probability of a symptomatic individual arriving into the UK then provides an upper bound for the probability of an infectious individual entering the UK. We split the population of individuals using the transport hub into symptomatic and not symptomatic.

To be determined as symptomatic an individual must meet certain criteria which are defined in the screening process. From the screening algorithm, we see that to register as infected a passenger must have a raised temperature (>37.5 °C) from two readings or a single raised temperature reading plus other significant signs or symptoms.

Although Ebola screening takes place at the departure hub, arrival hub, and at any interchange hubs we do not consider journeys with transfer stops and consider departure screening to be the screening event immediately prior to the UK arrival screening. We are also not concerned about secondary cases which result from transmission during a flight: the

minimum incubation period is 2 days (Public Health England 2015) so any secondary cases would not be symptomatic by arrival screening, thus would be picked up by alternative methods outside of the hub.

Individuals are screened prior to departure and any individual found to be symptomatic at departure screening will not be allowed to fly. Infected individuals who are not symptomatic (that is, are in the incubation phase) are able to pass departure screening. Individuals in the incubation period at arrival screening will not be detected and will leave the transport hub. These individuals will have been provided information on what to do if they develop symptoms, as stated in the algorithm presented in the supplementary information. So we wish to identify those individuals who develop symptoms during their flight. Thus our problem then becomes “what is the probability that an infected yet not symptomatic individual becomes symptomatic during a flight to the UK?"

We assume that the probability that a traveller is infected can be determined from existing literature, expert opinion, or from surveillance reports in country of origin (noting that this may be a biased estimate given that the propensity for an individual to travel is subject to socio-economic conditions in that country and other social drivers). This probability may be different for different groups of individuals. For instance, if the travelling population has a high proportion of healthcare workers who have been deployed to work on Ebola in West Africa compared with individuals who have travelled through the area for other reasons then the chance of them being infected may be higher than for other travellers.

We define two timelines: one related to infection (the time from exposure to symptom onset) and the other related to travel (the time from exposure to arrival, with a distinct time to departure occurring prior to arrival). The origin for the infection timeline is the incubation period which begins at the time of exposure to the infectious agent (here Ebola virus). If the incubation period is shorter than the time until departure then the individual will be symptomatic at departure screening. If the incubation period is longer than the time until departure screening then the person will board the flight. We assume that the time from exposure to departure can be described by a probability distribution, and similarly for the transit time (departure to arrival). We then want to know when symptom onset occurs relative to the transit time with the three possible cases shown in Figure 3. We want to know whether symptom onset occurs during the period of travel (so between departure and arrival). This is shown by case B in Figure 3.

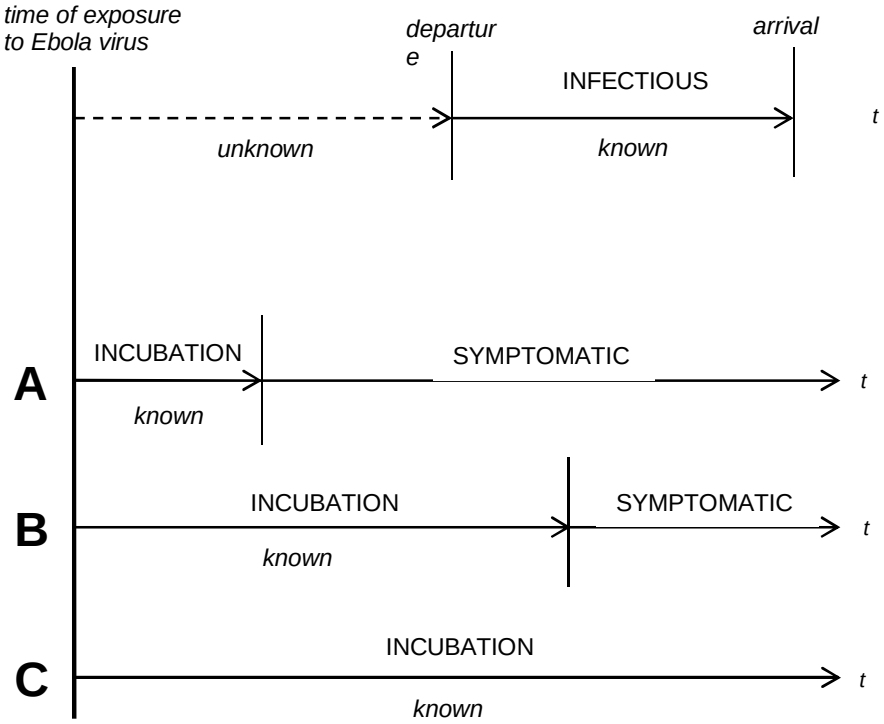


Figure 3: Schematic of the travel and infection timelines. We assume that we know both the transit time and the incubation period: these may be probability distributions as opposed to a single value. We do not know the time from exposure to Ebola virus to departure screening. A) The incubation period ends prior to departure and the traveller will be symptomatic at departure screening and denied boarding. B) Symptoms begin during transit. C) The incubation period is longer than the time from exposure to arrival screening. We are interested in case B.

We introduce the independent continuous random variables: $X \sim D_1$, the time from exposure to departure screening; $T \sim D_2$, the time from departure screening to arrival screening (travel time); $I \sim D_3$, the time from exposure to symptom onset (incubation period).

We need to know the time from exposure to arrival screening which is the time from exposure to departure, X, plus the transit time, T. As we have the addition of probability distributions we obtain a new distribution, $Y = X + T$, giving the time from exposure to arrival. We know the probability distribution function (pdf) of Y is the convolution of the pdfs of X and T (see supplementary information for detail). The probability P of an infected individual becoming symptomatic during transit is thus $P(X < I < Y)$. We now introduce the variables

$$W = X - I,$$
$$Z = Y - I$$

so that we can write

$$\begin{aligned} P(X < I < Y) &= P(W < 0 < Z) \\ &= P(Z > 0) - P(W > 0) \\ &= (1 - P(Z < 0)) - (1 - P(W > 0)) \\ &= F_W(0) - F_Z(0), \end{aligned}$$

where $F$is the cumulative density function (cdf) of the corresponding random variable.

Now we have the probability that an infected individual becomes symptomatic during a flight, given that they are infected. However, we have further real-world information that has not been utilised; namely, that individuals who become symptomatic prior to departure will be prevented from travelling. Thus what we wish to know is actually the probability of becoming symptomatic given infected as a proportion of those that pass through departure screening. Consequently, the probability that an individual becomes symptomatic during a flight given that they are infected and exit screening is perfect is

$$\frac{P(X < I < Y)}{P(X < I)} = \frac{F_W(0) - F_Z(0)}{F_W(0)}.$$

Further details of calculating this probability are given in the supplementary information.

2.6.5 Summary

We are aware that a passenger becoming symptomatic on a flight may lead to operational changes, such as diverting the aeroplane to a closer airport. This would affect our answer as the chance of a symptomatic individual arriving into the UK would be affected by the point at which they became symptomatic mid-flight and the flight path as these factors affect which is the closest airport. We could also develop the model to account for incorrect conclusions from the screening process (that is by introducing a chance that a symptomatic individual could pass departure screening or that a non-symptomatic individual could fail arrival screening). This may occur if an individual was attempting to hide their symptoms, for instance by taking medication to lower his or her temperature.

The model formulated above will be considered in D3.4 with worked simulations of the examples shown mathematically here and extended in D3.5. For a case that develops infectiousness whilst in a transport hub, the model given in section 2.2 may be further developed to consider second generation cases arising from this case in the transport environment. We present such assessment in D3.4.

3. Bioterrorist Diseases

3.1 Introduction

When assessing a scenario involving deliberate release of biological material one needs to consider the mechanism of delivery, the airborne dissemination and the effect of the infectious material on humans in vicinity. In this section we develop a modelling framework for this but first consider the natural transmission potential of the two case study diseases (pneumonic plague and anthrax). We develop a model for aerosol dispersion that may be applied to a generic environment and then consider the effect on nearby humans, the stages involved in such modelling are shown in Figure 4.

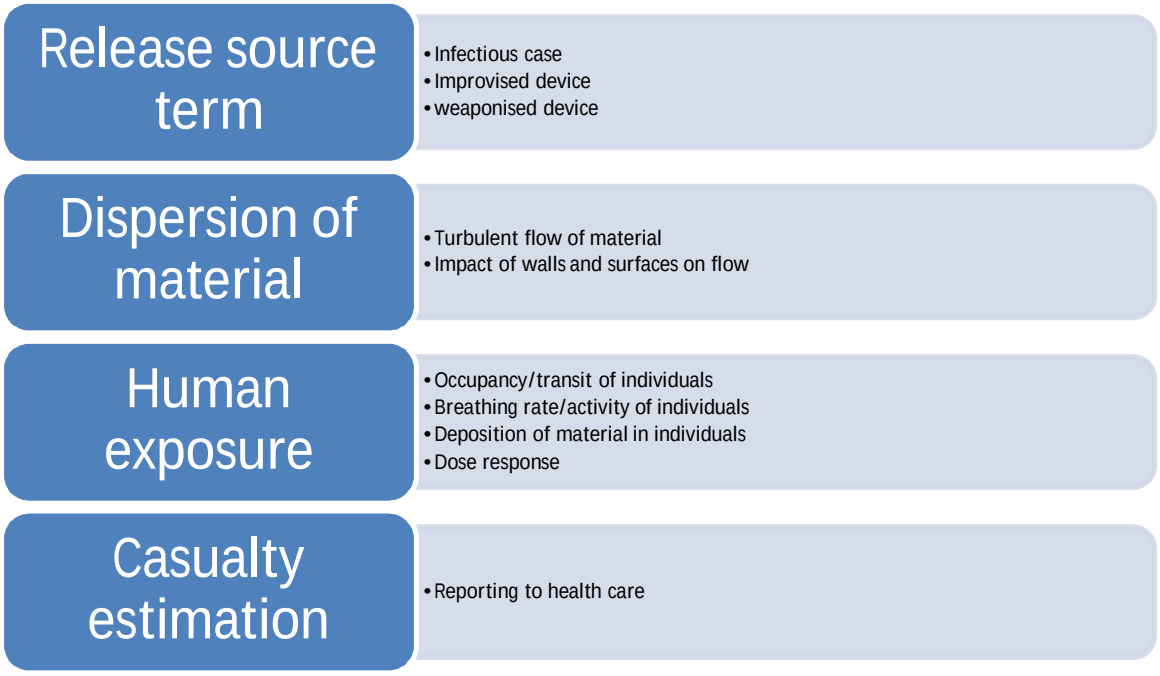


Figure 4: Schematic showing the modelling framework considered for infections of individuals due to deliberate pathogen release within a transport hub.

3.2 Natural reservoir of infection

The pandemic diseases considered in section 2 have the potential for zoonotic transmission (with for example avian reservoirs for flu and bat reservoirs for Ebola) and similarly the deliberate release agents chosen for PANDHUB can arise naturally too.

3.2.1 Anthrax

A disease caused by the spore-forming *Bacillus anthracis* bacterium, anthrax is another zoonosis from antiquity. Predominantly a disease of herbivores, humans can become infected directly or indirectly from contact with infected animals or animal products (World Health Organization 2008). Anthrax is not considered to be a communicable disease; however, person-to-person transmission (involving the cutaneous form) has been recorded but is extremely rare (Lalitha, Anandi et al. 1988) (Quinn and Turnbull 1998). In animals,

anthrax remains endemic and common in some Mediterranean countries, some countries of central and South America and central Asia, some sub-Saharan African countries, and western China. Elsewhere, sporadic cases and outbreaks in people continue to occur (World Health Organization 2008) where there is the disease in animal populations or where humans are exposed to contaminated animal products.

Humans are moderately resistant to infection and there is a reasonable evidence that prolonged exposure to widespread low-level environmental contamination (occupationally and non-occupationally) poses a negligible risk (World Health Organization 2008) (Bennett, Bennett et al. in preparation) (Doolan, Freilich et al. 2007) (Norman, Ray Jr et al. 1960) (Wattiau, Govaerts et al. 2009). There is also evidence that even exposure to significant amounts of aerosolised *Bacillus anthracis* spores results in very few cases of infection (Brachman, Gold et al. 1962) (Dahlgren, Buchanan et al. 1960). There are three main forms of human anthrax: cutaneous, inhalational (the most severe form) and gastrointestinal, the form depending on the route of infection. Only inhalational anthrax is considered within this project and the reader is guided to D2.3 for more detail on the transmission mode for this form.

B. anthracis exists in the environment as spores. These spores are very resistant to extremes of heat, cold, pH, desiccation, chemicals and irradiation. In the spore state, *B. anthracis* can persist and remain viable in the soil and other environments for many decades or even hundreds of years (World Health Organization 2008). However, it has been concluded that the risk of being infected with anthrax from contaminated soils or other environmental materials is exceedingly low, and the chances of infection with inhalational or intestinal forms are next to nil (Turnbull 1996).

3.2.2 Plague

The infectious disease plague, caused by the bacterium *Yersinia pestis*, occurs on every continent except Australia in tropical and sub-tropical latitudes and warmer parts of temperate latitudes (WHO Plague Manual 1999). According to the WHO, from 1989 to 2003, there were over 38 000 cases and 2845 deaths in 25 countries (WHO Plague Fact Sheet). In 2013 there were 783 cases reported worldwide, including 126 deaths (WHO Plague Fact Sheet).

Y. pestis circulates in animal reservoirs, primarily in rodents. Plague-infected rodent fleas can transmit the disease between rodents and to other animals, including humans. Plague can be divided into ‘wild plague’, where disease exists independent of humans, and ‘domestic plague’, where disease can exist as epidemics in humans and animals alike (WHO Plague Manual 1999). Human cases/outbreaks tend to be associated with risk factors such as occupational work or hunting/camping where contact is made with infected rodents, or where domestic rodents are infected. Human plague naturally occurs when humans come into contact with infected fleas or infected animal bodies or, less commonly, through contact with an infected person. For further information refer to D2.1. Humans are extremely susceptible to plague and there have been three notable plague pandemics in the last few millennia: the Justinian plague in the sixth century AD; the ‘Black Death’ pandemic of the fourteenth century; the third at the end of the 19th and beginning of the 20th centuries.

Humans can be infected indirectly through the bite of an infected flea (the most common route) or through direct contact when handling plague-infected animals or meat. Indirect infection is considered out of scope of PANDHUB. Human-to-human transmission can occur from a bite from an infected human flea or from contact of a susceptible person with either an infected person or contaminated object (less common). To cause infection in the latter

cases, the organism has to cross into the body via mucous membranes or breaks in the skin. Disease manifests as one of three forms (bubonic, septicaemic and pneumonic), depending on the route of infection, but only the pneumonic form is relevant for PANDUB. For further information on transmission refer to D2.3, section 5.2.

Y. pestis is susceptible to environmental factors such as sunlight, high temperatures and desiccation as well as many disinfectants such as Lysol and chlorine preparations (WHO Plague Manual 1999). Such factors that potentially impact on infection control mechanisms will be discussed in D4.4.

3.2.3 Summary

Natural importation via a transport hub cannot be ruled out given both diseases are endemic is parts of the world. Anthrax is not person-to-person transmissible and so such an importation would be managed in a similar way as a deliberate attack might be given the rarity of clinical cases (i.e. the suspicion would be that the case was infected deliberately until the follow up epidemiology suggested they had visited areas with infection hazard). Despite the potential for person-to-person transmission naturally occurring pneumonic plague outbreaks have historically been self-limiting and so a case that enters a transport hub would lead to similar questions about the transmission potential as for the pandemic agents discussed above. The relevant mode of transmission given the scope of PANDHUB is documented in D2.3 for both infections. Here we focus on a deliberate release of the causative agent at the hub, with symptomatic pneumonic cases appearing in the health system some days later, rather than the management of sick cases in the hub.

