Development of a Probabilistic Network for Clinical Detection of Classical Swine Fever

Geenen, P.L.¹, Elbers, A.R.W.², Van der Gaag, L.C.¹, Loeffen, W.L.A.²
¹ Dept. of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands
² Dept. of Virology, Central Institute for Animal Disease Control, WUR, Lelystad, The Netherlands

ABSTRACT

Presence of a classical swine fever (CSF) infection is usually first revealed by the occurrence of clinical signs. In order to supply veterinary practitioners with an additional tool to identify CSF suspect situations as early as possible, a decision-support system is being developed with the help of experts in the field of CSF. The system builds upon a probabilistic network that is used for computing the probability of the presence of CSF in individual pigs in herds with clinical problems. We performed a preliminary evaluation study of the network using data from experimental infection trials with CSF virus strains of different virulence. The large differences in the clinical signs observed resulted in distinct differences in the computed probabilities, thereby illustrating the potential of our system.

INTRODUCTION

Experience shows that a first outbreak of classical swine fever (CSF) in a country typically remains undetected for a relatively long period of time (e.g. Elbers et al., 1999). During this period, in farm-dense areas especially, the virus can spread to multiple farms before control measures are effectuated, which often results in an almost uncontrollable epidemic. The presence of CSF is usually first revealed by the occurrence of clinical signs (Elbers et al., 2002). Unfortunately, the signs associated with CSF are quite aspecific, which causes the clinical diagnosis of CSF to be highly uncertain. In order to supply veterinary practitioners with an additional tool to identify a CSF suspect situation as early as possible, a decision-support system is being developed with the help of experts. When a pig farm with serious clinical problems is visited, the system leads the veterinary practitioner through a list of questions related to the clinical diagnosis of CSF. The answers are entered into a probabilistic network, which returns the probability of the pigs having a CSF infection. This probability then is presented to the practitioner, together with an advice how to act.

In this paper, we present the results of a preliminary evaluation study of our probabilistic network for the clinical diagnosis of CSF. An assessment of the performance of the network and of the robustness of its diagnoses, are essential before the network can be used in real-life practice (Lucas et al., 2004). For our study, we used a Swiss data collection of experimental infections with CSF virus strains of different virulence. We would like to note that our evaluation study is a preliminary study of the network's clinical performance. For a more comprehensive study, field data on clinical signs of both infected and non-infected pigs are needed. For this purpose, we recently started a field trial in the Netherlands, in a non-CSF situation, involving 11 veterinary practitioners; further data will be collected during CSF outbreaks within member states of the European Union in the framework of the EPIZONE research project. In the field trial in the Netherlands, the collection of data is facilitated using a data-entry program that runs on a personal digital assistant. With this program, data are collected in a standardised way that adheres to the precise definitions of the stochastic variables and their values in our probabilistic network for CSF, thereby minimising biases in data collection. We hope to report more comprehensive evaluation studies of our network using field data in the near future.

THE CSF NETWORK

A probabilistic network is a graphical model of a probability distribution over a collection of stochastic variables (Jensen, 2001) and is used for solving classification problems in a wide variety of fields. It models the relevant variables in a field, along with their relations, in a graphical structure; the strengths of the relations are captured by conditional probabilities. For our decision-support system for CSF, we constructed such a probabilistic network with the help of two experts. The network includes 41 variables (with 2-5 values each), of which 24 can be observed upon clinical investigation. The variables capture processes in the underlying pathogenesis, risk factors, relevant clinical signs, and various alternative explanations for these signs. The network further includes 82 relations between the variables and 2113 conditional probabilities. Building a probabilistic network with the help of experts is a creative process, in which the knowledge of a field is carefully gathered and modelled. For the construction of our CSF network, we held a sequence of interviews with two experts, in which we elicited the knowledge that was to be modelled. In the first interview the experts were asked to describe the onset and progression of a CSF infection and the associated clinical
symptoms. This interview was followed by 13 structured interviews in which detailed questions were asked. The experts started with identifying the relevant variables in the field (e.g. Conjunctivitis), their possible values (e.g. present or absent) and the relationships between them (e.g. Mucositis may result in Conjunctivitis). The relations between the variables were elicited using questions like “What could cause cyanosis?” and “What manifestations could a high body temperature have?”. The knowledge about the variables, their values and their relations was modelled in a graphical structure. When this structure was considered reasonably robust, the focus of the interviews shifted towards the strengths of the relations between the variables. To estimate the conditional probabilities required, we used standardised forms with questions accompanied by a probability scale with words and numbers. The thus completed probabilistic network could now be used for computing any probability over its variables.

