



Can we get information on dairy cow chronic stress biomarkers from milk mid-infrared spectra?

C. Grelet,^{1*} H. Simon,^{1*} J. Leblois,² O. Christophe,¹ M. Jattiot,³ N. Gaudillère,⁴ R. Reding,⁵ J. Wavreille,¹ E. J. P. Strang,⁶ F. J. Auer,⁷ K. Goossens,⁸ E. Chevaux,⁹ N. Gengler,¹⁰ and F. Dehareng^{1†}

¹Walloon Agricultural Research Center, 5030 Gembloux, Belgium

²Walloon Breeders Association Group, 5590 Ciney, Belgium

³Innoval, Noyal-sur-Vilaine, 35530 Bretagne, France

⁴Eliaance, Pôle Données et Elevage, 75795 Paris, France

⁵Convis, 9042 Ettelbruck, Luxembourg

⁶LKV Baden-Württemberg, 70190 Stuttgart, Germany

⁷LKV-Austria, 1200 Vienna, Austria

⁸ILVO, 9820 Merelbeke, Belgium

⁹Lallemand Animal Nutrition, 31702 Blagnac, France

¹⁰Gembloux Agro-Bio Tech (Ulg-GxABT), University of Liège, 5030 Gembloux, Belgium

ABSTRACT

In recent years, animal welfare considerations have increased for citizens and public authorities. Within the several dimensions of welfare, chronic stress is particularly difficult to measure objectively, while in its chronic form, it may cause metabolic, inflammatory, and infectious diseases, fertility problems, or lower milk production. Therefore, the goal of this work was to evaluate the potential of milk mid-infrared (MIR) spectra for predicting at the individual cow level the content of 2 chronic stress biomarkers: hair cortisol and blood fructosamine. A dataset was specifically created by organizing a dedicated large-scale sampling. Sampling was performed in commercial dairy farms by several regional milk recording organizations in Austria, Belgium, France, Germany, and Luxembourg for a total of 1,412 individual cow records. Each of the 1,412 cows was sampled 1 time for milk, hair, and blood. Milk MIR analysis was conducted locally on a total of 30 standardized spectrometers. After removing the spectral and reference missing values and cleaning data, the final dataset comprised 1,071 hair cortisol values and 940 blood fructosamine values. Qualitative models were tested to discriminate low versus high contents of hair cortisol and blood fructosamine for each individual dairy cow from the milk MIR spectra through the use of partial least squares discriminant analysis. Four quantitative modeling strategies were also tested to predict the exact quantitative value of the 2 biomarkers. All models were evaluated in an external herd valida-

tion process by iteratively excluding 33% of herds (i.e., 26 herds out of 78). The best model discriminating high hair cortisol values for individual dairy cows showed a global accuracy of 71%, with a specificity of 74% and a sensitivity of 61%, and this model included MIR spectra, with milk yield, parity, square of parity, DIM, square of DIM, and breed as predictors. The best discrimination of high fructosamine values demonstrated a global accuracy of 73%, with a specificity of 75% and a sensitivity of 67% and was based on the combination of 2 models. When testing the 4 quantitative methodologies, the best R^2 were 0.13 for hair cortisol and 0.2 for blood fructosamine in external herd validation, showing a poor capacity of milk MIR spectra to predict the exact quantitative value of the 2 biomarkers. While the quantitative models did not supply satisfying results, the qualitative models provided accuracies of 71% and 73% in a robust external herd validation, respectively. Therefore, the accuracy obtained, even if not very high, can be considered very positive given the difficulty of predicting these totally indirect biomarkers of chronic stress. While there are currently no possibilities to get objective information on chronic stress in a routine frame, the prediction of both hair cortisol and blood fructosamine content, at the individual cow level, may facilitate large-scale chronic stress assessment and improvement by providing objective, cheap, and quantitative information.

Key words: Fourier transform mid-infrared spectrometry, cortisol, fructosamine, welfare

INTRODUCTION

In recent years, animal welfare considerations have increased for citizens and public authorities. In European Union countries, 91% of citizens consider it important

Received February 20, 2025.

Accepted June 25, 2025.

*These authors contributed equally to this work.

†Corresponding author: f.dehareng@cra.wallonie.be

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-25. Nonstandard abbreviations are available in the Notes.

to protect the welfare of farm animals, and most of them (60%) are also interested in labels identifying products from welfare-friendly farming systems and are willing to pay more for these products (Europa, 2023). Evaluating welfare is therefore of growing importance in the dairy sector but requires appropriate phenotypes, that is, observable and measurable characteristics or traits providing indication on cows' welfare (Brito et al., 2020). However, animal welfare is multidimensional and usually measured at the herd level by using complex and exhaustive methodologies. Particularly, objective assessment of fear and distress endured by animals needs tedious behavioral measurements (Welp et al., 2004; MacKay et al., 2014; Proudfoot et al., 2018). Nonetheless, the stress experienced by dairy cattle, and particularly in its chronic form, when stress is repeated or persists in the long term, is negative for cattle welfare. Conventionally, stress is defined as a process arising from real or perceived environmental demands that can be appraised as threatening or benign, depending on the availability of adaptive coping resources to an individual (McEwen and Gianaros, 2010). The stress mechanism can be adaptive in the short term (acute stress), as it promotes short-term adaptation and generally allows a quick and complete adaptation to recover physiological balance. When the stress turns chronic—repeated or continuous in the long term—without possible adaptation, the stress can be maladaptive by promoting wear-and-tear on the body and brain, long-term dysregulation of allostasis, and compromising stress resiliency and health (Fink, 2009; McEwen and Gianaros, 2010). It may cause serious problems, such as alterations in the psychological, immunological, metabolic, and endocrine status of cows (Trevisi and Bertoni, 2009). This induces metabolic, inflammatory, and infectious diseases (Moberg et al., 1980; Romero, 2004), fertility problems (Walker et al., 2008), growth disturbances (Elsasser et al., 2020), weight loss (Mormède et al., 2007), and lower milk production (Tallo-Parra et al., 2018). Chronic stress is detrimental to the welfare of cows, farm productivity, and economics, and harms the public perception of dairy production.

For these reasons, efforts to develop an objective, quantitative, and large-scale methodology to assess chronic stress in dairy cows would be highly relevant. Such advances would enable us to better objectify, monitor, and improve chronic stress in herds but also to conduct large-scale studies to better understand stress and its contributory factors.

During a previous phase of this work, hair cortisol and blood fructosamine were identified among several potential biomarkers as relevant chronic stress biomarkers (Grelet et al., 2022). Biomarkers are defined as a characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or biological responses to an exposure or intervention (FDA-NIF

Biomarker Working Group, 2025). Being an indication of chronic stress, biomarkers (hair cortisol and blood fructosamine) contain information on them but are not a real diagnosis and may vary according to other factors. Cortisol increases in blood due to the activation of the hypothalamic-pituitary-adrenocortical (HPA) axis (Mormède et al., 2007) and accumulates continuously in hair over time through vascular supply, from the surrounding tissues and fluids, or is synthesized by the hair follicle itself (Ito et al., 2005; Heimbürge et al., 2019). Therefore, hair cortisol highlights long-term HPA axis activation and is considered a useful marker to assess repeated or long-term stress over the last few weeks (Meyer and Novak, 2012; Tallo-Parra et al., 2017). It has been observed to increase in case of diseases (Comin et al., 2013; Burnett et al., 2015), change of environment (Comin et al., 2011; Peric et al., 2017), or increase of stocking density (Schubach et al., 2017). In the previous study of this work, dairy cows placed under stressful conditions over 4 wk, through overstocking and unusual events (i.e., transfers to unfamiliar barns and diffusion of stressful noises), exhibited significantly higher hair cortisol content compared with a control group, with 36.2 versus 21.5 pg/mg, respectively (Grelet et al., 2022). However, its extremely low concentration (i.e., ~20 ng/g of hair) and the tedious and laborious laboratory treatment of hairs preceding the analytical measurement make its use impossible for routine and cheap assessment of chronic stress.

Fructosamine is a general term for molecules resulting from protein glycosylation (Armbruster, 1987). More specifically, fructosamine is a keto-amine rearrangement product of sugar molecules, such as glucose or fructose, and the amino group of protein molecules, usually albumin or other plasma proteins such as Ig (Megahed et al., 2018). Blood fructosamine content, therefore, depends on glucose and total protein blood concentration and composition, as well as the plasma protein turnover rate (Armbruster, 1987; Kawamoto et al., 1992). Due to the life span of proteins (e.g., 16.5 d for albumin), it is considered that fructosamine provides a retrospective record of blood glucose but also protein levels during the previous 1 to 3 wk (Rahko et al., 1987; Jensen et al., 1993). Fructosamine will, therefore, increase with prolonged hyperglycemia or prolonged hyperproteinemia, and decrease with prolonged hypoglycemia or increased protein turnover (Megahed et al., 2018). Unlike glucose, fructosamine concentration is not affected by diurnal variations and not subject to acute changes potentially induced by sampling itself (Jensen et al., 1993). Caré et al. (2018) and Mostafavi et al. (2014) reported that low fructosamine level reflects undernutrition and associated energy deficits during the transition period. However, stress is inducing an increase of blood glucose, as a source of energy for muscles to enable the animal to escape (Sapolsky

et al., 2000), and high fructosamine level was associated with recurrent stress, through high, or frequent increase of blood glucose, or both (Trevisi and Bertoni, 2009). Caré et al. (2018) also observed a positive correlation between fructosamine and albumin, suggesting that inflammatory conditions may also affect fructosamine. In the chronic stress experiment previously performed, dairy cows in stressful conditions over 4 wk exhibited significantly higher blood fructosamine content compared with a control group, with 240 versus 228 $\mu\text{mol/L}$, respectively (Grelet et al., 2022). However, serum fructosamine routine measurement for chronic stress assessment currently involves blood sampling, which induces pain for animals and may negatively affect welfare conditions if done on a routine, for example, monthly, basis.

Mid-infrared (MIR) analysis of milk has been presented as a method of choice for large-scale phenotyping, for both fine milk composition and cow status, as it contains raw but massive information on global milk composition (Gengler et al., 2016). Bahadi et al. (2021) evaluated the effect of cattle housing conditions on MIR spectra but without the predictive purpose and development of MIR models. Recently, Ramirez Mauricio et al. (2025) tried to use milk MIR spectra to predict welfare traits related to animal-based measures, farm management, structure, and equipment, but model accuracy levels achieved were generally too low to be used. To the best of our knowledge, the possibility of obtaining information on chronic stress through MIR analysis of milk has not been explored. Therefore, the aim of this work was to evaluate the potential of milk MIR spectra in predicting the 2 previously mentioned chronic stress biomarkers at the individual cow level: hair cortisol and blood fructosamine. Both quantitative and qualitative approaches were tested, the quantitative one to predict the exact quantitative value of the biomarkers, and qualitative approach to discriminate the low biomarkers content versus high content (in hair or in blood). For modeling purposes, sampling of individual cows was oriented both to cows with supposed stressful conditions and cows with supposed good welfare conditions. This step was only performed on a screening purpose, to maximize the chance to have both low and high content of cortisol in hair and fructosamine in blood. The sampling recommendations were not considered further as a welfare indication, and the MIR models only aimed to predict the hair cortisol and the blood fructosamine contents, and not the welfare conditions in themselves. The 2 molecules being measured in other matrices than milk cannot be directly predicted by milk MIR spectra. However, chronic stress is known to affect both the energy and protein metabolisms (Klasing, 1985; Newsholme and Leech, 2010; van der Kooij, 2020), which, in turn, are likely to affect global milk composition and hence the milk MIR

spectra. Therefore, we hypothesized that indirect links between hair cortisol, blood fructosamine, and milk MIR spectra may exist, potentially enabling the development of predictive models. By providing objective, cheap, and quantitative information on chronic stress biomarkers of individual dairy cows through a simple milk analysis, the development of such tools would facilitate large-scale chronic stress assessment and improvement.

