B-4689
[1-5]

RESEARCH ARTICLE Indian Journal of Animal Research

Discriminant Analysis to Characterize Mastitis Resistant and Susceptible Karan Fries Cows and Response Operating Characteristics Curve for Comparing Regression/ Discriminant Model

Nisha Sharma, Sushil Kumar, D. Revanasiddu, Beena Sinha,
Ragini Kumari, Rebeka Sinha, I.D. Gupta, Archana Verma

10.18805/IJAR.B-4689

ABSTRACT

Background: The profitability in the dairy enterprises is strongly affected by udder health and, to a minor extent, by morphometry. The present study was conducted to differentiate mastitis susceptible and resistant animals based on udder and teat type traits infrequently exist.

Methods: The study was undertaken in ICAR-NDRI, Karnal of Haryana. A discriminate model was applied to differentiate mastitis susceptible and resistant Karan Fries cows for the study. For meeting the objective; 123 Karan Fries cows were subjected to developing a reliable mathematical function/equation (discriminant functions) for the provision of maximum separation among the mastitis resistant and susceptible animals. The hypothesis was designed to know if these udder and type traits appeal to mastitis-resistant and susceptible animals. The hypothesis for testing the equality of covariance matrices was depicted as: H0: $\sum_1 = \sum_2$ Vrs H1: At least one pair of sigma's $\sum$ is different.

Result: The model included contributory factors such as fore udder length, central ligament, udder balance, shortest distance of floor from fore teat (SDFT), shortest distance of floor from rear teat (SDRT), fore udder attachment (FUA), teat direction, Rear legs rearview, Hock development, udder depth, and udder width constructed and was found to demonstrate 69.1 percent accuracy with $p<0.001$ and the Karan Fries functions at group centroids are -0.169 and 0.443 respectively for mastitis resistant and mastitis infected animals. The model will facilitate 92.1 per cent of the cases to correctly classify for mastitis resistant animals and the remaining 7.9 per cent being misclassified into the mastitis group.

Key words: Discriminant function, Karan Fries, Mastitis, Udder and teat type traits, ROC curve.

INTRODUCTION

Crossbred animals are more susceptible to mastitis owing to their higher milk production thus mastitis is one of the key areas of research for sustainable milk production and animal welfare together with a reduction in economic losses due to treatment and culling of productive animals. Annual losses in the dairy industry due to mastitis was almost 2.37 thousand crore rupees in India (Lakshmi, 2016), and not only this but there is every chance that the better genetics may be lost permanently due to involuntary culling as a result of the mastitis in high genetic merit animals. The Karan Fries (KF) is a crossbred cattle (dairy strain) developed in India at the National Dairy Research Institute (NDRI) Karnal, Haryana. The present study was designed to fit a discriminate model where udder and teat type traits affecting udder health are integrated to differentiate a mastitis resistant and susceptible animal's condition. In India, Thirunavukkarasu (2003) designed a model to which can help to discriminate mastitic and non-mastitic milch animals based on udder morphology, non-genetic factors, and udder hygiene. Discriminate function models so far in India in animal sciences, have been applied successfully for discriminating the healthy and subclinical animals based on SCC in Holstein cows (Jadhav *et al*. 2019), and also it has been used to characterize three dairy cattle breeds on the basis of several milk characteristics (Leotta, 2004). The relative emphasis should be given to udder and teat type traits in selection indices along with production and reproduction performance. However, scanty literature is available about udder and teat type conformation traits in Karan-Fries cattle. In India, NDDB started recording udder

Division of Animal Genetics and Breeding , ICAR-National Dairy Research Institute, Karnal-132 001 Haryana, India.

Corresponding Author: Nisha Sharma, Division of Animal Genetics and Breeding, ICAR-National Dairy Research Institute, Karnal-132 001 Haryana, India. Email: nishasharma1777@gmail.com

How to cite this article: Sharma, N., Kumar, S.,Revanasiddu, D., Sinha, B., Kumari, R., Sinha, R., Gupta, I.D. and Verma, A. (2021). Discriminant Analysis to Characterize Mastitis Resistant and Susceptible Karan Fries Cows and Response Operating Characteristics Curve for Comparing Regression/ Discriminant Model. Indian Journal of Animal Research. DOI: 10.18805/ IJAR.B-4689.

