*Proceedings of the 12th International conference of the
Association of Institutions of Tropical Veterinary Medicine*

A BAYESIAN APPROACH IN VETERINARY EPIDEMIOLOGY

BERKVENS D.*, PRAET N., THYS E., SPEYBROECK N.

Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerp, Belgium

ABSTRACT

The interpretation of diagnostic test results serves as the basis from which to introduce the Bayesian view in statistics. The matter is familiar to all students and it offers the opportunity to explain the fundamental difference between information present in the data (the test results) and information, essential for a correct interpretation of the test results but entirely external to the data (information about the test characteristics). The predictive value of a positive/negative test result is a well-understood example of Bayes' formula. It is employed to explain the concepts of prior information, data likelihood and posterior information, whereby real-life situations are used as examples to ensure students fully grasp the link between the probability distributions and the biological meaning. The model is then expanded first to a multi-test situation and finally to more general risk evaluation problems. The necessary care is taken to ensure students recognise the role that prior information, be it hard data or expert opinion, plays in veterinary epidemiology and that they understand that a Bayesian approach is in fact the natural way to deal with it.

INTRODUCTION

Correct interpretation of diagnostic test results is an essential part of almost any epidemiological survey, study or experiment. It is very often not realised that test results on their own do not permit this interpretation and that external information in the form of the test sensitivity and test specificity is required to transform an apparent prevalence (proportion positive test results) in a true prevalence (proportion infected/diseased animals) (Enoe *et al.*, 2000). The combination of external (essentially *a priori*) information with observations forms the basis of the Bayesian statistical approach. A full overview of the use of a Bayesian framework in diagnostic testing can be found in Berkvens *et al.,* (2006). Diagnostic test interpretation forms a useful and intuitive way to introduce the Bayesian philosophy. The current paper outlines the different steps that are taken during a module on quantitative veterinary epidemiology and risk analysis

*Contact author: E-mail: dberkvens@itg.be

(Master of Science in Tropical Animal Health at the Institute of Tropical Medicine, Antwerp) to introduce this approach ensuring that the students fully grasp the link between the probability distributions and the biological meaning and recognise the role that prior information, be it hard data or expert opinion, plays in veterinary epidemiology and that they understand that a Bayesian approach is in fact the natural way to deal with this.

BAYES' THEOREM

Bayes' Theorem or Bayes' Formula (Bayes, 1763) introduces inverse probability to be used when calculating the probability of antecedent events based on the occurrence of consequent events. Think of the antecedent event as the point in time when the animal became infected and the consequent event as the result of a diagnostic test carried out at a later point in time. 'Infected' is used in the most general sense of the word and may be replaced by 'infested', 'diseased', 'carrier' or whatever term to distinguish this animal (a case) from a non-case.

The actual Bayes' Formula is:

$$P(Ante \mid Cons) = \frac{P(Cons \mid Ante) \cdot P(Ante)}{P(Cons)} \quad (1)$$

with: Ante = having become infected; Cons = positive test result

or, in terms of infection and test result:

$$P(D^+ \mid T^+) = \frac{P(T^+ \mid D^+) \cdot P(D^+)}{P(D^+)} \quad (2)$$

where D^+ refers to 'being infected' and T^+ to 'a positive test result'. The formula thus calculates the probability that a particular animal is infected given a positive test result. In the latter form, it becomes clear that this formula is what is commonly known in epidemiology as the predictive value of a positive test result and it is indeed exactly what inverse probability signifies: what is the probability that an animal was infected previously, given a positive test result now? Bayes' theorem is thus in fact a commonly used tool when having to decide whether or not to classify an animal as infected or not.

THE USE OF PRIOR INFORMATION

The need to combine prior information with observations introduces the concepts of probability and likelihood. The simplest system uses a beta distribution (Johnson *et al.*, 1995) to represent the prior probability distribution of (e.g.) test sensitivity and a binomial likelihood (Johnson *et al.*, 1992) for the test results (in a group of known infected animals) and shows how these are combined into an updated posterior beta distribution. The flexibility of the beta distribution is demonstrated and the influence and weighting of both prior information and data are explained.

