Characterising livestock system ‘zoonoses hotspots’

A. Kleczkowski, A. Breed, L. Matthews, D. Thronicker and F. de Vries
8 February 2012

Executive summary
A systematic review of the published literature was undertaken, to explore the ability of different types of model to help identify the relative importance of different drivers leading to the development of zoonoses hotspots. We estimated that out of 373 papers we included in our review, 108 papers touched upon the objective of ‘Assessment of interventions and intervention policies’, 75 addressed the objective of ‘Analysis of economic aspects of disease outbreaks and interventions’, 67 the objective of ‘Prediction of future outbreaks’, but only 37 broadly addressed the objective of ‘Sensitivity analysis to identify criteria leading to enhanced risk’. Most models of zoonotic diseases are currently capturing outbreaks over relatively short time and largely ignoring socio-economic drivers leading to pathogen emergence, spill-over and spread. In order to study long-term changes we need to understand how socio-economic and climatic changes affect structure of livestock production and how these in turn affect disease emergence and spread. Models capable of describing this processes do not appear to exist, although some progress has been made in linking social and economical aspects of livestock production and in linking economics to disease dynamics. Henceforth we conclude that a new modelling framework is required that expands and formalises the ‘one world, one health’ strategy, enabling its deployment in the re-thinking of prevention and control strategies. Although modelling can only provide means to identify risks associated with socio-economic changes, it can never be a substitute for data collection. Finally, we note that uncertainty analysis and uncertainty communication form a key element of modelling process and yet are rarely addressed.
Introduction

The aim of this study is to establish the contribution that models (mathematical and economic) can make to identifying the characteristics (e.g. location relative to centres of population, intensity of livestock numbers, type of management etc) of livestock systems in developing countries which are likely to lead to the emergence of “zoonoses hotspots”. Our report provides a short review of the current usage of models and particularly network and agent-based methods in studying zoonotic disease outbreaks, with emphasis on developing countries. Here we present initial findings. The attached annexes of the presentation at a workshop organised by DFID and the review paper provide the full findings.

Zoonoses hotspots have many diverse origins and mechanisms leading to emergence, spill-over, spread and persistence of diseases (1-3). Suitability of models to describe various factors affecting emergence of diseases can be assessed against the list of drivers leading to hot-spot appearance:

1. Location (country, region) and distance-based (using appropriate metrics)
2. Ecological (environmental conditions, populations)
3. Economic and trade (wet markets, commercial farms, backyard production)
4. Social (villages, periurban, urban)
5. Behavioural (travel, risk-awareness, changes in risk attitude)

The first two drivers have a long history of inclusion into modelling of zoonotic diseases. The remaining three, however, have largely been so far neglected. Conventional epidemiological models view human behavioural responses as external to the disease system. Social and economic approaches focus on these but often neglect the spatial, heterogeneous, stochastic, multi-host and multi-species nature of disease transmission (1-3). Lloyd-Smith et al. in their recent review (3) found that out of 442 modelling studies, only 4% included any economic analysis, despite claiming to deliver the policy implications of public health controls (4-7). Even fewer models embrace marketing (8) or social and behavioural (9, 10) aspects of zoonotic disease or consider the behaviour of individuals (public, farmers, policy-makers) (11, 12). Studies still typically neglect the socio-economic factors that promote particular industrial structures or approaches to livestock management (13). Our analysis corroborates these conclusions for more recent papers in years 2008-2011 i.e. after (3) was published (see below).
Review methods and literature

Literature was searched using Thompson ISI Web of Knowledge and Google Scholar. We pursued four lines of research into literature:

1. We first searched for all papers referring to ‘hot-spots’ and developing countries. This search resulted in only very few papers. We also searched for ‘network’, ‘agent based’ and ‘individual based’ models applied to epidemiology.

2. We repeated Lloyd-Smith et al. (2009) literature search for papers using models to address zoonotic diseases for years 2005-2011. As we expected that most advances in modelling would be available for influenza, we limited search for Topic=(influenza OR HPAI OR flu) leading to 1774 papers (ISI WoK now lists PLoS journals so we did not perform a separate search). This is broken down over time into: 2005 – 114, 2006 – 197, 2007 – 264, 2008 – 244, 2009 – 392, 2010 – 489, 2011 – 74 and reflects a rapid growth noticed by Lloyd-Smith et al. (2009). For the present review we only used papers from 2008 onwards, as most advances in agent-based modelling and social networks have been achieved since then. Full analysis will be presented in the review paper.

