



HAL
open science

A Diagnosis Support System for Veterinary Necropsy based on Bayesian Networks

Vianney Sicard, Sébastien Assié, Laetitia Dorso, Florian Chocteau, Sébastien
Picault

► **To cite this version:**

Vianney Sicard, Sébastien Assié, Laetitia Dorso, Florian Chocteau, Sébastien Picault. A Diagnosis Support System for Veterinary Necropsy based on Bayesian Networks. 13th International Conference on Agents and Artificial Intelligence, Feb 2021, Online Streaming, France. pp.645-654, 10.5220/0010223106450654 . hal-03342265

HAL Id: hal-03342265

<https://hal.inrae.fr/hal-03342265>

Submitted on 1 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

A diagnosis support system for veterinary necropsy based on Bayesian networks

Vianney Sicard^a, Sébastien Assié^b, Laëtitia Dorso^c, Florian Chocteau^d and Sébastien Picault^e

*INRAE, Oniris, BIOEPAR, 44300, Nantes, France
firstname.lastname@inrae.fr*

Keywords: Veterinarian autopsy, Bayesian network, Decision-support tool, Cattle diseases.

Abstract: Veterinary autopsy requires a high level of expertise and skills that not all veterinarians necessarily master, especially in the context of the desertification of rural areas. The development of support systems is a challenging issue, since such a tool, to be considered relevant and accepted by practitioners in their diagnosis process, must avoid any black box effect. The diagnosis support system we introduce here, IVAN (“Innovative Veterinary Assisted Necropsy”), aims to engage the user in an explicit, understandable, validable and reviewable process, able to cope with the specific issues of cattle necropsy. Besides, it provides uncertainty management to deal with approximate lesion descriptions. IVAN relies on a Bayesian network to infer relevant proposals at each step of the diagnostic process. IVAN was trained on a set of real autopsy cases from autopsy reports, and its performance was assessed using another set of reports. In addition, the tool had to provide results in short response time and be able to run the application on mobile device and web server. In addition to demonstrating the feasibility of the approach, IVAN is a first step towards other support systems in other species and in broader contexts than autopsy.

1 INTRODUCTION

Autopsy (or necropsy) is the macroscopic morphological examination of all organs, based on the dissection of a corpse to either determine the cause of death, or identify an ongoing pathological process that required euthanasia. In cattle, autopsy is often necessary, when sudden death of an animal or serial mortality occurs in a farm, to provide the essential elements for a documented diagnosis and to set up the most appropriate measures for the other animals in the herd, potentially impacting public veterinary health. Veterinary autopsy requires very specific skills (anatomy, diagnostics, etc.) that not all field veterinarians necessarily master. The increasing scarcity of veterinarians in rural areas makes this issue even more critical because of the potential severe sanitary and economic impact of flawed disease detection (even post-mortem) in livestock. Connected intelligent medical tools will provide a substantial assis-

tance to field veterinarians by guiding them throughout the diagnostic process, proposing the most relevant organ to look at, helping to identify potential diseases, and suggesting complementary analyses.

The use of expert systems is a serious opportunity to address the scarcity of veterinary practitioners with a specialization in necropsy. With the joint development of artificial intelligence (AI), data availability and increased information technology (IT) resources, it is possible to develop reliable systems to model a realistic medical decision-making process in order to assist non-expert veterinarians. However, the user must be able to validate each step of the process, hence, such a system must avoid a black box effect by making explicit, unambiguous and understandable proposals that can be assessed throughout the diagnostic process.

The transfer of veterinary autopsy expertise to field practitioners through a support system based on AI is a real challenge. Expert systems dedicated to diagnosis support have been developed in human medicine since the 60s, but they generally target a specific disease and require to integrate a large amount of human knowledge. The specificity of autopsy is the multiplicity of possible diseases and of intermediary steps (lesions, organs, morphological di-

^a <https://orcid.org/0000-0002-4909-5544>

^b <https://orcid.org/0000-0002-8291-9533>

^c <https://orcid.org/0000-0001-7790-951X>

^d <https://orcid.org/0000-0003-3276-4593>

^e <https://orcid.org/0000-0001-9029-0555>

agnosis...), which leads to a highly combinatory process. In this paper, we introduce IVAN (“Innovative Veterinary Assisted Necropsy”), a system based on Bayesian Networks (BN) with a compromise between veterinary expertise and data-based learning, to provide the clearest possible process and leave the choice and validation of each diagnostic step by the veterinarian. To the best of our knowledge, there is currently no similar solution in human nor veterinary medicine which implied to develop new methods able to address the specificities of necropsy.

The paper is structured as follows: first, we present related work and their limitation, together with the Bayesian approach; second, we introduce the principles and algorithms implemented in IVAN; third, we present how we carried out the evaluation of the system.

2 BACKGROUND

In this section, we provide an overview on existing work on diagnosis support methods. We also briefly present the principles and interest of BN, and finally emphasize the specificities of cattle necropsy.

