J. Dairy Sci. 107:476–494
https://doi.org/10.3168/jds.2023-23693

Combination of milk variables and on-farm data as an improved diagnostic tool for metabolic status evaluation in dairy cattle during the transition period

S. Heirbaut,[1] X. P. Jing,[1,2] B. Stefańska,[3] E. Pruszyńska-Oszmałek,[4] B. Ampe,[5] C. Umstätter,[6,7] L. Vandaele,[5] 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 Ghent, Belgium
[2]State Key Laboratory of Grassland and Agro-Ecosystems, International Centre for Tibetan Plateau Ecosystem Management, School of Life Sciences, Lanzhou University, Lanzhou 730000, China
[3]Department of Grassland and Natural Landscape Sciences, Poznań University of Life Sciences, 60-632 Poznań, Poland
[4]Department of Animal Physiology, Biochemistry, and Biostructure, Poznań University of Life Sciences, 60-637 Poznań, Poland
[5]Animal Science Unit, ILVO, 9090 Melle, Belgium
[6]Thünen Institute of Agricultural Technology, Thünen Institute, DE-38116 Braunschweig, Germany
[7]Automatisierung und Arbeitsgestaltung, Agroscope, 8356 Ettenhausen, Switzerland

ABSTRACT

Milk composition, particularly milk fatty acids, has been extensively studied as an indicator of the metabolic status of dairy cows during early lactation. In addition to milk biomarkers, on-farm sensor data also hold potential in providing insights into the metabolic health status of cows. While numerous studies have explored the collection of a wide range of sensor data from cows, the combination of milk biomarkers and on-farm sensor data remains relatively underexplored. Therefore, this study aims to identify associations between metabolic blood variables, milk variables, and various on-farm sensor data. Second, it seeks to examine the supplementary or substitutive potential of these data sources. Therefore, data from 85 lactations on metabolic status and on-farm data were collected during 3 wk before calving up to 5 wk after calving. Blood samples were taken on d 3, 6, 9, and 21 after calving for determination of β-hydroxybutyrate (BHB), nonesterified fatty acids (NEFA), glucose, insulin-like growth factor-1 (IGF-1), insulin, and fructosamine. Milk samples were taken during the first 3 wk in lactation and analyzed by mid-infrared for fat, protein, lactose, urea, milk fatty acids, and BHB. Walking activity, feed intake, and body condition score (BCS) were monitored throughout the study. Linear mixed effect models were used to study the association between blood variables and (1) milk variables (i.e., milk models); (2) on-farm data (i.e., on-farm models) consisting of activity and dry matter intake analyzed during the dry period ([D]) and lactation ([L]) and BCS only analyzed during the dry period ([D]); and (3) the combination of both. In addition, to assess whether milk variables can clarify unexplained variation from the on-farm model and vice versa, Pearson marginal residuals from the milk and on-farm models were extracted and related to the on-farm and milk variables, respectively. The milk models had higher coefficient of determination (R^2) than the on-farm models, except for IGF-1 and fructosamine. The highest marginal R^2 values were found for BHB, glucose, and NEFA (0.508, 0.427, and 0.303 vs. 0.468, 0.358, and 0.225 for the milk models and on-farm models, respectively). Combining milk and on-farm data particularly increased R^2 values of models assessing blood BHB, glucose, and NEFA concentrations with the fixed effects of the milk and on-farm variables mutually having marginal R^2 values of 0.608, 0.566, and 0.327, respectively. Milk C18:1 was confirmed as an important milk variable in all models, but particularly for blood NEFA prediction. On-farm data were considerably more capable of describing the IGF-1 concentration than milk data (marginal R^2 of 0.192 vs. 0.086), mainly due to dry matter intake before calving. The BCS [D] was the most important on-farm variable in relation to blood BHB and NEFA and could explain additional variation in blood BHB concentration compared with models solely based on milk variables. This study has shown that on-farm data combined with milk data can provide additional information concerning the metabolic health status of dairy cows. On-farm data are of interest to be further studied in predictive modeling, particularly because early warning predictions using milk data are highly challenging or even missing.

Key words: metabolic imbalance, milk biomarkers, sensor data, transition period

Received May 12, 2023.
Accepted August 13, 2023.
*Corresponding author: veerle.fievez@ugent.be

INTRODUCTION

Blood sampling is often considered as the reference test to determine the metabolic status during early lactation, based on nonesterified fatty acids (**NEFA**) or BHB concentrations (Oetzel, 2004; LeBlanc, 2010). However, it is obvious that blood sampling has its limitations in practice, which is illustrated by the broad range of literature on alternative biomarkers such as ketone bodies in urine and milk. Milk composition, in particular milk fatty acids, have been extensively investigated to monitor and provide further insight on the metabolic status of dairy cows during early lactation (Jorjong et al., 2015; Pires et al., 2022). Some milk fatty acids can be routinely determined by mid-infrared (**MIR**) spectra instead of an extensive laboratory analysis by GC (Rutten et al., 2009; Soyeurt et al., 2011; Maurice-Van Eijndhoven et al., 2013), which could facilitate implementation in practice. Heirbaut et al. (2023a) used MIR-predicted milk biomarkers to assess metabolic imbalance during early lactation with similar predictive performance as models based on milk fatty acids determined by GC. These MIR-predicted biomarkers contained milk BHB and fatty acid classes such as total milk C18:1. However, colostrum (Mann et al., 2016) or even early milk samplings at d 3 of lactation (Heirbaut et al., 2023a) were not suitable to evaluate the metabolic health status.

In addition to milk biomarkers, on-farm sensor data also could be informative regarding metabolic health status, and combine easy, practical on-farm implementation with potentially earlier decisive management information (e.g., data from the dry period). For example, Menichetti et al. (2020) found that the prepartum blood NEFA concentration was positively associated with the coefficient of variation of the lying time within 7 d before the blood sampling. Indeed, activity and rumination sensor data showed associations with or predictive value for metabolic health or resilience; for example, rumination time is used as an indicator of (subclinical) health status (Soriani et al., 2012; Calamari et al., 2014), including early warning around dry-off (Abuelo et al., 2021). Prepartum activity measured by accelerometers could be informative in relation to postpartum disease, although still with a high number of false positives (Belaid et al., 2021). Finally, sensor data also have been used to model (lifetime) resilience variables (e.g., early culling; van Dixhoorn et al., 2018; Adriaens et al., 2020; Ouweltjes et al., 2021).

Although all those studies showed the potential of collecting a wide range of cow sensor data, to our knowledge, few studies combined milk biomarkers with on-farm sensor data. Studying sensor data and milk biomarkers together in relation to metabolic blood variables could help to understand whether these data sources provide supplementary (i.e., making an addition; Merriam-Webster, 2022) or substitutive information (i.e., replacement; Merriam-Webster, 2022). The study of Xu et al. (2019) combined BW during lactation with milk variables, including protein, fat, lactose content, and SCC, but milk fatty acids or milk BHB were not included in their study and the milk biomarkers had the highest importance in the prediction. Mäntysaari et al. (2019) modeled blood NEFA concentration as a function of milk variables, BW, and BCS and the latter 2 variables were selected in the model having the best fit. The set of sensor data (i.e., BW and BCS), as well as blood variables (i.e., NEFA only), were limited in this study. Therefore, conducting further investigations into the associations between blood variables, milk variables, and on-farm sensor data is essential. Compared with milk biomarkers, on-farm sensor data have the potential advantage of providing prepartum information on the postpartum metabolic health status.

Therefore, the objective of this research is to (1) identify associations between metabolic blood variables, milk variables, and various on-farm sensor data and (2) study their supplementary or substitutive potential for explaining variation in various metabolic blood variables. It is hypothesized that the combination of both data sources will be a better proxy for assessing the concentrations of the metabolic blood variables than the single data sources. This research was set up as an initial screening in which predictive models were not developed because more data would be required to avoid the risk of overfitting.

MATERIALS AND METHODS

Animals, Samplings, and Sensor Measurements

The experiment (2018/329) was conducted at the research farm of ILVO (Flanders Research Institute for Agriculture, Fisheries, and Food, Melle, Belgium). It took place from October 2018 until October 2020 and was approved by the Ethical Committee of ILVO (2018/329). Detailed data about the animals, housing, and diets were previously described in Heirbaut et al. (2023a). In our previous study (Heirbaut et al., 2023a), a machine learning model for diagnosing metabolic impaired health status using routine MIR-based milk composition was developed and compared with a machine learning model using milk fatty acids determined via an extensive GC method. Based on different metabolic blood variables the cows were grouped into metabolically balanced and imbalanced cows via a k-means clustering procedure. The developed machine learning model had similar performance as a machine

learning model using milk fatty acids determined via an extensive GC method. A subset of the 117 multiparous lactations of Holstein-Friesian cows monitored in Heirbaut et al. (2023a) had behavior data available, allowing to identify associations between the different metabolic blood variables, milk variables, and various on-farm sensor data and to study their supplementary or substitutive potential. Accordingly, data have been further filtered, based on the availability of BCS, feed intake, and activity data, resulting in 85 lactations (77 unique cows). These cows had BCS, feed intake, and activity data during the period 3 wk before calving, as well as milk, feed intake, and activity data during the postpartum period until 3 wk after calving. In this study any cows with clinical health disorders were kept in the trial and not removed from the dataset (except in case the disease or death during the trial resulted in too limited data availability: 1 cow died during the first 3 wk, 1 cow died shortly after the 3 wk of trial, and 1 cow was removed a priori due to too many missing blood data). The disease definition was based on the reported cases of mastitis, clinical ketosis, metritis, hypocalcemia, displaced abomasum, or other clinical health problems (e.g., severe lameness), which required intervention by veterinarian or farm staff.

In the cleaned dataset of 85 lactations, there were 14 lactations (14 unique cows) which experienced at least 1 clinical disease during the first 3 wk after calving. In the filtered dataset, no cows died during the first 3 wk of the experiment.

The dry cows and lactating cows were housed separately in a freestall barn with a slatted floor, with a stocking density of less than one cow per cubicle. The maternity pens, which had straw bedding, were located within the same building. Cows stayed in these maternity pens from the time of imminent calving (such as pelvic ligament relaxation and teat filling) until 3 d after calving. If a cow developed a disease, it was kept in the maternity pen for a longer duration. Throughout the study, individual feed intake was tracked using RIC feeding bins (Insentec, Hokofarm Group, Marknesse, the Netherlands), except for the period around calving during which they remained in the maternity pen. During lactation, concentrate intake was monitored using automatic concentrate providers (Greenfeed, C-Lock Inc., Rapid City, US; DeLaval, Tumba, Sweden) and in the herringbone milking parlor (DeLaval). The cows were provided with ad libitum access to water.

Blood samples were collected from the cows according to the procedures previously reported in the study by Heirbaut et al. (2022, 2023a). Specifically, samples were taken in the morning on d 3 (mean ± SD, 3.1 ± 0.32), 6 (5.9 ± 0.55), 9 (9.1 ± 0.56), and 21 (20.9 ± 0.75) after calving. Feed was offered around 1.5 h before blood sampling. Samples were obtained from either the coccygeal vessels (on d 3, 6, and 9) or the jugular vein (on d 21). The collected blood samples were analyzed for BHB, NEFA, IGF-1, insulin, and fructosamine. The IGF-1 concentration was measured using a commercially available Bovine IGF-1 ELISA Kit (catalog number 201–04–0024, standard range 60 to 960 ng/mL, Shanghai Sunred Biological Technology Co. Ltd., China). The detection assay range was from 6.0 to 1,800 ng/mL for IGF-I. The analysis was performed according to the manufacturer's instructions. Samples were analyzed in duplicate and absorbance values were read at 450 nm using a microplate spectrophotometer (Biotek Synergy Neo2, BIOKOM, USA). The intra-assay coefficients of variation were less than 10%. The BHB, NEFA, insulin, and fructosamine analyses are described in Heirbaut et al. (2022, 2023a).

Daily milk samples (27 mL) were collected in a representative manner during the morning (0530 h; 2 h before feeding), from d 3 until 23 postpartum. The samples were analyzed by the laboratory of Qlip (the Netherlands) for milk fat, protein, lactose, urea (ISO, 2013), SCC (ISO, 2006), BHB, SFA, UFA, MUFA, and total milk C18:1 (estimated using Fourier transform infrared spectrometry). More details can be found in Heirbaut et al. (2023a).