   Title=(influenza OR HPAI OR flu OR AVI) AND Topic=(model* OR dynamic* OR simulat*) AND Topic=(mathematic* OR stochastic* OR determinis* OR compartmental OR transmission OR reproducti* OR R0 OR reservoir OR estimat* OR sensitivit* OR epidemi* OR endemi* OR epizooti* OR enzooti* OR spillover OR cross-species OR zoono* OR sylva*) AND Year Published=(2005-2011)

3. We traced all papers that cite a number of key review papers on modelling zoonotic diseases, particularly Ferguson et al. (2005), Woolhouse & Gaunt (2007), Jones et al. (2008) and Lloyd-Smith et al. (2009) (see references: (2, 3, 14, 15)). The search produced 775 records (with significant overlap).
4. We also added papers from our own publication database collected for research and grant applications (particularly for the ESEI initiative).

5. For selected papers we also performed snowball search by looking at papers citing them and in some cases we followed the citation path further.

6. From this large set of potentially interesting papers, we selected a number of papers we considered exemplary in their approach and relevant for the present review. We particularly looked for research papers using network, social network and agent-based modelling, papers addressing economic and behaviour aspects of emergence, spread and control of zoonotic diseases, and papers related to disease and their socio-economic drivers in developing countries (mainly Asia). We also included a number of review papers comparing different modelling techniques.

The list currently contains 373 papers and is publicly available via a searchable database on Mendeley (http://www.mendeley.com/). Mendeley also allows sharing PDF files via a separate restricted group (by invitation only) – PDF files are available for most papers on the list. Access to the restricted group can be requested from the authors.

**Overview of modelling techniques**

In addition to standard texts, e.g. (16, 17), there has recently been a number of excellent modelling reviews, see (18-20). The following is a brief introduction to various modelling techniques, illustrated by a selection of papers relevant for this review (this illustration is not comprehensive and we only included a small selection).

Epidemiological models are generally built on the SIR paradigm (whereby the population is divided into compartments, S for Susceptible, I for Infected and R for Recovered/removed) (16). Two broad groups of models have been used, deterministic (assuming no chance events) and stochastic (with chance events), with further subdivision of each group into continuous-time and discrete time models (21).
Independently of the details of the model, contact structure is its key element, quantifying contacts between individuals – a ‘who-contacts-whom’ term. Once contact between a healthy individual and an infected one is established, infectious agent can be passed on with a certain probability. The contact structure can be modelled in different ways, based on different assumptions about the dynamic processes surrounding infection (cf. a recent review by Koopman (19)):

The simplest model assumption follows the Kermack-McKendrick approach via chemical kinetics – probability of a contact is proportional to the product of densities of healthy and infectious individuals. The underlying assumption of an ideal mixing is clearly violated by most epidemic processes – despite this, the standard SIR model has been extremely successful (22, 23) and is still used in different versions, some described by an interacting set of more than 1000 differential equations, so including much detail (24, 25). SIR models without spatial structure are still used in epidemiology and even more widely in the context of economic analysis (e.g. (26-35)).

Diffusion equations have originally been used to describe spatial spread of epidemics (see e.g. (36)) and more recently in relation to agent-based modelling (37, 38).

Metapopulation models collect individuals in loosely interacting subgroups, assuming high levels of interaction within each subgroup (see e.g. (39) for a metapopulation modelling linked to economics and (40) for a paper on zoonotic diseases). Household models (41-44) are good examples of metapopulation models, and the technique is widely used in ecology.

Gravity and distance-based models (45, 46) assume that intensity of contacts is inversely related to distance between individuals or groups of individuals (47-54). This approach is often extended to integro-differential equations (55). Both metapopulation and gravity/distance models can be either deterministic or stochastic and can either describe populations in terms of ‘densities’ or ‘individuals’. Network models are inherently ‘individual-based’, although ‘individuals’ can represent groups of animals/humans.

Network models represent ‘who-contacts-whom’ patterns by assuming – in the simplest case – that each individual has a fixed number of links to other individuals and those contacts can lead to
disease transmission at a constant rate (19). Network structure can be simple or complicated (e.g. hierarchical), regular, random or mixed (56-59), static or dynamic (60) or even dependent on the state of individuals (adaptive networks, (61-64)). Many models use arbitrary networks – often regular (square or triangular), random (54, 64-72) or scale-free (66, 73). Advances in computing and data collection allow using realistic network patterns, often based on social networks and allowing modelling real systems in unprecedented detail (14, 18, 40, 50, 56, 59, 74-116).

As network models are often analytically intractable (but see (117-119)), many approximation techniques have been proposed, particularly moment closures; see e.g. (57, 58). Typically they can be more readily designed and parameterised than network models, though at a cost of losing detail and accuracy.

