JDS Communications®
TBC; TBC

https://doi.org/10.3168/jdsc.2026-1074
Short Communication
Health, Welfare, and Behavior

Predicting classes of β-Hydroxybutyrate concentrations in blood of dairy cows from milk mid-infrared spectra: application of a weighted discriminant model addressing class imbalance

M. Lesnoff,[1,2,3]† C. Grelet,[4]†* V. Wolf,[5] L. Millot,[5] A. Köck,[6] J. Leblois,[7,8] B. Bapst,[9] M. A. Crowe,[10] T. Larsen,[11] D. C. Wathes,[12] C. P. Ferris,[13] K. L. Ingvartsen,[11] C. Marchitelli,[14] F. Becker,[15] J. Hummel,[16] M. Gelé,[17] H. Soyeurt,[18] N. Gengler,[18] and F. Dehareng[4]

Abstract: Monitoring biomarkers in milk or blood is widely used to assess dairy cow health and to detect diseases such as (sub)clinical ketosis. Blood β-hydroxybutyrate (BHB) is a key indicator of metabolic imbalance, and its prediction from milk mid-infrared (FT-MIR) spectra has been extensively explored. However, most of the predictive models applied in spectrometry fail to achieve acceptable accuracy due to the highly skewed distribution of BHB concentrations (many blood samples have low concentrations while a few have high concentrations). Imbalanced distributions complicate both quantitative and class modeling. Thus, the objective of this work was to undertake discrete predictions of BHB (low versus high level) by implementing a specific weighting in a usual discrimination algorithm to overcome the class imbalance between healthy cows and those suffering of hyperketonemia. The threshold of 1.2 mmol/l was used to distinguish low from high BHB, i.e., absence or presence of hyperketonemia. The cleaned data set combined 4,221 records comprising milk FT-MIR spectra and blood BHB content, that were collected across 10 countries between 2013 and 2024, and in multiple breeds, and numerous research projects. Milk MIR spectra (selected wavenumber regions, second derivative Savitzky–Golay preprocessing) served as predictors. A replicated external herd validation strategy was implemented. The weighted model outperformed the standard model, achieving average intra-class error rates of around 18% on external test sets. It reached 82% accuracy, 83% sensitivity and 81% specificity, with these results comparable to or better than previously published discriminant models, despite using only raw MIR spectra. These findings demonstrate that the use of weighting within discriminant methods is a valid option for the class prediction of asymmetrically distributed health biomarkers.

Measuring biomarkers in milk or in blood to provide information on disease status of dairy cows is a common practice, both for routine herd management and genetic studies. Among others, the concentration of blood β-Hydroxybutyrate (**BHB**) is routinely used as an indicator of physiological imbalance and hyperketonemia (Moyes et al., 2013; Suthar et al., 2013; McArt et al., 2015). For rapid, large scale and cost-effective screening of BHB, its prediction from Fourier transform mid-infrared (**FT-MIR**) spectrum of milk samples has been widely studied. Most previous studies have attempted to predict BHB concentrations using partial least square regression (**PLS**) models (Belay et al., 2017; Benedet, Franzoi, et al., 2019; Grelet et al., 2019; Kostensalo et al., 2023; Luke et al., 2019; Walleser et al., 2021), with the external validation coefficients of determination of these models ranging from 0.21 to 0.70. Other quantitative predictive algorithms, such as elastic net, random forest, gradient boosting, artificial neural networks, and stacking ensembles have also been used, with comparable predictive performances (Giannuzzi et al., 2023; Pralle et al., 2018; Walleser et al., 2021). However, an alternative approach is to seek to identify if a cow has either a low or a high level of BHB as this discriminates between absence vs. presence of hyperketonemia. Therefore, transforming the quantitative concentrations of biomarkers into discrete variables which separate cows with low and high concentrations through a threshold, and using predictive discrimination algorithm, would be very useful (Luke et al., 2019).

