Approaches for evaluating veterinary epidemiological models: verification, validation and limitations

A. Reeves*, M.D. Salman & A.E. Hill

Animal Population Health Institute, Department of Clinical Sciences, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO 80523, United States of America

*Corresponding author: Aaron.Reeves@colostate.edu

Summary

The evaluation of models of the spread and control of animal diseases is crucial if these models are to be used to inform decisions about the control or management of such diseases. Two key steps in the evaluation of epidemiological models are model verification and model validation. Verification is the demonstration that a computer-driven model is operating correctly, and conforms to its intended design. Validation refers to the process of determining how well a model corresponds to the system that it is intended to represent. For a veterinary epidemiological model, validation would address such issues as how well the model represents the dynamics of the disease in question in the population to which this model is applied, and how well the model represents the application of different measures for disease control.

Just as the development of epidemiological models is a subjective, continuous process, subject to change and refinement, so too is the evaluation of models. The purpose of model evaluation is not to demonstrate that a model is a 'true' or 'accurate' representation of a system, but to subject it to sufficient scrutiny so that it may be used with an appropriate degree of confidence to aid decision-making.

To facilitate model verification and validation, epidemiological modellers should clearly state the purpose, assumptions and limitations of a model; provide a detailed description of the conceptual model; document those steps already taken to test the model; and thoroughly describe the data sources and the process used to produce model input parameters from those data.

Keywords

Evaluation of models – Model credibility – Model validation – Model verification – Verification of models.

Introduction

Computer-driven epidemiological modelling is an increasingly common technique for assessing the potential consequences and possible spread of animal diseases. Modelling of animal diseases has been used to estimate the possible magnitude of an outbreak and the resources needed for a response, and to inform policy decisions on measures for disease control (4, 6, 14, 15, 17, 28, 29, 47, 58). Epidemiological models may take several forms. Some are based on analytical formulas that describe the system of interest in a rigorously mathematical way (14, 15, 28, 29, 63). Others employ computer-driven simulation to mimic the actual mechanistic processes at work within a system (5, 16, 24).

Regardless of their form, all models – especially models which are intended for use by response planners and policy-makers – require careful evaluation. For models to be effectively used in these instances, a sufficiently high level of credibility of the model and its results must be achieved so that decision-makers and other stakeholders can have a justifiable degree of confidence in their application. By the same token, careful evaluation of models can identify and clarify their limitations and weaknesses, temper tendencies toward over-reliance on apparently 'objective' model-produced outcomes, and minimise their misapplication.

Methods for model evaluation are quite diverse; as several authors have noted, there is no single standard or approach that can be applied to all models (32, 41). At a very basic level, as the mathematical or computational complexity of epidemiological models increases, it is essential to demonstrate that the mathematical framework or software used for a model is free from major errors which would threaten the accuracy of the calculations that the model produces. Some approaches for evaluating models are, by necessity, qualitative. Any assessment of the conceptual quality of a model, for example, is fundamentally qualitative in nature. In some instances, it may be possible to use quantitative or statistical approaches to demonstrate correspondence between a model and a natural system, although the use of such quantitative methodologies does not necessarily ensure that a model is conceptually sound.

The aim of this paper is to describe approaches for evaluating epidemiological models intended to inform management or policy decisions on animal diseases, with an emphasis on two approaches that have been called 'verification' and 'validation'. The authors' specific objectives are as follows:

– to briefly define and describe the processes of model verification and validation

– to discuss several approaches used to address the challenging issue of validating epidemiological models intended to inform emergency response plans

– to illustrate practical approaches to model verification and validation, based on the authors' experiences as members of the research team behind the North American Animal Disease Spread Model (NAADSM) (24)

– and, finally, to present a set of suggestions for steps that could be taken to improve the credibility and acceptance of epidemiological models for the management of animal diseases.

Model context, development and evaluation

Figure 1 illustrates a conceptual series of steps in the process of model development and application. Several of these steps deal explicitly with the evaluation of models, but almost every stage in the figure implies some form of appraisal of the model under development. Decisions made at the outset of model development about the specific purpose of a model and the questions which it is being designed to answer will affect the ways in which the model's utility and credibility are assessed.

First and foremost, models must be evaluated in the context of the problems that they are intended to answer (35, 43, 53). The criteria for judging a model that is intended to inform broad questions in a qualitative way will be quite different from those used to evaluate a model that claims to offer specific predictive capabilities.

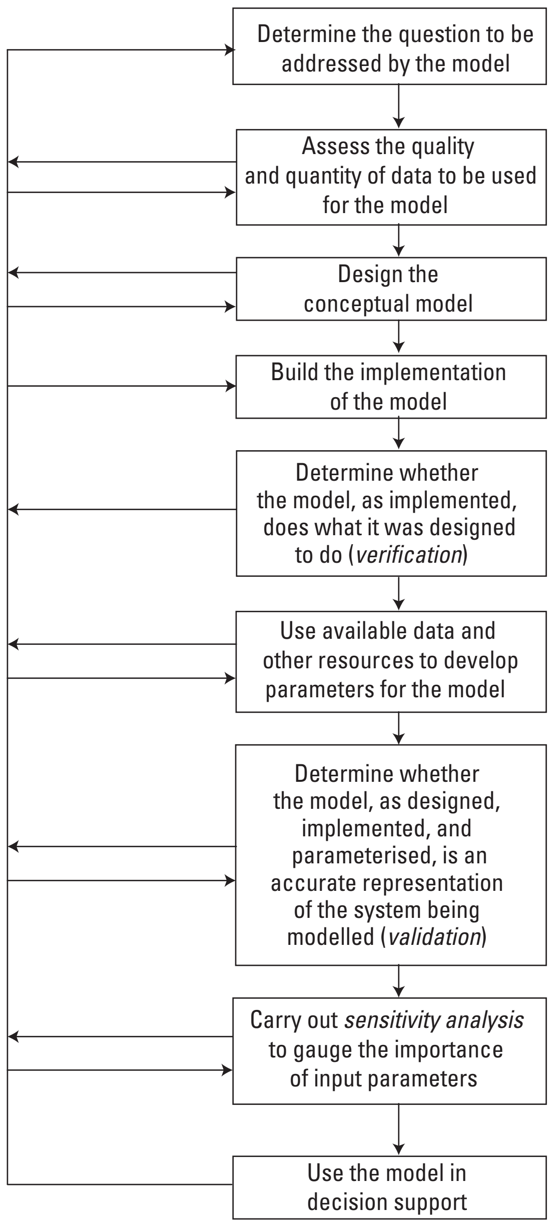

Fig. 1
Schematic diagram of the stages of model development, evaluation and application
Adapted from Dent and Blackie (8), Martin *et al.* (38), and Taylor (62)

Secondly, for results of modelling investigations to be credible, the models must be built upon reliable data (51, 62). Models based on incomplete or theoretical input data may yield useful hypotheses for further research and evaluation, but the limitations of such models should be clearly and expressly stated. The more complete the input data for a model are, the more likely it is that the model's output will be credible.