Finally, 2 types of sensor measurements were carried out in the study. The first one involved using a BCS camera from DeLaval to determine the cows' BCS. This measurement was taken roughly every 2 wk during the dry period. The BCS camera output data used in this study were retrieved in 0.1 increments. The provided data were obtained as a 7-d moving average. The system has been validated by Mullins et al. (2019) and showed the precision of the scores within the range of 3.0 to 3.75. However, the study of Mullins et al. (2019) revealed underestimation for lower BCS and overestimation for higher BCS. The second type of measurement used IceTag3D motion sensors from IceRobotics (Edinburgh, Scotland) to monitor the cows′ walking activity both during the dry period and lactation. The study of Heirbaut et al. (2022) contains more detailed information on the sensor measurements.

Data Processing

The data analysis was performed using R (version 4.2.2; R Core Team, 2022) and RStudio (version 2022.07.2). The data files were imported, wrangled, explored, and visualized using various packages, including readxl (version 1.3.1, Wickham and Bryan, 2019), tidyverse (version 1.3.1, Wickham et al., 2019), data.table

(version 1.14.0, Dowle and Srinivasan, 2021), ggrepel (version 0.8.2, Slowikowski, 2020), skimr (version 2.1.3, Waring et al., 2021), ggplot2 (version 3.3.5, Wickham, 2016), smplot (version 0.1.0, Min, 2022), scales (version 1.2.0, Wickham and Seidel, 2022), esquisse (version 1.1.0, Meyer and Perrier, 2022), patchwork (version 1.0.1, Pedersen, 2020), and ggthemr (version 1.1.0, Tobin, 2020). Package management was performed using the pacman package (version 0.5.0, Rinker and Kurkiewicz, 2017).

Data Treatment. Several data treatment methods were applied to ensure the quality and reliability of the data. Outliers in the data were identified and handled appropriately. The IGF-1 concentrations were winsorized (fixed upper threshold of 250 ng/mL) and the insulin concentrations below the detection limit were imputed by the minimum concentration (details in Heirbaut et al., 2023a). Milk composition records with milk fat concentrations outside the 1% and 99% quantiles for each sampling day, were removed from analysis. Milk BHB contained one biologically implausible result and hence was winsorized.

IceTag activity data consisted of steps, lying time, number of lying bouts, standing time, and motion index. The motion index was not used for further modeling due to the high correlation with steps. Standing time is directly related to lying time (1 − lying time) and was therefore not included.

Activity data, DMI, and milk variables were aligned based on corresponding sampling days during lactation. For each sampling day x, the average was taken of day x − 1 until day+1 (except milk samples taken on d 3: only average of day x and day x +1). The exact sampling day was afterward defined as a fixed number (d 3, 6, 9, and 21) factor variable for modeling. Before calculating the average, the number of lying bouts was winsorized based on the 95% quantile using the package DescTools (v0.99.38., Signorell, 2020).

For data during the dry period, average BCS was calculated during the period d −21 until d −1 for cows having at least 1 observation during this period. The average number of steps, lying time, and number of lying bouts during the last week before calving was calculated. Afterward the number of lying bouts was winsorized based on the 95% quantile using the package DescTools (v0.99.38., Signorell, 2020). The DMI during the last week before calving was summarized by a linear mixed effect model using package lme4 (v.1.1–31; Bates et al., 2015) with Nelder-Mead optimization of parameters. This linear mixed effect model described the DMI in function of DIM and a random slope and intercept for each cow. For each cow, the random intercept, slope, and, root mean squared error (**RMSE**) were extracted as DMI variables of interest.

Mixed Effect Modeling. Linear mixed effect models were used to study the association between blood variables (BHB, NEFA, glucose, insulin, IGF-1, and fructosamine; measured at d 3, 6, 9, and 21) and (1) milk variables; (2) on-farm data consisting of activity and DMI during the dry and lactation period, and BCS (dry period); and (3) the combination of milk variables and on-farm data. In total 324 observations were used (85 lactations, 4 sampling days). A schematic overview of the data processing is given in Figure 1.

Linear mixed effect models were built, using the package lme4 (v.1.1–31; Bates et al., 2015) and lmerTest (v3.1–3; Kuznetsova et al., 2017). The cow was defined as a random effect. All numeric dependent and independent variables were scaled and centered before entry into the models to (1) avoid any potential convergence issues due to very different scales of predictor variables; (2) increase the interpretability of the estimates. *P*-values were based on Satterthwaite approximation. The performance package (v0.10.0; Lüdecke et al., 2021) was used to calculate variance-inflation factors (VIF; based on the fixed effects model) to check linearity of residuals versus fitted values, the homogeneity of variance, and the normality of model residuals. In case the assumption of normality was violated, variables were transformed using the logarithm with base 10. The SCC, BHB, NEFA, IGF-1, insulin, and fructosamine were $\log_{10}$ transformed to meet the condition of normally distributed residuals. All milk variables, except SFA, UFA, and MUFA (because of correlated variables), were included as variables in the mixed effect models. In addition, the fat and protein content were removed to avoid multicollinearity issues, due to the correlation with the fat/protein ratio. All model terms and biologically relevant interactions were evaluated via backward elimination and omitted if the *P*-value >0.10. Days in milk, health status (binary variable), and parity were forced into the models. The factor health status was a clinical disease definition, based on the reported cases that required intervention by veterinarian or farm staff. This additional factor was used to avoid that a clinical disease state would be a factor of unknown variation (e.g., influencing the DMI, activity). Cows were treated according to the standard procedures of the farm, without any influence on the intervention based on results from this observational study.

The models studying the association between blood variables and milk variables were constructed as given in Equation [1]:

$$Y_{i,j} = \mu + a_1 FP_{i,j} + a_2 Ur_{i,j} + a_3 BHB_{i,j} + a_4 LogSCC_{i,j} + a_5 \text{C18:1}_{i,j} + a_6 DIM_j + a_7 Lac_j + a_8 H_j + U_j + \varepsilon_{i,j}, \quad [1]$$

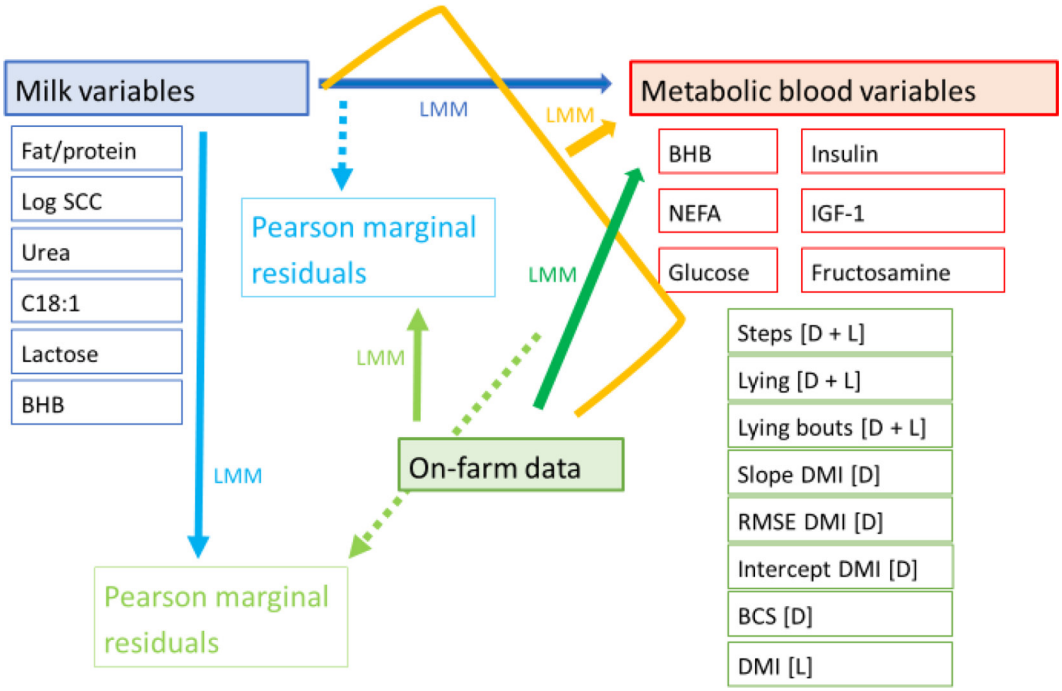

Figure 1. Schematic overview of the modeling approach followed to study the associations between milk data (variables in blue; determined by mid-infrared, except SCC) and on-farm data (variables in green; BCS, activity, and DMI) in relation to metabolic blood variables (variables in red; BHB, nonesterified fatty acids [NEFA], glucose, insulin, IGF-1, and fructosamine) and model residuals by the use of linear mixed models (LMM). Blood samples were collected on d 3, 6, 9, and 21 during lactation, while daily milk samples were obtained from d 3 to 23. Body condition scores were determined using a DeLaval BCS camera, and walking activity was monitored using IceTag3D motion sensors from IceRobotics. The study included 85 lactations from 77 unique cows. The arrows in dark blue and dark green refer to the study of the association between the raw blood variables with the milk and on-farm data, respectively. The dashed arrows refer to the extraction of model residuals from the milk model (blue) and on-farm model (green). The arrows in light blue and light green refer to the study of the association between the extracted residuals from the blood variables and the milk and on-farm data, respectively. The variable annotations "[D]" and "[L]" refer to the dry period and lactation, respectively. Slope DMI [D], RMSE DMI [D], and intercept DMI [D] are the slope, root mean squared error (RMSE), and intercept of a linear mixed effect model studying the DMI during the last week before calving (d −7 to −3) in function of DIM, with a random slope and intercept for each lactation.

where $Y_{i,j}$ represents the blood variable (BHB, NEFA, glucose, insulin, IGF-1 or fructosamine) of cow j at day i. The overall intercept is given by μ. $FP_{i,j}$, $Ur_{i,j}$, $BHB_{i,j}$, $LogSCC_{i,j}$, and C18:1$_{i,j}$ are the fat/protein ratio, urea, milk BHB, $\log_{10}$ SCC, and milk C18:1 concentration of cow j at lactation day i (average day $i-1$ until day $i+1$, except d 3), respectively. DIM_j and Lac_j are DIM (fixed as 3, 6, 9, or 21) and the lactation number (2, 3, >3). H_j is the health status of cow j. U_j is the random intercept, and $\varepsilon_{i,j}$ is the error term. The coefficients are given by a_1 to a_8.

The models studying the association between blood variables and on-farm data were constructed as given in Equation [2]:

$$Y_{i,j} = \mu + a_1 S_{L[i],j} + a_2 L_{L[i],j} + a_3 LB_{L[i],j} + a_4 DMI_{L[i],j} + a_5 S_{D,j} + a_6 L_{D,j} + a_7 LB_{D,j} + a_8 bDMI_{D,j} + a_9 aDMI_{D,j} + a_{10} RMSE\ DMI_{D,j} + a_{11} BCS_{D,j} + a_{12} DIM_j + a_{13} Lac_j + a_{14} H_j + U_j + \varepsilon_{i,j}, \quad [2]$$

where $Y_{i,j}$ represents the blood variable (BHB, NEFA, glucose, insulin, IGF-1, or fructosamine) of cow j at day i. The overall intercept is given by μ. $S_{L[i],j}$, $L_{L[i],j}$, $LB_{L[i],j}$, and $DMI_{[i],j}$ are the numbers of steps, lying time, number of lying bouts, and DMI of cow j at lactation day i (i: average day $i-1$ until day $i+1$) or the dry period (D). $bDMI_{D,j}$, $aDMI_{D,j}$, and $RMSE\ DMI_{D,j}$ are the random intercept, random slope, and root mean squared error of DMI for cow j during the last week before calving. $BCS_{D,j}$ is the average BCS of cow j during the last 3 wk before calving. DIM_j and Lac_j are DIM (fixed as 3, 6, 9, or 21) and the lactation number (2, 3, >3), respectively. H_j is the health status of cow j. U_j is the random intercept, and $\varepsilon_{i,j}$ is the error term. The coefficients are given by a_1 to a_{14}.

A third type of model was constructed as full model, by including all the fixed effects from Equations [1] and [2] in one model. Hence this model combined milk variables and on-farm variables to study the different metabolic blood variables.

Models [1] and [2] were constructed to conclude whether and to which extent the milk variables and on-farm data are associated with metabolic blood variables. To assess whether milk variables and on-farm data could explain unexplained variance in models [2] and [1], Pearson marginal model residuals from Equations [1] and [2] were extracted using the redres package (v0.0.0.9; Goode et al., 2022). Since models [1] and [2] included a random intercept for the cow level, marginal residuals were extracted to only account for fixed effects. Hence, these residual models were again fitted with a random intercept for cow and modeled in function of the explanatory variables in model [2] and [1], respectively.

