J. Dairy Sci. 108:13085–13098
https://doi.org/10.3168/jds.2025-26829

Going deeper in lipolysis: Prediction of individual free fatty acid contents in milk by mid-infrared spectroscopy

Octave Christophe,[1,2]* Romain Reding,[3] Julie Leblois,[4] Clément Grelet,[1] Denis Pittois,[5] Hélène Soyeurt,[2] Cedric Guignard,[5] and Frédéric Dehareng[1]*
[1]Walloon Agricultural Research Center (CRA-W), 5030 Gembloux, Belgium
[2]TERRA Research and Teaching Centre, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium
[3]CONVIS, 9084 Ettelbruck, Luxembourg
[4]Elevéo asbl, AWE Groupe, 5590 Ciney, Belgium
[5]Department Environmental Research and Innovation (ERIN), Luxembourg Institute of Science and Technology, 4422 Belvaux, Luxembourg

ABSTRACT

The dairy sector deals with a recurring issue: a taste alteration due to degradation of fat, commonly called lipolysis. Lipolysis occurs after milking, through the physical shocks induced by freezing, pumping, transfer, and storage of the milk. Physical breaking of fat globules makes triglycerides accessible to enzymes and degrades into free fatty acids (FFA). Among them, the volatile short-chain FFA lead to organoleptic issues through undesired tastes. Currently, methods to measure FFA are complex and time-consuming. Easier methods provide incomplete information on the total amount of FFA, or a combination of these. To bring a new way of preventive and corrective action for dairies and farmers, this study aims to develop predictive models based on milk Fourier transform mid-infrared spectroscopy (FT-MIR) to quantify individual FFA, with a special focus on short-chain FFA. A total of 792 milk samples were collected from individual cows and in bulk tanks in 4 different countries (Luxembourg, Belgium, France, and Germany) and analyzed by FT-MIR spectroscopy as well as GC coupled with tandem mass spectrometry. When performing a random sampling of individual cows, and therefore relying only on spontaneous lipolysis, most of the samples showed that short-chain FFA under the limit of quantification and derived FT-MIR models had poor performance. In a second step, lipolysis was induced by 4 different mechanical processes of milk. Storage after milk homogenization demonstrated a clear increase of short-chain FFA values leading to improved predictive performance of models. The predictions of short-chain FFA could be used for rough screening on the dairy population, especially C4 and total short-chain FFA as their models had a validation R^2 above 0.7. Unfortunately, the mid- and long-chain FFA models showed poor predictive performance. However, those results can still be considered positive because they may enable a better monitoring of taste defect mainly occurring from short-chain FFA.
Key words: milk, lipolysis, free fatty acid, MIR, chemometrics

INTRODUCTION

In the dairy sector, lipolysis is a well-known process degrading the flavor quality of milk (Connolly et al., 1979), butter (Mcdaniel et al., 1969), or cheese (Ivanov et al., 2021). Nevertheless, in the case of certain blue and hard cheeses, high lipolysis levels is considered positive and results in a more pronounced flavor profile, as the free fatty acids (**FFA**) serve as precursors to various flavor compounds (Collins et al., 2003). Lipolysis in milk is a reaction catalyzed by an enzyme called lipoprotein lipase (**LPL**; Delosière et al., 2023) that performs the hydrolysis of triglycerides (lipids). This reaction produces FFA that affect the flavor of dairy products (Deeth, 2019). This phenomenon can be attributed to 2 sources of LPL: The main one is an endogenous enzyme present in milk, although it may also originate from bacterial lipases produced by psychotropic bacteria (Deeth, 2022). The occurrence of microbial lipolysis may result in significant off-flavor effects when the milk colony exceeds 10^6 cfu/mL. When microbial quality is maintained at levels below 10^5 cfu/mL, microbial lipolysis is unlikely to take place (Santos et al., 2003).

