Research paper

Identifying risk factors for ovine respiratory processes by using Bayesian networks

Javier Galapero[a,*], Sara Fernández[a], Carlos J. Pérez[b], Fernando Calle-Alonso[b], Joaquín Rey[c], Luis Gómez[a]

[a] Histology and Pathological Anatomy, Department of Animal Medicine, Faculty of Veterinary Medicine, University of Extremadura, Avda de la Universidad s/n, 10003 - Cáceres (Spain).

[b] Biostatistics Unit, Department of Mathematics, Faculty of Veterinary Medicine, University of Extremadura, Avda de la Universidad s/n, 10003 - Cáceres (Spain).

[c] Infectious Pathology and Epidemiology Unit, Department of Animal Health, Faculty of Veterinary Medicine, University of Extremadura, Avda de la Universidad s/n, 10003 - Cáceres (Spain).

[*] *Corresponding author. Tel.:+34 927257106*
E-mail address: jagaar@unex.es

Highlights

- Risk factors in ovine respiratory processes were studied by using Bayesian networks.
- Lung consolidation was mainly associated to temperature and relative humidity.
- *Mycoplasma sp.* was recorded as a risk for lung consolidation.
- Risk factors in ovine respiratory processes were identified by bayesian networks.
- Bayesian networks may serve as a helpful tool for lamb production systems.

Abstract

A proposal is put forward to use Bayesian networks to identify risk factors for pulmonary consolidation. An experiment was conducted with 410 fattening lambs from five feedlots in Extremadura (southwestern Spain). Environmental conditions (temperature, relative humidity, and ammonia concentration) were recorded during the study period. In a microbiological study, *Mycoplasma* spp. and *Pasteurellaceae* were obtained by conventional culture and identified by nested polymerase chain reaction. After slaughter, lungs were collected and examined macro- and microscopically (histological type and pulmonary consolidation). To the best of the authors' knowledge, Bayesian networks have not before been used to relate the presence/absence of pulmonary consolidation to environmental conditions, *Mycoplasma* spp., *Pasteurella* spp., and histological changes. The results showed that the main factors causing ovine inflammatory respiratory processes and pulmonary consolidation were temperature, relative humidity, and *Mycoplasma* spp. Control of these factors may help reduce the incidence of pulmonary consolidation.

Keywords

Bayesian networks, lung pathology, ovine respiratory processes, preventive medicine, risk factors.

Introduction

Analysing animal health data is a complex task since the relationships between variables are usually not known. In fact, determining information about the way the variables are related is usually an objective of analysis. Lewis *et al.* (2011) discussed the potential of using Bayesian networks as analytical tools for processing complex animal health data. Specifically, they proposed the use of structure discovery to identify variables that may be associated with health status.

Bayesian networks belong to a family of probabilistic graphical models (see, e.g., Jensen, 1995). These graphical structures are mathematically rigorous and intuitively understandable, and can be used to represent knowledge about an uncertain domain. A Bayesian network is a directed acyclic graph (a graph with directed edges between vertices) with an associated set of probability distributions that enables an effective representation and computation of the joint probability distribution over a set of random variables. The visual representation of the random variables' dependency structure is especially useful. Bayesian networks combine principles from graph theory, probability theory, computer science, and statistics (Jensen and Nielsen, 2007).

These networks can be implemented to search for the structure between variables by finding an optimal directed acyclic graph for the dataset in hand, providing information about the possible relationships linking the variables involved. The network can then be used to infer conditional probabilities. Bayesian networks can also be built using reliable subjective information provided by an expert instead of searching for the structure. Moreover, a mixed approach can be taken by first searching for the structure based on data, and then adjusting the structure using subjective information.

Bayesian networks have been extensively used in many fields of study, especially in artificial intelligence. They are gradually being introduced for the analysis of data in the veterinary field. McKendrick *et al.* (2000) applied Bayesian networks to aid in the differential diagnosis of tropical bovine diseases. Otto and Kristensen (2004) proposed a biological model based on a Bayesian network to determine risk factors for infection with *Mycoplasma hyopneumoniae* in swine for slaughter. Ettema *et al.* (2009) used a Bayesian network to estimate the probability of claw and digital skin diseases by combining cow- and herd-level information. Jensen *et al.* (2009) used a Bayesian network to model the causes of leg disorders in finisher herds. McCormick *et al.* (2013) applied these networks to the identification of environmental conditions that influence disease in pigs, and Firestone *et al.* (2013) analysed the associations between risk factors and the infection status of horses in an equine influenza outbreak in Australia.

In the present study, we focus on pathological respiratory processes that lead to direct and indirect losses such as mortality, weight loss, low conversion rates, and greater numbers of pulmonary lesions at slaughter (Goodwin *et al*., 2004; Lacasta *et al*., 2008). The agents and risk factors that may influence these processes are numerous. Temperature and relative humidity have been identified as associated risk factors (Brogden *et al.*, 1998; Yener *et al.*, 2009). Indeed, climatic factors that increase relative humidity have a significant influence on the presence of pneumonia (Lacasta *et al*., 2008). High mortality and morbidity from these causes during the summer months have been reported (Plummer *et al.*, 2007). In the present work, we take a Bayesian network based approach to analyse conjointly the influence of some possible risk factors on pulmonary consolidation, rather than treating them separately.

To the best of our knowledge, Bayesian networks have not previously been used to relate the presence/absence of pulmonary consolidation to environmental conditions,

Mycoplasma spp., *Pasteurella* spp., and histological changes. A Bayesian network-based approach is considered to analyse some relevant scenarios for pulmonary consolidation.

Materials and Methods

Experimental design

Four hundred and ten Merina breed lambs and their commercial crossings of both sexes from five feedlots in Extremadura (southwestern Spain) formed the sample for study. The animals' ages ranged from 80 to 100 days, and they were monitored during the feedlot period (15-21 days) from February to November. They were held in pens in the different feedlots at a density of 0.5 m^2 per animal. Feed (pellet concentrate and straw) and water were administered *ad libitum*. The animals had been transported by road from the farms to the feedlots. Twenty-one days later, they were transported to the slaughterhouse at weights of 24-26 kg.

Environmental conditions were registered during each feedlot period. Temperature and relative humidity were recorded using a data logger (175H1, Testo®, Titisee, Germany). Two measurements per hour for each parameter were recorded, and the data were processed in a spreadsheet (Microsoft Excel®). The mean temperature and mean relative humidity during the fattening period were calculated. The ammonia concentration was recorded using an ammonia detector (Gastec GV-100S®, Kanagawa, Japan). Two weekly measurements were made during each feedlot period, and the mean was

calculated. The data were binned into intervals using, in part, the classification proposed by ITOVIC (1991).