Thirdly, just as the conceptual development of models is, in many respects, a subjective undertaking, so too is the evaluation of models. Individual modellers must weigh the relative importance of different aspects of epidemiological systems, and may come to different conclusions about how to represent various processes in their models, or even about which processes to represent. Any assessment of the credibility of a model must consider these subjective design decisions.

Fourthly, Figure 1 makes the distinction between a conceptual model or model framework, and a specific model that applies a particular conceptual framework, together with a particular data set or set of parameter values, to represent a specific situation. The North American Animal Disease Spread Model, for example, is a framework for the development of epidemiological simulation models, which has been used to build specific models of a variety of diseases in different settings and populations, such as foot and mouth disease (FMD) (46, 66), Aujezsky's disease (pseudorabies) (47), and highly pathogenic avian influenza (HPAI) (45), among others. Both the conceptual framework and the particular instances in which the framework are used need to be evaluated. The utility of the former does not necessarily rely upon the latter, but the quality of specific models is highly dependent on both the conceptual framework and the data used for their construction.

Finally, Figure 1 illustrates that the process of model development and evaluation is cyclical and iterative. Evaluation is not a single, discrete step, and 'is not something to be attempted after the simulation model has already been developed, and only if there is time and money remaining' (35). Model evaluation should instead be considered ongoing: model assumptions should be reassessed continually as new sources of information become available.

The assessment of the computational correctness of a model has been called 'verification'. Verification deals with questions such as: 'Does the computer program perform all calculations correctly?' and: 'Does the program match exactly what the designers intended?' The assessment of how well a model conforms to or exemplifies the system that it is intended to represent is sometimes referred to as 'validation' (32, 53, 57). Validation is intended to address the question, 'Is a model an adequate representation of the real system?' (For the remainder of this paper, the authors will follow these definitions for 'verification' and 'validation', but note that these definitions are not universally applied. For example, Oreskes *et al.* [42] use the terms 'verification' and 'validation' to denote somewhat different concepts.) Together, verification and validation efforts can help investigators to ascertain the overall quality and credibility of a model.

Model verification

Model verification refers to the process of determining whether the model, as implemented in software, conforms to the desired conceptual model (53). In other words, verification provides an assessment of whether the software implementation of the model is working correctly. Among the criteria by which a model's verification status might be assessed are its correctness (the 'extent to which a model meets its specifications') and its reliability (the 'extent to which a model can be expected to perform its intended function with required precision') (40, 59). Any model used for scientific research or for decision support should be expected to meet a high standard for such characteristics.

Model verification, although straightforward in concept, can be time-consuming, particularly as models become more complex. Sargent (53) and Scheller *et al.* (56) present useful discussions of some of the software engineering practices that can facilitate the construction of verified models, particularly for larger projects, and several authors have provided detailed descriptions of approaches to verification (33, 70). In this paper, the authors focus on two central aspects of model verification that have a direct impact upon the credibility of epidemiological models regardless of their form, size or scope: producing documentation that describes the conceptual model in detail, and thorough testing to ensure that the model is performing as intended.

Describing the conceptual model

As shown in Figure 1, designing the conceptual model is an early stage in model development. There is a great deal of value in explicitly documenting this conceptual model. Such documentation can be used to assess the conceptual validity of the model (see below), but, at a more basic level, it can provide a standard by which the correctness of a model can be judged (33, 56). The purpose of a written model specification is to describe, in clear, accessible language, the purpose, requirements and conceptual details of a model. The intended audience of such a document includes the modellers themselves, as well as any technical personnel who will be involved in implementing the model, among others (see 'Conceptual validity', below). The model specification can also provide a basis for model testing (23, 56).

In the case of NAADSM, the model specification document (23) describes every component of the modelling framework in detail: it is the authoritative source that describes how the conceptual model should operate, and is the standard by which the software implementation of the conceptual model is judged. Although the specification may be updated as needed, to correct ambiguities or to incorporate new features, the complete history of the specification is tracked, and every version is available for reference and evaluation by independent researchers (22, 23).

Model testing

Fairley (12) and Whitner and Balci (70) distinguish between two forms of model testing, which they refer to as 'static' and 'dynamic'. For simple models, static testing may be sufficient. This approach involves a structured examination of the formulas, algorithms and code used to implement a model, preferably by several reviewers who were not directly involved in writing the implementation themselves. Garner and Beckett (16) describe the use of this approach in the development of 'AusSpread', a simulation platform designed initially to model the spread and mitigation of FMD.

For more complex models, dynamic testing is often useful. During dynamic testing, a computer program is run repeatedly under different conditions to ensure that the output it produces is correct, according to the conceptual model, and consistent with expectations. Often, such tests are established to be run repeatedly and automatically, to ensure that any changes to the software implementation did not inadvertently introduce errors; this process is referred to as regression testing. Scheller *et al.* (56) describe several levels of testing, from simple unit tests that evaluate specific, individual functions; to broader system testing that assesses the interaction of all of the components of a model. The authors will illustrate these approaches in the following sections, with examples from the development of NAADSM.

Automated software testing of the North American Animal Disease Spread Model framework

To ensure that the NAADSM application correctly implements the conceptual model specification, NAADSM relies upon an automated regression-testing approach. Simple models have been constructed to test every aspect of the NAADSM application. There are currently well over 1,000 individual models in this suite of tests, and new tests are continually being developed. When the NAADSM application is compiled from program source code, every test is automatically run and results are tracked using a freely available framework for software testing (55). Prior to the public release of any new version of NAADSM, every test in the suite must be passed. Every simple model developed for testing is published, along with the complete source code for the NAADSM application.

Manual testing of the North American Animal Disease Spread Model

In addition to the automated use of simple tests, manual testing using more complex situations has been carried out for the NAADSM framework. Every aspect of the model framework is examined by analysts working independently of the programmers to confirm that the model conforms to the published specification. Any errors identified during manual testing are noted and must be corrected before public release.