Statistical reporting was performed using package sjPlot (v2.8.1.; Lüdecke, 2022). Calculation of coefficient of determination ($\mathbf{R^2}$) is a common practice in linear models, but very rare in mixed models since there are different complex ways to calculate due to the random effects. However, R^2 has important value in (biological) models. Hence in this study, marginal and conditional R^2 and partial R^2 were estimated using the Nakagawa and Schielzeth approach (Nakagawa and Schielzeth, 2013) via the r2beta function in the r2glmm package (v0.1.2; Jaeger, 2017). The partial R^2 quantifies the proportion of variation in the dependent variable (response variable) explained uniquely by a specific predictor variable while holding other predictors constant, while the marginal and conditional R^2 refer to the proportion of variation explained by the complete model, respectively excluding and including the random model effects.

Moreover, the RMSE was calculated in 2 ways, based on the model including random effects, as well as excluding the random effects, using the package Metrics (v0.1.4; Hamner and Frasco, 2018).

Sparse Partial Least Squares. Finally, to explore further the relationship between (1) metabolic blood variables and (2) milk and on-farm data, sparse partial least squares (**sPLS**) was performed by using package mixOmics (v6.12.2; Rohart et al., 2017). This method maximizes the covariance between the latent variables and is able to model multiple response variables (i.e., the blood variables). To enhance interpretability sPLS included the LASSO penalization on loading vectors to reduce the number of original variables used when constructing latent variables (Rohart et al., 2017). Variable plots and clustered heatmap of sPLS were constructed to assess the relationship of the variables among the different components.

RESULTS

Summary Descriptive Data

Cows in this study had a median parity of 3 (range 2 to 6). The average recorded milk yield (mean ± SD; d 3 to d 23) was 36.6 ± 6.83 kg/d. Milk samples had an average fat content of 5.1 ± 1.24 g/100 g of milk and an average protein content of 3.6 ± 0.52 g/100 g of milk. The milk samples had a median BHB concentration of 0.09 (range 0.0 to 2.5) mmol/L milk and an average total milk C18:1 concentration of 26.6 ± 5.39 g/100 g of fat. The cows had an average DMI of 14.8 ± 2.60 kg during the period 7 to 3 d before calving. During lactation the average daily DMI was 20.6 ± 3.91 kg. The cows had an average BCS of 3.4 ± 0.24 during the last 3 wk before calving. There were 14 cows in the dataset, which were considered as clinically diseased in the mixed model analysis. There were 5 cases (multiple cases per cow are possible) of hypocalcemia (1 of the cows with hypocalcemia also had a not classified disease), 4 cases of ketosis, 4 of displaced abomasum, 3 with mastitis, and 1 with uterine disorder. There was also 1 cow that had a cesarean section; due to the possible influence on the first postpartum days, this cow was also classified as diseased. Furthermore, there was 1 cow with retained placenta for more than 5 d and 1 cow with trauma. The mean ± standard deviation of the first occurrence of the disease cases was 4.5 ± 7.0 DIM.

Modeling Metabolic Blood Variables Using On-Farm and Milk Variables: Overall Comparison

Linear mixed effect models were constructed to assess BHB (mmol/L), NEFA (mmol/L), glucose (mmol/L), IGF-1 (ng/mL), insulin (ng/mL), and fructosamine (μmol/L) using milk variables and on-farm data. Model coefficients of the milk and on-farm models are reported in Supplemental Figures S1 and S2 (https://doi.org/10.6084/m9.figshare.22730858.v2; Heirbaut et al., 2023b).

Milk models had higher R^2 compared with the on-farm models (Figure 2A), except for IGF-1 and fructosamine. The highest model R^2 values were found for BHB, glucose, and NEFA. For these variables the milk models had a marginal R^2 of 0.508, 0.427, and 0.303, respectively, versus 0.468, 0.358, and 0.225 for the on-farm models. The fixed effects of the milk model explained 25.5% of the variance in the insulin concentration, whereas the on-farm model explained 22.4%. IGF-1 and fructosamine had lower model R^2 values. In total, only 8.6 and 6.4% of the variation in IGF-1 and fructosamine concentration could be explained by the fixed effects of the milk model, whereas the fixed effects of the on-farm model explained 19.2 and 15.5% of the variance in IGF-1 and fructosamine concentration, respectively.

The following parts of the results will focus on the importance of the individual variables in the milk and

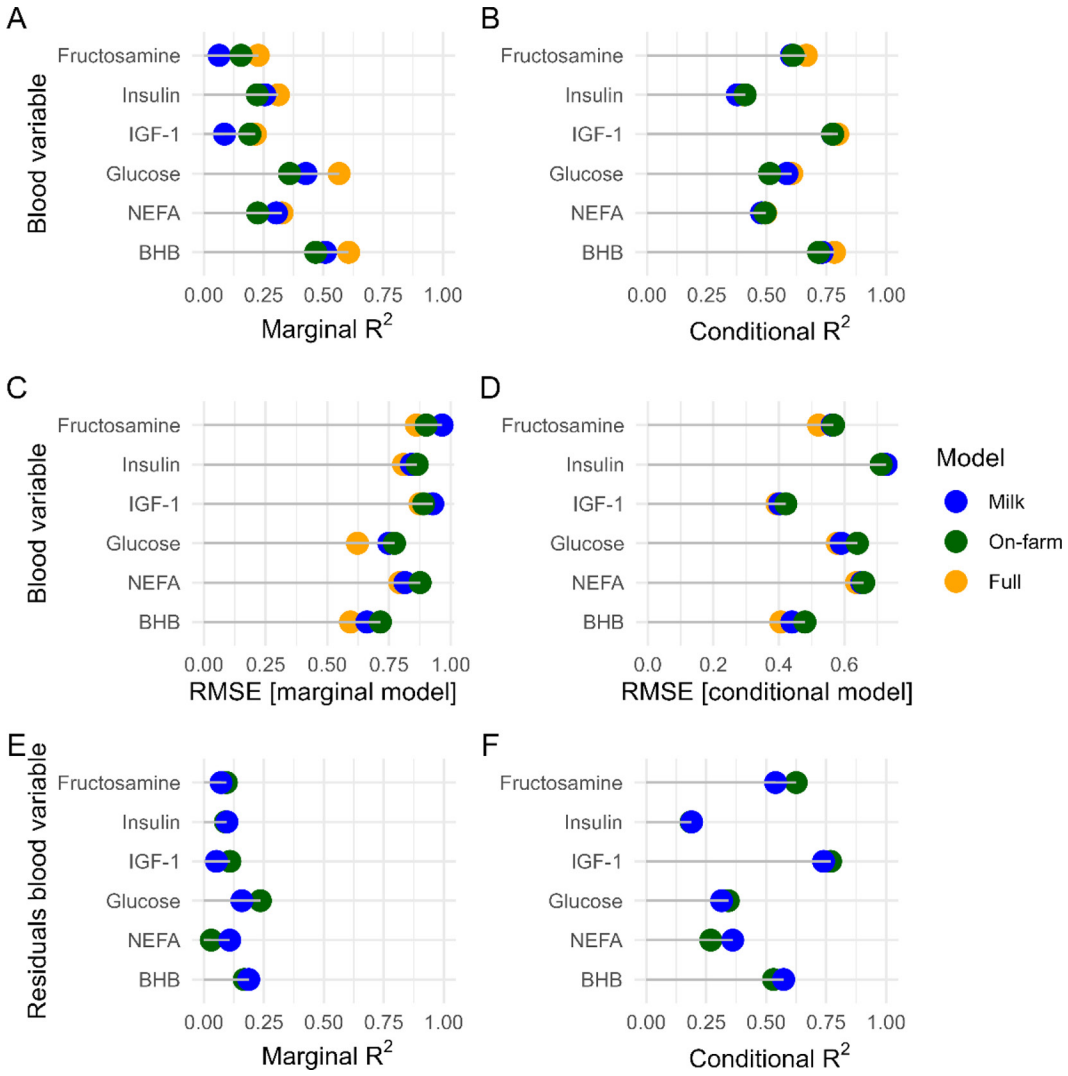

Figure 2. Marginal (A), conditional R^2 (B) values, and root mean squared error (RMSE) based on centered and scaled variables for the marginal (C) and conditional models (D) studying the associations between milk variables (determined by mid-infrared, except SCC), on-farm data (BCS, activity, and DMI), and milk and on-farm data combined (Full) in relation to metabolic blood variables BHB (mmol/L), nonesterified fatty acids (NEFA; mmol/L), glucose (mmol/L), insulin (ng/mL), IGF-1 (ng/mL), and fructosamine (μmol/L). (E) and (F) are the marginal and conditional R^2 of the mixed effect models using on-farm data predicting milk model residuals and using milk data predicting on-farm-model residuals of the blood variables. The BHB, NEFA, IGF-1, insulin, and fructosamine were $\log_{10}$ transformed. All numerically dependent and independent variables were centered and scaled before being entered into the model. Blood samples were collected on d 3, 6, 9, and 21 during lactation, whereas daily milk samples were obtained from d 3 to 23. Body condition scores were determined using a DeLaval BCS camera, and walking activity was monitored using IceTag3D motion sensors from IceRobotics.

on-farm models, first studying the raw metabolic blood variables and second the residuals of milk and on-farm models. Figure 3 reports the partial R^2 of the individual variables, and Figures 4 and 5 summarize the results of the sPLS in a variable plot and heatmap, respectively. Component 1 of the sPLS explained 16.2% and 21.4% of the variation in (1) the combined milk and on-farm dataset, and (2) the blood variables dataset, respectively. Component 2 explained 11.6% and 16.3% of the variation.

Modeling Metabolic Blood Variables: Description of Individual Variables in the Milk-Based Models

The significant milk variables (fixed effect, interaction, or both) in the different raw metabolic blood variable models are presented as full bars in Figure 3A. Green bars refer to a positive association and red to a negative association. In brief, all milk variables (C18:1, BHB, fat/protein ratio, urea, lactose, and logSCC) were related to one or multiple blood outcomes.