Bacteria identification

Nucleic acid purification and amplification of the 16S-23S intergenic spacer region was carried out using primers F2A and R2 to identify *Mycoplasma* spp., as described by Tang *et al*. (2000). This technique detected the presence of *Mycoplasma ovipneumoniae (M. ovipneumoniae)* and *Mycoplasma arginini (M. arginine).* Molecular identification of *Pasteurella multocida (P. multocida)* was carried out using a polymerase chain reaction. This technique was used to amplify a specific fragment of the gene *kmt1* of *P. multocida* with primers KMT1SP6 and KMT1T7, as described by Townsend *et al.* (1998). *Mannheimia haemolytica (M. haemolytica)* was identified by direct haemagglutination, and *Pasteurella pneumotropica* was also isolated.

Macroscopic and microscopic study

Lungs were collected at abattoir. They were photographed on both sides, and classified into two groups depending on the presence or absence of consolidation, i.e., consolidation greater than 0% or equal to 0%, respectively.

For the histological study, samples were fixed in neutral buffered formalin (3.5%, 0.1 M, and pH 7.2), routinely processed, and embedded in paraffin. Sections of 5 µm were stained with haematoxylin and eosin.

Four histological groups were established following the classification of Caswell and Williams (2007). (1) Diffuse alveolar damage: This group included ciliary necrosis and compaction and alveolar denudation, characterized by loss of type I pneumocytes and epithelial basement membrane. A mild inflammatory reaction was observed in the

peribronchial adjacent zone, with no bronchial associated lymphoid tissue (BALT) reaction. No fibrotic processes or pleural phenomena were found. (2) Interstitial pneumonia: This group was characterized by septal damage, loss of type I pneumocytes and proliferation of type II pneumocytes. Septa were thickened, with mononuclear cell inflammatory infiltrates, marked congestion, and oedema. Hyaline membranes were observed in interstitium and alveoli. They consisted of fibrin, eosinophilic proteinaceous material, and cell debris. The airways and pleural surface showed no significant changes. (3) Purulent bronchopneumonia: The predominant lesional pattern in this group was of an exudative type. Neutrophils, macrophages, and cell debris were observed in bronchial, bronchiolar, and alveolar lumina. BALT was enlarged as a multinodular structure. Perivascular inflammatory processes mainly involved mononuclear cells. Alveolar septa did not show any change. A severe fibrinous inflammatory reaction was present on the pleural surface. Pneumonia together with purulent bronchopneumonia was also described. (4) Mixed changes (bronchointerstitial pneumonia): Inflammatory reaction within the interalveolar septa and exudative processes were concomitant in all the samples included in this group ‘1:1 of the studied sample’. BALT was frequently enlarged, although displaying a diffuse pattern. Alveoli were filled with neutrophils and macrophages, and some bronchi and bronchioles showed this inflammatory cell infiltration. However, thickening of the interalveolar septum and a marked increase in the number of mononuclear cells were also found in areas where no exudate was observed. The pleural surface was affected in these lungs, showing lesions similar to the purulent bronchopneumonia group. Microphotographs were taken using a microscope (Eclipse 80i, NIKON®, Tokyo, Japan) with a digital video camera (DXMI200F, NIKON®, Tokyo, Japan).

Data analysis

Six algorithms were tested to build Bayesian networks using the GeNie/SMILE software (Druzdzel, 1999). Specifically, these were Greedy Thick Thinning, Bayesian Search, Essential Graph Search, Tree Augmented Naive Bayes, Augmented Naive Bayes, and Simple Naive Bayes. A cross-validation scheme was used to analyse model performance. The networks were compared using the area under the Receiver Operating Characteristic (ROC) Curve (AUC) and the accuracy rates. The AUC is a measure of how well the model can distinguish between the two groups of pulmonary consolidation (absence/presence). Its value ranges between 0 and 1.

The Greedy Thick Thinning search algorithm (Dash and Cooper, 2004) provided the best results. With this method, initially each node has no parents. The nodes providing the strongest increases in the score in the resulting structure (using the Bayesian Dirichlet with score equivalence and uniform priors, known as BDeu criterion) are added incrementally as parents (Cooper and Herskovits, 1992; Silander *et al.*, 2008). When the addition of parents does not increase the score, they are no longer added to the node. A weakly informative prior distribution is considered. The prior distribution of the parameters is associated with their corresponding nodes and all the possible combinations of the parents $P(\theta_{X_i|p_a(X_i)}) = Dirichlet(1, \ldots, 1)$, where θ is the parameter vector and $p_a(X_i)$ is the combination of states of the parents for node for node X_i. The total number of parameters, i.e., the length of the parameter vector, is the product of the number of states of the node under consideration multiplied by the number of states of every parent node. The maximum number of parent nodes was set to six (one less than the number of variables), allowing all possible relationships among variables.

With this algorithm, an optimized network structure is obtained using the BDeu criterion. The arcs represent the statistical dependence existing between the linked variables, but the relationships illustrated by the Bayesian network may be due to a causal connection or may be spurious from a practical point of view. Although the final graph is usually represented with directed arcs, the causality or direction of association cannot always be completely confirmed without detailed experimental analysis (Heckerman *et al.*, 1995). A sensitivity analysis was also performed.

After building the network, we analysed the most interesting risk scenarios in terms of probabilities by considering evidence propagation. Evidence propagation is one of the most powerful characteristics of Bayesian networks. It allows the probabilities of each node to be updated via bidirectional propagation of new information through the whole structure. In each scenario, a percentage of 100% is set for one category in one or more nodes in order to show how the environmental and/or bacterial evidence influences pulmonary consolidation. The probability estimates for the different scenarios will be reported in terms of percentages in two tables.

Results

With the specifications presented in the previous section, six Bayesian networks were built using the following variables: mean relative humidity, mean temperature, mean ammonia level, *Mycoplasma* spp., *Pasteurella* spp., histological type, and pulmonary consolidation. The environmental variables are continuous, and were binned into intervals according to pre-defined environmental criteria.

In order to compare the networks, a 4-fold cross-validation scheme was performed with 500 iterations. The AUC and accuracy rates were calculated. Figure 1 shows the ROC curves. The estimated AUC for the Greedy Thick Thinning algorithm was the best, giving a high value (0.9086). This means that the model fits the data quite well. The other algorithms achieved lower AUC values.

[Figure 1 about here]

In coherence with the AUC criterion, the greatest accuracy rate was also achieved with the Greedy Thick Thinning algorithm (see Table 1). Note that this result is not only obtained for the overall accuracy rate, but also for the presence and absence accuracy rates. Hence, this algorithm provides the best results using both the AUC and the accuracy rates with the cross-validation scheme that was considered.