Irrespective of methodology adopted, previous authors have highlighted difficulties in modeling blood biomarkers concentrations such as BHB and Nonesterified Fatty Acids (NEFA) due

[1] SELMET, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France, [2] CIRAD, UMR SELMET, Montpellier, France, [3] ChemHouse Research Group, Montpellier, France, [4] Walloon Agricultural Research Center (CRA-W), Gembloux, Belgium, [5] CEL25-90, Roulans, France, [6] ZuchtData, Vienna, Austria, [7] Eleveo, Awé group, Ciney, Belgium, [8] EEIG European Milk Recording (EMR), Ciney, Belgium, [9] Qualitas AG, Zug, Switzerland, [10] University College Dublin (UCD), Dublin, Ireland, [11] Dept Animal and Veterinary Sciences, Aarhus University, Tjele, Denmark, [12] Royal Veterinary College (RVC), London, United Kingdom, [13] Agri-Food and Biosciences Institute (AFBI), Belfast, Northern Ireland, [14] Research Center for Animal Production and Aquaculture (CREA), Roma, Italy, [15] Institute for Farm Animal Biology (FBN), Dummerstorf, Germany, [16] University of Göttingen, Göttingen, Germany, [17] Idele, Paris, France,, [18] TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium. † The authors contributed equally to the study *Corresponding author: Clément Grelet, 24 Chaussée de Namur, 5030 Gembloux, Belgium, Phone: 0032/81/87.52.26, Fax: 0032/81/620.388, E-mail: c.grelet@cra.wallonie.be.
 Received April 23, 2026. Accepted June 18, 2026.

The list of standard abbreviations for JDSC is available at adsa.org/jdsc-abbreviations-26. Nonstandard abbreviations are available in the Notes.

to their highly right-skewed distributions (Aernouts et al., 2020; Kostensalo et al., 2023). Indeed, in practice, the majority of samples have a low BHB concentration (from healthy cows) while a few samples from (sub)-clinical ketotic cows show an exponential increase of BHB. The resulting right-skewed distribution of the concentrations is particularly difficult to model from spectral data, as it goes beyond the linear relationship of milk molecules and spectral absorbance values. When transforming BHB in a discrete variable by using a threshold (i.e., "low" vs. "high" BHB), the issue remains and is translated in a class imbalance, i.e., between numbers of healthy *vs*. cow in hyperketonemia. Moreover, this issue is observed with many other disease biomarkers, including lactoferrin (Soyeurt et al., 2020) and N-acetyl-β-d-glucosaminidase activity (Grelet et al., 2024) for mastitis, and nonesterified fatty acids as an indicator of negative energy balance (Aernouts et al., 2020). In these scenarios with imbalanced data, modeling methodologies must be adapted (Kostensalo et al., 2023; Soyeurt et al., 2020).

Thus, the objective of this work was to undertake discrete predictions of BHB (low versus high level) by implementing a specific weighting in a usual discrimination algorithm to overcome the class imbalance between healthy cows and those suffering of hyperketonemia.

Data used in this study (total of 6,059 milk and blood samples) were collected within 2 international and 7 national projects conducted between 2013 to 2024. Data were collected from 1,799 cows across 83 herds in 10 countries. The data set involved several breeds (Holstein, Braunvieh, Fleckvieh, Simmental, Brown Swiss, Montbéliardes, Abondance and crossbred cows), parity ranged from 1 to 13, while cows were between 5 to 226 d in milk.

A total of 780 samples were collected within the GplusE project (EU FP7-Project) from AFBI (Agri-Food and Biosciences Institute, UK), Aarhus University (Denmark), CREA (Research Center for Animal Production and Aquaculture, Italy), CRA-W (Walloon Agricultural Research Centre, Belgium), FBN (Leibniz Institute for Farm Animal Biology, Germany) and UCD (University College Dublin, Ireland) and in 4 commercial farms in the UK. A total of 953 samples were collected within the OptiMIR project (EU, Interreg NWE) from 3 experimental farms (Poisy and les Trinottieres, France; Neumuhle, Germany). In national projects, 420 samples were collected from 23 commercial farms in France (within project of CEL25–90), 1,017 samples were collected from 49 commercial farms in Austria by Zuchtdata (D4Dairy, FFG comet, project 872039, https://d4dairy.com/), 812 samples from 10 commercial farms involved in 2 Belgian projects (Fertiwalim and Scorewelcow), 345 samples from 10 commercial farms in German Indikuh project (funding code: 2817905815), and 1732 samples were collected in Switzerland by Qualitas AG as described in Caldeira et al. (2020).

All milk samples were collected following the guidelines provided by the International Committee for Animal Recording (ICAR Dairy Cattle Milk Recording Working Group, 2017) using ICAR approved milk samplers. Samples were collected according to either AT or A4 milk recording protocols, namely morning or evening samples, or morning and evening samples pooled into a daily sample with a similar quantity of morning and evening milk as per normal milk recording routines (i.e., 50% from morning and 50% from evening). All milk samples were stored at 4°C until analysis, with 0.02–0.03% bronopol used as a preservative. Analyses were undertaken in local laboratories using a wide range of instruments with a total of 36 spectrometers utilized to analyze the samples: 35 Foss instruments (models FT2, FT6000, FT+ and FT7; Foss, Hillerød, Denmark) and one Standard Lactoscope FT-MIR automatic (Perkin Elmer, previously Delta Instruments, Drachten, The Netherlands). All the spectra from the different instruments were standardized to be merged into a common data set following the procedure described by Grelet et al., (2015).