2.1 Existing diagnosis support methods

Existing medical support systems (*e.g.* Munin (Andreassen et al., 2001), Prostanet (Lacave and Díez, 2003)) focus on a few diseases (only one most of the time) in living human beings, and generally lack explicit or understandable process. The first (de Dombal et al., 1972), developed in the 1960s, concerned only heart disease and acute abdominal pain. They implemented Bayes’ naive method and got good results on simple issues. However, they were limited because the observations were not always correlated. Subsequently, more recent systems used uncertain reasoning (*e.g.* Munin), but, even if the diagnosis proposals were very close to the expert’s, many inconsistencies suggested revising the underlying assumptions. Mycin (De Baets and Fodor, 1999) used BN to solve both issues: uncertain reasoning and consistency. Thereafter, softwares like Prostanet provided robust solutions with a very high level of expertise. More recently, McKendrick (McKendrick et al., 2000), Seidel (Seidel et al., 2003), Greenen (Greenen et al., 2011) and Aristoteles (Aristoteles et al., 2019) have confirmed that expert systems based on BN are relevant for diagnosis assistance, especially because of its usage potential for use in a wide range of epidemiological disease situations. However, the solutions mentioned above are applied to a single disease,

whereas the solution we propose in the context of veterinary necropsy handles hundreds of diseases.

Other methods, based on deep learning methods, particularly artificial neural network (Amato et al., 2013), also focus on a single disease. In addition, they do not enable explicit and understandable diagnostic process (black box effect), which, currently, remains a strong limitation to acceptability among veterinarians. Hence, we preferred to rely upon a Bayesian approach.

2.2 Bayesian networks

A BN can be defined as a probabilistic graphical model representing random variables (Ben-Gal, 2008). It is both a knowledge representation and reasoning frame, a system for calculating conditional probabilities, and the underlying architecture for developing an expert system.

A BN is a directed acyclic graph (DAG) composed of sets of nodes connected by edges. Nodes represent variables in the Bayesian sense: observable quantities, latent variables, unknown parameters or hypotheses. Nodes and variables are equivalent (a node represents a single variable and a variable can be represented by only one node). The parameters describe how each variable relates probabilistically to its parents. Edges represent direct causal relationships between nodes.

The DAG of a BN necessarily respects the Markov property, *i.e.* a node is independent of all its non-descendent conditionally on its parents (Pearl, 2009) so we have a joint probability density (Pearl, 1982) (1).

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \pi_{X_i}) \quad (1)$$

Where π_{X_i} is the parents of X_i .

Each node is conditionally independent of its non-descendent.

Each node is endowed with its own conditional probability table (CPT) (Figure 1) which gives the probabilities of a variable with respect to the others.

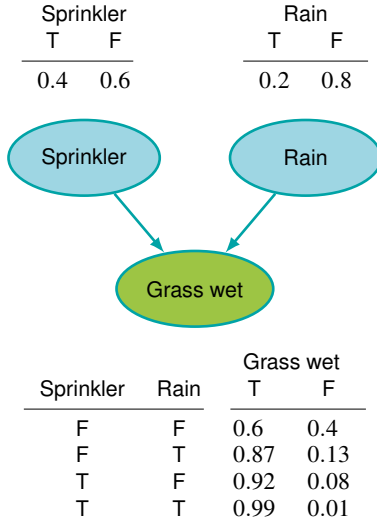


Figure 1: Example of a trivial BN and associated CPT

The marginal distribution for a given node with children can be calculated applying the marginalization algorithm (Pearl, 1982; Smail, 2004):

As an example, the calculation of the probability to have grass wet true ($P(G_T)$)

Example of calculation of $P(G_T)$ from the Fig. 1: (G = Grass wet, S = Sprinkler, R = Rain)

$$\begin{aligned}
 P(G_T) &= P(G_T|S_F, R_F) \cdot P(S_F) \cdot P(R_F) \\
 &+ P(G_T|S_F, R_T) \cdot P(S_F) \cdot P(R_T) \\
 &+ P(G_T|S_T, R_F) \cdot P(S_T) \cdot P(R_F) \\
 &+ P(G_T|S_T, R_T) \cdot P(S_T) \cdot P(R_T)
 \end{aligned}$$

$$\begin{aligned}
 P(G_T) &= (0.6 * 0.6 * 0.8) \\
 &+ (0.87 * 0.6 * 0.2) \\
 &+ (0.92 * 0.4 * 0.8) \\
 &+ (0.99 * 0.4 * 0.2)
 \end{aligned} \tag{2}$$

$$P(G_T) = 0.766$$

The probability of non-observed data can be inferred in the BN, following the direction opposite to causal effect by applying Bayes' theorem (BT) (3) or its generalization to multiple nodes (4).

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} \tag{3}$$

$$\begin{cases}
 P(A|X_1, X_2, \dots, X_n) = \alpha \left[\prod_{i=1}^n P(X_i|A) \right] \cdot P(A) \\
 \text{with } \alpha = \frac{1}{\sum_{i=1}^n P(A|X_i)}
 \end{cases} \tag{4}$$

where $\{X_1, \dots, X_n\}$ is the vector of child nodes of A (all X_i assumed independent one from each other).