[Table 1 about here]

The network structure and conditional probabilities were calculated based on the collected dataset, using the Greedy Thick Thinning algorithm (see Figure 2). The percentages given on the nodes are estimated conditional probabilities that illustrate how the state of a variable influences the probability distribution for the states of other variables. Pulmonary consolidation was directly influenced by the mean temperature, mean relative humidity, and histological type, and indirectly by *Mycoplasma* spp. The joint probability distribution obtained was:

$$P(T,H,A,M,P,HI,C) = P(T)P(H|T)P(A|T,H)P(M|T)P(P|T,M)P(HI|M,H)P(C|T,H,HI)$$

with *T* being mean temperature, *H* mean relative humidity, *A* mean ammonia level, *M* *Mycoplasma*, *P* *Pasteurella*, *HI* histological changes, and *C* pulmonary consolidation.

[Figure 2 about here]

A sensitivity analysis was performed to provide information on how small changes in the preliminary factors impact the terminal events (absence/presence of pulmonary consolidation). In Figure 2, the intensity of the red colouring of the nodes different from that of pulmonary consolidation indicates which variables have more impact in leading to changes in the probability of the target node states when small changes are made in the probability of their states (the more intense the red, the more impact). Mean temperature and mean relative humidity produce the most important changes in pulmonary consolidation.

In order to make a deeper sensitivity analysis, all the possible cases (single states and combinations) were studied. The impact of the top ten cases on pulmonary consolidation was displayed in a sensitivity tornado chart (Figure 3). The variation displayed in this chart is obtained by adjusting the probability of occurrence for the different states ±20% of their respective values, either several factors at once or one at a time. The interpretation of the tornado chart is simple. Green indicates an increase of 20% in the factor studied, and red a decrease of 20%. The effect of these variations is reflected with a bar, indicating whether the effect is directly proportional (green on the right) or inversely proportional (red on the right). For example, the original probability of consolidation presence was 48.6%, and an increase (see Bar 6 in Figure 3) of 20% in the probability of the mean temperature (state 10-13 degrees), would increase the probability of consolidation presence to 49%. On the other hand, a decrease of 20% in the probability of the mean temperature (state 10-13 degrees) would decrease the probability of pulmonary consolidation presence to 48.2%. Although the model is robust and the appreciated changes are very small, temperature and relative humidity can be

pointed to as the two variables causing the greatest variations in pulmonary consolidation when small changes are instituted.

[Figure 3 about here]

Now we shall focus on analysing various scenarios. Evidence propagation about the state of a given node sheds light on the states of the nodes pointed to by that given node. The probabilities for each state of the nodes can be estimated with the conditional distribution and the joint distribution. For example, if the histology is set to the value diffuse alveolar damage, then the probability of the absence of pulmonary consolidation given that diffuse alveolar damage occurs is:

$$
\begin{aligned}
&P(C - Absence|HI - IrritativeChanges)\\
&= \frac{P(C = Absence, HI = IrritativeChanges)}{P(HI = IrritativeChanges)}\\
&= \frac{\sum_{T,H,A,M,P} P(T,H,A,M,P,HI = IrCh, C = Absence)}{\sum_{T,H,A,M,P,C} P(T,H,A,M,P,HI = IrCh, C)}\\
&= \frac{\sum_{T,H,A,M,P} P(T)P(H|T)P(A|T,H)P(M|T)P(P|T,M)P(HI = IrCh|M,H)P(C = Abs|T,H,HI = IrCh)}{\sum_{T,H,A,M,P,C} P(T)P(H|T)P(A|T,H)P(M|T)P(P|T,M)P(HI = IrCh|M,H)P(C|T,H,HI = IrCh)}
\end{aligned}
$$

In this case, the probability of the absence of pulmonary consolidation given that diffuse alveolar damage was produced was estimated as 63.3%. The evidence propagation concept is used to show how one or more variables influence the absence or presence of pulmonary consolidation. Table 2 presents the evidence propagation for some possible environmental scenarios.

[Table 2 about here]

Mean temperature and mean relative humidity were negatively correlated (r =-0.704, $p<0.001$), i.e., when the temperature increased, the relative humidity decreased (see Figure 4).

[Figure 4 about here]

The Bayesian network shows a relationship between temperature and humidity, and both these variables showed a direct relationship with ammonia levels. According to ITOVIC (1991), ideal conditions of a feedlot correspond to a mean temperature in the range 13-16°C and a mean relative humidity of 70-80%. We define a scenario (Scenario 1) with these two conditions and, as one observes in Table 2, the estimated probabilities of absence in the *Pasteurellaceae* and *Mycoplasma* families are 85.4% and 91.8%, respectively. In addition, the mildest histological type (diffuse alveolar damage) provides the greatest probability (70.7%). In this scenario, the estimated probability of presence of consolidation was low (14.1%). Subsequent scenarios will be compared with this one as reference.

If the optimal temperature (13-16°C) is considered, but with a lower relative humidity, i.e., 50-60% (Scenario 2), the estimated probability of the presence of consolidation increases to 88.6%. In this case, the histology showed an increase in the estimate of interstitial pneumonia with respect to the previous scenario (50.3%). When the temperature is ideal and the ammonia level is in the range 15-20 ppm (Scenario 3), there is mainly associated a high relative humidity (>80). The estimated probability of the presence of pulmonary consolidation is 43.9%, a low percentage with respect to the previous scenario. The estimated probabilities in interstitial pneumonia and suppurative bronchopneumonia are greater than the ones obtained in Scenario 1. When the temperature is above 28°C (Scenario 4), the two most probable states of relative humidity are lower than 50 (<40 and 40-50). The probabilities of *M. arginini* and presence of consolidation increase to 40.8% and 43.1%, respectively. Considering the histology, there is a decrease in the estimated probability of diffuse alveolar damage (41.9%), and increases in the estimated probabilities of interstitial pneumonia (28.3%),

bronchopneumonia (19.4%), and mixed changes (10.4%). When the temperatures considered are below 10°C (Scenario 5), the probabilities of the presence of the *Mycoplasma* spp. and *Pasteurellaceae* families increase with respect to the ideal conditions of Scenario 1. Moreover, there were high estimated probabilities of severe lung inflammation.

Finally, the optimum relative humidity and ammonia levels of 15-20 ppm are considered (Scenario 6). The most likely temperature is 16-19°C, and the estimated probability for the presence of consolidation is 36.6%, which is higher than in the ideal scenario.