MATERIALS AND METHODS

Large-Scale Sampling

A specific dataset was created with the objective of developing MIR models, by organizing a dedicated large-scale sampling. It has been previously reported that MIR models provide bad prediction quality in case of extrapolation, that is, when a model is applied to samples that are unknown from calibration dataset variability (Grelet et al., 2021). Therefore, to develop sufficiently robust MIR models for real routine use, sampling was done to integrate a maximum of variability in the dataset in terms of spectral data, sample characteristics, and range of variables to predict. Consequently, the large-scale sampling protocol aimed to collect (1) a wide range of values regarding hair cortisol and blood fructosamine and (2) as much variability as possible within the region of the HappyMoo project (Interreg Northwest Europe). Sampling was performed in commercial dairy farms by the regional milk recording organizations (MRO) partners from July 2021 to June 2022 in Austria, Belgium, France, Germany, and Luxembourg for a total of 1,412 individual cow records.

To maximize the variability of hair cortisol and blood fructosamine values, local MRO partners were asked to collect samples from ~25% of cows with supposedly good welfare conditions, and 75% of cows potentially suffering from chronic stress for at least 1 mo. This ratio was decided empirically to optimize the probability of having high values for the MIR modeling stage. Indeed, high values need to be integrated into the model to avoid further extrapolation in case of use in real life. However, the relationship between the 2 biomarkers and the various dimensions of welfare is still unknown or not clear, and the potential stress sources may not affect the individual cows and their biomarker levels in the same way. Consequently, to increase probabilities to obtain high biomarker values, it was decided to overweight the proportion of cows potentially suffering from chronic stress in the sampling (i.e., 75% of sampling). In the absence of a gold standard method or proxy for chronic stress, local MRO were asked to look for potential chronic stress sources in an empirical way. They were asked to identify cows in stressful situations, such as overstocking (25% more cows than feeding places, 25% more cows than lying places), chronic diseases

es (individual recording of lameness, mastitis, or others), bad floor quality (slippery floor), difficulties in accessing food or water, inadequate barn design or location (ventilation, angles, inadequate milking equipment settings [vacuum], electrical troubles, high voltage lines, airport, wind turbine, and so on), permanent sources of pain and fear, observations by milk recording staff (stressed cows during milking, and so on), inappropriate or aversive human handling or management, environmental noises, and discomfort. The individual disease events were gathered from the MRO database. The barn and environment conditions were collected by local MRO advisers from their frequent farm visits and exchanges with farmers. These sampling recommendations did not aim to precisely study the link between biomarkers and potential sources of chronic stress in this work but were only considered as a rough screening tool to increase the probability of obtaining real high values of both biomarkers for modeling purposes. The sampling recommendations were not considered further as a welfare indication, and only the biomarkers' contents were targeted for MIR modeling.

To obtain as much variability as possible in terms of milk composition, MIR instruments, zootechnical, and local specificities, cow samples were taken from each geographic region as described below. The number of collected samples was proportional to the number of cows within the region of each MRO partner: 166 in Austria; 97 in Wallonia, Belgium; 149 in Brittany, France; 128 in Doubs, France; 158 in Normandy, France; 60 in Puy de Dôme, France; 60 in the Rhône, France; 169 in the North, France; 184 in Pays de la Loire, France; 201 in Baden-Württemberg, Germany; and 40 in Luxembourg. A total of 78 herds were involved in the sampling, leading to an average of 18 cows sampled per herd. As stage of lactation, parity, and breed may affect both milk composition and biomarker contents, cows were sampled at different stages of lactation and parities (i.e., 1 to 11) and included 7 breeds (Brown Swiss, Holstein, Jersey, Montbéliarde, Simmental, Vorderwälder, and crossed breeds), to cover a maximum of variability and avoid further extrapolation when using the model. The data also covered different seasons, with grazing and winter diets, and mountain and plain geographic regions.

Each of the 1,412 cows was sampled 1 time for milk, hair, and blood. The milk samples were collected the day before or at the same time as hair and blood samples. Due to missing milk MIR spectra, cortisol, or fructosamine values, the dataset contained 1,362 and 1,254 complete records for hair cortisol and blood fructosamine, respectively.

In addition to large-scale sampling, a few samples from 2 existing trials were added to the dataset. The first experiment was run at Centre Wallon de Recherches Agronomiques (CRA-W), Belgium, before large-scale

sampling and involved 30 Holstein cows. This trial aimed to induce overstocking during a 4 wk period on 15 cows, with 15 control cows under normal conditions. The data comprised 29 individual hair cortisol values and 100 blood fructosamine values (control cows being sampled for fructosamine at wk 0 and 4, and control cows being sampled weekly from wk 0 to 4). Further description can be found in Grelet et al. (2022). The second trial was run at the Institute for Agriculture, Fisheries, and Food Research (ILVO), Belgium, after large-scale sampling, and involved 39 Holstein cows with individual hair cortisol values. These samples were taken during the transition period, at 35 d postcalving for each of the 39 cows. Samples from these cows were collected with the same sampling interval as for the large-scale sampling. With the inclusion of these additional samples, the final dataset contained 1,430 records with hair cortisol and 1,354 with blood fructosamine reference values.

Hair Sampling and Cortisol Analysis

Hair samples were taken from the tail switch. As shown in Figure 1, only the upper part of the tail switch was sampled to avoid any discomfort for the cow (fly chasing). To target long-term stress periods but also to enable the link with milk composition and MIR spectra, it was decided to analyze the hair grown within 30 d before the milk sampling. At the tail switch, hairs grow about 0.5 mm/d (Burnett et al., 2014; Heimbürge et al., 2020), meaning that 30 d of growth corresponds to 1.5 cm of hair from the skin. Hair was therefore precut with scissors and clippers with a limitation comb at ~1.5 cm from the skin to remove excess hair (Figure 1). The remaining hair was trimmed with electric clippers and a high limitation comb fixed at 2 to 3 mm to avoid collecting skin particles, and collected in clean containers such as aluminum baskets.

The clippers were cleaned with a dry brush between each sampling. Hair was dried at room temperature for 1 wk, stored in plastic vials, and sent to CRA-W, Belgium, for analysis. Samples were stored in the absence of light pending their analysis. Samples were treated in line with a protocol adapted from Tallo-Parra et al. (2015) and described in Grelet et al. (2022). In short, hairs were sieved to separate them from skin follicles, dirt, and feces, then weighed to achieve 250 mg and washed 3 times with isopropanol (2-propanol 99.5%) for 2.5 min to remove saliva, sweat, and sebum, as diffusion of cortisol to these fluids is influenced by acute stress (Nedić et al., 2017). The hair samples were later dried, ground using a ball mill, and cortisol was extracted with pure methanol by agitating at 100 rpm for 18 h at 30°C. After centrifugation, the supernatant was placed in an oven at 38 °C for 24 h to evaporate methanol. The dried extracts were

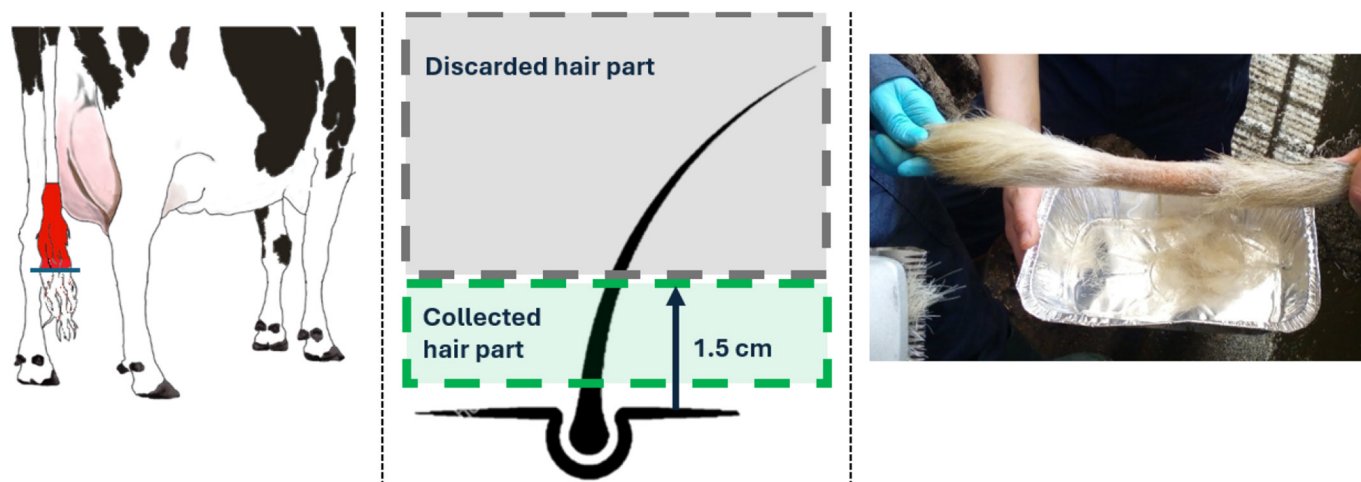


Figure 1. Hair sampling for cortisol analysis. Samples were taken from the upper half of the tail switch (in red in the image on the left). Hair was precut at ~1.5 cm from the skin to gather only hair grown in the last 30 d, and remaining hair was trimmed 2 to 3 mm from the skin and collected for analysis (center). Tail switch shown after hair sampling in the photo on the right.

reconstituted with 0.25 mL buffer provided in the ELISA kit and stored at -20°C . Cortisol concentration was determined using Salimetrics extended range salivary cortisol ELISA kit (1-3002, Salimetrics, State College, PA) following manufacturer protocol. Intra-assay repeatability was 6% CV (22 samples analyzed in triplicates), and interassay reproducibility of the ELISA was 4% CV (based on the plate averages of the 22 samples analyzed in triplicate on 2 plates).

Blood Sampling and Fructosamine Analysis

Blood samples were collected from the tail vein (vena caudalis), on the same day as hair specimens, in 3.5-mL yellow dry tubes with serum separating gel. The tubes were gently mixed and stored vertically at 4°C or in a tube rack on top of ice, until centrifugation. Tubes were centrifuged between 20 min and a few hours after sampling. Centrifugation was performed in laboratories or at farms, at $2,500 \times g$ for 10 min at 2 to 8°C to harvest serum. Five hundred microliters of serum was transferred to 1-mL Eppendorf vials and stored at -20°C . Frozen samples were sent to CRA-W with dry ice, and analysis was carried out at Synlab (Liège, Belgium) using the spectrophotometric method described in Westgard et al. (2017) following standard procedures by Abbott (Alinity C).