2.3 Animal necropsy and available data

The process of veterinary necropsy followed by specialized practitioners is composed of several inference steps (Fig. 2). First, after recording general information on the animal (sex, breed, age), the veterinarian decides to target several organs (step 1), on which to search for lesions. Lesions consist in abnormal shape, size, color, consistency, content or distribution of either parts of organs or of whole organs, each combination of these features resulting in a specific morphological diagnostic (MD). Then, the practitioner deduces possible diagnoses from one or various observed MD (step 2). Most of the time, making the final diagnosis is possible only after gaining additional discriminant information (step 3), e.g. by selecting other organs to look at, or performing biological tests. The diversity of possible diseases and MD, as well as the difficulty to identify proper MD from observed lesions, are a key issue in animal necropsy and rely both on a vast theoretical knowledge and on daily experience.

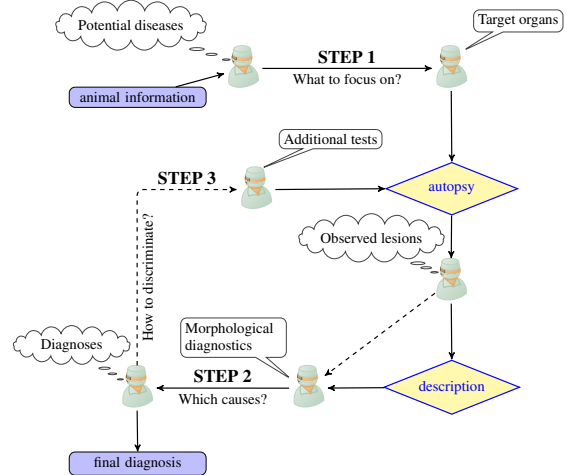


Figure 2: Main steps of the diagnostic process. Information deduced by the practitioner can either stay implicit or be mentioned explicitly in the final necropsy report.

To model this process, two data sources were immediately available. The first one comes from four years of autopsy reports performed on cattle by veterinary necropsy specialists in a necropsy service in a French veterinary school. The second is a corpus of theoretical knowledge compiled by a veterinary necropsy specialist concerning the description of lesions, morphological diagnostic (MD) and dis-

eases. Recorded data is composed of epidemiological and health information from reports for 783 cases of bovine autopsies, involving 543 distinct MD, covering a total of 152 diseases. In each report, the cause of death is attributed to a single disease based on the observed MD. Theoretical knowledge is indeed larger with a description of 379 diseases, involving potentially 1914 MD (some of them being quite rare). Hence, theoretical data can be used to complement observations.

3 PRINCIPLES AND METHOD

In this section, we present the architecture and algorithms used to implement our diagnostic support system for veterinary necropsy (IVAN).

3.1 Intrinsic complexity and technical constraints

Using a classical BN structure (with each disease represented by a Boolean node) would involve 925 nodes (i.e. 1850 Boolean parameters). Their combination would lead to a huge computation time while the final application must be used on the field and therefore must produce results in reasonable time. Since autopsy reports conclude with a single disease, it was possible for diseases to replace the multiple Boolean nodes by a single node with multiple values. Besides, several optimization methods were applied in the processing of BN calculations.

The challenge was to address the following issues simultaneously:

- no black box: the diagnosis process is not a straight deduction from an initial set of input data; it must instead be decomposed into several steps which are all explicit, understandable, validable and reviewable by the user and are expected to guide the user and propose a prioritization of possible options
- uncertainty management: assist and guide non-experts by providing them the most relevant proposals for accurate input information
- reactivity: short response time even with highly combinatorial process and mobile application
- versatility: the application must run on both a mobile device and a web server

To solve versatility issues, we use the Meteor.js framework which enables developing once in a single language, here Javascript (JS), and then compile the code for each platform. This comes with a limitation

due to JS, which is not a language dedicated to scientific computing, and provides currently no specific library for BN.

3.2 Calculation of conditional probability tables

To ensure that the diagnostic process followed by IVAN mimics the usual diagnostic process as expected by the practitioner, the structure of the BN was built with the help of veterinary necropsy specialists, not by learning the structure from the data. The data provided by necropsy reports were restructured into tables (in xls format) directly by the veterinary necropsy specialists, according to the specific structure of the global BN (Naïm et al., 2011). To calculate each CPT corresponding to the different nodes of the BN, we used the R library *bnlearn* (Scutari and Denis, 2015) which provides outputs in its specific format (*.bif*).

Examples :

- description of the node *age_group* defined from veterinary necropsy expertise:

```
variable age_group {
  type discrete [ 5 ] {
    ADULT, YOUNG, NEWBORN, CALF, OLD
  };
}
```

- declaration of the CPT between the disease *Mannheimia (BR10)* and the MD *Pulmonary emphysema (RP2430)* calculate from data (the values below give, on each line representing the presence of MD *RP2430*, the probability of having or not the disease *BR10*):

```
probability ( BR10 | RP2430 ) {
  (present) 0.2403846, 0.7596154;
  (absent) 0.01150121, 0.98849879;
}
```

To calculate CPT *bnlearn* needs the following inputs:

- the graph structure in a specific syntax, e.g. for BN1: `[sex][breed][age_group][diseases|sex:breed:age_group]`,
- the data table structure according to the BN structure,
- the method to use (here Bayesian method with an imaginary sample size (iss) of 10 (Scutari and Denis, 2015)).