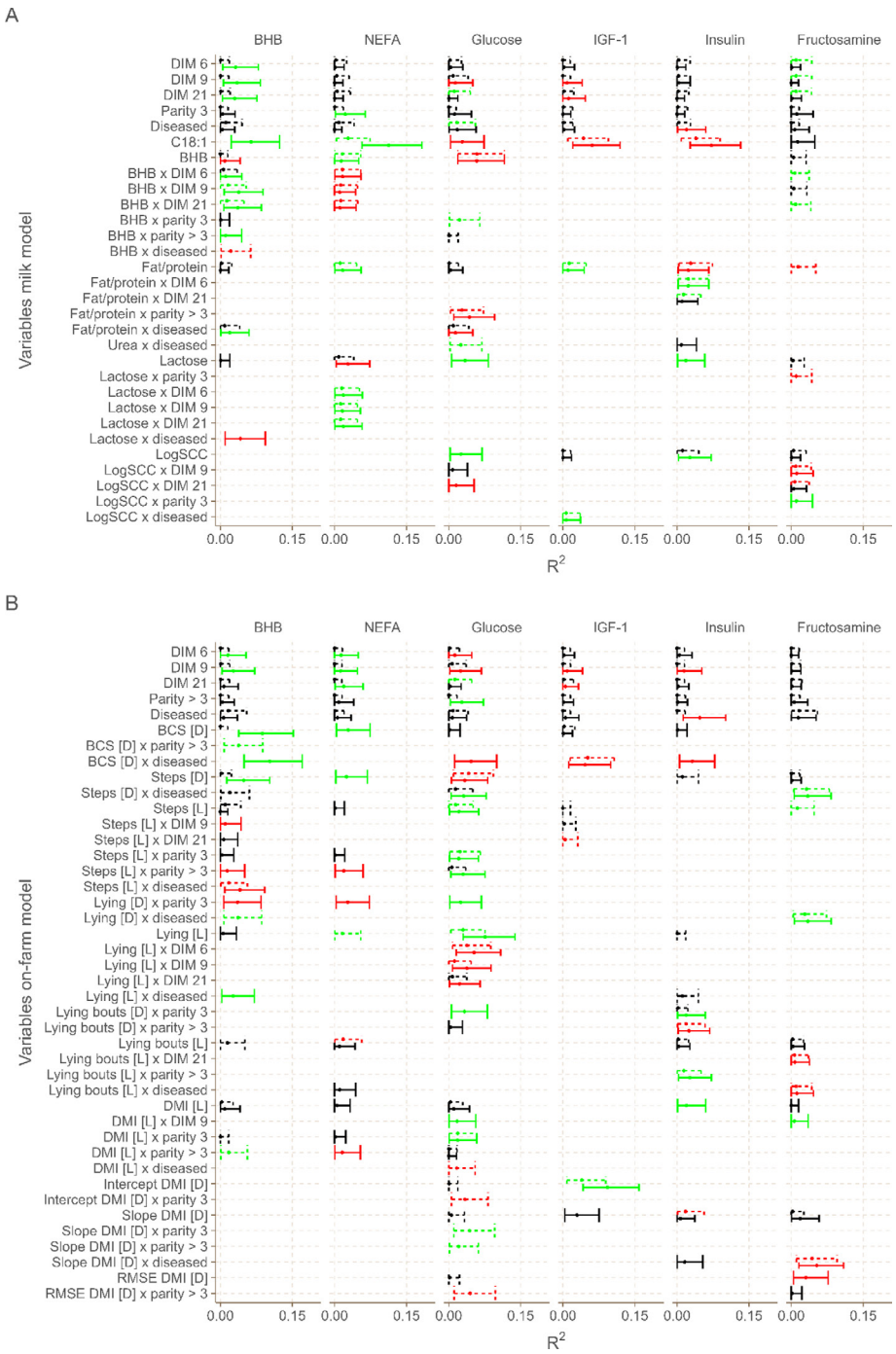

Figure 3. Partial R^2 values with 95% confidence limits of variables in the models studying the associations between (A) milk variables (determined by mid-infrared, except SCC) and (B) on-farm data (BCS, activity, and DMI) in relation to concentrations of metabolic blood variables BHB (mmol/L), nonesterified fatty acids (NEFA; mmol/L), glucose (mmol/L), insulin (ng/mL), IGF-1 (ng/mL), and fructosamine (μmol/L; presented by full lines), as well as in relation to the on-farm model residuals and milk model residuals of the different blood variables (presented by the dotted lines). Green and red lines refer to variables with a significant positive and negative estimate ($P < 0.05$), respectively. Lines in black refer to nonsignificant ($P > 0.05$) variables in the model. Blood samples were collected on d 3, 6, 9, and 21 during lactation, whereas daily milk samples were obtained from d 3 to 23. Body condition scores were determined using a DeLaval BCS camera, and walking activity was monitored using IceTag3D motion sensors from IceRobotics. Slope DMI [D], RMSE DMI [D], and intercept DMI [D] refer to the slope, root mean squared error (RMSE), and intercept of a linear mixed effect model studying the DMI during the last week before calving (d −7 to −3) in function of DIM, with a random slope and intercept for each lactation. Variables denoted with [L] and [D] refer to the lactation and dry period, respectively. For visual purposes, only variables with at least one P-value < 0.05 are displayed. The BHB, NEFA, IGF-1, insulin, and fructosamine were $\log_{10}$ transformed. All numerically dependent and independent variables were centered and scaled before being entered into the model.

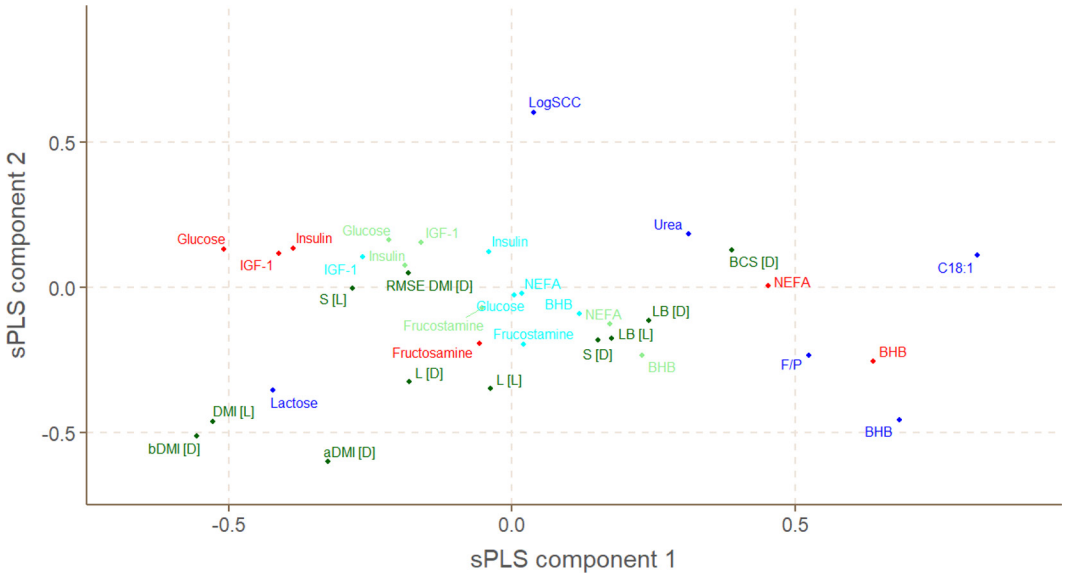

Figure 4. Variable plot of sparse partial least squares (sPLS) regression relating (1) blood variables (BHB, NEFA, glucose, insulin, IGF-1, and fructosamine; raw blood variables indicated in red) and (2) milk variables (determined by mid-infrared, except SCC; indicated in dark blue) and on-farm variables (variables related to the BCS, activity, and DMI; indicated in dark green). In addition to the concentrations of the different blood variables, 2 sets of Pearson marginal model residuals were also included in the sPLS, that is, residuals from the mixed models studying the blood variables (1) in function of milk variables (i.e., milk-model; indicated in light blue) and (2) in function of on-farm variables (i.e., on-farm model; indicated in light green). Blood samples were collected on d 3, 6, 9, and 21 during lactation, whereas daily milk samples were obtained from d 3 to 23. Body condition scores were determined using a DeLaval BCS camera, and walking activity was monitored using IceTag3D motion sensors from IceRobotics. The plot shows the 2 sPLS components (latent variables) and the loading vectors for each variable. F/P is the fat/protein ratio. LB, L, and S are the number of lying bouts, lying time, and the number of steps during lactation ([L]) or dry period ([D]). BCS [D] is the BCS during dry period. DMI [L] is the dry matter intake during lactation and bDMI [D], aDMI [D], and RMSE DMI [D] refer to the intercept, slope, and root mean squared error (RMSE) of a linear mixed effect model studying the DMI during the last week before calving (d −7 to −3) in function of DIM, with a random slope and intercept for each lactation. Blood BHB, NEFA, IGF-1, insulin, and fructosamine were $\log_{10}$ transformed. All numerically dependent and independent variables were centered and scaled before being entered into the model.

C18:1. C18:1 was a common significant variable in the milk models of all the blood variables, except fructosamine ($P = 0.07$). Among all milk variables in the models, C18:1 had the highest partial R^2, except for the glucose model (Figure 3A). The direction of the association between C18:1 and the different blood variables can also be derived from Figures 4 and 5. Figure 4 shows the positive loading of milk C18:1 as well as blood BHB and NEFA according to the first sPLS component. Milk C18:1 was negatively associated with glucose, IGF-1, and insulin (Figures 4 and 5).

In contrast to the overall significant (BHB, NEFA, IGF-1, and insulin: $P < 0.001$; glucose: $P = 0.002$) main effects observed for milk C18:1 and the various blood variables (excluding fructosamine: $P = 0.074$), the associations of milk BHB, lactose, urea, fat/protein ratio, and log SCC with the studied blood variables exhibited greater heterogeneity. Moreover, the direction of these associations frequently relied on interactions with DIM, parity, or disease state. Interactions with DIM sometimes resulted in neutralization of the main effect, which was tested with linear hypothesis tests. Some in-depth interpretation of these interactions is given below. Furthermore, partial R^2 of these milk variables were smaller than of milk C18:1, except for milk BHB, which was the milk variable showing the highest partial R^2 in the blood glucose model.

Fat/Protein Ratio. Similar to milk C18:1, the fat/protein ratio, either in or not in interaction with other variables, was significant in all the models except for the fructosamine model (Figure 3A). A higher fat/protein ratio was positively associated with the NEFA and IGF-1 concentrations ($P = 0.019$ and $P = 0.005$, respectively). In case of a clinical disease, a positive association with BHB ($P = 0.009$) and a negative association with glucose was found ($P = 0.04$). A negative interaction of the fat/protein ratio with parity >3 was observed in the glucose model ($P < 0.001$). For insulin, a negative main effect was observed ($P = 0.005$), but this was alleviated by positive interactions at d 6, 9, and 21 (Supplemental Figure S1).

Milk BHB. Milk BHB was negatively associated with the blood glucose concentration ($P < 0.001$) and had the highest partial R^2 in this model (0.059; Figure 3A). In the models studying IGF-1, insulin, and fructosamine, milk BHB was not retained during model se-

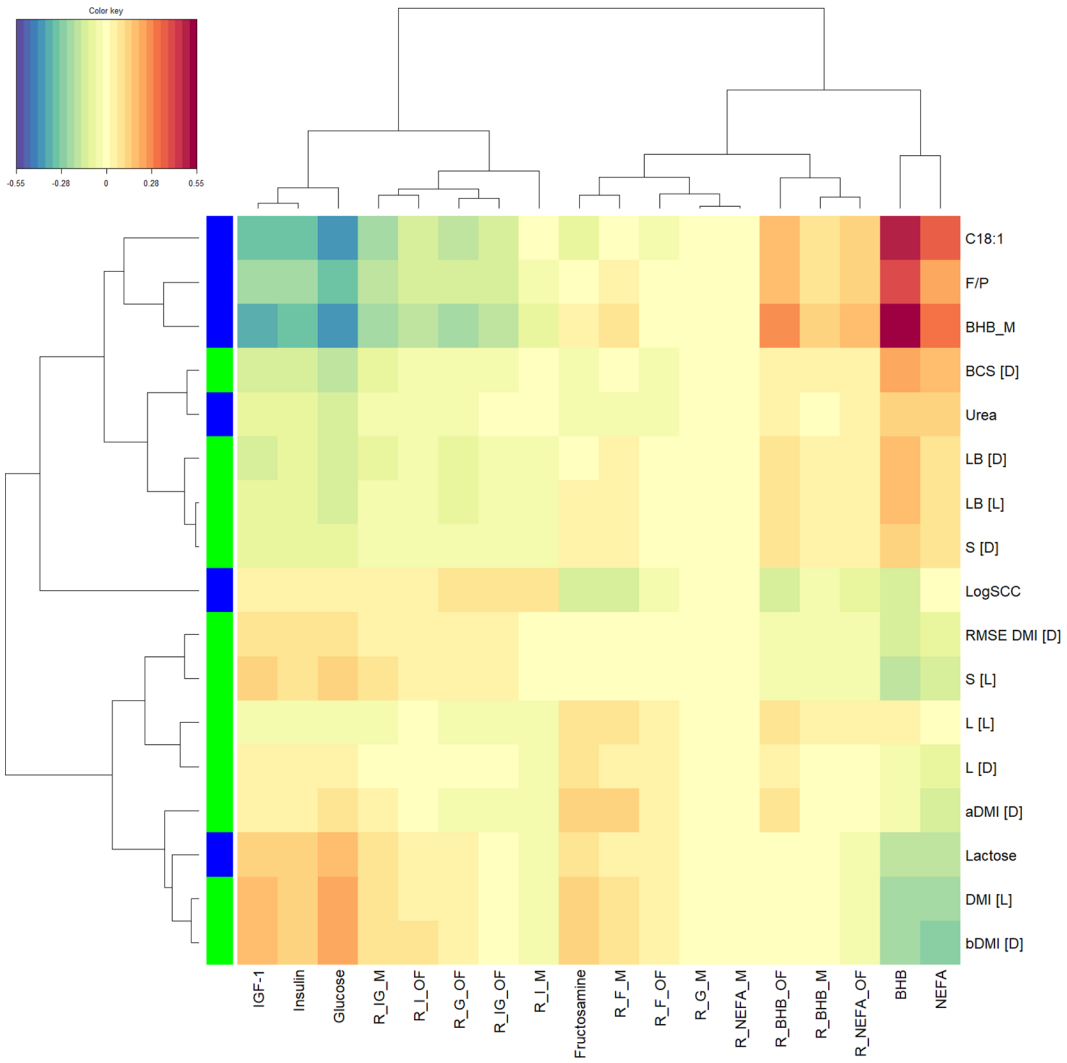

Figure 5. Clustered heatmap of sparse partial least squares (sPLS) regression relating the blood variables (BHB, NEFA, glucose, insulin, IGF-1, and fructosamine) and milk variables (determined by mid-infrared, except SCC; blue bars) and on-farm variables (BCS, activity, and DMI; green bars). Blood samples were collected on d 3, 6, 9, and 21 during lactation, whereas daily milk samples were obtained from d 3 to 23. Body condition scores were determined using a DeLaval BCS camera, and walking activity was monitored using IceTag3D motion sensors from IceRobotics. The plot represents the correlation structure extracted from (1) the dataset with blood variables and (2) the dataset with milk and on-farm variables. The correlation of each original variable pair is determined by each of their correlation with the components from the sPLS. F/P is the fat/protein ratio. R_ refers to model residuals for the blood variables BHB, NEFA, glucose (R_G), insulin (R_I), IGF-1 (R_IG) or fructosamine (R_F) from the milk model (_M) or the on-farm model (_OF). LB, L, and S are the number of lying bouts, lying time, and the number of steps during lactation ([L]) or dry period ([dry]). BCS [D] is the BCS during dry period. DMI [L] is the dry matter intake during lactation, and bDMI [D], aDMI [D], and RMSE DMI [D] refer to the intercept, slope, and root mean squared error (RMSE) of a linear mixed effect model studying the DMI during the last week before calving (d −7 to −3) in function of DIM, with a random slope and intercept for each lactation. BHB_M is the BHB concentration in the milk. The BHB, NEFA, IGF-1, insulin, and fructosamine were $\log_{10}$ transformed. All numerically dependent and independent variables were centered and scaled before being entered into the model.