The limitations of model verification

Model verification procedures can be quite objective and thorough. Many techniques developed in the field of software engineering can be rigorously applied to the programming of models (7, 56). Model verification offers no answer, however, to the crucial questions: 'Is the model useful?' and 'Is the model adequate for the purposes for which it was designed?' Questions like these can be addressed by a variety of approaches that fall under the general heading of 'model validation'.

Model validation

Validation refers to the process of determining whether a model is an acceptable representation of the system that it is intended to represent, given the purpose of the model or study (35, 53). A more elaborate definition is provided by Schlesinger (57): model validation is the 'substantiation that a ... model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model'. It is important to note that 'acceptable representation' in the definition above does not mean an 'accurate' or a 'true' representation: Oreskes *et al.* (42) convincingly argued that it is impossible to establish whether any particular model is an accurate representation of a natural system, and that the use of the term 'validation' in this sense is highly misleading.

The problem of model validation

In contrast to the process of model verification, establishing the validity of models is not clear cut, and can be quite problematic. As McCarl (41) observed, 'there is not, and never will be, a totally objective and accepted approach to model validation'. The standards by which a model's validation is judged are partly dependent upon the purpose of the model. The validation of models designed strictly to address research questions (for example, to generate and test hypotheses concerning population or disease dynamics or to identify new areas of research) does not have to be as stringent as the evaluation of models that will be used to inform operational management decisions. When such decisions are made on the basis of the results of modelling studies, it is important to know that these studies are appropriate, accurate and correct. Given the difficulties associated with the study of very complex multifactorial problems, the subjective elements of modelling itself, and philosophical issues like those presented by Oreskes *et al.* (42), the threshold for accepting a model cannot be 'proof' of its accuracy or validity. Rather, this threshold should be that of reasonable confidence in the results produced by the model. As Holling (25) stated, 'provisional acceptance of any model implies not certainty, but rather a sufficient degree of belief to justify further action'. The task of model validation, as described here, is that of evaluating models in order to have a justifiable level of confidence in their results before they influence policy or management decisions.

It is often constructive to think of a model in a similar way to a scientific hypothesis. An epidemiological model, for example, represents the modellers' hypotheses about the interactions among members of a population, the dynamics of disease in that population, the mechanisms of disease spread, and the efficacy of different disease control measures. As with any hypothesis, models should be tested and challenged. As models are subjected to and withstand increasing levels of scrutiny in diverse situations, their credibility is increased. Such models can then be applied to problems of management and policy with greater confidence, provided that it is always clearly understood that no model truly represents physical reality, and that the acceptance of any model must be subject to continuing evaluation.

What follows is not a set of methods that will prove that a model represents a real system, but rather a set of activities that might be undertaken to provide evidence which may either support or refute the hypothesis presented by a model. Several authors present descriptions and detailed taxonomies of the methods used to assess model validity (33, 34, 51, 53). The authors' intention, in the following sections, is to present and discuss the usefulness of some of these methods, together with examples of their application, both from the authors' own experiences and from other published reports of animal disease modelling. The authors also refer readers to several excellent discussions of model validation, including those presented by Oreskes *et al.* (42), Rykiel (51) and Taylor (62).

Conceptual validity

A particularly useful – and a foundational – criterion for the validation of an epidemiological model is the answer to the question, 'Does the structure of the model make logical and biological sense?' This has been referred to as 'conceptual validity' (51, 53). For a model to have conceptual validity, its theoretical underpinnings should be shown to be based on known and scientifically accepted properties of the system of interest, or at least on reasonable and justifiable assumptions about such properties. Among some of the questions that might be addressed in assessing the conceptual validity of a model are the following:

– does the model fit the purpose or purposes for which it was designed?

– does the structure of the model sufficiently capture the relationships and interactions among components of the system being modelled?

– given the purpose of the model, are key components of the system absent from the model, or oversimplified? Is additional detail necessary for any component?

– based on existing knowledge and experience, are the outcomes produced by the model reasonable?

Review by independent experts on the subject matter concerned – sometimes referred to as establishing 'face validity' (51) – can be used as a means of assessment. In this case, it is quite helpful to have a detailed document that describes the conceptual model, as noted earlier. Such a document can provide a basis for discussion and evaluation of the details of the model's operation. The publication of model descriptions (5, 24, 26, 61) greatly facilitates the assessment of the conceptual validity of models.

Reliance on the peer-reviewed literature provides one avenue for the conceptual assessment of epidemiological models. The NAADSM Development Team has also taken a more direct approach and sponsored a series of meetings of subject-matter experts, including epidemiologists, virologists, economists, policy-makers and other modellers, to review the NAADSM modelling framework (10, 64, 65). The structure and assumptions of the modelling platform have been described in detail during these workshops, and discussion, suggestions and advice are solicited from all participants. The results of these expert panel evaluations are then used to guide future research and development.

The use of data in model validation

As noted in 'Model context, development and evaluation', it is possible to assess the conceptual framework separately from the data used to inform a model. Empirical data are generally used in two ways during modelling:

– input data are used to develop parameters that will influence model outcomes

– data that represent the outcomes or results of a system (output data) are used to provide a basis for comparison with model-produced outcomes.

In a few cases, particularly for endemic disease situations, large amounts of both types of data may be available for models of disease spread in populations. In many instances, however, we have access to information pertaining to only a single outbreak of disease in a particular set of circumstances. Information collected during the 2001 outbreak of FMD in the United Kingdom (UK), which has been widely used for modelling studies (14, 15, 28, 29, 54), represents one such data set. In still other instances, models are developed to explore hypothetical situations (5, 6, 16, 45). In these cases, some information is generally established to inform model inputs, but there can be no data on the (non-existent) system outcomes.

Whatever the form or source of data used to inform models, their correctness and validity should also be considered. As Rykiel (51) points out, there is no guarantee that available data necessarily provide a better or more accurate depiction of a real system than a conceptual model. The process of ensuring so-called data validity (51, 53) can in itself be complex.

Several authors have emphasised the notion that, in order to demonstrate validity, models should be tested against data not used during their construction (30, 60). Green and Medley (20) indicated that such a step should be a requirement before a model is used to inform policy decisions. This is one of several possible approaches that fall into the general category of 'operational validation' (53).

Although this suggestion seems straightforward, its implementation for incompletely understood biological and epidemiological systems is problematic. First, it implies that reliable, valid data exist for at least two situations, for both the development of parameters and for comparison to actual system outcomes. Secondly, this approach would require the existence of a suitable means of evaluation by which the similarity of model-produced outcomes to system outputs can be assessed. Thirdly, it implies that these situations are sufficiently dissimilar from one another that they represent unique tests of a model, but are still similar enough that exactly the same approach to modelling developed for one situation can be legitimately applied to the others. The authors have already mentioned the first difficulty. The remaining two problems are discussed below.