The output files generated by *bnlearn* are parsed by the application to build the different BN needed.

3.3 Belief propagation

Belief propagation in a BN can be calculated using either polytree propagation (PP) (Rebane and Pearl, 1987) or junction tree (JT) (Madsen and Jensen, 1998). JT is usually used when the BN cannot be expressed as a polytree. In our case the BN is a polytree (Fig. 3), so we used the PP algorithm.

To optimize calculation, marginal distributions are computed “just in time” and results are stored for later use. The propagation is calculated only for the branch of the BN concerned with the Depth-First Search algorithm. Thus, marginal distribution goes up to the considered node and the path can be reused without recalculation.

Belief propagation is based on the opposite effect to causal effect. As mentioned above, generalised BT, which involves only a single layer of the BN, limits the number of calculations to perform. The computational cost of generalised BT is proportional to the number of variables involved.

3.4 Subgraph division

The formalisation of the diagnosis process is composed of three main steps corresponding to the different stages of the veterinarian’s reflection process to make a diagnosis (Fig. 4):

1. determine a set of potential diseases (and therefore their concerned organs to be autopsied) based on animal information
2. diagnose diseases from a set of MD observed during autopsy
3. propose additional tests to discriminate between possible diseases

Each of these steps are conditioned by the choice of the veterinarian from the list of proposals given by the system. The respect of these validation breakpoints is crucial to ensure that the system is not perceived as a monolithic black box which would provide final outputs directly from inputs. On the contrary, each step is seen as mimicking the corresponding reasoning step in the veterinary necropsy specialist’s diagnosis process.

It is therefore possible to divide the whole BN into three subgraphs representing these steps. This significantly reduces the calculation times for BN processing, making it possible to implement this method in JS, hence enabling the final application to be usable in the field.

To make the divisions, we had to adapt the structure of the BN with a specific approach consisting of grouping nodes that have similar specification (Smail

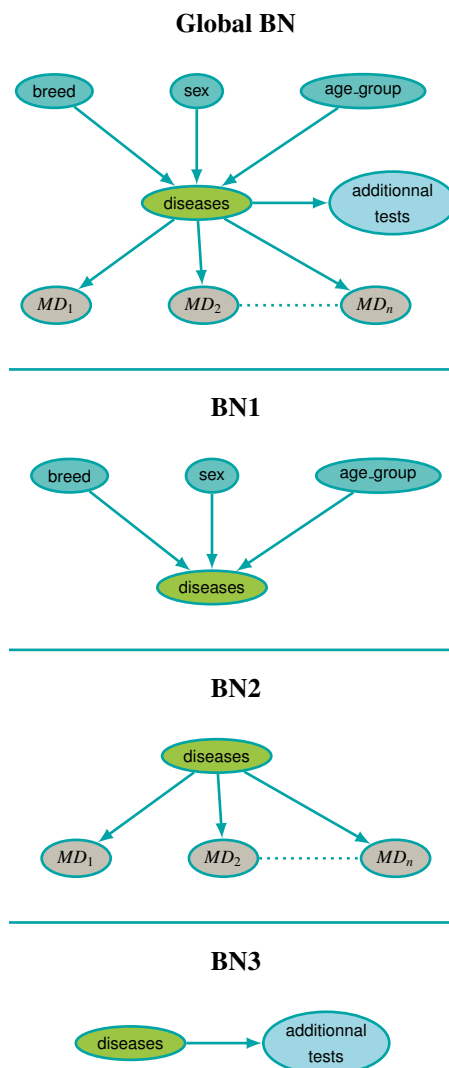


Figure 3: Global architecture of the BN after node grouping, and the different sub-graphs: BN1 → animal information to diseases, BN2 → diseases to MD, BN3 → disease to additional tests. These subgraphs correspond to the successive steps of the diagram process (Fig. 4)

and Raoult, 2005), *i.e.* grouping each disease node into a single one, compound of the parameters corresponding to each possible disease, and similarly for additional tests. The disease node is used as a pivot node for each sub-graph (Fig. 3).