Milk Sampling and MIR Analysis

Because animal handling for hair and blood sampling may have an effect on milk composition, milk samples were taken the same day, or the day before, as hair and blood samples to avoid any effect of hair and blood sampling on milk composition. Individual milk samples were

collected by local MRO partners according to the guidelines edited by the International Committee for Animal Recording (ICAR; ICAR Dairy Cattle Milk Recording Working Group, 2017) and with ICAR-approved milk samplers. Samples were collected in 24-h representative milk or alternative (morning or evening) milk recording protocols, so were morning samples, or evening samples, or morning and evening samples pooled into a daily sample with a similar quantity of morning and evening milk as in usual milk recording routines (i.e., 50:50). All samples were stored at 4°C with 0.02% to 0.03% bronopol until analysis. Analyses were executed locally using a wide range of instruments, with a total of 30 spectrometers used to analyze the samples: 21 Foss instruments—models FT6000, FT+, and FT7 (Foss, Hillerød, Denmark), 7 Bentley FTS (Bentley, Chaska), and 2 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 dataset following the procedure described in Grelet et al. (2015). Milk MIR spectra were visually explored through a principal component analysis.

Data Cleaning

From the dataset comprising 1,430 hair cortisol and 1,354 blood fructosamine records, 4 cleaning steps were performed. To avoid erroneous associations between spectra and samples, as well as analytical issues during analysis, a local 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/100 mL between local and labo-

ratory predictions (Zhang et al., 2021) were discarded to prevent incorrect association between spectra and samples or other analytical errors ($n = 107$ records deleted, 7% of the dataset). Only records with DIM between 5 and 365 were retained ($n = 80$ records deleted, 6% of the dataset). As advised by the kit manufacturer, hair cortisol ELISA analyses showing a CV higher than 12% between 2 replicates were removed ($n = 139$ records deleted, 10% of the dataset). Fructosamine results from blood samples showing a hemolysis index of 2 or higher were eliminated ($n = 163$ records deleted, 12% of the dataset). Spectra were not filtered according to Mahalanobis distances, on purpose, to maintain spectral variability in the dataset. Finally, cortisol data from 1 partner and fructosamine records from 2 other partners were excluded from the dataset due to errors in sampling or analytical steps, or both (e.g., length of hair, temperature-controlled storage of blood samples). The final dataset comprised 1,071 hair cortisol and 940 blood fructosamine values. All samples were later retained without removing outliers during the model development.

Data Editing

Computations and models were carried out with programs developed in MATLAB 2022 (MathWorks, Inc., Natick, MA) and PLS Toolbox v. 8.5.1 (Eigenvector Research, Inc., Wenatchee, WA). The MIR spectra were pre-treated by a first derivative with a gap of 5 wavenumbers to sharpen the signal, as tested by Soyeurt et al. (2011). The selected spectral area consisted of 212 wavenumbers from 968.1 to 1,577.5, 1,731.8 to 1,762.6, 1,781.9 to 1,808.9, and 2,831.0 to 2,966.0 cm^{-1} to exclude spectral areas not reproducible between instruments and which may cause prediction inaccuracies when transferring the models for real use on various instruments (Grelet et al., 2021). The spectral variables were mean-centered for further concatenation with extra predictors. Several

variables were added sequentially, after also being mean-centered, to the MIR spectra as predictors: DIM, DIM², parity, parity², milk yield (MY), breed, and hair color. These extra predictors were first added individually to the MIR spectra to assess their respective added values, and then classified on the basis of the obtained root mean square error (RMSE) into an order of integration in the models, the predictors leading to the lowest RMSE integrated first. Hair color was only added to the models to predict hair cortisol, as it is known to affect the cortisol contents (Burnett et al., 2014), and was added as a dummy variable (1 for dark and 0 for light hair). Given the imbalance between the number of Holstein cows and the number of animals from other breeds, breed was only considered as Holstein versus other breeds, and also added in a dummy variable (1 for Holstein and 0 for others).

MIR Discriminant Models

The qualitative MIR models aimed at discriminating the individual dairy cows having low or high content of the 2 chronic stress biomarkers from the individual milk MIR spectra. They did not aim at classifying stress conditions or welfare situations in themselves but only if hair cortisol and blood fructosamine were low or high. Qualitative models were tested through the use of partial least squares discriminant analysis (PLS-DA). However, no threshold was identified in the literature to make the distinction between nonstressed and stressed cows. Therefore, the model was optimized statistically by testing several potential thresholds of cortisol values (i.e., 14, 16, 18, 20, 22, and 24 pg/mg). The threshold enabling the best discrimination between low and high values was selected, looking specifically at sensitivity to find a good balance between sensitivity and specificity, because specificity was always higher than sensitivity due to the imbalance of the dataset.

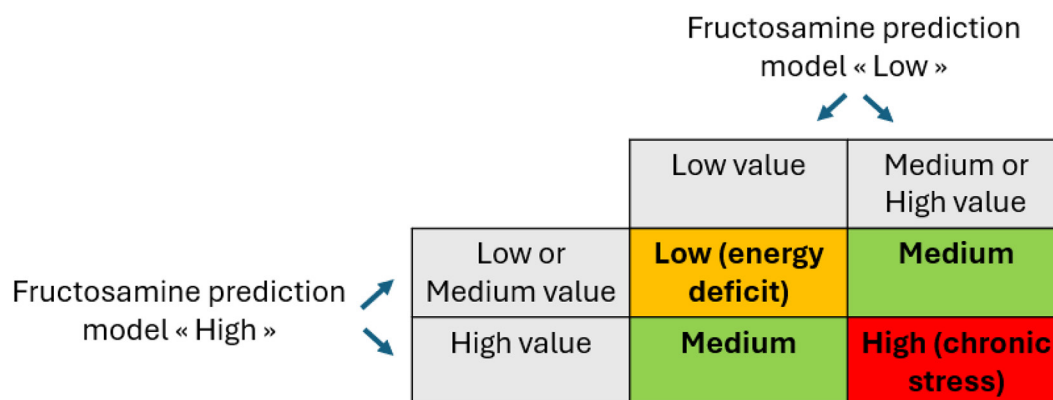


Figure 2. Methodology of combining the 2 fructosamine models, 1 discriminating low fructosamine content (from medium and high) and 1 discriminating high fructosamine content (from low and medium), respectively.

As low fructosamine values are considered to reflect energy deficit, and high fructosamine values appear to mirror chronic stress, the same process was applied to blood fructosamine, but 2 thresholds were considered to distinguish between low, intermediate, and high values. To do this, 2 separate models were developed: the first one to discriminate low values (from intermediate and high values together) and a second one to discern high values (from low and intermediate together). The first threshold was investigated within the low fructosamine range (i.e., 170, 180, 190, and 200 $\mu\text{mol/L}$) and the high threshold within the high fructosamine range (i.e., 220, 230, 240, and 250 $\mu\text{mol/L}$). The predictions of the 2 models were later combined following the decision rule as explained in Figure 2. Another modeling strategy with only 1 model directly discriminating 3 classes was also tested, but performances were less interesting, especially in terms of sensitivity, and results are therefore not displayed.

MIR Quantitative Models

The quantitative MIR models were aimed at predicting the exact quantitative value of both cortisol in hair and fructosamine in blood, for each individual cow, based on milk MIR spectra. Four quantitative modeling strategies were tested, each time with all predictors being sequentially added to MIR spectra as described previously. In the first strategy, all the variables were mean-centered and concatenated before the use of PLS regression. For the second strategy, the same treatment of variables was executed in combination with support vector machine regression (SVM). In these 2 first strategies, the high number of variables from the MIR spectra (i.e., 212) may have diluted the information added by the other predictors. Therefore, as suggested by Fearn (2010), the MIR spectral variables in strategies 3 and 4 were first reduced into 14 latent variables through a PLS and later concatenated to the other predictors. In strategy 3, the reduced spectra and other predictors were entered into a PLS-multiple linear regression model. In strategy 4 (PLS-SVM), the same variables were entered into an SVM model.

Validation

Hyperparameters, such as latent variables in the PLS or cost, epsilon, and gamma in the SVM, were first optimized during a cross-validation process including

10-fold. Models were then evaluated in an external herd validation process. To do so, the dataset was split into calibration and validation datasets by randomly excluding one-third of herds (i.e., 26 herds out of 78) from the different countries and partners to constitute the validation dataset. The remaining two-thirds of herds were used as the calibration dataset. This process was conducted 3 times to run all the herds in external validation, and the metrics were calculated over the entire dataset by concatenating the 3 thirds validation sets. In both the calibration and validation datasets, all samples were retained without removing outliers during model development. By doing so, the models were evaluated on cows and herds that were not present in the calibration dataset.

Discriminant model statistics, both in the calibration and validation steps, were expressed in terms of sensitivity (percentage of good classification in the high content group), specificity (percentage of good classification in the low content group), and global accuracy (global percentage of correct classification) over the entire data. In the preliminary threshold and variable optimization phases, the best threshold and best variables for integration in the model were simply identified by looking at average sensitivity and specificity over the 3 validation sets (one-third of herds each), meaning that these metrics were not computed over the entire dataset.

Quantitative model statistics, both in the calibration and validation steps, were expressed in terms of R^2 (coefficient of determination) and RMSE. The models were evaluated using the external herd validation sets by calculating the mean R^2 and mean RMSE over the 3 external herd validation datasets. Ratio of RMSE/SD of global dataset was not calculated because it was deemed irrelevant for models with asymmetric distribution, for which SD did not correctly describe the spread of the population (Bellon-Maurel et al., 2010). As a last step, final models were developed using all the data; calibration and validation as well as internal cross-validation R^2_{cv} and RMSE_{cv} (10 subsets constituted by cyclic iteration).

