J. Dairy Sci. 106:690–702
https://doi.org/10.3168/jds.2022-22217

Diagnostic milk biomarkers for predicting the metabolic health status of dairy cattle during early lactation

S. Heirbaut,[1] X. P. Jing,[1,2] B. Stefańska,[3] E. Pruszyńska-Oszmałek,[4] L. Buysse,[1] P. Lutakome,[1,5,6] M. Q. Zhang,[1] M. Thys,[7] L. Vandaele,[7] and V. Fievez[1]*

[1]Laboratory for Animal Nutrition and Animal Product Quality, Department of Animal Sciences and Aquatic Ecology, Faculty of Bioscience Engineering, Ghent University, 9000 Gent, Belgium
[2]State Key Laboratory of Grassland and Agro-Ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China
[3]Department of Grassland and Natural Landscape Sciences, Poznań University of Life Sciences, Dojazd 11 Street, 60-632 Poznań, Poland
[4]Department of Animal Physiology, Biochemistry, and Biostructure, Poznań University of Life Sciences, Wołyńska 35 Street, 60-637 Poznań, Poland
[5]School of Agricultural and Environmental Sciences, Mountains of the Moon University, PO Box 837, Fort Portal, Uganda
[6]Department of Agricultural Production, College of Agricultural and Environmental Sciences, Makerere University, PO Box 7062, Kampala, Uganda
[7]ILVO, Scheldeweg 68, 9090 Melle, Belgium

ABSTRACT

Data on metabolic profiles of blood sampled at d 3, 6, 9, and 21 in lactation from 117 lactations (99 cows) were used for unsupervised k-means clustering. Blood metabolic parameters included β-hydroxybutyrate (BHB), nonesterified fatty acids, glucose, insulin-like growth factor-1 (IGF-1) and insulin. Clustering relied on the average and range of the 5 blood parameters of all 4 sampling days. The clusters were labeled as imbalanced (n = 42) and balanced (n = 72) metabolic status based on the values of the blood parameters. Various random forest models were built to predict the metabolic cluster of cows during early lactation from the milk composition. All the models were evaluated using a leave-group-out cross-validation, meaning data from a single cow were always present in either train or test data to avoid any data leakage. Features were either milk fatty acids (MFA) determined by gas chromatography (MFA [GC]) or features that could be determined during a routine dairy herd improvement (DHI) analysis, such as concentration of fat, protein, lactose, fat/protein ratio, urea, and somatic cell count (determined and reported routinely in DHI registrations), either or not in combination with MFA and BHB determined by mid-infrared (MIR), denoted as MFA [MIR] and BHB [MIR], respectively, which are routinely analyzed but not routinely reported in DHI registrations yet. Models solely based on fat, protein, lactose, fat/protein ratio, urea and somatic cell count (i.e., DHI model) were characterized by the lowest predictive performance [area under the receiver operating characteristic curve (AUC_{ROC}) = 0.69]. The combination of the features of the DHI model with BHB [MIR] and MFA [MIR] powerfully increased the predictive performance (AUC_{ROC} = 0.81). The model based on the detailed MFA profile determined by GC analysis did not outperform (AUC_{ROC} = 0.81) the model using the DHI-features in combination with BHB [MIR] and MFA [MIR]. Predictions solely based on samples at d 3 were characterized by lower performance (AUC_{ROC} DHI + BHB [MIR] + MFA [MIR] model at d 3: 0.75; AUC_{ROC} MFA [GC] model at d 3: 0.73). High predictive performance was found using samples from d 9 and 21. To conclude, overall, the DHI + BHB [MIR] + MFA [MIR] model allowed to predict metabolic status during early lactation. Accordingly, these parameters show potential for routine prediction of metabolic status.
Key words: metabolic status, milk composition, predictive modeling, dairy cattle

INTRODUCTION

During early lactation cows commonly encounter negative energy balance resulting in mobilization of body reserves as nonesterified fatty acids (**NEFA**; Drackley, 1999). An excessive mobilization of adipose tissue is associated with metabolic diseases such as ketosis (Ospina et al., 2010). Measuring the concentration of NEFA and BHB in blood is considered as the gold standard to monitor the metabolic health status (Oetzel, 2004; LeBlanc, 2010). When body reserves are being mobilized, NEFA concentration in the blood will increase. If this NEFA concentration exceeds the oxidizing ability of the liver, ketone bodies (e.g., BHB) are formed as a result of incomplete oxidation (LeB-

Received April 22, 2022.
Accepted August 9, 2022.
*Corresponding author: Veerle.Fievez@UGent.be

lanc, 2010). In addition to BHB and NEFA, additional biochemical blood indices, such as glucose, insulin, and IGF-1 hormones, are important when evaluating success or failure of homeorhetic adaptation in early lactation (Lucy, 2008). When single blood metabolites, such as BHB or NEFA, are being used for health monitoring, specific threshold values can be used. However, concomitant use of multiple biochemical blood indices prevents this threshold approach. Hence, unsupervised learning through cluster analysis recently got wide attention to combine multiple blood biomarkers holistically (Tremblay et al., 2018; De Koster et al., 2019; Xu et al., 2019; Foldager et al., 2020). As such, cows are clustered according to their metabolic health status based on blood sampling and determination of biomarkers. Afterward, this grouping based on cluster analysis can be targeted by predictive models. This is in particular of interest because under practical farming conditions routine blood sampling for laboratory analysis is labor intensive, invasive, and expensive, hence limiting the application to individual animal cases or as herd monitoring tool. Therefore, predicting metabolic health status from milk parameters is of major interest. Such studies particularly used blood BHB or NEFA concentrations to diagnose metabolic health (e.g., van der Drift et al., 2012; Jorjong et al., 2014; Mann et al., 2016), whereas only few more recent studies included a more extended set of reference blood parameters (e.g., De Koster et al., 2019; Xu et al., 2019). However, those studies do not address the effect of early milk sampling, which particularly should be assessed in light of early warning of metabolic disorders. The prediction models developed by Xu et al. (2019) included on-farm cow data, such as dry period length, parity, milk production traits, and BW, whereas the models by De Koster et al. (2019) relied on different sets of milk biomarkers, including milk mid-infrared (**MIR**) wavelengths, metabolites, and glycans. Although milk fatty acids (**MFA**) are associated with negative energy balance and metabolic status (Jorjong et al., 2014; Mann et al., 2016), they were not included as predictors in the studies of De Koster et al. (2019) and Xu et al. (2019). Despite their physiological correlation with metabolic status and hence their potential to improve predictions, gas-chromatographic analysis of MFA hampers their inclusion in practical models. An accurate quantification of some (categories of) MFA, routinely determined using MIR spectra (Soyeurt et al., 2011), could eliminate this analytical obstacle. However, MIR-based fatty acid (**FA**) quantification does not cover all FA, which could be determined by GC and is still under development. Hence, it is of interest to compare the model performances of MFA [GC]- and MFA [MIR]-based models.

Obviously, an accurate diagnosis of metabolic health is paramount in this kind of studies. Estimation of the incidence (rather than prevalence) of metabolically impaired health requires repeated sampling and analysis of the reference blood parameters, particularly during the first 2 wk of lactation, which is considered the main risk period (Benedet et al., 2019). However, the study by De Koster et al. (2019) is solely based on one blood sample during this critical period.

Currently, the frequency of milk sampling for DHI is too low and not adapted to measuring during the critical period of metabolic problems. Targeting early lactating cows and a modified sampling frequency could solve this issue. In line with this the effect of days in milk at sampling on the predictive performance of the milk-based models should be assessed. Hence, the current research particularly focused on this early lactation period and responded to the current scientific gaps by inclusion of (1) repeated blood sampling as well as (2) repeated milk samplings to study potential differences in predictive performance based on sampling time.

In this regard, the hypothesis of this research was that biomarkers in milk can be used to predict the metabolic status in early lactation, but predictive performance depends on the sampling day. Also, we hypothesized predictions using MIR-based features have similar predictive performance compared with GC features. Accordingly, the objective of this research was to build and compare different classes of models to predict metabolic status based on cluster analysis during early lactation using milk composition. Features used in the predictive models are either MFA determined by GC or routinely determinable features during a DHI analysis.

MATERIALS AND METHODS

Animals and Management

The experiment was performed at the research farm of ILVO (Melle, Belgium) from October 2018 until October 2020. The experiment (2018/329) was approved by the Ethics Committee of ILVO. Only multiparous Holstein-Friesian cows (n = 120 lactations) were monitored in this study. After the experimental phase, the data from 3 cows were excluded because of missing data or death during the experiment, resulting in 117 lactations. These lactations contained data from 99 unique cows, from which 18 cows were monitored during 2 lactations. Cows were studied until d 23 in lactation.

