The Kenya Veterinarian 18(2) (1994) 264-266

- 6.4 -

DESCRIPTIVE AND PREDICTIVE MODELS OF THE EQUINE SARCOID IN THE DONKEY

REID, S.W.J[a]., GETTINBY, G[b].

Epidemiology embraces a number of disciplines in the study of disease and the factors that influence disease occurrence and distribution in a population. The concept of the use of integrated information technology in the storage, analysis and dissemination of knowledge is at the forefront of epidemiological thinking (Thrusfield, 1988). The holistic approach to disease causality and multivariate quantitative techniques applied to population data are the tools of the modern epidemiologist (Davies, 1983; Schwabe, 1993).

The equine sarcoid is a fibropapilloma that affects the skin of horses, donkeys and other equidae (Jackson, 1936; Ragland et al., 1970). There have been several investigations into the epidemiology of the disease. Results from these studies indicate that breed, sex, age and MHC type are significantly associated with susceptibility (Angelos et al., 1988; Mohammed et al., 1992). In distinction to the horse, we have recently reported on the disease in a large population of donkeys and described the results of analyses comparing epidemiological features of affected and unaffected animals (Reid et al., 1994) The aim of the present study was to use a large clinical database to develop multivariate models identifying risk factors associated with equine sarcoids in a large population of donkeys.

MATERIALS AND METHODS

The study population was defined using the criteria described by Reid et al. (1994), and Reid and Gettinby (1994). The variables used in this present study were those which had been found to differ significantly in their distribution in the sarcoid affected and unaffected populations (Reid et al., 1994), that is gender, age at entry to the population and duration of exposure. Three models were developed. First, the contingency tables generated in the previous study were combined to give a (gender x age at first exposure x duration of exposure) grouped data set based on the numbers of diseased animals alone. Secondly, tables were generated combining the diseased animal data with the numbers of healthy animals to provide proportions of diseased animals within each group. These data were analysed using a logistic model with a group of animals (healthy and diseased) as the experimental unit. Thirdly, the raw data were modelled using logistic regression, wherein the individual animal, healthy or diseased, was the experimental unit.

RESULTS

Applying the log-linear model with a Poisson distribution and log link function, the first model, consisting of the main effects and all two-factor interactions, produced a residual scaled deviance of 8.94 with 6 degrees of freedom. Only the removal of the gender x age at first exposure interaction and the age at first exposure x duration of exposure interaction produced a model with no significant increase in residual mean deviance. The linear predictor of the parsimonious model describing the diseased animals thus included the variables:

gender + age at first exposure + duration of exposure + gender x duration of exposure

[a]Veterinary Informatics and Health Group, Department of Veterinary Medicine, University of Glasgow Veterinary School, Bearsden, Glasgow, G61 1QH, UK.
[b]Veterinary Informatics and Health Group, Department of Statistics and Modelling Science, University of Strathclyde, George Street, Glasgow, G1 1XH, UK.

This model has been discussed in detail previously (Reid and Gettinby, 1994). When proportions of animals affected in each category were included as dependent variables in a model with a binomial distribution and logistic link function, the model including all two-factor interactions had a residual scaled deviance of 8.19 with 6 degrees of freedom. Removal of the gender x duration of exposure interaction indicated that this was the only significant two-factor interaction.

In the third model, in which the individual animal was the experimental unit, the model building process produced a more complex predictor. Age and duration of exposure variables were continuous and were measured in years. Gender was dichotomous with female as the reference category. The final model contained gender, age at first exposure, age at first exposure **2, duration of exposure and age at first exposure x duration of exposure. Because of the presence of continous variables, goodness-of-fit was assessed by the Hosmer-Lemeshow statistic. The statistic value was 7.04 which had a P-value equal to 0.53 with 8 degrees of freedom. Thus the model showed a good fit of the data.

DISCUSSION

The three models applied in the present study differ in the choice of experimental unit. Ignoring the gender x age at first exposure and age at first exposure x duration of exposure interactions did not significantly affect closeness of fit of either of the first two models, whilst removal of the gender x duration of exposure interaction did. Although the first model was constructed from counts of cases and was a description of diseased animals' relationships one with another it did not take unaffected animals into account. The model arising from the analysis of the proportions of affected to healthy animals produced similar results to the first model, i.e., only the gender x duration of exposure interaction was significant. Although both models are a description of a clinical state, the second model is better suited to predicting risks relative to the healthy population.

In the third model, quadratic and interaction terms were required in addition to main effect terms. The interaction term in the third model pointed more towards an age related then a gender/exposure effect evidenced in the first two models. However, despite the satisfactory fit of the third model, inspection of its diagnostic capability suggested that this may have been predominantly due to the large number of non-affected animals. It was concluded that although a good fit was obtained with the third model, other discriminatory factors such as genotype may substantially improve the predictive power. Thus, it was deduced that the second model, based on proportions of groups of donkeys affected, was likely to be an adequate, an effective herd management tool.

REFERENCES

Angelos, J.A., Oppenheim, Y., Rebhun, W., Mohammed, H. and Antczak, D.F.(1988) Evaluation of breed as a risk factor for sarcoid and uveitis in horses. Anim. Genet., 19: 417-425.

Davies, G. (1983) Development of veterinary epidemiology. Vet. Rec., 112:51-53.

Jackson, C. (1936) The incidence and pathology of tumours of domesticated animals in South Africa. Onderstepoort J. Vet. Sci. Anim. Ind., 6: 248.

Mohammed, H.O., Rebhun, W.C. and Antczak, D.F. (1992) Factors associated with the risk of developing sarcoid tumours in horses. Equine Vet. J. 24: 165-168.

Ragland, W.L., Keown, G.H. and Spencer, G.R. (1970) Equine sarcoid. Equine Vet. J., 2: 2-11.

Reid, S.W.J. and Gettinby, G. (1994) Modelling the equine sarcoid. Proc. Society for Veterinary Epidemiology and Preventive Medicine. Belfast, Northern Ireland, 1994. pp.167-173.

Reid, S.W.J., Gettinby, G., Fowler, J.N. and Ikin, P. (1994). Epidemiological observations on the equine sarcoid in a population of Equus asinus. Vet. Rec., 134: 207-211.

Schwabe, C.W. (1993) The current epidemiological revolution in veterinary medicine. Part II. Prev. Vet. Med., 18: 3-16.

Thrusfield, M. (1988) The application of epidemiological techniques to contemporary veterinary problems. Br. Vet. J., 144: 455-469.