Blood samples were collected in the morning (within 3 h of milking) from the tail vein (coccygeal vein). Blood BHB concentration from coccygeal vein was reported to be similar with jugular vein concentration (Alvarenga et al., 2019). Blood samples within the GplusE, OptiMIR, Indikuh and Qualitas AG projects (3,810 samples) were collected in heparin coated tubes, centrifuged to harvest plasma, stored at −20°C, with plasma subsequently analyzed by chemical reference methods (e.g., GplusE samples were analyzed at Aarhus University and BHB was determined by measuring absorbance at 340 nm due to the production of NADH at alkaline pH in the presence of BHB dehydrogenase). Samples from CEL25–90, D4Dairy and the Belgian projects (2,249 samples) were analyzed for BHB using capillary strip tests WellionVet Belua and Freestyle Optium β-Ketone.

From the 6,059 records, only those with an interval of less than 2 d between milk and blood sampling were retained (693 samples removed). Furthermore, a total of 459 samples were identified with laboratory analytical issues and were removed. To avoid erroneous associations between spectra and blood samples, as well as analytical issues during analysis, a homemade fat model was applied to the spectra and the generated predictions were compared with the predictions provided by the laboratories. Records with a difference above 0.3 g/100mL (Zhang et al., 2021) between local and laboratory predictions were discarded (47 samples removed). Finally, records from 4 experiments were excluded due to errors in sampling, storage or analytical steps and the final study data set comprised of n_{tot} = 4,221 samples in 64 herds. MIR spectra (data $\boldsymbol{X}$) were used as predictors. The spectral areas were selected from ranges (968.1–1,577.5 cm^{-1}, 1,731.8–1762.6 cm^{-1}, 1791.9–1,808.9 cm^{-1} and 2,831.0–2,966.0 cm^{-1}) that excluded areas not reproducible between instruments. The spectra were preprocessed by a Savitzky-Golay filter (Luo et al., 2005) with a filter size = 9, a smoothing polynomial order = 3, and a second derivative (Figure 1).

The response variable to predict (data $\boldsymbol{y}$) was the BHB discrete level defined by 2 classes: "low" (concentration ≤ 1.2 mmol/L), and "high" (concentration > 1.2 mmol/L). Cows classified as the "high" level are considered to suffer from hyperketonemia. Cows were not clinically inspected for presence of ketosis or other diseases, therefore the classification only targets the absence vs the presence of hyperketonemia. The data set was highly imbalanced in favor of class "low," with only 8% of the samples classified as class "high."

The prediction algorithm used as a comparison basis was the simplest Partial Least Squares Discriminant Analysis (**PLSDA**), often referred as the "usual" or "normal" PLSDA as defined by (Brereton & Lloyd, 2014). It consists in first transforming the univariate class membership variable to predict to a dummy table (Y_{dummy}; i.e., if there are C classes, a table with C columns of 0/1 binary variables representing the absence/occurrence of each class). Then a PLS regression between the spectra and this dummy table is made: a multivariate PLS is run on data ($\boldsymbol{X}$, $\boldsymbol{Y}_{dummy}$) to com-