RESULTS

Description of Reference Values and Spectra

Descriptive statistics of reference values are shown in Table 1, and distributions are plotted in Figure 3. Individual hair cortisol values ranged from 1.0 to 53.9

Table 1. Descriptive statistics of the measured hair cortisol and blood fructosamine

Item	n	Minimum	Maximum	Mean	Median	SD	Skewness
Hair cortisol (pg/mg)	1,071	1.0	53.9	15.7	14.2	7.3	1.4
Blood fructosamine ($\mu\text{mol/L}$)	940	123.0	283.0	211.9	212.0	27.4	-0.2

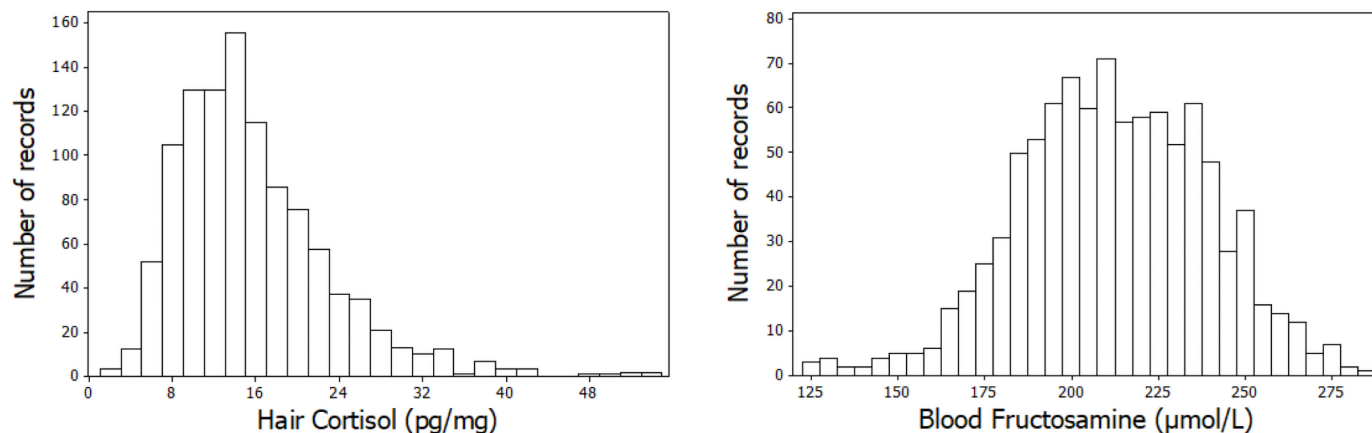


Figure 3. Distribution of the measured hair cortisol and blood fructosamine.

pg/mg, with an average of 15.7 pg/mg and presented a positively skewed distribution. Individual blood fructosamine values ranged from 123.0 to 283.0 $\mu\text{mol/L}$, with an average of 221.9 $\mu\text{mol/L}$ and showed a normal distribution. Hair cortisol and blood fructosamine reference values were not correlated (P -value = 0.174). Figure 4 shows the milk MIR spectra from the dataset plotted into a principal component analysis (PCA) and displays the latent variables from 1 to 4, with these 4 variables representing 89.25% of spectral variability.

Qualitative MIR Models

Hair Cortisol. The results of the milk MIR qualitative models predicting if the hair cortisol was low or high are displayed in the form of confusion matrices, comparing

the real hair cortisol classification (low or high according to the thresholds tested) versus the hair cortisol classification predicted by the models, for all individual dairy cow records. During optimization of the threshold, the best balance between sensitivity and specificity was found with a threshold of 22 pg/mg hair, in combination with the model including MIR spectra, with MY, parity², DIM, breed, parity, and DIM², respectively, in the PLS-DA (all details in Supplemental Table S1, see Notes). Hair color surprisingly did not provide added value and reduced the discrimination performances of hair cortisol. Knowing that PLS-DA is sensitive to unbalanced proportions between the groups to discriminate, it is worth noting that the threshold providing the best balance between low and high cortisol, that is, the lowest one (14 pg/mg hair) was not selected. Therefore, the threshold selection seems to

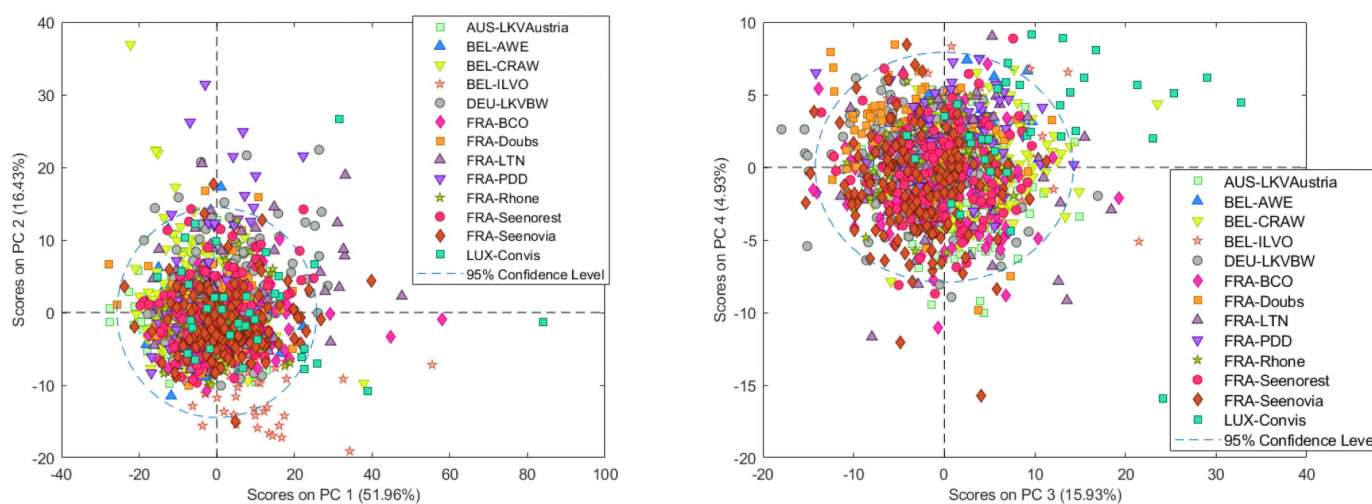


Figure 4. Entire spectral dataset represented through a principal component analysis on spectra pretreated with first derivative and reduced to 212 wavenumbers. Principal components (PC) from 1 to 4 are represented. Each symbol-point corresponds to regions of sampling. First 3 letters in the legend correspond to countries (AUS = Austria; BEL = Belgium; DEU = Germany; FRA = France; LUX = Luxembourg) and then comes the name of the different milk recording organization partners: AWE = Association Wallonne de l'Elevage; BCO = Bretagne Ouest; CRA-W = Centre Wallon de Recherches Agronomiques; LKVBW = Landesverband Baden-Württemberg.

Table 2. Results of the partial least square discriminant analysis (PLS-DA) model discriminating high hair cortisol values (i.e., >22 pg/mg hair) from low hair cortisol values, for individual cows¹

Item	Real low cortisol	Real high cortisol	Total	Specificity, %	Sensitivity, %	Accuracy, %
Cross validation						
Predicted low cortisol	719	73	792	81		77
Predicted high cortisol	172	107	279		59	
Total	891	180	1,071			
External herd validation						
Predicted low cortisol	656	71	727	74		71
Predicted high cortisol	235	109	344		61	
Total	891	180	1,071			

¹The confusion matrix displays the real hair cortisol classification (low or high according to the threshold) versus the hair cortisol classification predicted by the mid-infrared (MIR) model, both in 10-fold cross-validation and in external herd validation over the entire dataset (n = 1,071). External herd validation was done by randomly excluding 33% of herds (i.e., 26 herds out of 78) from the different countries and partners, and this process was executed 3 times to run all the herds in external validation. Model made with 12 latent variables.

be based on real low versus high cortisol discrimination rather than on balance between groups, which could have been a potential mathematical artifact. Table 2 shows the confusion matrix of hair cortisol classification, both in cross-validation and in external herd validation over the entire dataset; the final model was done with 12 latent variables. In cross-validation, the model succeeds in correctly classifying 77% of records, with a specificity of 81% and a sensitivity of 59%. In external herd validation, the capacity of the model to correctly classify samples logically decreased to 71%, with a specificity of 74% and a sensitivity of 61%. The model shows a sensitivity of 61% only, and a high rate of false positives, which are probably due to the imbalance between low and high cortisol, which represent 83% and 17% of the dataset, respectively.

Blood Fructosamine. The results of the milk MIR qualitative models predicting if the blood fructosamine was low or high are displayed in the form of confusion matrices, comparing the real blood fructosamine classification (low or high according to the thresholds tested)

versus the blood fructosamine classification predicted by the models, for all individual dairy cow records. The first step was to optimize the best threshold and the best model to discriminate the low fructosamine values, reflecting energy deficit. During optimization, the best balance between sensitivity and specificity was found with a threshold of 200 $\mu\text{mol/L}$, in combination with the model including MIR spectra, with MY, parity, parity², DIM, DIM², and breed, respectively, in the PLS-DA (all details in Supplemental Table S2, see Notes). Table 3 shows the confusion matrix for the best selected model, both in cross-validation and in external herd validation over the entire dataset, the final model was done with 14 latent variables. In cross-validation, the model succeeds in correctly classifying 78% of records, with a specificity of 79% and a sensitivity of 75%. In external herd validation, the performance of the model decreased to 67% accuracy, 59% sensitivity, and 71% specificity.

The second step aimed to optimize the best threshold and the best predicting variables to discriminate the high

Table 3. Results of the partial least square discriminant analysis (PLS-DA) model discriminating low blood fructosamine values (i.e., <200 $\mu\text{mol/L}$) from intermediate to high values, for individual dairy cows¹

Item	Real low fructosamine	Real medium or high fructosamine	Total	Sensitivity, %	Specificity, %	Accuracy, %
Cross validation						
Predicted low fructosamine	246	126	372	75		78
Predicted medium or high fructosamine	82	486	568		79	
Total	328	612	940			
External herd validation						
Predicted low fructosamine	195	177	372	59		67
Predicted medium or high fructosamine	133	435	568		71	
Total	328	612	940			

¹The confusion matrix displays the real blood fructosamine classification (low or high according to the threshold) versus the blood fructosamine classification predicted by the mid-infrared (MIR) model, both in 10-fold cross-validation and in external herd validation over the entire dataset (n = 940). External herd validation was done by randomly excluding 33% of herds (i.e., 26 herds out of 78) from the different countries and partners, and this process was executed 3 times to run all the herds in external validation. Model made with 14 latent variables.

Table 4. Results of the partial least square discriminant analysis (PLS-DA) model discriminating high blood fructosamine values (i.e., >230 µmol/L) from low to intermediate values, for individual dairy cows¹

Item	Real low and medium fructosamine	Real high fructosamine	Total	Specificity, %	Sensitivity, %	Accuracy, %
Cross validation						
Predicted low and medium fructosamine	513	58	571	75		75
Predicted high fructosamine	175	194	369		77	
Total	688	252	940			
External herd validation						
Predicted low and medium fructosamine	467	77	544	68		68
Predicted high fructosamine	221	175	396		69	
Total	688	252	940			

¹The confusion matrix displays the real blood fructosamine classification (high or low according to the threshold) versus the blood fructosamine classification predicted by the mid-infrared (MIR) model, both in 10-fold cross-validation and in external herd validation over the entire dataset (n = 940). External herd validation was done by randomly excluding 33% of herds (i.e., 26 herds out of 78) from the different countries and partners, and this process was executed 3 times to run all the herds in external validation. Model made with 14 latent variables.

