An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle

T. K. Belay,*[1] B. S. Dagnachew,* Z. M. Kowalski,† and T. Ådnøy*
*Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway
†Department of Animal Nutrition and Dietetics, University of Agriculture in Krakow, Krakow 30-059, Al. Mickiewicza 24/28, Poland

ABSTRACT

Fourier transform mid-infrared (FT-MIR) spectra of milk are commonly used for phenotyping of traits of interest through links developed between the traits and milk FT-MIR spectra. Predicted traits are then used in genetic analysis for ultimate phenotypic prediction using a single-trait mixed model that account for cows' circumstances at a given test day. Here, this approach is referred to as indirect prediction (IP). Alternatively, FT-MIR spectral variable can be kept multivariate in the form of factor scores in REML and BLUP analyses. These BLUP predictions, including phenotype (predicted factor scores), were converted to single-trait through calibration outputs; this method is referred to as direct prediction (DP). The main aim of this study was to verify whether mixed modeling of milk spectra in the form of factors scores (DP) gives better prediction of blood β-hydroxybutyrate (BHB) than the univariate approach (IP). Models to predict blood BHB from milk spectra were also developed. Two data sets that contained milk FT-MIR spectra and other information on Polish dairy cattle were used in this study. Data set 1 (n = 826) also contained BHB measured in blood samples, whereas data set 2 (n = 158,028) did not contain measured blood values. Part of data set 1 was used to calibrate a prediction model (n = 496) and the remaining part of data set 1 (n = 330) was used to validate the calibration models, as well as to evaluate the DP and IP approaches. Dimensions of FT-MIR spectra in data set 2 were reduced either into 5 or 10 factor scores (DP) or into a single trait (IP) with calibration outputs. The REML estimates for these factor scores were found using WOMBAT. The BLUP values and predicted BHB for observations in the validation set were computed using the REML estimates. Blood BHB predicted from milk FT-MIR spectra by both approaches were regressed on reference blood BHB that had not been used in the model development. Coefficients of determination in cross-validation for untransformed blood BHB were from 0.21 to 0.32, whereas that for the log-transformed BHB were from 0.31 to 0.38. The corresponding estimates in validation were from 0.29 to 0.37 and 0.21 to 0.43, respectively, for untransformed and logarithmic BHB. Contrary to expectation, slightly better predictions of BHB were found when univariate variance structure was used (IP) than when multivariate covariance structures were used (DP). Conclusive remarks on the importance of keeping spectral data in multivariate form for prediction of phenotypes may be found in data sets where the trait of interest has strong relationships with spectral variables.
Key words: direct prediction, indirect prediction, β-hydroxybutyrate, milk spectra, dairy cattle

INTRODUCTION

Subclinical ketosis (**SCK**) is an economically important metabolic disorder in early-lactation dairy cows. It is associated with reduced milk production (Duffield et al., 2009), reduced reproductive performance (Walsh et al., 2007), and increased risk of displaced abomasum (LeBlanc et al., 2005; Duffield et al., 2009) and clinical ketosis (Seifi et al., 2011). The disorder is closely related to the negative energy balance occurring in early lactation. Prevalence of SCK can vary between farms; reported prevalence rates range from 8.9 to 43% in the first 2 mo of lactation (McArt et al., 2012; van der Drift et al., 2012; Suthar et al., 2013).

Clinical and SCK are characterized by increased concentrations of ketone bodies (BHB, acetoacetate, and acetone) in milk and blood (Andersson, 1988). Blood BHB concentration has been used as a gold standard for detection of SCK and several studies have used a threshold of 1.2 (Duffield et al., 1997; van der Drift et al., 2012) to 1.4 mmol/L (Geishauser et al., 2000; Oetzel, 2004; Denis-Robichaud et al., 2014) to discriminate between cows with and without SCK. However,

the gold standard method does not allow for routine testing of all animals at risk at regular intervals due to some practical limitations, such as difficulty in blood sampling, especially for farmers or testing many blood samples at a time. Determination of ketone bodies in milk could make the sampling easier (Enjalbert et al., 2001; de Roos et al., 2007). Milk sampling is performed monthly in the milk recording procedures. More routinely, measurements of milk BHB can be done by Fourier transform mid-infrared (**FT-MIR**) spectroscopy analysis in milk samples at test days (de Roos et al., 2007; van der Drift et al., 2012). Blood or milk BHB predicted from milk spectra could be used for detection of SCK in farm management for dairy cows.

The FT-MIR spectra acquisition of the milk sample is multivariate (e.g., 1,060 variables per sample). Hundreds of these spectral variables are used for phenotyping of traits of interest (e.g., BHB) through links developed between the traits and milk spectra. The predicted phenotypes are then together with pedigree information and variance component estimates used in BLUP to calculate individual breeding values (EBV) and other random components included into the model; this is the conventional method used today for genetic evaluation of animals. Dagnachew et al. (2013b) referred to such an approach as indirect prediction (**IP**) and also proposed an alternative approach called direct prediction (**DP**), where genetic analyses are directly applied on the milk FT-MIR spectral variables or on factor scores (latent traits). The BLUP predictions (EBV, herd test day, permanent environment, and residual) for the traits of interest are predicted as correlated traits to the corresponding random components of spectra. Milk FT-MIR spectral variables exhibit strong correlations among each other (Soyeurt et al., 2010; Dagnachew et al., 2013a), and a direct genetic analyses on such correlated spectral variables may result in better accuracy of genetic evaluations (Dagnachew et al., 2013b). In our study, the 2 approaches, IP and DP, were used to predict phenotypes using BLUP predictions of the random and fixed effects part of the models.

The 2 approaches (IP and DP) have been used to predict EBV for milk fat, protein, and lactose contents using Norwegian dairy goat data (Dagnachew et al., 2013b). Those authors showed that accuracy of EBV were improved by 3 to 5% when DP was used compared with the IP approach; they also reported high rank correlation coefficients (0.93 to 0.96) between EBV predicted using IP and DP. However, independent chemical analyses (reference values) for the milk content were not available in that study (i.e., the study relied on phenotypes predicted based on the same spectra for both model calibration and evaluation). Recently, Bonfatti et al. (2017) compared the IP and DP approaches

to estimate EBV for several traits related to fine composition and technological properties of milk and reported rank correlations ranging from 0.07 to 0.96, but <0.5 for most traits. In the present study, we adopted the IP and DP approaches to predict phenotype (not EBV) for BHB having an independent reference value for this trait. Our hypothesis was that keeping spectra multivariate in the form of factor scores or latent traits throughout REML and BLUP analyses instead of converting the spectra into single-trait before the genetic analyses should keep more information, and possibly also give a better prediction of the derived single-trait BHB after multiple-trait mixed modeling accounting for the cows' circumstances. We hypothesized that with multivariate information, one variable may carry information about another variable and thus improve the predictions.

The main objective of our study was to verify whether multivariate mixed modeling of milk FT-MIR spectra that are in the form factor scores (DP) gives better prediction of blood BHB than the univariate (IP) approach, where traits are first predicted from the spectra and then the predicted traits used in genetic analysis for ultimate phenotypic prediction. To do so, the current study developed prediction models for blood BHB from milk spectra and blood BHB measured by reference method.

MATERIALS AND METHODS

Data

Two data sets (referred to as data set 1 and data set 2) were used in our study, made available by the Polish Federation of Cattle Breeders and Dairy Farmers, which provides the monthly milk recording of cows in Poland. Both data sets contained FT-MIR spectra of individual milk samples, pedigree information, milk yield, and other cow and farm information. The milk samples had been analyzed by the MilkoScan FT6000 instrument (Foss Analytical A/S, Hillerød, Denmark). Major milk components, such as protein, fat, lactose, fat composition (both group and individual fatty acids), and ketone bodies (acetone and BHB), had been predicted using the Foss calibration and were available in the data sets.