3.3 Aerosol dispersal

In this section we discuss the assumptions for dispersion of material arising from three mathematical approaches and discuss the merits of these to the PANDHUB work.

3.3.1 Mean field approach

An approach to modelling aerosol contamination is to assume that the air in a room is very well mixed and so a release of a number of bacteria, Q, is immediately dispersed throughout the room to some average concentration per unit volume of air, though we allow for the following assumption which will be useful later:

I. There is some external decay of concentration (for example due to gravitational settling, surface deposition, fresh air ingress or loss of viability due to UV, temperature or humidity), this is modelled with the super-parameter γ. This means that the concentration $C(t)$ changes like $\frac{dC}{dt} = -\gamma C$.

From this assumption, integrating the differential equation in a room of height H_c, width L_W and length $L_L$the average concentration would be $C = {Qe^{-\gamma t}}/{H_c L_W L_L}$and so the exposure for an individual in the room, assuming some dwelling time T_D in the room and entry time T_S , would be $D = {Qe^{-\gamma T_S}(1 - e^{-\gamma T_D})}/{(\gamma H_c L_W L_L)}$. This is similar to the formulation given in Cooper and Horowitz where $\gamma = {V}/{(H_c L_W L_L)}$ where V is the ventilation rate $[m^3/s]$ (Cooper and Horowitz 1986). This model may be extended to consider interconnected rooms in buildings with internal ventilation (Parker and Bowman 2011).

Human exposure may then be seen as a Poisson process with dose given by D. Extension of this may be allowed by assuming some aggregation of particles in space and so the dose D itself being a random variable with a gamma distribution then the inhaled dose will be a negative binomial random variable (Pratt, Bennett et al. 2016).

3.3.2 Turbulent flow: the Gaussian Puff Model

In practice the contaminant will not be immediately well mixed in the air and there will be some transient period of flow dispersion and mixing potentially on a scale similar to the dwelling time of individuals in the area. When bacteria are aerosolised they become subject to local turbulent forces (causing random movement of the air packet within which the aerosol is contained) as well as longer range advective processes (transport of the packet by the dominant airflow). There are a number of models for the atmospheric dispersion of aerosol material. It is not within the scope of PANDHUB to perform a systematic review of such techniques because we are interested in a hub setting specifically.

Stockie provides a helpful review of the assumptions made and the derivation of the classical Gaussian *plume* model and the reader is referred to this article and references therein for detail (Stockie 2011). The aim is to solve the three-dimensional advection-diffusion equation

$$\frac{\partial C}{\partial t} + \nabla.(\boldsymbol{u}C) = \nabla.(\boldsymbol{K}\,\nabla C) + S - \gamma C$$

where $C(t, \boldsymbol{x})$ is the concentration of the infectious agent in time and space. For illustration at this stage we make the further assumptions to I above:

II. A number of bacteria, Q, are released from a fixed point $\boldsymbol{x} = (x_0, y_0, H)$ where H is the height above ground level at time $t = 0$. The source term function takes the form $S = Q\delta(t)\delta(x - x_0)\delta(y - y_0)\delta(z - H)$.

III. Variation in topography is negligible so that the ground surface is taken to be the plane $z = 0$. However, we assume at this point that the domain is infinite for simplicity. Human exposure will occur at the height of inhalation (so approximately the standing or seated head height).

This equation is useful but before we can consider solutions a number of questions remain. For example, what happens to material that makes contact with surfaces (most importantly floor, ceiling and walls but also moveable items such as chairs), what form should be used for the eddy diffusion terms, what is the role of natural or mechanical ventilation, and how do people interact with the contamination (i.e. by movement)?

To answer these questions we searched the PUBMED and Scopus databases with the terms “indoor” and “dispersion model”. 75 results were returned on Scopus and 20 on PUBMED. After removing duplicates and deselecting based on title, the abstracts of 13 articles were reviewed leading to six articles being read in full of which two were relevant. Deselection at each stage was based on applications not of relevance to PANDHUB (articles were based on odour, pollution or specific (non-biological) contaminant or looked at indoor/outdoor exchange of material or at urban flows more generally) or model used (box models and CFD models were excluded because they are discussed in sections 3.3.1 and 3.3.3). A similar search with terms “indoor” and “eddy” and “diffus*” yielded 37 articles on the two databases with 10 with relevant titles given the same inclusion and exclusion criteria. On reading abstracts only one article was kept for full review. There are thus three articles that provide the basis of the modelling in this section, along with the references therein. These articles

provide potential solutions and are discussed where relevant in the following three subsections.

3.3.2.1 Strong advection, very distant or non-reflective boundary

In this situation we make the following assumptions:

IV. The air flow ingress velocity is constant and aligned with the x-axis, $\boldsymbol{u} = (U, 0, 0)$. Furthermore the wind velocity in that direction U is sufficiently large that diffusion in the x -direction is smaller than advection.

V. The turbulent eddy diffusion is a function of downwind distance only and is isotropic in the x, y plane (so $\boldsymbol{K} = (K(x), K(x), K_z(x))$).

VI. We are not interested in negative value of x and so $x \in [0, \infty)$ but that $y, z \in (-\infty, \infty)$. Furthermore we wish the concentration to be zero at the infinite boundaries so $C \to 0$ as $x, y, z \to \infty$ and $y, z \to -\infty$.

Given these assumptions in items I to VI above, we may solve the advection-diffusion equation analytically to find the classical solution

$$C = \frac{Qe^{-\gamma t}}{(4\pi t)^{3/2} K(x)\sqrt{K_z(x)}} \exp\left(-\frac{(x - Ut)^2 + y^2}{4K(x)t} - \frac{(z - H)^2}{4K_z(x)t}\right).$$

Moreover it is usual practise in the literature of outdoor releases to redefine the eddy diffusivity in terms of the standard deviation of the concentration such that $\sigma^2(x) = 2K(x)t$ and $\sigma_z^2(x) = 2K_z(x)t$. (Rasouli and Williams 1995) modelled a ground level release of gas in a room with ceiling mounted ventilation with an alternative form for $\boldsymbol{K}$. However, it is unlikely that strong advection will be present over the entire domain for an indoor release and so we do not continue with this approach.

3.3.2.2 No dominant advection, distant or non-reflective walls

In this case, the previous assumptions change such that

IV. The air flow velocity is constant and zero, $\boldsymbol{u} = (0, 0, 0)$.

V. The turbulent eddy diffusion is constant and is isotropic in all directions, $\boldsymbol{K} = (K, K, K)$.

VI. We are interested in the entire spatial domain (so walls absorb the contaminant) so $x, y, z \in (-\infty, \infty)$. Furthermore we wish the concentration to be zero at the infinite boundaries so $C \to 0$ as $x, y, z \to \pm\infty$.

Given the assumptions IV to VI above we may solve the advection-diffusion equation to find

$$C = \frac{Qe^{-\gamma t}}{(4\pi K t)^{3/2}} \exp\left(-\frac{r^2}{4Kt}\right)$$

where $r^2 = (x - x_0)^2 + (y - y_0)^2 + (z - H)^2$ (Cooper and Horowitz 1986). The exposure from time T_s for a person dwelling for time T_D is then given by the integral of C over time,

$$D_0 = \int_{T_s}^{T_s+T_D} \frac{Qe^{-\gamma t}}{(4\pi Kt)^{\frac{3}{2}}} \exp\left(-\frac{r^2}{4Kt}\right) dt$$
$$= -\frac{Qe^{-r\sqrt{\frac{\gamma}{K}}}}{8\pi rK}\left(1 + \text{erf}\left(\frac{r-2\sqrt{\gamma K}t}{2\sqrt{Kt}}\right) + e^{2r\sqrt{\frac{\gamma}{K}}}\left[\text{erf}\left(\frac{r+2\sqrt{\gamma K}t}{2\sqrt{Kt}}\right) - 1\right]\right)_{T_s}^{T_s+T_D}$$
$$= \frac{Qe^{-r\sqrt{\frac{\gamma}{K}}}}{8\pi rK}\left(\text{erf}\left(\frac{r-2\sqrt{\gamma K}T_s}{2\sqrt{KT_s}}\right) - \text{erf}\left(\frac{r-2\sqrt{\gamma K}(T_s+T_D)}{2\sqrt{K(T_s+T_D)}}\right)\right.$$
$$\left. + e^{2r\sqrt{\frac{\gamma}{K}}}\left[\text{erf}\left(\frac{r+2\sqrt{\gamma K}T_s}{2\sqrt{KT_s}}\right) - \text{erf}\left(\frac{r+2\sqrt{\gamma K}(T_s+T_D)}{2\sqrt{K(T_s+T_D)}}\right)\right]\right).$$

People may also move whilst interacting with the contaminant field and this will be considered later in more detail.

We may now consider some simplification to this equation. If $T_S = 0$ (so the individuals were present at the start of the release) and defining $w = r/\sqrt{4KT_D}$ and $v = \sqrt{\gamma T_D}$ then

$$D_0 = \frac{Q}{8\pi rK}([1 - \text{erf}(w - v)]e^{-2wv} + [1 - \text{erf}(w + v)]e^{2wv}).$$

Then if $w \ll 1$ so $T_D \gg r^2/4K$ (and so individuals dwelt in the area for sufficient time to get a the maximum dose possible for that location) then

$$D_0 \approx \frac{Q}{8\pi rK}(e^{-2wv} + e^{2wv} + \text{erf}(v)[e^{-2wv} - e^{2wv}])$$

Let us consider the assumption on T_D above. For a room 10 metres wide and 20 metres long, a person sat (so head is 1 metre high) in the middle of the room (at $(x, y, z) = (10,5,1)$) from the start of the release with release source at $(x_0, y_0, H) = (0,5,2)$ we have $r^2 = 101$ (and assuming 15 minutes' dwelling time (900 s)) and the simplification is valid provided $K \gg 0.028$. We return to the question of K in section 3.3.2.4. Moreover if $uv \ll 1$ so that $\gamma T_D \sim O(1)$ or smaller (and so the dwelling time is similar to the deposition/air exchange rate) then we recover the limiting form (Cooper and Horowitz 1986):

$$D_0 \approx \frac{Q}{4\pi rK}.$$

3.3.2.3 No dominant advection, with reflection from surfaces

If the surface is reflective and so particles bounce back off walls and floors, the dose can be modelled with an infinite number of imaginary sources being placed on the other side of the wall (Drivas, Valberg et al. 1996). Changing the boundaries to being reflective as in Drivas *et al.* leads to

$$C = \frac{Qe^{-\gamma t}}{(4\pi Kt)^{3/2}} C_x C_y C_z,$$

hereafter referred to as the Drivas model for brevity, where

$$C_j = \sum_{m=-\infty}^{\infty} \exp\left(-\frac{(j + 2mL_j - j_0)^2}{4Kt}\right) + \exp\left(-\frac{(j + 2mL_j + j_0)^2}{4Kt}\right).$$

For dummy variable j interchanged as appropriate with x, y or z. However, there is limited evidence of reflection for biological material and good evidence that resuspension (aerosols re-entering the atmosphere after settling) is surface dependent. This form was considered by (Cheng, Acevedo-Bolton et al. 2011), for a naturally ventilated, domestic room, where they simplified the Drivas model to consider only the closest surfaces.

Extending previous work, let us assume that some fraction θ of bio-aerosol is deposited on walls and other surfaces during contact. This fraction may vary by type of surface, but for model parsimony we assume for a given room this value is sufficiently similar to be assumed the same. In this case we have a modification of the Drivas model with

$$C_j = \sum_{m=-\infty}^{\infty} \theta^{2|m|} \exp\left(-\frac{(j + 2mL_j - j_0)^2}{4Kt}\right) + \theta^{2|m|+1-a} \exp\left(-\frac{(j + 2mL_j + j_0)^2}{4Kt}\right)$$

where $a = 2$ if $m \geq 0$ and zero otherwise. Thus if $\theta = 1$ we have the Drivas model and if $\theta = 0$ that proposed by (Cooper and Horowitz 1986).

Alternatively the solution may be reformulated with an absorptive boundary condition. However, contaminant will settle on the surfaces and remain there until the agent loses viability, is cleaned or is re-aerosolised. Thus surface deposition can be considered as the fraction $1 - \theta$ that does not reflect with the additional contribution on the floor of the loss of aerosol due to the gravitational component of γ.

Each of infinite sums given by C_x, C_y, C_z can be truncated depending on the size of the room and the eddy-diffusion value and the surface reflection fraction. For example to order θ^2we find

$$C_j = \exp\left(-\frac{(j - j_0)^2}{4Kt}\right) + \theta \exp\left(-\frac{(j + j_0)^2}{4Kt}\right) + \theta \exp\left(-\frac{(j - 2L_j + j_0)^2}{4Kt}\right) + \theta^2 \exp\left(-\frac{(j + 2L_j - j_0)^2}{4Kt}\right) + \theta^2 \exp\left(-\frac{(j - 2L_j - j_0)^2}{4Kt}\right) + \cdots$$

In this situation the dose received will be

$$D = \int_0^{T_D} \frac{Qe^{-\gamma t}}{(4\pi Kt)^{3/2}} C_x C_y C_z \, dt = D_0 + \frac{\theta Q}{8\pi K}\left(\sum_n \frac{[1 - erf(u_n - v)]e^{-2u_n v} + [1 - erf(u_n + v)]e^{2u_n v}}{r_n}\right) + O(\theta^2)$$

where $u_n = r_n / \sqrt{4KT_D}$ and $r_n = \sqrt{(x - x_n)^2 + (y - y_n)^2 + (z - z_n)^2}$ and $x_1 = x_0, y_1 = y_0, z_1 = -H;\ x_2 = x_0, y_2 = y_0, z_2 = 2L_z - H;\ x_3 = x_0, y_3 = -y_0, z_3 = H;\ x_4 = x_0, y_4 = -y_0 + 2L_y, z_4 = H;\ x_5 = -x_0, y_5 = y_0, z_5 = H;\ x_6 = -x_0 + 2L_x, y_6 = y_0, z_6 = H.$

For long dwelling times or short distances from the source

$$D \approx \frac{Q}{4\pi r K}\left(1 + \sum_n \frac{\theta r}{r_n}\right)$$

which will be the same as D_0 if $\theta r \ll r_n$.

3.3.2.4 Eddy diffusion parameterisation K

The dispersion coefficients, σ and σ_z, (where $\sigma^2(x) = 2K(x)t$ and $\sigma_z^2(x) = 2K_z(x)t$) are usually empirically derived functions of downwind distance when considering an outdoor release. Here we illustrate this using the so called 'Briggs' formulation (see table 4.2 of (Hanna, Briggs et al. 1982)) We then have, in general form, $\sigma = {ax}/{(1 + bx)^c}$, $\sigma_z = {a_z x}/{(1 + b_z x)^{c_z}}$ where the coefficients a, b, c (and z subscript counterparts) vary by Pasquill stability class (a categorisation based on wind speed and daytime insolation or nighttime cloudiness) and by whether the environment being modelled is urban or rural and are predefined (see for example (Hanna, Briggs et al. 1982)).

For an indoor release, as might be expected in a hub, the distances travelled by aerosol before hitting a barrier are likely to be 10's of metres and so $\sigma \cong ax$, $\sigma_z \cong a_z x$ because $b, b_z \ll 1$. An alternative parameterisation for standard deviation arising from outdoor release literature is to assume a power law in distance $\sigma^2 = \alpha x^\beta$ (Hanna, Briggs et al. 1982) (Stockie 2011). This is not explored further here but plausible parameterisations have $\beta \approx 1$. These outdoor parameterisations will not be valid for internal rooms but may have application for bus or rail stations, in particular for large outdoor platforms.