Table 5. Results of the combination of the 2 fructosamine models, 1 discriminating low fructosamine content (from moderate and high), and 1 discriminating high fructosamine content (from low and moderate), respectively¹

Item	Real low fructosamine (<200 µmol/L; undernutrition)	Real medium fructosamine	Real high fructosamine (>230 µmol/L; chronic stress)	Total	High fructosamine specificity, %	High fructosamine sensitivity, %	High fructosamine accuracy, %
External herd validation							
Predicted low	169	110	38	317	75		73
Predicted medium	100	136	46	282			
Predicted high	59	114	168	341		67	
Total	328	360	252	940			

¹The final confusion matrix displays the real blood fructosamine classification (low, medium, or high according to the thresholds) versus the combined blood fructosamine classification predicted from the 2 mid-infrared (MIR) models. Results are only shown for the external herd validation process over the entire dataset (n = 940).

Table 6. Results of the quantitative models predicting hair cortisol from milk mid-infrared (MIR) spectra and additional predictors, with different variables and 4 modeling methodologies^{1,2}

Variable in the model\algorithm	RMSE (in pg/mg hair)				R ²			
	PLS	SVM	PLS-MLR	PLS-SVM	PLS	SVM	PLS-MLR	PLS-SVM
Calibration								
MIR	6.0	5.9	6.0	6.0	0.32	0.35	0.33	0.35
MIR+MY	6.0	5.9	6.0	6.0	0.32	0.35	0.33	0.34
MIR+MY+parity ²	6.0	5.9	5.9	5.9	0.32	0.36	0.34	0.35
MIR+MY+parity ² +DIM	6.0	5.8	5.9	5.9	0.32	0.37	0.34	0.36
MIR+MY+parity ² +DIM+breed	6.0	5.9	5.9	5.9	0.33	0.35	0.34	0.35
MIR+MY+parity ² +DIM+breed+parity	6.0	5.9	5.9	5.9	0.33	0.37	0.34	0.36
MIR+MY+parity ² +DIM+breed+parity+DIM ²	6.0	5.8	5.9	5.9	0.33	0.37	0.34	0.35
MIR+MY+parity ² +DIM+breed+parity+DIM ² +color	6.0	5.7	—	—	0.33	0.40	—	—
External herd validation								
MIR	7.0	6.9	7.0	6.9	0.11	0.12	0.11	0.12
MIR+MY	7.0	7.0	7.0	6.9	0.11	0.11	0.11	0.12
MIR+MY+parity ²	7.0	7.0	7.0	7.0	0.10	0.11	0.11	0.11
MIR+MY+parity ² +DIM	7.0	6.9	7.0	6.9	0.11	0.12	0.11	0.12
MIR+MY+parity ² +DIM+breed	7.1	7.0	7.0	7.0	0.10	0.11	0.11	0.12
MIR+MY+parity ² +DIM+breed+parity	7.1	7.0	7.0	6.9	0.10	0.11	0.11	0.12
MIR+MY+parity²+DIM+breed+parity+DIM²	7.1	6.9	7.0	7.0	0.11	0.13	0.11	0.12
MIR+MY+parity ² +DIM+breed+parity+DIM ² +color	7.1	9.8	—	—	0.10	0.04	—	—

¹The calibration and external herd validation are displayed. External herd validation was done by randomly excluding 33% of herds (i.e., 26 herds out of 78) from the different countries and partners, and this process was executed 3 times to run all the herds in external validation. Results are displayed in terms of global RMSE and R². n = 1,071 individual records. Bold indicates model obtaining the best performance in external herd validation.

²MY = milk yield; PLS = partial least square regression; PLS-MLR = multiple linear regression model applied on concatenation of PLS reduced MIR spectra and other predictors; PLS-SVM = support vector machine model applied on concatenation of PLS reduced MIR spectra and other predictors; RMSE = root mean square error; SVM = support vector machines regression.

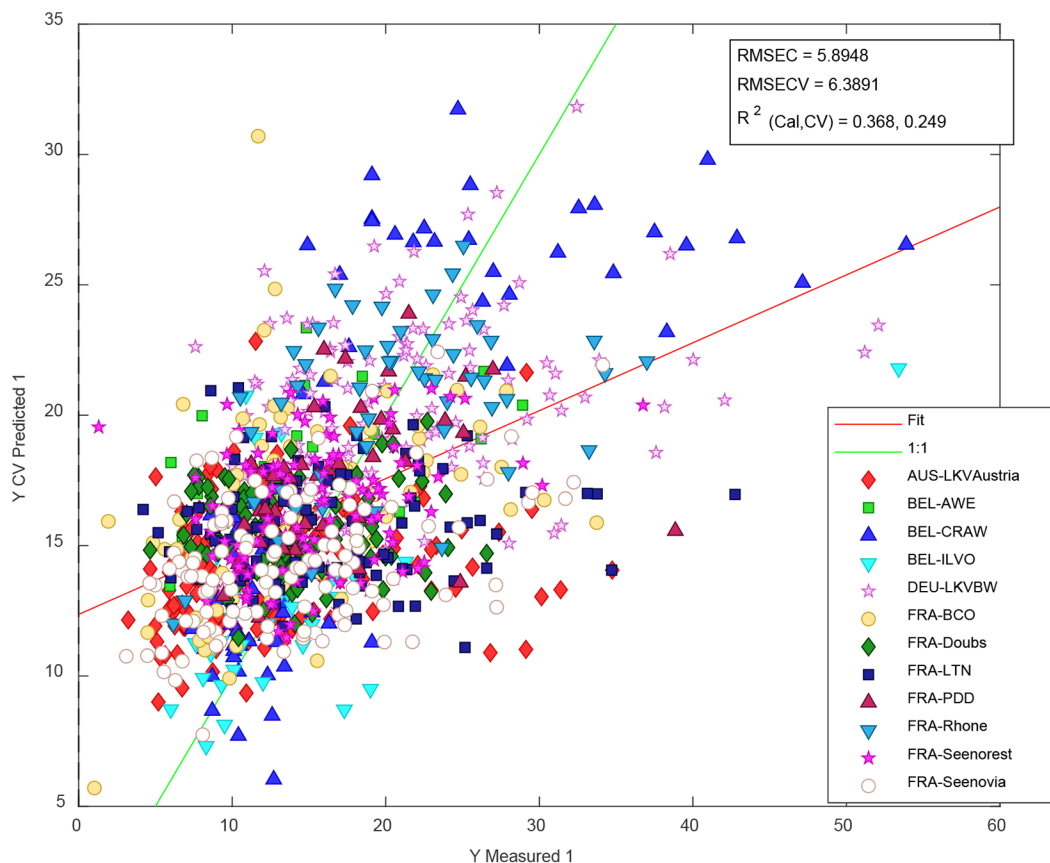


Figure 5. The 10-fold cross-validation performances of the final hair cortisol model developed on the full dataset ($n = 1,071$). PLS = partial least square regression; PLS-MLR = multiple linear regression model applied on concatenation of PLS reduced mid-infrared (MIR) spectra and other predictors; PLS-SVM = support vector machines model applied on concatenation of PLS reduced MIR spectra and other predictors; RMSE = root mean square error; RMSEC = root mean square error of calibration; RMSECV = root mean square error of cross validation; SVM = support vector machines regression. Each symbol-point corresponds to regions of sampling. First 3 letters in the legend correspond to countries (AUS = Austria; BEL = Belgium; DEU = Germany; FRA = France; LUX = Luxembourg) and then comes the name of the different milk recording organization partners: AWE = Association Wallonne de l'Elevage; BCO = Bretagne Ouest; CRA-W = Centre Wallon de Recherches Agronomiques; LKVBW = Landesverband Baden-Württemberg.

fructosamine values, mirroring chronic stress. During optimization, the best balance between sensitivity and specificity was found with a threshold of $230 \mu\text{mol/L}$, in combination with the model including only MIR spectra in the PLS-DA (all details in Supplemental Table S3, see Notes). Table 4 shows the confusion matrix for the best selected model, both in cross-validation and in external herd validation over the entire dataset, the final model was done with 14 latent variables. In cross-validation, the model succeeded in correctly classifying 75% of records, with a specificity of 75% and a sensitivity of 77%. In external herd validation the performance of the model decreased to 68% accuracy, 69% sensitivity, and 68% specificity.

In the third step, the predictions of the 2 models, the model discriminating low values and the model discriminating high values, were combined into a single result, as described in Figure 2. Table 5 displays the classifications and the performances over the 3 classes in external herd

validation. To discriminate the high fructosamine samples, reflecting chronic stress, the combination of models enabled sensitivity, specificity, and accuracy to increase to 67%, 75%, and 73%, respectively. The combination of models did not change the accuracy for discriminating low fructosamine samples, still at 67% but increased the imbalance between sensitivity, which dropped to 52%, and specificity, which rose to 76%. However, this is not the primary objective of this study, which is chronic stress and, consequently, high fructosamine values. Moreover, several studies have already focused on predicting energy balance or related traits with good performances (McParland and Berry, 2016; Aernouts et al., 2020).

Quantitative MIR Models

Hair Cortisol. The quantitative milk MIR models aimed at predicting the exact quantitative value of hair

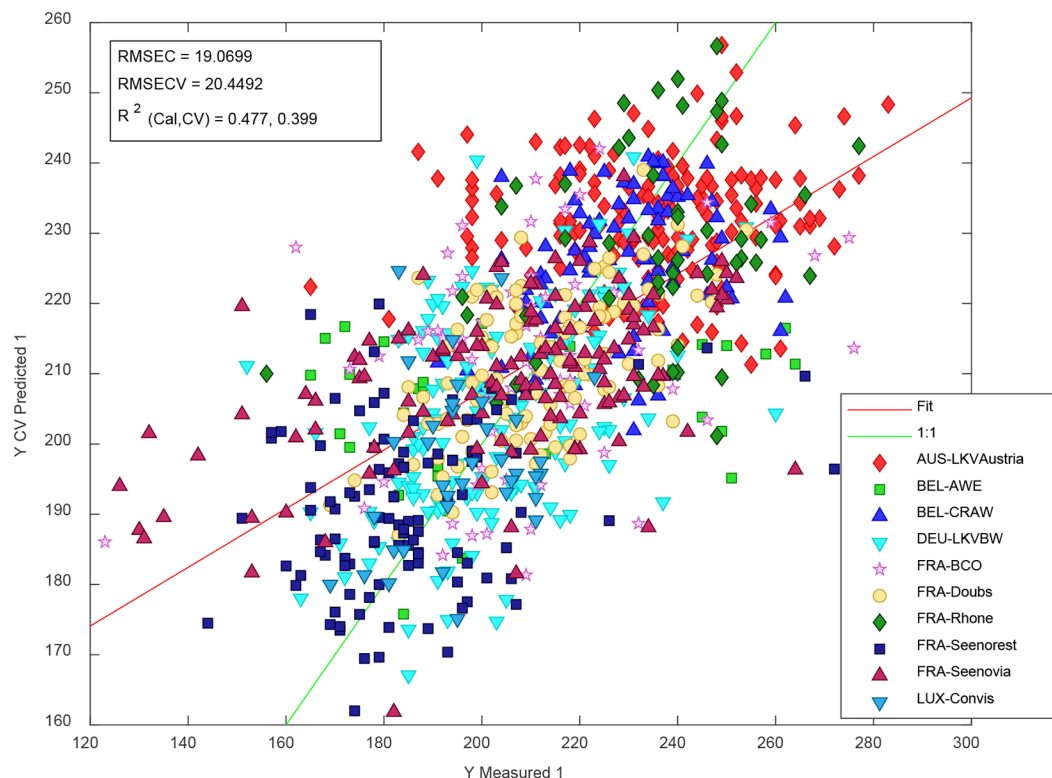


Figure 6. The 10-fold cross-validation performances of the final blood fructosamine model developed on the full dataset ($n = 940$). PLS = partial least square regression; PLS-MLR = multiple linear regression model applied on concatenation of PLS reduced mid-infrared (MIR) spectra and other predictors; PLS-SVM = support vector machines model applied on concatenation of PLS reduced MIR spectra and other predictors; RMSE = root mean square error; RMSEC = root mean square error of calibration; RMSECV = root mean square error of cross validation; SVM = support vector machines regression. Each symbol-point corresponds to regions of sampling. First 3 letters in the legend correspond to countries (AUS = Austria; BEL = Belgium; DEU = Germany; FRA = France; LUX = Luxembourg) and then comes the name of the different milk recording organization partners: AWE = Association Wallonne de l'Elevage; BCO = Bretagne Ouest; CRA-W = Centre Wallon de Recherches Agronomiques; LKVBW = Landesverband Baden-Württemberg.