Data Set 1. After merging the measured blood BHB and phenotypes predicted from the spectra with their corresponding spectral data, data set 1 consisted of data on 832 Polish Holstein Friesian cows (1,914 observations; i.e., at least 2 records per cow) that had been milked 2 or 3 times per day. The spectra and other phenotypes that were predicted from the spectra were recorded for each milking, whereas blood BHB
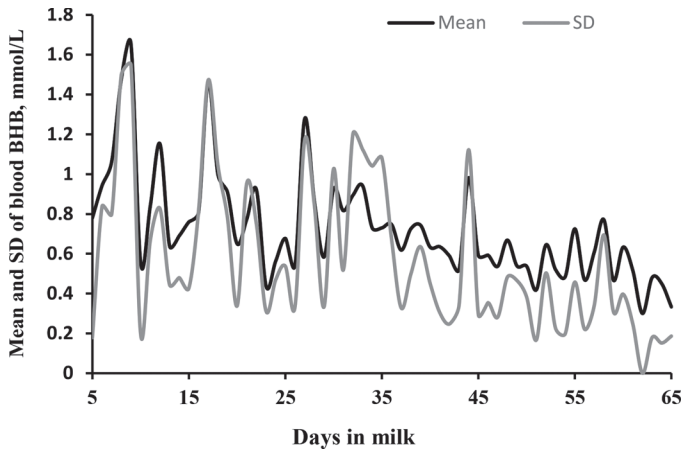
Figure 1. Mean and SD of blood BHB (mmol/L) by DIM from 826 lactating Polish dairy cows.

was measured only once using the glucometer Optium Xido (Abbott, Winey, UK) on test day at 1000 to 1400 h. For better correlation between blood BHB and milk spectra, milk and blood samples from the same milking were used. The data were collected between September 2013 and June 2014, and the cows were from 2 to 127 DIM. Cows with <5 (n = 1) or >65 DIM (n = 4) or with duplicated records (n = 1) were excluded from analysis, resulting in 826 cows kept in 55 herds. Mean blood BHB concentrations at each DIM were calculated for the cows (826) and is depicted in Figure 1.

Cows in data set 1 were randomly divided into a calibration and a validation set. The calibration set (n = 496 cows from 31 herds) was used to develop a link between milk spectra and blood BHB, whereas the validation set (n = 330 cows from 24 herds) was used to validate the prediction model and for evaluation of the IP and DP approaches. Descriptive statistics of the calibration and validation sets of data set 1 are presented in Table 1. Phenotypic associations of the measured blood BHB with Foss-predicted milk BHB and predicted blood BHB from spectra of validation set by models developed in our study were estimated.

Data Set 2. The large data set (data set 2) originally contained 1,173,141 observations recorded from September to December 2014. Unlike data set 1, data set 2 did not contain BHB measured from blood. All

cows with <5 (n = 67 observations) or >65 DIM (n = 934,600 observations) were excluded from analysis, resulting in 238,474 observations. Furthermore, cows with no pedigree information or with an unknown age at test day were removed from the data set. For better separation of animal effects from herd effects, herd test days (**HTD**) with less than 2 records were also excluded from the data set. Finally, 158,028 observations from 107,988 cows (daughters of 8,339 sires and 100,423 dams) kept in 9,102 herds remained for estimation of (co)variance components of the spectra. A pedigree file containing animals with records and their ancestors was available. The total number of animals in the pedigree file that had a link to the data file were 469,751.

Selection of Spectral Variables

The major proportion of milk is water, hence the water spectrum influences the milk spectra. Both the O-H bending region (approximately between 1,620 and 1,700 cm^{-1}) and the O-H stretching region (above 3,025 cm^{-1}) of water are more or less opaque for the infrared light used, resulting in noise-like regions (Afseth et al., 2010; Dagnachew et al., 2013a). Therefore, the 2 regions comprising 536 spectral data points were excluded and only the remaining 524 spectral data points (926–1,617 cm^{-1} and 1,701–3,025 cm^{-1}) were considered for further analysis. These 524 spectral variables are referred to as region I.

Spectral region between 1,803 and 2,800 cm^{-1} (262 wave numbers) has been reported to have no specific bands or useful chemical information (Andersen et al., 2002; Iñón et al., 2004; Dagnachew et al., 2013a). Region I without these spectral variables (between 1,803–2,800 cm^{-1}) is referred to as region II (i.e., region II is a subset of region I).

Preprocessing of Spectra

Calibrations of prediction models, for relationship of milk spectra and blood BHB, and genetic analyses were performed on both unprocessed and preprocessed spectra. The selected spectral variables were preprocessed by 2 methods. First, second derivatives of the spectra by the Savitzky-Golay (**SG**) numerical algorithm with

Table 1. Descriptive statistics for reference blood BHB (mmol/L) in data set 1 and its subsets: calibration, subset of calibration (extreme values ≤0.5 or ≥1.4 mmol/L), and validation sets

Data	No. of cows	Mean	SD	Minimum	Maximum
Data set 1	826	0.760	0.743	0.1	6.3
Calibration set	496	0.734	0.725	0.1	6.3
Calibration subset	296	0.716	0.928	0.1	6.3
Validation set	330	0.800	0.768	0.1	5.5

9 window-size and second-order polynomials were calculated. Second, the spectra preprocessed by the SG were further preprocessed using extended multiplicative signal correction (**EMSC**). Preprocessing was performed on both region I and region II.

Multivariate Calibration of Prediction Models

The calibration data set (n = 496) was used to develop a link between blood BHB and milk spectra using the pls package (Mevik and Wehrens, 2007) implemented in R (R Core Team, 2016). Partial least squares (**PLS**) regression analyses were done on all 496 blood BHB values in the calibration set, and on a subset with 296 observations with extreme blood BHB values (low: <0.6 mmol/L, high: ≥1.4 mmol/L). In the analyses, blood BHB was used as a response variable (y) whereas unprocessed or preprocessed spectra (region I or II) were used as predictor variables (X).

The calibration models were cross-validated using 10 random segments, and the optimum number of PLS factors were determined based on the first local minimum value in root mean squared error of prediction of the cross-validation (**RMSE$_{cv}$**). The calibration models were then applied to the validation data set. The PLS regression parameters, such as regression coefficients $(\hat{\beta}_{PLS})$, matrices of weights (**W**) that reflect covariance structures between y and X, matrices of factor scores (**T**), matrices of y-loadings (**Q**), and matrices of X-loadings (**P**), were used in the subsequent predictions and calculations. Predictions were performed following the DP or IP approaches. Figure 2 shows a schematic representation of these 2 prediction approaches.

DP

The DP approach has several steps: spectral variables dimension reduction into few factors, estimation of covariance components for the factor scores from data set 2, prediction of random components for factor scores from the validation set using the estimated covariance components, and conversion of predicted factor scores into predicted blood BHB. The steps are described in detail in subsequent sections.

Spectral Variables Dimension Reduction. Direct genetic analysis to estimate (co)variance components for the random effects of the mixed model of all the selected spectral variables (i.e., 524 or 262 spectral data points) simultaneously was not possible with currently available analytical packages used in quantitative genetics. They are limited to fewer traits in multi-trait analysis (Meyer, 2007; Madsen and Jensen, 2008; Gilmour et al., 2009). Moreover, many of the spectral variables are highly collinear and the redundancy needs to be removed or absorbed. Dimension reduction is usually done by principal component analysis (**PCA**), PLS regression, or factor analysis. In the current study, we reduced the dimension of the spectral variables into factor scores through a weight matrix (**W**) obtained from the PLS regression with respect to blood BHB, as described above. The PLS factors mainly contain information related to the response variable(s) in the regression, and hence are expected to give better prediction than PCA components that contain general info in spectra. Previously, PCA was used on the same data set and results from such an analysis has been reported (Belay et al., 2015).

Estimation of Covariance Components for Factor Scores. A matrix of factor scores (**T**) were calculated for observations of spectra in data set 2 using the weight matrix (**W**) that had been obtained by PLS regression on the calibration part of data set 1:

$$\mathbf{T = XW}, \qquad [1.1]$$

where **T** is an $n \times c$ matrix of factor scores, with n being number of observations (n = 158,028) and c the number of factors. Numbers of factors were chosen by cross-validation in PLS regression for the calibration part of data set 1. **X** is an $n \times k$ spectral data matrix for data set 2, with k being number of spectral variables. **W** is a $k \times c$ weight matrix that reflects the covariance structure between milk spectra (**X**) and blood BHB (**y**); **W** is determined as a function of **X** and **y** by PLS regression in the calibration part of data set 1.