Cheng and colleagues (Cheng, Acevedo-Bolton et al. 2011) conducted trials and found that for a domestic sized room with natural ventilation $K = (0.52\gamma + 0.31)(H_c L_W L_L)^{2/3}$, where γ is defined as the air exchange rate only and the units of time is hours but this is unlikely to be useful in a transport hub setting as any ventilation will be mechanically driven and rooms will be larger.

Drivas *et al.* follow the derivation by Karlsson where the eddy diffusivity K is related to the mechanical ventilation, temperature profiles in the room and turbulent processes (Karlsson, Sjostedt et al. 1994) such that K is given by the solution of the cubic equation

$$\frac{K^3}{k^2 H_C^4} + \frac{g\Delta T\, K}{H_c T} = \frac{a u_0^2}{2}.$$

Where von Karman's constant $k = 0.4$, T is the room temperature [Kelvin], ΔT the temperature difference between ceiling and floor [Kelvin], $g = 9.8$ is acceleration due to gravity [m/s], H_c is the height of the room [m], a the fresh air ventilation rate [1/s] and u_0 is the velocity at the ventilation intake register [m/s].

Cooper and Horowitz conducted a trial in a room 7 metres by 10.7 metres by 2.6 metres with no dominant advection with 15 air exchanges per hour (Cooper and Horowitz 1986). In developing the model they expected $K = 0.1$ but in order for the model to explain the trial data, for short distances/long dwelling times a value of $K = 0.04$ is inferred but the empirical results suggest in fact that $K = 1$ and so the experimental data is not completely described by the model. The discrepancy is suggested by Drivas *et al*. as being due to the lack of surface reflection in the Cooper and Horowitz model. Note that the Cheng formula above for natural ventilation predicts $K = 0.07$.

The model presented by Drivas worked reasonably well to predict experimental studies conducted by VTT in a space with a volume of approximately of 1600 m^3 when the effective diffusion coefficient K is fitted to the data. These results are shown in Figure 5. Here a value of K about 0.01 m^2/s agreed quite well with the experiments. This is somewhat lower than the value of 0.04 m^2/s found by (Cooper and Horowitz 1986) in their studies.

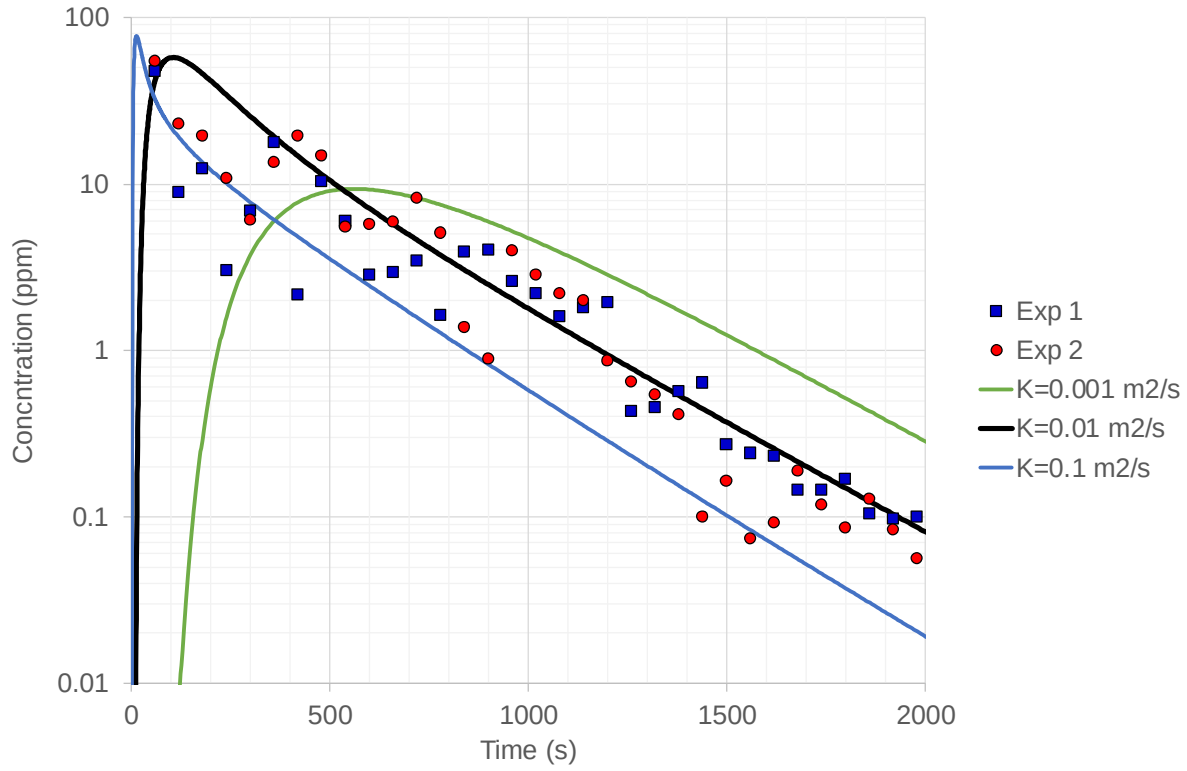


Figure 5: Comparison of Drivas model outputs with experimental results conducted by VTT. Points on the graph show the experimental data arising from the two trials whilst the two curves show different model evaluations for the given K values.

It is likely that the evaluation of K will be specific to the room conditions at the time of the release and so not parameterisable ahead of time and may vary with distance from the source. Indeed the Karlsson formula allows for the potential for variation with changes in temperature and intake rates around the room space. A sensitivity analysis will be necessary to test the impact of any choice of K. It should also be noted that the reported experiments were conducted in relatively small rooms with no people within them. People moving will cause additional turbulence and their thermal plumes will cause greater uplift of material towards the ceiling. This complexity is beyond the scope of this type of modelling.

3.3.3 Computational fluid dynamics

An order-of-magnitude estimate of airborne bio-particle concentrations can be achieved with models assuming continuous generation and uniform contaminant distribution in the confined space to be considered. They have been used to estimate particulate and microbial concentrations in spacecraft (Lobascio 1997). However, these simple models are not able to calculate concentration gradients which may be important in short term releases, such as incidents where the duration of the release is short compared to the nominal time constant t defined as the volume of the ventilated space divided by the air exchange rate: $t = {}^{1}/_{\gamma}$.

In principle, computational fluid dynamics (CFD) models, which numerically solve equations describing air flows and particle motions, are capable of more detailed analysis to find high contamination risk areas. Such models have been recently developed to study infection risk in hospital isolation rooms (Choi and Edwards 2008), operating theatres (Brohus, Balling et al. 2006) and in aircraft (Wan, Sze To et al. 2009). CFD has also been used to calculate chemical/biological (CB) agent spread in metro systems (Policastro and Gordon 1999) and metro trains (Engman (ed.) 2002).

In CFD, the geometric domain of interest is subdivided into a large number of small cells over which the equations of conservation of mass, energy and momentum are discretised and solved iteratively. The major factors affecting the accuracy of numerical simulations are roughly: the calculation grid size; the underpinning turbulence model; and the accuracy of boundary conditions (Allard, Awbi et al. 2007).

3.3.3.1 Calculation grid and geometry

The size and quality of the calculation grid effects the results, through the convergence of the calculations. In general, a non-uniform grid with a higher density of grid points near the interfaces where concentration and velocity gradients are large is needed, with particular attention required near walls and other barriers in the domain. The geometry of the simulated space should be close to the real one.

3.3.3.2 Turbulence models in the CFD code

The derivation of time-averaged fluid motion equations leads to the introduction of turbulent stresses and turbulent heat fluxes which act as additional diffusion terms due to correlations between the fluctuating velocities and temperature. These terms are unknown and can be modelled by different turbulent closures such as the simple eddy-viscosity model, in which the turbulent stresses and the turbulent heat fluxes are replaced by introducing turbulent viscosities. Despite the limitations of the eddy-viscosity model, it generally provides acceptable results with good computational economy. A more advanced approach, which provides more detailed and perhaps more accurate predictions for indoor airflows, is large eddy simulation, where small eddies are modelled, while larger eddies are assumed to affect the flow. However, this approach requires at least two orders longer computational times than the eddy-viscosity models making it often an impractical choice (Allard, Awbi et al. 2007).

3.3.3.3 Accuracy of boundary conditions

Boundary conditions affect the flow and are therefore crucial for the accuracy of the simulation results. In enclosed spaces, the typical boundary conditions include air supplies which modify the flow and input parameters, exhaust openings, contaminant and/or heat sources, and surfaces. Especially important is the accurate and correct description of air supply openings, because the flow characteristics of the diffusers usually dominate the air flow in an enclosed space.

An additional challenge pertinent to the PANDHUB project is the movement of people inside the transport hubs. A moving person creates air currents like wakes which enhance the dispersion of nearby contaminants. Therefore movements are most often ignored in simulations due to the complexity of the phenomenon although they can have a significant influence on the contamination field.

Each transport hub is unique and therefore the CFD simulation results are pertinent only to the specific case. The generalisation of results is usually not possible.

3.3.4 Summary of dispersion modelling

The eddy diffusion model proposed in section 3.3.2.3 allows for partial reflection. Critical to the model is the interaction between the eddy diffusion parameter, air exchange rates and dwelling time. The model allows for deposition on surfaces (the proportion $1-\theta$ of material that is not reflected) and additional deposition on floors due to the gravitational settling component of γ. This will then be the candidate model taken forward in later WP3 deliverables.

Particle size is also important, with larger bio-aerosols settling due to gravity rather than dispersing. The causative agents of plague and anthrax have different particle sizes, but, depending on the technology used in the dissemination device, it may be that particles are differently sized from that expected.

Models such as those presented here could be used to predict the impact of a cough or sneeze in the transport hub environment. Such human releases would cause smaller releases than those imagined for a deliberate release, so wall effects may tentatively be neglected and so D_0 may be used. However, a simpler transmission model given in section 2.2 may be sufficient better given the likely available contact data within transport hubs.

We recommend at this stage of research that for a detailed risk assessment of a bio-aerosol release in a transport hub computational fluid dynamics simulations be conducted. However, given the mismatch in apparent precision between the fluid flow calculations and individual location and movement patterns in a hub it is sufficient to use the coarser scale modelling with assessment of γ and sensitivity analysis conducted on the choice of K and θ.

3.4 Individual Human Exposure

Under the assumption of well-mixed air within a volume, the particle concentration profile is spatially uniform. When all particles are released into the air at a single moment the described dispersion scenario has exponential decay in the count of viable particles within space. For those presented scenarios where the air is not well mixed, the concentration of particles of interest within the volume is not uniform in space and in general will not be uniform in time. We must thus build a modelling framework that can account for non-uniform exposure.

3.4.1 Breathing rate

The exposure dose *D,* with dimension [bacteria s/m^3], must be scaled by the breathing rate of an individual β, with dimension [m^3/s], to provide an estimate of the number of bacteria inhaled. This is called the inhaled dose.

An individual's breathing rate is usually assumed not to change while a packet of air passes over him or her. The breathing rate of an average person may be assumed to be 0.012 metres cubed per minute as a baseline (light intensity activity) but this may increase or decrease depending on age, activity and stress levels (see Table 6.2 (Moya, Phillips et al. 2011)). Sedentary or resting travellers would be more likely to have breathing rates of 0.004 or 0.005 metres cubed per minute.

Movements of the population around the area of the dispersed contaminant may need to be considered as the duration of the release increases. That is, the assumption that a person will be static for the time that contamination passes over him or her from the various sources is potentially untenable and this is explored below in section 4.

3.4.2 Deposition

Only a fraction of the inhaled dose will be deposited in the lung, so $d = \varphi \beta D$ is the average number of organisms deposited in the lung for an individual (the retained or deposited dose). This reduction, φ, is due to some material settling in the upper respiratory tract, some being swallowed into the gut from which infection would be more difficult, or other clearance of material away from the lung. These processes may take time but should occur on a shorter timescale than the dispersal and the subsequent infection process. The parameter φ maybe derived from multi-pathway particle dosimetry modelling[5] for a given breathing rate.

More formally, a person challenged with some inhaled dose βD of bacteria has a fraction deposited in the lung space d (so $d \sim B(\varphi, P(\beta D))$ where B is a binomial distribution with probability of success φ and number of trials given by a Poisson distribution $P(\beta D)$). A binomial random variable with number of trials drawn from a Poisson distribution is itself then a Poisson distribution such that $d \sim P(\varphi \beta D)$. This deposition fraction will vary depending on activity level of host and particle size inhaled, but will be assumed not to vary with individual spores.

3.4.3 Competing risks framework

Recent work (Pratt, Bennett et al. 2016) has shown that, once deposited in the lung, both plague and anthrax infection may be modelled by a competing risks framework. This means that there are two critical events in the infection process. If one event occurs then subsequent infection is impossible, but if the other event occurs then subsequent infection is inevitable. For anthrax, the latter is germination of a spore and the former is its clearance, whilst for plague the latter is evasion of the host immune responses and former is clearance. If these events are random variables with exponential distributions with respective means of λ and θ then the probability that the case is not infected is given by $p = {\lambda}/{\lambda + \theta} = {1}/{1 + \alpha}$ where $\alpha = \theta / \lambda$.

We may make various assumptions about the behaviour of the terms in the competing risks model (Pratt, Bennett et al. 2016). For example, if there is little immune response variation between hosts but the deposition varies between hosts as a beta random variable then the probability P of infection is then given by

$$P = 1 - M\left(\mu\omega, \omega, -\frac{\beta D}{1 + \alpha}\right).$$

Where M is the Kummer function, and μ and ω are parameters of the deposition. The parameters α and β may then vary between individuals depending on activity and age, whilst the inhaled dose D is a function of space and time. Now we have a candidate model for assessing the risk of infection we must now consider the role of passenger movement in the environment.

[5] http://www.ara.com/products/mppd.htm

4. Passenger flows

We now have a model for dispersion of contaminant, but to assess subsequent human exposure we must consider the role of a human movement in the hub. This can be done by data analysis or by model simulation.

4.1 Room occupancy and transit

For the well-mixed scenario for aerosol dispersion (section 3.3.1), the deposited dose for a passenger will be determined by the time interval over which the passenger occupies the contaminated volume (the dwelling time,T_D, discussed above given some arrival time T_S as described in section 3.3.1).The deposited dose is then determined by the occupancy time of the room and the time of arrival in that area since the release.

To assess the profile of passenger dose exposure, it is sufficient to determine the distribution of arrival and length-of-stay times for passengers. It should be noted, however, that length-of-stay and arrival time are not certainly independent. For example, departure from a transport hub is often at a similar time for all passengers in a certain area, independent of arrival time in that area.

For static occupancy, such as in a gate waiting zone, of a contaminated area where the air volume is not well mixed (and so needing a model such as that described in section 3.3.2), the exposed---and consequently deposited---dose depends on the concentration profile of the contaminant at the point of occupation. To assess the profile of passenger dose exposure it is necessary to determine not the distribution of arrival and length-of-stay times for passengers, but the spatial distribution of passengers within the area of contamination.

Unlike the case of transit and occupancy for a well-mixed contamination zone, transit through a volume which is not well-mixed is not a special case of occupancy. Of particular note is the movement of passengers, making the dose exposure dependent not only on the time of arrival since release and continued exposure but on the path in space of the individual.

4.2 Determination of passenger exposure profiles

The three cases noted above for determinants of dose can all be addressed by a determination of passenger behaviour at the population level. To make this determination we need to briefly review past work on passenger flow simulation.