Dry cows and lactating cows were housed separately in a naturally ventilated freestall barn with slatted floor. Stocking density was always <1 cow/cubicle. From imminent calving (e.g., pelvic ligament relax-

ation, teat filling) until d 3 after calving, cows were housed in maternity pens with straw bedding within the same building. In case of severe disease (reported cases which required intervention by veterinarian or farm staff), cows stayed for a longer period in these pens. During the trial 24 cows were diagnosed as suffering from one or multiple of the following clinical health issues: displaced abomasum (n = 10), hypocalcemia (n = 9), clinical ketosis (n = 9), uterine infection (n = 6), mastitis (n = 5), and other (n = 6).

From 3 wk before calving, cows received the partial mixed ration of the lactating cows supplemented with a dry cow mineral premix and on average 1 kg of balanced concentrate per cow per day. Belgian-Dutch energy and protein evaluation systems were used: protein digestible at the level of the small intestine requirements and supply were assessed according to the DVE-system (Van Duinkerken et al., 2011) and net energy requirements and supply according to the VEM-system (Van Es, 1975). The partial mixed ration of the lactating cows was calculated for an average adult cow of 650 kg, producing 26 kg of fat- and protein-corrected milk and was based on maize silage, prewilted grass silage, pressed beet pulp, and soybean meal to balance digestible protein (**DVE**)- and net energy (**VEM**)-requirements and supplemented with a VEM-DVE-balanced concentrate. The supply of the balanced concentrate changed according to lactation stage and slightly changed during the course of the experiment (running over 2 yr) in relation to the quality and feed values of the used silages. Concentrate intake started at 1.7 kg of balanced concentrates, 0.2 kg formaldehyde-treated soybean meal (Covasoy Braz., FeedValid), 0.3 kg of soybean meal at d 3 after calving. Formaldehyde-treated soybean meal was increased over a period of 7 d to 1 kg, whereas the balanced concentrate was increased to 6 kg over a period of 20 d. Detailed information about the chemical composition and concentrate increment over time is given in Supplemental Tables S1 and S2 (https://doi.org/10.6084/m9.figshare.19626474; Heirbaut et al., 2022). Individual feed intake was monitored throughout the trial using roughage intake control feeding bins (Insentec, Hokofarm Group), except during the period around calving. During lactation, concentrate intake was monitored at the automatic concentrate providers (Greenfeed, C-Lock Inc.; DeLaval) and in the herringbone milking parlor (DeLaval). The cows had access to water ad libitum.

Blood and Milk Samples

Blood sampling took place at d 3 (mean ± SD; 3.1 ± 0.28), 6 (5.9 ± 0.55), 9 (9.1 ± 0.49), and 21 (21.0 ± 0.70) after calving using an 18-gauge needle and Venoject system (BD Diagnostics). Samples were taken from the coccygeal vessels (d 3, 6, and 9) or the jugular vein (d 21) in the morning between 0915 and 0945 h (around 1.5 h after feeding). Blood was collected in plasma NaF (4 mL) and serum blood tubes (10 mL; SSTTM II Advance tubes; BD Diagnostics). Serum tubes were kept at room temperature for 30 min before centrifuging and NaF tubes were kept cooled until centrifuging. The NaF tubes were centrifuged at 1,000 × g for 10 min and serum tubes at 1,500 × g for 15 min at room temperature. Serum and plasma samples were divided into aliquots and stored at −20°C [analyses serum: BHB, NEFA, and insulin; analyses plasma (NaF): glucose] or −80°C (analysis serum: IGF-1). Glucose, BHB, and NEFA were analyzed using a Gallery Discrete Analyzer (ThermoFisher Scientific). The BHB was analyzed using kit RB1007 (Randox Laboratories Ltd.), NEFA using kit FA115 (Randox Laboratories Ltd.), and glucose with kit 981779 (ThermoFisher Scientific) by the laboratory of Dierengezondheidszorg Vlaanderen (Belgium). In brief, the determination of BHB concentration is based on the oxidation of D-3-hydroxybutyrate to acetoacetate by the enzyme 3-hydroxybutyrate dehydrogenase. During this reaction NAD^+ is reduced to NADH and the associated change of absorbance, measured at 340 nm, can be directly correlated with the D-3-hydroxybutyrate concentration. The analysis of NEFA is based on the reaction with co-enzym A and ATP, in a reaction catalyzed by acyl CoA synthetase. The formed acyl CoA is further oxidized in the presence of acyl CoA oxidase, leading to the formation of hydrogen peroxide. In the presence of 4-aminoantipyrine and TOOS [N-ethyl-N-(2hydroxy-3-sulphopropyl)m-toluidine], the hydrogen peroxide reacts, resulting in a purple color, of which the absorbance is read at 540 nm. Analysis of glucose is based on a phosphorylation by ATP, in a reaction catalyzed by hexokinase. The glucose-6-phosphate formed is oxidized to 6-phosphogluconate by glucose-6-phosphate dehydrogenase. In this same reaction an equimolar amount of NAD is reduced to NADH, with a resulting increase in absorbance at 340 nm. Serum IGF-I was analyzed by an RIA method using the nonextraction IGF-1 IRMA DSL-2800 (LifeSpan Biosciences) by Poznań University of Life Sciences. The concentration of this hormone was determined with isotope J125 by using the automatic gamma radiation reader (Wizard2 2-Detector Gamma Counter, Perkin Elmer). In this method the peptide being determined is "sandwiched" between 2 antibodies. The tubes were coated by the first antibody. The second one [anti-IGF-1 (J-125) reagent] was radiolabeled for detection.

The analytic peptide (IGF-1) present in blood serum, standards and controls were bound by both of the antibodies to form a "sandwich" complex. Unbound compounds were removed by decanting. The Mercodia bovine insulin ELISA kit (Bio-connect Diagnostics) was used for insulin analysis.

Dairy cows were milked twice daily, starting at 0530 h and 1630 h in a 2 × 7 herringbone milking parlor, and their milk yield (kg/d) was recorded electronically. To determine milk performance, milk samples (27 mL) were collected from the cows in a representative way during the morning milking, at daily basis from d 3 to 23 after calving. Samples were stored at 4°C and contained preservatives (sodium azide, maximum concentration 0.02% m/m and bronopol, maximum concentration 0.005% m/m). Milk fat, protein, lactose, urea, BHB, SCC, SFA, UFA, MUFA, and total C18:1 were determined by Qlip laboratory (Zutphen, the Netherlands), which performed routine DHI analysis by means of Fourier transform infrared spectrometry (Milkoscan FT6000, Foss Electric). Milk fat, protein, lactose, and urea were determined according to ISO 9622:2013 (ISO, 2013). The SCC was analyzed according to ISO 13366–2:2006 (ISO, 2006). Milk BHB and FA were estimated based on the MIR spectra from in-house established equations. Previously MFA predictions based on similar versions of those in-house models by Qlip have also been used in van Gastelen and Dijkstra (2016) and Tremblay et al. (2019). In Supplemental Table S3 (https://doi.org/10.6084/m9.figshare.19626474; Heirbaut et al., 2022), a comparison between the FA classed predicted by MIR and analyzed by GC is reported for the current data set.

In addition to the analysis done by Qlip, a second sample, containing the same preservatives as the sample for DHI and stored at −20°C, was analyzed by GC at d 3, 6, 9, and 21 to determine the detailed MFA composition as described by Jing et al. (2018). Quantification of FA methyl esters was based on the conversion of peak areas to the mass proportion of FA by a theoretical response factor for each FA (Ackman and Sipos, 1964; Wolff et al., 1995).

Data Processing

Data were processed using R (R Core Team, 2020), version 4.0.2 and RStudio, version 1.3.959. Data reading, manipulation, exploration, tabular reporting, and visualization were done with the packages readxl (v1.3.1; Wickham and Bryan, 2019), tidyverse (v1.3.1; Wickham et al., 2019), ggplot2 (v3.3.5; Wickham, 2016), Ggally (v1.5.0; Schloerke et al., 2020), skimr (v2.1.3; Waring et al., 2021), data.table (v1.14.0; Dowle and Srinivasan, 2021), and flextable (v0.6.6; Gohel, 2021).