The factor scores characterize the relationship of the milk information to the blood BHB and give the relationship among the spectral variables. Factor scores were then treated as traits in multiple-trait mixed model analyses. A repeatability test day animal model was used to estimate variance components for the factor scores, **T**, from data set 2 (only spectra, no blood BHB). The model in matrix notation was

$$\mathbf{t = Xb + Za + Wp + Hd + e}, \qquad [1.2]$$

where **t** is a vector of factor scores (the **t** of 1 milk sample above the other); **b** is a vector of fixed effects of breed (2 levels), lactation number (1 to 4), herd size (3 levels) × lactation stage (6 levels), and months of test (4 levels); **a** is a vector of random animal genetic effects; **p** is a vector of random permanent environmental effects; **d** is a vector of random HTD effects; and **e** is vector of random residual effects. **X**, **Z**, **W**, and **H** are design matrices that relate records to the corresponding (fixed and random) effects. The 2 breed levels are

Polish Holstein-Friesian (black-white), which accounted for 86.4% of the records, and the other 15 breeds (such as Polish Holstein-Friesian red-white, Simental, Polish red-white, and so on), which accounted for 13.6% of the records all together. Herd size was defined based on number of cows with records per herd in the original data set (before edition) and categorized as small (<35 cows), medium (35–99 cows), and large (≥100 cows). Each group contained similar numbers of cows. Days in milk (lactation stage) were categorized into 6 levels, each with 10 test days. Each of the 4 test months were modeled. The assumed (co)variance structure was

$$\mathrm{var}\begin{bmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{d} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \otimes \mathbf{A} & 0 & 0 & 0 \\ 0 & \mathbf{P} \otimes \mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{H} \otimes \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{R} \otimes \mathbf{I} \end{bmatrix},$$
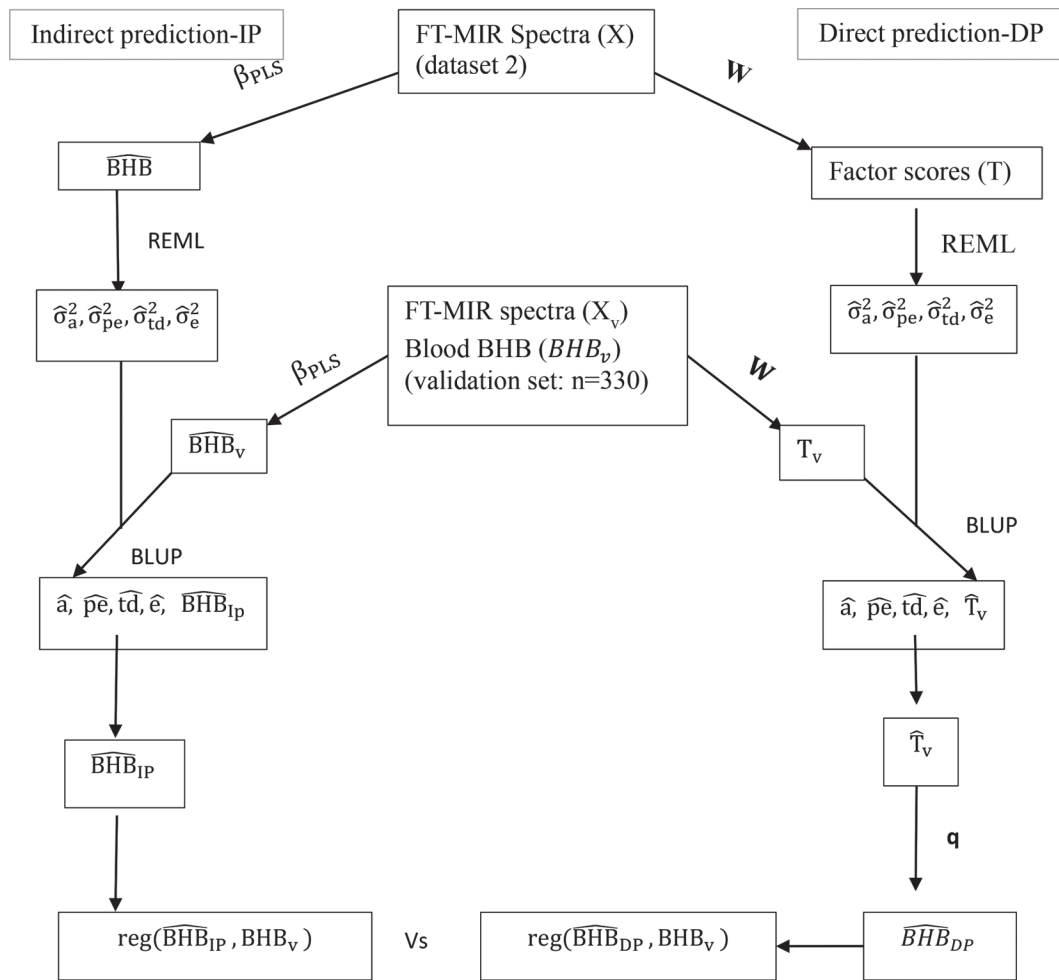
Figure 2. Schematic representation of the indirect (IP) and direct (DP) prediction approaches. In the IP, the phenotype for trait of interest (e.g., BHB) was predicted from milk spectra using regression coefficient (β_{PLS}), and this predicted trait was analyzed with a genetic model for ultimate phenotypic prediction. In the DP, multiple predicted factor scores that were obtained through a weight matrix (W) were analyzed with a genetic multivariate model before predicted model components were combined through the y-loading (q) to eventually predict the phenotype. Reg = regression; T_v = factor scores for observations in the validation set; $\sigma_a^2, \sigma_{pe}^2, \sigma_{td}^2,$ and σ_e^2 = estimates of variance components for genetic, permanent animal environment, herd test-day and residual, respectively; a, pe, td, and e = additive genetic, permanent animal environmental, herd test-day, and residual effects, respectively; $\widehat{BHB}$ = predicted BHB from spectra of data set 2 using the PLS regression-based prediction equation; $\widehat{BHB}_v$ = predicted BHB from spectra of validation set using the PLS regression-based prediction equation; $\widehat{BHB}_{IP}$ = predicted BHB using the IP approach; $\widehat{BHB}_{DP}$ = predicted BHB using the DP approach; BHB_v = measured/reference blood in the validation set.

where $\mathbf{G}$ is genetic covariance (5×5) matrix, $\mathbf{P}$ is the covariance (5×5) matrix for within-cow permanent environmental effects, $\mathbf{H}$ is the covariance (5×5) matrix for HTD effects, and $\mathbf{R}$ is the residual covariance (5×5) matrix, $\mathbf{I}$ and $\mathbf{A}$ are identity and additive relationship matrices, respectively, and $\otimes$ is the Kronecker product.

The (co)variance component estimates were obtained by the REML method using the multivariate average information-REML algorithm of WOMBAT (Meyer, 2007). These estimated variance components are population parameters that should characterize any data coming from the population. Preliminary bivariate analyses were performed and (co)variance component estimates from the bivariate analyses of the factor scores were pooled using the iterative summing of expanded part matrices approach (Mäntysaari, 1999) implemented in WOMBAT (Meyer, 2007). The pooled covariance matrices were priors in the multivariate REML.

BLUP Analyses for Factor Scores from Spectra of Validation Set. Once (co)variance components were estimated for factor scores from the large spectral data set (data set 2), BLUP values could be calculated for the random components of any new data set from the population using the estimated (co)variance components and the structural circumstances of the new data set (genetic relationship, permanent environment, and HTD design). Factor scores $(\mathbf{T_v})$ for observations in the validation set (n = 330) were calculated using the weight matrix $\mathbf{W}$ from the model calibration and milk spectra of the validation set as follows. Neither blood BHB nor milk spectra in the validation set were used in the model development:

$$\mathbf{T_v} = \mathbf{X_v}\mathbf{W}, \qquad [1.3]$$

where $\mathbf{T_v}$ is an $n_v \times c$ matrix of factor scores, $\mathbf{W}$ is as defined in Eq. [1.1], and $\mathbf{X_v}$ is an $n_v \times k$ spectral data matrix for the validation set, with k being number of spectral variables. The subscript v is used to indicate validation set.

A model similar to Eq. [1.2], with some modification in the fixed part of the model, was used to run BLUP for the $\mathbf{T_v}$ using the covariance components estimated with Eq. [1.2] and the $\mathbf{I}$ and $\mathbf{A}$ relevant for the validation set. In this model [1.4], fixed effects of lactation number (4 levels), lactation stage (6 levels), and year (2 levels) × season (2 levels: April to September and October to March) were fitted:

$$\mathbf{t_v} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}, \qquad [1.4]$$

where $\mathbf{t_v}$ is a vector of factor scores for observations in the validation set (with the $\mathbf{t_v}$ of 1 milk sample above the other), and other model elements were as defined in the Eq. [1.2].

Conversion of the Predicted Factor Scores into Predicted Blood BHB. In addition to predictions of the random effects $(\hat{\mathbf{a}}, \hat{\mathbf{p}}, \text{and } \hat{\mathbf{d}})$, predicted factor scores $(\hat{\mathbf{T}}_v)$ were given directly by WOMBAT from the BLUP run for the factor scores of the validation set. These predicted factor scores in multivariate form were converted into predicted blood BHB $(\widehat{\mathrm{BHB}}_{\mathrm{DP}})$ through the $\mathbf{Y}$-loading matrix $(\mathbf{Q})$ used in transposed vector form as

$$\widehat{\mathrm{BHB}}_{\mathrm{DP}} = \hat{\mathbf{T}}_v\mathbf{q}',$$

where $\mathbf{q}$ is a vector, not a matrix, as only a single response variable was in the PLS regression analysis. It had dimension $1 \times c$, where c is the number of factors retained and it relates factors to response variables.