We wish to focus on the hub environment alone and to movements within that environment, but also limit to English articles in peer reviewed journals. Using the search terms *("transport" AND ("hub" OR "station" OR "termin*")) AND "passenger" AND ("flows" OR "move*")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j") OR LIMIT-TO (SRCTYPE , "k") OR LIMIT-TO (SRCTYPE , "b"))* within the Scopus document database returns 196 results. Some of these papers consider the movement of vehicles, consider economic rationale for hub placement and design, or interconnectedness of hubs and so do not consider the actual movement of individuals within a hub. Some also are focussed on evacuation of people from the hub for acute emergencies like a fire and so are no relevant in this context. This left 15 articles for full text review. After reading these articles six of these articles considered the flow of passengers on the transport network, not in the hub, two considered only hub demand/usages as a whole and one considering the transit time between platforms at a station in a simplistic manner and so all were rejected.

Understanding of passenger movement around a transport hub has been considered within the remaining six articles illustrating that individual passengers can simulated through being treated as agents in an agent-based model, or captured data can be analysed to statistically infer patterns of behaviour. This illustration of past work is sufficient for PANDHUB and we do not need to be more comprehensive in searching other reference databases.

4.2.1 Agent-based models for passenger movement

Three articles from the search considered simulation, two gave fairly superficial modelling frameworks (Yao, Sun et al. 2012) (Xie, Jia et al. 2013) though the third article presented a simulation model called PEDSIM for a metro station (Anagnostopoulos, Karlaftis et al. 2003). Through discussion with subject matter experts (in the UK Dept of Transport and the UK Defence Science and Technology Laboratory) there are established commercial modelling systems which may capture movement of passengers around a transport hub. Two of note are CAST[6] and AnyLogic[7]. We do not review both of those here, though note that CAST have been used to analyse both airports and train stations. We instead provide a conceptual overview.

In such an agent-based system, the floor plan of the modelled area will be provided. It will display functional aspects such as check-in areas, security zones, meeting points and waiting areas, entry and departure points (platforms, gates, doors), and walls.

Agents within the model will have characteristics and goals. Individuals will be characterized by activity, such as: passenger, stationary or patrolling staff, escorts and greeters, and visitors. A goal for a passenger may be to get to a particular zone (such as a departure point). Alternatively, it may be waiting for a specific time or until a particular event (such as queuing at check-in or security, or waiting to greet an arriving in-bound passenger). Vehicles may also be modelled as agents, but for this discussion they will be discounted.

Within the model, agents occupy space. The route an individual takes to move from one point to another features collision avoidance (both with other agents and the geometry) in addition to other routing choices. The preferred density of passengers may vary by activity or zone. From some experience of the CAST system, we believe that there may be application of these systems to an assessment of risk of exposure to the biological agents considered in the PANDHUB project.

Within a CAST model, various statistics for agents may be recorded. Further, patches of floor may be marked as sensors to capture passenger flow rates and speeds. Using such patches to correspond to various densities of an exposure footprint (perhaps determined by the earlier dispersion calculations) and agent statistics, population exposure may be established. Such calculations are beyond the true intention of the modelling system and this method has not been fully explored.

In addition to the weakness given above, any such simulations will be closely tied to the input. In particular the results will be, to varying degrees, sensitive to the layout of the transport hub, the release location, and passenger volumes (affected by season, day and time).

Whether the effort applying such specific modelling and data to understand overall risk, using general proprietary software with the associated cost overhead for software licenses, is

[6] http://www.airport-consultants.com/cast-simulation
[7] http://www.anylogic.com/areas/airports-stations-shopping-malls

sufficiently beneficial is to be determined by transport hub operators or governments. As an alternative, in the rest of this section we consider approaches which do not use this form of detailed agent-based modelling.

4.2.2 Data analysis for passenger movement

For passengers within a transport hub there are possibilities to trace their movements. For example, gross flow measures can be assessed from known end-points and volumes (Bing and Zhang 2013). Known ingress points at transport hubs couple with departure points in many situations. That is, passengers at an airport leaving on a specific flight will have a known departure time, egress location and an expected waiting area. Further, passage through other marked points, such as check-in area or security, can be captured through human observation or video capture (Weston and Marshall 1973) (Xie, Jia et al. 2015).

Alternatively, many contemporary transport hub operators offer free Wi-Fi connections to those within the hub. By consideration of the particular access points accessed by a single device, tracked through a unique identifier during the course of a visit (and indeed, across many visits), the route a passenger takes---including timing information---can be approximately determined. We can also see that transport hub operators provide distinct networks for staff and passengers, so that we can separately regard the different risk and behaviour profiles of those two classes of individuals. Being able to track across multiple visits further allows consideration of types of passenger behaviour.

Wi-Fi connection data are appealing only when there is a representative sample of passenger type and behaviour captured. From a survey at Helsinki airport, it was assessed that 70% of passengers not resident in Finland, and 30% of those resident in Finland, access the Wi-Fi network there during a visit. Of all passengers, 50% access the network during a visit.

A spatially referenced dataset exists for visitors accessing Helsinki airport. A device is registered in the dataset when it pings an access point. The interval between pings will vary by device, but an average time across devices of 30 seconds has been assumed by Finavia, the operator of Helsinki airport.

For the purpose of reporting and analysing the connection data, the various access points are collected into zones. Zones differ in size and one zone may cover several gate areas. Certain zones cover transit areas where one would mainly expect passengers to pass through rather than wait in them.

To facilitate analysis of the dataset, Finavia have a commercial software product, which produces reports based on user-defined queries. Broken down by zones, a report may detail passenger volumes over time. Alternatively, queries may be prepared to see flows of passengers (including routes taken) between specified start and end zones. Occupancy time within zones may also be determined, although as noted below there may be significant errors in this measure when the reported times are short.

Interpretation of the dataset is not without problems. For example, although the sampling of passengers seem reasonable, the accuracy of passenger movement determination is not perfect. For most analyses it is required that a device pings twice within the same zone: for transit or with devices with large ping intervals, there may be missing zones, although finer methods using inference of such discontinuities may mitigate this. Further, a device may be in the supposed footprint of one zone but ping an access point belonging to another (and so

be recorded there). This may occur even when there is an impassable barrier between the two zones.

In general, the limitations mentioned above are not themselves sufficient to lead to a preference to ignore this dataset.

4.2.3 Using passenger tracks to calculate exposure

As noted, passenger tracking data can be used to estimate room occupancy times for waiting areas by zone and time of day. For a well-mixed air volume with contaminant, the number of passengers and the distribution of occupancy times give, through the simple application of the dose exposure profile, an indication of the dose of contaminant deposited in the lung by passenger. It should be noted, though, that some zones may consist of several distinct enclosed areas. For a temporally varying level of contaminant it is further necessary to consider the distribution of entry times into the affect area.

Where the level of contamination may vary spatially, it is necessary to consider not only occupancy numbers but, in more detail, the waiting points in and routes through the zone. Careful consideration of the geometry of the area under consideration will identify possible routes, waiting areas and points of attraction. Detailed knowledge held by the transport hub operator will cover factors such as loitering behaviours and average transit speeds. For example, in a waiting lounge, the locations of seats may be considered to ensure that the model improves on an assumed uniform spatial distribution of present passengers. A preferred distancing or clumping behaviour of travellers or groups of travellers may be well understood by the operator.

For a release in a security area or corridor there are likely to be established well-defined routes taken by passengers, again at a fairly well understood travel speed. Data about flow rates over time can be used with a proposed risk field to estimate counts and times of exposure. Much of the effort of calculating applicable routes and waiting behaviour can be addressed using the modelling systems mentioned above, but with the level of uncertainty of the dispersion modelling and developed release scenario, a cruder estimate based on the simpler features of the captured population data may be sufficient to inform contingency planning for releases of this nature.

4.3 Evaluating the disease outbreak in wider community

Section 3.4 examines inhalation, deposition, and effect of the agent given a concentration in the air surrounding an individual. The previous parts of this section have elaborated on methods for estimating the overall exposure to the agent of those people in areas of transport hubs with risk.

Once infected the progression of disease in individuals will vary by disease and is discussed in D2.1, we restate this here briefly. An evaluation of symptom onset and progression for pandemic influenza and Ebola will be considered in D3.4 as part of the screening model parameterisation. For pneumonic plague we propose characterising the distributions based on the findings of Gani and Leach, who suggest a mean of 4.3 days (log-normally distributed with variance of 1.8) for the time between infection and symptom onset (incubation period) and then a mean of 2.5 days (log-normally distributed with variance 1.2) for the time between symptom onset and death (symptomatic period) (Gani and Leach 2004). Cases that survive will require hospitalisation for a longer period.

For inhalational anthrax the incubation period is dose dependent and be governed by the inhaled dose described in section 3, and can be mathematically given by the Wilkening model. Anthrax infection progresses in an individual with initial non-specific symptoms which last for some period (which we call the prodromal period) and with subsequently more specific symptoms (the fulminant period) before death. Holty *et al.* showed that the prodromal period can be descried with a log-normal distribution with mean 4.16 and standard deviation 2.22, whilst the fulminant period can be characterised by a log-normal distribution with parameters mean 0.667 and standard deviation 0.75, (Holty et al. 2006) with the dimension of the parameters given in days. Figure 7 illustrates a simulation of an anthrax attack on a population with these probability distributions governing transit between stages of infection.

These delays will impact on the delay between infection and eventual presentation to healthcare settings which may be an important consideration as we consider the spatial distribution of cases in D3.4 and mitigation in D3.5.

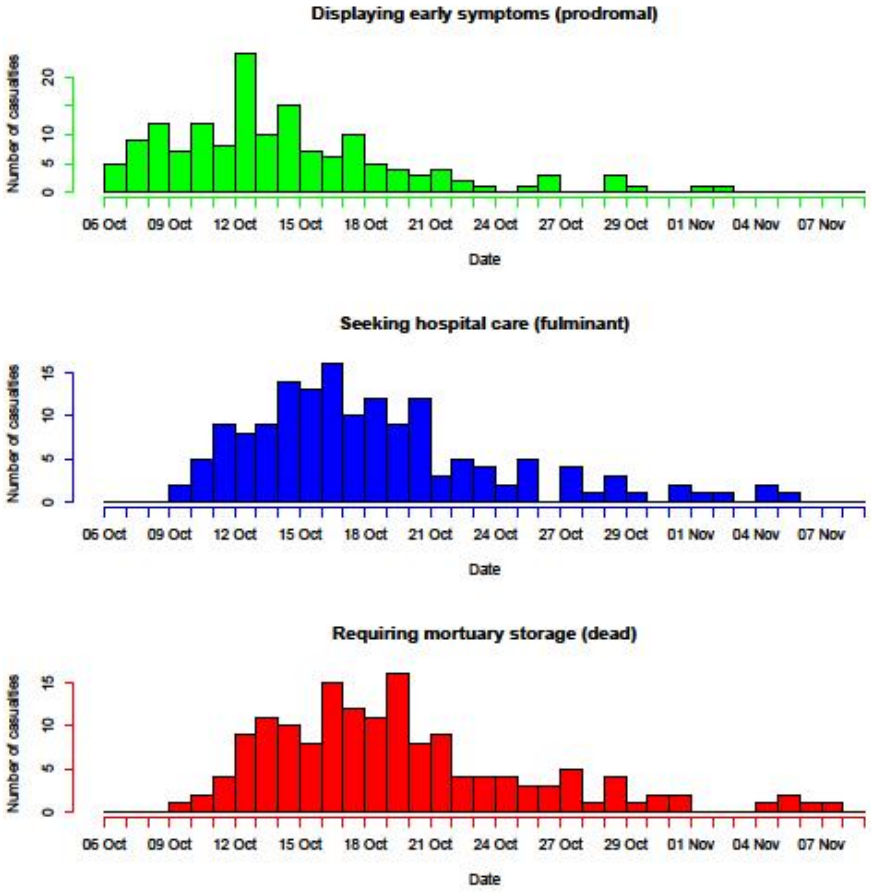


Figure 7: An example of determining anthrax outbreak characteristics given a population exposed to a defined passing plume of agent.

4.4 Discussion

This report presents a modelling strategy for assessing the likely burden of disease following a deliberate release of an airborne agent in a transport hub. The proposed release scenario suggests the appropriate dispersion model to use from section 3.3 and outputs from this model are used to assess exposure.

Assessing exposure may involve custom-built models in a full agent-based modelling system or a highly simplified estimate of passenger volumes and flows. Here we have described the possible application of the CAST modelling system and detailed data systems which capture passenger data.

Each of these methods to assess population movement within a transport hub takes the results of the dispersion modelling to calculate potential exposure of inhalable agent. From the number of exposed individuals, and the deposited dose for those individuals, characteristics of the disease outbreak may be calculated. Among these characteristics are the number of patients exhibiting classes of symptoms, the time to onset of symptoms, and the expected number of deaths (perhaps with particular intervention strategies).

Although the calculated resulting disease outbreak contains counts of patients in numerous classes, there are further complications to be expected around the reporting of disease. In general, the time to onset of symptoms will be of the order of days: it is likely that patients will seek healthcare away from the location of the transport hub. D3.4 is concerned with the ultimate presentation to the healthcare system of the diseased. This dispersion of cases from a large transport hub, possibly across many countries, complicates the detection and intervention aspects of any response. Two of the authors of this deliverable have been granted access to the commercial system used by Finavia. Indicative extracts of the Helsinki Wi-Fi connection dataset will be used in preparing results presented in D3.4.

It is recommended that this project takes a data driven approach to passenger movement inference rather than an agent based simulation based one.

5. Next steps and Recommendations

In this report we have developed a mathematical modelling framework to consider infection of individuals within the transport hub environment by pandemic diseases and through deliberate release using the specific agents used as scenarios in the earlier D2.1. This report therefore delivers task 3.3 of the PANDHUB programme of work by taking the scenarios from D2.1 and building on outputs from D2.3 about disease transmission modes for the diseases considered and considers the impact of passenger flows with a framework for simulation modelling.

Limitations in the modelling have been discussed throughout the report. In D3.4 we will present the results of the simulations to show the spatial extent both within the hub and external to the hub. In D3.5 we will further develop the infection models presented here to consider mitigation strategies for contingency planning.