For the first division, the disease node contains the parameters corresponding to each possible disease. This node has three parent nodes corresponding to the animal’s information (sex, age, breed). The second division is the principal BN which produces the diagnosis. The disease node contains the parameters corresponding to each possible disease. This node has several children corresponding to the different possible MD. The third division is the simplest, with only

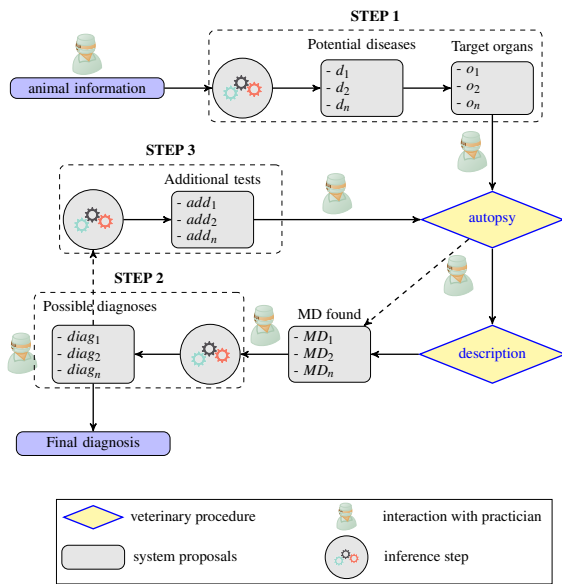


Figure 4: Main steps of the diagnostic process in IVAN. The veterinarian chooses whether to acknowledge or revise proposals at each key step identified previously.

two nodes connecting diseases and additional tests.

At this stage, the practitioner can select the diseases which seems the more probable to get a proposal of additional test for further discrimination.

3.5 Clustering of subgraph inferences

To facilitate the interpretation of outputs inferred in each subgraph (hence user's subsequent choice), the proposals are not displayed directly with their calculated probability value, which is expected mainly to provide a confidence level and a relative order of proposals. Thus, results are clustered into 1 to 3 categories (High, Medium and Low probabilities) without changing their order (Fig. 5), using the K-means algorithm (Lloyd, 1982) with Euclidean distance. The relevant number of classes is determined by the Silhouettes method (Rousseeuw, 1987), to ensure that *e.g.* three medium-probability outputs are gathered into the "Medium probability" class rather than dispatched into High, Medium and Low probability classes respectively.

This approach is a way to cope with the uncertainty of observations or descriptions, which results in the relatively low relevance of providing the user with absolute probability values. According to veterinary necropsy specialists' evaluation of the tool, this qualitative classification was much more meaningful than a list sorted by probability value.

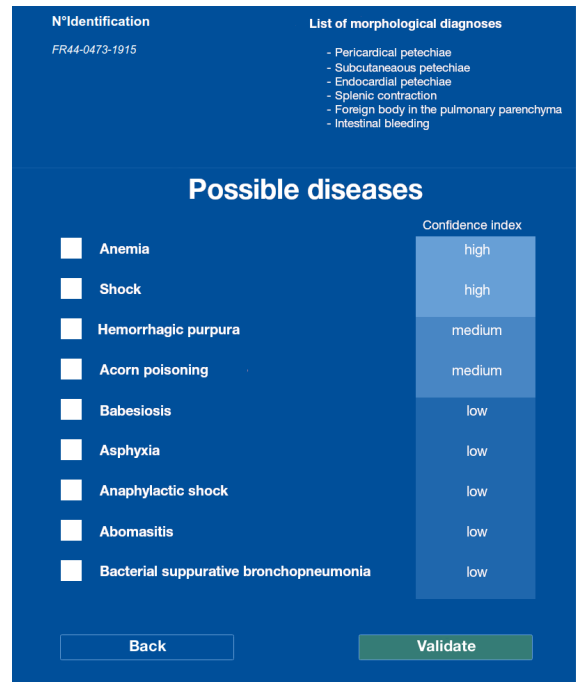


Figure 5: Screenshot of IVAN proposing diagnoses (center, left) grouped by probability levels (center, right) ranging from high to low, based on observed MD (upper right corner)

3.6 Morphological diagnosis discrimination

To perform the diagnostic process with the BN, the veterinarian must enter the appropriate MD name into the system. This can be an issue if the veterinarian is not familiar with MD identification. To solve this problem, MD can be inferred from the description of organ lesions, even imprecise. Each MD was described by veterinary necropsy specialists according to the following criteria: size, shape, colour, consistency, content and distribution. In order to manage the uncertainty of descriptions, veterinary necropsy specialists have established a list of valid values for each criterion and determined a distance between those values to qualify their differences.

Each criterion k is associated with a set of n_k valid values a_i^k (*e.g.* $k = \text{colour} \implies a_i^k = \text{white}$) with a distance measure between each pair of valid values ($|a_i^k - a_j^k|$) defined by veterinary necropsy specialists. The theoretical description of a lesion on an organ, associated to a specific MD, is a vector : $V^* = (a_i^{*k})_{i=1, n_k}$ with $k = \{\text{shape, size, color, consistency, content, distribution}\}$ and $(a_i^{*k})_{i=1, n_k}$ being the set of valid values for criterion k . The description of an observed lesion, entered by the veterinary during the diagnosis process, is a vector $V = (a_i^k)_{i=1, n_k}$. The