Clustering

K-means clustering was used to cluster cows based on blood BHB, NEFA, glucose, insulin, and IGF-1. Due to the outlier sensitivity of the k-means clustering algorithm a winsorizing strategy was used for IGF-1, using a fixed upper threshold of 250 ng/mL (96.5 percentile). Insulin concentrations below the detection limit were imputed by the minimum concentration in the data set. The k-means clustering was performed using the algorithm of Hartigan and Wong (1979). Before clustering, all variables [20 variables: 5 blood parameters × 4 sampling points (d 3, 6, 9, and 21)] were transformed by taking the natural logarithm to adjust skewed distributions. Rather than using the daily blood values for clustering, the mean and range across the different sampling days were determined for each blood parameter to reduce the dimensionality (to 10 parameters) but to keep the information about the within-animal variation. Finally, the clustering parameters were standardized (mean zero, unit variance), because the Euclidean distance is sensitive to the scale and units of the variables included. As such, each cow was allocated to a single cluster based on the 10 blood parameters. Cows monitored during 2 lactations were clustered separately for each lactation. The number of clusters was based on the elbow method, a visual inspection of the total within sum of the squares as a function of the number of clusters, using the package factoextra (v1.0.7; Kassambara and Mundt, 2020). Clustering was done using the base R k-means function (R Core Team, 2020) with 100 random start centers and a maximum number of iterations of 100. Afterward, clustering stability was assessed using a bootstrap resampling scheme with 100 iterations and by calculating the Jaccard similarities of the original clusters to the clusters based on the bootstrapped data (package fpc; v2.2–5; Hennig, 2020). During clustering all lactations were considered as independent. Two different clusters were found, postclustering denoted as clusters metabolically "balanced" and "imbalanced" cluster.

Mixed Effect Modeling

Linear mixed effect models using the package lme4 (v1.1-26; Bates et al., 2015) were used to evaluate the effect of cluster on the different milk parameters. Cow was defined as a random effect, resulting in a compound symmetry structure of the model residuals. The days of sampling, cluster, and parity were used as fixed effects and forced into the models. The inclusion of the 2-way interaction cluster × sampling day was evaluated based on Akaike information criterion. The *P*-values were obtained via Satterthwaite's degrees of freedom method

using the package lmerTest (v3.1-2; Kuznetsova et al., 2017). Packages emmeans (v1.4.7; Lenth et al., 2020) and multcomp (v1.4-13; Hothorn et al., 2008) were used for Tukey post hoc tests and calculating least squares means. Normality of model residuals was checked visually using Q-Q plots constructed using package car (v3.0-10; Fox et al., 2019). In case the assumption of normality was violated, variables were square root transformed (BHB, C12:1). Reported parameters were expressed as least squares means ± standard error and were back-transformed values if necessary.

Predictive Modeling

To compare the predictive value of the different milk biomarkers, different classes of models were built using data from d 3, 6, 9, and 21 (n = 464 observations). These models were trained to determine whether a cow is considered imbalanced or balanced based on individual milk samples (i.e., d3, 6, 9, or 21). Reference classification of the balanced and imbalanced group was based on the k-means clustering procedure, which relied on the average and range of the 5 blood parameters of all 4 sampling days as explained before. The following classes of predictive models were constructed: (1) MFA [MIR] (MFA based on MIR: saturated, unsaturated, monounsaturated MFA and total C18:1); (2) DHI (fat, protein, fat/protein ratio, lactose, urea, SCC); (3) DHI + BHB [MIR] (features of the DHI model + BHB [MIR]); (4) DHI + MFA [MIR] (features of the DHI model + MFA [MIR]); (5) DHI + BHB [MIR] + MFA [MIR] (features of the DHI model + BHB [MIR] + MFA [MIR]); (6) MFA [GC] model (all 62 MFA based on GC analysis, as well as the saturated, unsaturated, poly-unsaturated, de novo, total C18:1, and odd- and branched-chain FA and the ratio *cis*-9 C18:1/C15:1 were included). All models also included parity information and day of sampling. No variable selection step was performed: all features were kept in the models.

Random forest models were constructed using the packages caret (v6.0–84; Kuhn et al., 2021) and ranger (v0.13.1; Wright and Ziegler, 2017). To increase computational speed the models were trained using a local parallel backend using base R parallel package (R Core Team, 2020).

Two modeling approaches were followed. The first modeling approach aimed at comparing the performance for the different types of models. For this purpose, a cross-validation approach with a default parameter tuning was performed for all models. There was no separate test set, because the number of available datapoints for the MFA [GC] model was limited (d 3, 6, 9, and 21). Models were evaluated using 10-fold cross-validation, 15 repeats. Folds were stratified using total $C18:1_{MIR}$ concentration in the milk to ensure a physiologically representative data distribution for model training. Number of trees was set fixed at 500. Minimal node size was kept constant at a value of 1. Typically, random forests do not account for whether different observations are repeated measures, increasing the risk of data leakage (i.e., information from the training data set leaking to the test set and increasing the predictive performance). To avoid such issues, all data from a single cow (repeated samplings within one lactation, but also multiple lactations from the same cow) were blocked either in the training or in the test data. As such, the model was always evaluated on data from cows which were not present during the training process. Model performance was calculated global across all sampling days as well as for each sampling day separately. For each model receiver operating characteristic (**ROC**) and precision recall (**PR**) curves were constructed and the AUC was calculated to evaluate the model performance using the ROCR package (v1.0–11; Sing et al., 2005).

The second modeling approach was a more extended modeling for the DHI + BHB [MIR] + MFA [MIR] model, combined with the use of a separate test set (from the same experiment) to evaluate the model performance in a more robust way. MIR-data were available on a daily basis in the period d 3 till 23 (n = 2,312). Those daily data were all used as individual data records to train the MIR-based models. Again, data from a single cow were always included either in the training or in the test data to avoid any data leakage. This modeling approach concerned an extended parameter tuning on train data using adaptive resampling of the tuning parameter grid and a total of 100 different tuning combinations. Number of trees was set fixed at 500. Adaptive resampling was performed to tune: (1) number of variables randomly sampled as candidates at each split, (2) splitting rule (gini, extratrees and hellinger), and (3) minimal node size. Finally, the best performing model was selected and evaluated on a separate test. This approach permitted to train the best model and evaluate it on independent test data. The test set contained 20% of the balanced cows and 20% of the imbalanced cows.

RESULTS

Summary Descriptive Data

The cows enrolled in the study had a median parity of 3 (range 2–7). The mean daily milk yield (d 3 to 23) was 36.7 ± 6.91 kg (mean ± SD), with an average fat content of 5.1 ± 1.33 g/100 g of milk and an average

Table 1. Summary of the blood parameters of the metabolically imbalanced (IMB) and balanced (BAL) clusters at d 3, 6, 9, and 21 in lactation (LSM ± SE); *P*-values for the fixed effects of cluster and DIM, and the possible interaction cluster × DIM, are reported as well[1]

DIM	Cluster	BHB[2] (mmol/L)	NEFA[2] (mmol/L)	Glucose (mmol/L)	Insulin[2] (ng/mL)	IGF-1[2] (ng/mL)
3	IMB	1.02 ± 0.059*	0.71 ± 0.052	3.08 ± 0.076*	0.13 ± 0.013	71.4 ± 5.46*
	BAL	0.75 ± 0.033	0.45 ± 0.027	3.31 ± 0.057	0.22 ± 0.019	106.7 ± 6.11
6	IMB	1.25 ± 0.072*	0.74 ± 0.053	3.02 ± 0.076	0.12 ± 0.012	62.2 ± 4.77*
	BAL	0.81 ± 0.035	0.46 ± 0.028	3.12 ± 0.057	0.21 ± 0.018	100.1 ± 5.74
9	IMB	1.57 ± 0.091*	0.76 ± 0.055	2.60 ± 0.076*	0.11 ± 0.011	69.4 ± 4.94*
	BAL	0.78 ± 0.034	0.48 ± 0.028	3.08 ± 0.057	0.20 ± 0.017	89.0 ± 5.10
21	IMB	1.58 ± 0.091*	0.74 ± 0.053	2.88 ± 0.076*	0.16 ± 0.016	57.1 ± 4.37*
	BAL	0.79 ± 0.034	0.46 ± 0.028	3.23 ± 0.057	0.27 ± 0.023	97.0 ± 5.56
Fixed effects						
Cluster		$P < 0.001$	$P < 0.001$	$P < 0.001$	$P < 0.001$	$P < 0.001$
DIM		$P < 0.001$	$P = 0.788$	$P < 0.001$	$P < 0.001$	$P < 0.001$
Cluster × DIM		$P < 0.001$	—	$P = 0.014$	—	$P = 0.017$

[1]Day of blood sampling and cluster were forced into the models. Two-way interaction was evaluated and omitted if $P > 0.10$. In case of significant interaction, a Tukey post hoc test was performed.