Submitted: 22-06-2021 **Accepted:** 07-09-2021 **Online:** 29-10-2021

and teat traits in the INAPH project. Realizing the importance of these functional traits and as a research priority, the study designed for future analysis by taking a large sample size. By giving weightage to these traits, we may probably increase the disease resistance. The hypothesis was designed to know if these udder and type traits appeal to mastitis resistant and susceptible animals. The study's main objective was specifically based on developing a reliable mathematical function/equation (discriminant functions) for the provision of maximum separation among the mastitis-resistant and susceptible animals.

MATERIALS AND METHODS

Data

Data on 123 Karan Fries cows pertaining to udder and teat traits *viz.*, fore udder length, central ligament, udder balance, SDFT, SDRT, fore udder attachment, teat direction, Rear legs rearview, Hock development, udder depth, and udder width was generated from the year 2017-2019 at Livestock Research Centre (LRC) of ICAR-National Dairy Research Institute (NDRI) Karnal, Haryana, India. Factors involved in the present study are of practical significance at the field level. Most of the traits were measured using ICAR Guidelines for conformation recording of dairy cattle (Anon, 2018) except central ligament which was measured at the base of the rear udder, udder balance; the level of the rear udder was assessed in relation to the depth of the front udder, shortest distance of floor from fore teat (SDFT), shortest distance of floor from rear teat (SDRT); Mean of distance from pair of the front teat ends to the floor was taken as SDFT and from rear teats was SDRT, fore udder attachment which was in degrees (the soundness of attachment of fore udder to the abdominal wall) and udder width; The udder width was measured as a distance linking two lateral lines of attachment of the udder to the abdominal wall, below the flank.

Incidence of mastitis

Incidence of mastitis among the lactating Karan Fries cows was recorded by considering the animals which were affected by mastitis at least once in their lifetime. Data were recorded from the mastitis register maintained at the Animal Health Complex of ICAR-National Dairy Research Institute, Karnal, and data comprised of records of 20 years since 2000.

Analysis

The data for udder and teat type traits were corrected for the non-genetic factors prior to the start of analysis. The sources of variation were the stage of lactation which was divided into 4 stages *i.e*, 0-90 days, 90-180 days, 180-270 days and >270 days; parity (1- 5 or >5); the season of calving (Summer: April to June, Rainy: July to August, Autumn: September to November, Winter: December to March). Discriminate function analysis to develop a functional model to discriminate mastitis resistant and susceptible animals was performed by usingSPSS version 22.

The dichotomous logistic regression model used for the ROC curve can be expressed as:

$$\ln\left(\frac{P_{ijklmne}}{1-P_{ijklmne}}\right) = \beta_0 + \beta_j Y_j + \beta_k N_k + \beta_l P_l + \beta_m B_m + \beta_n D_n + \beta_o G_o + \beta_p H_p + \beta_q I_q + \beta_r J_r + \beta_s K_s + \beta_t L_t + \beta_u M_u + e\text{-}ijklmnopqe$$

Where,

$P_{ijklmnoe}$ = Probability of odds of disease incidence; Y_j is the effect of j^{th} fore udder length, N_k is the effect of k^{th} central ligament, P_l is the effect of l^{th} udder balance, B_m is the effect of m^{th} udder circumference, D_n is the effect of the n^{th} SDFT, G_o is the o^{th} SDRT. H_p is the effect of p^{th} fore udder attachment group, I_q is the effect of the q^{th} teat direction group, J_r is the effect of the r^{th} is the effect of rear legs rear view, k_s is the s^{th} hock development. L_t is the effect of t^{th} udder depth group, M_u is the effect of the u^{th}udder width group. Corresponding regression coefficients are indicated by β. $e_{ijklmnopqe}$ is the residual error corresponding to the responding variable.

ROC curve

Response operating characteristics (ROC) curves were plotted for the discriminant model. A ROC curve graphically displays sensitivity and 100% minus specificity (false positive rate) at several cut-off points. ROC curve summarizes the confusion matrix at all possible thresholds. ROC curves are drawn using true positive rate on Y- axis and false positive rate on X- axis. It is as a result natural to do a hypothesis test to assess whether the AUC differs significantly from 0.5. Specifically, the null and alternate hypotheses are defined as H0: AUC = 0.5 versus H1: AUC ≠ 0.5. This test statistic given by [AÛC - 0.5/SE(AÛC)] is approximately normally distributed and has favorable statistical properties.