lection. For blood BHB, positive estimates were found for the interaction between DIM (6, 9, and 21) and milk BHB (Figure 3A; $P = 0.02$, $P < 0.001$, and $P < 0.001$, respectively). These positive interactions had the highest magnitude at d 9 and 21 (Supplemental Figure S1). The significant main effect of milk BHB on blood BHB, however, was negative ($P = 0.04$). The linear hypothesis test showed the net effect of the negative main effect in combination with these positive interactions was different from 0 at d 9 and 21 ($P < 0.001$). Hence, at d 9 and 21, milk BHB was positively associated with blood BHB. This effect was more pronounced for parity >3 ($P = 0.04$). For blood NEFA a positive main effect was observed ($P = 0.02$) and negative interactions at d 6 ($P = 0.006$), 9 ($P = 0.046$), and 21 ($P = 0.04$; Figure 3A), which neutralized the positive main effect. The net

effect of the main effect and these interactions was not different from 0 ($P = 0.48$, 0.31, and 0.48, respectively).

Lactose. The main effect of lactose had the third highest partial R^2 value in the glucose model (0.034; Figure 3A). High milk lactose concentrations were associated with higher blood glucose concentrations ($P = 0.002$). The same was observed for insulin ($P = 0.03$). The lactose concentration had a negative estimate in the NEFA models ($P = 0.002$), although this effect was neutralized at d 6, 9, and 21 through the positive interaction terms ($P = 0.006$, $P = 0.010$, and $P = 0.007$, respectively), since linear hypothesis test showed the net effect of these interactions was not different from 0. In the case of a clinical disease status, lactose was negatively associated with the blood BHB concentration ($P < 0.001$).

SCC. The log SCC had a positive main effect on glucose ($P = 0.003$); however, at d 6, 9, and 21 this was alleviated by negative interactions ($P = 0.15$, 0.07, and 0.02, respectively; Supplemental Figure S1). Also, a positive main association was found for insulin ($P = 0.006$).

Modeling Metabolic Blood Variables: Description of Individual Variables in the On-Farm Models

The significant on-farm variables in the raw metabolic blood models are presented as full bars in Figure 3B. As in the milk models, green bars refer to a positive association and red to a negative association. In brief, less uniformity in the significance of the different on-farm variables over the different blood variables was found compared with the milk models.

BCS [D]. The average BCS [D] during the last 3 wk before calving was significant (as fixed effect or interaction) in all the models, except for the fructosamine model, where it was not retained during model selection. The BCS [D] was positively associated with the blood BHB and NEFA concentrations ($P < 0.001$ and $P = 0.03$, respectively) and had the highest partial R^2 in the model (BHB: 0.088 for the main effect and 0.103 for the interaction with health status; NEFA: 0.028 for the main effect; Figure 3B).

DMI During Lactation and the Dry Period. Dry matter intake during lactation was significant in all models (as main effect or in the interaction terms; $P < 0.05$), except in the BHB (tended to be negatively associated, $P = 0.08$) and IGF-1 model (not retained during model selection; Figure 3B). However, in the blood IGF-1 model, the intercept of DMI during the last week before calving was positively associated with the IGF-1 concentration ($P = 0.002$) and had the highest partial R^2 (0.093; Figure 3B). The negative coefficients of DMI during lactation for blood NEFA were not significant ($P = 0.20$), but a negative interaction with parity >3 was observed ($P = 0.03$; Supplemental Figure S2). Dry matter intake during lactation was also positively associated with the insulin concentration ($P = 0.018$). The slope of DMI (during the dry period) tended to be negatively associated with the IGF-1 concentration ($P = 0.081$) and it had the third highest partial R^2 (0.029; Figure 3B).

Interactions With DIM and Disease. Further, various associations between blood and on-farm variables showed an interactive effect with DIM, parity, or disease state. The net effect often depended on DIM, which is important because this sometimes resulted in neutralization or counteraction of the main effect, depending on the DIM. Some in depth interpretation of these interactions is given below. In general, health status often interacted with different on-farm variables. The positive association between BCS [D] and blood BHB concentration ($P < 0.001$) was amplified in case of clinical disease ($P < 0.001$). Moreover negative associations between the BCS [D] and the glucose, insulin, and IGF-1 concentrations were only observed in case of the occurrence of a clinical disease in early lactation ($P = 0.003$, $P = 0.013$, and $P = 0.02$, respectively).

Number of Steps. Elevated number of steps during lactation was positively associated with the glucose concentration ($P = 0.011$; Figure 3B). The number of steps at d 9 was negatively associated with the BHB concentration at d 9 ($P = 0.02$). Increased number of steps in combination with parity >3 or disease was associated with lower blood BHB concentration ($P = 0.03$ and $P < 0.001$, respectively). Increased number of steps during the dry period was associated with increased blood BHB concentrations ($P = 0.006$). For glucose, a negative association was found between blood BHB and the number of steps during the dry period ($P = 0.011$).

Lying Time. Lying time during lactation had the highest partial R^2 in the model studying glucose (0.076; Figure 3B). Higher lying time at d 3 was associated with higher glucose concentrations ($P < 0.001$), whereas at d 6, 9 and 21 this was alleviated by the negative interaction (net effect of main and interaction terms did not differ from 0; $P = 0.66$, $P = 0.79$, and $P = 0.099$, respectively).

DMI. For glucose the effect depended on DIM; at d 9 a positive interaction was observed ($P = 0.008$; Figure 3B), which neutralized the tendency for a negative main effect ($P = 0.057$; Supplemental Figure S2).

Modeling Residual Variation in Metabolic Blood Variables Using On-Farm and Milk Variables

General Supplementary and Substitutive Aspects of On-Farm and Milk Models. To assess

whether milk variables and on-farm data could account for unexplained variance in the milk and on-farm models, Pearson marginal model residuals were extracted and modeled in function of on-farm and milk variables, respectively. Model coefficients of the on-farm and milk model are reported in Supplemental Figures S3 and S4 (https://doi.org/10.6084/m9.figshare.22730858.v2; Heirbaut et al., 2023b). In addition to modeling the residuals, a full model was also constructed combining the milk and on-farm variables in one model (Figure 2). Combining the milk variables and on-farm data resulted in higher R^2 values, in particular for BHB, glucose, and IGF-1, but the blood variables kept the same order of R^2 values as the on-farm model (Figure 2A), except IGF-1 and fructosamine. A marginal R^2 of 0.606 was observed for BHB, followed by glucose (0.566), NEFA (0.327), insulin (0.312), fructosamine (0.228), and IGF-1 (0.216).

Milk C18:1 Versus On-Farm Variables. The variable plot of the sPLS (Figure 4) shows the orthogonal position of milk C18:1 with respect to lying time during lactation, as well as during the dry period. In line with this, milk C18:1 remained significant in the models relating blood NEFA, IGF-1, and insulin residuals of the on-farm models (independent variables) to milk variables (dependent variables; Figure 3A; Supplemental Figure S3). However, the partial R^2 values of milk C18:1 always decreased compared with the milk models with raw blood data as dependent variables. Further comparison of full and dotted lines in Figure 3A revealed that milk variables which showed significant associations with the raw blood variables in most cases remained significantly associated with their residuals of the on-farm models.

BCS and DMI During the Lactation Versus Milk Variables. In the on-farm models studying the raw blood variables, BCS [D] and DMI (during lactation, as well as the dry period) were generally considered relatively important variables. In contrast, their role fluctuated more in the residual on-farm models. The BCS [D] remained the variable with the highest partial R^2 relating BHB residuals of the milk model (independent variables) to on-farm variables (dependent variables; BCS [D] × parity >3: partial R^2 0.038; Figure 3B; Supplemental Figure S4). However, whereas it was the most important variable when studying the raw NEFA values, it was not retained anymore in the model with NEFA residuals as dependent variable. With a partial R^2 of 0.019 for DMI × parity 3, the importance of DMI during lactation remained quite similar in the model studying the residuals of glucose concentration as compared with the original on-farm model. However, DMI during lactation was no longer retained in the models targeting the residuals of blood NEFA and insulin (Figure 3B; Supplemental Figure S4). On the other hand, the interaction DMI × parity >3 was positively associated with the residuals of blood BHB ($P = 0.012$), whereas in case of the raw blood BHB it was not retained.

DMI During the Dry Period Versus Milk Variables. The intercept of DMI during the dry period remained positively associated in the model describing IGF-1 residuals ($P = 0.044$; Figure 3B; Supplemental Figure S4). This is in line with the fact that C18:1 is the most important milk variable in relation to IGF-1 ($P < 0.001$; Figure 3A) and the independence of C18:1 and DMI during the dry period based on the partial orthogonal position in Figure 4. As such the intercept of DMI during the dry period had the second highest partial R^2 in the model targeting IGF-1 residuals (Figure 3B; Supplemental Figure S4).

Some exceptions of on-farm variables that became significant in the models studying the milk model residuals, although not significant in the models of the raw blood variables, are discussed below. Different DMI-related variables from the dry period became significantly associated with the glucose residuals from the milk model. For instance, the interaction of the slope of DMI during the dry period and parity 3 and >3 became positively associated with the blood residual glucose concentration ($P = 0.002$ and $P = 0.03$, respectively) and the interaction RMSE DMI during dry period × parity >3 negatively ($P = 0.002$), whereas before they were not associated (not selected during model selection).

DISCUSSION

The objective of this research was to (1) identify associations between metabolic blood variables, milk variables, and various on-farm sensor data and (2) study their supplementary or substitutive potential for explaining variation in various metabolic blood variables. It was hypothesized that the combination of both data sources would be a better proxy for assessing the concentrations of the metabolic blood variables than the single data sources. In the present discussion, first the general performance of the milk and on-farm models will be discussed, then the most contributing individual variables in the milk and on-farm models will be examined and finally the results from the supplementary and substitutive potential of these data sources will be interpreted.

General Situation of Milk and On-Farm Models

In the literature, milk variables have been extensively used for (predictive) classification or regression of blood

BHB and NEFA concentrations (van der Drift et al., 2012; Dórea et al., 2017; Aernouts et al., 2020), whereas associations with other blood variables and on-farm data remain rather unexplored. In our study, milk-based models assessing blood BHB, NEFA, and glucose obtained the highest R^2 values. The combination of the fixed effects of the milk and the on-farm variables explained 60.6, 56.6, and 32.7% of the variation in BHB, glucose, and NEFA concentration, respectively. In addition, associations with fructosamine and the hormones insulin and IGF-1 were found. This diversity of associations with the multiple metabolic blood variables is not surprising, since these metabolic blood variables are all physiologically involved in the regulation of homeostasis and homeorhesis during early lactation. Although the R^2 values are still moderate, the overall model outcome for BHB is roughly in line with literature results. Bonfatti et al. (2019) found a calibration R^2 of 0.56 and a validation R^2 of 0.50 when predicting BHB using milk infrared spectra. However, the milk model R^2 value of NEFA was rather low (0.322) compared with literature. Aernouts et al. (2020) found an R^2 value of 0.502 when associating blood NEFA to milk MIR spectra. The lower R^2 value of the milk model in our research could be attributed to (1) the scope of our study (i.e., associative modeling, rather than predictive modeling); (2) the more limited number of variables included in our models compared with the inclusion of the full milk spectra; and (3) focus on the lactation start (DIM 3, 6, and 9 represented about 75% of the data). To our knowledge, no single other study has put this much emphasis on the early lactation, which is crucial for diagnosis in practice, but can also impair the model performance. For instance, removing data from d 3 and 6 from our analysis and refitting the models, powerfully increased the marginal R^2 of the milk model for BHB from 0.508 to 0.605 and a lower but still important increase from 0.303 to 0.350 for NEFA. Additionally, sampling time also could have contributed to this result as Mäntysaari et al. (2019) and Aernouts et al. (2020) reported lower model performance when using morning milk samples (as done in our study) compared with evening milk to predict blood NEFA. Indeed, as shown in the study of Seely et al. (2021), there exist important diurnal variations in BHB and NEFA concentrations. It is important to note that the time of sampling in relation to the diurnal dynamics may influence the strength of the association. For instance in the study of Seely et al. (2021) NEFA peaks were observed 2 h before morning feeding, whereas BHB peaks were observed 4 h after morning feeding. Moreover, and in line with the studies of Mäntysaari et al. (2019) and Aernouts et al. (2020), Seely et al. (2022) showed that the morning milk sampling may result in underestimation of the association because the lowest relative concentration of preformed fatty acids was observed at 0600 h and the highest at 1400 h. In the current study we have chosen to be consistent regarding the time of milk and blood sampling relative to feeding to avoid any undesirable variation. Although not the focus of this study, taking into account specific variable related diurnal optimal sampling times might improve the absolute strength of the association in future.