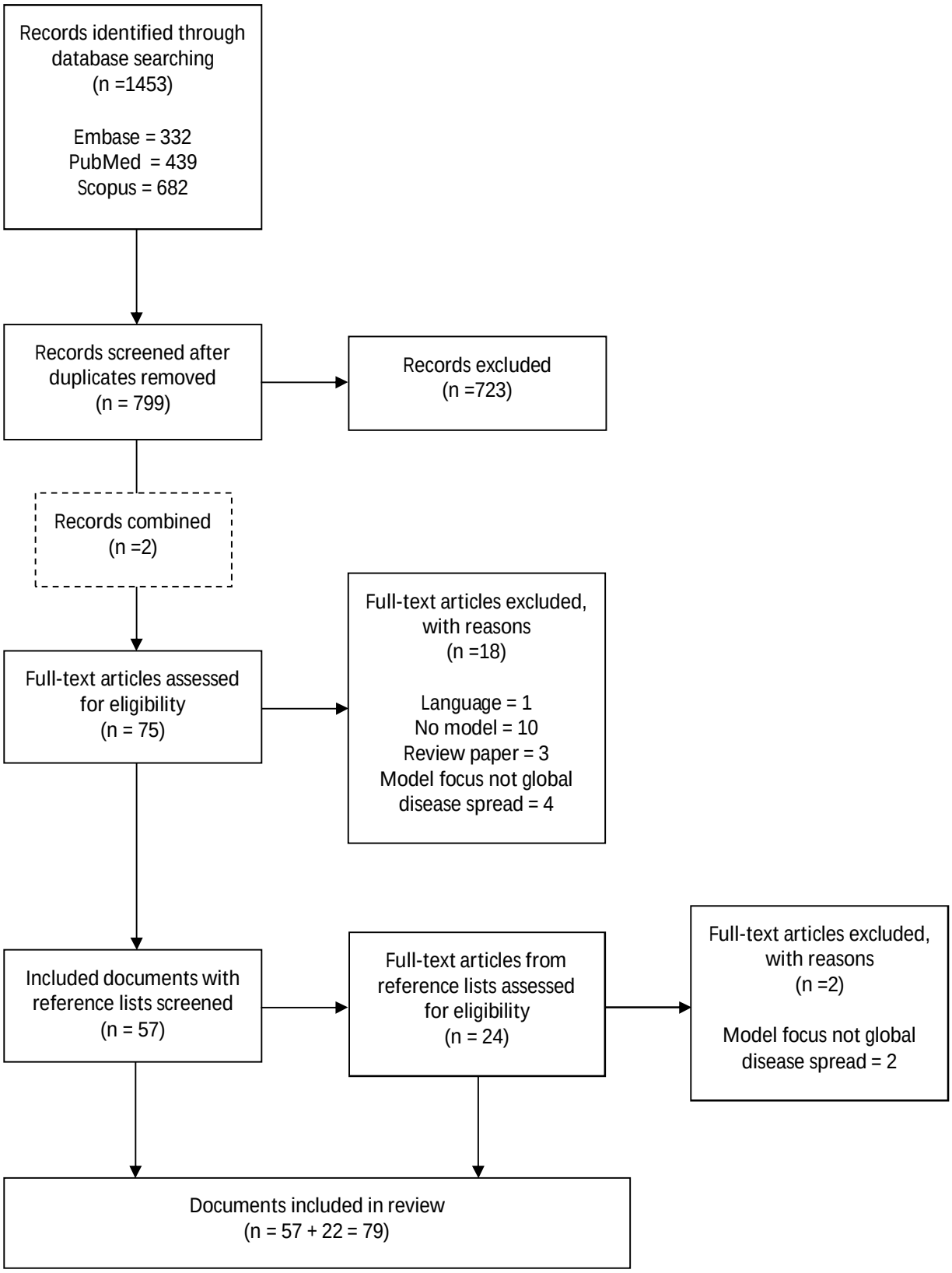
Specifically we have conducted a thorough review of the literature regarding international spread of infection and have identified the GLEaM model[8] as a good candidate for modelling such spread. It is a well-established and published model. Whilst this is a potential candidate we are not recommending that it is necessarily used by operators or planners: this should be a judgement for them given the specific scenario they are considering, we are simply advocating that we do not need to develop another global disease simulation tool in this project.

We have developed a framework for simulating the appearance of a symptomatic traveller inside a hub given the time elapsed since their exposure to the infection. This has necessary simplifications but for rare importations such appearance will occur randomly within the transport hub and so is independent of the precise passenger flows. We also have presented a candidate model for infections directly arising from cases in a hub. Specific parameterisation will appear in D3.4 to illustrate the risks of contact with other users and the hub staff for the case study infections.

We have developed a simulation method for deliberate release of biological hazard. Given the number of simulation tools identified we do not recommend creating another tool but will instead recommend a data driven approach or analysis of specific data arising from a hub context that purports to measure passenger flows. Further analysis of the data arising from passenger flows around Helsinki airport is necessary and so simulation results will appear in D3.4 to show the spatial risks within the hub environment.

[8] http://www.gleamviz.org/model/

Supplementary information - Global disease spread scoping review



1. Apolloni A, Poletto C, Colizza V. Age-specific contacts and travel patterns in the spatial spread of 2009 H1N1 influenza pandemic. BMC Infectious Diseases. 2013;13(1).

Study Goal	Investigate the role of mixing and travel behaviours of a population to determine the effect during a major epidemic.
Model Type	Multi-host stochastic metapopulation model with two age classes. A gravity model is used for population movement.
Input Data	Epidemiological data: H1N1 data – taken from existing literature Population (age): Eurostat, census Travel: airport-specific stats, Eurostat.
Validation Data	none
Notes	Input data taken from 8 European countries plus Mexico. Age travel statistics from 8 European countries plus Mexico were collected, typically obtained from travel surveys at airports.

2. Apolloni A, Poletto C, Ramasco JJ, Jensen P, Colizza V. Metapopulation epidemic models with heterogeneous mixing and travel behaviour. Theoretical Biology and Medical Modelling. 2014;11(1).

Study Goal	Determine the pandemic potential of a disease considering heterogeneous mixing patterns.
Model Type	Multi-host stochastic epidemic metapopulation model.
Input Data	Epidemiological data: existing literature Population: Polymod data. UN statistics. Travel: existing literature
Validation Data	none
Notes	Spatially structured population with non-homogeneous mixing. Different population partitions, mixing patterns and mobility structures are Considered.

3. Arino J, Portet S. Epidemiological implications of mobility between a large urban centre and smaller satellite cities. J Math Biol. 2015.

Study Goal	Consider a central urban area surrounded by satellite cities. Can a satellite city infect the urban centre? Can a satellite city protect itself from infection stemming from the urban centre? How does the size of a city affect its vulnerability to invasion by an infectious disease or its role in facilitating the spread of a disease?
Model Type	SIR-type metapopulation model
Input Data	Epidemiological data: existing literature Population: census Travel: University of Manitoba Transport Information Group
Validation Data	none
Notes	Article focus is on analysing mathematical system.

4. Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. PLoS ONE. 2011;6(1).

Study Goal	Assess the effect of different travel restrictions on stopping or delaying a pandemic.
Model Type	GLEaM (Global Epidemic and Mobility) model. Metapopulation model with network modelling for subpopulations. SEIR model.
Input Data	Epidemiological data: existing literature (H1N1pdm) Population: census Travel: IATA
Validation Data	None
Notes	Defined a baseline situation based on 2009 H1N1pdm data and then did comparisons with theoretical scenarios representing different interventions on travel.

5. Balcan D, Colizza V, Goncalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc Natl Acad Sci U S A. 2009;106(51):21484-9.

Study Goal	Determine how local and global population mobility combine to affect the global spread of an infectious disease
Model Type	GLEaM
Input Data	Epidemiological data: Existing literature. Population: none Travel: IATA (number of available seats per year for each direct connection between two of these airports. The coverage of the dataset is estimated to be 99% of the global commercial traffic.) Offices of Statistics for multiple countries.
Validation Data	2001-2002 seasonal influenza data for regions within the USA.
Notes	

6. Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C, Ramasco JJ, et al. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): A Monte Carlo likelihood analysis based on human mobility. BMC Medicine. 2009;7:45.

Study Goal	Estimate the transmission potential of H1N1pdm during the winter of 2009 within the Northern hemisphere . Also assess the seasonal effects on the virus and predict the spatiotemporal pattern of the pandemic.
Model Type	Global structured metapopulation model. (GLEaM)
Input Data	Epidemiological data: Existing literature; initial conditions of epidemic set to fit La Gloria, Mexico. Population: Socioeconomic Data and Application Center (SEDAC) of Columbia University dataset Travel: IATA, OAG
Validation Data	No validation data as the article was written in July 2009 and sought to predict the spread of H1N1pdm for the northern hemisphere winter months of 2009.
Notes	

7. Bobashev G, Morris RJ, Goedecke M. Sampling for global epidemic models and the topology of an International airport network. PLoS ONE. 2008;3(9).

Study Goal	Determine an appropriate sample of cities to be included in a network model so that it is most representative of the real world whilst also being a tractable model.
Model Type	Network model
Input Data	Epidemiological data: Not clearly stated. Population: US census bureau Travel: OAG 2000
Validation Data	None
Notes	Claim: A model including the largest airports, the largest cities, the most-connected cities, and the most central cities is enough to describe the dynamics of the global spread of influenza. Analysis conducted suggests that a relatively small number of cities (200 or 300 out of alround 3000) can capture enough network information to adequately describe the global spread of a disease such as influenza.

8. Bogoch, II, Creatore MI, Cetron MS, Brownstein JS, Pesik N, Miniota J, et al. Assessment of the potential for international dissemination of Ebola virus via commercial air travel during the 2014 west African outbreak. Lancet. 2015;385(9962):29-35.

Study Goal	Assess the potential for Ebola to spread across international borders via commercial air travel. Assess the relative efficiency of exit versus entry screening.
Model Type	Data analysis - analysis of IATA flight data.
Input Data	Epidemiological data: WHO reports of the number of confirmed, probable, suspected cases of Ebola. Population: World Bank 2013 population estimates. Travel: International Air Transport Association data: Information on future flight schedules - passenger carrying capacity as seats on flights between directly connected airports; dataset including monthly, passenger-level flight itinerary data on past flights.
Validation Data	none
Notes	No clear description of how any modelling was done.

9. Bonabeau E, Toubiana L, Flahault A. The geographical spread of influenza. Proceedings of the Royal Society B: Biological Sciences. 1998;265(1413):2421-5.

Study Goal	Model how an infectious disease spreads within one epidemic cycle
Model Type	Time series analysis
Input Data	Epidemiological data: Weekly reports from general practitioners regarding influenza from across French territory (1984 – 1995). Population: Spatial distribution of French population from French Institut Géographique National. Travel: none
Validation Data	none
Notes	Test a null hypothesis against data set. Hypothesis is that the spread of an epidemic is statistically uniform in geographic space.

10. Brent Daniel W, Hengartner NW, Rivera MK, Powell DR, McPherson TN. An epidemiological model of spatial coupling for trips longer than the infectious period. Mathematical Biosciences. 2013;242(1):1-8.

Study Goal	Modelling the effects of transportation systems on infectious disease spread when the travel is not much shorter than the infectious period and daily travels are not small when compared to overall population size.
Model Type	SEIR model
Input Data	Epidemiological data: existing literature – Hong Kong 1968 influenza outbreak Population: existing literature Travel: existing literature
Validation Data	None
Notes	Difficult to determine what was done/ what data were used.

11. Brockmann D, Helbing D. The hidden geometry of complex, network-driven contagion phenomena. Science. 2013;342(6164):1337-42.

Study Goal	Demonstrate that complex network systems in the real world can be accurately described by a comparatively simple wave propagation model
Model Type	SIRmodel on a global mobility network.
Input Data	Epidemiological data: Not clear – existing literature? WHO flunet. Population: none Travel: OAG (Difficult to determine what data was used and where it came form.)
Validation Data	none
Notes	

12. Brockmann D, Hufnagel L, Geisel T. Dynamics of modern epidemics. SARS: A Case Study in Emerging Infections2007.

Study Goal	Forecast the geographical spread of an epidemic.
Model Type	Stochastic model. Stochastic SIR model.
Input Data	Epidemiological data: existing literature for SARS in Hong Kong. Population: none Travel: 2003 IATA, 2003 AOG (is this OAG?)
Validation Data	WHO outbreak data
Notes	Modelling initial phase of an epidemic where SIR not appropriate because number of infected is very small. Probabilistic model considered instead.

13. Brownstein JS, Wolfe CJ, Mandl KD. Empirical Evidence for the Effect of Airline Travel on Inter-Regional Influenza Spread in the United States. PLoS Med. 2006;3(10):e401.

Study Goal	Model the spread of influenza across a large population (the USA)
Model Type	Data analysis
Input Data	Epidemiological data: CDC pneumonia and influenza mortality data, climate effects from National Climatic Data Center. Population: existing literature - climate effects on human behaviour Travel: 2005 air carrier traffic statistics, US Department of Transportation Bureau of Transportation Statistics
Validation Data	Viral surveillance data from the WHO/NREVSS collaborating laboratories (1997–2005). Weekly influenza-like illness data for France (1996–2005), French Sentinel Network.
Notes	

14. Camitz M, Liljeros F. The effect of travel restrictions on the spread of a moderately contagious disease. BMC Medicine. 2006;4.

Study Goal	The effect of different travel restrictions on a SARS-like disease.
Model Type	Data analysis
Input Data	Epidemiological data: existing literature Population: None)possibly in the survey: not clear) Travel: Survey by Statistics Sweden (1999-2001)
Validation Data	None
Notes	Allowed for a small fraction of the population to not comply with travel restrictions. Cannot access Appendix 1, where more information on the data is available.

15. Chao DL, Halloran ME, Obenchain VJ, Longini IM, Jr. FluTE, a Publicly Available Stochastic Influenza Epidemic Simulation Model. PLoS Comput. Biol. 2010;6(1):e1000656.

Study Goal	Description of a computer simulation model which can be used to model the dynamics of influenza outbreaks across a large population.
Model Type	Agent-based
Input Data	Epidemiological data: existing literature Population: US census (2000), existing literature. Travel: Census Transportation Planning Package - information on the home and destination census tracts of US workers. Existing literature which references U. S. Department of Transportation, Bureau of Transportation Statistics. Existing literature for airport figures.
Validation Data	none
Notes	Model source code available.

16. Colizza V, Barrat A, Barthelemy M, Valleron AJ, Vespignani A. Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. PLoS Medicine. 2007;4(1):0095-110.

Study Goal	Assess possible containment strategies for a global outbreak of influenza H5N1.
Model Type	Metapopulation stochastic epidemic model. SLIR model.
Input Data	Epidemiological data: existing literature Population: census data Travel: International Air Transport Association (IATA) data (airport pairings where direct flights occur, number of seats on flights)
Validation Data	None
Notes	Model predicts future outbreak so requires a new outbreak for validation.

17. Colizza V, Barrat A, Barthelemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci U S A. 2006;103(7):2015-20.

Study Goal	Determine the role of the worldwide transportation network on the global diffusion of infectious diseases and assess the reliability of forecasts.
Model Type	metapopulation
Input Data	Epidemiological data: unclear. Existing literature? Population: census data Travel: International Air Transport Association (IATA) data (world list of airport pairs connected by direct flights and the number of available seats on any given connection for 2002)
Validation Data	none
Notes	

18. Colizza V, Barrat A, Barthélemy M, Vespignani A. The modeling of global epidemics: Stochastic dynamics and predictability. Bulletin of Mathematical Biology. 2006;68(8):1893-921.

Study Goal	Determine the effect of the air travel network on the global spread of an emerging infectious disease.
Model Type	Stochastic model using epidemic Langevin equation. SIR compartmental model.
Input Data	Epidemiological: existing literature Population: census Travel: IATA data (2002)
Validation Data	none
Notes	Example case: influenza pandemic.

19. Colizza V, Barrat A, Barthélemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: The SARS case study. BMC Medicine. 2007;5.

Paper	V. Colizza, A. Barrat, M. Barthélemy and A. Vespignani	2007	Predictability and epidemic pathways in global outbreaks of infectious diseases: The SARS case study
Study Goal	Develop a model of global infectious disease spread which makes use of transportation data.		
Model Type	Stochastic meta-population model with SIR-type compartmental model.		
Input Data	Epidemiological: existing literature - Hong Kong SARS outbreak Population: census Travel: IATA		
Validation Data	WHO SARS reported cases		
Notes			

20. Colizza V, Barrat A, Barthélemy M, Vespignani A. Epidemic predictions and predictability in complex environments. Biophysical Reviews and Letters. 2008;3(1-2):217-26.

Study Goal	Model the global spread of emerging infectious diseases which explicitly, incorporating transportation network and census data.
Model Type	Computational model. Compartmental model. Stochastic metapopulation model.
Input Data	Epidemiological: existing literature - Hong Kong SARS outbreak Population: census Travel: IATA
Validation Data	WHO SARS reported cases
Notes	Model details do not appear in the paper. No obvious mention of supplementary information.

