J. Dairy Sci. 99:1–16
http://dx.doi.org/10.3168/jds.2016-10935

A prognostic model to predict the success of artificial insemination in dairy cows based on readily available data

C. J. Rutten,*[1] W. Steeneveld,*† J. C. M. Vernooij,* K. Huijps,‡[2] M. Nielen,* and H. Hogeveen*†
*Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, 3584 CL, Utrecht, the Netherlands
†Business Economics Group, Wageningen University, 6706 KN, Wageningen, the Netherlands
‡CRV, 6800 AL Arnhem, the Netherlands

ABSTRACT

A prognosis of the likelihood of insemination success is valuable information for the decision to start inseminating a cow. This decision is important for the reproduction management of dairy farms. The aim of this study was to develop a prognostic model for the likelihood of successful first insemination. The parameters considered for the model are readily available on farm at the time a farmer makes breeding decisions. In the first step, variables are selected for the prognostic model that have prognostic value for the likelihood of a successful first insemination. In the second step, farm effects on the likelihood of a successful insemination are quantified and the prognostic model is cross-validated. Logistic regression with a random effect for farm was used to develop the prognostic model. Insemination and test-day milk production data from 2,000 commercial Dutch dairy farms were obtained, and 190,541 first inseminations from this data set were used for model selection. The following variables were used in the selection process: parity, days in milk, days to peak production, production level relative to herd mates, milk yield, breed of the cow, insemination season and calving season, log of the ratio of fat to protein content, and body condition score at insemination. Variables were selected in a forward selection and backward elimination, based on the Akaike information criterion. The variables that contributed most to the model were random farm effect, relative production factor, and milk yield at insemination. The parameters were estimated in a bootstrap analysis and a cross-validation was conducted within this bootstrap analysis. The parameter estimates for body condition score at insemination varied most, indicating that this effect varied most among Dutch dairy farms. The cross-validation showed that the prognosis of insemination success closely resembled the mean insemination success observed in the data set. Insemination success depends on physiological conditions of the cow, which are approximated indirectly by production and reproduction data that are routinely recorded on the farm. The model cannot be used as a detection model to distinguish cows that conceive from cows that do not. The model validation indicates, however, that routinely collected farm data and test-day milk yield records have value for the prognosis of insemination success in dairy cows.
Key words: dairy, prognostic model, insemination success, reproduction

INTRODUCTION

Reproduction management has a large influence on dairy farm profitability (Inchaisri et al., 2010b; Giordano et al., 2011). Economically important factors include the decision to start inseminating, the voluntary waiting period (**VWP**; Inchaisri et al., 2011b), and the decision to stop inseminating a dairy cow (Inchaisri et al., 2012). The goal underlying the choice of VWP is to realize a defined calving interval and milk yield at dry off, which requires a cow to conceive before a certain DIM. The optimal VWP within a lactation depends on factors such as daily milk yield, production level, estrus detection rate, and the cow's health status (Inchaisri et al., 2011b). A prognosis of the likelihood of insemination success is valuable information for the decision to start with insemination, and can therefore improve reproduction management.

Although one study focused on prognosis of the likelihood of insemination success (Shahinfar et al., 2014), most studies have focused on identifying risk factors associated with insemination success (Caraviello et al., 2006; Gábor et al., 2008; Inchaisri et al., 2010a, 2011a; Lane et al., 2013; Löf et al., 2014). These studies have identified the following risk factors: milk yield, parity, cow breed, insemination number, lactation stage, calving season, insemination season, use of sexed semen,

Received January 22, 2016.
Accepted April 14, 2016.
[1]Corresponding author: c.j.rutten@uu.nl
[2]Current address: kirstenvanhees.nl, 5451 NR Mill, the Netherlands.

and information on the course of the previous calving (e.g., whether dystocia occurred) for insemination success.