The linear discriminate function model can be expressed as follows;

$$D = a + b_1Z_1 + b_2Z_2 + b_3Z_3 + \ldots\ldots\ldots\ldots + b_{12}M_{12}$$

Where,

i = 1, 2, 3, ….., 12 D-total discriminant score for mastitis resistant and susceptible animals.

In this model, b_1, b_2 ... b_{12} are estimated canonical coefficients for the data and Z_1, Z_2 ... Z_{12} represent independent variables, which are udder and teat conformation traits. Discriminant analysis was adopted to classify animals into statistically different groups as mastitis susceptible and mastitis resistant, based on udder and teat characteristics.

The hypothesis for testing the equality of covariance matrices was depicted as: H0: $\sum_1 = \sum_2$ Vrs H1: At least one pair of sigma's $\sum$ is different.

Box and plots

Box plots are substitute ways of demonstrating the distribution of the discriminant function scores for each group. Box plots are non-parametric. They display variation in samples of a statistical population without making any assumptions of the underlying statistical distribution.

RESULTS AND DISCUSSION

Among all these factors considered, it was found that fore udder length and central ligament contributed significantly ($p<0.01$) to calculate the difference between the mastitis resistant and susceptible animals. Wilk's lambda value, F value and p values are indicated in Table 1. Eigen value for the above factors was calculated as 0.076 and the overall Wilks' lambda value was 0.929. The Chi-square value was 8.83 at 2 degrees of freedom and the model significantly ($p<0.01$) classified 69.1 per cent of original grouped cases correctly. Further, a linear discriminant function model was constructed by taking into consideration the significant factors only. In this case, the log determinants appear similar and Box's M is 0.62 with F = 0.20 which is significant at p<.0001 (Table 2). The unstandardized and standardized canonical discriminant function coefficients for the respective factors along with their F value Wilk's lambda and the p values after reevaluating the significant factors only were represented in Table 3. The Eigenvalue for the later analysis was 0.076 with a canonical correlation value of 0.27 suggests the model explains 7.29% of the variation in the grouping variable. The canonical correlation is the multiple correlations between the predictors and the discriminant function. Unstandardized Canonical Discriminant Coefficients(b) are used to generate the discriminant function (equation). The original observations were correctly classified as 92.1% resistant to mastitis and the remaining 7.9% being misclassified into the mastitis group (Table 4). The discriminant function was able to correctly classify 69.1% of the mastitis and resistance to mastitis in Karan Fries cows. The discriminate function fitted was: D = 0.586+ 0.262 Z1– 0.37 Z2. The function fitted demonstrated 69.1 per cent accuracy with p< 0.001. The functions at group centroids were -0.169 and 0.443 respectively for mastitis resistant and susceptible animals. Centroids are actually the group means of canonical variables. Cases with scores close to a centroid were predicted as belonging to that group. Several different procedures of accuracy have been developed, but the uncomplicated is area under the ROC curve and Box and plots were depicted in Table 5 and Fig 1. In the present study, the area under the ROC curve is 0.68 which is considered acceptable with the standard error of 0.05 in the discriminant analysis model whereas in the regression model it is 0.84.Results indicated that there is 68% chance that the researcher will correctly distinguish animals with mastitis or mastitis resistant based on the udder and teat type traits in discriminant analysis whereas in the regression model chance is more *i.e.* 84%.

Wilk's lambda statistic is the test for univariate equality of group means. Large values of lambda designated by the factors *viz*., SDFT, fore udder attachment, and udder depth

Table 1: Test of equality of group means table in karan fries cows.

Variables	Mean±SD	F	Wilk's Lambda	P value
Fore udder length	3.89±3.08	4.97	0.96	0.028
Central ligament	4.30±1.84	3.43	0.97	0.06
Udder balance	-0.63±5.12	0.34	0.99	0.55
Udder circumference	129.61±18.37	1.42	0.98	0.23
SDFT	48.56±6.58	2.15	0.98	0.14
SDRT	48.55±6.65	0.04	1.00	0.84
Fore udder attachment	123.00±14.12	0.08	0.99	0.77
Teat direction	4.57±3.23	0.41	0.99	0.52
Rear legs rear view	8.41±1.50	2.56	0.97	0.11
Hock development	8.57±1.23	2.91	0.97	0.09
Udder depth	36.63±8.50	0.17	0.99	0.67
Udder width	69.83±10.28	0.90	0.99	0.34

Table 2: Test for equality of covariances matrices for karan fries cows.