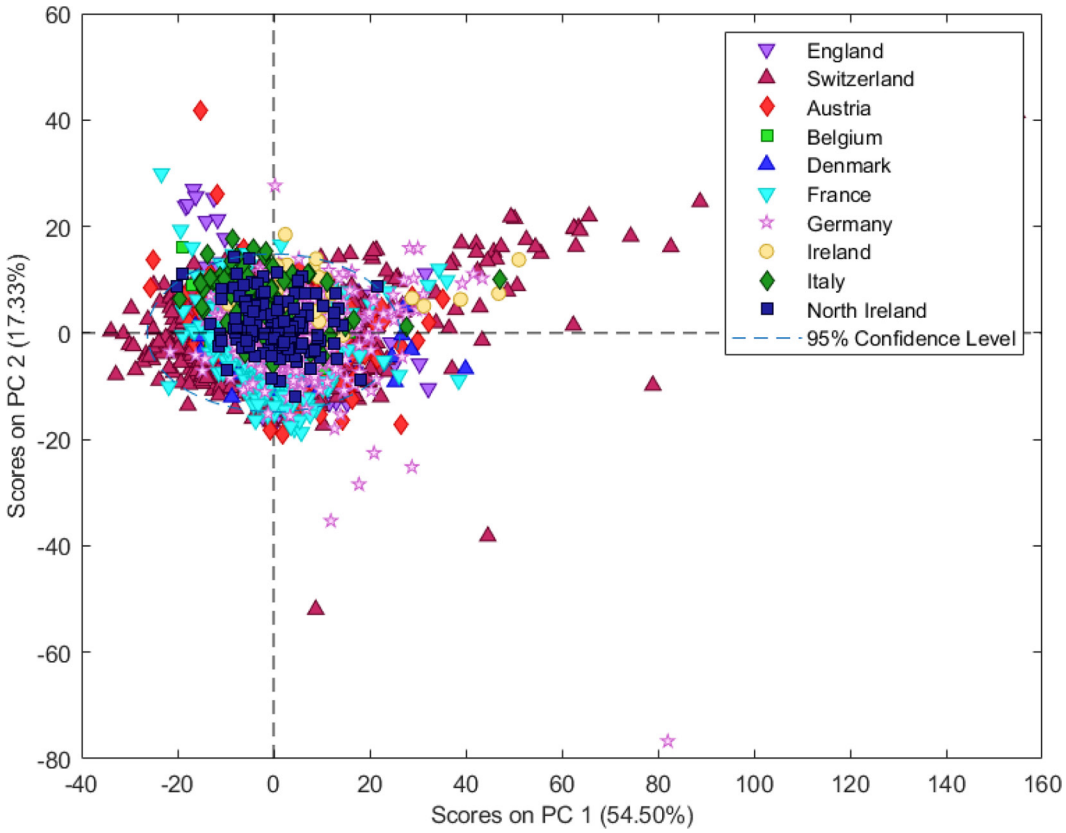

Figure 1. Principal Component Analysis on the milk mid-infrared spectra (n = 4221). Spectra are plotted after standardization, first derivative, selection of 212 repeatable wavenumbers and mean-centering and scaling of each wavenumbers.

pute the PLS-score matrix ($\boldsymbol{T}$) and a multiple linear regression (**MLR**) model is run on data $\{\boldsymbol{T}, \boldsymbol{Y}_{\text{dummy}}\}$, giving predictions $\hat{\boldsymbol{Y}}_{\text{dummy}}$. For each observation, predicting the final class by the one corresponding to the column that shows the highest value in the predicted dummy table. The usual PLSDA has the advantage to be fast, and quite efficient when the number of classes to discriminate is not too large (≤3–4 classes). Nevertheless, it is highly sensitive to class imbalance and generates biased predictions in such case: the algorithm favors systematically the prediction of the classes that are predominant in the data (Brereton & Lloyd, 2014). A simple approach to remove this bias consists in internally weighting the PLSDA (**WPLSDA**) to give the same importance to each class. Noting n the number of the samples in the training set, the principle in WPLSDA is to embed a vector of weights $\boldsymbol{w} = (w_1, \ldots w_n)$ in the PLSDA equations. A detailed description of the general WPLS algorithm used in the study is given in Lesnoff et al. (2020). To control imbalanced class data, vector $\boldsymbol{w}$ can be defined such as the sums of the weights by class are equal. In the present case of 2 classes, and noting p_{low} and p_{high} the sum of the weights w_i of the samples in class "low" and "high," respectively, vector $\boldsymbol{w}$ is defined such as $p_{low} = p_{high} = 1/2$.

A double replicated K-fold cross-validation (Filzmoser et al., 2009) was implemented to optimize the PLSDA and WPLSDA models and estimate their predictive classification error rates. The following process was repeated $n_{\text{rep}} = 100$ times:

- 21 herds (~33% of the total 64 herds) are randomly sampled and define the test set for the ongoing process. The remaining 43 herds define the training set.
- On the training set, a replicated K-Fold cross-validation (**CV**) ($K = 3$ and 50 replications) is implemented, and the optimal numbers of latent variables (**LVs**) for PLSDA and WPLSDA are selected based on the lower classification error rate averaged over the 50 replications.
- The optimal model is fitted again on the (entire) training set and used to predict the test set. The test predictions are used to compute the predictive classification error rates for the given process: global accuracy (overall percentage of well classified samples), sensitivity (percentage of good classification in actual class "high"), specificity (percentage of good classification in actual class "low"), positive predictive value (**PPV**; percentage of actual class "high" in predictions "high") and negative predicted values (**NPV** ; percentage of actual class "low" in predictions "low")