[2]Variables were ln-transformed.

*Significant differences ($P < 0.05$) at each DIM, which were tested when cluster × DIM interaction was significant.

protein content of 3.6 ± 0.50 g/100 g of milk. The average daily total DMI was 20.5 ± 4.18 kg.

In Supplemental Tables S4 and S5 (https://doi.org/10.6084/m9.figshare.19626474; Heirbaut et al., 2022) the descriptive information about the blood values and clinical health disorders are given respectively.

Cluster Analysis

K-means clustering resulted in 2 clusters with a high cluster stability (Jaccard indices after resampling with bootstrap scheme: 0.81 and 0.89). Based on the values of the blood parameters, the clusters were labeled as imbalanced (n = 42) and balanced (n = 75). An overview of the blood parameters and clusters is given in Table 1. The imbalanced cluster was characterized by high BHB and NEFA and low glucose, insulin, and IGF-1 concentrations. At d 6, 9, and 21 the average BHB concentration exceeded 1.20 mmol/L in the imbalanced cluster, whereas the average BHB did not exceed this threshold at any of the time points in the balanced cluster. The imbalanced cluster was characterized by 31, 50, 64, and 69% observations with a BHB concentration above 1.20 mmol/L at d 3, 6, 9, and 21 respectively, whereas for the balanced cluster it was 0, 0, 4, and 1% respectively. On all postpartum sampling days, the average NEFA concentration was higher than 0.60 mmol/L in the imbalanced cluster and not in the balanced cluster. The imbalanced cluster was characterized by 69, 71, 67, and 64% observations with a NEFA concentration above 0.60 mmol/L at d 3, 6, 9, and 21 respectively, whereas for the balanced cluster this was 32, 45, 31, and 31% respectively. A chi-squared contingency table test did not show differences in parity distribution (defined as parity 2, 3, and >3) across the clusters ($P = 0.13$). The imbalanced and balanced cluster contained 7 versus 3 cases of displaced abomasum, 6 versus 3 cases of hypocalcemia, 6 versus 1 case(s) of clinical ketosis (based on clinical observations without considering blood BHB concentrations), 4 versus 1 case(s) of uterine infection, 1 versus 3 case(s) of mastitis and 3 versus 3 cases of other diseases/health problems (data are represented as disease cases, multiple disease cases for the same cow are possible), respectively. Generally, cows in the metabolically imbalanced cluster had a 2.97 higher relative risk (95% CI: 1.43–6.21) to develop one of the mentioned health problems.

Predictive Models at Days 3, 6, 9, and 21

In Figures 1 and 2 the ROC and PR curves for the different models predicting cows in the metabolically imbalanced cluster are reported. The DHI model was characterized by the lowest AUC_{ROC} (0.69) and AUC_{PR} (0.53). The inclusion of BHB or MFA [MIR] resulted in important predictive improvement (DHI + BHB [MIR] model: AUC_{ROC} 0.79 and AUC_{PR} 0.69; DHI + MFA [MIR] model: AUC_{ROC} 0.77 and AUC_{PR} 0.65). Combining BHB [MIR] as well as MFA [MIR] with DHI resulted in additional predictive gain (AUC_{ROC} 0.81 and AUC_{PR} 0.72). A model based on MFA determined by GC (AUC_{ROC} 0.81 and AUC_{PR} 0.70) performed better compared with the MFA [MIR] model (AUC_{ROC} 0.75 and AUC_{PR} 0.62), but did not outperform the model combining DHI, BHB [MIR] and MFA [MIR]. In the latter model the 4 most important features were: BHB [MIR] and the FA classes (predicted from MIR spectra): saturated, monounsaturated, and total C18:1 (Figure 3). In Figure 4, the 20 most important features of the GC model are reported. The 5 most important

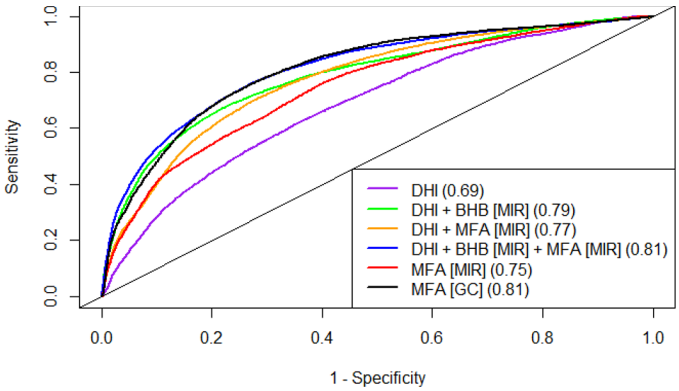

Figure 1. Receiver operating characteristic (ROC) curves and area under the ROC curve (values presented between brackets in the legend) of different random forest prediction models predicting metabolically imbalanced cows using milk composition. K-means clustering, based on averages and ranges of the 4 blood samplings, was used to determine metabolic imbalance, whereas milk models to predict imbalanced cows relied on features of individual milk samples (i.e., d 3, 6, 9, or 21). Results are based on a 10-fold leave-group-out cross-validation with 15 repeats, and include all the hyperparameters investigated during the training process. Models are trained and evaluated based on data from d 3, 6, 9, and 21 (n = 464 observations). The following prediction models were evaluated: (1) DHI-based features (fat, protein, fat/protein ratio, lactose, urea, SCC); (2) DHI and BHB based on mid-infrared (MIR) spectra (DHI + BHB [MIR]); (3) DHI and milk fatty acids (MFA) based on MIR spectra (DHI + MFA [MIR]); (4) DHI, BHB [MIR] and MFA based on MIR spectra (DHI + BHB + MFA [MIR]); (5) MFA based on MIR spectra (MFA [MIR]); and (6) MFA based on GC (MFA [GC]).

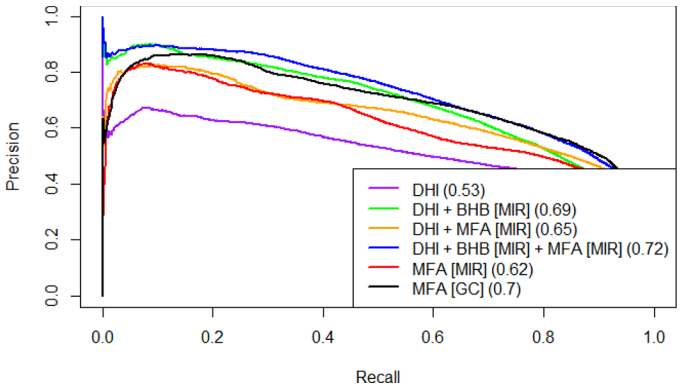

Figure 2. Precision recall (PR) curves and area under the PR curve (values presented between brackets in the legend) of different random forest prediction models predicting metabolically imbalanced cows using milk composition. K-means clustering, based on averages and ranges of the 4 blood samplings, was used to determine metabolic imbalance, whereas milk models to predict imbalanced cows relied on features of individual milk samples (i.e., either d 3, 6, 9, or 21). Results are based on a 10-fold leave-group-out cross-validation with 15 repeats, and include all the hyperparameters investigated during the training process. Models are trained and evaluated based on data from d 3, 6, 9, and 21 (n = 464 observations). The following prediction models were evaluated: (1) DHI-based features (fat, protein, fat/protein ratio, lactose, urea, SCC); (2) DHI and BHB based on mid-infrared (MIR) spectra (DHI + BHB [MIR]); (3) DHI and milk fatty acids (MFA) based on MIR spectra (DHI + MFA [MIR]); (4) DHI, BHB [MIR] and MFA based on MIR spectra (DHI + BHB + MFA [MIR]); (5) MFA based on MIR spectra (MFA [MIR]); and (6) MFA based on GC (MFA [GC]).

predictors were C11:0, C12:0, C14:0, C15:0, and C10:0. The relative variable importance of the ratio *cis*-9 C18:1/C15:0 and C4:0 also exceeded 50.

In Supplemental Table S6 (https://doi.org/10.6084/m9.figshare.19626474; Heirbaut et al., 2022) the milk composition and MFA are compared between metabolically imbalanced and balanced cows. The milk BHB [MIR] concentration was higher in metabolically imbalanced cows at d 6, 9, and 21 (Tukey post hoc test $P < 0.05$), but not at d 3 in lactation. Regarding FA determined by MIR, cluster had a significant effect on all the FA classes ($P < 0.001$), and there were no interaction effects. Concentrations of UFA, MUFA, and total C18:1 were found in higher concentrations in metabolically imbalanced cows. The concentration of SFA was higher in the balanced cluster. In order of variable importance in the GC model, the 5 most important variables all showed a significant cluster effect (P-value cluster <0.001). The interaction term cluster × DIM was kept in all statistical models of these 5 MFA [GC] (except for C14:0), but all sampling points differed when performing Tukey post hoc tests, resulting in higher concentrations of C11:0, C14:0, C12:0, C10:0, and C15:0 in the balanced cluster across all sampling days.