A variety of quantitative, statistical approaches to show the correspondence between model-produced outputs and outcomes generated by biological systems have been devised and applied in a few situations (13, 32, 36, 39, 48, 49, 50, 69). Most of these approaches to what has been called statistical validation rely upon the existence of a large amount of data (i.e. many observations) pertaining to the outcome of the natural system, which limits their applicability to most situations of interest to animal disease modellers.

Waller *et al.* (69) proposed the use of Monte Carlo hypothesis tests, which, in essence, compare a single set of outcome data from a real system to multiple model-generated outcome data sets, and seek to answer the question, 'Do the observed data appear consistent with the model?' rather than the more typical question, 'Does the model appear consistent with the observed data?' Although this approach is not without value, it raises an additional question: how representative is any single outcome? When considering recent outbreaks of FMD in the UK, for example, is the 2001 outbreak, which resulted in the infection of over 2,000 herds (1), more or less representative than the 2007 outbreak, which produced only eight infected herds (2)? How 'consistent' would each of these two outcomes have to be with model-produced data to conclude affirmatively that the data are consistent with the model? Efforts to compare outcomes from epidemiological models to data generated by individual outbreaks should be undertaken with care: such comparisons are potentially informative, but an over-reliance on quantitative approaches for evaluation of models may well be misleading.

The disparity between these two recent FMD outbreaks in the UK also illustrates the third potential problem raised above: the dissimilarities among outbreaks of even the same diseases in generally the same types of populations make it difficult to test a model against data not used during its construction. As described in 'Model context, development and evaluation', the use of data is integral to model construction. Although the conceptual framework of a model and the data used to inform this model are distinct and can (and should) be evaluated individually, output generated by a model is inseparable from the combination of these two elements. The correspondence of model output to a natural system cannot be evaluated without considering the conceptual model and the source data simultaneously.

Validation of model components

Although it is difficult to demonstrate the validity of an entire model by the means described above, especially in the absence of relevant data, it may be possible to assess the validity of some individual components of a more complex model. This component-based approach to validation is sometimes recommended (38). An example is a recently completed validation of the process used in NAADSM to simulate animal movements and contacts among farm premises (9).

Briefly, the objective of this study was to validate the contact component used in NAADSM by comparing simulated movements to real-world, farm-to-farm movements that had been recorded for adult milking cows in Ontario, Canada. The study concluded that the approach used in NAADSM performed reasonably well in simulating average network characteristics observed in real-world movement data, but did not perform as well in simulating extreme upper percentiles of movement network components, involving rare but observed farms with excessively high shipment frequencies. The results of this study will be used to inform future development, with the objective of providing better representations of actual events and thus leading to greater confidence in the results of modelling studies.

Comparison of models

Comparison of the results from several independently developed models may be used to improve the level of confidence in the models tested. This process has been called 'relative validation' (11).

Dubé *et al*. (11) conducted a comparison of three simulation models using relatively simple disease scenarios. Among the findings of this comparison was that, although statistically significant differences were observed among model outputs, results from all three models supported the same or very similar conclusions on approaches for disease control. This finding could be used to increase the confidence of end users and decision-makers in modelling results (11). The results of a follow-up investigation that considered more complex scenarios are reported elsewhere in this issue (52).

Several similar comparisons of models of the spread and control of animal disease have also been undertaken. Vigre (68) reported on a comparison of mathematical and simulation-based models. The differences identified were more substantial than those reported by Dubé *et al*. (11), and may reflect the broader distinctions between the fundamental assumptions made by the individual models. Continued investigations in this vein would be quite helpful. Gloster *et al.* (19) also recently reported on the comparison of several models of airborne dispersion of FMD virus. Like Dubé *et al*. (11), they reported that the results of the models evaluated were broadly similar but, of course, highly dependent on the assumptions made and the data used by different groups of modellers.

Loehle (36) identified the comparison of models as a component of the larger process of what he called structural analysis, or an evaluation of the inherent assumptions and identification of the deficiencies of various models. Loehle argued that, because of the existence of such structural differences among models, and because comparing multiple models is the most effective way to identify and determine the effects of such differences, it is essential to direct multiple modelling efforts towards any important policy or management problem.

Sensitivity analysis as a form of validation

When data from real systems are limited, sensitivity analysis is sometimes suggested to inform model validation efforts (6, 27, 32). Sensitivity analysis is used to determine the amount of influence that particular parameters have on the outcome produced by a model. Sensitivity analysis can also be used to assess the conceptual validity of a model: if certain parameters are expected to be important in a system, based on prior knowledge of that system, then sensitivity analysis should bear out these expectations (32).

Of greater value is the use of sensitivity analysis to determine which parameters in a model are important. If a model includes parameters about which there is a high degree of uncertainty, but which are shown by sensitivity analysis to have a substantial impact on model results, such parameters are good targets for additional research. An example of applying such sensitivity analysis to animal disease modelling can be found elsewhere in this issue (44).

Suggestions for constructing useful, credible models of animal disease

As discussed in the preceding sections, the primary objective of model verification and validation is not to demonstrate that a model is a true or even a highly accurate representation of a real system, but rather to provide a set of approaches and criteria by which a model can be evaluated. For models that might be used as a partial basis for policy or management decisions, it is essential that such evaluation establishes a foundation of support and credibility. To that end, the authors suggest the following practical steps that members of the veterinary epidemiological community can take to produce credible, useful models of the spread and control of disease in animal populations. These suggestions are drawn from the authors' own experience, as well as from many of the other valuable sources cited throughout this article; in particular, those written by Bart (3), Rykiel (51), Law and McComas (35) and Sargent (53).

Clearly and precisely state the purpose for which a model was designed

The importance of the first step, illustrated in Figure 1: that of determining and then clearly and precisely stating the questions to be asked of a model, may seem self evident, but this step is often overlooked (3). Overton (43) remarked that: 'the great majority of criticisms of models relate to a capacity for which the model was not designed in the first place'. A clear understanding of the purpose of a model is a prerequisite for any further evaluation.

Provide a detailed description of the conceptual model and document the assumptions and limitations of the model

Virtually every paper on techniques for the verification and validation of models stresses the importance of documentation for the conceptual model (3, 33, 35, 53, 56). A model description should not be produced solely, or even primarily, for the developers of an individual model. Those who will derive the most benefit from the existence of such documents will be other model users, in the broadest sense of the term: other researchers, analysts and decision-makers, who will be expected to apply or evaluate the model and its results. Such documentation is particularly useful when it includes discussions of the model's assumptions and limitations, presented in ways that are clear and biologically relevant (21).