cortisol for each individual cow, based on milk MIR spectra. The best model, both in terms of RMSE and R^2 , was the SVM model with milk spectra, MY, parity², DIM, breed, parity, and DIM² as variables (Table 6). In calibration, this model has an RMSE of 5.8 pg/mg hair which, when compared with mean cortisol (15.7 pg/mg hair), represents a relative error of 37% and R^2 at 0.37. In external herd validation, which represents the ability of the model to face real field data, the RMSE was still in the same order of magnitude, with 6.9 pg/mg hair, but R^2 drastically decreased to 0.13. The final model made with the entire dataset (i.e., calibration and validation sets) is shown in Figure 5, and in the 10-fold cross-validation, the RMSE was 6.4 pg/mg hair and R^2 was 0.25.

Blood Fructosamine. The quantitative MIR models were aimed at predicting the exact quantitative value of blood fructosamine, for each individual cow, based on milk MIR spectra. The best model, both in terms of RMSE and R^2 , was the SVM model with milk MIR spectra, MY, parity, parity², DIM, and DIM² as variables. In calibration, this model had an RMSE of 19.6 $\mu\text{mol/L}$ serum

which, when compared with mean fructosamine (211.9 $\mu\text{mol/L}$ serum), represents a relative error of 9.2% and R^2 at 0.49. In external herd validation, which is representative of the ability of the model to face real field data, the RMSE was still in the same order of magnitude, with 24.4 $\mu\text{mol/L}$ serum, but R^2 drastically decreased to 0.20. The final model made with the entire dataset is shown in Figure 6, and in the 10-fold cross-validation, the RMSE was 20.4 $\mu\text{mol/L}$ serum and R^2 was 0.40 (Table 7).

DISCUSSION

Description of Reference Values and Spectra

Measured hair cortisol values were in the same order of magnitude as observed by Ghassemi Nejad et al. (2017), del Rosario González-de-la-Vara et al. (2011), Heimbürge et al. (2020), and Otten et al. (2022) and higher than findings observed by Burnett et al. (2014), Comin et al. (2013), and Fischer-Tenhagen et al. (2018). The blood fructosamine values were generally in the same order of

Table 7. Results of the quantitative models predicting blood fructosamine from milk mid-infrared (MIR) spectra and additional predictors, with different variables and 4 modeling methodologies^{1,2}

Variable in the model\algorithm	RMSE (μmol/L serum)				R ²			
	PLS	SVM	PLS-MLR	PLS-SVM	PLS	SVM	PLS-MLR	PLS-SVM
Calibration								
MIR	20.5	19.5	20.5	19.4	0.44	0.49	0.44	0.50
MIR+MY	20.4	19.8	20.4	18.1	0.44	0.48	0.44	0.56
MIR+MY+parity	20.3	19.8	20.3	17.1	0.45	0.48	0.45	0.61
MIR+MY+parity+parity ²	20.3	18.7	20.3	17.1	0.45	0.53	0.45	0.62
MIR+MY+parity+parity ² +DIM	20.2	19.2	20.2	18.4	0.46	0.50	0.45	0.55
MIR+MY+parity+parity ² +DIM+DIM ²	20.2	19.6	20.2	17.6	0.46	0.49	0.45	0.58
MIR+MY+parity+parity ² +DIM+DIM ² +breed	19.9	19.0	20.0	18.4	0.47	0.52	0.47	0.55
External herd validation								
MIR	26.6	26.1	26.5	26.1	0.15	0.18	0.15	0.17
MIR+MY	26.3	25.5	26.4	25.7	0.18	0.20	0.17	0.20
MIR+MY+parity	26.1	25.5	26.2	25.8	0.19	0.20	0.17	0.19
MIR+MY+parity+parity ²	26.1	26.3	26.2	25.8	0.19	0.19	0.17	0.19
MIR+MY+parity+parity ² +DIM	26.1	25.9	26.1	25.5	0.18	0.21	0.18	0.19
MIR+MY+parity+parity²+DIM+DIM²	26.2	25.4	26.2	25.7	0.19	0.20	0.18	0.19
MIR+MY+parity+parity ² +DIM+DIM ² +breed	26.6	26.2	26.1	25.6	0.18	0.20	0.18	0.19

¹The calibration and external herd validation are displayed. Validation was done by randomly excluding 33% of herds (i.e., 26 herds out of 78) from the different countries and partners, and this process was executed 3 times to run all the herds in external validation. Results are displayed in terms of global root mean square error (RMSE) and R². n = 940 individual records. Bold indicates model obtaining the best performance in external herd validation.

²MY = milk yield; PLS = partial least square regression; PLS-MLR = multiple linear regression model applied on concatenation of PLS reduced MIR spectra and other predictors; PLS-SVM = SVM model applied on concatenation of PLS reduced MIR spectra and other predictors; SVM = support vector machines regression.

magnitude as values in literature, being similar to Mostafavi et al. (2014), slightly lower than Megahed et al. (2018) and Sorondo and Cirio (2009), and slightly higher than Caré et al. (2018). Hair cortisol and blood fructosamine reference values were not correlated (P -value = 0.174), meaning that even if the 2 variables are biomarkers of chronic stress, they do not represent the same information and potentially further complement each other's. A visual analysis of the spectra through PCA indicates that spectra within the entire dataset are homogeneous among partners, geographic regions, and instruments. No individual samples could be identified as obvious outliers, which indicates a good global quality and homogeneity in the MIR analysis in the respective local laboratories.

Quantitative MIR Models

The quantitative MIR models were aimed at predicting the exact quantitative value of both hair cortisol and blood fructosamine, for each individual cow, based on milk MIR spectra. In external herd validation, the R² were 0.13 and 0.2 for hair cortisol and blood fructosamine, respectively. To our knowledge, this is the first work aiming to predict these 2 biomarkers from the individual milk MIR spectra of cows, therefore, comparison with other studies is not possible. In light of the predictive statistical performance, it is likely that the hair cortisol model cannot be used for any quantitative aspects. According to the model

performance classification made in Grelet et al. (2021), the use of the model is not recommended. At least, herd management and monitoring should not be realized with this quantitative model, and the qualitative discriminant model should be considered instead. As for cortisol, the predictive statistical performance of the quantitative fructosamine models is likely to prevent the use of this model for quantitative aspects, and the qualitative model should be considered as an alternative.

Qualitative MIR Models

The qualitative MIR models were aimed at discriminating the individual dairy cows having low or high content of the 2 chronic stress biomarkers, using the individual milk MIR spectra. They did not aim at classifying stress conditions or welfare situations in themselves but only if hair cortisol and blood fructosamine were low or high. While the quantitative models do not provide satisfying results and are, therefore, unsuitable for monitoring chronic stress in dairy farms, the qualitative models successfully distinguished high hair cortisol and high serum fructosamine levels with accuracies of 71% and 73%, respectively. These accuracies reflect the ability of models to determine if individual cows have low or high biomarker content, and not their direct welfare status. Despite the moderate accuracies, the models may provide valuable new insights that were previously un-

available. These performances should also be observed in light of the expected difficulty in predicting these molecules. Indeed, both hair cortisol and blood fructosamine are not directly observed through milk analysis, so models rely only on indirect links, and cortisol concentration in hair is extremely low. The chronic stress is known to affect both the energy and protein metabolism (Klasing, 1985; Newsholme and Leech, 2010; van der Kooij, 2020), and molecules in milk related to energy and protein metabolisms are probably the indirect links between hair cortisol, blood fructosamine, and the milk MIR spectra. Therefore, the obtained accuracy, even if not very high, can be considered satisfactory in regard to the challenge of predicting these biomarkers of chronic stress. Compared with the hair cortisol model, the prediction of high fructosamine content to reflect chronic stress performs slightly better. Accuracy and specificity are in the same order of magnitude, but fructosamine discrimination shows a higher sensitivity, with 67% and 61%, respectively, for fructosamine and cortisol. This is not extremely surprising, as fructosamine is directly related to energy metabolism which is a main driver in cow physiology and therefore more susceptible to having an effect on milk composition. However, blood fructosamine is less well-documented, and its association with chronic stress is less obvious than hair cortisol.

The external herd validation is performed on unknown herds, not contained in the calibration dataset. Therefore, these statistics can be considered robust and reflecting the discrimination performances expected when using the model with real field data, in a routine large-scale process. The use of such models, at least in European regions covered by the sampling, can be expected to result in similar performances. To our knowledge, this is the first work aimed at achieving such discrimination from milk composition and additional variables, making comparison with previous studies impossible.

Perspectives

It seems preferable to consider only the qualitative models for practical use, but also to use these models at a herd level rather than at an individual level to cope with the imprecision of the models, for example, by looking at the proportion of cows above the thresholds within the herd. The models only predict biomarkers of chronic stress and not chronic stress in itself. Therefore, the results should be used with caution, especially for herd management, and high content of the biomarkers should rather be considered as a “risk” of chronic stress and not as a real gold standard diagnosis. Additionally, the models do not aim to substitute current welfare evaluations but rather to provide a large-scale screening tool

or a preliminary alert regarding chronic stress, and not welfare as a whole. Animal welfare is multidimensional, and the freedom from fear and distress is only one of these dimensions (Webster, 2001). For example, moving animals to grazing, by increasing social interaction, daily movement, and change in diet, has been shown to increase hair cortisol (Comin et al., 2011), while the possibility for the animal to graze is usually considered as positive for general welfare (Burow et al., 2013). Therefore, the developed tools do not provide information on the global welfare status of animals but on a part of it. It could be integrated into a global welfare assessment if combined with measures or observations of the other welfare dimensions.