Log determinants		Log determinant	Test results	0.621
Mastitis/ Resistant to mastitis	Rank		Approx.	0.202
0	2	3.44	df1	3
1	2	3.32	df2	70643.4
Pooled within groups	2	3.41	Sig.	0.89

Table 3: Represented values for the significant variables considered for discriminating mastitis resistant and susceptible animals.

Variables	Wilk's lambda	Unstandardized canonical discriminant coefficients function	Standardized canonical Discriminant coefficients function	F	P-value	Structure matrix
Fore udder length	0.96	0.262	0.79	4.97	0.028	0.73
Central ligament	0.92	-0.373	-0.68	4.58	0.012	-0.60

indicated that the group means did not appear to be different, while small values indicated that group means appeared to be different. Thus it was found that the fore udder length and central ligament factors played a major role in discriminating Mastitis resistant and susceptible animals. The fore udder should be moderate in length and firmly attached to the body wall, since large fore udder length leads to Decreasing teat-end-to-floor distance serve as risk factors for clinical mastitis and periparturient udder edema (Slettback *et al*., 1995) since, there is a negative genetic correlation between clinical mastitis and teat-tip-to-floor distance (Jensen *et al*., 1985) and associated with higher SCC (Slettbakk *et al*., 1990), leads to an increasing proportion of teat lesions (Grommers *et al*. 1971). The stronger central ligament can lead to the minimization of the potential for injuries and maximizing milking management by keeping the teats in place and udder elevated and can have a longer stay in the herd. Box's M tests the null hypothesis that the covariance matrices do not differ between groups formed by the dependent. This test was found to be nonsignificant in the present study so we fail to reject the null hypothesis of no difference and conclude that, covariance matrices are equal. Wilks' lambda indicates the significance of the discriminant function indicating a highly significant function ($p < 0.01$) and provides the proportion of total variability not explained, *i.e.* it is the converse of the squared canonical correlation. So we have 92.9% unexplained. A highly significant classificatory variable *i.e.*, fore udder length was however introduced in the present model which was not considered by previous workers (Thirunavukkarasu 2003; Jadhav *et al*., 2019). From the discriminant equation, it could be inferred that the variables considered in the present analysis together were able to classify effectively normal and infected animals. In the ROC curve, An area of 0.5 represents the diagonal, which means no discrimination exists *i.e.* ability to diagnose animals with mastitis or mastitis resistant. In ROC curve, Area under the curve in the regression model was slightly more above this diagonal line than the discriminant model, so the regression model considered to have a better discriminating ability to diagnose animals with mastitis or mastitis resistant. An area of 1 represents the perfect indicator and accurate test. The present study relates with

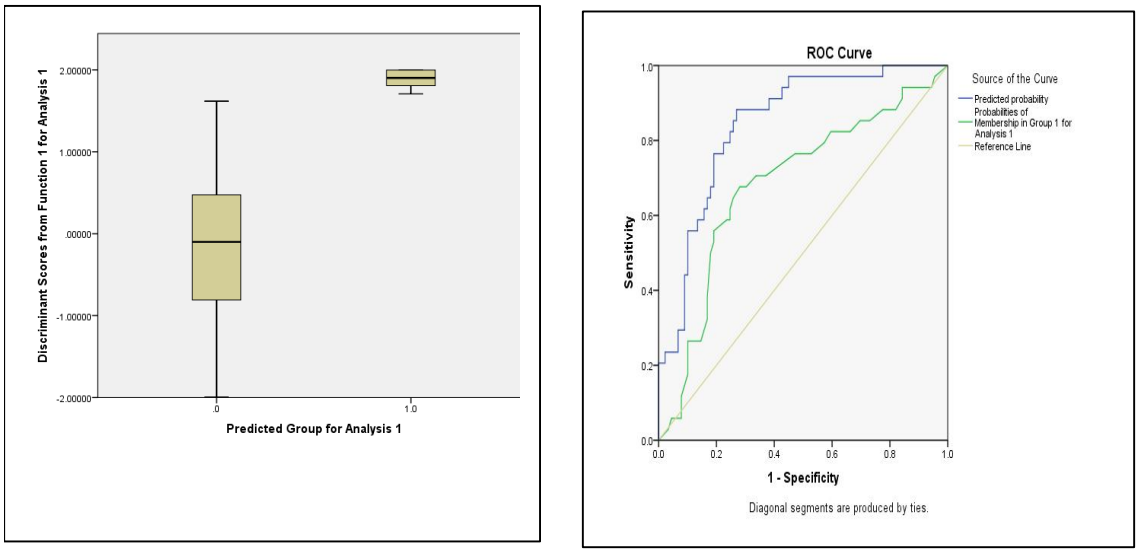
Table 4: Classification results for the discriminant analysis of mastitic and resistant to mastitis Karan Fries cows.

