J. Dairy Sci. 101:1–10
https://doi.org/10.3168/jds.2017-14076

Predicting blood β-hydroxybutyrate using milk Fourier transform infrared spectrum, milk composition, and producer-reported variables with multiple linear regression, partial least squares regression, and artificial neural network

R. S. Pralle, K. W. Weigel, and H. M. White[1]
Department of Dairy Science, University of Wisconsin, Madison 53706

ABSTRACT

Prediction of postpartum hyperketonemia (HYK) using Fourier transform infrared (FTIR) spectrometry analysis could be a practical diagnostic option for farms because these data are now available from routine milk analysis during Dairy Herd Improvement testing. The objectives of this study were to (1) develop and evaluate blood β-hydroxybutyrate (BHB) prediction models using multivariate linear regression (MLR), partial least squares regression (PLS), and artificial neural network (ANN) methods and (2) evaluate whether milk FTIR spectrum (mFTIR)–based models are improved with the inclusion of test-day variables (mTest; milk composition and producer-reported data). Paired blood and milk samples were collected from multiparous cows 5 to 18 d postpartum at 3 Wisconsin farms (3,629 observations from 1,013 cows). Blood BHB concentration was determined by a Precision Xtra meter (Abbot Diabetes Care, Alameda, CA), and milk samples were analyzed by a privately owned laboratory (AgSource, Menomonie, WI) for components and FTIR spectrum absorbance. Producer-recorded variables were extracted from farm management software. A blood BHB ≥1.2 mmol/L was considered HYK. The data set was divided into a training set (n = 3,020) and an external testing set (n = 609). Model fitting was implemented with JMP 12 (SAS Institute, Cary, NC). A 5-fold cross-validation was performed on the training data set for the MLR, PLS, and ANN prediction methods, with square root of blood BHB as the dependent variable. Each method was fitted using 3 combinations of variables: mFTIR, mTest, or mTest + mFTIR variables. Models were evaluated based on coefficient of determination, root mean squared error, and area under the receiver operating characteristic curve. Four models (PLS–mTest + mFTIR, ANN–mFTIR, ANN–mTest, and ANN–mTest + mFTIR) were chosen for further evaluation in the testing set after fitting to the full training set. In the cross-validation analysis, model fit was greatest for ANN, followed by PLS and MLR. Diagnostic strength after cross-validation was poorest for MLR and was similar for ANN and PLS. Models that used mTest + mFTIR variables performed marginally better than models that used only mFTIR or mTest variables. These results suggest that blood BHB prediction models that use mFTIR + mTest variables may be useful additions to existing HYK diagnostic and management programs.

Key words: ketosis, transition cow, diagnostic tools, neural network

INTRODUCTION

Hyperketonemia (**HYK**) is a metabolic disorder that impairs milk production, reproduction, and health outcomes in lactating dairy cows (Duffield, 2000; McArt et al., 2012). The postpartum prevalence of HYK ranges worldwide from 15 to 22%, although it can vary greatly between farms (Suthar et al., 2013; Santschi et al., 2016; Chandler et al., 2018). Measurement of blood BHB concentration early postpartum is the predominant diagnostic method, and HYK is commonly defined as blood BHB ≥1.2 mmol/L (Iwersen et al., 2009; McArt et al., 2012; Gordon et al., 2017). Milk ketone body concentrations have been considered for HYK diagnosis because they are correlated with blood concentrations (Marstorp et al., 1983; Andersson, 1984; Enjalbert et al., 2001). The most accurate methods for detecting milk ketone bodies are flow-injection analysis, GLC, and enzymatic assays; however, these techniques are expensive, time consuming, and difficult to automate. Fourier transform infrared (**FTIR**) spectrometry could provide a practical alternative for predicting concentrations of milk ketone bodies because it is already

Received November 1, 2017.
Accepted January 8, 2018.
[1]Corresponding author: heather.white@wisc.edu

used to evaluate milk composition in DHI milk testing programs (Rutten et al., 2009, 2011).