The final external predictive performances of the 2 models were assessed based on the distributions of the $n_{\text{rep}} = 100$ classification test set predictions. For imbalanced data, the naïve global classification error rate (i.e., overall % of misclassified samples, noted **ERRP**) computed on CV or test predictions is a biased estimate of the model performance since it gives too much importance to the dominant class. The arithmetical mean of the intra-class clas-

sification error rate (**M-ERRP**) equalizing the class importance, is a better choice and was used in the present study.

All the computations were implemented with the chemometric package Jchemo (Lesnoff M., 2021) in the free Julia language (Bezanson et al., 2017).

BHB values ranged from 0.09 to 5.13 mmol/L, with an average of 0.71 mmol/L, a standard deviation of 0.49 mmol/L and a skewness of 3.81. The data set was highly imbalanced in favor of class "low," with only 8% of the samples classified as class "high" (concentration > 1.2 mmol/L). This prevalence of hyperketonemia can be considered low in regards of other studies. Benedet et al. (2019) in a large review reported hyperketonemia prevalence ranging from 11.2% to 47.2%, across several countries and using similar thresholds (1.2 mmol/L in majority of cases). However, most of those studies focused on early lactation only (e.g., from 7 to 14 DIM in Vanholder et al. (2015) in which the maximum prevalence was reported) while our data set merged records with large DIM range, with an average of 54 DIM and a standard deviation of 48 DIM.

Cross-validations results on the 100 repeated training sets are presented in Figure 2. As expected for imbalanced data, PLSDA was inefficient, with M-ERRP always higher than 40%. The weighted model WPLSDA showed better performances. The optimal dimensions ranged from 5 to 20 LVs (mean = 13 LVs) with average M-ERRP around 19% (std = 1.6%).

On the test sets, WPLSDA (for the optimal number of LVs) also showed relevant performances, with average M-ERRP of 17.9% (min = 11.7%, max = 25.4%, SD = 3.9%), with average accuracy of 81.3%, average sensitivity of 83.2%, average specificity of 81.1%, averaged NPV of 98.2% and average PPV of 28.0% (Table 1). The bias of the usual PLSDA is clearly apparent in the confusion matrix (Table 1): predictions were highly favored toward class "low" independently of the true class, inducing a very low sensitivity (average = 10.2%). The classification performances of WPLSDA can therefore be considered as good to discriminate both low and high BHB records. However, the approach only partially overcomes the class imbalance issue as the PPV (the real percentage of "high" values within the total of "high" predictions) is only 28%, meaning that because of class imbalance, majority of samples being suspected of hyperketonemia are actually from healthy cows. Profile of the misclassified records should be deeply investigated, to see whether these cows are real negative or close to 1.2 mmol/L threshold and should rather be considered in a gray zone. Impact on this "over-alerting" on the practical use of such tool should also be investigated.

Table 1. Confusion matrices for PLSDA and WPLSDA (weighted model) for the blood BHB class predictions from the milk FT-MIR spectra on the external herd validation test sets. Each cell of the table represents the average number of samples over the 100 repetitions (averages were rounded to simplify the presentation)

Actual true values	PLSDA predictions			WPLSDA Predictions		
	Low	High	total	Low	High	total
Low	1278	6	1284	1037	242	1279
High	97	11	108	19	94	113
total	1375	17	1392	1056	336	1392
Spec.			99.5%			81.1%
Sens.			10.2%			83.2%
Glob. Acc.			92.6%			81.3%

PLSDA = partial least squares regression discriminant analysis; WPLSDA = weighted partial least squares regression discriminant analysis; Spec. = Specificity; Sens. = sensitivity; Glob. Acc. = global accuracy.