Findings related to non-environmental conditions are presented in Table 3. If the absence of the *Mycoplasma* spp. and *Pasteurellaceae* families (Scenario 7) is considered, the estimated probabilities for diffuse alveolar damage and interstitial pneumonia are 52.6% and 32.9%, respectively. The estimated probability of presence of pulmonary consolidation is 46.9%. When only *M. ovipneumoniae* is considered (Scenario 8), suppurative bronchopneumonia and mixed changes reach high probabilities with respect to the previous scenario. However, when *M. arginini* is considered (Scenario 9), there was a mild increase in diffuse alveolar damage (29.2%) with respect to Scenario 8. Both this and the previous scenario indicate the important influence of *M. ovipneumoniae* and *M. arginini* in pathological respiratory processes.

[Table 3 about here]

In regards to histology, the greatest estimated risk of pulmonary consolidation (68.3%) is obtained when bronchopneumonia is considered (Scenario 12). The estimated probabilities of presence of pulmonary consolidation are 62.0% for mixed changes (Scenario 13) and 54.2% for interstitial pneumonia (Scenario 11). However, when only

diffuse alveolar damage is considered, the presence of pulmonary consolidation is reduced to 36.7% (Scenario 10). In these four scenarios, the absence of *Pasteurella* is prevalent, with estimated probabilities greater than 90%, whereas the probabilities of absence of *Mycoplasma* range between 45.1% and 83.8%.

Discussion

It is necessary to consider the relationships among variables and their influence on health status in epidemiological studies of animal diseases (Geenen et al., 2011; Lewis *et al*., 2011). Some authors have used Bayesian networks to determine risk factors by using different parameters as markers of animal health (McKendrick *et al*., 2000; Ettema *et al.*, 2009; Jensen *et al*., 2009; Lewis *et al.*, 2011, among others). Bayesian networks allow the use of information from relationships among all the variables (Pearl, 1988; Neapolitan, 2004). Moreover, evidence propagation is a useful tool to build scenarios providing information for decision-making processes to help to improve animal health. Bayesian networks allow the incorporation of expert knowledge, if required, in subsequent stages of the model at multiple levels. The prior distribution for the first experiment may be weakly informative, but with prior information from experts, the results of the model can change when combining it with the available data. However, especial care must be taken when incorporating expert information since it may not always be reliable.

Several scenarios were especially defined to analyse the effect of the main environmental risks for ovine respiratory syndrome. The evidence was set to represent the best and worst conditions defined in the veterinary literature (ITOVIC, 1991). The

relationships between environmental conditions and the occurrence of respiratory problems have been widely studied (Nash *et al.,* 1997; Lacasta *et al*., 2008). The aim of the present study was to analyse the relationship between potential risk factors such as certain environmental variables and microorganisms implicated in respiratory processes. In our study, the results show that the most important variables influencing the presence of pulmonary consolidation are mainly environmental, specifically, temperature and relative humidity. If these variables are fixed at their optimal values according to ITOVIC (1991) (temperature 13-16 °C and relative humidity 70-80%), the estimated probability of the presence of pulmonary consolidation is very low, and the estimated probability of inflammatory processes (interstitial pneumonia, suppurative bronchopneumonia, and mixed changes) is small. In this case, the bacterial involvement is lower, in agreement with other authors (Hervás *et al.,* 1996; Niang *et al.,* 1998). However, if these conditions change, then the estimated probabilities of the presence of pulmonary consolidation and of different inflammatory types rise.

The presence of microorganisms appears to be linked to temperature. The Bayesian network shows a direct relationship between temperature and *Mycoplasma* spp., with an increase in the probability of presence of this agent when temperatures higher than the optimal are recorded. These findings agree with those of other authors (Hervás *et al.,* 1996; Niang *et al.,* 1998) who observed an increase in respiratory pathologies associated with the presence of *M. ovipneumoniae* in the warm months (with high temperatures and low relative humidity). Relative humidity also seems to affect pulmonary consolidation. Thus, when the mean relative humidity ranges within 50-60% and the mean temperature is maintained at optimum levels (13-15ºC), an increase in the estimated probabilities of interstitial pneumonia and pulmonary consolidation is

obtained. This is possibly associated with irritative phenomena in the respiratory mucosa caused by environmental dryness.

Caswell and Williams (2007) described lung inflammatory processes associated with toxic irritants. In our study, ammonia concentration seems to have an effect on the presence of consolidation as well. Ammonia values are mainly mediated by temperature and humidity. This is in agreement with the results provided by other authors who noted the association between respiratory processes and high concentrations of ammonia in old buildings with poor ventilation (Lacasta et al., 2008). When the ammonia level is fixed at a range between 15 and 20 ppm, an increase in the probability of the presence of pulmonary consolidation, interstitial pneumonia, bronchopneumonia, and mixed changes is obtained. This corroborates this parameter's involvement in inflammatory processes.

According to other authors (Alley *et al*., 1975; Jones *et al.*, 1982; Sheehan *et al*., 2007), *M. arginini* and *M. ovipneumoniae* seem to be involved in pneumonic processes in sheep. Furthermore, Lin *et al*. (2008) and Nicholas *et al*. (2008) found a relationship between *M. arginini* and *Pasteurella* genus bacteria (*M. haemolytica*). In our study, *Mycoplasma* spp. and histological types are associated with each other, and the same is true for *Mycoplasma* spp. and *Pasteurellaceae* genus bacteria. However, *Pasteurellaceae* play a less important role for pulmonary consolidation than *Mycoplasma* spp. Oruç (2006) and Lacasta *et al.* (2008) identified *M. haemolytica* as one of the main microorganisms in sheep respiratory processes, with this being the main agent isolated in the pulmonary consolidation processes that they studied.

Suppurative bronchopneumonia is the most frequent microscopical lesion associated with pulmonary consolidation (Gázquez *et al*., 2001, Gonçalves *et al.*, 2010). This is

supported by the results obtained in our study. However, Mawhinney *et al*., (2010), point to interstitial pneumonia as the main lesion of red hepatization. According to our scenarios, the estimated probability of the presence of pulmonary consolidation is 54.2% when interstitial pneumonia is present. This probability is lower than that obtained for suppurative processes. Therefore, pulmonary consolidation can be associated with interstitial and suppurative processes, with interstitial pneumonia being the predominant lesion.

To the best of our knowledge, Bayesian networks have not previously been used to relate the presence/absence of pulmonary consolidation with environmental conditions, *Mycoplasma* spp., *Pasteurella* spp., and histological changes together. Given a model that includes interaction among all the variables, and the fact that evidence propagation allows hypothetical scenarios to be defined, this tool is interesting in the identification and understanding of all the features involved in the study of pulmonary consolidation.

Conflict of interest statement

None of the authors of this paper have a financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of the paper.