True	Mastitis type	Predicted group membership		
		0	**1**	**Total**
Original	0	82	7	89
	1	31	3	34
%	0	92.1	7.9	100
	1	91.2	8.8	100

69.1% of originally grouped cases correctly classified.

Table 5: Area under the curve in Karan Fries cows.

Model	Area	Std error[a]	Asymptotic Sig	95% confidence interval	
				Lower bound	Upper bound
Regression model	0.84	0.037	0.000	0.773	0.917
Discriminant model	0.68	0.056	0.002	0.575	0.793



a. Box plot

b. ROC curve

Fig 1a, b: Box plots depicting the distribution of discriminant scores and ROC curve for Regression and discriminant analysis model in Karan Fries cows.

Montgomery *et al.* (1986) who prefers the logistic regression model for the prediction of coliform mastitis as compared to logistic regression. However, this was a pilot study and we meticulously recorded the data so as to avoid any biasedness. With this data, we could see definite trends and hence we aimed at such analysis. In the future, with larger dataset, the study will be repeated.

CONCLUSION

The most tedious work for a livestock holder is the exact differentiation of the animal as mastitis resistant/ mastitis susceptible. A discriminate function model can be successfully used, including field applicable factors such as fore udder length and central ligament to determine the exact status of the animal. The function fitted in the present study demonstrated 69.1 per cent accuracy with $p< 0.001$ and the functions at group centroids were -0.169 and 0.443 respectively for mastitis resistant/ susceptible animals. To increase the economic value of animals, we need to increase the profitability of farm by focusing on durability (udder and teat type traits), production and reproduction traits by creating a selection index. Using type traits in combination with production traits can help to achieve increased longevity, persistency, and herd life of animals. It is also possible to select cows for their strength, mastitis resistance, production, and reproductive efficiency and look for morphometric markers in type traits that can define the overall genetic merit of the cow.

ACKNOWLEDGEMENT

We are grateful to Director ICAR-NDRI, Karnal (Haryana) for providing the facilities for the above work during the study.

REFERENCES

Anonymous (2018). ICAR Guidelines for conformation Recording of Dairy cattle, Beef cattle, Dual Purpose Cattle and Dairy Goats. https://www.icar.org

Grommers, F.J., Van de Braak, A.E., Antonisse, H.W. (1971). Direct trauma of the mammary glands in dairy cattle I. Variations in incidence due to animal variables. British Veterinary Journal. 127: 271-82.

Jadhav, P.V., Tarate, S.B., Das, D.N., Suresh, K.P. (2019). Screening of animals for subclinical mastitis: A discriminate function analysis. Indian Journal of Animal Research. 53.

Jensen, N.E., Madsen, P., Larsen, B., Klastrup, O., Nielsen, S.M., Madsen, P.S. (1985). Heritability of and markers of resistance against mastitis in the Danish RDM breed. Kieler Milchwirtschaftliche Forschungsberichte. 7: 506-10.

Lakshmi, R. (2016). A review on Mastitis. International Journal of Medicine Research. 1: 118-123.

Leotta, R. (2004).Use of linear discriminant analysis to characterize three dairy cattle breeds on the basis of several milk characteristics. Italian Journal of Animal Science. 3: 377-83

Montgomery, M.E., White, M.E., Martin, S.W. (1987). A comparison of discriminant analysis and logistic regression for the prediction of coliform mastitis in dairy cows. Canadian Journal of Veterinary Research. 51: 495.

Slettbakk, T., Jørstad, A., Farver, T.B., Hird, D.W. (1990). Impact of milking characteristics and teat morphology on somatic cell counts in first-lactation Norwegian cattle. Preventive Veterinary Medicine. 8: 253-67.

Slettbakk, T., Jørstad, A., Farver, T.B., Holmes, J.C. (1995). Impact of milking characteristics and morphology of udder and teats on clinical mastitis in first-and second-lactation Norwegian cattle. Preventive Veterinary Medicine. 24: 235-44.

Thirunavukkarasu, M. (2003). A model to discriminate mastitic and non-mastitic milch animals. Indian Veterinary Journal. 80: 890-894.