The inherent lipolytic enzyme (LPL) found in milk is present in adequate amounts to facilitate significant hydrolysis of milk fat. However, the absence of this phenomenon in practical scenarios indicates that the actual level of lipolytic enzymes, regardless of whether they are of native or microbial origin, is not the main factor

Received April 24, 2025.
Accepted August 22, 2025.
*Corresponding authors: o.christophe@cra.wallonie.be and 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.

influencing the tendency of different milks to undergo lipolysis (Downey, 1980). This is attributed to the milk fat globule membrane (**MFGM**), which is a distinctive structure encasing the triglycerides in milk, offering both stability and protection. This membrane is composed of a complex mix of proteins, phospholipids, and other bioactive components. In normal conditions, milk has a natural protection against the catalytic reaction caused by lipase; MFGM acts as a barrier and limits the access of lipases to the triglycerides contained within the globule. Lipolysis increases when this barrier weakens, and this leads to the distinction of the 2 types of lipolysis: induced lipolysis and spontaneous lipolysis. Induced lipolysis refers to the process of lipolysis that occurs in raw milk due to mechanical agitation or turbulence. Lipoprotein lipase is primarily associated with the protein fraction of milk. Typically, this enzyme is unable to hydrolyze triglycerides and release FFA given that the triglycerides are protected by their size and the phospholipid membrane of the fat globule (Shimizu et al., 1982). Mechanical action damages the MFGM, thereby allowing LPL to access triglycerides, which are subsequently hydrolyzed. Breaking these membranes or decreasing the size through a mechanical induction of milk (or both) should be the best way to enhance milk FFA concentration. Spontaneous lipolysis is defined as lipolysis without any mechanical agitation of the raw milk.

Factors affecting spontaneous lipolysis are related to individual cow characteristics such as the stage of lactation (Ortiz et al., 1970), feeding and nutrition (Gholson et al., 1966; Kodgev and Rachev, 1970; Astrup et al., 1980; Vanbergue et al., 2017), estrus cycle (Wells et al., 1969), season (Kodgev and Rachev, 1970), milk production (Jellema and Schipper, 1975), and mastitis (Luhtala and Antila, 1968; Murphy et al., 1989). This natural phenomenon could be exacerbated by induced lipolysis, particularly due to rapid or extensive cooling (Johnson and Von Gunten, 1962).

Nowadays, lipolysis may present a greater challenge due to the increased use of automatic milking systems (**AMS**). Indeed, the increased milking frequency associated with AMS can lead to higher lipolysis levels when compared with conventional milking methods, and there is a noticeable increase in the occurrence of lipolysis (Wiking, 2005). This phenomenon can be attributed, in part, to the fact that more frequent milking decreases the duration available for milk fat globules to stabilize within the mammary gland, thereby increasing their vulnerability to mechanical disruption during the milking process (Abeni et al., 2005).

The resulting products of the lipolysis, FFA, have been globally recognized as important contributors to rancid flavor defects in dairy products (Kintner and Day, 1965). Incomplete lipolysis, which leads to the formation of partial glycerides (either mono- or di-glycerides), may influence the foaming and creaming properties of a dairy product (Deeth, 2006). Short-chain free fatty acids (**SCFFA**) including butyric, caproic, and caprylic acids are known to have a strong impact on the rancid taste of the milk, whereas this is not the case for long-chain free fatty acids (**LCFFA**; Scanlan et al., 1965; Duncan et al., 1991). Notably, SCFFA possess a lower sensory threshold compared with their long-chain counterparts (Duncan et al., 1991). Furthermore, the perception of rancidity is significantly influenced by FFA within the C4 to C12 range (González-Córdova and Vallejo-Cordoba, 2003). Depending on the person, increased FFA concentration in milk could be described as rancid, butyric, bitter, unclean, soapy, or astringent (Ray et al., 2013).

However, the absolute quantification of these individual SCFFA, responsible for taste alteration, is very difficult. Historically, lipolysis was quantified with the method referenced by the Bureau of Dairy Industries (Evers, 2003) based on the measurement of the fat acidity. This method recovers about 50% to 70% of the milk total FFA content.

In contrast, the analysis of a wide range of FFA is now possible by GC coupled with tandem mass spectrometry (**GC-MS/MS**). Despite its performance, this method is time-consuming, expensive, and difficult to apply for routine analysis on a large set of samples. Alternatively, infrared-based models have been recognized for their effectiveness in the dairy sector for over a decade. Predictive models are regularly employed to enhance farm management practices. Specifically, some models are used to analyze milk fatty acid composition (Soyeurt et al., 2006), minerals (Christophe et al., 2021), acetone (Grelet et al., 2016), protein fraction (Macedo Mota et al., 2023), and other indirect animal traits such as methane (Dehareng et al., 2012; Vanlierde et al., 2021), nitrogen use efficiency (Grelet et al., 2020), and DMI (Tedde et al., 2021). The current routine methodology to detect lipolysis is also based on Fourier transform mid-infrared (**FT-MIR**) analysis of milk through manufacturer equations. However, these models only enable the estimation of the total amount of FFA. Because of their different impacts on flavor and taste of products, there is an interest for the industry to quantify individual FFA rather that the entire lipolysis. This study will focus on predicting individual FFA in milk using FT-MIR spectroscopy from C4 to C18:1. The developed models are expected to enhance the understanding of lipolysis and improve the differentiation of undesirable flavors in milk associated with SCFFA.