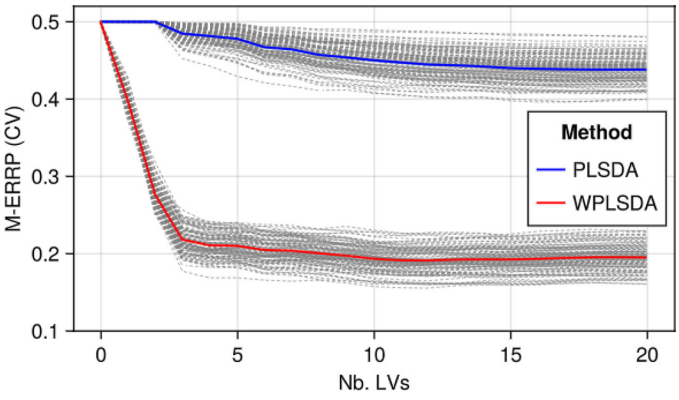

Figure 2. Cross validation results on the training set for PLSDA and WPLSDA. For each model, the dashed lines represent the 100 repetitions (for each, average of the 50 internal CV replications), and the lines in plain represent their averages. PLSDA = partial least squares discriminant analysis; WPLSDA = weighted partial least squares discriminant analysis. Nb. LVs = number of latent variables.

Only a few studies have used discrimination modeling to predict low vs. high BHB concentrations from milk MIR spectra. Gelé et al. (2015) sought to predict the ketosis risk using BHB and nonesterified fatty acids concentrations and obtained a sensitivity of 81% and a specificity of 69%. Walleser et al. (2023) obtained similar performances with deep learning, gradient boosting machine models and stacking ensembles. Luke et al. (2019) attempted to model low vs. high BHB concentrations, but due to severe imbalance between groups (high BHB represented only 2% of records) the model was statistically not significant. Pralle et al. (2018) also discriminated low vs. high BHB concentrations and obtained a sensitivity of 83% and a specificity of 81%. However, their model included not only the FT-MIR milk spectra but also many other covariables (e.g., milk fatty acids, protein percentage, fat percentage, fat:protein ratio, SCC, urea, milk acetone, milk BHB, calving date, parity, days in milk, age at first calving, previous lactation length, previous lactation 305-d mature equivalent milk, gestation length, and dry period length) making their results difficult to compare with the performance of our model. The discrimination with implemented class weighting approach proposed in the present study provided similar performance, but using only the FT-MIR spectra. The current results suggest that combining data from multiple sources and countries, as well as using weighted modeling algorithms, is a good option to predict traits with specific distributions such as BHB. The model was tested in a robust external herd validation, and results are expected to be similar in real field conditions.

References

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.

Alvarenga, E. A., G. H. F. A. Moreira, E. J. Facury Filho, F. O. P. Leme, T. F. Moreira, and A. U. Carvalho. 2019. Evaluation of different sites for blood collection for β-hydroxybutyrate assessment in dairy cows using handheld meter. Arq. Bras. Med. Vet. Zootec. 71:1425–1427. https://doi.org/10.1590/1678-4162-10164.

Belay, T. K., B. S. Dagnachew, Z. M. Kowalski, and T. Ådnøy. 2017. An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. J. Dairy Sci. 100:6312–6326. https://doi.org/10.3168/jds.2016-12252.

Benedet, A., M. Franzoi, M. Penasa, E. Pellattiero, and M. De Marchi. 2019. Prediction of blood metabolites from milk mid-infrared spectra in early-lactation cows. J. Dairy Sci. 102:11298–11307. https://doi.org/10.3168/jds.2019-16937.

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

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah. (2017). Julia: A Fresh Approach to Numerical Computing. Https://Doi.Org/10.1137/141000671, *59*(1), 65–98. https://doi.org/10.1137/141000671

Brereton, R. G., and G. R. Lloyd. 2014. Partial least squares discriminant analysis: taking the magic away. J. Chemometr. 28:213–225. https://doi.org/10.1002/cem.2609.

Caldeira, M. O., D. Dan, A. L. Neuheuser, R. Stürmlin, C. Weber, D. L. Glauser, M. Stierli, U. Schuler, J. Moll, S. Wegmann, R. M. Bruckmaier, and J. J. Gross. 2020. Opportunities and limitations of milk mid-infrared spectra-based estimation of acetone and β-hydroxybutyrate for the prediction of metabolic stress and ketosis in dairy cows. J. Dairy Res. 87:196–203. https://doi.org/10.1017/S0022029920000230.

Filzmoser, P., B. Liebmann, and K. Varmuza. 2009. Repeated double cross validation. J. Chemometr. 23:160–171. https://doi.org/10.1002/cem.1225.

Gelé, M., M. Ferrand-Calmels, L. Brun-Lafleur, A. Werner, and F. Gollé-Leidreiter. 2015. Predicting the risk of ketosis using mid infrared spectrometry. ICAR Tech. Ser. 19–24. http://www.icar.org/wp-content/uploads/2015/11/ICAR-Technical-Series-19-Krakow-2015-Proceedings.pdf.

