Livestock Science 153 (2013) 20–26

Contents lists available at SciVerse ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Comparison of Poisson, probit and linear models for genetic analysis of number of inseminations to conception and success at first insemination in Iranian Holstein cows

R. Abdollahi-Arpanahi [a,*], F. Peñagaricano [b], H. Aliloo [a], H. Ghiasi [c], J.I. Urioste [d]

[a] Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, 31587-77871 Karaj, Iran
[b] Department of Animal Sciences, University of Wisconsin, Madison, WI, USA
[c] Department of Animal Science, Faculty of Agriculture, Payme Noor University, Iran
[d] Departamento de Producción Animal y Pasturas, Facultad de Agronomía, Universidad de la Republica, Montevideo, Uruguay

ARTICLE INFO

Article history:
Received 18 May 2012
Received in revised form
21 December 2012
Accepted 18 January 2013

Keywords:
Dairy cows
Genetic parameters
Model comparison
Nonlinear models
Fertility traits

ABSTRACT

The goals of this study were to estimate genetic parameters and to assess alternative models for genetic evaluation of number of inseminations to conception (INS) and success at first insemination (SF) in Iranian Holstein cows. Two models were considered for each trait: linear and probit models for SF, and linear and Poisson models for INS. Data consisted of 72,124 records of parities 1 to 6 from 27,113 cows having lactation between 1981 and 2007 and distributed over 15 large Holstein herds. Genetic parameters and goodness of fit statistics were estimated using the whole data set and predictive ability of models was assessed via a 4-fold cross-validation based on mean squared error of prediction (MSEP) and correlation between observed and fitted values. Estimates of heritability ranged from 0.039 to 0.062 for SF and 0.040 to 0.165 for INS. The performance of linear and probit models was very similar for SF. Predictions of random effects from these models were highly correlated, and both models exhibited similar predictive ability. For INS, the linear model performed better than the Poisson model according to goodness of fit statistics, but these two models showed the same predictive ability. Overall, nonlinear models did not outperform linear models for genetic evaluations of SF and INS in Iranian Holstein cows.

1. Introduction

Reproductive performance is economically important in livestock. Economic losses due to poor fertility can be attributed to the cost of prolonged calving interval, increased insemination costs, reduced returns from calves born and consequently increasing replacement costs (Chang et al., 2006). By including some fertility traits such as number of inseminations to conception and success at first insemination in the selection index, dairy producers can reduce the cost of production and as a result increase the profit. Furthermore, reliable estimations of genetic parameters are needed for defining a breeding strategy for these traits.

According to Gonzalez-Recio and Alenda (2005), proper female fertility performance could be defined as showing timely manner heat and becoming pregnant with a low number of inseminations. For instance, reducing the number of inseminations to conception and increasing the rate of success at first insemination will lead to increased fertility, and pregnant cows when dairy producers desire, with reduced costs of used semen, hormonal treatments and veterinary fees. Therefore, these

* Corresponding author. Tel.: +98 2632248082; fax: +98 2612246752.
E-mail address: r.abdollahi@ut.ac.ir (R. Abdollahi-Arpanahi).

1871-1413/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.livsci.2013.01.009

fertility traits should be considered in genetic evaluations and included in the breeding goals (Gonzalez-Recio and Alenda, 2005).

Number of services to conception (INS) as an indicator of reproductive efficiency has been defined as the number of services required for a successful conception. The INS is directly related to open days, pregnancy rate, non-return rate and the reciprocal of services per conception. In contrast to open days and days to first service which are interval traits that do not account for elapsed time from first to last insemination and veterinary costs, the variable INS takes partly these important issues into consideration. In this sense, Gonzalez-Recio and Alenda (2005) reported that by including INS in an index with combination of production traits, the profit would be increased by 229 €/cow. Therefore, INS should be registered and considered in genetic evaluations since it would maximize profitability at balancing economically production and fertility.

Success at first insemination (SF) is a trait that is similar to pregnancy rate (Bormann, 2006) and which is often included in dairy female selection indices. SF is an economically important trait due to the cost of the semen, labors involved in heat checking and breeding for multiple AI, and the difference in the quality and value between AI calves and natural calves. Moreover, heifers that became pregnant at first breeding are going to calve earlier, and hence they will have a better chance to breed back on the following year. Heritability estimates for first-service conception range from 0.03 to 0.22 (Cammack et al., 2009).