21. Cooley P, Brown S, Cajka J, Chasteen B, Ganapathi L, Grefenstette J, et al. The role of subway travel in an influenza epidemic: A New York city simulation. Journal of Urban Health. 2011;88(5):982-95.

Study Goal	Effect of subway travel on the spread of an infectious disease
Model Type	Agent-based computer simulation with SEIR framework
Input Data	Epidemiological: existing literature [9] . Influenza from longini, 1957-1958 influenza pandemic from fergusion et al. H1N1pdm data. Population: National statistics, census, business analyst data, schools database data, household surverys. Travel: Census, travel patterns surveys
Validation Data	none
Notes	Simulation of 5 boroughs of New York City with each person residing in one of these boroughs. Each person has a set of socio-demographic characteristics. Interactions with households, workplaces, schools and community activities are included with subway riders stratified as commuters, shoppers and miscellaneous travellers. Concluded that interventions aimed at subways riders would provide "very limited benefits on overall attack rates and epidemic peaks".

22. Eichner M, Schwehm M, Wilson N, Baker MG. Small islands and pandemic influenza: Potential benefits and limitations of travel volume reduction as a border control measure. BMC Infectious Diseases. 2009;9:160.

Study Goal	Quantify the value of travel restrictions imposed in Pacific island nations during an influenza pandemic to determine their efficacy in delaying or preventing the arrival of the disease.
Model Type	Probabilistic model
Input Data	Epidemiological: unclear Population: unclear Travel: country statistics departments
Validation Data	None
Notes	Assumed that the disease would become global through aircraft travel facilitating spread. Cannot access supplementary information which may contain more information on the data.

[9] Longini I, Nizam A, Xu S, et al. Containing pandemic influenza at the source. Science. 2005; 309: 1083–1087.
Ferguson NM, Cummings D, Cauchemez S, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature. 2005; 437: 209–214.
Germann T, Kadau K, Longini IJ, Macken C. Mitigation strategies for pandemic influenza in the United States. PNAS. 2006; 103(15): 5935–5940.
Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature. 2006; 442(7101): 448–526.
Halloran ME, Ferguson NM, Eubank S, et al. Modeling targeted layered containment of an influenza pandemic in the United States. Proc Natl Acad Sci USA. 2008; 105(12):
4639–44.

23. Epstein JM, Goedecke DM, Yu F, Morris RJ, Wagener DK, Bobashev GV. Controlling pandemic flu: The value of international air travel restrictions. PLoS ONE. 2007;2(5).

Study Goal	Determine the effects of travel restrictions and vaccination on the global transmission of pandemic influenza, also considering economic costs of interventions.
Model Type	Stochastic difference equation compartmental model
Input Data	Epidemiological: existing literature - H5N1 influenza Population: census, existing literature Travel: OAG
Validation Data	none
Notes	

24. Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature. 2006;442(7101):448-52.

Study Goal	Model the effectiveness of different intevention options in a novel influenza pandemic scenario.
Model Type	Individual based simulation model
Input Data	Epidemiological: existing literature – 1968 influenza pandemic; existing literature; household datasets Population: Landscan 2003 (constructed from census, remote sensing, land use and transport network data) – gives instantaneous population density; census data; existing literature Travel: Origin-Destination Survey of Airline Passenger Traffic dataset collected by the Bureau of Transportation Statistics of the US Department of Transportation; 2001 National Household Travel Survey
Validation Data	none
Notes	

25. Flahault A, Deguen S, Valleron AJ. A mathematical model for the European spread of influenza. Eur J Epidemiol. 1994;10(4):471-4.

Study Goal	Simulate an influenza epidemic in Europe.
Model Type	metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: International Civil Aviation Organization (1991)
Validation Data	none
Notes	

26. Flahault A, Letrait S, Blin P, Hazout S, Menares J, Valleron AJ. Modelling the 1985 influenza epidemic in France. Stat Med. 1988;7(11):1147-55.

Study Goal	Simulate the 1984-1985 influenza epidemic in France.
Model Type	metapopulation
Input Data	Epidemiological: Computer Network for Surveillance of Communicable Diseases, surveys, existing literature (1968-1969 Hong Kong pandemic, mortality data from England) Population: none Travel: National Railroad Network (France)
Validation Data	none
Notes	

27. Flahault A, Valleron A. A method for assessing the global spread of HIV-1 infection based on air travel. Mathematical population studies. 1992;3(3):161-71, 227.

Study Goal	Quantify the impact of air traffic on the global spread of HIV.
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: existing literature - L. A. Rvachev and I. M. Longini Jr (1985) used 1968-1969 mean daily number of airline passengers data.
Validation Data	Authors states that the model was validated by comparing results with accepted geographic patterns of HIV-1 spread.
Notes	Abstract mentions that this paper aplies an influenza global disease spread model but has been adapted for HIV.

28. Flahault A, Vergu E, Boëlle P-Y. Potential for a global dynamic of Influenza A (H1N1). BMC Infectious Diseases. 2009;9(1):1-11.

Study Goal	Predict the worldwide spread of influenza A H1N1 across 52 cities across the world.
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: existing literature
Validation Data	none
Notes	

29. Flahault A, Vergu E, Coudeville L, Grais RF. Strategies for containing a global influenza pandemic. Vaccine. 2006;24(44-46):6751-5.

Study Goal	Determine the efficacy of 5 interventions (isolation of infectious patients, reduction of air traffic, therapeutic antiviral treatment, prophylactic antiviral treatment, vaccination) on the spread of a pandemic disease across the world through the air traffic network.
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: existing literature
Validation Data	none
Notes	Development of the Rvachev and Longini model (see, for instance, A mathematical model for the global spread of influenza (1985))

30. Gautreau A, Barrat A, Barthélemy M. Arrival time statistics in global disease spread. Journal of Statistical Mechanics: Theory and Experiment. 2007(9).

Study Goal	Look at the arrival time of a disease given starting seed and travel patterns
Model Type	Metapopualtion
Input Data	Epidemiological: unclear - none Population: uinlclear – existing literature as Colizza thanked in acknowledgements Travel: unclear – IATA mentioned in acknowledgements but not in bibliography
Validation Data	none
Notes	

31. Gautreau A, Barrat A, Barthélemy M. Global disease spread: Statistics and estimation of arrival times. Journal of Theoretical Biology. 2008;251(3):509-22.

Study Goal	Determine the speed at which an epidemic arrives in a cities linked by transportation networks
Model Type	Metapopualtion
Input Data	Unclear where data is from
Validation Data	Same as input data
Notes	

32. Goedecke DM, Bobashev GV, Yu F, editors. A stochastic equation-based model of the value of international air-travel restrictions for controlling pandemic FLU. Proceedings - Winter Simulation Conference; 2007.

Study Goal	Look at the spread of influenza and potential effects of air travel restrictions using an equation-based model
Model Type	Metapopualtion
Input Data	Epidemiological: existing literature Population: census, surveys, existing data Travel: census, surveys, existing data
Validation Data	None
Notes	

33. Gonçalves B, Balcan D, Vespignani A. Human mobility and the worldwide impact of intentional localized highly pathogenic virus release. Scientific Reports. 2013;3.

Study Goal	Describe the (theoretical) global spread of smallpox being released before any intervention is set up
Model Type	Large scale structured metapopulation model: Global Epidemic and Mobility model (GLEaM)
Input Data	Epidemiological: existing data Population: Center for International Earth Science Information – Gridded Population of the World and Global Rural-Urban Mapping Project Travel: IATA, OAG, existing literature
Validation Data	none
Notes	

34. Grais RF, Ellis JH, Glass GE. Assessing the impact of airline travel on the geographic spread of pandemic influenza. European Journal of Epidemiology. 2003;18(11):1065-72.

Study Goal	Putting the 1968-69 pandemic fly strain in today's air travel volumes and see how it would spread globally
Model Type	Metapopualtion
Input Data	Epidemiological: existing literature (Rvachev and Longini (1985)) Population: US Department of State's International Population Database Travel: OAG; IATA; US Department of Transportation (Origin-Destination database, T100 and T100(f) databases); International Civil Aviation Organization; Back Aviation Solutions Inc.; existing literature
Validation Data	none
Notes	

35. Han XP, Zhao ZD, Hadzibeganovic T, Wang BH. Epidemic spreading on hierarchical geographical networks with mobile agents. Communications in Nonlinear Science and Numerical Simulation. 2014;19(5):1301-12.

Study Goal	Explore the patterns of spread of disease in relation to traffic systems and mobility behaviours
Model Type	Metapopualtion
Input Data	unclear
Validation Data	none
Notes	Compared mobility patterns to literature: Gonzalez *et al* (2008); Brockmann *et al* (2006); Jiang *et al* (2009)

36. He D, Lui R, Wang L, Tse CK, Yang L, Stone L. Global spatio-temporal patterns of influenza in the post-pandemic era. Scientific Reports. 2015;5.

Study Goal	Using FluNet data to determine interactions between influenza virus strains and determine global seasonal patterns of pandemic virus activities.
Model Type	Data analysis
Input Data	Epidemiological: FluNet (2010-2013), existing data Population: World borders dataset Travel: none
Validation Data	Fits of time series data to influenza trends in ten different countries.
Notes	

37. Hollingsworth TD, Ferguson NM, Anderson RM. Will travel restrictions control the international spread of pandemic influenza? Nat Med. 2006;12(5):497-9.

Study Goal	Analyse the impact of restricting international travel during the 2003 SARS epidemic.
Model Type	Probabilitstic
Input Data	Epidemiological: WHO confirmed case annoucenments, existing literature Population: none Travel: IATA, individual airport data
Validation Data	none

Notes	

38. Hollingsworth TD, Ferguson NM, Anderson RM. Frequent travelers and rate of spread of epidemics. Emerging Infectious Diseases. 2007;13(9):1288-94.

Study Goal	Simulating high frequency travellers spreading SARS and influenza-like illnesses
Model Type	Metapopulation
Input Data	Epidemiological: none Population: none Travel: existing literature
Validation Data	none
Notes	Population split into frequent and infrequent fliers

39. Hsu CI, Shih HH. Transmission and control of an emerging influenza pandemic in a small-world airline network. Accid Anal Prev. 2010;42(1):93-100.

Study Goal	Look at the effects of air travel on the evolution of a flu pandemic in a small world network, by looking at human to human transmission
Model Type	Network model
Input Data	Epidemiological: existing literature Population: none Travel: flight data from Amadeus website, Airports Council Internationl (2004 Worldwide Airport Traffic statistics)
Validation Data	none
Notes	

40. Huang CY, Wen TH, editors. A multilayer epidemic simulation framework integrating geographic information system with traveling networks. Proceedings of the World Congress on Intelligent Control and Automation (WCICA); 2010.

Study Goal	Describe the MEDSIM (Multilayer Epidemic Dynamics Simulation model) network that integrates population and network approaches
Model Type	Metapopulation
Input Data	unclear Epidemiological: Population: census Travel: census; transportation data
Validation Data	none
Notes	Four layered model using: GIS to model the human travel network, model the effect of difference age groups moving across cities, model contact between age groups and individuals grouped by age SEIR model of transmission.

41. Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(42):15124-9.

Study Goal	Introduce a new model describing the global spread of disease and prove that forecasting of spread is possible
Model Type	Probabilisitc two part model: 1 – local infection dynamic: SIR model 2 – aviation network dispersal: matrix of transition probabilities among populations
Input Data	Epidemiological: WHO data, CDC data Population: none Travel: IATA, OAG
Validation Data	none
Notes	Authors state that model outputs show an almost one to one correspondance with WHO data, however this is the same data set used to parametrise the model.

42. Hwang GM, Mahoney PJ, James JH, Lin GC, Berro AD, Keybl MA, et al. A model-based tool to predict the propagation of infectious disease via airports. Travel Medicine and Infectious Disease. 2012;10(1):32-42.

Study Goal	Determine time trends for new influenza to arrive in the US from best connected international cities to the US in February.
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: census, united nations data, World Gazetteer Travel: DIIO (Data In. Intelligence Out.)
Validation Data	Outputs are close to the reported H1N1pdm mean incidence rate.
Notes	

43. Hyman JM, LaForce T. Modeling the Spread of influenza Among Cities. In: Banks HT, Castillo-Chavez c, editors. Bioterrorism: Mathematical Modeling Applications in Homeland Security. Philadelphia, USA: SIAM; 2003. p. 211-36.

Study Goal	Analyse US influenza and pneumonia data, searching for correlations between the number of cases in a city and other epidemiological parameters.
Model Type	metapopulation
Input Data	Epidemiological: existing literature, fits to CDC mortality data Population: US census (2000) Travel: US Department of Transportation (airline flight data)
Validation Data	CDC data (1996-2001)
Notes	

44. Johansson MA, Arana-Vizcarrondo N, Biggerstaff BJ, Staples JE, Gallagher N, Marano N. On the treatment of Airline travelers in mathematical models. PLoS ONE. 2011;6(7).

Study Goal	Use two models to determine the speed and pattern of spread of disease by airlines.
Model Type	Metapopulation models: SARS direct transmission model, dengue vector-borne transmission model
Input Data	Epidemiological: existing literature Population: United Nations Statistics Division (2005 Demographic Yearbook) and Population Division (World Urbanisation Prospects: the 2007 Revision Population Database) Travel: OAG, US Department of Transportation. Three different travel network parameteriza Tions assessed: seat-based direct-travel only, connecting-travel-inclusive,and a skewed version of the connecting-travel.
Validation Data	none
Notes	

45. Kenah E, Chao DL, Matrajt L, Halloran ME, Longini IM, Jr. The global transmission and control of influenza. PLoS One. 2011;6(5):e19515.

Study Goal	Model the seasonality of influenza in the tropics and implications for a public health response if a pandemic were to occur.
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: existing literature Travel: existing literature
Validation Data	Fot to H1N1pdm and Hong kong influenza.
Notes	

46. Kerneis S, Grais RF, Boelle PY, Flahault A, Vergu E. Does the effectiveness of control measures depend on the influenza pandemic profile? PLoS One. 2008;3(1):e1478.

Study Goal	Determine trends in geographical and temporal global influenza spread, taking into account the variability of the parameters involved
Model Type	metapopulation
Input Data	Epidemiological: existing literature Population: International population database, Central Intelligence Agency - The World Factbook 2009. Travel: none
Validation Data	none
Notes	Many of the model parameters were varied so only some had values taken from existing literature

47. Khan K, Arino J, Hu W, Raposo P, Sears J, Calderon F, et al. Spread of a Novel Influenza A (H1N1) Virus via Global Airline Transportation. New England Journal of Medicine. 2009;361(2):212-4.

Study Goal	Model how travellers disseminated during the beginning of the influenza H1N1 2009 pandemic and consequent;ly predict the spread of the disease.
Model Type	Data analysis
Input Data	Epidemiological: unclear Population: unclear Travel: IATA
Validation Data	none
Notes	

48. Knipl DH, Röst G, Wu J. Epidemic spread and variation of peak times in connected regions due to travel-related infections-dynamics of an antigravity-type delay differential model. SIAM Journal on Applied Dynamical Systems. 2013;12(4):1722-62.

Study Goal	Formulate a model that best describes the temporal spread of influenza between places connected by long–distance travel
Model Type	SEAIR based antigravity model
Input Data	Epidemiological: existing literature, WHO surveillance data Population: existing literature Travel: Statistics Canada (http://www.statcan.gc.ca);
Validation Data	none
Notes	

49. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. PLoS Pathog. 2014;10(2):e1003932.