IP

In this approach, BHB values were predicted from milk spectra using the PLS regression coefficient $(\hat{\beta}_{\mathrm{PLS}})$ estimated above, and the predicted phenotypes used in further analyses. This is the conventional approach used for genetic evaluation and other purposes in animal sciences or in other fields. The BHB was predicted as

$$\widehat{\mathrm{BHB}} = \mathbf{X}\hat{\beta}_{\mathrm{PLS}}, \qquad [2.1]$$

where $\widehat{\mathrm{BHB}}$ is predicted BHB phenotype from spectra $\mathbf{X}$ of data set 2 through PLS regression coefficient $(\hat{\beta}_{\mathrm{PLS}})$ found in the calibration part of data set 1.

Covariance components and the corresponding variance ratios were estimated by REML for the predicted BHB by fitting single-trait animal model considering the same effects as in Eq. [1.2]:

$$\widehat{\mathrm{BHB}} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}. \qquad [2.2]$$

The model elements are as defined in Eq. [1.2], but with univariate variance structure. We assumed $\mathrm{var}(\mathbf{a}) = \mathbf{A}\sigma_a^2$, $\mathrm{var}(\mathbf{p}) = \mathbf{I}\sigma_{pe}^2$, $\mathrm{var}(\mathbf{d}) = \mathbf{I}\sigma_d^2$, and $\mathrm{var}(\mathbf{e}) = \mathbf{I}\sigma_e^2$, where σ_a^2 is additive genetic variance, σ_{pe}^2 is permanent environmental variance, σ_d^2 is HTD variance, and σ_e^2 is residual variance. The BHB were then predicted for observa-

tions in the validation set $\left(\widehat{\mathrm{BHB}}_v\right)$ using the $\hat{\beta}_{\mathrm{PLS}}$ that was used in Eq. [2.1], but using spectra from the validation set ($\mathbf{X}_v$; i.e., $\widehat{\mathrm{BHB}}_v = \mathbf{X}_v\hat{\beta}_{\mathrm{PLS}}$).

Assuming similar effects as in Eq. [1.4], but with a single-trait animal model, BLUP solutions for fixed and random effects were found for $\widehat{\mathrm{BHB}}_v$ from validation set:

$$\widehat{\mathrm{BHB}}_v = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}, \qquad [2.3]$$

where model elements were as defined in Eq. [1.2] and Eq. [2.2]. A similar variance structure as in Eq. [2.2] was assumed.

For this BLUP, run on $\widehat{\mathrm{BHB}}_v$, the variance components used were estimated either from (1) single-trait REML (i.e., the one estimated in Eq. [2.2]), or (2) multiple-trait REML, as estimated in Eq. [1.2] after converting from multivariate covariance to univariate variance structures. The multivariate covariance structure from Eq. [1.2] for additive, permanent environmental, HTD, and residual covariance were converted into the corresponding univariate variance structure as

$$\hat{\sigma}_a^2 = \mathbf{qGq}',$$

$$\hat{\sigma}_{pe}^2 = \mathbf{qPq}',$$

$$\hat{\sigma}_d^2 = \mathbf{qHq}', \text{ and}$$

$$\hat{\sigma}_e^2 = \mathbf{qRq}. \qquad [2.4]$$

Predicted blood BHB $\left(\widehat{\mathrm{BHB}}_{\mathrm{IP}}\right)$ were directly obtained from WOMBAT together with predicted random effects and solutions for random residuals. Thus, we got 2 vectors of predicted BHB for observations in validation set, 1 from DP $\left(\widehat{\mathrm{BHB}}_{\mathrm{DP}}\right)$ and the other from IP $\left(\widehat{\mathrm{BHB}}_{\mathrm{IP}}\right)$. In addition to these predicted BHB, we measured blood BHB (reference values), which had not been used in calibration, from observations in the validation set.

Evaluation of the IP and DP Approaches

The 2 sets of predicted blood BHB $\left(\widehat{\mathrm{BHB}}_{\mathrm{DP}} \text{ and } \widehat{\mathrm{BHB}}_{\mathrm{IP}}\right)$ are 2 different predictions of blood BHB. Performance of the 2 approaches for prediction of BHB was evaluated based on adjusted coefficient of determination ($\mathbf{R}^2$) estimated by regressing the IP or DP predicted blood BHB against measured blood BHB (reference

values). Prediction with the IP or DP approach was also compared with prediction of BHB by PLS (using the PLS regression found in calibration on the milk spectra of the validation set).

RESULTS

Description of Reference Blood BHB

Table 1 shows descriptive statistics for reference blood BHB data. Content of BHB in the 826 blood samples analyzed ranged from 0.1 to 6.3 mmol/L, with an average of 0.760 mmol/L and a standard deviation of 0.743 mmol/L. More than 80% of the samples had <1.0 mmol/L. The most frequent concentrations observed were 0.3, 0.4, and 0.6 mmol/L. Out of the 826 cows sampled, 114 of them had a concentration $\geq$1.2 mmol/L of blood BHB. Mean blood BHB concentration for all cows (826) at each DIM was calculated and is given in Figure 1. The mean blood BHB concentration was generally high around the beginning of lactation and decreased as DIM progressed. Mean and standard deviation of BHB in calibration set were lower than the corresponding values in the validation set (Table 1). We also found a difference in ranges of BHB values between calibration and validation set.

Cross-Validation and Validation Results

The link between untransformed or log-transformed blood BHB and milk spectra was developed using PLS regression analysis on the calibration set (all, n = 496) and on its subset (n = 296 with extreme values). The results from such analyses are presented in Tables 2 and 3. Based on the 3 sets of spectra (unprocessed and preprocessed by SG and EMSC), the 2 spectral regions (regions I or II), and the 2 categories of blood BHB values (all or extreme), 9 different prediction models were developed. The idea was to find models with the better fit. Five to 10 PLS factors were retained based on the first local minimum value in $\mathrm{RMSE}_{\mathrm{cv}}$. Table 2 shows cross-validation and validation statistics for untransformed blood BHB predicted using the 9 prediction models developed. In the cross-validations, averages of predicted values were the same as corresponding mean reference values, but with smaller standard deviations and ranges than the reference values. These results (i.e., the low variation and the reduced range of predicted values) indicate lack of precision of the models on high values. For untransformed blood BHB, the R^2 in cross-validation ($\mathbf{R}^2_{\mathbf{cv}}$) ranged from 0.217 to 0.316, with $\mathrm{RMSE}_{\mathrm{cv}}$ ranging from 0.630 to 0.787 (Table 3). The $\mathrm{RMSE}_{\mathrm{cv}}$ were relatively high, which might be due to the lack of precision of the models on high values

Table 2. Cross-validation and validation descriptive statistics[1] for untransformed blood BHB (mmol/L) predicted from milk spectra using different calibration models

Model[2]	Cross-validation				Validation			
	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Unprocessed spectra								
All BHB values with region I	0.734	0.389	0.014	2.721	0.756	0.443	−0.059	2.612
Extreme BHB with region II	0.716	0.565	−0.156	3.304	0.756	0.598	−0.389	3.082
All BHB values with region II	0.734	0.390	−0.017	2.506	0.745	0.444	−0.085	2.629
Second derivative spectra (SG)								
All BHB values with region I	0.734	0.397	−0.119	2.551	0.746	0.425	−0.056	2.535
Extreme BHB with region II	0.716	0.608	−0.436	2.905	0.774	0.623	−0.426	3.327
All BHB values with region II	0.734	0.403	−0.118	2.257	0.733	0.429	−0.075	2.509
EMSC preprocessed spectra								
All BHB values with region I	0.734	0.389	−0.118	2.114	0.732	0.403	−0.175	2.119
Extreme BHB with region II	0.716	0.586	−0.845	2.900	0.730	0.593	−0.730	2.722
All BHB values with region II	0.734	0.392	−0.260	2.159	0.724	0.405	−0.206	2.084

[1]Mean, SD, minimum, and maximum of predicted blood BHB values (mmol/L) are presented.

[2]Spectra were preprocessed by Savitzky-Golay (SG) algorithm and extended multiplicative signal correction (EMSC).

of the data sets (Table 2) that had a high proportion of low values. This is evident from models developed based on extreme BHB values, where a majority of them were low with few high values. The logarithmic transformation of blood BHB values increased the R^2_{cv}, ranging from 0.313 to 0.381. The $RMSE_{cv}$ for the log-transformed blood BHB values were between 0.222 and 0.278.