Conclusion

Bayesian networks are generic tools with great potential for use in a wide range of epidemiological disease situations. In the present study, a Bayesian network model has been able to identify risk factors in ovine respiratory processes.

The main factors causing inflammatory processes and pulmonary consolidation in ovine respiratory processes are temperature, relative humidity, and *Mycoplasma* spp. The control of these variables may help to prevent this ovine pathology. The proposed model can be applied to improve conditions on farms, and thus enhance productivity.

Acknowledgements

This work was mainly supported by the *Gobierno de Extremadura* in the 2010 Call for Strategic Sectors Cooperation Projects between research groups and companies (Project PCE1007). It was partially supported by the *Consejería de Economía y Competitividad* (Projects MTM2011-28983-C03-02 and MTM2014-56949-C3-3-R), *Gobierno de Extremadura* (Projects GRU10142 and GR15106), and the *European Union* (European Regional Development Funds).

References

Alley, M. R., Quinlan, J. R., and Clarke, J. K., 1975. The prevalence of Mycoplasma ovipneumoniae and Mycoplasma arginini in the respiratory tract of sheep. New Zeal. Vet. J. 23, 137–141.

Brogden K.A., Lehmkuhl H.D., Cutlip R.C. 1998. *Pasteurella haemolytica* complicated respiratory infections in sheep and goats. Vet. Res. 29, 233- 254.

Caswell J. L., and Williams K. J., 2007. Respiratory System. In: Maxie G.M. (ed.) pathology of domestic animals. Saunders Elsevier, London, pp. 524-650.

Cooper, G. F., and Herskovits, E., 1992. A Bayesian method for the induction of probabilistic networks from data. Mach. Learn. 9 (4), 309–347.

Dash, D., and Cooper, G. F., 2004. Model averaging for prediction with discrete Bayesian networks. J. Mach. Learn. Res. 5, 1177–1203.

Druzdzel, M. J., 1999. In: SMILE: Structural modeling, inference, and learning engine and GeNIe: A development environment for graphical decision-theoretic models. American Association for Artificial Intelligence, 902–903

Ettema, J. F., Østergaard, S., and Kristensen, A. R., 2009. Estimation of probability for the presence of claw and digital skin diseases by combining cow- and herd-level information using a Bayesian network. Prev. Vet. Med. 92, 89–98.

Firestone, S.M., Lewis, F.I., Schemann, K., Ward, M.P., Toribio, J.A., Dhand, N.K., 2013. Understanding the associations between on-farm biosecurity practice and equine influenza infection during the 2007 outbreak in Australia. Prev. Vet. Med. 110, 28-36.

Gázquez, A., Redondo, E., Martínez, S., and Gómez, L., 2001. Ultrastructural study of ovine pulmonary pasteurellosis: involvement of neutrophils and macrophages. Histol. Histopathol. 16, 855-860.

Geenen, P.L., Van der Gaag, L.C., Loeffen, W.L.A., Elbers, A.R.W., 2011. Constructing naive Bayesian classifiers for veterinary medicine: A case study in the clinical diagnosis of classical swine fever. Res. Vet Sci. 91, 64-70.

Goodwin, K.A., Jackson, R., Brown, C., Davies, P.R., Morris, R.S., Perkins, N.R., 2004. Pneumonic lesions in lambs in New Zealand: patterns of prevalence and effects on production. New Zeal.Vet. J. 52, 175-179.

Gonçalves, R., Mariano, I., Núñez, A., Branco, S., Fairfoul, G., and Nicholas, R., 2010. Atypical non-progressive pneumonia in goats. Vet. J. 183, 219-221.

ITOVIC (Institut Technique de l'Élevage Ovin et Caprin), 1991. La ventilation thermique des bergeries et chevreries, Technipel.

Heckerman, D., Geiger, D. and Chickering, D.M., 1995. Learning Bayesian networks: the combination of knowledge and statistical data. Mach. Learn. 20, 197–243.

Hervás, J., Méndez, A., Gómez-Villamandos, J.C., Villalba, E., Díaz, E., Cano, T., Carrasco, L., Padró, J.M., Fernández, A., and Sierra, M.A., 1996. Etiologic and pathologic study of respiratory disease in lambs from intensive breeding facilities in southern Spain. Zbl. Vet. Med. C. 43(4), 221-231.

Jensen, F. V., 1995. Introduction to Bayesian Networks, University-College London Press, UK.

Jensen, F. V., and Nielsen, T. D., 2007. Bayesian Networks and Decision Graphics, Springer-Verlag, New York.

Jensen, T. B., Kristensen, A. R., Toft, N., Baadsgaard, N. P., Østergaard, S., and Houe, H., 2009. An object-oriented Bayesian network modeling the causes of leg disorders in finisher herds. Prev. Vet. Med. 89, 237–248.

Jones, G. E., Field, A. C., Gilmour, J. S., Rae, A. G., Nettleton, P. F., and McLauchlan, M., 1982. Effects of experimental chronic pneumonia on bodyweight, feed intake and carcase composition of lambs. Vet. Rec. 110, 168-173.7

Lacasta, D., Ferrer, L. M., Ramos, J. J., González, J. M., and De las Heras, M., 2008. Influence of climatic factors on the development of pneumonia in lambs. Small Ruminant Res. 80, 28-32.

Lewis, F. I., Brülisauer, F., and Gunn, G. J., 2011. Structure discovery in Bayesian networks: an analytical tool for analysing complex animal health data. Prev. Vet. Med. 100 (2): 109-115.

Lin, Y.C., Miles, R.J., Nicholas, R.A.J., Kelly, D.P., Wood, A.P., 2008. Isolation and immunological detection of Mycoplasma ovipneumoniae in sheep with atypical pneumonia, and lack of a role for Mycoplasma arginini. Res. Vet.. Sci. 84, 367–373.

Mawhinney, I., Woodger, N., and Knudsen, S., 2010. A typical interstitial pneumonia in grazing adult red deer (Cervus elaphus). J. Comp. Pathol. 143 (2-3), 209-212.

McCormick, B. J. J., Sánchez-Vázquez, M. J., and Lewis, F. I. 2013. Using Bayesian networks to explore the role of weather as a potential determinant of disease in pigs. Prev. Vet. Med. 110(1): 54-63.

McKendrick, I. J., Gettinby, G., Gu, Y., Reid, S. W. J., and Revie, C.W., 2000. Using a Bayesian belief network to aid differential diagnosis of tropical bovine diseases. Prev. Vet. Med. 47: 141–156.

Nash, M. L., Hungerford, L. L., Nash, T. G., and Zinn, G. M., 1997. Risk factors for respiratory disease mortality in lambs. Small Ruminant Res. 26, 53-60.