Studies in which milk FTIR data have been used to predict milk ketone body concentrations have reported moderate correlations between predicted and assayed milk concentrations (Hansen, 1999; Heuer et al., 2001; de Roos et al., 2007). However, correlation of FTIR-based milk ketone body concentration predictions with blood BHB concentration is low (Chandler et al., 2018), and use of FTIR-predicted milk ketone bodies for HYK diagnosis has produced tests with insufficient sensitivity to diagnose individual cows (van der Drift et al., 2012; Chandler et al., 2018). Marginal improvement of FTIR-predicted milk ketone bodies used as an HYK diagnostic was made by including FTIR-predicted milk ketone bodies with cow test-day information in logistic (van der Drift et al., 2012) and multiple linear regression (**MLR**; Chandler et al., 2018) models predicting blood BHB concentration. At this time, these predictions of blood BHB concentration are recommended as a tool for monitoring HYK prevalence at the herd level, but sensitivity and specificity are insufficient for use as an individual cow diagnostic tool (van der Drift et al., 2012; Chandler et al., 2018).

Inclusion of milk FTIR spectrum data, combined with FTIR-predicted milk ketone body concentrations and milk composition in models, could generate more robust models with enhanced prediction of blood BHB and HYK diagnosis by capturing residual spectrum predictive abilities. Furthermore, advancements in computing technology have made flexible and computationally demanding methods, such as machine learning, increasingly available for such applications. Artificial neural networks (**ANN**) are a type of machine-learning prediction method with the ability to self-learn relationships from labeled experimental data and generalize to unlabeled situations. This advanced modeling technique may allow optimization of blood BHB concentration prediction from milk FTIR spectrum, milk component, and producer-reported variables.

We hypothesized that the use of more advanced modeling techniques, such as partial least squares regression (**PLS**) and ANN, would result in improved predictions of HYK status from producer-reported and milk composition variables compared with MLR. Furthermore, we hypothesized that predictions of blood BHB could be optimized by inclusion of milk FTIR spectrum, milk component, and producer-reported variables. Based on these hypotheses, our objectives were to (1) develop and evaluate blood BHB prediction models using MLR, PLS, and ANN methods and (2) evaluate whether milk FTIR spectrum–based models are improved by the inclusion of milk composition and producer-reported variables.

MATERIALS AND METHODS

Cows from 2 privately owned Holstein dairy farms in southern Wisconsin and the University of Wisconsin–Madison Emmons Blaine Dairy Cattle Research Center were enrolled in the study. Dairy farms were selected based on the following criteria: early-lactation cows grouped in a separate pen, availability of headlocks for blood sampling, capability of proportional milk sampling, use of management software, and willingness to participate. All animal use and handling protocols were approved by the University of Wisconsin–Madison College of Agricultural and Life Sciences Animal Care and Use Committee. Farms were visited twice a week for regular sampling of paired blood and milk samples from multiparous Holstein cows (n = 1,013) between 5 and 18 DIM. This provided an opportunity to collect up to 4 paired blood and milk samples per cow within the 2-wk sampling period. Routine HYK testing within the first 10 DIM at the University of Wisconsin–Madison farm allowed some cows to contribute 5 samples as described by Rathbun et al. (2017).

Sample Collection and Analysis

Morning milk samples were collected by an International Committee on Animal Recordings–approved sampling system using a proportional sampler that had been calibrated within the previous 12 mo. Animal identification numbers were recorded by automatic radio-frequency identification collection and verified by visual identification of animal identification tags to prevent inaccurate identification recording. Samples were preserved with 2-bromo-2-nitropropane-1,3-diol (Advanced Instruments Inc., Norwood, MA) and transported to AgSource Cooperative Services (Menomonie, WI) for analysis according to standard test-day procedures. In brief, all milk samples were preheated to 40°C and mixed before analysis of milk fat and milk protein by FTIR using the Foss MilkoScan FT+ (Foss Analytical, Hillerød, Denmark) in accordance with the instrument manufacturer's instructions and ISO 9622/IDF 141:2013 (AOAC official method 972.16; AOAC International, 2016). Analysis of SCC was performed using Fossomatic FC (Foss Analytical). Milk BHB and milk acetone concentrations were predicted by FTIR using Foss Ketolab (Foss Analytical) based on the calibrations of de Roos et al. (2007). Additionally, the Foss MilkoScan FT+ analysis of milk samples provided predictions for the proportions of SFA, UFA, *trans* fatty acids (**FA**), and short-, medium-, and long-chain FA in milk based on Foss FTIR FA prediction models (Foss Analytical). Per the DHI's standard operating procedures, milk samples were analyzed on equipment that