In the validation, predicted BHB contents (Table 2) were smaller than the corresponding reference values (Table 1). For untransformed blood BHB, R^2 in validation (R^2_v) ranged from 0.308 to 0.374, with root mean square error of validation ($RMSE_v$) ranging from 0.607 to 0.638 (Table 3). Similar to the cross-validation,

log-transformation of BHB values increased R^2 in the validation, except for the extreme blood BHB values. For both untransformed and log-transformed blood BHB, R^2_v were generally higher than the corresponding estimates in cross-validation, except models developed on log-transformed extreme blood BHB values (Table 3). The $RMSE_v$ for untransformed BHB were also lower than the corresponding $RMSE_{cv}$, whereas the reverse was true for the log-transformed BHB. This indicates that prediction ability of models based on log-transformed BHB could be compromised compared with the untransformed BHB.

In both cross-validation and validation, preprocessing of spectra either by SG or EMSC generally increased

Table 3. Summary of partial least squares (PLS) regression prediction models for untransformed and $\log_{10}$-transformed blood BHB in cross-validation and validation under unprocessed, second derivative (SG), and EMSC preprocessed[1] spectra

Model	No. of factors[2]	Cross-validation[3]				Validation[4]			
		Untransformed BHB		Transformed BHB		Untransformed BHB		Transformed BHB	
		$RMSE_{cv}$	R^2_{cv}	$RMSE_{cv}$	R^2_{cv}	$RMSE_v$	R^2_v	$RMSE_v$	R^2_v
Unprocessed spectra									
All BHB values with region I	6 (8)	0.6396	0.2109	0.2318	0.3198	0.6065	0.3738	0.2469	0.3964
Extreme BHB with region II	5	0.7865	0.2760	0.2776	0.3130	0.6327	0.3186	0.2792	0.2277
All BHB values with region II	5 (7)	0.6397	0.2172	0.2326	0.3169	0.6153	0.3554	0.2468	0.3966
Second derivative spectra (SG)									
All BHB values with region I	5	0.6383	0.2201	0.2223	0.3730	0.6153	0.3554	0.2462	0.3999
Extreme BHB with region II	5	0.7748	0.2875	0.2628	0.3814	0.6199	0.3457	0.2765	0.2421
All BHB values with region II	5 (10)	0.6384	0.2186	0.2210	0.3807	0.6150	0.3562	0.2446	0.4075
EMSC preprocessed spectra									
All BHB values with region I	5 (10)	0.6302	0.2368	0.2227	0.3701	0.6312	0.3217	0.2414	0.4228
Extreme BHB with region II	5	0.7622	0.3159	0.2662	0.3718	0.6378	0.3075	0.2707	0.2741
All BHB values with region II	5 (10)	0.6301	0.2351	0.2228	0.3690	0.6313	0.3216	0.2397	0.4309

[1]Spectra were preprocessed by Savitzky-Golay (SG) algorithm and extended multiplicative signal correction (EMSC).

[2]Number of PLS factors, and numbers of factors in parentheses were for models based on log-transformed blood BHB.

[3]$RMSE_{cv}$ = root mean squared error of the cross-validation; R^2_{cv} = coefficient of determination of the cross-validation.

[4]$RMSE_v$ = root mean squared error of the validation; R^2_v = coefficient of determination of the validation.

Table 4. Estimates of variance ratios for genetic, permanent environment (PE), herd test days (HTD), and residual random effects for the factor scores calculated from raw spectra in region II and all (or extreme[1]) blood BHB values

Factor score	Variance ratio[2]			
	Genetic	PE	HTD	Residual
1	0.093 (0.093)	0.143 (0.143)	0.169 (0.169)	0.595 (0.595)
2	0.221 (0.227)	0.212 (0.213)	0.082 (0.080)	0.485 (0.480)
3	0.176 (0.180)	0.119 (0.122)	0.158 (0.166)	0.547 (0.531)
4	0.163 (0.156)	0.165 (0.162)	0.137 (0.130)	0.534 (0.552)
5	0.053 (0.056)	0.070 (0.075)	0.504 (0.480)	0.374 (0.388)

[1]Numbers in parentheses are variance ratio for factor scores calculated from raw spectra in region II and extreme blood BHB values.

[2]Ratio is relative to total phenotypic variance for each factor score. Standard error of variance ratios due to genetic, PE, HTD, and residual were 0.004–0.012, 0.004–0.01, 0.003–0.004, and 0.004, respectively.

R^2 or reduced prediction errors, except for some models with untransformed BHB in the validation analyses. In validation, better results (high R^2_v or low $RMSE_v$) were obtained with unprocessed spectra for untransformed BHB and with EMSC for log-transformed BHB. In the cross-validation, better results were obtained with EMSC for untransformed BHB and with EMSC or SG for log-transformed BHB. Despite the large number of spectral variables contained in region I, it had no effect on the R^2 of validation or cross-validation, except for validation of unprocessed spectra with untransformed BHB values (Table 3). Comparing the models with respect to the 2 sets of BHB values, prediction models with extreme BHB values (both untransformed and log-transformed) had generally higher R^2_{cv} and $RMSE_{cv}$, but had lower R^2_v than models with all BHB values (Table 3).

Genetic Parameters for the Factor Scores and BHB

Out of the 9 prediction models that were developed based on untransformed blood BHB, 4 of them were selected to be used in the genetic analysis for ultimate phenotypic prediction: 2 models from raw and 2 from SG preprocessed spectra of region II with all or extreme BHB values. Models based on spectra of region I were not selected, as they did not give better accuracy despite the large number of spectral variables in region I. Models developed based on log-transformed BHB were not used for genetic analyses, as IP and DP approaches can be evaluated independent of the BHB scale. Moreover, models based on log-transformed BHB had slightly higher prediction error in the validation than in the cross-validation, whereas the reverse is true for models from untransformed BHB (Table 3). Estimates of variance ratios for each factor score, calculated from raw and preprocessed spectra using the 4 selected calibration models, are presented in Tables 4 and 5. Genetic variance ratios (heritabilities) for the 5 factor scores calculated from unprocessed spectra ranged from 0.053 to 0.227 (Table 4) and from 0.081 to 0.158 (Table 5) for SG preprocessed spectra. The corresponding variance ratios of the permanent environmental effects ranged from 0.070 to 0.213 and from 0.074 to 0.153. Variance ratios of the HTD ranged from 0.080 to 0.504 and from 0.130 to 0.374 for the factors from raw and

Table 5. Estimates of variance ratios for genetic, permanent environment (PE), herd test day (HTD), and residual random effects for the factor scores calculated from Savitzky-Golay (SG) preprocessed spectra in region II and all (or extreme[1]) blood BHB values

Factor score	Variance ratio[2]			
	Genetic	PE	HTD	Residual
1	0.097 (0.095)	0.140 (0.140)	0.169 (0.169)	0.595 (0.595)
2	0.081 (0.084)	0.118 (0.111)	0.257 (0.259)	0.544 (0.545)
3	0.158 (0.144)	0.114 (0.106)	0.209 (0.231)	0.519 (0.519)
4	0.102 (0.096)	0.086 (0.074)	0.376 (0.349)	0.437 (0.481)
5	0.113 (0.118)	0.118 (0.153)	0.173 (0.130)	0.595 (0.599)

[1]Numbers in parentheses are variance ratio for factor scores calculated from Savitzky-Golay (SG) preprocessed spectra in region II and extreme blood BHB values.

[2]Ratio is relative to total phenotypic variance for each factor. Standard error of variance ratios due to genetic, PE, HTD, and residual is 0.006–0.008, 0.000–0.007, 0.003–0.004, and 0.004, respectively.

Table 6. Multifactor (direct prediction) REML based estimates of variance ratios and variance components for genetic, permanent environment (PE), herd test day (HTD), and residual random effects for BHB found from raw or Savitzky-Golay (SG) preprocessed milk spectra from data set 2 (n = 158,028) using the 4 selected calibration models[1]

Model	Variance ratio (variance component)			
	Genetic	PE	HTD	Residual
Unprocessed spectra				
All BHB values with region II	0.111 (0.018)	0.142 (0.023)	0.279 (0.045)	0.468 (0.076)
Extreme BHB with region II	0.109 (0.036)	0.144 (0.047)	0.275 (0.090)	0.473 (0.156)
Second derivative spectra (SG)				
All BHB values with region II	0.083 (0.017)	0.158 (0.032)	0.342 (0.070)	0.416 (0.085)
Extreme BHB with region II	0.088 (0.036)	0.145 (0.060)	0.337 (0.139)	0.430 (0.177)

[1]Estimated multivariate covariances have been converted into one-trait variance structure relevant for BHB prediction. Ratios are relative to total phenotypic variance for BHB from each model. Values in parentheses are estimates for variance components.

preprocessed spectra, respectively. The corresponding variance ratios for the residual effects varied from 0.374 to 0.595 and from 0.437 to 0.595.