Agent-based models attempt to realistically capture aspects of individual behaviour (86, 120-123). Rules for actions and interactions of autonomous agents are associated with each individual (which could be an animal/human being or a group of animals/humans) combined with simulations of simultaneous operations and interactions of multiple agents are used to re-create and predict the emergence and spread of infectious diseases. Most agent-based models include specification of: (i) agents specified at various scales; (ii) decision-making rules (behaviour of agents); (iii) learning rules or adaptive processes; (iv) an interaction topology (either non-spatial or spatial); and (v) an environment. Given enough good-quality data this approach can represent a ‘real’ system in which it is possible to identify each individual (124). Agent-based models are increasingly being used to describe spread of infections, see e.g. (14, 47, 48, 51, 53, 76, 79, 86, 101, 108-110, 113, 120-139). Lattice-gas approximations are similar to agent-based models, but movement and interactions might not be motivated by realism but rather by numerical schemes (140, 141).

Parameter estimation forms an essential part of modelling and recent years have brought significant advances through MCMC and Bayesian techniques – for an excellent recent review see (20). Parameter values can be found in literature or they can be estimated from various data sets; however they often come from sources with varying reliability. In assessing the suitability of models, we need to look at trade-offs between realism, accuracy and specificity of models as well as data requirements for parameter estimation. What is possible in terms of data collection in the developed world (USA, Europe) might be very difficult in South East Asia.
Incorporation of uncertainty from various factors increases the complexity of the modelling process and requires advanced methodological tools; both precise inference using MCMC and alternatives such as Approximate Bayesian Computation techniques are currently used (142, 143). The richness of inputs from multiple sources necessitates model assessment and selection; this is not straightforward in epidemics, and methodology involving latent tests in a Bayesian framework (144) can be explored. Epidemic case-reporting behaviour (as shaped by economic and social factors, e.g. (145)) has an impact on collected data and resulting estimation. Uncertainty in parameter estimation and in model structure needs to be reflected in uncertainty about model outcomes, but particularly in the case of model structure, this step is rarely taken (28, 141).

Once parameters are estimated, analytical or numerical methods can be used to make predictions from the model and test them against the data. Often, the estimation and testing are done in-sample, i.e. testing is done on the same data set from which parameters have been obtained. In-sample inference is notoriously unreliable as it often produces models that fit data spuriously well (something that economists are well aware of, see (146)). However, out-of-sample testing is very rarely done and it is not clear how we should approach differences in parameters and model structure between different epidemic outbreaks (147, 148).

The model structure and parameters are unlikely to be known exactly. This leads to the following key questions: (i) how wrong is the model? and (ii) does it matter? This clearly has important implications for predictions and yet this problem is very rarely addressed (28). Wearing et al. (149), for example, show how unrealistic model assumptions systematically impair our modelling and prediction ability. Such papers are, unfortunately, rare.

Economic modelling is only rarely linked to epidemiology (150-153). Most papers research cost-effectiveness of intervention strategies and are based either on direct calculations of costs (34, 154-157) or on using large-scale economic models (157-162). Individual decisions are described by maximising utility (with or without discounting) but real options approach (where utility is evaluated based upon future dynamics of a model, thus taking into account uncertainty about outcomes) is only rarely used (163). A game theoretical approach has only recently made its way into epidemiology, largely in the context of voluntary vaccination (151, 153, 164, 165).
Evolutionary game theory has also been applied to changes in behaviour during epidemic outbreaks (35, 63, 166, 167). Farm economics and implications for disease emergence and persistence are also addressed, but rarely linked to epidemiological modelling (168-170). Epidemiological problems are often framed in the absence of perspectives from those with relevant expertise such as farming communities, constraining the design and successful implementation of viable solutions in the framing of epidemiological problems (171-173).

Results

Papers in the database were assigned different tags representing a number of concepts we identified in the literature analysis. This allows readers of this report to quickly find papers relating to a particular subject by using Mendeley tagging system. The sample is not random and hence the proportions below reflect partly our selection of papers (with the sample biased in favour of developing countries, network- and agent-based models and influenza).