Heritability estimates for fertility traits are typically low, ranging from 1% to 5% when using linear models for statistical analysis (e.g. Wall et al., 2003). Nonlinear models could be better alternatives for more accurate assessment of genetic parameters and genetic merit of the animals in these traits. Several probability models can be used for describing the number of inseminations to conception (INS) and success at first insemination (SF) in dairy cattle. For categorical or counted variables such as INS, several alternatives are available, e.g., Poisson, linear or ordinal threshold models (Peñagaricano et al., 2011; Vazquez et al., 2009a, 2009b). In general, analysis of counts with models based on the Poisson distribution tends to be more appropriate than models based on normal distribution. For binary outcomes such as SF, the Bernoulli distribution parameterization using either a probit or a logit link is a natural approach. Another alternative, despite the binary nature of the variable, is the use of a linear model.

The objectives of the current study were, firstly, to estimate genetic parameters; and, secondly, to assess alternative models for genetic analysis of SF (probit and linear) and INS (Poisson and linear). Models were compared based on goodness of fit statistics and their predictive ability in a 4-fold cross-validation analysis.

2. Materials and methods

2.1. Data

A total of 72,124 records of parities 1 to 6 from 27,113 cows collected from 1981 to 2007 in 15 large Holstein herds of Iran were used in the analysis. Only artificial insemination mating records were used. Response variables studied were number of inseminations to conception (INS) and success to first insemination (SF). These 15 dairy herds are under registration of the Dairy Herd Improvement Program of the Animal Breeding Centre of Iran. Artificial insemination (AI) technicians record all insemination events and are in charge of maintaining an accurate dataset. All cows had records for both traits. Reproductive traits in different parities were treated as repeated measurements. For INS, all the values greater than 9 were set to 9 in order to reduce possible recording errors. SF was a binary trait defined as 1 if the cow became pregnant at first insemination and 0 otherwise.

2.2. Statistical Models

Preliminary least squares analyses were conducted for both traits using the GLM procedure of Statistical Analysis System (SAS) (SAS, 2004) to decide which non-genetic effects had to be included in the final models. The following linear predictor was common to both traits:

$$\boldsymbol{\eta} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{hys} + \mathbf{Wpe}$$

where $\boldsymbol{\eta}$ is a function of the expected value of SF or INS; **b** is a vector for fixed effects of parity (6 levels), age at previous calving (10 levels) and months of first insemination (12 levels); **a** is the additive genetic effect; **hys** is herd-year-season effect; **pe** is permanent environmental effect for cows; **e** is the residual term; and $\mathbf{X}$, $\mathbf{Z_1}$, $\mathbf{Z_2}$ and $\mathbf{W}$ are incidence matrices relating data to the corresponding effects.

Random effects were assumed to follow the multivariate normal distribution,

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{hys} \\ \mathbf{pe} \end{pmatrix} \sim N\left[0, \begin{pmatrix} \mathbf{A}\sigma_a^2 & 0 & 0 \\ 0 & \mathbf{I}\sigma_{hys}^2 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{pe}^2 \end{pmatrix}\right]$$

where **a**, **hys** and **pe** are the vectors of additive genetic animal, herd-year-season and permanent environmental effects, respectively; σ_a^2, σ_{hys}^2 and σ_{pe}^2 are the additive genetic, herd-year-season and permanent environmental effects variances, respectively; **A** is the additive relationship matrix (with order $32{,}447 \times 32{,}447$) and **I** is the identity matrix of order 838 for herd-year-season and 27,113 for permanent environmental effects. Animal, herd-year-season and permanent environmental effects were assumed to be independent of residual effects in the models. General structure of the data and the pedigree is presented in Table 1.

2.2.1. Probit model for SF

The Probit model (Gianola, 1982) describes the observable outcome (SF) using an underlying linear model $\mathbf{z} = \boldsymbol{\eta} + \mathbf{e}$, where **e** is independent and identically distributed standard normal random vector. The scoring rule in this model is:

$$\mathbf{SF} = \begin{cases} 1 \text{ (success)} \\ 0 \text{ (failure)} \end{cases}$$