Provide details of the steps taken for model verification

At its most basic level, the credibility of a model relies upon the demonstration that the model, as implemented in software, does what it is supposed to do. Anyone asked to evaluate a model, particularly if it will be used to influence policy, should have access to a computational implementation of the model and details of the verification procedure employed, as well as to any tests used for verification, so that he or she can reproduce and evaluate the computational correctness of the model.

Describe the data used to develop model parameters and document the approaches and assumptions used to produce model parameters from these data

The process of translating raw data into parameters suitable for use in models is seldom straightforward. An understanding of this process, however, is essential if reviewers are to have an adequate basis for judging the model's results. Two recent reports illustrate this suggestion quite nicely: Mardones *et al.* (37) conducted a meta-analysis based on 21 research papers and documented in detail the procedures that they used to estimate the durations of different disease states for FMD. In a different study, Patyk *et al.* (45) produced a model of the spread and control of HPAI in South Carolina in the United States. This study included an online supplement that described in detail all the sources of information used for the study, as well as the computational tools that the authors developed and used for parameter development.

Involve independent experts in the evaluation of models and their outcomes

Veterinary epidemiological modelling is an interdisciplinary undertaking. Modellers can take advantage of a great deal of expertise in different fields by involving experts from these fields. For models to be used for decision-making, it is also essential to involve other stakeholders in this process; for example, those who are responsible for decision-making or for implementing policies in the field. In the authors' own experience with NAADSM, they have found that, through its widespread application, they have benefited substantially from the efforts of others to use and evaluate it.

A variety of forums have become available for sharing and discussing veterinary epidemiological modelling work over the last few years (10, 64, 65, 67). The authors encourage anyone involved with the construction, use or evaluation of models to seek out and take advantage of such opportunities when they occur.

When possible, use existing information for data-driven validation of models or their components

The authors have discussed the limitations and advantages of this approach in 'The use of data in model validation' and 'Validation of model components' above. Such approaches should be undertaken with care, and with the recognition that the results will not be definitive: a poor conceptual model may still produce a good fit to observed data and *vice versa*. In situations where appropriate information is available, however, the comparison of model-produced outcomes to real data can still be enlightening. Retrospective analysis of past outbreaks is crucial to understanding them, and modelling can be a very useful tool in this pursuit (18, 31).

Present a range of possible outcomes, including 'best-case' and 'worst-case' scenarios

As discussed above, models are not definitive representations of reality. We are often uncertain about the ways in which at least some components of our systems operate, and also about specific parameter values. Presenting a range of results is one way to capture some of this uncertainty.

Use sensitivity analysis to determine the importance of the parameters used in a model

In addition to the benefits discussed in 'Sensitivity analysis as a form of validation', evaluating the importance of model parameters – especially those for which data are limited – can be used to estimate the potential effects of parameters about which the modellers are uncertain.

Compare the purposes, conceptual bases and outcomes of various models

During the modelling process, different modellers make different subjective decisions and assumptions. Qualitative agreement among several models may lend credibility to the conclusions drawn from model-based studies. Areas of disagreement among models should prompt additional research and investigation to improve our level of understanding of the system components in question.

Finally, treat model evaluation as an ongoing process, not as settled fact

Every epidemiological model is a work in progress, informed and updated by existing and new knowledge about the dynamics of disease; changes in agricultural and social practices; and changes in the forms, sources and quality of available data. The validity of any epidemiological model should be continually reassessed under new conditions or as the state of our knowledge improves.

Conclusions

The careful evaluation of any model intended to inform management or policy decisions is an essential activity. Two key steps in assessing the quality and usefulness of epidemiological models are verification and validation. Unfortunately, there are no purely quantitative, strictly objective means by which to evaluate models. Each model, and each situation to which modelling will be applied, is unique, and unique means may be necessary to evaluate a model and its particular applications. Holling (25) pointed out that, 'provisional acceptance of any model implies not certainty, but rather a sufficient degree of belief to justify further action'. The authors have outlined a set of recommendations that can be used by epidemiological modellers to cultivate confidence in applying this technique to important problems in animal population health. Individual models will continue to be developed and compared, and will evolve as they are scrutinised. Through these exercises, our collective aim of providing useful tools to assist in decision-making processes can be met.

To achieve a sufficient level of credibility for model outcomes, it is essential not to involve solely modellers in their evaluation. As Rykiel (51) observed, 'to the extent that a model is a scientific experiment and theoretical development, its testing and validation are within the purview of the scientific community'. The authors agree, and would add that, in the case of models for animal diseases, the evaluation of models is also within the purview of field epidemiologists and veterinary practitioners, policy planners and decision-makers, and animal industry representatives.

Acknowledgements

The authors would like to thank Ric Hupalo and Marian Talbert for their contributions to the literature search that accompanied the writing of this paper. They would also like to thank the many participants in their modelling courses and workshops who have made valuable contributions to discussions of model evaluation and validation. Partial financial support for this work was provided by the Colorado State University Program for Economically Important Infectious Animal Diseases, which is funded by a special grant from the US Department of Agriculture National Institute of Food and Agriculture. This work was conducted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Clinical Sciences from Colorado State University (AR).

■

Méthodes d'évaluation des modèles épidémiologiques vétérinaires : vérification, validation et examen des contraintes

A. Reeves, Mo D. Salman & A.E. Hill

Résumé

Il est indispensable de procéder à l'évaluation des modèles de propagation et de prophylaxie des maladies animales avant de les utiliser pour étayer les décisions concernant la lutte contre les maladies animales ou leur gestion. L'évaluation des modèles épidémiologiques comporte deux étapes importantes, la vérification et la validation. La vérification consiste à démontrer qu'un modèle issu d'un programme informatique fonctionne correctement et qu'il se comporte conformément aux objectifs visés lors de sa conception. La validation est le processus consistant à déterminer si le modèle est approprié par rapport au système qu'il est censé représenter. La validation d'un modèle épidémiologique vétérinaire porte sur des aspects tels que, par exemple, la justesse des représentations fournies par le modèle de la cinétique de la maladie étudiée dans une population donnée, ou de l'application des différentes mesures de prophylaxie envisageables.