Neapolitan, R. E. 2004. Learning Bayesian networks (Vol. 1). Prentice Hall, Chicago IL.

Niang, M., Rosenbusch, R. F., Andrews, J. J. & Kaeberle, M. L., 1998. Demonstration of a capsule on Mycoplasma ovipneumoniae. Am. J. Vet. Res. 59, 557–562.

Nicholas, R. A. J., Ayling, R. D., and Loria, G. R., 2008. Ovine Mycoplasmal infections. Small Ruminant Res. 76, 92–98.

Otto, L., and Kristensen, C. S., 2004. A biological network describing infection with Mycoplasma hyopneumoniae in swine herds. Prev. Vet. Med. 66 (1–4): 141–161.

Oruç, E., 2006. The Pathologic and Bacteriologic Comparison of Pneumonia in Lambs. Turk. J. Vet. Anim. Sci. 30, 593-599.

Pearl, J., 1988. Probabilistic Reasoning In Intelligent Systems. Morgan Kaufmann, San Mateo, CA.

Plummer, P.J., Plummer, C.L., Still, K.M. 2012. Diseases of the Respiratory System. Sheep and Goat Medicine, Second Ed. Saunders Elsevier, Missouri, 126-149.

Sheehan, J. P. C., Joseph Brady, J., Ball, H., Doherty, M. L., Quinn, P. J., Robin A. J. N., and Markey, B. K., 2007. An aetiopathological study of chronic bronchopneumonia in lambs in Ireland. Vet. J. 173, 630-637.

Silander, T., Roos T., Kontkanen, P. and Myllymaki, P., 2008. Factorized normalized maximum likelihood criterion for learning Bayesian network structures. Proceedings of the 4th European Workshop on Probabilistic Graphical Models (PGM-08), 257-272.

Tang, J., Hu, M., Lee, S., Roblin, R., 2000. A polymerase chain reaction based method for detecting *Mycoplasma/Acholeplasma* contaminants in cell culture. J. Microbiol. Meth. 39, 121–126.

Yener, Z., Ilhan, F., Ilhan, Z., Saglam, Y.S., 2009. Immunohistochemical detection of Mannheimia (Pasteurella) haemolytica antigens in goats with natural pneumonia. Vet. Res. Commun. 33, 305-313.

Figure 1. ROC curves and AUC for Bayesian networks built using six algorithms: (a) Greedy Thick Thinning; (b) Bayesian Search; (c) Essential Graph Search; (d) Tree Augmented Naive Bayes; (e) Augmented Naive Bayes; (f) Naive Bayes.

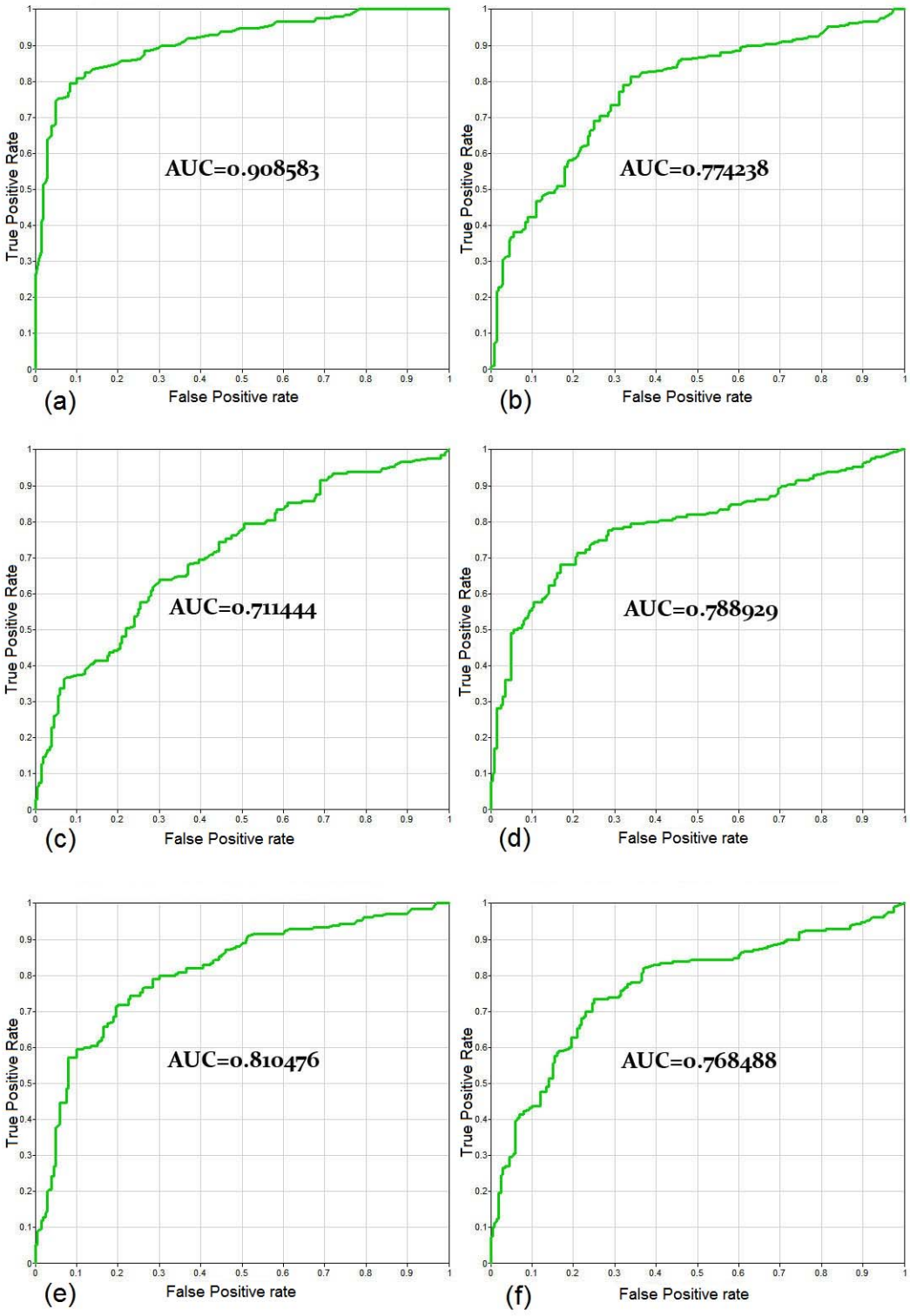

Figure 2. Bayesian network structure and estimated conditional probabilities.