The study will include 2 different datasets, a first based on spontaneous lipolysis and the second based on induced lipolysis. Both datasets will be merged to develop several FFA models from milk samples sourced from 4 different

countries (Luxembourg, Belgium, Germany, and France), from 5 different breeds and the 2 main milking systems.

Given the challenges associated with sourcing milk samples characterized by high spontaneous lipolysis, mechanical induction is mandatory to generate samples exhibiting high concentration in FFA to improve the variability of the training set used to build predictive models. This induced lipolysis will be achieved by applying various mechanical treatments, including vortexing, bubbling, and varying storage durations posthomogenization.

By providing a practical and efficient method for individual FFA quantification, this study aims to support dairy producers in monitoring and controlling lipolysis-related flavor defects, ultimately contributing to improved product quality and consumer satisfaction.

MATERIALS AND METHODS

Study Design

The initial phase involved the formulation of an equation aimed at predicting FFA levels based on samples derived from spontaneous lipolysis only. The subsequent phase sought to enhance the training set's variability by incorporating samples subjected to induced lipolysis through 4 mechanical perturbations.

Milk Sampling and General Methodology

All samples were collected following the guidelines edited by the International Committee for Animal Recording (ICAR, 2022) and with ICAR-approved milk samplers. Milk samples were collected from 5 breeds (Montbéliarde, Holstein, Brown Swiss, Dual Purpose, Belgian Blue, and Simmental) in 4 different countries (Luxembourg, Belgium, France, and Germany) by different local DHI organizations. A significant number of individual cow milk samples were collected to account for variability across various factors, including breed, country of origin, DIM, and parity, in relation to spontaneous lipolysis. Additionally, one bulk milk sample per farm was sampled to include the impact of the farm's milking infrastructure. The complete dataset resulted in 739 individual cow samples and 53 bulk milk samples, 314 of them allocated to spontaneous lipolysis and 478 related to the induced lipolysis.

The samples, which concentrate on spontaneous lipolysis, were divided into 2 aliquots: one for reference analysis of FFA and the other for FT-MIR analysis. Aliquoting was realized directly in the milking parlor when milk was still at udder temperature to prevent fat-aqueous phasing and unperfect aliquoting. One aliquot was analyzed in the laboratory of the country's DHI using a milk mid-infrared spectrometer and the second aliquot was frozen at −18°C and sent at −18°C to the Luxembourg Institute of Science and Technology laboratory (Ettelbruck, Luxembourg) for reference FFA analysis.

In contrast, a distinct procedure was employed for the samples undergoing an induced lipolysis step. These samples were not aliquoted; instead, they were sent fresh to the CRA-W laboratory in Gembloux, Belgium, where they were heated to a temperature of 40°C ± 2°C before undergoing mechanical treatments. Following the mechanical treatment, the samples were subsequently divided into 2 aliquots to facilitate simultaneous reference and spectral analyses. Both analyses are performed at the same time. This approach was essential to mitigate any potential increase in lipolysis over time. Further details regarding the analysis and mechanical treatment are provided in the subsequent sections.

Mechanical Induction of Lipolysis

Different treatments of the milk were induced to enhance the FFA concentration and to be at realistic levels. To improve the FFA concentration, a thorough comprehension of the induced lipolysis principle through different mechanical induction are studied to identify the optimal approach for inducing lipolysis to elevate the concentration of FFA. The first way to induce the mechanical disruption of milk was the shaking of the milk with a vortex. A subset consisting of 10 samples, each derived from 10 distinct cows, was used. Before shaking, the samples were aliquoted into 2 samples: One was frozen for 1 h at −20°C and the other was kept fresh at 4°C. Different shaking durations of 30, 60, and 120 s were applied on each sample (which had been frozen or not) with an RSlab-6 vortex (Rogo-Sampaic laboratory, Wissous, France).