Other factors thought to influence insemination success in dairy cows include diseases (Fourichon et al., 2000; Canu et al., 2010), BCS, ketosis, SCC, and breed of the cow (Caraviello et al., 2006; Lane et al., 2013; Löf et al., 2014). An important factor is BCS, which can be seen as a proxy of the dairy cow's body reserves (Löf et al., 2014). In particular, a postpartum decrease in BCS, indicating mobilization of body reserves, is associated with lower reproductive performance (Roche et al., 2009; Hernandez et al., 2012). In case of extreme mobilization of body reserves, ketosis may occur, which is also associated with lower reproductive performance (Roche et al., 2009; Schefers et al., 2010). Ketosis could be included in an analysis of the likelihood of insemination success by using a ketosis score from milk recordings, which is based on ketone bodies in milk (de Roos et al., 2007), or on the ratio of fat to protein content. Many variables with potential prognostic value for the likelihood of successful insemination, such as fat and protein contents, SCC, lactose content, and ketosis score, have been evaluated as risk factors for insemination success.

Several studies have shown the importance of farm factors for the reproductive performance of dairy herds (Schefers et al., 2010; Shahinfar et al., 2014). For insemination success in individual cows, herd conception rate in the past 3 mo has been suggested as an important predictor (Shahinfar et al., 2014). Other farm factors found to be influential are VWP, estrus detection accuracy, stocking density, udder health level of the herd, and use of a resynchronization program (Schefers et al., 2010). It is important, therefore, that models which give a prognosis of insemination success account for farm effects and quantify these effects.

It is logical to assume that risk factors having a strong association with insemination success are useful for making a prognosis. However, the strength and causality of association give no indication of the prognostic power of a parameter. To implement a prognosis of the likelihood of insemination success in a decision support system, it is important to understand the accuracy of the prognosis. This requires knowledge of the robustness of the estimated parameters in a prognostic model, given the variation among farms and individual cows. Furthermore, it is important to conduct an independent validation of the prognosis. Shahinfar et al. (2014) conducted a validation for machine learning algorithms and reported an area under the receiver operating characteristic curve (AUC) in the range of 0.6 to 0.75.

The aim of this study was to develop a prognostic model for the likelihood of successful first insemination. In the first step, variables were selected for the prognostic model that had prognostic value for the likelihood of a successful first insemination. The selected variables are readily available on-farm when a farmer makes breeding decisions and are obtained from data that is currently routinely recorded on Dutch dairy farms. In the second step, the farm effects on the likelihood of a successful insemination were quantified and the prognostic model was cross-validated.

MATERIALS AND METHODS

Available Data

Data on inseminations, calving dates, parity and age of the cow, pregnancy checks, lactation length, relative production factor (a factor that represents the production level of an individual cow relative to her herd mates), and test-day milk yields from 2,000 randomly selected dairy farms, covering the period 2009 to 2012, were obtained from the database of the Dutch cattle breeding and milk production recording company (CRV, Arnhem, the Netherlands). In 2009 and 2012, 20,268 and 18,682 farms participated in the milk production recording program of CRV, representing 81 and 85% of the total number of farms, respectively. Selected farms were required to have a farm size of at least 50 cows and to make use of the CRV AI service. Records were removed from the data set when the data were considered biologically implausible. For example, one test-day was removed as the lactose percentage was 15%.

The data set contained 1,403,194 recorded AI and 613,668 individual lactations. Only records containing inseminations with a known preceding calving date (so the calving date that started the lactation in which the cow was inseminated at a certain DIM) were selected; 653,694 recorded AI were removed from the initial data set because no preceding calving date was registered, leaving 749,500 records. The focus in this study was on the first insemination of a lactation. A farmer will choose to start breeding a cow after a certain DIM. If insemination is unsuccessful, the farmer decides either to inseminate the cow again or to stop inseminating and cull, rather than choosing to postpone a reinsemination. A prognosis of the likelihood of insemination success is therefore most relevant for the first insemination of a lactation.

Definition of Successful and Unsuccessful First Inseminations

Inseminations recorded in the database were considered unsuccessful when, within the current lactation,