Giannuzzi, D., L. F. M. Mota, S. Pegolo, F. Tagliapietra, S. Schiavon, L. Gallo, P. A. Marsan, E. Trevisi, and A. Cecchinato. 2023. Prediction of detailed blood metabolic profile using milk infrared spectra and machine learning methods in dairy cattle. J. Dairy Sci. 106:3321–3344. https://doi.org/10.3168/jds.2022-22454.

Grelet, C., J. A. Fernández Pierna, P. Dardenne, V. Baeten, and F. Dehareng. 2015. Standardization of milk mid-infrared spectra from a European dairy network. J. Dairy Sci. 98:2150–2160. https://doi.org/10.3168/jds.2014-8764.

Grelet, C., T. Larsen, M. A. Crowe, D. C. Wathes, C. P. Ferris, K. L. Ingvartsen, C. Marchitelli, F. Becker, A. Vanlierde, J. Leblois, U. Schuler, F. J. Auer, A. Köck, L. Dale, J. Sölkner, O. Christophe, J. Hummel, A. Mensching, J. A. Fernández Pierna, H. Soyeurt, M. Calmels, R. Reding, M. Gelé, Y. Chen, N. Gengler, and F. Dehareng. 2024. Prediction of key milk biomarkers in dairy cows through milk mid-infrared spectra and international collaborations. J. Dairy Sci. 107:1669–1684. https://doi.org/10.3168/jds.2023-23843.

Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K. L. Ingvartsen, M. Crowe, M. T. Sorensen, E. Froidmont, C. P. Ferris, C. Marchitelli, F. Becker, T. Larsen, F. Carter, F. Dehareng, A. Santoro, C. Byrne, P. Rudd, and R. O'Flaherty. 2019. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. Animal 13:649–658. https://doi.org/10.1017/S1751731118001751.

ICAR Dairy Cattle Milk Recording Working Group. (2017). *Section 2 -Guidelines for Dairy Cattle Milk Recording (*https://www.icar.org/Guidelines/02-Overview-Cattle-Milk-Recording.pdf*)*.

Kostensalo, J., M. Lidauer, B. Aernouts, P. Mäntysaari, T. Kokkonen, P. Lidauer, and T. Mehtiö. 2023. Short communication: Predicting blood plasma nonesterified fatty acid and beta-hydroxybutyrate concentrations from cow milk—addressing systematic issues in modelling. Animal 100912:100912. https://doi.org/10.1016/j.animal.2023.100912.

Lesnoff, M. (2021). *Jchemo: Chemometrics and machine learning for high-dimensional data with Julia*. UMR SELMET, Univ Montpellier, CIRAD, INRA, Institut Agro. https://github.com/mlesnoff/Jchemo.jl

Lesnoff, M., M. Metz, and J. M. Roger. 2020. Comparison of locally weighted PLS strategies for regression and discrimination on agronomic NIR data. J. Chemometr. 34:e3209. https://doi.org/10.1002/cem.3209.

Luke, T. D. W., S. Rochfort, W. J. Wales, V. Bonfatti, L. Marett, and J. E. Pryce. 2019. Metabolic profiling of early-lactation dairy cows using milk mid-infrared spectra. J. Dairy Sci. 102:1747–1760. https://doi.org/10.3168/jds.2018-15103.

Luo, J., K. Ying, and J. Bai. 2005. Savitzky–Golay smoothing and differentiation filter for even number data. Signal Processing 85:1429–1434. https://doi.org/10.1016/j.sigpro.2005.02.002.

McArt, J. A. A., D. V. Nydam, and M. W. Overton. 2015. Hyperketonemia in early lactation dairy cattle: A deterministic estimate of component and total cost per case. J. Dairy Sci. 98:2043–2054. https://doi.org/10.3168/jds.2014-8740.

Moyes, K. M., T. Larsen, and K. L. Ingvartsen. 2013. Generation of an index for physiological imbalance and its use as a predictor of primary disease in dairy cows during early lactation. J. Dairy Sci. 96:2161–2170. https://doi.org/10.3168/jds.2012-5646.

Pralle, R. S., K. W. Weigel, and H. M. White. 2018. Predicting blood β-hydroxybutyrate using milk Fourier transform infrared spectrum, milk composition, and producer-reported variables with multiple linear regression, partial least squares regression, and artificial neural network. J. Dairy Sci. 101:4378–4387. https://doi.org/10.3168/jds.2017-14076.