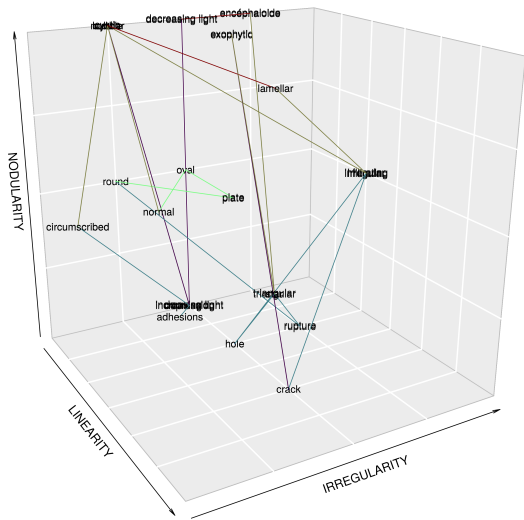


Figure 6: Example of the description of distances between parameters for the shape of the lesion

distance matrix $D = (|a_i^{*k} - a_j^k|)_{i,j}$ allows to identify the most relevant MD associated to the lowest value of $\sum_k D_{ik}$. (Fig. 6), proposed for validation to the user.

3.7 Supervised learning

IVAN is enriched by supervised learning based on its usage. The results of each diagnostic step are stored in the database in separate table. These records are periodically validated or discarded by veterinary necropsy specialists and then re-injected into the process of the calculation of conditional probability tables to update them. This process is semi-automated, the validation step by a veterinary necropsy specialist being the only section executed by humans. For now, this step is mandatory to ensure that user-generated information is reliable enough (especially because IVAN user keeps the full responsibility for the final diagnosis).

4 TESTS AND VALIDATION

4.1 Execution time

The use of the application on the field requires low response times. We have tested it on a server with 2 GB RAM and 2 CPUs. The test of performance was done on the different steps separately to determine which one takes the more execution time (Table 1).

As expected, the more MD there are, the longer the execution time of the second step is, but this remains within an acceptable time range for real-world

Table 1: Typical execution time for each BN corresponding to the three main steps. The crucial step (from MD to disease) was assessed with both a simple case (1 MD) and the most complex case found on reports (6 MD).

BN1 (organs)	BN2 1 MD	BN3 6 MD	BN3 < 1s
< 2s	< 1s	< 15s	< 1s

use, compared to the time spent by the veterinary for necropsic exams (about 1h for an adult cow).

4.2 Protocol for assessing diagnosis proposals

To carry out a preliminary evaluation of IVAN, we used a set of 40 new autopsy reports that were not included in the set of learning reports. We considered two criteria: a score with respect to the best proposal, and the relevance of the diagnostic. Both can be assessed for each step of the process.

The score refers to the relative position (based on the rank from 1 to number of proposals $nbProp$) of a real diagnosis item (organ, MD or disease mentioned in the autopsy report) in the list of IVAN proposals for the corresponding step, ordered by decreasing probability (5). A score between 1 and 0.5 means that the diagnosis was found in the first half of the list proposed by the software.

$$\begin{cases} score = 0 & \text{if real diagnosis not proposed} \\ score = \frac{nbProp - rank + 1}{nbProp} & \text{otherwise} \end{cases} \quad (5)$$

The relevance refers to the presence or absence of each diagnosis in the list of results, which means that IVAN's proposals are considered relevant if one of them matches the real-case diagnosis as mentioned in the final report. This measure was defined in accordance with veterinary necropsy specialists' point of view, to account for the lack of explicit intermediary hypotheses, possibly formulated during the necropsic process, but discarded from the final report. Indeed, when used by a non-expert veterinarian in real situations, the correct diagnosis may have a poor score, because alternative diagnoses that are highly probable without prior information may be discarded later due to additional exams. Thus, the relevance is a way to assess whether IVAN is able to propose the right diagnosis among its proposals.

4.3 Results

Figure 7 shows that most of the results are in first half of the list for all subgraphs outputs. The low score for

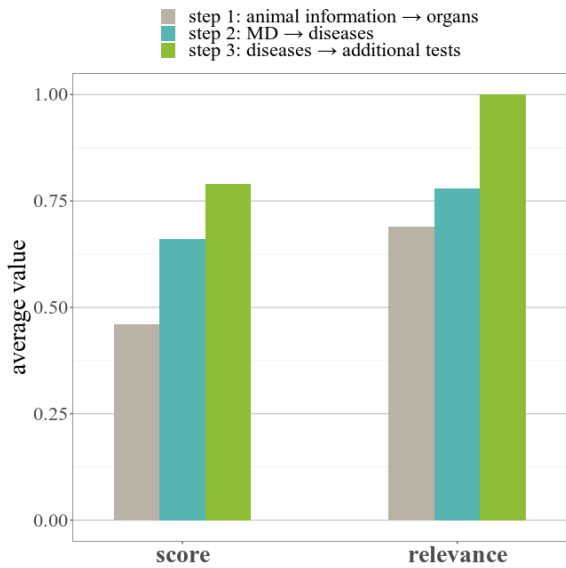


Figure 7: Average score and relevance of IVAN on a sample of 40 new reports

the first step (0.46) is due to the wide range of possibilities at this stage of the diagnosis, but in terms of performance, it can be considered that the application is able to provide sound advice. In contrast, the third step has very high score. This is due to the limited range of additional tests.