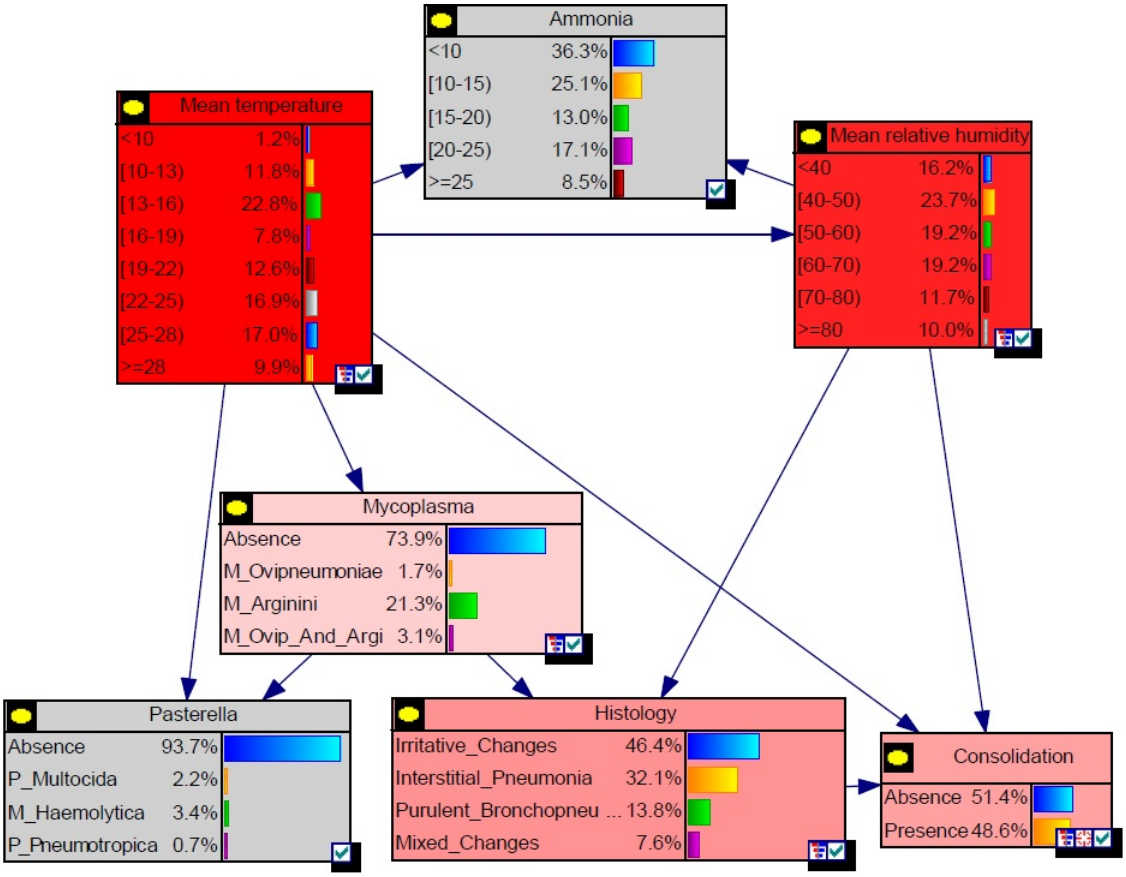

Figure 3. Sensitivity tornado chart for the presence of pulmonary consolidation.

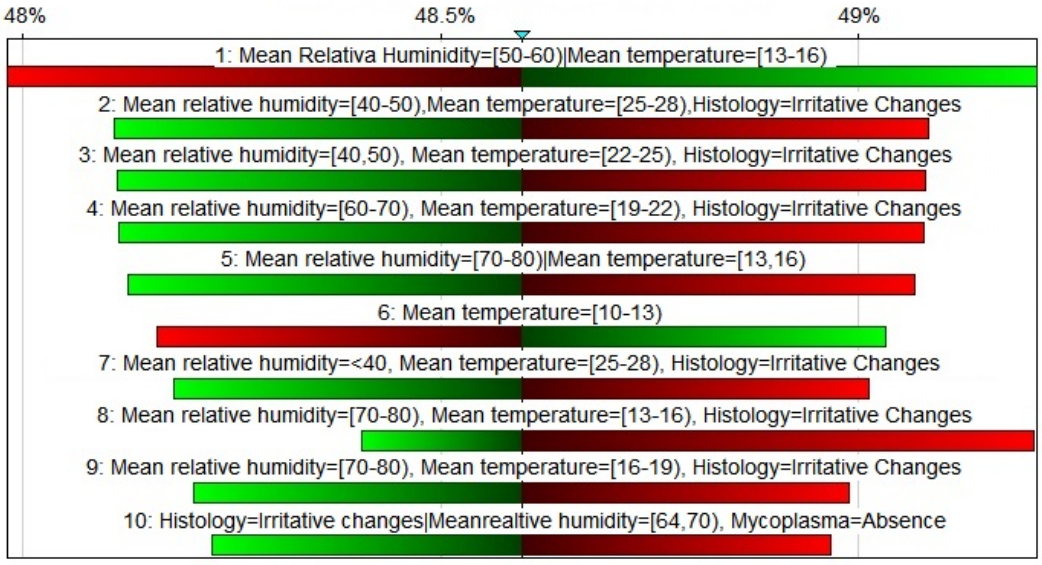

Figure 4. Monthly mean temperature and mean relative humidity.

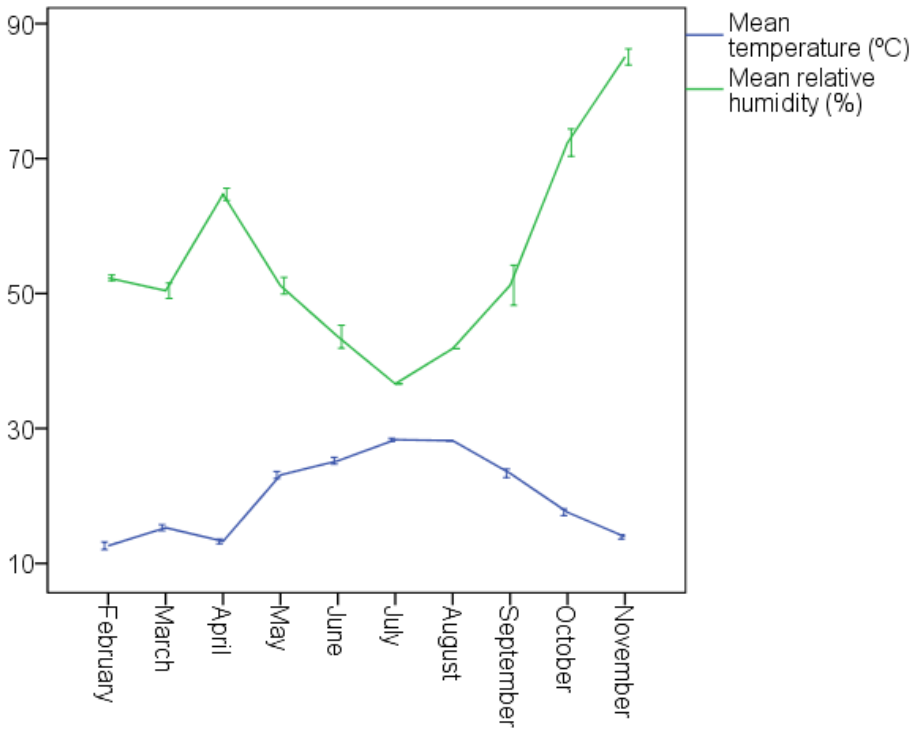

Table 1. Accuracy rates (in percentages) for Bayesian networks built using different algorithms.