L'évaluation des modèles épidémiologiques, tout comme leur développement constituent un processus subjectif et continu, voué à connaître des phases de changement et de perfectionnement. L'évaluation d'un modèle n'a pas pour but de démontrer que celui-ci fournit une représentation « vraie » ou « juste » d'un système, mais de le soumettre à un examen minutieux afin qu'il puisse être utilisé, avec un niveau de confiance approprié, en appui de la prise de décisions. Afin de faciliter la vérification et la validation des modèles, les concepteurs des modèles épidémiologiques devraient indiquer clairement les objectifs, hypothèses de départ et contraintes de ces modèles, fournir un descriptif détaillé de leur cadre conceptuel, indiquer les phases de validation déjà entreprises pour tester le modèle et exposer précisément les sources d'où procèdent les données ainsi que les procédures employées pour configurer les paramètres d'entrée à partir de ces données.

Mots-clés

Crédibilité d'un modèle – Évaluation d'un modèle – Validation d'un modèle – Vérification d'un modèle.

■

Métodos de evaluación de los modelos epidemiológicos veterinarios: verificación, validación y limitaciones

A. Reeves, Mo D. Salman & A.E. Hill

Resumen

Para que los modelos de propagación y control de enfermedades animales aporten elementos de utilidad a la hora de tomar decisiones en la lucha contra esas enfermedades es imprescindible evaluarlos previamente. En el proceso de evaluación de un modelo epidemiológico hay dos etapas básicas: la verificación y la validación del modelo. La verificación consiste en demostrar que un modelo informatizado funciona correctamente y conforme al objetivo con que fue concebido. La validación es el proceso por el que se determina hasta qué punto

un modelo se corresponde con el sistema que en principio representa. En el caso de un modelo epidemiológico veterinario, para validarlo se estudian por ejemplo el grado de exactitud con que da cuenta de la dinámica de la enfermedad en cuestión dentro de la población a la que se aplica el modelo y el grado de fidelidad con que éste refleja la aplicación de distintas medidas de lucha zoosanitaria.

Al igual que la elaboración de modelos epidemiológicos es un proceso subjetivo, continuo y sometido a cambios y perfeccionamientos, otro tanto cabe decir de la evaluación de esos modelos, que no tiene por finalidad demostrar que un modelo es una representación “verdadera” o “exacta” de un sistema, sino más bien someterlo a un análisis lo bastante preciso como para que los responsables de adoptar decisiones puedan utilizarlo con un adecuado nivel de confianza.

Para facilitar la verificación y validación de modelos epidemiológicos, sus conceptores deben dejar claramente sentadas las finalidades, premisas y limitaciones del modelo, proporcionar una detallada descripción de su marco teórico, exponer los procesos ya instituidos para ponerlo a prueba y especificar en detalle las fuentes de datos empleadas y el proceso seguido para generar los parámetros de entrada a partir de esos datos.

Palabras clave
Credibilidad de los modelos – Evaluación de modelos – Validación de modelos – Verificación de modelos.

■

References

1. Anderson I. (2002). – Foot and mouth disease 2001: lessons to be learned inquiry report. Available at: webarchive.nationalarchives.gov.uk/ 20100807034701/ archive.cabinetoffice.gov.uk/ fmd/fmd_report/ documents/index.htm (accessed on 26 June 2011).

2. Anderson I. (2008). – Foot and mouth disease 2007: a review and lessons learned. Available at: webarchive.nationalarchives.gov.uk/ 20100807034701/ archive.cabinetoffice.gov.uk/ fmdreview/ (accessed on 26 June 2011).

3. Bart J. (1995). – Acceptance criteria for using individual-based models to make management decisions. *Ecol. Applic.*, **5** (2), 411-420.

4. Bates T.W., Carpenter T.E. & Thurmond M.C. (2003). – Benefit-cost analysis of vaccination and preemptive slaughter as a means of eradicating foot-and-mouth disease. *Am. J. vet. Res.*, **64** (7), 805-812.

5. Bates T.W., Thurmond M.C. & Carpenter T.E. (2003). – Description of an epidemic simulation model for use in evaluating strategies to control an outbreak of foot-and-mouth disease. *Am. J. vet. Res.*, **64** (2), 195-204.

6. Bates T.W., Thurmond M.C. & Carpenter T.E. (2003). – Results of epidemic simulation modeling to evaluate strategies to control an outbreak of foot-and-mouth disease. *Am. J. vet. Res.*, **64** (2), 205-210.

7. Baxter S.M., Day S.W., Fetrow J.S. & Reisinger S.J. (2006). – Scientific software development is not an oxymoron. *PLoS Comput. Biol.*, **2** (9), e87.

8. Dent J.B. & Blackie M.J. (1979). – Systems simulation in agriculture. Applied Science Publishers, London.

9. Dubé C. (2009). – Network analysis of dairy cattle movements in Ontario to support livestock disease simulation modelling. Ph.D. dissertation submitted to the University of Guelph, Ontario, Canada.

10. Dubé C., Geale D. & Sanchez J. (2008). – NAADSM orientation workshop and project plan for pilot studies, software development, and oversight of NAADSM application in South America. Canadian Food Inspection Agency, Ottawa, Ontario, Canada.

11. Dubé C., Stevenson M.A., Garner M.G., Sanson R.L., Corso B.A., Harvey N., Griffin J., Wilesmith J.W. & Estrada C. (2007). – A comparison of predictions made by three simulation models of foot-and-mouth disease. *N.Z. vet. J.*, **55** (6), 280-288.

12. Fairley R.E. (1978). – Tutorial: static analysis and dynamic testing of computer software. *Computer*, **11** (4), 14-23.

13. Fay T.H., Greeff J.C., Eisenberg B.E. & Groeneveld H.T. (2006). – Testing the model for one predator and two mutualistic prey species. *Ecol. Modell.*, **196** (1-2), 245-255.

14. Ferguson N.M., Donnelly C.A. & Anderson R.M. (2001). – The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science,* **292** (5519), 1155-1160. E-pub.: 12 April 2001.

15. Ferguson N.M., Donnelly C.A. & Anderson R.M. (2001). – Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature,* **413** (6855), 542-548. Erratum: *Nature,* **414** (6861), 329.

16. Garner M.G. & Beckett S.D. (2005). – Modelling the spread of foot-and-mouth disease in Australia. *Aust. vet. J.*, **83** (12), 758-766.

17. Garner M.G., Cowled B., East I.J., Moloney B.J. & Kung N.Y. (2011). – Evaluating the effectiveness of early vaccination in the control and eradication of equine influenza – a modelling approach. *Prev. vet. Med.*, **99** (1), 15-27. E-pub.: 16 March 2010.