The real interest lies in the second step, which corresponds to the diagnostic step. Most of the results are at the top of the first half of the list (0.79) and provide the correct diagnosis in 78% of cases. The result is in the high average of a veterinarian practitioner who is not a specialist in cattle necropsy. We can probably achieve better results with a larger learning sample.

5 DISCUSSION AND PERSPECTIVES

To the best of our knowledge IVAN is the first system to support the diagnosis of veterinary necropsy based on Bayesian Networks.

Necropsic diagnosis requires a high level of expertise, which is not necessarily mastered by field veterinarians. IVAN is a helpful assistant for practitioners by providing relevant proposals at each step of the necropsic procedure without any black box effect and leaving the final decision to the veterinarian. The diagnosis process of the system is fully understandable, validable and reviewable by the user, including uncertainty management by guiding the veterinarian in describing the lesion to give the system the exact mor-

phological diagnosis required.

The first performance tests are encouraging, both in terms of execution time and relevance of the diagnostic proposals. According to veterinary necropsy specialists, who are also lecturers at a public veterinary school, IVAN results would be higher in average than students results in final year of veterinary school, and at least equal or better than field veterinarians not expert but familiar with autopsies.

5.1 Further assessment

The purpose of a tool like IVAN is to enhance the reliability of the necropsies performed by field veterinarians, especially in the context of an increased scarcity of veterinary necropsy specialists. Thus, the tool should either make non-experts more confident in their diagnosis, or come to a diagnosis at all.

Thus, to evaluate IVAN in more realistic conditions as an intelligent assistant, we need to compare the outcomes of an interaction between a non-expert practitioner and the software throughout a necropsy, with two extremes: on the one hand, the conclusions of the veterinary necropsy specialist (used as a gold standard), and on the other hand, the conclusions of the non-expert, assumed incorrect in many cases (according to veterinary necropsy specialists).

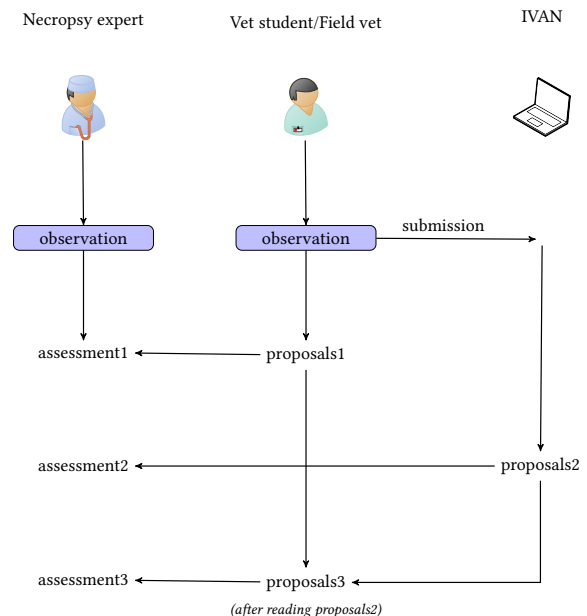


Figure 8: Evaluation protocol to be carried out for each of the three steps of the diagnostic process, to assess the added-value of IVAN for assisting a non-expert veterinarian during on-field necropsy

We plan to do so in the forthcoming months with the help of field veterinarians or of veterinary thesis

students engaged in a specialization in necropsy. For each step of the necropsic process, proposals made by the non-expert practitioner and by IVAN will be assessed by the veterinary necropsy specialist (in terms of score and relevance), then the non-expert practitioner will be provided with IVAN proposals and asked either to keep its conclusions or revise them, which will lead to a third assessment by the veterinary necropsy specialist (Fig. 8).

Thus, we will be able to evaluate to what extent the conclusions of the non-expert veterinarians may either be comforted by IVAN proposals, or enhanced by its suggestions, as a measure of the quality of the support provided by the machine. The study of how frequently erroneous conclusions of a non-expert are modified when confronted to correct conclusions by the software will also provide insights on the acceptability of the software as an “intelligent assistant”. Conversely, cases where IVAN proposals are incorrect will highlight cases underrepresented in learning data.

5.2 Perspectives

In addition to the intended experimental assessment, IVAN will be integrated in the coming months within a professional software platform to be proposed as an e-service to field veterinarians. This will be an opportunity to carry out a thorough assessment of the tool usage and provide data from real use cases. A procedure will be set up to assess the reliability of necropsies carried out with IVAN and integrate validated reports into the learning database.

The first results of IVAN on bovine necropsy make it possible, firstly, to consider other breeds and in a further step, to apply the process on living animals, including clinical signs instead of lesion description. Supporting necropsy for other breeds can rely upon the autopsy reports for bovine. That can be applied quite readily and quickly in the continuity of the current work. The post-mortem diagnosis can also benefit from knowledge regarding the ante mortem condition. The issues to address to do so are similar for the diagnosis of living animals, since they rely on the same kind of hypothetico-deductive inferences. This will be the major challenge for the next stages of IVAN, and for the development of a specific tool to support living animal diagnosis.