As opposed to the overall model performance in Figures 1 and 2, Table 2 shows the variation in model performance (cross-validated AUC and AUC_{PR}) across the different sampling days. Generally, the predictive models of d 3 were characterized by a lower AUC_{PR} which never exceeded 0.60. The DHI + BHB [MIR] + MFA [MIR] had the best predictions at d 9 and 21. AUC_{PR} values above 0.80 were found for the DHI + BHB [MIR] + MFA [MIR] model at d 9 and 21 and the GC model at d 21.

DHI + BHB [MIR] + MFA [MIR] Model, Day 3 to 23

Based on the performance of the DHI + BHB [MIR] + MFA [MIR] model using data from d 3, 6, 9, and 21 only, a second modeling with separate test set and adaptive resampling was performed using the full data set in the range d 3 to 23. Across the 10 cross-validation runs, 15 repeats and all different hyperparameter runs, the ROC curve had an AUC_{ROC} of 0.79 and the PR curve an AUC_{PR} of 0.69 on training data. The predictive performance for each day separately is given in Supplemental Table S7 (https://doi.org/10.6084/m9.figshare.19626474; Heirbaut et al., 2022). Refitting the

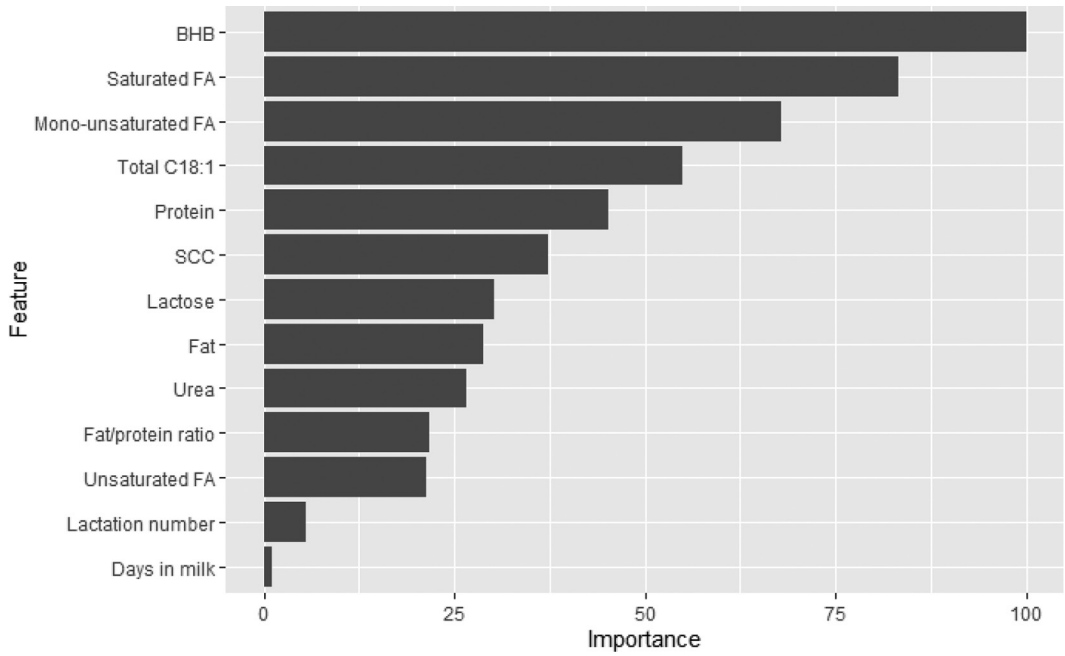

Figure 3. Scaled variable importance of the features in the DHI + BHB [MIR]+ MFA [MIR] model. MIR = mid-infrared; MFA = milk fatty acids.

best model on training data resulted in an AUC_{ROC} of 0.91 and AUC_{PR} of 0.87. When evaluating the model on test data, the performance slightly decreased to an AUC_{ROC} of 0.90 and an AUC_{PR} of 0.82.

DISCUSSION

Classifying dairy cows during early lactation according to their metabolic status is often challenging and debatable. Traditionally, the metabolic status is assessed by using threshold values for BHB or NEFA (LeBlanc, 2010). However, as indicated by De Koster et al. (2019) and Xu et al. (2019), other parameters without previously determined thresholds (e.g., the hormone IGF-1) may also be important to assess the energy balance and metabolic status in early lactation. These blood parameters then can be combined in an unsupervised learning approach to cluster cows according to their metabolic health status (Tremblay et al., 2018; De Koster et al., 2019; Xu et al., 2019). In the current research, the metabolic clusters were based on 5 parameters (BHB, NEFA, glucose, insulin, and IGF-1), which also have been used by De Koster et al. (2019) (except for insulin) and Xu et al. (2019).

Before performing a clustering, it is crucial to assess and interpret the variation in the data set, because the clustering algorithm is solely a mathematical approach which groups observations based on a specific distance metric. Supplemental Table S4 shows there is substantial variation in the different blood parameters. Depending on the sampling day, up to 26 and 55% of the cows exceed the 1.2 or 0.6 mmol/L threshold for BHB and NEFA, respectively. Based on this substantial variation, the data set is well suited for applying a clustering, which afterward was also compared by the stable results when performing a bootstrap resampling.

In our study 36% of the cows were in the imbalanced cluster, which is higher than in other studies (De Koster et al., 2019; Xu et al., 2019). Interherd differences might be at the origin of this difference (Benedet et al., 2019), although the repeated blood sampling in the current study also may have contributed. Indeed, this approach aimed to estimate the incidence of metabolically imbalanced cows rather than the prevalence based on a single or limited number of blood samples. In our study 37% of the cows had an elevated blood BHB concentration (BHB >1.2 mmol/L) at least once. Only 6.0% of the cows had an elevated blood BHB across all 4 samplings, whereas only 16.2% of the cows were characterized by one single elevated blood BHB event. This variation addresses the importance of repeated samplings (Mann et al., 2016; Benedet et al., 2019).

Despite these methodological differences across studies (Carrier et al., 2004; Oetzel, 2004; Brunner et al., 2019), all report an unequal distribution of metabolically balanced versus imbalanced cows within a herd. Also in our study, this unequal ratio (i.e., 36% of the cows were in the imbalanced cluster) has been observed.

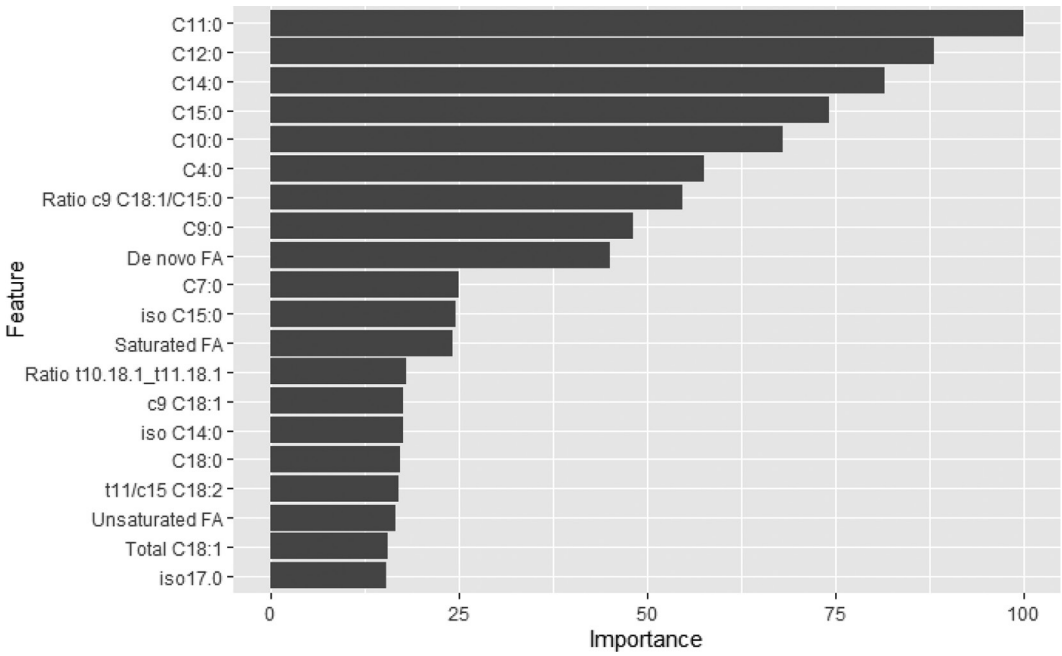

Figure 4. Scaled variable importance of the 20 most important features in the MFA [GC] model. MFA = milk fatty acids.