Corresponding variance components for blood BHB were calculated from the estimated covariance structures of factor scores using Eq. [2.4]. Table 6 presents estimated variance ratios and variance components for genetic, PE, HTD, and residual of BHB. For BHB from unprocessed spectra, average estimates of variance ratios (variance ratios for BHB from all and extreme BHB values were averaged within spectral sets) for genetic, PE, HTD, and residual were 0.110, 0.143, 0.277, and 0.471, respectively. The corresponding values for BHB from SG preprocessed spectra were 0.086, 0.152, 0.340, and 0.423. Variance components for BHB were also estimated by univariate REML using Eq. [2.2], where spectral variables had first been converted into a single trait (BHB) through the PLS regression coefficient. Variance components and variance ratios for such BHB were slightly lower than the genetic parameters presented in Table 6, except the estimates for HTD, and estimated residual effects that were the same (Tables 6 and 7).

Most of the factor scores and BHB that were predicted from unprocessed spectra had higher estimates of heritability and proportion of variance due to PE and HTD effects than those from SG preprocessed spectra (Tables 4–7). The larger estimates for factors and BHB from unprocessed spectra may be due to unprocessed spectra possibly containing unwanted heritable variation, which could be removed by preprocessing. Spectral preprocessing removes not only unwanted variations (such as variation in intensity of light sources, scattering, contaminants, optical path length, and so on) in spectra, but also some real molecular structures or milk constituents, which might be heritable.

Prediction Ability of the IP and DP Approaches

Performance of the IP and DP approaches were evaluated based on R^2 estimated by regressing the IP- or DP-predicted blood BHB on the reference blood BHB values of the validation data set that had not been used in model calibrations. Table 8 presents the estimated R^2 for the IP and DP approaches. The R^2 for the DP method were intermediate and ranged from

Table 7. Univariate (indirect prediction) REML based estimates of variance ratios and variance components for genetic, permanent environment (PE), herd test day (HTD), and residual random effects for BHB found from raw or Savitzky-Golay (SG) preprocessed milk spectra from data set 2 (n = 158,028) using the 4 selected calibration models[1]

Model	Variance ratio (variance component)			
	Genetic	PE	HTD	Residual
Unprocessed spectra				
All BHB values with region II	0.103 (0.017)	0.141 (0.023)	0.288 (0.047)	0.468 (0.076)
Extreme BHB with region II	0.101 (0.033)	0.142 (0.047)	0.283 (0.094)	0.473 (0.156)
Second derivative spectra (SG)				
All BHB values with region II	0.075 (0.015)	0.155 (0.032)	0.353 (0.072)	0.416 (0.085)
Extreme BHB with region II	0.081 (0.033)	0.142 (0.059)	0.347 (0.143)	0.430 (0.178)

[1]Ratio is relative to total phenotypic variance for BHB from each model. Values in parentheses are estimates for variance components.

Table 8. Coefficient of determination between reference blood BHB values and blood BHB predicted by the direct and indirect prediction approaches from milk spectra[1]

Calibration model	Indirect prediction (IP)		Direct prediction (DP)
	Variances from single REML	Variances from multiple REML	
Unprocessed spectra			
All BHB values with region II	0.2865	0.2898	0.2692
Extreme BHB with region II	0.2775	0.2805	0.2631
Second derivative spectra (SG)			
All BHB values with region II	0.2943	0.2972	0.2804
Extreme BHB with region II	0.3061	0.3091	0.2978

[1]In the IP approach, where spectral variables first converted into single-trait and then genetic analysis was applied on the trait for ultimate phenotypic prediction, variances estimated from single-trait REML or multiple REML (after converting into variance structure) were used. In the DP approach, spectral variables reduced to factor scores that were analyzed using multitrait genetic analysis and eventually combined into the phenotypic trait.

0.263 to 0.298, whereas the corresponding estimates for IP method, when variance components from multiple REML were used, ranged from 0.281 to 0.301 and from 0.278 to 0.306 when variance components from single-trait REML were used (Table 8). The predictability of the IP approach was slightly higher compared with the predictability of the DP approach; this means that a more accurate prediction of BHB was found when univariate variance structure was used than when multivariate covariance structures were used. Predictability of the 2 approaches were compared with the predictability of models given in Table 3 (PLS regression based predictions equations) for untransformed BHB in the validation analyses. The PLS regression-based prediction equations are the commonly used methods for phenotyping of trait of interest from milk FT-MIR spectra. Predictability of the IP and DP approaches were lower than predictability of equations developed based on the classical PLS regression in validation (Table 3).

As in the calibration models, preprocessing of spectra slightly improved accuracy of BHB prediction in both the IP and DP approaches. The improvement in accuracy due to preprocessing was slightly better in DP than in IP approaches. This indicated that the DP approach could perform better with spectra that contain less noisy information; noisy information in multivariate form could result in inferior performance.

DISCUSSION

Multivariate Calibration Models

The distribution of the data in the calibration set was slightly different than in the validation set, mainly due to lower mean and standard deviation of the reference values (Table 1). This could explain some difference in cross-validation and validation statistics. That the

R^2_v was generally higher than the R^2_{cv} might be due to higher mean and standard deviation of the reference values in validation. It has been shown that R^2 is highly dependent on distribution of the data and especially on the range of data (Grelet et al., 2016). Because of the way in which blood BHB was measured (i.e., values with few digits: 0.1, 0.2, ..., 6.3), many samples had the same BHB values; this resulted in a large number of few distinct values. Such duplication in BHB values (not in the corresponding spectra) could also influence the R^2_{cv} by reducing variation or range of values in the random segments used for the cross-validation. In the validation set, data were not divided into random segments, so existing variation in the blood BHB was available. That could possibly result in the higher observed R^2_v than R^2_{cv}.

The R^2_{cv} of the prediction models developed in our study were low, but in the range of estimates reported for untransformed milk BHB (0.10 to 0.64) or for log-transformed milk BHB (0.09 to 0.63; de Roos et al., 2007). Grelet et al. (2016) found R^2_{cv} of 0.71 and R^2_v of 0.63 for milk BHB, larger than estimates found in the current study. With blood BHB used as a reference value in calibration, Broutin (2015, 2016) also found higher R^2_{cv} (0.7360 or 0.5999) than that observed in our study. The predictive ability of calibration models developed in the present study may not be sufficient to determine exact values of blood BHB, but may allow for a rough screening to distinguish cows with high or low values. It has been concluded that FT-MIR–predicted ketone bodies may be promising as screening tool for ketosis at herd level, but not accurate enough for management decisions at an individual animal level (de Roos et al., 2007; van der Drift et al., 2012; Grelet et al., 2016).

The correlations between reference BHB and predicted BHB obtained by the models developed in our study (averaged to 0.584) were higher than the correla-

tion between reference blood BHB and Foss-predicted milk BHB (0.567). This indicates that these models may be more appropriate to indicate ketosis, as they predict blood BHB instead of milk BHB. It also shows the interest of predicting blood values directly from FT-MIR spectra rather than using milk BHB from FT-MIR spectra. The R^2 between reference and predicted blood BHB (Table 3) also indicate that milk spectra would contain substantial amount of information about BHB. Reported phenotypic correlations between reference blood BHB and reference milk BHB vary widely, ranging from 0.66 to 0.89 (Enjalbert et al., 2001; Denis-Robichaud et al., 2014; Friedrichs et al., 2015); correlation coefficients found in the current study were in the lower range of the values reported in those studies. However, only Broutin (2016) reported on the correlation between reference blood BHB and predicted blood BHB from milk spectra, finding a correlation of 0.7370.

Several factors could contribute to the degree of accuracy of prediction models observed in our study. Relationship between blood BHB and milk spectra might not be linear, which could in part explain the observed low R^2. The R^2 of prediction models and concentration of analyte (e.g., fat composition and so on) are known to have direct relationships (Soyeurt et al., 2006; Rutten et al., 2009). Infrared absorbance is directly proportional to concentration of analyte or substance (Beer's law), indicating that analytes with very low concentrations (e.g., BHB) are difficult to detect by the FT-MIR spectroscopy. The concentration of BHB in milk is very low (21.7 mg/L given its molar mass of 104.11 g/mol), which is below the detection limit (100 mg/L) of the FT-MIR spectroscopy (Dardenne et al., 2015). Therefore, it is important to note that calibration of BHB in milk can only be done by indirect links with global milk composition, not by the specific spectral responses of BHB in milk (Grelet et al., 2016).