Metabolic Status and the Individual Milk Model-Based Variables

Milk C18:1 was an important variable associated with all blood variables, which corroborates with the predominant contribution of C18:1 *cis*-9 to the blood NEFA composition during body mobilization (Hostens et al., 2012). Accordingly, milk C18:1 *cis*-9 generally is considered an important biomarker of body reserves' mobilization (Jorjong et al., 2014). Furthermore, when the blood NEFA concentration is higher than the oxidative capacity of the liver, NEFA are incompletely oxidized to ketone bodies (BHB, acetone, acetoacetate) or NEFA can be re-esterified into triglycerides within the liver, which are then stored in lipid droplets or transported out of the liver for energy utilization or storage in adipose tissue. (LeBlanc, 2010). This is also in line with the research of Jorjong et al. (2015) showing via a logistic regression model the milk C18:1 *cis*-9-to-C15:0 ratio as the most important variable to diagnose hyperketonemia (BHB >1.20 mmol/L). In this study 70% of cows with hyperketonemia had C18:1 *cis*-9-to-C15:0 ratios exceeding 40, whereas 90% of nonhyperketonemia cases had a C18:1 *cis*-9-to-C15:0 ratio below this threshold. Finally, milk C18:1 was negatively associated with the blood IGF-1 and insulin concentration. Low IGF-1 can be used as a biomarker for negative energy balance (Wathes et al., 2007), whereas insulin inhibits lipolysis and decreases NEFA concentration (Butler et al., 2003). On the other hand, the incomplete oxidation of these NEFA, as measured by milk BHB, was not associated with blood insulin concentrations. In line with this, Zarrin et al. (2017) did not find any postpartum associations between BHB and insulin. Milk BHB has been studied extensively as biomarker to detect elevated blood BHB concentrations (Tatone et al., 2017; Renaud et al., 2019), which was reflected in the positive associations between milk and blood BHB at d 9 and 21. Milk BHB also was significantly associated with the blood glucose concentration. In addition to an indirect association, Zarrin et al. (2017) showed that BHB infusion in blood decreased the plasma glucose concentration during early lactation. Moreover one-third of hyperketonemic cows are simultaneously

hypoglycemic (Dubuc and Buczinski, 2018). Hence monitoring ketone bodies in milk can give information on the blood glucose status.

The glucose concentration was positively associated with the lactose content. Indeed, glucose is known as the precursor of lactose and induces the cell growth in dairy cow mammary cells (Lin et al., 2016). Given its relation with energy balance (Ouweltjes et al., 2007), milk lactose has been suggested as a biomarker for ketosis (Costa et al., 2019; Yang et al., 2019). In line with this, the lactose concentration was negatively associated with blood BHB concentration in case of disease and negatively associated with NEFA at d 3 in lactation.

Metabolic Status and the Individual On-Farm Model-Based Variables

Dry matter intake and BCS [D] have already been widely studied in relation to different metabolic blood variables. In line with our results, different studies have shown a negative association between (1) NEFA and DMI (Piantoni et al., 2015), and (2) hyperketonemia and DMI (González et al., 2008; Goldhawk et al., 2009; Yang et al., 2019). In our study DMI did not show a significant association with blood BHB, only a tendency was observed ($P = 0.083$), which might be due to the model correction for clinical disease events in our study. Monitoring the BCS [D] explained a relatively high proportion of the variation in postpartum BHB and NEFA concentrations. In line with our results, Pires et al. (2013) and Gillund et al. (2001) have shown that high BCS [D] at calving was associated with increased hyperketonemia risk postpartum, but no differences in glucose and insulin concentrations were found (Pires et al., 2013). In general high BCS before calving is associated with a condition of increased catabolic state postpartum, a negative association with IGF-1 would be expected. However, as opposed to BHB, the IGF-1 concentration was only negatively associated with the BCS [D] in case of disease ($P = 0.02$). Meikle et al. (2004) showed that high BCS was not significantly associated with the IGF-1 concentration postpartum, but the decay in IGF-1 concentration from prepartum until after calving was higher for cows with high BCS. In line with this, Wathes et al. (2007) found a positive association between IGF-1 and BCS both measured at wk 2 and 4 of lactation, which would be attributed to the fact that cows with better energy status—and hence less mobilization of body reserves—have a higher IGF-1 status. Indeed, the low insulin status during negative energy balance is assumed to be responsible for the downregulation of the liver growth hormone receptor 1A resulting in less stimulation by growth hormone on the liver IGF-1 production (Butler et al., 2003).

The direction of the association between activity and the blood BHB and glucose concentration often depended on the DIM. At d 9, a negative interaction effect was observed between the BHB concentration and the number of steps ($P = 0.02$). Edwards and Tozer (2004) showed that walking activity is generally lower in sick cows. In line with this, Najm et al. (2020) found a lower activity for cows suffering from hyperketonemia. On the other hand, it should be noted that the lactation stage is also important. In our research, no uniform association was found, but rather some fluctuation over the DIM (interaction effects). For instance, lying time had a positive main estimate for the glucose concentration ($P < 0.001$); however, significant negative interaction effects with decreasing magnitude over the period d 6 to 21 were found, alleviating this positive effect. In our study the metabolic status was only monitored until d 21, but van Hoeij et al. (2019) found an increased lying time for cows with a better metabolic status in wk 4 of lactation. The pathways of these associations are largely unknown and should be further investigated. Kaufman et al. (2016) suggested that the reduced activity in cows with hyperketonemia might be a result of an energy saving mechanism. Itle et al. (2015) hypothesized a lower hierarchical position of these animals and consequently longer waiting times at the feed bunk were at the (causal) origin of these differences in activity levels. Significant effects of activity were indeed often accompanied by effects of feed intake in our study. To assess the association between activity and DMI, an additional mixed effect model was constructed studying the postpartum DMI in function of the activity variables (average D_{X-1} to D_{X+1}, following the same methodology as described in Materials and Methods). The model had an R^2 of 0.47, with a partial R^2 of 0.31 and 0.23 for factors DIM 21 and clinical disease, respectively. The number of steps and lying time tended to positively associated ($P = 0.06$ and $P = 0.07$, respectively) with the DMI, which could be in line with the theory that cows with lower DMI are the cows which have to wait longer time at the feeding bunk (lower number of steps and lower lying time).

Interestingly differences in activity were already noticeable during the dry period. Increased number of lying bouts during the dry period × parity 3 was associated with increased glucose and insulin concentrations ($P = 0.014$ and $P = 0.004$, respectively). In contrast, the number of steps during the dry period was positively associated with blood BHB and NEFA and negatively with glucose concentrations ($P = 0.006$, $P = 0.039$, and $P = 0.011$, respectively). This seems to imply that high activity during the dry period is associated with impaired metabolic health status. However, as opposed to postpartum behavior, conflicting

results are reported in literature. Rodriguez-Jimenez et al. (2018) found reduced standing time during the prepartum period for cows with postpartum subclinical hyperketonemia. In contrast, Itle et al. (2015) found a reduced lying time for cows with clinical hyperketonemia during the week before calving, whereas Kaufman et al. (2016) did not find any prepartum differences in lying time, lying bout length, and lying bout frequency. However, reduced lying time during the postpartum period was observed for hyperketonemia in the study of Kaufman et al. (2016). These results, contradictory at first, actually emphasize the complexity of activity as an early warning tool. Interanimal differences in activity suggest that considering individual-specific normal and healthy behaviors may offer more comprehensive insights into the (deviations of) health status rather than solely modeling at group level (Wagner et al., 2020). In our study the lower marginal R^2 values compared with the high conditional R^2 values, that account the animal-specific random intercept, seem to support this. Factors further complicating the activity–health status interaction include time to occurrence of the disease event and type of (multifactorial) disease. For instance, Edwards and Tozer (2004) found higher walking activity in cows with hyperketonemia compared with healthy animals, 8 d before the diagnosis of hyperketonemia. Hence, for cows diagnosed during wk 1 of lactation, this could also affect the activity during the dry period. It should also be noted that in our study, diseased versus nondiseased cows were considered, and further differentiation in type of disease was not made. This an important source of unknown variation because, depending on the type of disease, the activity levels can be influenced positively or negatively. For instance, activity is lower in cows diagnosed with metritis (Liboreiro et al., 2015; Stangaferro et al., 2016). However, in case of hypocalcemia the increased lying time postpartum is preceded by a reduced lying time 24 h before calving (Jawor et al., 2012). Moreover, it should also be noted that the number of lying bouts in our study is rather approached as a "restless indicator" as opposed to the real physical number of lying bouts. Indeed, Kok et al. (2015) have shown that the converting algorithm for IceTags data generates false positive lying bouts due to horizontal leg movements, which drastically increases the number of lying bouts reported. As such the "number of lying bouts" can be much higher in the period around calving. Applying time filter criteria can help to improve the accuracy of the lying behavior monitoring, but could have the danger of filtering restless lying behavior related to (sub)optimal conditions.

Supplementary and Substitutive Aspects of On-Farm and Milk Variables

Combining milk and on-farm data consistently increased the marginal R^2 values, although with different magnitudes. Interestingly, compared with the milk model, on-farm data were considerably better in describing the IGF-1 concentration and powerfully increased the marginal R^2 value. Figure 4 shows the orthogonal position of (1) milk C18:1 and the fat/protein ratio versus (2) the intercept of DMI during the dry period, which suggests (partial) independence of information. Indeed, the intercept of DMI during the dry period remained significant when the IGF-1 residuals from the milk model were studied ($P = 0.044$), and milk C18:1 and the fat-to-protein ratio remained significant when the on-farm IGF-1 residuals were studied ($P < 0.001$ and $P = 0.003$, respectively). Based on our results, on-farm data can give more information about the IGF-1 status compared with milk data, which is remarkable because the significant on-farm variables are mainly related to the period before calving. As such, the period before calving gave more information than the milk biomarkers assessed during early lactation. This could suggest a link between IGF-1 and the energy intake before calving. Lower IGF-1 concentrations limit its negative feedback on the growth hormone production in the pituitary gland and consequently lipolysis is stimulated. In line with this, in humans, IGF-1 also has been considered as a biomarker for malnutrition and energy shortage (Caregaro et al., 2001). Piechotta et al. (2015) found that cows suffering from clinical ketosis had lower prepartum IGF-1 concentrations. They hypothesized that the increased GH production and lipolysis during the end of the pregnancy increases the risk for ketosis. Additionally, Wathes et al. (2021) also showed that low IGF-1 concentration is associated with lower DMI. Although not studied in cattle, Hawkes and Grimberg (2015) discussed that a reduced leptin concentration and consequently increased hypothalamic production of neuropeptide Y (NPY) may affect the GH/IGF-1 axis by reducing the pituitary GH secretion. In literature impaired DMI before calving has been extensively associated with hyperketonemia. For instance, Goldhawk et al. (2009) found that a 1-kg decrease in DMI during the last week before calving increased the risk of subclinical hyperketonemia by a factor of 2.2. Despite the unquestionable biological importance of DMI before calving in relation to the health status, our results do not suggest a direct association between DMI during the dry period and increased blood BHB and NEFA concentrations in early lactation. In line with

this, Dann et al. (2005) studied the effect of energy restriction during the dry period and did not find any effect on serum BHB postpartum. Horst et al. (2021) argued that high BHB and NEFA are not indicators of metabolic imbalance, but could reflect a normal and biologically healthy state when not associated with a drop in pre- or postpartum DMI. As opposed to DMI before calving, the BCS [D] was quite well associated with the blood BHB and NEFA concentrations ($P < 0.001$ and $P = 0.029$, respectively). Even taking into account that milk variables such as milk BHB (Ježek et al., 2017) and milk-infrared-estimated ketone bodies are well suited to predict BHB (Chandler et al., 2018), BCS [D] still succeeded to explain additional variation in blood BHB concentration (BCS [D] × parity >3; $P = 0.018$). Moreover, it should be noted that in our research a BCS camera has been used, which by some has been considered as inaccurate at determining the high and low BCS scores (Mullins et al., 2019). Nevertheless, the BCS [D] still remained of importance to assess blood BHB concentrations.