Study Goal	Incorporating viral genetic data to human travel patterns to evaluate predictors in the global spread of flu
Model Type	Generalised Linear Model (GLM), with: Monte Carlo to determine communities; SIR to determine migration; continuous time Markov chain for phylogenetic movement
Input Data	Epidemiological: World Health Organization database (FluNet) Population: Existing literature (Geographica: the complete illustrated atlas of the world), www.citypopulation.de, World bank (2008) Travel: OAG
Validation Data	none
Notes	

50. Lessler J, Kaufman JH, Ford DA, Douglas JV. The Cost of Simplifying Air Travel When Modeling Disease Spread. PLoS ONE. 2009;4(2):e4403.

Study Goal	Compare a simple model infectious disease spread (including air travel) with a more complex one to assess differences in model results.
Model Type	Pipe model, gravity model
Input Data	Epidemiological: none Population: none Travel: U. S. Department of Transportation Research and Innovative Technology administration Bureau of Transportation Statistics (RITA-BTS) - individual ticket data from 2007.
Validation Data	none
Notes	

51. Li X, Tian H, Lai D, Zhang Z. Validation of the gravity model in predicting the global spread of influenza. International Journal of Environmental Research and Public Health. 2011;8(8):3134-43.

Study Goal	Evaluate the spatial and temporal scales for which the gravity model is sufficiently accurate
Model Type	Gravity model
Input Data	Epidemiological: WHO (number of cases)(April to July 2009); Centre for Disease Control and Prevention (April to July 2009) Population: US Census Bureau; International Monetary Fund (IMF) World Economic Outlook Databases, U.S. Department of Commerce Travel: none
Validation Data	Comparison of model results to number of actual flu cases (6th July 2009) from WHO for 168 countries
Notes	

52. Liu L, Liu X. Global stability of a transport-related infection model with general incidence rate in two heterogeneous cities. BioSystems. 2014;126:41-51.

Study Goal	Assess the effect of travel on disease spread via a two-city transport-related disease spread model with heterogeneity and different demographic parameters.
Model Type	Metapopulation
Input Data	No data
Validation Data	No data
Notes	

53. Longini IM. A mathematical model for predicting the geographic spread of new infectious agents. Mathematical Biosciences. 1988;90(1):367-83.

Study Goal	Model the temporal and geographic spread of an epidemic
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: existing literature
Validation Data	WHO morbidity data
Notes	

54. Longini IM, Jr., Fine PE, Thacker SB. Predicting the global spread of new infectious agents. Am J Epidemiol. 1986;123(3):383-91.

Study Goal	Review a model of global disease spread, including its application to Hong Kong influenza A(H3N2)
Model Type	Metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: Canberra Department of Transportation - Air transport statistics (1981); ABC Travel Guide - ABC World airways guide (1981); International Civil Aviation Organisation (1968-1982)); existing literature.
Validation Data	WHO morbidity data
Notes	Builds on work from: O. V. Baroyan, G. A. Mironov and L. A. Rvachev, An algorithm modeling global epidemics of mutant origin, *Programming and Comput. Software* 6(5):272-277 (1981), English transl. from *Programmirovanie*, No. 5, pp. 73-79 (1980) (in Russian).

55. Marcelino J, Kaiser M. Reducing influenza spreading over the airline network. PLoS Currents. 2009(AUG).

Study Goal	Study the effect of removing particular flight routes rather that shuitting down transport hubs on the spread of influenza.
Model Type	Agent-based
Input Data	Epidemiological: none Population: none Travel: OAG
Validation Data	none
Notes	

56. Marcelino J, Kaiser M. Critical paths in a metapopulation model of H1N1: Efficiently delaying influenza spreading through flight cancellation. PLoS Currents. 2012.

Study Goal	Assess the effect of targeted flight restrictions to delay the spread of influenza.
Model Type	metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: OAG
Validation Data	None
Notes	

57. Meloni S, Perra N, Arenas A, Gómez S, Moreno Y, Vespignani A. Modeling human mobility responses to the large-scale spreading of infectious diseases. Scientific Reports. 2011;1:62.

Study Goal	Assess the effect of self-initiated behavioral changes into the mobility patterns of Individuals on an epidemic.
Model Type	metapopulation
Input Data	Epidemiological: none Population: none Travel: OAG
Validation Data	None
Notes	

58. Merler S, Ajelli M. The role of population heterogeneity and human mobility in the spread of pandemic influenza. Proceedings of the Royal Society B: Biological Sciences. 2010;277(1681):557-65.

Study Goal	Understand how the sociodemographic variations across Europe and heterogeneous mobility patterns affect the spread of flu across these countries
Model Type	individual-based simulation model
Input Data	Epidemiological: unclear Population: Eurostat; National Statistical Offices (for countries not covered by Eurostat); existing literature Travel: Eurostat
Validation Data	Outcomes seen in the H1N1 pandemic in Europe; comparision of attack rates to those observed (by age) during 1918-19 US Spanish flu outbreak
Notes	

59. Merler S, Ajelli M, Pugliese A, Ferguson NM. Determinants of the spatiotemporal dynamics of the 2009 h1n1 pandemic in europe: Implications for real-time modelling. PLoS Computational Biology. 2011;7(9).

Study Goal	Determine what factors influenced the spread of H1N1 in Europe (2009 pandemic)
Model Type	SEIR
Input Data	Epidemiological: European Centre for Disease Prevention and Control (2009) Population: Eurostat Travel: Eurostat
Validation Data	Health Protection Agency data on H1N1pdm
Notes	

60. Meyers LA, Pourbohloul B, Newman ME, Skowronski DM, Brunham RC. Network theory and SARS: predicting outbreak diversity. J Theor Biol. 2005;232(1):71-81.

Study Goal	Offer insight into the heterogeneity of SARS outbreaks and applications to public health strategies.
Model Type	Agent-based
Input Data	Epidemiological: existing literature Population: Statistics Canada, BC Stats, Vancouver School Board December 2002 Ready Reference; existing literature (The British Columbia Health Atlas). Travel: none
Validation Data	none
Notes	

61. Mugglin AS, Cressie N, Gemmell I. Hierarchical statistical modelling of influenza epidemic dynamics in space and time. Stat Med. 2002;21(18):2703-21.

Study Goal	Data analysis method for modelling the spread of an infectious disease via contact between infected and susceptible individuals.
Model Type	Bayesian statistical model
Input Data	Epidemiological: personal communication, 1991 census (Scotland) – reference not clearly stated , temperatgure – no reference Population: 1991 census (Scotland) – references not clearly stated, existing literature Travel: none
Validation Data	none
Notes	

62. Nakata Y, Röst G. Global analysis for spread of infectious diseases via transportation networks. Journal of Mathematical Biology. 2014.

Study Goal	Study the spread of infection during transport between hetrogeneous regions connected in a transportation network
Model Type	metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: none
Validation Data	none
Notes	

63. Nicolaides C, Cueto-Felgueroso L, González MC, Juanes R. A Metric of Influential Spreading during Contagion Dynamics through the Air Transportation Network. PLoS ONE. 2012;7(7):e40961.

Study Goal	Highlight that different metrics (network topology, geography, traffic structure and individual mobility patterns) are important to consider to predict the spread of disease
Model Type	Agent-based model
Input Data	Epidemiological: none Population: none Travel: Federal Aviation Administration (US, January 2007 to July 2010); existing literature (Barnhart C, Fearing D, Vaze V (2011) Modeling passenger travel and delays in the National Air Transportation System.)
Validation Data	none
Notes	Empirical model of human mobility

64. Ohkusa Y, Sugawara T. Application of an individual-based model with real data for transportation mode and location to pandemic influenza. Journal of Infection and Chemotherapy. 2007;13(6):380-9.

Study Goal	Using real data to trace day by day an outbreak diffusion in a metropolitan area
Model Type	Data analysis
Input Data	Epidemiological: existing literature Population: none Travel: Tokyo Metropolitan Traffic Planning Council (2004) - Guide for person-trip data in Tokyo metropolitan area
Validation Data	none
Notes	

65. Poletto C, Tizzoni M, Colizza V. Heterogeneous length of stay of hosts' movements and spatial epidemic spread. Scientific Reports. 2012;2.

Study Goal	Model disease dynamics in a spatially structured population where mobility is influenced by heterogeneous lengths of stay.
Model Type	Metapopulation model with spatially structure population. The mobility process has a heterogeneous length of stay. SIR infection dynamics at nodes.
Input Data	Epidemiological: none Population: unclear Travel: Eurostat; UK Travel Trends data; Australian Department of Resources, Energy and Tourism - Tourism Research Australia (TRA)
Validation Data	none
Notes	

66. Poletto C, Tizzoni M, Colizza V. Human mobility and time spent at destination: Impact on spatial epidemic spreading. Journal of Theoretical Biology. 2013;338:41-58.

Study Goal	Model for infectious disease spread which incorporates heterogeneous movement of individuals.
Model Type	metapopualtion
Input Data	None
Validation Data	None
Notes	Individuals have memory of their origin subpopulation and time at other locations is distributed to model the heterogeneity of human mobility.

67. Ruan S, Wang W, Levin SA. The effect of global travel on the spread of sars. Math Biosci Eng. 2006;3(1):205-18.

S. Ruan 2009 The effect of global travel on the spread of SARS
Math Biosci Eng 6 1 207-8

Study Goal	Global spread of SARS using medical geography theory
Model Type	Metapopulation
Input Data	Epidemiological: existing listerature (Gumel *et al* (2004)) Population: none Travel: none
Validation Data	none
Notes	A correction was made in 2009 with reference to Litaker, J. R., Chou, J. Y., Novak, S. and Wilson, J. P. (2003), Implications of SARS: Medicalgeography and surveillance in disease detection, *Ann. Pharmacotherapy* 37:1841-1848.

68. Rvachev AL, Longini Jr IM. A mathematical model for the global spread of influenza. Mathematical Biosciences. 1985;75(1):3-22.

Study Goal	Model the pandemic spread of a single strain of influenza across 52 of the world's major cities, located on all continents.
Model Type	metapopulation
Input Data	Epidemiological: existing literature Population: Demographic reference books – no references given Travel: OAG, Canberra Department of Transportation - Air transport statistics; ABC Travel Guide - ABC World airways guide (1981); International Civil Aviation Organisation (1968-1982)); existing literature.
Validation Data	WHO morbidity data
Notes	Builds on work from: O. V. Baroyan, G. A. Mironov and L. A. Rvachev, An algorithm modeling global epidemics of mutant origin, *Programming and Comput. Software* 6(5):272-277 (1981), English transl. from *Programmirovanie*, No. 5, pp. 73-79 (1980) (in Russian).

69. Sattenspiel L, Dietz K. A structured epidemic model incorporating geographic mobility among regions. Math Biosci. 1995;128(1-2):71-91.

Study Goal	Model infectious disease spread across dicrete geographic regions with mobility between these regions.
Model Type	metapopulation
Input Data	No clear references for data
Validation Data	Same as for input data.
Notes	

70. Shi P, Keskinocak P, Swann JL, Lee BY. The impact of mass gatherings and holiday traveling on the course of an influenza pandemic: A computational model. BMC Public Health. 2010;10.

Study Goal	Explore how social mixing and contact pattern variations impact the course of a flu pandemic
Model Type	Agent based model, temporally and spatially explicit
Input Data	Epidemiological: existing literature Population: 2000 US Census Data, existing literature Travel: none
Validation Data	none
Notes	

71. Tizzoni M, Bajardi P, Poletto C, Ramasco JJ, Balcan D, Gonçalves B, et al. Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm. BMC Medicine. 2012;10(1):1-31.

Study Goal	Validate predictions from a real-time model against the 2009 influenza A/H1N1 pandemic.
Model Type	metapopulation
Input Data	Epidemiological: existing literature, official reports – not listed, country/region-specific surveillance data Population: U.S. Census Bureau and Department of Transportation, existing literature Travel: U.S. Census Bureau and Department of Transportation, existing literature, national statistics
Validation Data	Comparison with H1N1 data
Notes	

72. Tuncer N, Le T. Effect of air travel on the spread of an avian influenza pandemic to the United States. International Journal of Critical Infrastructure Protection. 2014;7(1):27-47.

Study Goal	Spread of avian influenza by air travel
Model Type	Metapopulation 2 city model; consider only susceptible and infected individuals
Input Data	Epidemiological: existing literature (WHO) Population: unclear – existing literature? Travel: United States Department of Transportation (Office of the Secretary for Aviation and International Affairs) - International airline flight data
Validation Data	none
Notes	*Takes control measures into consideration*

73. Van den Broeck W, Gioannini C, Goncalves B, Quaggiotto M, Colizza V, Vespignani A. The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. BMC Infect Dis. 2011;11:37.

Study Goal	Development of a computer simulation model for the worldwide spread of a person to person transmissible disease.
Model Type	metapopulation
Input Data	Epidemiological data: existing results from GLEaM analysis of influenza H1N1pdm data Population: Socioeconomic Data and Applications Center (SEDAC). Columbia University Travel: IATA, OAG
Validation Data	none
Notes	Article focuses on describing the GLEaMviz software rather than on a specific outbreak scenario.

74. Viboud C, Bjornstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science. 2006;312(5772):447-51.

Study Goal	Analyse interpandemic spread of influenza across the US
Model Type	Gravity model
Input Data	Epidemiological: existing literature Population: U.S. Census Bureau and Department of Transportation, existing literature Travel: U.S. Census Bureau and Department of Transportation, existing literature
Validation Data	none
Notes	

75. Yoneyama T, Krishnamoorthy MS. Simulating the spread of influenza pandemic of 2009 considering international traffic. Simulation. 2012;88(4):437-49.

Study Goal	Model how 2009 flu pandemic spread through the world
Model Type	metapopulation
Input Data	Unclear where data is from. Authors mention the use of 'real data' but there is no accompanying citation in the text.
Validation Data	As with Input data.
Notes	

76. Yoneyama T, Krishnamoorthy MS. Simulating the spread of pandemics with different origins considering international traffic. International Journal of Advancements in Computing Technology. 2012;4(22):201-9.

Study Goal	Determine global spread of pandemic through hybrid model
Model Type	metapopulation
Input Data	World Almanac Books (populations) and WTO (global connections) Epidemiological: Unclear where the claimed 'real data' comes from. No reference given. Population: None. Travel: World Tourism Organization - Compendium of Tourism Statistics 2008, Yearbook of Tourism Statistics 2008
Validation Data	Unclear
Notes	

77. Zhang Y, Liu Z, Zhang Y, Yang H, Bo Y, Fang L, et al., editors. Spatially explicit epidemiological simulation system of influenza A (H1N1) in China. 2010 18th International Conference on Geoinformatics; 2010 18-20 June 2010.

Study Goal	Study the transmission of influenza A H1N1 within mainland China.
Model Type	metapopulation
Input Data	Epidemiological: existing literature Population: none Travel: existing literature
Validation Data	none
Notes	

78. Zhang Y, Zhang Y, Liu Z, editors. The role of different transportation in the spreading of new pandemic influenza in mainland China. Proceedings - 2011 19th International Conference on Geoinformatics, Geoinformatics 2011; 2011.