The results of PLS-DA models provide a classification into low or high content, but use for quantitative applications (i.e., genetics and so on) is also possible by observing the predicted probabilities to belong to the 2 groups. Because stress is a transient gradient, the distinction of stressed versus nonstressed cows using a cutoff might not be optimal, making preferable the use of the quantitative probabilities. Additionally, in the absence of known thresholds, values were found through statistical means, by searching for the best discrimination between low and high classes. Therefore, these thresholds do not originate from a physiological basis, and this should be further validated or adapted. For such reasons, it is necessary to test the models on large-scale datasets to observe how they behave with real external data, and fine-tune the different parameters to provide relevant information to the dairy sector and highlight the best way to apply such models. Similarly, understanding chronic stress mechanisms and the links between the 2 biomarkers and animal or environmental effects is still incomplete, and further work should focus on deeper and broader-scale studies of such relationships. Other factors might potentially affect the level of hair cortisol and blood fructosamine, and there is the need for deeper physiological understanding. For example, biomarkers in the first month of lactation could be affected by other mechanisms than chronic stress (e.g., fetal cortisol and so on), consequently, the relevance of the predictive models according to the stage of lactation should be investigated and should be treated cautiously. The complementarity between hair cortisol and blood fructosamine should also be better understood to produce a combination of the 2 biomarkers for better phenotyping. Finally, given the imbalance between sample sizes between low and high groups, for both biomarkers, PLS-DA models provide low sensitivity and high false positive rates, and additional efforts should be made to test innovative algorithms which are better adapted to address this disparity and ultimately minimize this issue.

CONCLUSIONS

The goal of this work was to assess the potential of milk MIR spectra for predicting 2 chronic stress biomarkers: hair cortisol and blood fructosamine. While the quantitative models did not provide satisfying results, the qualitative models enabled us to discriminate high hair cortisol and high serum fructosamine levels with 71% and 73% accuracy in a robust external herd validation, respectively. Therefore, the obtained accuracy, despite not being very high, can be considered very positive considering the difficulty in predicting these totally indirect biomarkers of chronic stress. The development of such tools may facilitate large-scale chronic stress, and ultimately participate in welfare assessment and improvement by providing objective, cheap, and quantitative information on the stress dimension of it.

NOTES

HappyMoo project (NWE730) received funding from INTERREG IVB NWE and the Walloon Region (Belgium). Authors gratefully thank Petimat Kitaeva, Mohamed El Morabit, Matthieu Dubuisson, Raphaël Vivegnis, Lora Renault, Olivier Genard, Maxence Didelez, Maxime Druez, and Sophie Mathieux (CRA-W, Gembloux, Belgium) for their invaluable work. Authors also thank the members of the HappyMoo Consortium: Carlo Bertozzi, Marie-Nguyet Tran, Romain Reding, Jean-Pierre Allard, Guillaume Courtois, Christophe Lecomte, Luc Manciaux, Corine Portemont, Morgane Hoenen, Nicolas Gaudillère, Pauline Woehrle, Manon Jattiot, Marion Calmels, Valérie Wolf, Fatima Ezzahra Bouzidi Idrissi, Daragh Matthews, Mike Lynch, Marine Gelé, Mahommed El Jabri, Laura Dale, Andreas Werner, Tatjana Heim, Elisa Strang, Kevin Klever, Robert Strumpfen, Peter Höckels, Martin Busfield, Urs Schuler, Anthony Tedde, Hélène Soyeurt, and Sébastien Franceschini (CRA-W, Gembloux, Belgium). The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. Supplemental material for this article is available at <https://www.cra.wallonie.be/fr/can-we-get-information-on-dairy-cows-chronic-stress-biomarkers-from-milk-mid-infrared-spectra>. The authors have not used generative artificial intelligence. The experiment conducted at CRA-W was carried out in accordance with ARRIVE guidelines, the EU Directive 2010/63/EU for animal experiments, and the protocol (19-2181) was approved by Liège University Ethics Committee. Samples collected in commercial dairy farms were taken by farm veterinarians within the context of routine health diagnosis. Samples collected at ILVO were not subject to ethical

approval as only noninvasive sampling was done, that is, milk and hairs. The authors have not stated any conflicts of interest.

Nonstandard abbreviations used: CRA-W = Centre Wallon de Recherches Agronomiques; HPA = hypothalamic-pituitary-adrenocortical; ICAR = International Committee for Animal Recording; ILVO = Institute for Agriculture, Fisheries, and Food Research; MIR = mid-infrared; MLR = multiple linear regression; MRO = milk recording organizations; MY = milk yield; PC = principal component; PCA = principal component analysis; PLS-DA = partial least square discriminant analysis; RMSE = root mean square error; RMSEC = root mean square error of calibration; RMSECV = root mean square error of cross validation; SVM = support vector machine.

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>.
- Armbruster, D. A. 1987. Fructosamine: Structure, analysis, and clinical usefulness. *Clin. Chem.* 33:2153–2163. <https://doi.org/10.1093/clinchem/33.12.2153>.
- Bahadi, M., A. A. Ismail, and E. Vasseur. 2021. Fourier transform infrared spectroscopy as a tool to study milk composition changes in dairy cows attributed to housing modifications to improve animal welfare. *Foods* 10:450. <https://doi.org/10.3390/foods10020450>.
- Bellon-Maurel, V., E. Fernandez-Ahumada, B. Palagos, J. M. Roger, and A. McBratney. 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends Anal. Chem.* 29:1073–1081. <https://doi.org/10.1016/j.trac.2010.05.006>.
- Brito, L. F., H. R. Oliveira, B. R. McConn, A. P. Schinckel, A. Arrazola, J. N. Marchant-Forde, and J. S. Johnson. 2020. Large-scale phenotyping of livestock welfare in commercial production systems: A new frontier in animal breeding. *Front. Genet.* 11:793. <https://doi.org/10.3389/fgene.2020.00793>.
- Burnett, T. A., A. M. L. Madureira, B. F. Silper, A. Nadalin, A. Tahmasbi, D. M. Veira, and R. L. A. Cerri. 2014. Short communication: Factors affecting hair cortisol concentrations in lactating dairy cows. *J. Dairy Sci.* 97:7685–7690. <https://doi.org/10.3168/jds.2014-8444>.
- Burnett, T. A., A. M. L. Madureira, B. F. Silper, A. Tahmasbi, A. Nadalin, D. M. Veira, and R. L. A. Cerri. 2015. Relationship of concentrations of cortisol in hair with health, biomarkers in blood, and reproductive status in dairy cows. *J. Dairy Sci.* 98:4414–4426. <https://doi.org/10.3168/jds.2014-8871>.
- Burow, E., T. Rousing, P. T. Thomsen, N. D. Otten, and J. T. Sorensen. 2013. Effect of grazing on the cow welfare of dairy herds evaluated by a multidimensional welfare index. *Animal* 7:834–842. <https://doi.org/10.1017/S1751731112002297>.
- Caré, S., E. Trevisi, A. Minuti, A. Ferrari, J. J. Loor, and L. Calamari. 2018. Plasma fructosamine during the transition period and its relationship with energy metabolism and inflammation biomarkers in dairy cows. *Livest. Sci.* 216:138–147. <https://doi.org/10.1016/j.livsci.2018.08.003>.
- Comin, A., T. Peric, M. Corazzin, M. C. Veronesi, T. Meloni, V. Zufferli, G. Cornacchia, and A. Prandi. 2013. Hair cortisol as a marker of hypothalamic-pituitary-adrenal axis activation in Friesian dairy cows clinically or physiologically compromised. *Livest. Sci.* 152:36–41. <https://doi.org/10.1016/j.livsci.2012.11.021>.