GENERALISATION OF BAYES' FORMULA

Conceptually the most difficult step is to generalise Bayes' formula from a one-parameter situation (virtually not applicable in real-life) to the more general multi-parameter situation. This transition is achieved using multi-testing as an example. Multi-testing implies the application of different diagnostic tests to each individual and results in a rather complex model relating probabilities of the various possible test result combinations to the individual test characteristics including terms describing the interactions between the tests. The full model is described in Berkvens *et al.* (2006). Model development and computations are done in R and WinBUGS, two freely downloadable packages. Once more, particular attention is paid to ensure students fully appreciate where in the model the prior information is introduced and how the data are linked to this information through the probability equations. Interpretation of posterior probability distributions and various goodness-of-fit statistics is also dealt with, as well as model identifiability. Students should be able to construct their own models within the multi-test environment (e.g. using a second test conditional on the result of the first) and fully understand the Bayesian analysis approach. Throughout real-life examples are used to illustrate the different concepts (Dorny *et al.*, 2004; Geurden *et al.*, 2004).

APPLICABILITY OF THE BAYESIAN ANALYSIS

Once the students know how to develop a model in WinBUGS, the final step is to show that the Bayesian approach can also be used in more traditional statistical analyses, such as general and generalised linear models (Congdon, 2003). Also here, the distinction between external information and information contained in the data is clearly explained. Lastly, the usefulness of Bayesian estimation is demonstrated through the use of spatio-temporal analyses, using the various standard models as summarised in Lawson *et al.* (2003).

DISCUSSION AND CONCLUSION

Using the well-known example of diagnostic test result interpretation allows the introduction of the principles of Bayesian analysis, combination of prior information with data, in a relatively simple and intuitive way. The mathematical background can be included or not, depending on the level of pre-knowledge of the students. Real-life examples help to establish the link between the different components of the models and experimental data and eternal information. This approach allows the students to come to terms with the fact that data alone do not suffice in certain circumstances and that a Bayesian approach is the natural and preferred way to deal with this need to incorporate information external to the data.

REFERENCES

Bayes T., 1763. An Essay towards solving a Problem in the Doctrine of Chances. Philosophical Transactions of the Royal Society, 53 : 370-418.

Berkvens D., Speybroeck N., Praet N., Adel A., Lesaffre E., 2006. Estimating disease prevalence in a Bayesian framework using probabilistic constraints. Epidemiology, 17 : 145-153.

Congdon P., 2003. Applied Bayesian Modelling. John Wiley & Sons Ltd., Chichester.

Dorny P., Phiri I.K., Vercruysse J., Gabriel S., Willingham III, A.L., Brandt J., Victor B., Speybroeck N., Berkvens D., 2004. A Bayesian approach for estimating values for prevalence and diagnostic test characteristics of porcine cysticercosis. International Journal for Parasitology, 34 : 569-576.

Enoe C., Georgiadis M.P., Johnson W.O., 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true diseases state is unknown. Preventive Veterinary Medicine, 45 : 61-81.

Geurden T., Claerebout E., Vercruysse J., Berkvens D., 2004. Estimation of diagnostic test characteristics and prevalence of Giardia duodenalis in dairy calves in Belgium using a Bayesian approach. International Journal for Parasitology, 34 : 1121-1127.

Johnson N.L., Kotz S., Kemp A.W., 1992. Univariate Discrete Distributions. John Wiley & Sons Inc., New York.

Johnson N.L., Kotz S., Balakrishnan N., 1995. Continuous Univariate Distributions, Volume 2. John Wiley & Sons Inc., New York.

Lawson A.B., Browne W.J., Vidal Rodeiro C.L., 2003. Disease Mapping with WinBUGS and MLWIN. John Wiley & Sons Ltd, Chichester.