Study Goal	Simulate the spread of H1N1 (2009) within and between provinces in China
Model Type	SEIR, gravity model
Input Data	Epidemiological: Ministry of Health of China Population: none Travel: National Bureau of Statistics of China (railway traffic), Tibet Bureau of Statistics (railway traffic)
Validation Data	none
Notes	

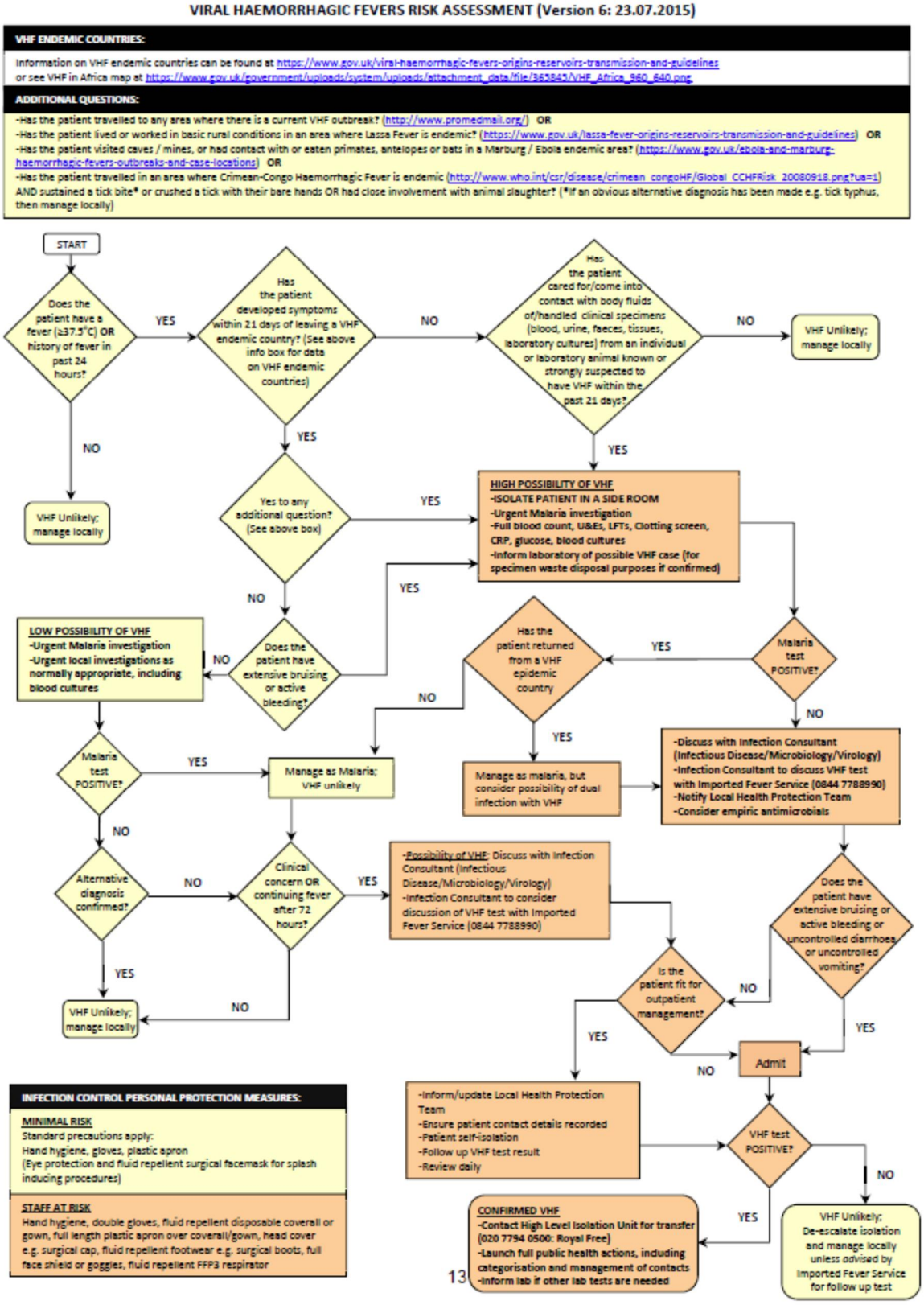
79. Zhong SB, Huang QY, Song DJ. Simulation of the spread of infectious diseases in a geographical environment. Science in China, Series D: Earth Sciences. 2009;52(4):550-61.

Study Goal	Simulate the spread of infectious disease at local and global level
Model Type	Theoretical using several model types : SIRS and Cellular Automata (CA)
Input Data	Epidemiological: WHO (SARS data) Population: none Travel: none
Validation Data	None
Notes	Population: SIR with further sub classification by states; geography: defined by cells (CA model)

Supplementary Information– Direct intervention at transport hub through screening

Adobe Acrobat Document

Supplementary Information– screening questionnaire



Screening algorithm for VHFs (viral haemorrhagic fevers), from page 13 of (Advisory Committee on Dangerous Pathogens 2015). Contains public sector information licensed under the Open Government Licence v3.0.

Bibliography

Advisory Committee on Dangerous Pathogens (2015). Management of Hazard Group 4 viral haemorrhagic fevers and similar human infectious diseases of high consequence.

Allard, F., H. Awbi, L. Davidson and A. Schälin (2007). Computational Fluid Dynamics in Ventilation Design. REHVA Guidebook No 10., REHVA: Federation of European Heating and Air-Conditioning Associations.

Anagnostopoulos, D., M. G. Karlaftis and M. Nikolaidou (2003). "Development of an object-oriented pedestrian traffic flow simulation environment for transport terminal planning." Journal of Advanced Transportation **37**(1): 83-103.

Baize, S., D. Pannetier, L. Oestereich, T. Rieger, L. Koivogui, N. F. Magassouba, B. Soropogui, M. S. Sow, S. Keïta, H. De Clerck, A. Tiffany, G. Dominguez, M. Loua, A. Traoré, M. Kolié, E. R. Malano, E. Heleze, A. Bocquin, S. Mély, H. Raoul, V. Caro, D. Cadar, M. Gabriel, M. Pahlmann, D. Tappe, J. Schmidt-Chanasit, B. Impouma, A. K. Diallo, P. Formenty, M. Van Herp and S. Günther (2014). "Emergence of Zaire Ebola Virus Disease in Guinea." New England Journal of Medicine **371**(15): 1418-1425.

Bennett, E., A. Bennett, N. Silman, T. Pottage and I. Hall (in preparation). "Human exposure and infection risk from low levels of indoor *Bacillus anthracis*."

Bing, L. and L. Zhang (2013). Research on Simulation of Pedestrian in Passenger Transfer Hub Based on Commuter Trans Tech Publications. **409:** 1407--1414.

Brachman, P. S., H. Gold, S. A. Plotkin, F. R. Fekety, M. Werrin and N. R. Ingraham (1962). "Field Evaluation of a Human Anthrax Vaccine." American Journal of Public Health and the Nations Health **52**(4): 632-645.

Briand, S., E. Bertherat, P. Cox, P. Formenty, M.-P. Kieny, J. K. Myhre, C. Roth, N. Shindo and C. Dye (2014). "The International Ebola Emergency." New England Journal of Medicine **371**(13): 1180-1183.

Brohus, H., K. D. Balling and D. Jeppesen (2006). "Influence of movements on contaminant transport in an operating room." Indoor Air **16**(5): 356-372.

Carson, J. S. (2002). Model verification and validation. Simulation Conference, 2002. Proceedings of the Winter.

Cheng, K.-C., V. Acevedo-Bolton, R.-T. Jiang, N. E. Klepeis, W. R. Ott, O. B. Fringer and L. M. Hildemann (2011). "Modeling Exposure Close to Air Pollution Sources in Naturally

Ventilated Residences: Association of Turbulent Diffusion Coefficient with Air Change Rate." Environmental Science & Technology **45**(9): 4016-4022.

Choi, J. I. and J. R. Edwards (2008). "Large eddy simulation and zonal modeling of human-induced contaminant transport." Indoor Air **18**(3): 233-249.

Cooper, D. W. and M. Horowitz (1986). "Exposures from indoor powder releases: models and experiments." American Industrial Hygiene Association Journal **47**(4).

Dahlgren, C. M., L. M. Buchanan, H. M. Decker, S. W. Freed, C. R. Phillips and P. S. Brachman (1960). "BACILLUS ANTHRACIS AEROSOLS IN GOAT HAIR PROCESSING MILLS." American Journal of Epidemiology **72**(1): 24-31.

Diekmann, O. and J. A. P. Heesterbeek (2000). Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. Chichester, UK, Wiley.

Doolan, D. L., D. A. Freilich, G. T. Brice, T. H. Burgess, M. P. Berzins, R. L. Bull, N. L. Graber, J. L. Dabbs, L. L. Shatney, D. L. Blazes, L. M. Bebris, M. F. Malone, J. F. Eisold, A. J. Mateczun and G. J. Martin (2007). "The US Capitol Bioterrorism Anthrax Exposures: Clinical Epidemiological and Immunological Characteristics." Journal of Infectious Diseases **195**(2): 174-184.

Drivas, P. J., P. A. Valberg, B. L. Murphy and R. Wilson (1996). "Modeling Indoor Air Exposure from Short-Term Point Source Releases." Indoor Air **6**(4): 271-277.

Engman (ed.), L. K. (2002). Chemical Weapons—Threat, Effects and Protection. Sweden, Swedish Defence Research Agency.

Ferguson, N. M., D. A. T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn and D. S. Burke (2005). "Strategies of containing an emerging influenza pandemic in Southeast Asia." Nature **437**: 209-214.

Ferguson, N. M., D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley and D. S. Burke (2006). "Strategies for mitigating an influenza pandemic." Nature **442**(7101): 448-452.

Fischer, W. A. and D. A. Wohl (2016). "Confronting Ebola as a Sexually Transmitted Infection." Clinical Infectious Diseases **62**(10): 1272-1276.

Fowkes, N. D. and J. J. Mahony (1994). An Introduction to Mathematical Modelling, Wiley.

Gani, R. and S. Leach (2004). "Epidemiological determinants for modelling pneumonic plague outbreaks." Emerging Infectious Disease **10**(4): 608-614.

Hanna, S. R., G. A. Briggs and R. P. J. Hosker (1982). Handbook on Atmospheric Dispersion, U.S. Department of Energy.

Hethcote, H. W. (2000). "The Mathematics of Infectious Diseases." SIAM Rev. **42**(4): 599-653.

House, T. (2014). "Epidemiological determinants of Ebola outbreaks." Elife.

IATA. Retrieved 13/07/2016, from http://www.iata.org/Pages/default.aspx.

Karlsson, E., A. Sjostedt and S. Hakkansson (1994). "Can weak turbulence give high concentrations of carbon dioxide in baby cribs." Atmospheric Environment **28**: 1297-1300.

Keeling, M. J. and P. Rohani (2008). Modeling Infectious Diseases in Humans and Animals, Princeton University Press.

Lalitha, M. K., V. Anandi, N. Walter, J. O. Devadatta and B. M. Pulimood (1988). "Primary anthrax presenting as an injection "abscess"." Indian J Pathol Microbiol **31**(3): 254-256.

Legrand, J., R F. Grais, P Y. Boelle, A J. Valleron and A. Flahault (2007). "Understanding the dynamics of Ebola epidemics." Epidemiology and Infection **135**(4): 610-621.

Lobascio, C. (1997). Optimizing the airborne particulate contamination control function in manned spacecraft. In the proceedings of the Sixth European Symposium on Space Environmental Control Systems, Noordwijk, the Netherlands.

MOBS Lab. "Global Epidemic and Mobility Model." Retrieved 13/07/2016, from http://www.mobs-lab.org/global-epidemic-and-mobility-model.html.

Moya, J., L. Phillips, L. Schuda, P. Wood, A. Diaz, R. Lee, R. Clickner, R. J. Birch, N. Adjei, P. Blood, K. Chapman, R. Castro and K. Mahaffey (2011). Inhalation rates Exposure Factors handbook, Environmental Protection Agency.

Norman, P. S., J. G. Ray Jr, P. S. Brachman, S. A. Plotkin and J. S. Pagano (1960). "Serologic testing for anthrax antibodies in workers in a goat hair processing mill." Am J Hyg **72**: 32-37.

OAG. Retrieved 13/07/2016, from http://www.oag.com/.

Osterholm, M. T., K. A. Moore, N. S. Kelley, L. M. Brosseau, G. Wong, F. A. Murphy, C. J. Peters, J. W. LeDuc, P. K. Russell, M. Van Herp, J. Kapetshi, J.-J. T. Muyembe, B. K. Ilunga, J. E. Strong, A. Grolla, A. Wolz, B. Kargbo, D. K. Kargbo, P. Formenty, D. A. Sanders and G. P. Kobinger (2015). "Transmission of Ebola Viruses: What We Know and What We Do Not Know." mBio **6**(2): e00137-00115.

Parker, S. and V. Bowman (2011). "State-space methods for calculating concentration dynamics in multizone buildings." Building and Environment **46**: 1567-1577.

Policastro, A. J. and S. P. Gordon (1999). The Use of Technology in Preparing Subway Systems for Chemical/ Biological Terrorism. Proceedings of the 1999 Commuter Rail/Rapid Transit Conference Toronto, Canada.

Pratt, A., E. Bennett, S. Leach and I. Hall (2016). "Dose-response modelling: extrapolating from experimental data to wider populations." In preparation.

Public Health England (2015). Ebola Q&A for the public.

Quinn, C. and P. Turnbull (1998). Anthrax. Topley and Wilson's microbiology and microbial infections. L. Collier. London, UK, Arnold. **3**.

Rasouli, F. and T. A. Williams (1995). "Applicaiton of dispersion modelling to indoor gas release scenarios." Journal of the air and waste management association **45**(3): 191-195.

Rodrigue, J. P., C. Comtois and B. Slack (2013). The Geography of Transport Systems, Taylor & Francis.

Sargent, R. G. (2011). Verification and validation of simulation models. Proceedings of the Winter Simulation Conference. Phoenix, Arizona, Winter Simulation Conference**:** 183-198.

Stockie, J. M. (2011). "The mathematics of atmospheric dispersion modeling." SIAM Review **53**(2): 349-372.

Turnbull, P. (1996). "Guidance on Environments Known to Be or Suspected of Being Contaminated with Anthrax Spores." Land Contamination and Reclamation **4**: 37--46.

Wan, M. P., G. N. Sze To, C. Y. H. Chao, L. Fang and A. Melikov (2009). "Modeling the Fate of Expiratory Aerosols and the Associated Infection Risk in an Aircraft Cabin Environment." Aerosol Science and Technology **43**(4): 322-343.

Wattiau, P., M. Govaerts, D. Frangoulidis, D. Fretin, E. Kissling, M. Van Hessche, B. China, M. Poncin, Y. Pirenne and G. Hanquet (2009). "Immunologic response of unvaccinated workers exposed to anthrax, Belgium." Emerg Infect Dis **15**(10): 1637-1640.

Weston, J. G. and J. Marshall (1973). "Pedestrian movement in crossing flows." Transportation Planning and Technology **2**(1): 49-54.

WHO Plague Fact Sheet. "Plague Fact Sheet Number 267." Retrieved 25/07/2016, from http://www.who.int/mediacentre/factsheets/fs267/en/.

WHO Plague Manual (1999). Plague manual: epidemiology, distribution, surveillance and control.

Wolfram Math World. "Cellular Automata." Retrieved 20/10/2016, from http://mathworld.wolfram.com/CellularAutomaton.html.

World Health Organization (2008). Anthrax in humans and animals.

Xie, Z., L. Jia, Y. Qin and L. Wang (2013). "A Hybrid Temporal-Spatio Forecasting Approach for Passenger Flow Status in Chinese High-Speed Railway Transport Hub." Discrete Dynamics in Nature and Society **2013**: 7.

Xie, Z., L. Jia and L. Wang (2015). "Passenger Flow Detection of Video Surveillance: A Case Study of High-Speed Railway Transport Hub in China." Elektronika ir Elektrotechnika **21**.

Yao, L., L. Sun, Z. Zhang, S. Wang and J. Rong (2012). "Research on the Behavior Characteristics of Pedestrian Crowd Weaving Flow in Transport Terminal." Mathematical Problems in Engineering **2012**: 9.