- Comin, A., A. Prandi, T. Peric, M. Corazzin, S. Dovier, and S. Bovolenta. 2011. Hair cortisol levels in dairy cows from winter housing to summer highland grazing. *Livest. Sci.* 138:69–73. <https://doi.org/10.1016/j.livsci.2010.12.009>.
- del Rosario González-de-la-Vara, M., R. A. Valdez, V. Lemus-Ramirez, J. C. Vázquez-Chagoyán, A. Villa-Godoy, and M. C. Romano. 2011. Effects of adrenocorticotrophic hormone challenge and age on hair cortisol concentrations in dairy cattle. *Can. J. Vet. Res.* 75:216.
- Elsasser, T. H., N. C. Steele, and R. Fayer. 2020. Cytokines, stress, and growth modulation. Pages 261–290 in *Cytokines in Animal Health and Disease*. CRC Press. <https://doi.org/10.1201/9781003067382-10>.
- Europa. 2023. Attitudes of Europeans towards animal welfare - Octobre 2023 Eurobarometer Public Opinion in the European Union survey. Accessed Feb. 18, 2025. <https://europa.eu/eurobarometer/surveys/detail/2996>.
- FDA-NIF Biomarker Working Group (Food and Drug Administration-National Institutes of Health Biomarker Working Group). 2025. Glossary - BEST (Biomarkers, EndpointS, and other Tools) Resource - NCBI Bookshelf. Accessed June 24, 2025. <https://www.ncbi.nlm.nih.gov/books/NBK338448/#IX-B>.
- Fearn, T. 2010. Combining other predictors with NIR spectra. *NIR News* 21:13–16. <https://doi.org/10.1255/nirn.1175>.
- Fink, G. (2009). Stress: Definition and history. Pages 549–555 in *Encyclopedia of Neuroscience*. Academic Press. <https://doi.org/10.1016/B978-008045046-9.00076-0>.
- Fischer-Tenhagen, C., M. Ladwig-Wiegard, W. Heuwieser, and C. Thöne-Reineke. 2018. Short communication: Is hair cortisol a potential indicator for stress caused by chronic lameness in dairy cows? *J. Dairy Sci.* 101:5439–5443. <https://doi.org/10.3168/jds.2017-13967>.
- Gengler, N., H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, M. L. Vanrobays, A. Lainé, S. Vanderick, C. Grelet, A. Vanlierde, E. Froimont, and P. Dardenne. 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. *J. Dairy Sci.* 99:4071–4079. <https://doi.org/10.3168/jds.2015-10140>.
- Ghassemi Nejad, J., B. W. Kim, B. H. Lee, and K.-I. Sung. 2017. Coat and hair color: Hair cortisol and serotonin levels in lactating Holstein cows under heat stress conditions. *Anim. Sci. J.* 88:190–194. <https://doi.org/10.1111/asj.12662>.
- Grelet, C., P. Dardenne, H. Soyeurt, J. A. Fernandez, A. Vanlierde, F. Stevens, N. Gengler, and P. Dehareng. 2021. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. *Methods* 186:97–111. <https://doi.org/10.1016/j.ymeth.2020.07.012>.
- 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., V. Vanden Dries, J. Leblos, J. Wavreille, L. Mirabito, H. Soyeurt, S. Franceschini, N. Gengler, Y. Brostaux, and F. Dehareng. HappyMoo Consortium. 2022. Identification of chronic stress biomarkers in dairy cows. *Animal* 16:100502. <https://doi.org/10.1016/j.animal.2022.100502>.
- Heimbürge, S., E. Kanitz, and W. Otten. 2019. The use of hair cortisol for the assessment of stress in animals. *Gen. Comp. Endocrinol.* 270:10–17. <https://doi.org/10.1016/j.ygcen.2018.09.016>.
- Heimbürge, S., E. Kanitz, A. Tuchscherer, and W. Otten. 2020. Is it getting in the hair? – Cortisol concentrations in native, regrown and segmented hairs of cattle and pigs after repeated ACTH administrations. *Gen. Comp. Endocrinol.* 295:113534. <https://doi.org/10.1016/j.ygcen.2020.113534>.
- ICAR Dairy Cattle Milk Recording Working Group (International Committee for Animal Recording Dairy Cattle Milk Recording Working Group). 2017. Section 2 - Guidelines for Dairy Cattle Milk Recording. Accessed Jan. 13, 2025. <https://www.icar.org/Guidelines/02-Overview-Cattle-Milk-Recording.pdf>.
- Ito, N., T. Ito, A. Kromminga, A. Bettermann, M. Takigawa, F. Kees, R. H. Straub, and R. Paus. 2005. Human hair follicles display a functional equivalent of the hypothalamic-pituitary-adrenal (HPA) axis and synthesize cortisol. *FASEB J.* 19:1332–1334. <https://doi.org/10.1096/fj.04-1968fje>.
- Jensen, A. L., M. B. Petersen, and H. Houe. 1993. Determination of the fructosamine concentration in bovine serum samples. *J. Vet. Med. A* <https://doi.org/10.1111/j.1439-0442.1993.tb00607.x>.
- Kawamoto, M., J. J. Kaneko, A. A. Heusner, E. C. Feldman, and I. Koizumi. 1992. Relation of fructosamine to serum protein, albumin, and glucose concentrations in healthy and diabetic dogs. *Am. J. Vet. Res.* 53:851–855. <https://doi.org/10.2460/ajvr.1992.53.05.851>.
- Klasing, K. C. (1985). Influence of stress on protein metabolism. Pages 269–280 in *Animal Stress*. Springer Nature. https://doi.org/10.1007/978-1-4614-7544-6_15.
- MacKay, J. R. D., M. J. Haskell, J. M. Deag, and K. van Reenen. 2014. Fear responses to novelty in testing environments are related to day-to-day activity in the home environment in dairy cattle. *Appl. Anim. Behav. Sci.* 152:7–16. <https://doi.org/10.1016/j.applanim.2013.12.008>.
- McEwen, B. S., and P. J. Gianaros. 2010. Central role of the brain in stress and adaptation: Links to socioeconomic status, health, and disease. *Ann. N. Y. Acad. Sci.* 1186:190–222. <https://doi.org/10.1111/j.1749-6632.2009.05331.x>.
- McParland, S., and D. P. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. *J. Dairy Sci.* 99:4056–4070. <https://doi.org/10.3168/jds.2015-10051>.
- Megahed, A. A., M. W. H. Hiew, and P. D. Constable. 2018. Clinical utility of plasma fructosamine concentration as a hypoglycemic biomarker during early lactation in dairy cattle. *J. Vet. Intern. Med.* 32:846–852. <https://doi.org/10.1111/jvim.15049>.
- Meyer, J. S., and M. A. Novak. 2012. Minireview: Hair cortisol: A novel biomarker of hypothalamic-pituitary-adrenocortical activity. *Endocrinology* 153:4120–4127. <https://doi.org/10.1210/en.2012-1226>.
- Moberg, G. P., C. O. Anderson, and T. R. Underwood. 1980. Ontogeny of the adrenal and behavioral responses of lambs to emotional stress. *J. Anim. Sci.* 51:138–142. <https://doi.org/10.2527/jas1980.511138x>.
- Mormède, P., S. Andanson, B. Aupérin, B. Beerda, D. Guémené, J. Malmkvist, X. Manteca, G. Manteuffel, P. Prunet, C. G. van Reenen, S. Richard, and I. Veissier. 2007. Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. *Physiol. Behav.* 92:317–339. <https://doi.org/10.1016/j.physbeh.2006.12.003>.
- Mostafavi, M., H. A. Seifi, M. Mohri, and A. Jamshidi. 2014. Evaluation of fructosamine as a new biomarker for diagnosis of hepatic lipidosis in dairy cows. *Anim. Prod. Sci.* 55:1005–1010. <https://doi.org/10.1071/AN14108>.
- Nedić, S., M. Pantelić, S. Vranješ-Durić, D. Nedić, L. Jovanović, N. Cebulj-Kadunc, S. Kobal, T. Snoj, and D. Kirovski. 2017. Cortisol concentrations in hair, blood and milk of Holstein and Busha cattle. *Slov. Vet. Res.* 54:163–172. <https://doi.org/10.26873/SVR-398-2017>.
- Newsholme, E., and A. Leech. 2010. *Functional Biochemistry in Health and Disease*. Accessed Mar. 20, 2025. <https://www.wiley.com/en-us/Functional+Biochemistry+in+Health+and+Disease-p-9780471931652>.
- Otten, W., S. Heimbürge, A. Tuchscherer, and E. Kanitz. 2022. The age of hair matters – the incorporation of cortisol by external contamination is enhanced in distal hair segments of pigs and cattle. *Animal* 16:100495. <https://doi.org/10.1016/j.animal.2022.100495>.
- Peric, T., M. Corazzin, A. Romanzin, S. Bovolenta, A. Prandi, M. Montillo, and A. Comin. 2017. Cortisol and DHEA concentrations in the hair of dairy cows managed indoor or on pasture. *Livest. Sci.* 202:39–43. <https://doi.org/10.1016/j.livsci.2017.05.020>.
- Proudfoot, K. L., D. M. Weary, S. J. LeBlanc, L. K. Mamedova, and M. A. G. von Keyserlingk. 2018. Exposure to an unpredictable and competitive social environment affects behavior and health of transition dairy cows. *J. Dairy Sci.* 101:9309–9320. <https://doi.org/10.3168/jds.2017-14115>.
- Rahko, T., T. V. Kalima, and H. Saloniemi. 1987. Serum fructosamine levels in dairy cows related to metabolic status in early lactation. *Acta Vet. Scand.* 28:291–298. <https://doi.org/10.1186/BF03548596>.

- Ramirez Mauricio, M. A., M. Berton, N. Amalfitano, D. Giannuzzi, S. Pegolo, S. Raniolo, M. Nocetti, R. Negrini, M. Coppa, B. Martin, S. Schiavon, L. Gallo, E. Sturaro, and A. Cecchinato. 2025. Leveraging milk mid-infrared spectroscopy to authenticate animal welfare, farming practices, and dairy systems of Parmigiano Reggiano cheese. *J. Dairy Sci.* 108:2642–2657. <https://doi.org/10.3168/jds.2024-25466>.
- Romero, L. M. 2004. Physiological stress in ecology: Lessons from biomedical research. *Trends Ecol. Evol.* 19:249–255. <https://doi.org/10.1016/j.tree.2004.03.008>.
- Sapolsky, R. M., L. M. Romero, and A. U. Munck. 2000. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr. Rev.* 21:55–89. <https://doi.org/10.1210/edrv.21.1.0389>.
- Schubach, K. M., R. F. Cooke, A. P. Brandão, K. D. Lippolis, L. G. T. Silva, R. S. Marques, and D. W. Bohnert. 2017. Impacts of stocking density on development and puberty attainment of replacement beef heifers. *Animal* 11:2260–2267. <https://doi.org/10.1017/S1751731117001070>.
- Sorondo, M. L., and A. Cirio. 2009. Evaluation of the serum fructosamine test to monitor plasma glucose concentration in the transition dairy cow. *J. Dairy Res.* 76:173–178. <https://doi.org/10.1017/S0022029908003750>.
- 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>.
- Tallo-Parra, O., A. Carbajal, L. Monclús, X. Manteca, and M. Lopez-Bejar. 2018. Hair cortisol and progesterone detection in dairy cattle: Interrelation with physiological status and milk production. *Domest. Anim. Endocrinol.* 64:1–8. <https://doi.org/10.1016/j.domaniend.2018.02.001>.
- Tallo-Parra, O., M. Lopez-Bejar, A. Carbajal, L. Monclús, X. Manteca, and M. Devant. 2017. Acute ACTH-induced elevations of circulating cortisol do not affect hair cortisol concentrations in calves. *Gen. Comp. Endocrinol.* 240:138–142. <https://doi.org/10.1016/j.ygcen.2016.10.007>.
- Tallo-Parra, O., X. Manteca, M. Sabes-Alsina, A. Carbajal, and M. Lopez-Bejar. 2015. Hair cortisol detection in dairy cattle by using EIA: Protocol validation and correlation with faecal cortisol metabolites. *Animal* 9:1059–1064. <https://doi.org/10.1017/S1751731115000294>.
- Trevisi, E., and G. Bertoni. 2009. Some physiological and biochemical methods for acute and chronic stress evaluation in dairy cows. *Ital. J. Anim. Sci.* 8(Suppl. 1):265–286. <https://doi.org/10.4081/ijas.2009.s1.265>.
- van der Kooij, M. A. 2020. The impact of chronic stress on energy metabolism. *Mol. Cell. Neurosci.* 107:103525. <https://doi.org/10.1016/j.mcn.2020.103525>.
- Walker, S. L., R. F. Smith, D. N. Jones, J. E. Routly, and H. Dobson. 2008. Chronic stress, hormone profiles and estrus intensity in dairy cattle. *Horm. Behav.* 53:493–501. <https://doi.org/10.1016/j.yhbeh.2007.12.003>.
- Webster, A. J. F. 2001. Farm animal welfare: The five freedoms and the free market. *Vet. J.* 161:229–237. <https://doi.org/10.1053/tvjl.2000.0563>.
- Welp, T., J. Rushen, D. L. Kramer, M. Festa-Bianchet, and A. M. B. De Passillé. 2004. Vigilance as a measure of fear in dairy cattle. *Appl. Anim. Behav. Sci.* 87:1–13. <https://doi.org/10.1016/j.applanim.2003.12.013>.
- Westgard, S., V. Petrides, S. Schneider, M. Berman, J. Herzogenrath, and A. Orzechowski. 2017. Assessing precision, bias and sigma-metrics of 53 measurands of the Alinity ci system. *Clin. Biochem.* 50:1216–1221. <https://doi.org/10.1016/j.clinbiochem.2017.09.005>.
- Zhang, L., C. Li, F. Dehareng, C. Grelet, F. Colinet, N. Gengler, Y. Brostaux, and H. Soyeurt. 2021. Appropriate data quality checks improve the reliability of values predicted from milk mid-infrared spectra. *Animals (Basel)* 11:533. <https://doi.org/10.3390/ani11020533>.

ORCIDS

- C. Grelet, <https://orcid.org/0000-0003-3313-485X>
H. Simon, <https://orcid.org/0009-0008-7430-8813>
J. Leblois, <https://orcid.org/0000-0001-5112-0262>
O. Christophe, <https://orcid.org/0000-0002-9037-8079>
K. Goossens, <https://orcid.org/0000-0001-9064-0324>
E. Chevaux, <https://orcid.org/0000-0002-2375-5487>
N. Gengler, <https://orcid.org/0000-0002-5981-5509>
F. Dehareng, <https://orcid.org/0000-0002-6733-4334>