18. Garner M.G., Dubé C., Stevenson M.A., Sanson R.L., Estrada C. & Griffin J. (2007). – Evaluating alternative approaches to managing animal disease outbreaks – the role of modelling in policy formulation. *Vet. ital.*, **43** (2), 285-298.

19. Gloster J., Jones A., Redington A., Burgin L., Sørensen J.H., Turner R., Dillon M., Hullinger P., Simpson M., Astrup P., Garner G., Stewart P., D'Amours R., Sellers R. & Paton D. (2010). – Airborne spread of foot-and-mouth disease – model intercomparison. *Vet. J.*, **183** (3), 278-386. E-pub.: 12 January 2009.

20. Green L.E. & Medley G.F. (2002). – Mathematical modelling of the foot and mouth disease epidemic of 2001: strengths and weaknesses. *Res. vet. Sci.*, **73** (3), 201-205.

21. Guitian J. & Pfeiffer D. (2006). – Should we use models to inform policy development? *Vet. J.*, **172** (3), 393-395. E-pub.: 16 May 2006.

22. Harvey N. & Reeves A. (eds) (2008). – Model description: North American Animal Disease Spread Model 3.1, version 1.1.1. Available at: www.naadsm.org/documentation/specification (accessed on 10 September 2010).

23. Harvey N. & Reeves A. (eds) (2010). – Model description: North American Animal Disease Spread Model 3.1, version 1.1.2. Available at: www.naadsm.org/documentation/specification (accessed on 10 September 2010).

24. Harvey N., Reeves A., Schoenbaum M.A., Zagmutt-Vergara F.J., Dubé C., Hill A.E., Corso B.A., McNab W.B., Cartwright C.I. & Salman M.D. (2007). – The North American Animal Disease Spread Model: a simulation model to assist decision making in evaluating animal disease incursions. *Prev. vet. Med.*, **82** (3-4), 176-197. E-pub.: 5 July 2007.

25. Holling C.S. (1978). – Adaptive environmental assessment and management. John Wiley & Sons, New York.

26. Jalvingh A.W., Nielen M., Maurice H., Stegeman A.J., Elbers A.R.W. & Dijkhuizen A.A. (1999). – Spatial and stochastic simulation to evaluate the impact of events and control measures on the 1997-1998 classical swine fever epidemic in the Netherlands. I. Description of simulation model. *Prev. vet. Med.*, **42** (3-4), 271-295.

27. Karsten S., Rave G. & Krieter J. (2005). – Monte Carlo simulation of classical swine fever epidemics and control. II. Validation of the model. *Vet. Microbiol.*, **108** (3-4), 199-205.

28. Keeling M.J., Woolhouse M.E.J., May R.M., Davies G. & Grenfell B.T. (2003). – Modelling vaccination strategies against foot-and-mouth disease. *Nature*, **421** (6919), 136-142. E-pub.: 22 December 2002.

29. Keeling M.J., Woolhouse M.E.J., Shaw D.J., Matthews L., Chase-Topping M., Haydon D.T., Cornell S.J., Kappey J., Wilesmith J. & Grenfell B.T. (2001). – Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, **294** (5543), 813-817. E-pub.: 3 October 2001.

30. Kitching R.P., Taylor N.M. & Thrusfield M.V. (2007). – Veterinary epidemiology: vaccination strategies for foot-and-mouth disease. *Nature*, **445** (7128), E12-E13.

31. Kitching R.P., Thrusfield M.V. & Taylor N.M. (2006). – Use and abuse of mathematical models: an illustration from the 2001 foot and mouth disease epidemic in the United Kingdom. *In* Biological disasters of animal origin. The role and preparedness of veterinary and public health services (M. Hugh-Jones, ed.). *Rev. sci. tech. Off. int. Epiz.*, **25** (1), 293-311.

32. Kleijnen J.P.C. (1999). – Validation of models: statistical techniques and data availability. *In* Proc. of the 1999 Winter Simulation Conference (P.A. Farrington, H.B. Nembhard, D.T. Sturrock & G.W. Evans, eds), 5-8 December, Phoenix, Arizona, 647-654. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

33. Knepell P.L. & Arangno D.C. (1993). – Simulation validation: a confidence assessment methodology. IEEE Computer Society Press, Los Alamitos, California.

34. Law A.M. & Kelton W.D. (2000). – Simulation modeling and analysis. McGraw-Hill, Boston, Massachusetts.

35. Law A.M. & McComas M.G. (2001). – How to build valid and credible simulation models. *In* Proc. of the 2001 Winter Simulation Conference (B.A. Peters, J.S. Smith, D.J. Medeiros & M.W. Rohrer, eds), 9-12 December, Arlington, Virginia, 22-29. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

36. Loehle C. (1997). – A hypothesis testing framework for evaluating ecosystem model performance. *Ecol. Modell.*, **97** (3), 153-165.

37. Mardones F.O., Perez A.M., Sanchez J., Alkhamis M.A. & Carpenter T.E. (2010). – Parameterization of the duration of infection stages of serotype O foot-and-mouth disease virus: an analytical review and meta-analysis with application to simulation models. *Vet. Res.*, **41** (4), 45. E-pub.: 8 March 2010.

38. Martin S.W., Meek A.H. & Willeberg P. (1987). – Veterinary epidemiology: principles and methods. Iowa State University, Ames, Iowa.

39. Mayer D.G. & Butler D.G. (1993). – Statistical validation. *Ecol. Modell.*, **68** (1-2), 21-32.

40. McCall J.A., Richards P.K. & Walters G.F. (1977). – Factors in software quality. Air Development Center, Rome.

41. McCarl B.A. (1984). – Model validation: an overview with some emphasis on risk models. *Rev. Market. agric. Econ.*, **52** (3), 153-173.

42. Oreskes N., Shrader-Frechette K. & Belitz K. (1994). – Verification, validation, and confirmation of numerical models in the Earth sciences. *Science*, **263** (5147), 641-646.

43. Overton W.S. (1977). – A strategy of model construction. *In* Ecosystem modeling in theory and practice (C.A.S. Hall & J.W. Day Jr, eds). John Wiley & Sons, New York, 49-74.

44. Owen K., Stevenson M.A. & Sanson R.L. (2011). – A sensitivity analysis of the New Zealand standard model for foot-and-mouth disease. *In* Models in the management of animal diseases (P. Willeberg, ed.). *Rev. sci. tech. Off. int. Epiz.,* **30** (2), 513-526.