Despite the high importance of lactose when studying the glucose concentrations, lactose was not retained during mode selection when describing on-farm glucose residuals. Hence, it seems that information about the animal's glucogenic status provided by the lactose concentration also is reflected in sensor data. According to the sPLS, lactose was positioned in the proximity of DMI-related variables (both in lactation as well as the dry period). This could confirm the relation with the glucose status. In ruminants gluconeogenesis is crucial to produce glucose and is typically highest during and after conditions of high feed intake since at that time the highest production and absorption of propionate occurs (Aschenbach et al., 2010). It should be noted that the total variation explained by the milk model was still higher than the variation explained by the on-farm model. Nevertheless, combining the 2 different sets of variables in a single model notably increased R^2. Furthermore, in relation to the on-farm glucose residuals, the fat/protein ratio remained important (parity >3 × fat/protein ratio: $P = 0.003$). According to Jenkins et al. (2015) the fat/protein ratio can be used as a screening indicator of hyperketonemia, but not as final diagnostic criterion due its limitations regarding low sensitivity, specificity, or both (Duffield et al., 1997; Jenkins et al., 2015). The fat/protein ratio was indeed positioned in the proximity of blood BHB according to the first axis of the sPLS plot (Figure 4). Variables related to feeding were positioned opposite according to the first axis and additionally had a stronger negative score according to the second axis. As such, combining the fat/protein ratio and DMI variables could be valuable as a first warning indicator of metabolic status.

CONCLUSIONS

This study has shown that on-farm data combined with milk data can provide additional information concerning the metabolic health status of dairy cows. In general, milk biomarkers can be used to assess the metabolic status, but early stage predictions are still challenging and have not been systematically addressed in literature. According to this study, the inclusion of on-farm data could help to facilitate the diagnosis of the metabolic status of dairy cows postpartum. Moreover, several on-farm variables measured during the prepartum period were linked to the postpartum metabolic status of the dairy cow. As such, these data could be of major interest to identify and select animals at risk of metabolic disruption at an early stage. These potentially could then be monitored more strictly in the early period after calving (e.g., through milk variables). Hence, future predictive modeling research could benefit from taking into account additional on-farm data.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of Flanders Innovation and Entrepreneurship (VLAIO; Brussels, Belgium; LA170830). The research of Xiaoping Jing was supported by the Chinese Scholarship Council (CSC, Beijing, 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, Ghent, Belgium). The authors gratefully acknowledge the staff from the Animal Science Unit, Flanders Research Institute for Agriculture, Fisheries and Food (Melle, Belgium), 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, Ghent, Belgium), Ilke van Hese (Animal Sciences Unit, Flanders Research Institute for Agriculture, Fisheries and Food [ILVO], Scheldeweg, Melle, Belgium; Department of Reproduction, Obstetrics and Herd Health Faculty of Veterinary Medicine, Ghent University, Salisburylaan, Merelbeke, Belgium), and Mirjan Thys (Animal Sciences Unit, ILVO, Scheldeweg, Melle, Belgium) for assistance during sampling. Finally, we express our gratitude to the two anonymous reviewers for their input, which has significantly contributed to the enhancement of this manuscript. The authors have not stated any conflicts of interest.

REFERENCES

Abuelo, A., L. Wisnieski, J. L. Brown, and L. M. Sordillo. 2021. Rumination time around dry-off relative to the development of diseases in early-lactation cows. J. Dairy Sci. 104:5909–5920. https://doi.org/10.3168/jds.2020-19782.

Adriaens, I., N. C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. J. Dairy Sci. 103:7155–7171. https://doi.org/10.3168/jds.2019-17826.

Aernouts, B., I. Adriaens, J. Diaz-Olivares, W. Saeys, P. Mäntysaari, T. Kokkonen, T. Mehtiö, S. Kajava, P. Lidauer, M. H. Lidauer, and M. Pastell. 2020. Mid-infrared spectroscopic analysis of raw milk to predict the blood nonesterified fatty acid concentrations in dairy cows. J. Dairy Sci. 103:6422–6438. https://doi.org/10.3168/jds.2019-17952.

Aschenbach, J. R., N. B. Kristensen, S. S. Donkin, H. M. Hammon, and G. B. Penner. 2010. Gluconeogenesis in dairy cows: The secret of making sweet milk from sour dough. IUBMB Life 62:869–877. https://doi.org/10.1002/iub.400.

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.

Belaid, M. A., M. Rodriguez-Prado, M. López-Suárez, D. V. Rodríguez-Prado, and S. Calsamiglia. 2021. Prepartum behavior changes in dry Holstein cows at risk of postpartum diseases. J. Dairy Sci. 104:4575–4583. https://doi.org/10.3168/jds.2020-18792.

Bonfatti, V., S.-A. Turner, B. Kuhn-Sherlock, T. D. W. Luke, P. N. Ho, C. V. C. Phyn, and J. E. Pryce. 2019. Prediction of blood β-hydroxybutyrate content and occurrence of hyperketonemia in early-lactation, pasture-grazed dairy cows using milk infrared spectra. J. Dairy Sci. 102:6466–6476. https://doi.org/10.3168/jds.2018-15988.

Butler, S. T., A. L. Marr, S. H. Pelton, R. P. Radcliff, M. C. Lucy, and W. R. Butler. 2003. Insulin restores GH responsiveness during lactation-induced negative energy balance in dairy cattle: Effects on expression of IGF-I and GH receptor 1A. J. Endocrinol. 176:205–217. https://doi.org/10.1677/joe.0.1760205.

Calamari, L., N. Soriani, G. Panella, F. Petrera, A. Minuti, and E. Trevisi. 2014. Rumination time around calving: An early signal to detect cows at greater risk of disease. J. Dairy Sci. 97:3635–3647. https://doi.org/10.3168/jds.2013-7709.

Caregaro, L., A. Favaro, P. Santonastaso, F. Alberino, L. Di Pascoli, M. Nardi, S. Favaro, and A. Gatta. 2001. Insulin-like growth factor 1 (IGF-1), a nutritional marker in patients with eating disorders. Clin. Nutr. 20:251–257. https://doi.org/10.1054/clnu.2001.0397.

Chandler, T. L., R. S. Pralle, J. R. R. Dórea, S. E. Poock, G. R. Oetzel, R. H. Fourdraine, and H. M. White. 2018. Predicting hyperketonemia by logistic and linear regression using test-day milk and performance variables in early-lactation Holstein and Jersey cows. J. Dairy Sci. 101:2476–2491. https://doi.org/10.3168/jds.2017-13209.

Costa, A., C. Egger-Danner, G. Mészáros, C. Fuerst, M. Penasa, J. Sölkner, and B. Fuerst-Waltl. 2019. Genetic associations of lactose and its ratios to other milk solids with health traits in Austrian Fleckvieh cows. J. Dairy Sci. 102:4238–4248. https://doi.org/10.3168/jds.2018-15883.

Dann, H. M., D. E. Morin, G. A. Bollero, M. R. Murphy, and J. K. Drackley. 2005. Prepartum intake, postpartum induction of ketosis, and periparturient disorders affect the metabolic status of dairy cows. J. Dairy Sci. 88:3249–3264. https://doi.org/10.3168/jds.S0022-0302(05)73008-3.

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. https://cran.r-project.org/package=data.table.

Dubuc, J., and S. Buczinski. 2018. Short communication: Cow- and herd-level prevalence of hypoglycemia in hyperketonemic postpartum dairy cows. J. Dairy Sci. 101:3374–3379. https://doi.org/10.3168/jds.2017-13773.

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.

Edwards, J. L., and P. R. Tozer. 2004. Using activity and milk yield as predictors of fresh cow disorders. J. Dairy Sci. 87:524–531. https://doi.org/10.3168/jds.S0022-0302(04)73192-6.

Gillund, P., O. Reksen, Y. T. Gröhn, and K. Karlberg. 2001. Body condition related to ketosis and reproductive performance in Norwegian dairy cows. J. Dairy Sci. 84:1390–1396. https://doi.org/10.3168/jds.S0022-0302(01)70170-1.

Goldhawk, C., N. Chapinal, D. M. Veira, D. M. Weary, and M. G. von Keyserlingk. 2009. Prepartum feeding behavior is an early indicator of subclinical ketosis. J. Dairy Sci. 92:4971–4977. https://doi.org/10.3168/jds.2009-2242.

González, L. A., B. J. Tolkamp, M. P. Coffey, A. Ferret, and I. Kyriazakis. 2008. Changes in feeding behavior as possible indicators for the automatic monitoring of health disorders in dairy cows. J. Dairy Sci. 91:1017–1028. https://doi.org/10.3168/jds.2007-0530.

Goode, K., K. McClernon, J. Zhao, Y. Zhang, and Y. Huo. 2022. redres: Residuals and Diagnostic Plots for Mixed Models. R package version 0.0.0.9. https://github.com/goodekat/redres.git.

Hamner, B., and M. Frasco. 2018. Metrics: Evaluation Metrics for Machine Learning. R package version 0.1.4. https://CRAN.R-project.org/package=Metrics.

Hawkes, C. P., and A. Grimberg. 2015. Insulin-like growth factor-I is a marker for the nutritional state. Pediatr. Endocrinol. Rev. 13:499–511.

Heirbaut, S., D. Børge Jensen, X. P. Jing, B. Stefańska, P. Lutakome, L. Vandaele, and V. Fievez. 2022. Different reticuloruminal pH metrics of high-yielding dairy cattle during the transition period in relation to metabolic health, activity, and feed intake. J. Dairy Sci. 105:6880–6894. https://doi.org/10.3168/jds.2021-21751.

Heirbaut, S., X. P. Jing, B. Stefańska, E. Pruszyńska-Oszmałek, L. Buysse, P. Lutakome, M. Q. Zhang, M. Thys, L. Vandaele, and V. Fievez. 2023a. Diagnostic milk biomarkers for predicting the metabolic health status of dairy cattle during early lactation. J. Dairy Sci. 106:690–702. https://doi.org/10.3168/jds.2022-22217.

Heirbaut, S., X. P. Jing, B. Stefańska, E. Pruszyńska-Oszmałek, B. Ampe, C. Umstätter, L. Vandaele, and V. Fievez. 2023b. Supplemental tables and figures. https://doi.org/10.6084/m9.figshare.22730858.

Horst, E. A., S. K. Kvidera, and L. H. Baumgard. 2021. Invited review: The influence of immune activation on transition cow health and performance-A critical evaluation of traditional dogmas. J. Dairy Sci. 104:8380–8410. https://doi.org/10.3168/jds.2021-20330.

Hostens, M., V. Fievez, J. L. M. R. Leroy, J. Van Ranst, B. Vlaeminck, and G. Opsomer. 2012. The fatty acid profile of subcutaneous and abdominal fat in dairy cows with left displacement of the abomasum. J. Dairy Sci. 95:3756–3765. https://doi.org/10.3168/jds.2011-5092.

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.

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

Itle, A. J., J. M. Huzzey, D. M. Weary, and M. A. G. von Keyserlingk. 2015. Clinical ketosis and standing behavior in transition cows. J. Dairy Sci. 98:128–134. https://doi.org/10.3168/jds.2014-7932.

Jaeger, B. 2017. r2glmm: Computes R Squared for Mixed (Multilevel) Models. R package version 0.1.2. https://CRAN.R-project.org/package=r2glmm.

Jawor, P. E., J. M. Huzzey, S. J. LeBlanc, and M. A. G. von Keyserlingk. 2012. Associations of subclinical hypocalcemia at calving with milk yield, and feeding, drinking, and standing behaviors around parturition in Holstein cows. J. Dairy Sci. 95:1240–1248. https://doi.org/10.3168/jds.2011-4586.