Moreover, the 2 information sources that were used in our study, milk spectra and blood BHB, were from different media (milk and blood). Genetic differences between cows in udder ketone body metabolism may exist and could influence excretion of ketone bodies from blood to milk (van der Drift et al., 2012). van der Drift et al. (2012) also found that the random effect of herd explained considerable variation in the probability of hyperketonemia for cows. Those authors explained the herd differences in the association between blood and milk ketone body concentrations by time of milk sampling, feeding, and blood sampling that were not identical on the different farms. Time of sampling (before or after feeding, morning or evening milking) could result in variation of BHB in blood and milk, as there might be difference in metabolism of BHB production in milk and blood.

Evaluation of the IP and DP Approaches

The slightly better prediction of blood BHB from milk FT-MIR spectra by the IP than the DP approach was not in line with our expectation. It is also in contrast to the work of Dagnachew et al. (2013b), who reported better prediction in accuracy of EBV for milk contents in goat by DP than IP approaches. Bonfatti et al. (2017) also reported results that are mostly in contrast with the work of Dagnachew et al. (2013b), who found high rank correlations (>0.94) between IP and DP predicted EBV of all traits investigated. Bonfatti et al. (2017) reported <0.5 rank correlations between EBV predicted by IP and DP for most traits included in their study. Reasons why the DP approach resulted in better prediction for EBV (e.g., Dagnachew et al., 2013b), but not for phenotypic are not clear, but could be due to difference in methods of comparison (correlation vs prediction error variance) and type of parameters compared (phenotype vs EBV). Genetic parameter estimates (e.g., heritability) for BHB using covariance components (DP) after converting into univariate variance structure were higher (Table 6) than corresponding estimates using variance components (IP; Table 7), indicating better information content in the DP approach. However, neither phenotypic prediction from the multivariate mixed model using spectral variables that were reduced into few components by PLS (Table 8) nor principal component analysis (**PCA**; Belay et al., 2015) were promising. It is therefore important to verify if such results from the current study or previous studies (Dagnachew et al., 2013b; Belay et al., 2015) will be reproducible and to look for reasons behind the reported results for example using simulated data.

In the DP approach, dimension of spectral variables can be reduced into a few factor scores by PLS regression, as in the current paper, or into latent traits by PCA. Covariance components for the latent traits from PCA are population parameters that characterize any information coming from the population, as they represent any information available in the milk spectra, whereas covariance components for factor scores from PLS regression mainly contain information related to the particular trait used in the calibration. The PLS factors are thus expected to give better prediction of the trait than the one with latent traits from PCA. However, with PLS, information about other milk composition traits not included in the calibration may not be retained in the factor scores, as also indicated by Dagnachew et al. (2013b) and Bonfatti et al. (2017). The expected better prediction of traits with PLS model was confirmed. For example, prediction accuracy of DP was much lower than the IP approach when PCA was used (Belay et al., 2015) compared with when PLS

was used for spectral dimension reduction (Table 8). One possible reason for this could be that the retained 8 latent traits from PCA, which explained 99% of the total spectral variation (Belay et al., 2015), did not contain as much relevant information about the blood BHB as those 5 PLS factors used in our study did. Dagnachew et al. (2013b) also used 8 latent traits to extract genetic component of the FT-MIR spectra and indicated the possible existence of relevant information in the remaining 1% of the total spectral variation. Interestingly, Bonfatti et al. (2017) also showed that a considerable amount of information needed to predict phenotypes is lost when using 99% of original spectral variability, and loss of such information could affect prediction of EBV from spectral information. Those authors further showed that information left in 0.01% of original spectral variability is fundamental for prediction of some of the traits included.

Several possible reasons exist for the slightly lower prediction accuracy in the DP compared with the IP approach found in our study. It could be due to the low genetic correlations observed among the latent traits (factor scores). Expected improvement in accuracy of prediction from multivariate analysis would be due to its ability to account for the covariance among the traits. When the covariance or correlation among the traits are very low, multivariate analysis might not perform better than the univariate one. Any errors in the covariance estimation or the modeling of the observations may also have reduced the accuracy. The genetic and environmental parameters used in BLUP analysis are estimates and possibly contain errors. Use of estimates with errors in multivariate analysis would affect accuracy of prediction (Schaeffer, 1984; Thompson and Meyer, 1986). Response to selection, which is directly proportional to accuracy of predicted breeding values, highly depends on the precision of the estimates and the applied variance components (Villanueva et al., 1993). Under such conditions, univariate models can provide more precise estimates than multitrait models. Accuracies will also depend on the relevance of the models used, and whether or not anything can be gained by using the mixed model.

Lack of enough information about contemporary cows in the validation data set with blood BHB samples could be another major contributing factor to the poor performance observed in multivariate analysis. If no structure of the random effects of the model exists in the data to be predicted, there may be no benefit from using a multivariate mixed model. The number of cows in the validation set were small and each cow had only 1 measured blood BHB; hence, the cows in validation set could possibly be not well connected to each other genetically. In addition, almost all of the HTD classes

contained only 1 cow, which might impose difficulty in separating herd effect from genetic effect of cows and contribute to low accuracy in the multivariate analysis. With FT-MIR spectra, for more contemporary cows in the validation, more info that is multivariate could have been available. An attempt was made to merge the validation set with data set 2 to increase amount of information in the validation set (such as increasing number cows in the HTD classes); however, as these 2 data sets were collected in different years, they had only 1 cow in common. Hence, we failed to use information in both data sets for better accuracy in multiple model.

In our study, correlations of blood BHB with milk fat (0.35), protein (−0.06), and lactose (−0.20) contents were low to medium, which might have contributed to the lower accuracy in DP. These milk contents were prediction from spectra by Foss calibration. For example, Bonfatti et al. (2017) found a positive relationship between rank correlation and correlations of traits of interest with major sources of spectral variation (such as milk protein, fat, and lactose contents), and that variability of the traits of interest is better explained when they are highly correlated with the major sources of spectral variation. In addition, for traits more correlated with the major sources of spectral variation, the DP approach is more likely to be effective (Bonfatti et al., 2017). Those authors indicated that the better accuracy of EBV from the DP than from the IP that Dagnachew et al. (2013b) reported would be due to large contribution of milk protein, fat, and lactose contents to spectral variation.

In addition, R^2_{cv} of the calibration models could affect accuracy of the DP approach. For example, in Dagnachew et al. (2013b), R^2_{cv} was very high (>0.94) and the multivariate model performed better than the univariate one, possibly as a consequence of this. In the current study, the predictive ability of the calibration models was much lower (explained less than 45% of the variation in the blood BHB, Table 3), and the multivariate models performed slightly inferior to the univariate ones (Table 8). From these 2 studies, we can see that accuracy of predicting breeding values or phenotype seems to depend on predictive ability of the calibration model. This could lead us to the conclusion that, for DP to work better, there should be a strong relationship between the trait of interest and spectral variables. In other words, we should first be sure that the univariate method (IP) is working well with the data on hand before embarking on the DP method. Under such conditions, where the univariate method was found to be working, the multivariate method might perform better, but this should be an assumption that needs to be made. Bonfatti et al. (2017) reported that rank correlations between EBV obtained by the IP and

the DP approach are not related to the accuracy the calibration equations; however, the relationships between accuracies of EBV obtained by the 2 approaches and accuracy of calibration equations are not well established.

Both the IP and DP approaches had lower predictability for the phenotype than the predictability of equations developed based on the classical PLS regression. This indicated that inclusion of cows' circumstances at a given test day into the IP or DP model did not improve prediction of blood BHB from milk FT-MIR spectra. Therefore, for phenotypic prediction, the classical PLS regression-based prediction equation seems to be the method of choice. In our study, information related to the cow was added into the models after PLS regression or in the mixed model analysis of predicted trait or factor scores. On the other hand, it has been shown that it is possible to directly add the information to the spectra before PLS. For example, Vanlierde et al. (2015) included DIM directly into spectra using Legendre polynomial to predict methane, and prediction equations developed in such a way were shown to be more robust than equations that did not integrate the DIM information. Similarly, Shetty et al. (2017) used milk yield and live weight as predictors along with spectral variables to predict residual feed intake and DMI. They showed improvement in accuracy of model that included spectral information along with milk yield and live weight as predictors for DMI. Therefore, inclusion of cows' circumstances directly into spectra before PLS or using them as predictors along with spectral information during PLS can be an alternative to improve prediction accuracy for blood BHB from milk FT-MIR spectra.