There were broadly nine modelling goals, although some papers aimed at more than one goal:

a. Biological studies (analysis of epidemiological and environmental processes and drivers): 23%
b. Behaviour and social aspects linked to disease emergence and spread: 10%
c. Analysis of economic aspects of disease outbreaks and interventions (we did not include a large number of papers describing cost-effectiveness of vaccination strategies): 20%
d. Model development (developing modelling techniques, theoretical analysis of models): 21%
e. Model comparison (comparing and using different modelling techniques): 13%
f. Description of a past outbreak with view to identify processes and parameters: 12%
g. Prediction of future outbreaks (often in a context of analysing scenarios): 18%
h. Sensitivity analysis to identify criteria leading to enhanced risk; not only by modelling (hot-spots): 10%
i. Assessment of interventions and intervention policies (any aspect related to culling, vaccination, design, sensitivity, costs): 29%
A number of modelling techniques have been used to achieve those tasks:

A. Data analysis (statistical model; we only selected a small number of all papers, largely those relevant for hot-spot identification and for developing countries): 16%
B. Simple models (i.e. compartmental, deterministic or stochastic SIR-class models without explicit spatial structure): 12%
C. Metapopulation models (including household models): 8%
D. Networks (general, not specifically driven by data): 17%
E. Social networks (driven by data): 15%
F. Adaptive and dynamic networks: 2%
G. Agent-based models: 10%
H. Economic analysis: 12%

Discussion

1. We found no social network or agent-based modelling literature on characteristics of livestock systems that increase risk of zoonoses hotspot developments in developing countries.
2. Such characteristics are identified in a number of regression/statistical models (which we have, however, largely excluded from the review). Thus, identification of hotspots is currently addressed statistically rather than by dynamic modelling (174-184), although more advanced modelling techniques have also been used (83, 128, 185, 186).
3. While there seems to be very little social network or agent-based modelling on zoonotic outbreak risk in developing countries, there are a number of papers that successfully use network and/or agent-based modelling to predict disease dynamics and to research the effectiveness of containment and control. For example, social networks can be used to identify highly connected hubs and network components that are potentially sources of infection (see e.g. (14, 77, 81, 84, 88, 90, 101, 110, 111). Although large and detailed data sets are needed to parameterise such models, progress recently has been made in developing countries in this context (78, 83, 91).
4. A number of papers research the role of human behaviour in this context (53, 61, 63, 99, 110, 130, 166, 187, 188). Those authors stress importance to include behaviour in modelling approaches as it has strong impact on model conclusions.
5. Economic modelling is almost exclusively used to research cost-effectiveness of stockpiling, vaccination strategies and optimal resource allocation (we have not included these papers in the overview as there’s such a large number of them and it’s not exactly related to our subject). Since 2003 there has been a number of studies linking epidemiological modelling and (micro-)economics, see e.g. (29, 32, 33, 104, 125, 129, 139, 141, 150-156, 159-163, 165, 167, 168, 189-203), although significant progress has only been achieved since 2009 (33, 141, 150, 151, 159, 163). Shifts in production systems, marketing aspects of zoonotic diseases and analysis of food chain and consumer preferences has not even been considered as part of epidemiological modelling (204).

6. Uncertainty about important biological, social and economic parameters – or even model structure – is very important and yet often ignored, which hampers outbreak predictions (28, 34, 147, 197, 205, 206). Similar uncertainty prevails over human behaviour, structure of contact networks etc. Detailed data are difficult to obtain, but essential e.g. for social network models to be applicable. Under-reporting is notorious for many diseases, including influenza, yet very little modelling/statistical work has been done to address relevance of the reporting bias (145).

Conclusions

1. Most models of zoonotic diseases are currently capturing outbreaks over relatively short time and largely ignoring socio-economic drivers leading to pathogen emergence, spill-over and spread.

2. To study long-term changes we need to understand how socio-economic and climatic changes affect structure of livestock production and how these in turn affect disease emergence and spread.

3. Models capable of describing this processes do not appear to exist, although some progress has been made in linking social and economical aspects of livestock production (see e.g. (125)) and in linking economics to disease dynamics (see e.g. (150, 151, 168, 200)). Although initially progress is likely to be made for simple models (as in e.g. (33, 150, 151)), linking socio-economic agent-based models as in (125) to disease spread is potentially a promising avenue (120).

4. A new modelling framework is required that expands and formalises the ‘one world, one health’ strategy (207, 208), enabling its deployment in the re-thinking of
prevention and control strategies. It is at present unclear what form those models are likely to take, but social networks, dynamic (adaptive) networks and agent-based models are good candidates.

5. **Modelling can only provide means to identify risks associated with socio-economic changes but can never be a substitute for data collection.** The choice of the model is likely to be determined by a trade-off between data availability and modelling goals.

6. **Uncertainty analysis and uncertainty communication form a key element of modelling process and yet are rarely addressed.**
References

(only papers directly referred to in the report are listed here)