Jenkins, N. T., G. Peña, C. Risco, C. C. Barbosa, A. Vieira-Neto, and K. N. Galvão. 2015. Utility of inline milk fat and protein ratio to diagnose subclinical ketosis and to assign propylene glycol treatment in lactating dairy cows. Can. Vet. J. 56:850–854.

Ježek, J., M. R. Cincović, M. Nemec, B. Belić, R. Djoković, M. Klinkon, and J. Starič. 2017. Beta-hydroxybutyrate in milk as screening test for subclinical ketosis in dairy cows. Pol. J. Vet. Sci. 20:507–512. https://doi.org/10.1515/pjvs-2017-0061.

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.

Kaufman, E. I., S. J. LeBlanc, B. W. McBride, T. F. Duffield, and T. J. DeVries. 2016. Short communication: Association of lying behavior and subclinical ketosis in transition dairy cows. J. Dairy Sci. 99:7473–7480. https://doi.org/10.3168/jds.2016-11185.

Kok, A., A. T. M. van Knegsel, C. E. van Middelaar, H. Hogeveen, B. Kemp, and I. J. M. de Boer. 2015. Technical note: Validation of sensor-recorded lying bouts in lactating dairy cows using a 2-sensor approach. J. Dairy Sci. 98:7911–7916. https://doi.org/10.3168/jds.2015-9554.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest package: 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.

Liboreiro, D. N., K. S. Machado, P. R. B. Silva, M. M. Maturana, T. K. Nishimura, A. P. Brandão, M. I. Endres, and R. C. Chebel. 2015. Characterization of peripartum rumination and activity of cows diagnosed with metabolic and uterine diseases. J. Dairy Sci. 98:6812–6827. https://doi.org/10.3168/jds.2014-8947.

Lin, Y., X. Sun, X. Hou, B. Qu, X. Gao, and Q. Li. 2016. Effects of glucose on lactose synthesis in mammary epithelial cells from dairy cow. BMC Vet. Res. 12:81. https://doi.org/10.1186/s12917-016-0704-x.

Lüdecke, D. 2022. sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.1. https://CRAN.R-project.org/package=sjPlot.

Lüdecke, D., M. S. Ben-Shachar, I. Patil, P. Waggoner, and D. Makowski. 2021. performance: An R package for assessment, comparison and testing of statistical models. J. Open Source Softw. 6:3139. https://doi.org/10.21105/joss.03139.

Mann, S., D. V. Nydam, A. L. Lock, T. R. Overton, and J. 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.

Mäntysaari, P., E. A. Mäntysaari, T. Kokkonen, T. Mehtiö, S. Kajava, C. Grelet, P. Lidauer, and M. H. Lidauer. 2019. Body and milk traits as indicators of dairy cow energy status in early lactation. J. Dairy Sci. 102:7904–7916. https://doi.org/10.3168/jds.2018-15792.

Maurice-Van Eijndhoven, M. H. T., H. Soyeurt, F. Dehareng, and M. P. L. Calus. 2013. Validation of fatty acid predictions in milk using mid-infrared spectrometry across cattle breeds. Animal 7:348–354. https://doi.org/10.1017/S1751731112001218.

Meikle, A., M. Kulcsar, Y. Chilliard, H. Febel, C. Delavaud, D. Cavestany, and P. Chilibroste. 2004. Effects of parity and body condition at parturition on endocrine and reproductive parameters of the cow. Reproduction 127:727–737. https://doi.org/10.1530/rep.1.00080.

Menichetti, B. T., J. M. Piñeiro, A. A. Barragan, A. E. Relling, A. Garcia-Guerra, and G. M. Schuenemann. 2020. Association of prepartum lying time with nonesterified fatty acids and stillbirth in prepartum dairy heifers and cows. J. Dairy Sci. 103:11782–11794. https://doi.org/10.3168/jds.2020-18629.

Merriam-Webster. 2022. Dictionary by Merriam-Webster: America's Most-Trusted Online Dictionary. Accessed Dec. 15, 2022. https://www.merriam-webster.com/.

Meyer, F., and V. Perrier. 2022. esquisse: Explore and Visualize Your Data Interactively. R package version 1.1.0. https://CRAN.R-project.org/package=esquisse.

Min, S. H. 2022. smplot: smplot - A Package for Statistical Data Visualization. R package version 0.1.0. https://www.smin95.com.

Mullins, I. L., C. M. Truman, M. R. Campler, J. M. Bewley, and J. H. C. Costa. 2019. Validation of a commercial automated body condition scoring system on a commercial dairy farm. Animals (Basel) 9:287. https://doi.org/10.3390/ani9060287.

Najm, N.-A., L. Zimmermann, O. Dietrich, A. Rieger, R. Martin, and H. Zerbe. 2020. Associations between motion activity, ketosis risk and estrus behavior in dairy cattle. Prev. Vet. Med. 175:104857. https://doi.org/10.1016/j.prevetmed.2019.104857.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4:133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x.

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.

Ouweltjes, W., B. Beerda, J. J. Windig, M. P. L. Calus, and R. F. Veerkamp. 2007. Effects of management and genetics on udder health and milk composition in dairy cows. J. Dairy Sci. 90:229–238. https://doi.org/10.3168/jds.S0022-0302(07)72624-3.

Ouweltjes, W., M. Spoelstra, B. Ducro, Y. de Haas, and C. Kamphuis. 2021. A data-driven prediction of lifetime resilience of dairy cows using commercial sensor data collected during first lactation. J. Dairy Sci. 104:11759–11769. https://doi.org/10.3168/jds.2021-20413.

Pedersen, T. L. 2020. patchwork: The Composer of Plots. R package version 1.0.1. https://CRAN.R-project.org/package=patchwork.

Piantoni, P., C. M. Ylioja, and M. S. Allen. 2015. Feed intake is related to changes in plasma nonesterified fatty acid concentration and hepatic acetyl CoA content following feeding in lactating dairy cows. J. Dairy Sci. 98:6839–6847. https://doi.org/10.3168/jds.2014-9085.

Piechotta, M., W. Mysegades, U. Ligges, J. Lilienthal, A. Hoeflich, A. Miyamoto, and H. Bollwein. 2015. Antepartal insulin-like growth factor 1 and insulin-like growth factor binding protein 2 concentrations are indicative of ketosis in dairy cows. J. Dairy Sci. 98:3100–3109. https://doi.org/10.3168/jds.2014-8885.

Pires, J. A., T. Larsen, and C. Leroux. 2022. Milk metabolites and fatty acids as noninvasive biomarkers of metabolic status and energy balance in early-lactation cows. J. Dairy Sci. 105:201–220. https://doi.org/10.3168/jds.2021-20465.

Pires, J. A. A., C. Delavaud, Y. Faulconnier, D. Pomiès, and Y. Chilliard. 2013. Effects of body condition score at calving on indicators of fat and protein mobilization of periparturient Holstein-Friesian cows. J. Dairy Sci. 96:6423–6439. https://doi.org/10.3168/jds.2013-6801.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Renaud, D. L., D. F. Kelton, and T. F. Duffield. 2019. Short communication: Validation of a test-day milk test for β-hydroxybutyrate for identifying cows with hyperketonemia. J. Dairy Sci. 102:1589–1593. https://doi.org/10.3168/jds.2018-14778.

Rinker, T. W., and D. Kurkiewicz. 2017. pacman: Package Management for R. version 0.5.0. Buffalo, NY. http://github.com/trinker/pacman.

Rodriguez-Jimenez, S., K. J. Haerr, E. Trevisi, J. J. Loor, F. C. Cardoso, and J. S. Osorio. 2018. Prepartal standing behavior as a parameter for early detection of postpartal subclinical ketosis as-

sociated with inflammation and liver function biomarkers in peripartal dairy cows. J. Dairy Sci. 101:8224–8235. https://doi.org/10.3168/jds.2017-14254.
Rohart, F., B. Gautier, A. Singh, and K.-A. Lê Cao. 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. PLOS Comput. Biol. 13:e1005752. https://doi.org/10.1371/journal.pcbi.1005752.
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.
Seely, C. R., K. D. Bach, D. M. Barbano, and J. A. McArt. 2021. Effect of hyperketonemia on the diurnal patterns of energy-related blood metabolites in early-lactation dairy cows. J. Dairy Sci. 104:818–825. https://doi.org/10.3168/jds.2020-18930.
Seely, C. R., K. D. Bach, D. M. Barbano, and J. A. A. McArt. 2022. Diurnal variation of milk fatty acids in early-lactation Holstein cows with and without hyperketonemia. Animal 16:100552 https://doi.org/10.1016/j.animal.2022.100552.
Signorell, A. 2020. DescTools: Tools for Descriptive Statistics. R package version 0.99.38. https://CRAN.R-project.org/package=DescTools.
Slowikowski, K. 2020. ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.8.2. https://CRAN.R-project.org/package=ggrepel.
Soriani, N., E. Trevisi, and L. Calamari. 2012. Relationships between rumination time, metabolic conditions, and health status in dairy cows during the transition period. J. Anim. Sci. 90:4544–4554. https://doi.org/10.2527/jas.2011-5064.
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.
Stangaferro, M. L., R. Wijma, L. S. Caixeta, M. A. Al-Abri, and J. O. Giordano. 2016. Use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part III. Metritis. J. Dairy Sci. 99:7422–7433. https://doi.org/10.3168/jds.2016-11352.
Tatone, E. H., T. F. Duffield, S. J. LeBlanc, T. J. DeVries, and J. L. Gordon. 2017. Investigating the within-herd prevalence and risk factors for ketosis in dairy cattle in Ontario as diagnosed by the test-day concentration of β-hydroxybutyrate in milk. J. Dairy Sci. 100:1308–1318. https://doi.org/10.3168/jds.2016-11453.
Tobin, C. 2020. ggthemr: Themes for 'ggplot2'. R package version 1.1.0.
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 Dixhoorn, I. D. E., R. M. de Mol, J. T. N. van der Werf, S. van Mourik, and C. G. van Reenen. 2018. Indicators of resilience during the transition period in dairy cows: A case study. J. Dairy Sci. 101:10271–10282. https://doi.org/10.3168/jds.2018-14779.
van Hoeij, R. J., A. Kok, R. M. Bruckmaier, M. J. Haskell, B. Kemp, and A. T. M. van Knegsel. 2019. Relationship between metabolic status and behavior in dairy cows in week 4 of lactation. Animal 13:640–648. https://doi.org/10.1017/S1751731118001842.
Wagner, N., V. Antoine, M.-M. Mialon, R. Lardy, M. Silberberg, J. Koko, and I. Veissier. 2020. Machine learning to detect behavioural anomalies in dairy cows under subacute ruminal acidosis. Comput. Electron. Agric. 170:105233. https://doi.org/10.1016/j.compag.2020.105233.
Waring, E., M. Quinn, A. McNamara, E. A. de la Rubia, H. Zhu, and S. Ellis. 2021. skimr: Compact and Flexible Summaries of Data. R package version 2.1.3. https://CRAN.R-project.org/package=skimr.
Wathes, D. C., F. Becker, L. Buggiotti, M. A. Crowe, C. Ferris, L. Foldager, C. Grelet, M. Hostens, K. L. Ingvartsen, C. Marchitelli, M. Salavati, F. Signorelli, H. Takeda, L. Tang, and Z. Cheng, and the GplusE Consortium. 2021. Associations between circulating IGF-1 concentrations, disease status and the leukocyte transcriptome in early lactation dairy cows. Ruminants 1:147–177. https://doi.org/10.3390/ruminants1020012.
Wathes, D. C., Z. Cheng, N. Bourne, V. J. Taylor, M. P. Coffey, and S. Brotherstone. 2007. Differences between primiparous and multiparous dairy cows in the inter-relationships between metabolic traits, milk yield and body condition score in the periparturient period. Domest. Anim. Endocrinol. 33:203–225. https://doi.org/10.1016/j.domaniend.2006.05.004.
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.
Wickham, H., and D. Seidel. 2022. scales: Scale Functions for Visualization. R package version 1.2.0. https://CRAN.R-project.org/package=scales.
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.
Yang, W., B. Zhang, C. Xu, H. Zhang, and C. Xia. 2019. Effects of ketosis in dairy cows on blood biochemical parameters, milk yield and composition, and digestive capacity. J. Vet. Res. (Pulawy) 63:555–560. https://doi.org/10.2478/jvetres-2019-0059.
Zarrin, M., L. Grossen-Rösti, R. M. Bruckmaier, and J. J. Gross. 2017. Elevation of blood β-hydroxybutyrate concentration affects glucose metabolism in dairy cows before and after parturition. J. Dairy Sci. 100:2323–2333. https://doi.org/10.3168/jds.2016-11714.