In order for the model to be trained on data with sufficient variation, stratification of total C18:1 was performed during cross-validation. This ensures a more accurate sampling compared with a random sampling. Standard predictive models often fail when predicting minority classes (i.e., imbalanced cows). However, this potentially low predictive performance of the minority class is not reflected in the widely used accuracy metric (i.e., proportion of correct predictions), because the accuracy of a model that predicts only the majority class equals the proportion of the majority class (He and Garcia, 2009). This challenges the construction of high-performance models predicting the minority class and tempts to focus on the prediction of the majority class (e.g., cows in good or average metabolic status; Xu et al., 2019). Obviously, this does not contribute to a better identification of metabolically imbalanced cows, whereas the (early) identification of metabolically imbalanced cows is of particular importance under practical farming conditions. De Koster et al. (2019) built prediction models, addressing metabolically balanced as well as metabolically imbalanced cows. Unfortunately, only the accuracy metric has been used, which challenges direct (numeric) comparisons with our results. Moreover, the value of this metric is highly questionable with skewed class distributions as indicated before (He and Garcia, 2009). Surprisingly, De Koster et al. (2019) found highest prediction accuracy for prediction of imbalanced cows. The best prediction model had an accuracy of 81% for imbalanced cows; however, the number of imbalanced cows was 19 out of 107, which means a naive model predicting all cows as balanced would already achieve an accuracy of 82%. This questions the precision of the prediction. The precision itself or data to calculate precision were not reported. However, Grelet et al. (2019) [study using data of the same experiments as reported by De Koster et al. (2019)], but also extended by data from heifers and additional herds) reported the confusion matrix of a partial least squares discriminant analysis, allowing a more detailed comparison: global accuracy was 87% and sensitivity 76%, but the precision was only 56%. In our study a sensitivity of 76% corresponded to a precision of 61%. If a higher sensitivity would be pursued, a precision of 56% would result in a sensitivity of 81%. The values in our study are somewhat higher, but still in a similar range, despite the differences in sampling days and the inclusion of heifers in the study of Grelet et al. (2019). Low precision has been observed in the study of Mann et al. (2016), using a *cis*-9 C18:1 to C15:0 ratio of $\geq$45.0 [as proposed by Jorjong et al. (2015)] for detection of hyperketonemia, which resulted in a sensitivity, specificity, and accuracy of respectively 84.0, 63.2, and 69.5%. However, the precision was only 50.0%. Accordingly, only half of the cows predicted with elevated blood NEFA concentrations effectively experienced this metabolic status, whereas the other half were incorrectly alerted. As discussed by He and Garcia (2009) the use of alternative metrics such as precision could be beneficial in situations with an un-

Table 2. Area under the receiver operating characteristic curve (AUC_{ROC}) and area under the precision recall curve (AUC_{PR}) for d 3, 6, 9, and 21 in lactation of different random forest models predicting metabolically imbalanced cows using milk composition[1]

Model features[2]	AUC_{ROC}	AUC_{PR}	DIM
DHI	0.70	0.53	3
DHI + BHB [MIR]	0.70	0.53	3
DHI + MFA [MIR]	0.75	0.58	3
DHI + BHB [MIR]+ MFA [MIR]	0.75	0.58	3
MFA [MIR]	0.72	0.57	3
MFA [GC]	0.73	0.56	3
DHI	0.67	0.54	6
DHI + BHB [MIR]	0.79	0.75	6
DHI + MFA [MIR]	0.76	0.61	6
DHI + BHB [MIR]+ MFA [MIR]	0.81	0.69	6
MFA [MIR]	0.75	0.59	6
MFA [GC]	0.79	0.66	6
DHI	0.66	0.52	9
DHI + BHB [MIR]	0.83	0.75	9
DHI + MFA [MIR]	0.79	0.71	9
DHI + BHB [MIR]+ MFA [MIR]	0.85	0.81	9
MFA [MIR]	0.78	0.71	9
MFA [GC]	0.84	0.77	9
DHI	0.71	0.54	21
DHI + BHB [MIR]	0.81	0.74	21
DHI + MFA [MIR]	0.80	0.72	21
DHI + BHB [MIR]+ MFA [MIR]	0.85	0.80	21
MFA [MIR]	0.74	0.64	21
MFA [GC]	0.88	0.83	21

[1]K-means clustering, based on averages and ranges of the 4 blood samplings, was used to determine metabolic imbalance, whereas milk models to predict imbalanced cows relied on features of individual milk samples (i.e., either d 3, 6, 9, or 21).

[2]The following prediction models were evaluated: (1) DHI-based features (fat, protein, fat/protein ratio, lactose, urea, SCC); (2) DHI and milk BHB based on mid-infrared (MIR) spectra (DHI + BHB); (3) DHI and milk fatty acids (MFA) based on MIR spectra (DHI + MFA [MIR]); (4) DHI, BHB, and MFA based on MIR spectra (DHI + BHB + MFA [MIR]); (5) MFA based on MIR spectra (MFA [MIR]); and (6) MFA based on GC (MFA [GC]). Results are based on a 10-fold leave-group-out cross-validation with 15 repeats, and include different model hyperparameters using a limited grid search of 6 combinations (i.e., number of variables sampled as candidates at each split and splitting rule) investigated during the training process. Models are trained and evaluated based on data from d 3, 6, 9, and 21 (n = 464 observations).

equal class distribution. Indeed, from a practical point of view, not only the proportion of imbalanced cows identified by a model (i.e., sensitivity) is of importance, but also the precision [i.e., true positives/(true positives + false positives)]. Obviously, reliable and practically applicable predictions require a limited number of false positives. It has been shown that a model which generates a (too) high number of positives can impair the correct follow-up by dairy farmers (Eckelkamp and Bewley, 2020). From this point of view external model validation is very important before models could be implemented in practice. The models in our study were solely evaluated on data from a single farm and hence require further validation.

The MFA [GC] model highlighted the predictive power of some minor FA, such as C11:0 and C15:0. Given their low prediction accuracy by MIR, these MFA were not included in the DHI + BHB [MIR] + MFA [MIR] model. Indeed, accurate FA predictions based on MIR are generally lacking for FA at low concentrations (Rutten et al., 2009; Soyeurt et al., 2011). Rutten et al. (2009) suggested a minimum FA concentration of 2.45 g/100 g of FA for obtaining a R^2 of minimum 0.6 or higher. Nevertheless, despite their physiological relevance (Vlaeminck et al., 2006; Dorea et al., 2017) and high variable importance in the MFA [GC] model, the inclusion of these MFA did not seem necessary to accurately predict metabolic status by the DHI + BHB [MIR] + MFA [MIR] model. As opposed to the study of De Koster et al. (2019) the DHI + BHB [MIR] + MFA [MIR] model did not use the raw MIR spectra, but the FA classes predicted based on MIR. Although this could be considered an unnecessary intermediate step, it should be realized that the use of MIR-data in predictive models require dimension reduction. As such the use of MIR-based FA predictions could be considered as a preprocessing step to reduce the number of parameters, which may be of particular interest because the number of observations is often limited in this type of animal studies with transition animals and relying on blood-based reference metabolites.

Because DIM and often also parity affect the measured milk parameters, the construction of a model seems more appropriate than using single threshold values to capture these additional factors of variation. Regarding early warning predictions, further research focusing on the construction of specific models targeting this period and possible incorporation of additional data sources could be of interest (e.g., sensor data, data of dry period), because the predictions of the model performed at d 3 in lactation were generally characterized by lower predictive performance compared with the other days. De Koster et al. (2019) and Xu et al. (2019) did not focus on this early period. Based on our results early prediction of metabolic status using milk data should not be (solely) based on samples from d 3 in lactation. The lower BHB concentrations in the beginning of the lactation could be at the origin of this lower performance, only 11.1% of the samples was characterized by an elevated BHB concentration at d 3. In line with these results, Mann et al. (2016) showed FA in colostrum could not be used to predict elevated BHB and NEFA concentrations during the first 2 wk of lactation. Therefore, they recommended to target the period between 3 and 14 DIM for further studies. Important features in our models generally agree with the relevant milk biomarkers reported in literature. In line with Mann et al. (2016), several de novo synthesized FA (in order of importance: C14:0, C12:0, and C10:0) showed high variable importance.