Soyeurt, H., C. Grelet, S. McParland, M. Calmels, M. Coffey, A. Tedde, P. Delhez, F. Dehareng, and N. Gengler. 2020. A comparison of 4 different machine learning algorithms to predict lactoferrin content in bovine milk from mid-infrared spectra. J. Dairy Sci. 103:11585–11596. https://doi.org/10.3168/jds.2020-18870.

Suthar, V. S., J. Canelas-Raposo, A. Deniz, and W. Heuwieser. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. J. Dairy Sci. 96:2925–2938. https://doi.org/10.3168/jds.2012-6035.

Vanholder, T., J. Papen, R. Bemers, G. Vertenten, and A. C. B. Berge. 2015. Risk factors for subclinical and clinical ketosis and association with production parameters in dairy cows in the Netherlands. J. Dairy Sci. 98:880–888. https://doi.org/10.3168/jds.2014-8362.

Walleser, E., J. F. Mandujano Reyes, K. Anklam, M. Höltershinken, P. Hertel-Boehnke, and D. Döpfer. 2021. Developing a predictive model for beta-hydroxybutyrate and non-esterified fatty acids using milk fourier-transform infrared spectroscopy in dairy cows. Prev. Vet. Med. 197:105509. https://doi.org/10.1016/j.prevetmed.2021.105509.

Walleser, E., J. F. M. Reyes, K. Anklam, R. S. Pralle, H. M. White, S. Unger, N. Panne, M. Kammer, S. Plattner, and D. Döpfer. 2023. Novel prediction models for hyperketonemia using bovine milk Fourier-transform infrared spectroscopy. Prev. Vet. Med. 213:105860. https://doi.org/10.1016/j.prevetmed.2023.105860.

Zhang, L., Li, C., Dehareng, F., Grelet, C., Colinet, F., Gengler, N., Brostaux, Y., & Soyeurt, H. (2021). Appropriate Data Quality Checks Improve the Reliability of Values Predicted from Milk Mid-Infrared Spectra. *Animals 2021, Vol. 11, Page 533, 11*(2), 533. https://doi.org/10.3390/ani11020533

Notes

C. Grelet, https://orcid.org/0000-0003-3313-485X
A. Köck, https://orcid.org/0000-0002-1611-8017
B. Bapst, https://orcid.org/0000-0001-6807-6712
M. A. Crowe, https://orcid.org/0000-0003-1999-815X
T. Larsen, https://orcid.org/0000-0003-3058-5938
D. C. Wathes, https://orcid.org/0000-0002-8206-6091
C. P. Ferris, https://orcid.org/0000-0001-8623-5983
K. L. Ingvartsen, https://orcid.org/0000-0003-2372-563X
C. Marchitelli, https://orcid.org/0000-0001-5160-0717
F. Becker, https://orcid.org/0000-0001-9832-3091
J. Hummel, https://orcid.org/0000-0002-8876-7745
M. Gelé, https://orcid.org/0000-0002-8186-1851
H. Soyeurt, https://orcid.org/0000-0001-9883-9047
N. Gengler, https://orcid.org/0000-0002-5981-5509
F. Dehareng https://orcid.org/0000-0002-6733-4334

The authors gratefully thank the CRA-W milk lab team for their precious work.

OptiMIR project has received funding from INTERREG IVB NEW and the Walloon Region. GplusE project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. IndiKuh project 'Evaluation of Animal Welfare in Dairy Farming—Indicators for the Metabolism and Feeding' (IndiKuh, funding code: 2817905815) was supported by funds of the German Federal Ministry of Food and Agriculture (BMEL).

COMET-Project D4Dairy (Digitalisation, Data integration, Detection and Decision support in Dairying, project number: 872039), which was supported by the Federal Ministry of Mobility, Innovation, and Infrastructure (BMINI, Vienna, Austria) and the Federal Ministry of Economy, Energy, and Tourism (BMWET, Vienna, Austria) from the Republic of Austria, and the provinces of Lower Austria and Vienna in the framework of the Competence Centers for Excellent Technologies (COMET). The COMET program is handled by the Austrian Research Promotion Agency (FFG, Vienna, Austria; grant number 872039).

The authors have not stated any conflicts of interest. The authors did not use generative artificial intelligence.

The experiments within the GplusE project were conducted in accordance with ARRIVE guidelines and the EU Directive 2010/63/EU for animal experiments. Samples collected in commercial dairy farms were taken by farm veterinaries within the context of routine health diagnosis.