6 CONCLUSION

We consider our contribution a first step towards the development of diagnosis support systems for

veterinary necropsy, able to assist non-expert practitioners with relevant, reviewable, step-by-step recommendations mixing recorded cases and theoretical knowledge. IVAN’s database will be enriched by, on the one hand, new necropsy reports performed at the autopsy service, and, on the other hand, on-field records stored in IVAN performed by non-expert practitioners, after validation by veterinary necropsy specialists.

Though initiated in the context of bovine necropsy, our approach is generic enough to apply to other species and to living animals. The choice of a Bayesian Network for learning and probabilistic inference enables to cope with uncertainties encountered in the descriptions of lesions and resulting from the scarcity of data, but also to provide a natural decomposition of the diagnostic process, mimicking the inferences made by a veterinary necropsy specialist, hence facilitating the acceptance of the system by the user. Such systems could be an opportunity for AI to provide a substantial assistance to field veterinarians in the context of the medical desertification of rural areas.

ACKNOWLEDGEMENTS

This work is being carried out as part of the Telemedicine Business Chair, co-financed by MSD Animal Health and Oniris - Veterinary Medicine School of Nantes.

REFERENCES

- Amato, F., López, A., Peña-Méndez, E. M., Vañhara, P., Hampl, A., and Havel, J. (2013). Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine*, 11(2):47–58.
- Andreassen, S., Suojanen, M., Falck, B., and Olesen, K. G. (2001). Improving the Diagnostic Performance of MUNIN by Remodelling of the Diseases. In Quaglini, S., Barahona, P., and Andreassen, S., editors, *Artificial Intelligence in Medicine*, Lecture Notes in Computer Science, pages 167–176. Springer Berlin Heidelberg.
- Aristoteles, A., Adhianto, K., Andrian, R., and Nuhricha, Y. (2019). Comparative analysis of cow disease diagnosis expert system using bayesian network and dempster-shafer method. 10(4).
- Ben-Gal, I. (2008). Bayesian Networks. In *Encyclopedia of Statistics in Quality and Reliability*. American Cancer Society.
- De Baets, B. and Fodor, J. (1999). Van Melle’s combining function in MYCIN is a representable uninorm: An alternative proof. *Fuzzy Sets and Systems*, 104(1):133–136.

- de Dombal, F. T., Leaper, D. J., Staniland, J. R., McCann, A. P., and Horrocks, J. C. (1972). Computer-aided Diagnosis of Acute Abdominal Pain. *British Medical Journal*, 2(5804):9–13.
- Geenen, P., van der Gaag, L., Loeffen, W., and Elbers, A. (2011). Constructing naive bayesian classifiers for veterinary medicine: A case study in the clinical diagnosis of classical swine fever. 91(1):64–70.
- Lacave, C. and Díez, F. J. (2003). Knowledge Acquisition in PROSTANET – A Bayesian Network for Diagnosing Prostate Cancer. In Goos, G., Hartmanis, J., van Leeuwen, J., Palade, V., Howlett, R. J., and Jain, L., editors, *Knowledge-Based Intelligent Information and Engineering Systems*, volume 2774, pages 1345–1350. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137.
- Madsen, A. L. and Jensen, F. V. (1998). Lazy Propagation in Junction Trees. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, UAI'98, pages 362–369, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- McKendrick, I., Gettinby, G., Gu, Y., Reid, S., and Revie, C. (2000). Using a bayesian belief network to aid differential diagnosis of tropical bovine diseases. 47(3):141–156.
- Naïm, P., Wuillemin, P.-H., Leray, P., Pourret, O., and Becker, A. (2011). *Réseaux bayésiens*. Editions Eyrolles.
- Pearl, J. (1982). 1982 - reverend bayes on inference engines: A distributed hierarchical approach. page 4.
- Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146.
- Rebane, G. and Pearl, J. (1987). The Recovery of Causal Poly-trees from Statistical Data. In *Proceedings of the Third Conference on Uncertainty in Artificial Intelligence*, UAI'87, pages 222–228, Arlington, Virginia, United States. AUAI Press.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65.
- Scutari, M. and Denis, J.-B. (2015). *Bayesian Networks: With Examples in R*. Texts in Statistical Science. Chapman & hall/crc edition.
- Seidel, M., Breslin, C., Christley, R. M., Gettinby, G., Reid, S. W. J., and Revie, C. W. (2003). Comparing diagnoses from expert systems and human experts. 76(2):527–538.
- Smail, L. (2004). Algorithmique pour les réseaux bayésiens et leurs extensions.
- Smail, L. and Raoult, J. P. (2005). Successive Restrictions Algorithm in Bayesian Networks. In Famili, A. F., Kok, J. N., Peña, J. M., Siebes, A., and Feelders, A., editors, *Advances in Intelligent Data Analysis VI*, Lecture Notes in Computer Science, pages 409–418. Springer Berlin Heidelberg.