Algorithms	**Accuracy**	**Absence**	**Presence**
Greedy Thick Thinning	85.6	90.0	79.5
Bayesian Search	71.9	71.0	72.9
Essential Graph Search	66.8	70.0	63.8
Tree Augmented Naive Bayes	75.6	80.5	71.0
Augmented Naive Bayes	75.1	80.0	70.0
Naive Bayes	71.7	78.5	65.2

Table 2. Evidence propagation for some possible environmental scenarios. Each environmental scenario gives a percentage of 100% for one category in one or more variables. The resulting probabilities are expressed in percentages.

Scenarios						
Variables	**1**	**2**	**3**	**4**	**5**	**6**
Temperature						
<10ºC					100.0	2.7
10-13ºC						2.5
13-16ºC	100.0	100.0	100.0			11.3
16-19ºC						50.1
19-22ºC						11.9
22-25ºC						10.6
25-28ºC						8.4
>28				100.0		2.5
Relative humidity						
<40			4.8	54.1	16.7	
40-50			4.8	23.5	16.7	
50-60		100.0	4.8	5.6	16.7	
60-70			4.8	5.6	16.7	
70-80	100.0		4.8	5.6	16.7	100.0
>80			76.2	5.6	16.7	
Ammonia						
<10	45.4	61.4		24.4	20.0	
10-15	3.2	2.3		24.4	20.0	
15-20	3.6	2.5	100.0	2.2	20.0	100.0
20-25	44.5	31.5		46.7	20.0	
>25	3.2	2.3		2.2	20.0	
Mycoplasma						
Absence	85.4	85.4	85.4	50.6	44.6	81.4
Ovipneumoniae	1.1	1.1	1.1	3.8	11.2	1.8
Arginini	10.8	10.8	10.8	40.8	29.9	13.9
Ovipneumoniae and *arginini*	2.7	2.7	2.7	4.7	14.3	2.9
Pasteurella						
Absence	91.8	91.8	91.8	95.1	60.1	90.1
Multocida	2.8	2.8	2.8	1.8	14.9	2.4
M. haemolytica	5.1	5.1	5.1	2.3	18.6	6.5
Pneumotropica	0.3	0.3	0.3	0.8	6.4	1.00
Histology						
Diffuse alveolar damage	70.7	30.2	52.3	41.9	39.4	68.7
Interstitial pneumonia	15.0	50.3	26.2	28.3	29.8	15.1
Purulent bronchopneumonia	9.7	13.9	13.4	19.4	17.9	11.1
Mixed changes	4.6	5.6	8.1	10.4	12.9	5.1
Pulmonary consolidation						
Absence	85.9	11.4	56.1	56.9	50.0	63.4
Presence	14.1	88.6	43.9	43.1	50.0	36.6

Table 3. Evidence propagation for some possible non-environmental scenarios (presence of microorganisms and histological groups). Each non-environmental scenario gives a percentage of 100% for one category in one or more variables. The resulting probabilities are expressed in percentages.

Scenarios							
Variables	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Temperature							
<10ºC	0.3	8.0	1.7	1.0	1.1	1.6	2.0
10-13ºC	11.6	15.4	12.1	10.6	14.0	10.3	12.8
13-16ºC	25.8	15.1	11.5	24.0	23.5	19.2	19.1
16-19ºC	8.8	8.0	2.7	10.2	5.7	6.1	5.5
19-22ºC	15.7	8.0	4.8	12.9	14.34	8.6	10.4
22-25ºC	15.7	8.0	23.1	15.6	17.8	18.8	18.3
25-28ºC	15.3	15.2	25.1	16.8	14.9	21.6	18.3
>28	6.8	22.4	19.0	8.9	8.7	13.9	13.5
Relative humidity							
<40	14.0	22.0	23.8	15.3	14.6	21.1	19.9
40-50	22.8	22.2	28.1	25.6	18.5	30.6	22.0
50-60	19.9	16.7	16.9	12.5	28.8	19.6	18.5
60-70	20.6	17.0	15.0	18.5	25.0	8.4	18.3
70-80	12.5	11.6	7.9	17.3	5.5	9.5	7.9
>80	10.3	10.5	8.5	10.9	7.6	10.8	13.5
Ammonia							
<10	36.03	32.1	37.5	36.7	36.0	36.6	34.2
10-15	25.6	27.1	23.9	25.4	25.0	23.8	26.3
15-20	14.0	10.8	10.0	13.4	12.4	13.3	12.4
20-25	16.2	22.4	19.2	16.6	17.4	18.1	17.6
>25	8.2	7.7	9.4	7.9	9.2	8.3	9.4
Mycoplasma							
Absence	100.0			83.8	75.6	52.5	45.1
Ovipneumoniae		100.0		0.8	1.4	2.9	6.0
Arginini			100.0	13.4	20.5	39.4	39.6
Ovipneumoniae and *arginini*				2.0	2.4	5.2	9.3
Pasteurella							
Absence	100.0	93.7	93.7	93.6	93.78	93.9	93.8
Multocida		2.2	2.2	2.2	2.2	2.1	2.2
M. haemolytica		3.4	3.4	3.5	3.3	3.3	3.3
Pneumotropica		0.7	0.7	0.7	0.7	0.7	0.7
Histology							
Diffuse alveolar damage	52.6	22.0	29.2	100.0			
Interstitial pneumonia	32.9	27.3	31.0		100.0		
Purulent bronchopneumonia	9.8	23.6	25.6			100.0	
Mixed changes	4.7	27.0	14.2				100.0
Pulmonary consolidation							
Absence	53.1	45.8	47.1	63.3	45.8	31.7	38.0
Presence	46.9	54.2	52.9	36.7	54.2	68.8	62.0