Indeed, imbalanced cows were characterized by lower de novo FA concentrations. As discussed in Craninx et al. (2008) increased concentrations of long-chain MFA as a result of body mobilization concomitantly result in lower concentration of de novo MFA. In addition, C15:0 and the ratio *cis*-9 C18:1/C15:0 were identified as important predictors (cf. Jorjong et al., 2015; Mann et al., 2016). Interestingly, in our study, the odd short-chain FA C11:0 had the highest importance in the MFA [GC] model. In line with our results, Dórea et al. (2017) found high sensitivity for odd short-chain FA (e.g., C11:0, to detect elevated NEFA concentrations). C11:0, such as other odd-chain FA, requires propionyl-CoA as precursor (Fievez et al., 2012). A higher availability of propionate in the rumen might therefore have resulted in the higher odd-chain FA observed in the balanced cluster.

CONCLUSIONS

Both the DHI + BHB [MIR] + MFA [MIR] model and the MFA [GC] model accurately predict metabolic status during early lactation. As such, GC analysis is not required, which extends the opportunity to routinely screen the metabolic status. Nevertheless, the predictive performance varies considerably across sampling days, with models solely based on data from d 3 in milk resulting in low reliability. Considering this trade-off between predictive performance and early sampling, predictions could be performed on data from d 6 onwards. Therefore, the use of early milk samplings should be further investigated together with the possible incorporation of additional data sources to improve early warning predictions. Finally, before application in practice, additional data from different farms should be added and evaluated to make the model robust to differences in management and rations.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of Flanders Innovation & Entrepreneurship (VLAIO; Brussels, Belgium; LA170830). The research of Xiaoping Jing and Mingqi Zhang was supported by the Chinese Scholarship Council (CSC, China); the PhD research of Stijn Heirbaut was funded by a PhD grant from the Special Research Fund of the Ghent University (Bijzonder Onderzoeksfonds, BOF, Belgium). The authors gratefully acknowledge the staff from the Animal Science Unit, Flanders Research Institute for Agriculture, Fisheries and Food for their help during this experiment. We also thank Charlotte Melis (Laboratory for Animal Nutrition and Animal Product Quality, Department of Animal Sciences and Aquatic Ecology, Ghent University, Gent, Belgium) and Ilke van Hese [Animal Sciences Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Scheldeweg, Melle, 9090, Belgium; Department of Reproduction, Obstetrics and Herd Health Faculty of Veterinary Medicine, Ghent University, Salisburylaan, Merelbeke, 9820, Belgium] for assistance during sampling. The authors have not stated any conflicts of interest.

REFERENCES

Ackman, R., and J. Sipos. 1964. Application of specific response factors in the gas chromatographic analysis of methyl esters of fatty acids with flame ionization detectors. J. Am. Oil Chem. Soc. 41:377–378. https://doi.org/10.1007/BF02654818.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67:1–48. https://doi.org/10.18637/jss.v067.i01.

Benedet, A., C. L. Manuelian, A. Zidi, M. Penasa, and M. De Marchi. 2019. Invited review: β-hydroxybutyrate concentration in blood and milk and its associations with cow performance. Animal 13:1676–1689. https://doi.org/10.1017/S175173111900034X.

Brunner, N., S. Groeger, J. Canelas Raposo, R. M. Bruckmaier, and J. J. Gross. 2019. Prevalence of subclinical ketosis and production diseases in dairy cows in Central and South America, Africa, Asia, Australia, New Zealand, and Eastern Europe. Transl. Anim. Sci. 3:84–92. https://doi.org/10.1093/tas/txy102.

Carrier, J., S. Stewart, S. Godden, J. Fetrow, and P. Rapnicki. 2004. Evaluation and use of three cowside tests for detection of subclinical ketosis in early postpartum cows. J. Dairy Sci. 87:3725–3735. https://doi.org/10.3168/jds.S0022-0302(04)73511-0.

Craninx, M., A. Steen, H. Van Laar, T. Van Nespen, J. Martín-Tereso, B. de Baets, and V. Fievez. 2008. Effect of lactation stage on the odd- and branched-chain milk fatty acids of dairy cattle under grazing and indoor conditions. J. Dairy Sci. 91:2662–2677. https://doi.org/10.3168/jds.2007-0656.

De Koster, J., M. Salavati, C. Grelet, M. A. Crowe, E. Matthews, R. O'Flaherty, G. Opsomer, L. Foldager, and M. Hostens. 2019. Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers. J. Dairy Sci. 102:2631–2644. https://doi.org/10.3168/jds.2018-15533.

Dórea, J. R. R., E. A. French, and L. E. Armentano. 2017. Use of milk fatty acids to estimate plasma nonesterified fatty acid concentrations as an indicator of animal energy balance. J. Dairy Sci. 100:6164–6176. https://doi.org/10.3168/jds.2016-12466.

Dowle, M., and A. Srinivasan. 2021. data.table: Extension of "data.frame." R package version 1.14.0. Accessed Oct. 26, 2022. https://CRAN.R-project.org/package=data.table.

Drackley, J. K. 1999. ADSA Foundation Scholar Award: Biology of dairy cows during the transition period: The final frontier? J. Dairy Sci. 82:2259–2273. https://doi.org/10.3168/jds.S0022-0302(99)75474-3.\

Eckelkamp, E. A., and J. M. Bewley. 2020. On-farm use of disease alerts generated by precision dairy technology. J. Dairy Sci. 103:1566–1582. https://doi.org/10.3168/jds.2019-16888.

Fievez, V., E. Colman, J. M. Castro-Montoya, I. Stefanov, and B. Vlaeminck. 2012. Milk odd- and branched-chain fatty acids as biomarkers of rumen function—An update. Anim. Feed Sci. Technol. 172:51–65. https://doi.org/10.1016/j.anifeedsci.2011.12.008.

Foldager, L., C. Gaillard, M. T. Sorensen, T. Larsen, E. Matthews, R. O'Flaherty, F. Carter, M. A. Crowe, C. Grelet, M. Salavati, M. Hostens, K. L. Ingvartsen, and M. A. Krogh. 2020. Predicting physiological imbalance in Holstein dairy cows by three different sets of milk biomarkers. Prev. Vet. Med. 179:105006. https://doi.org/10.1016/j.prevetmed.2020.105006.

Fox, J., S. Weisberg, B. Price, D. Adler, D. Bates, G. Baud-Bovy, B. Bolker, S. Ellison, D. Firth, M. Friendly, G. Gorjanc, S. Graves, R. Heiberger, P. Krivitsky, R. Laboissiere, M. Maechler, G. Monette,

D. Murdoch, H. Nilsson, D. Ogle, B. Ripley, W. Venables, S. Walker, D. Winsemius, A. Zeileis, and R-Core. 2019. Car: Companion to applied regression. Accessed Oct. 26, 2022. https://cran.r-project.org/web/packages/car/index.html.

Gohel, D. 2021. Flextable: Functions for tabular reporting. R package version 0.6.6. Accessed Oct. 26, 2022. https://CRAN.R-project.org/package=flextable.

Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K. L. Ingvartsen, M. Crowe, M. T. Sorensen, E. Froidmont, C. P. Ferris, C. Marchitelli, F. Becker, T. Larsen, F. Carter, and E. Gplus. Consortium, and F. Dehareng. 2019. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. Animal 13:649–658. https://doi.org/10.1017/S1751731118001751.

Hartigan, J. A., and M. A. Wong. 1979. Algorithm AS 136: A k-means clustering algorithm. Appl. Stat. 28:100–108. https://doi.org/10.2307/2346830.

He, H., and E. A. Garcia. 2009. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 21:1263–1284. https://doi.org/10.1109/TKDE.2008.239.

Heirbaut, S., J. Xiaoping, B. Stefanska, E. Pruszyńska-Oszmałek, L. Buysse, P. Lutakome, M. Zhang, M. Thys, L. Vandaele, and V. Fievez. Supplemental tables and figures. Figshare. https://doi.org/10.6084/m9.figshare.19626474.

Hennig, C. 2020. Fpc: Flexible procedures for clustering. Accessed Oct. 26, 2022. https://CRAN.R-project.org/package=fpc.

Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. Biom. J. 50:346–363. https://doi.org/10.1002/bimj.200810425.

ISO. (International Organization for Standardization). 2013. Milk and liquid milk products—Guidelines for the application of mid-infrared spectrometry. ISO 9622:2013. ISO. https://www.iso.org/standard/56874.html.