CONCLUSIONS

In this study, predictive ability of the DP and IP approaches were evaluated using measured blood BHB and milk spectra-predicted blood BHB. A calibration and an independent validation data set was used. Accuracy of prediction with the 2 approaches were similar. Slightly better prediction of BHB was found when univariate variance structure was used (IP) than when multitrait covariance structures were used (DP) in mixed models. Prediction accuracies of the developed calibration models were also low, which could partly be due to a weak relationship between milk spectra and blood BHB. Blood BHB log-transformation, spectral preprocessing, and use of extreme blood BHB values have improved prediction accuracy of the calibration models and the 2 approaches. Conclusive remarks on the importance of keeping spectral data in a multivariate form for prediction of phenotype and model components (EBV, HTD, and so on) may be found in data sets where the trait of interest has strong relationships with spectral variables.

REFERENCES

Afseth, N. K., H. Martens, Å. Randby, L. Gidskehaug, B. Narum, K. Jørgensen, S. Lien, and A. Kohler. 2010. Predicting the fatty acid composition of milk: A comparison of two Fourier transform infrared sampling techniques. Appl. Spectrosc. 64:700–707.

Andersen, S. K., P. W. Hansen, and H. V. Andersen. 2002. Vibrational spectroscopy in the analysis of dairy products and wine. Accessed Jul. 14, 2016. www.foss.dk/~/media/B4E6F6119A7548F4B380C3E6C958DBD1.ashx.

Andersson, L. 1988. Subclinical ketosis in dairy cows. Vet. Clin. North Am. Food Anim. Pract. 4:233–251.

Belay, T. K., K. Słoniewski, Z. Kowalski, and T. Adnoy. 2015. Predicting ketosis from milk mid infrared (MIR) spectra using multivariate mixed models. In Proc. Third DairyCare Conference, Croatia, Zadar. Accessed Jun. 20, 2016. www.dairycareaction.org/uploads/2/4/2/6/24266896/belay_tuesday.pdf.

Bonfatti, V., D. Vicario, L. Degano, A. Lugo, and P. Carnier. 2017. Comparison between direct and indirect methods for exploiting Fourier transform spectral information in estimation of breeding values for fine composition and technological properties of milk. J. Dairy Sci. 100:2057–2067.

Broutin, P. 2015. Blood BHB determination by mid infrared spectroscopy for the monitoring of the cows metabolic activity and detection of subclinical ketosis—A new approach. ICAR Technical Workshop. Accessed Sep. 20, 2016. www.icar.org/wp-content/uploads/2015/09/4-Pierr.pdf.

Broutin, P. 2016. Determination of the concentration of a component in one fluid of an animal by spectroscopic analysis of another fluid. United States, Bentley Instrument (Lille, FR) 20160238520. Accessed Oct. 12, 2016. http://www.freepatentsonline.com/y2016/0238520.html.

Dagnachew, B. S., A. Kohler, and T. Ådnøy. 2013a. Genetic and environmental information in goat milk Fourier transform infrared spectra. J. Dairy Sci. 96:3973–3985.

Dagnachew, B. S., T. H. E. Meuwissen, and T. Ådnøy. 2013b. Genetic components of milk Fourier-transform infrared spectra used to predict breeding values for milk composition and quality traits in dairy goats. J. Dairy Sci. 96:5933–5942.

Dardenne, P., J. Fernandez, O. Baeten, and B. Lecler. 2015. Untargeted contaminant detection by IR. Wallon Agricultural Research Centre. Accessed Dec. 15, 2016. http://www.namur2015.be/docs/pdf/Session%205%20-%202%20-%20Dardenne.pdf.

de Roos, A., H. Van Den Bijgaart, J. Hørlyk, and G. De Jong. 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. J. Dairy Sci. 90:1761–1766.

Denis-Robichaud, J., J. Dubuc, D. Lefebvre, and L. DesCôteaux. 2014. Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. J. Dairy Sci. 97:3364–3370.

Duffield, T. F., K. Lissemore, B. McBride, and K. Leslie. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. J. Dairy Sci. 92:571–580.

Duffield, T. F., D. F. Kelton, K. E. Leslie, K. D. Lissemore, and J. H. Lumsden. 1997. Use of test day milk fat and milk protein to detect subclinical ketosis in dairy cattle in Ontario. Can. Vet. J. 38:713–718.

Enjalbert, F., M. Nicot, C. Bayourthe, and R. Moncoulon. 2001. Ketone bodies in milk and blood of dairy cows: Relationship between concentrations and utilization for detection of subclinical ketosis. J. Dairy Sci. 84:583–589.

Friedrichs, P., C. Bastin, F. Dehareng, B. Wickham, and X. Massart. 2015. Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools. Palais des Congrès, Namur, Belgium, 16-17 April 2015. Biotechnol. Agron. Soc. Environ. 19:97–124.

Geishauser, T., K. Leslie, J. Tenhag, and A. Bashiri. 2000. Evaluation of eight cow-side ketone tests in milk for detection of subclinical ketosis in dairy cows. J. Dairy Sci. 83:296–299.

Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson, and D. Butler. 2009. ASReml user guide release 3.0. VSN International Ltd., Hemel Hempstead, UK.

Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. F. Pierna, F. Colinet, and P. Dardenne. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825.

Iñón, F. A., S. Garrigues, and M. de la Guardia. 2004. Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. Anal. Chim. Acta 513:401–412.

LeBlanc, S. J., K. Leslie, and T. Duffield. 2005. Metabolic predictors of displaced abomasum in dairy cattle. J. Dairy Sci. 88:159–170.

Madsen, P., and J. Jensen. 2008. DMU: A User's Guide. A package for analysing multivariate mixed models. Version 6, release 4.7. DJF, Foulum, Denmark.

Mäntysaari, E. 1999. Derivation of Multiple Trait Reduced Random Regression (RR) Model for the First Lactation Test Day Records of Milk, Protein and Fat. Page 8 in 50th Annual Meeting. Europ. Ass. Anim. Prod. Mimeo. Zurich, Switzerland, Aug. 23–26, 1999.

McArt, J. A., D. Nydam, and G. Oetzel. 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. J. Dairy Sci. 95:5056–5066.

Mevik, B.-H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18:1–24.

Meyer, K. 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Univ. Sci. B 8:815–821.

Oetzel, G. R. 2004. Monitoring and testing dairy herds for metabolic disease. Vet. Clin. North Am. Food Anim. Pract. 20:651–674.

R Core Team. 2016. R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rutten, M. J., H. Bovenhuis, K. Hettinga, H. van Valenberg, and J. Van Arendonk. 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. J. Dairy Sci. 92:6202–6209.

Schaeffer, L. 1984. Sire and cow evaluation under multiple trait models. J. Dairy Sci. 67:1567–1580.

Seifi, H. A., S. J. LeBlanc, K. E. Leslie, and T. F. Duffield. 2011. Metabolic predictors of post-partum disease and culling risk in dairy cattle. Vet. J. 188:216–220.

Shetty, N., P. Løvendahl, M. Lund, and A. Buitenhuis. 2017. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. J. Dairy Sci. 100:253–264.

Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J. Dairy Sci. 89:3690–3695.

Soyeurt, H., I. Misztal, and N. Gengler. 2010. Genetic variability of milk components based on milk infrared spectral data. J. Dairy Sci. 93:1722–1728.

Suthar, V. S., J. Canelas-Raposo, A. Deniz, and W. Heuwieser. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. J. Dairy Sci. 96:2925–2938.

Thompson, R., and K. Meyer. 1986. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest. Prod. Sci. 15:299–313.

van der Drift, S., R. Jorritsma, J. Schonewille, H. Knijn, and J. Stegeman. 2012. Routine detection of hyperketonemia in dairy cows using Fourier transform infrared spectroscopy analysis of β-hydroxybutyrate and acetone in milk in combination with test-day information. J. Dairy Sci. 95:4886–4898.

Vanlierde, A., M.-L. Vanrobays, F. Dehareng, E. Froidmont, H. Soyeurt, S. McParland, E. Lewis, M. H. Deighton, F. Grandl, and M. Kreuzer. 2015. Hot topic: Innovative lactation-stage-dependent prediction of methane emissions from milk mid-infrared spectra. J. Dairy Sci. 98:5740–5747.

Villanueva, B., N. Wray, and R. Thompson. 1993. Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. Anim. Sci. 57:1–13.

Walsh, R. B., J. Walton, D. Kelton, S. LeBlanc, K. Leslie, and T. Duffield. 2007. The effect of subclinical ketosis in early lactation on reproductive performance of postpartum dairy cows. J. Dairy Sci. 90:2788–2796.