STUDY DESIGN

For a preliminary study of the performance of our probabilistic network for CSF, we used a Swiss collection of experimental data on infections with CSF virus strains of different virulence. The data were kindly made available by Dr. Martin Hofmann of the Institute of Virology and Immunoprophylaxis, Mittelhäusern, Switzerland; some of the data were published before by Mittelholzer et al. (2000). The data are composed of daily findings on the absence or presence of ten clinical signs and body temperature measurements of 16 SPF pigs experimentally infected with a highly virulent (Brescia, Eysterup and Koslov), a moderately virulent (vA187-1 and cp vA187-1), or an avirulent (vA187-Ubi) CSF strain. Presence of a clinical sign was recorded in three different categories of severity: slightly altered, distinct clinical sign, or severe CSF sign. For each of the signs and categories, precise definitions were provided on a checklist to minimise variation in judgment by the various people involved. Before we could use the data in our evaluation study, we had to translate the various clinical signs and categories into values for the stochastic variables of our probabilistic network. The translation resulted in data for a total of 17 variables observed in 16 pigs.

To evaluate its performance, the clinical signs per pig per day were entered into our probabilistic network. Variables for which no data were available were treated as not having been observed; the same applied for missing values. In the network, the prior probability that a pig for which no findings are available is infected with CSF, equals 0.000077. For each individual pig on each separate day, we computed from the network the posterior probability of this pig having a CSF infection. The posterior probabilities were expressed in terms of a multiplication factor for the prior probability. A posterior probability of 0.0077, for example, means that the probability that the pig under study is infected with CSF is 100 times larger than that for a pig for which no information is available. For the various strains of CSF, the time before detection and the number of detected infections were compared, assuming a threshold multiplication factor of 10. Only pigs for which the probability of having CSF has increased by a factor 10 or more, are considered as having been detected by the decision-support system. This detection threshold was chosen for explanation purposes only; an appropriate detection threshold has to be established from field data of infected and non-infected pigs.

RESULTS

The results of the evaluation study of our probabilistic network are shown in Figures 1a and b. In both figures, a multiplication factor equal to 1 expresses that the posterior probability of CSF being present in a particular pig on a particular day is equal to the prior probability of having a CSF infection. If all clinical signs are absent, which is the case for all pigs on the day of inoculation, the posterior probability of a CSF infection lies below the prior probability and a multiplication factor smaller than 1 is found. The two figures now show that the posterior probabilities computed for the various strains of different virulence, are markedly different. The severe clinical signs observed in the pigs that had been inoculated with a highly virulent strain, lead to a swift increase in the computed probabilities over time. Assuming a detection threshold of 10, the eight pigs all are detected as having a CSF infection, the first three even as soon as on day 4 (pigs 220, 309 and 310). In contrast, the pigs that had been inoculated with the avirulent strain hardly show any clinical signs and, as expected, are not detected by the network as possibly having a CSF infection. The computed probabilities even remain below the prior probability of CSF. The moderately virulent strains of CSF show a wide spectrum of severity in clinical signs, which is reflected in the output of our network. Four out of the six pigs that had been inoculated with such a strain, were detected as having a CSF infection.

CONCLUSION

With the help of two experts, we constructed a probabilistic network for the clinical detection of CSF in individual pigs. Our network is aimed specifically at detecting infections with moderately virulent strains as these are involved in most outbreaks of the disease; the high variation in associated clinical patterns, moreover, makes an infection with such a strain hard to detect. In a preliminary evaluation study, our network showed satisfactory performance on clinical data of experimentally inoculated pigs. The more severe the
clinical signs found in a pig, the higher was the probability yielded by the network that the pig is infected with CSF. The variation in clinical patterns of different strains of CSF thus was reflected in the computed probabilities, as a result of which some infected pigs were identified sooner than others. We expect that, by combining the findings for a number of diseased pigs from a herd, the performance of our probabilistic network can be further improved, especially for the moderately virulent strains of CSF.

**Figures 2a & b:** Posterior probabilities of CSF (expressed as multiplication factor of the prior probability) for pigs inoculated with a highly virulent (a), avirulent (a, in grey), or moderately virulent strain (b).

**REFERENCES**


**ACKNOWLEDGEMENTS**

We are grateful to Dr. Martin Hofmann who kindly provided us the original data of the Swiss infection trials. The research was supported by the Netherlands Organisation for Scientific Research (NWO) and by the Central Institute for Animal Disease Control (CIDC-Lelystad).