ISO. (International Organization for Standardization). 2006. Milk—Enumeration of somatic cells—Part 2: Guidance on the operation of fluoro-opto-electronic counters. ISO 13366-2:2006. ISO. https://www.iso.org/standard/40260.html.

Jing, L., L. Dewanckele, B. Vlaeminck, W. M. Van Straalen, A. Koopmans, and V. Fievez. 2018. Susceptibility of dairy cows to subacute ruminal acidosis is reflected in milk fatty acid proportions, with C18:1 trans-10 as primary and C15:0 and C18:1 trans-11 as secondary indicators. J. Dairy Sci. 101:9827–9840. https://doi.org/10.3168/jds.2018-14903.

Jorjong, S., A. T. M. van Knegsel, J. Verwaeren, R. M. Bruckmaier, B. De Baets, B. Kemp, and V. Fievez. 2015. Milk fatty acids as possible biomarkers to diagnose hyperketonemia in early lactation. J. Dairy Sci. 98:5211–5221. https://doi.org/10.3168/jds.2014-8728.

Jorjong, S., A. T. M. van Knegsel, J. Verwaeren, M. V. Lahoz, R. M. Bruckmaier, B. De Baets, B. Kemp, and V. Fievez. 2014. Milk fatty acids as possible biomarkers to early diagnose elevated concentrations of blood plasma nonesterified fatty acids in dairy cows. J. Dairy Sci. 97:7054–7064. https://doi.org/10.3168/jds.2014-8039.

Kassambara, A., and F. Mundt. 2020. Factoextra: Extract and visualize the results of multivariate data analyses. Accessed Oct. 26, 2022. https://CRAN.R-project.org/package=factoextra.

Kuhn, M., J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, Z. Mayer, B. Kenkel, R Core Team, M. Benesty, R. Lescarbeau, A. Ziem, L. Scrucca, Y. Tang, C. Candan, and T. Hunt. 2021. Caret: Classification and regression training. Accessed Oct. 26, 2022. https://CRAN.R-project.org/package=caret.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. LmerTest: Tests in linear mixed effects models. J. Stat. Softw. 82:1–26. https://doi.org/10.18637/jss.v082.i13.

LeBlanc, S. 2010. Monitoring metabolic health of dairy cattle in the transition period. J. Reprod. Dev. 56(Suppl.):S29–S35. https://doi.org/10.1262/jrd.1056S29.

Lenth, R. V., P. Buerkner, M. Herve, J. Love, F. Miguez, H. Riebl, and H. Singmann. 2020. Emmeans: Estimated marginal means, aka least-squares means. Accessed Oct. 26, 2022. https://cran.r-project.org/package=emmeans.

Lucy, M. C. 2008. Functional differences in the growth hormone and insulin-like growth factor axis in cattle and pigs: Implications for post-partum nutrition and reproduction. Reprod. Domest. Anim. 43:31–39. https://doi.org/10.1111/j.1439-0531.2008.01140.x.

Mann, S., D. V. Nydam, A. L. Lock, T. R. Overton, and J. A. A. McArt. 2016. Short communication: Association of milk fatty acids with early lactation hyperketonemia and elevated concentration of nonesterified fatty acids. J. Dairy Sci. 99:5851–5857. https://doi.org/10.3168/jds.2016-10920.

Oetzel, G. R. 2004. Monitoring and testing dairy herds for metabolic disease. Vet. Clin. North Am. Food Anim. Pract. 20:651–674. https://doi.org/10.1016/j.cvfa.2004.06.006.

Ospina, P. A., D. V. Nydam, T. Stokol, and T. R. Overton. 2010. Evaluation of nonesterified fatty acids and β-hydroxybutyrate in transition dairy cattle in the northeastern United States: Critical thresholds for prediction of clinical diseases. J. Dairy Sci. 93:546–554. https://doi.org/10.3168/jds.2009-2277.

R Core Team. 2020. R: The R project for statistical computing. Accessed Oct. 26, 2022. https://www.R-project.org/.

Rutten, M. J. M., H. Bovenhuis, K. A. Hettinga, H. J. F. van Valenberg, and J. A. M. 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. https://doi.org/10.3168/jds.2009-2456.

Schloerke, B., D. Cook, J. Larmarange, F. Briatte, M. Marbach, E. Thoen, A. Elberg, O. Toomet, J. Crowley, H. Hofmann, and H. Wickham. 2020. GGally: Extension to "Ggplot2." Accessed Oct. 26, 2022. https://CRAN.R-project.org/package=GGally.

Sing, T., O. Sander, N. Beerenwinkel, and T. Lengauer. 2005. ROCR: Visualizing the classifier performance in R. Bioinformatics 21:3940–3941. https://doi.org/10.1093/bioinformatics/bti623.

Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. J. Dairy Sci. 94:1657–1667. https://doi.org/10.3168/jds.2010-3408.

Tremblay, M., M. Kammer, H. Lange, S. Plattner, C. Baumgartner, J. A. Stegeman, J. Duda, R. Mansfeld, and D. Döpfer. 2018. Identifying poor metabolic adaptation during early lactation in dairy cows using cluster analysis. J. Dairy Sci. 101:7311–7321. https://doi.org/10.3168/jds.2017-13582.

Tremblay, M., M. Kammer, H. Lange, S. Plattner, C. Baumgartner, J. A. Stegeman, J. Duda, R. Mansfeld, and D. Döpfer. 2019. Prediction model optimization using full model selection with regression trees demonstrated with FTIR data from bovine milk. Prev. Vet. Med. 163:14–23. https://doi.org/10.1016/j.prevetmed.2018.12.012.

van der Drift, S. G. A., R. Jorritsma, J. T. Schonewille, H. M. Knijn, and J. A. 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. https://doi.org/10.3168/jds.2011-4417.\

Van Duinkerken, G., M. C. Blok, A. Bannink, J. W. Cone, J. Dijkstra, A. M. Van Vuuren, and S. Tamminga. 2011. Update of the Dutch protein evaluation system for ruminants: The DVE/OEB_{2010} system. J. Agric. Sci. 149:351–367. https://doi.org/10.1017/S0021859610000912.

Van Es, A. J. H. 1975. Feed evaluation for dairy cows. Livest. Prod. Sci. 2:95–107. https://doi.org/10.1016/0301-6226(75)90029-9.

van Gastelen, S., and J. Dijkstra. 2016. Prediction of methane emission from lactating dairy cows using milk fatty acids and mid-infrared spectroscopy. J. Sci. Food Agric. 96:3963–3968. https://doi.org/10.1002/jsfa.7718.

Vlaeminck, B., V. Fievez, S. Tamminga, R. J. Dewhurst, A. van Vuuren, D. De Brabander, and D. Demeyer. 2006. Milk Odd- and branched-chain fatty acids in relation to the rumen fermentation pattern. J. Dairy Sci. 89:3954–3964. https://doi.org/10.3168/jds.S0022-0302(06)72437-7.

Waring, E., M. Quinn, A. McNamara, E. A. de la Rubia, H. Zhu, J. Lowndes, S. Ellis, H. McLeod, H. Wickham, K. Müller, C. Kirkpatrick, S. Brenstuhl, and P. Schratz. lbusett, M. Korpela,

J. Thompson, H. McGehee, M. Roepke, P. Kennedy, D. Possenriede, D. Zimmermann, K. Butts, B. Torges, and R. Saporta. 2021. Skimr: Compact and flexible summaries of data. Accessed on Oct. 26, 2022. https://CRAN.R-project.org/package=skimr.

Wickham, H. 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the Tidyverse. J. Open Source Softw. 4:1686. https://doi.org/10.21105/joss.01686.

Wickham, H., and J. Bryan. 2019. Readxl: Read Excel files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl.

Wolff, R. L., C. C. Bayard, and R. J. Fabien. 1995. Evaluation of sequential methods for the determination of butterfat fatty acid composition with emphasis ontrans-18:1 acids. Application to the study of seasonal variations in French butters. J. Am. Oil Chem. Soc. 72:1471–1483. https://doi.org/10.1007/BF02577840.

Wright, M. N., and A. Ziegler. 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J. Stat. Softw. 77:1–17. https://doi.org/10.18637/jss.v077.i01.

Xu, W., A. T. M. van Knegsel, J. J. M. Vervoort, R. M. Bruckmaier, R. J. van Hoeij, B. Kemp, and E. Saccenti. 2019. Prediction of metabolic status of dairy cows in early lactation with on-farm cow data and machine learning algorithms. J. Dairy Sci. 102:10186–10201. https://doi.org/10.3168/jds.2018-15791.