45. Patyk K.A., Helm J., Martin M.K., Forde-Folle K.N., Olea-Popelka F.J., Hokanson J., Fingerlin T. & Reeves A. (Submitted for publication). – Establishing input parameters for an epidemiologic simulation model scenario of the spread and control of highly pathogenic avian influenza (H5N1) among commercial and backyard poultry flocks in South Carolina.

46. Pendell D.L., Leatherman J., Schroeder T.C. & Alward G.S. (2007). – The economic impacts of a foot-and-mouth disease outbreak: a regional analysis. *J. agric. appl. Econ.*, **39**, 13-33.

47. Portacci K., Reeves A., Corso B. & Salman M. (2009). – Evaluation of vaccination strategies for an outbreak of pseudorabies virus in US commercial swine using the NAADSM. *In* Proc. 12th International Symposium on Veterinary Epidemiology and Economics (ISVEE), 10-14 August, Durban, South Africa, 78. Available at: www.sciquest.org.nz/node/68285 (accessed on 26 June 2011).

48. Power M. (1993). – The predictive validation of ecological and environmental models. *Ecol. Modell.*, **68** (1-2), 33-50.

49. Reynolds Jr. M.R., Burkhart H.E. & Daniels R.F. (1981). – Procedures for statistical validation of stochastic simulation models. *Forest Sci.*, **27** (2), 349-364.

50. Robinson A.P. & Froese R.E. (2004). – Model validation using equivalence tests. *Ecol. Modell.*, **176** (3-4), 349-358.

51. Rykiel E.J. (1996). – Testing ecological models: the meaning of validation. *Ecol. Modell.*, **90** (3), 229-244.

52. Sanson R.L., Harvey N., Garner M.G., Stevenson M.A., Davies T.M., Hazelton M.L., O'Connor J., Dubé C., Forde-Folle K.N. & Owen K. (2011). – Foot-and-mouth disease model verification and 'relative validation' through a formal model comparison. *In* Models in the management of animal diseases (P. Willeberg, ed.). *Rev. sci. tech. Off. int. Epiz.,* **30** (2), 527-540.

53. Sargent R.G. (2009). – Verification and validation of simulation models. *In* Proc. of the 2009 Winter Simulation Conference (M.D. Rossetti, R.R. Hill, B. Johansson, A. Dunkin & R.G. Ingalls, eds), 13-16 December, Austin, Texas, 162-176. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

54. Savill N.J., Shaw D.J., Deardon R., Tildesley M.J., Keeling M.J., Woolhouse M.E.J., Brooks S.P. & Grenfell B.T. (2007). – Effect of data quality on estimates of farm infectiousness trends in the UK 2001 foot-and-mouth disease epidemic. *J. roy. Soc., Interface*, **4** (13), 235-241.

55. Savoye R. (2004). – DejaGnu: the GNU testing framework. Free Software Foundation, Boston, Massachusetts. Available at: www.gnu.org/ software/ dejagnu/ manual/ (accessed on 10 September 2010).

56. Scheller R.M., Sturtevant B.R., Gustafson E.J., Ward B.C. & Mladenoff D.J. (2010). – Increasing the reliability of ecological models using modern software engineering techniques. *Front. Ecol. Environ.*, **8** (5), 253-260.

57. Schlesinger S. (1979). – Terminology for model credibility. *Simulation*, **32** (3), 103-104.

58. Schoenbaum M.A. & Terry Disney W. (2003). – Modeling alternative mitigation strategies for a hypothetical outbreak of foot-and-mouth disease in the United States. *Prev. vet. Med.*, **58** (1-2), 25-52.

59. Scholten H. & Udink ten Cate A.J. (1999). – Quality assessment of the simulation modeling process. *Comput. Electron. Agric.*, **22** (2-3), 199-208.

60. Spedding C.R.W. (1988). – An introduction to agricultural systems. Elsevier Applied Science, London.

61. Stärk K.D.C., Pfeiffer D.U. & Morris R.S. (2000). – Within-farm spread of classical swine fever virus – a blueprint for a stochastic simulation model. *Vet. Q.*, **22** (1), 36-43.

62. Taylor N. (2003). – Review of the use of models in informing disease control policy development and adjustment. Department of Environment, Food, and Rural Affairs, London. Available at: epicentre.massey.ac.nz/ resources/ acvsc_grp/ docs/ Taylor_2003.pdf (accessed on 10 September 2010).

63. Thornley J.H.M. & France J. (2009). – Modelling foot and mouth disease. *Prev. vet. Med.*, **89** (3-4), 139-154. E-pub.: 27 March 2009.

64. United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Veterinary Services (VS), Centers for Epidemiology and Animal Health (CEAH) (2002). – North American Animal Health Committee Conference. Fort Collins, Colorado. USDA-APHIS-VS-CEAH, Fort Collins, Colorado.

65. United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Veterinary Services (VS), Centers for Epidemiology and Animal Health (CEAH) (2004). – Validation of SpreadModel: minutes of the subject matter expert team meeting. USDA-APHIS-VS-CEAH, Fort Collins, Colorado.

66. United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Veterinary Services (VS), Centers for Epidemiology and Animal Health (CEAH) (2009). – Vaccine bank requirements for foot-and-mouth disease in southwest Kansas: a preliminary report. USDA-APHIS-VS-CEAH, Fort Collins, Colorado.

67. United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Veterinary Services (VS), Centers for Epidemiology and Animal Health (CEAH) & World Organisation for Animal Health (OIE) (2008). – Proc. CEAH-OIE Epidemiological Modeling Workshop, 11-13 August, Fort Collins, Colorado. USDA-APHIS-VS-CEAH, Fort Collins, Colorado.

68. Vigre H. (2008). – A comparison of three simulation models – the EpiLab project. *In* Proc. Centers for Epidemiology and Animal Health (CEAH) World Organisation for Animal Health (OIE) Epidemiological Modeling Workshop, 11-13 August, Fort Collins, Colorado. United States Department of Agriculture, Animal and Plant Health Inspection Service, Veterinary Services & CEAH, Fort Collins, Colorado.

69. Waller L.A., Smith D., Childs J.E. & Real L.A. (2003). – Monte Carlo assessments of goodness-of-fit for ecological simulation models. *Ecol. Modell.*, **164** (1), 49-63.

70. Whitner R.B. & Balci O. (1989). – Guidelines for selecting and using simulation model verification techniques. *In* Proc. of the 1989 Winter Simulation Conference (E.A. MacNair, K.J. Musselman & P. Heidelberger, eds), 4-6 December, Washington, DC. IEEE, Piscataway, New Jersey, 559-568.