Another tested method of inducing lipolysis was bubbling. It is known in the dairy industry than the injection of air in the milking process leads to increasing lipolysis. Bubbling was carried out in a nitrogen environment to prevent oxidation of the milk fat. Different pressures were tested, but only one pressure was used to avoid spitting and the samples were later analyzed for individual FFA concentration. A total of 25 milk samples were collected in different farms from Luxembourg and Belgium. The samples were subjected to bubbling for a duration of 2 min at a pressure of 40 KPa.

The last tested way to induce the lipolysis was high-pressure homogenization associated with a long time of conservation between homogenization and analysis. The homogenization step was performed with the homogenizer of the Delta instrument FTIR lactoscope (PerkinElmer, Drachten, the Netherlands). This technique involves pumping milk through tiny openings at extremely high pressure (20 MPa), resulting in a reduction in the size

distribution of the milk fat globules. The opening size of the homogenizer and the pressure are not detailed by the supplier, but particle sizes after homogenization were observed between 0.2 and 2 µm.

This apparatus works with 2 steps of homogenization. The first step of the homogenization allows decreasing the size of milk fat particles, and the second homogenization allows spreading out the particles to avoid agglomeration. This process also damaged the MFGM (Kiełczewska et al., 2021). To check the particle size distribution and the agglomeration of the particles, some pictures were taken of several milks before and after homogenization. The milk was heated at 40°C and a droplet was placed on a microscope slide. The slide was observed through a contrast phase with a confocal microscope (Zeiss Axio Imager A1m, Jena, Germany). Different time conservation between homogenization and analysis was tested on the samples: 0, 0.5, 1, 2, 4, 12, 24, and 48 h, and the samples were stored at 4°C.

Reference Analysis for Free Fatty Acids in Milk

In parallel to the recording of the milk FT-MIR spectra, the reference analysis was conducted. A series of 5 steps was required to analyze the FFA content in milk. The first step was the preparation of the sample by collecting 2 mL of the milk at 40°C to ensure homogeneity, then transferring them to a 15-mL tube. Subsequently, 50 µL of a solution containing 1 g/L heptanoic (C7) and nonanoic (C9) acids (internal standards) in n-hexane was added, followed by a gentle homogenization for 1 min. The second step was a liquid-liquid extraction, where 5 mL of a 1:1 mixture of diethylether and n-hexane was added to the sample and gently mixed for 30 s. Afterward, 150 µL of 2.5 *M* sulfuric acid solution was added and gently mixed for another 30 s. The sample was then centrifuged at 3,000 × g for 300 s at room temperature. The supernatant was carefully transferred using a positive displacement pipette into a separate tube containing anhydrous sodium sulfate to remove any trace of water. A second extraction (without the sulfuric acid addition) was then repeated, and the supernatant was collected and added on top of the first extract. The third step consisted of the isolation of FFA from the extracts by solid-phase extraction (**SPE**) using an Aspec GX-274 (Gilson, Middleton, WI). Six-milliliter aminopropyl SPE columns (Sep-Pak Vac 6cc 500 mg NH_2 cartridges, Waters) were conditioned with 6 mL of n-hexane, then loaded with 9 mL of extract. A first elution with 6 mL of a 20:80 diethylether and n-hexane mixture was performed to remove the triglycerides. The FFA were eluted with 5 mL of a diethylether and 2% formic acid solution, and this fraction was collected in polytetrafluoroethylene–capped tubes. The fourth step was the derivatization of the FFA in 1.5-mL vials. To 50 µL of the extracted sample, 10 µL of 10 mg/L undecanoic acid (as internal standard) and 50 µL of N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide with 1% tert-butyldimethylchlorosilane were added. The vials were mixed and heated for 30 min at 70°C. The FFA analysis was performed on a triple-quadrupole GC-MS/MS (Agilent GC-MSMS 7000). The injection and oven parameters are presented in the Supplemental File S1 (see Notes).

Milk Spectral Analysis

Analyses were performed locally, in regions of the different DHI partners, on different brands and models of spectrometers: FT+ and FT7 (Foss, Hillerød, Denmark) and Standard Lactoscope FT-MIR automatic (PerkinElmer [previously Delta Instruments], Drachten, the Netherlands). The spectra were standardized using the European Milk Recording method, as outlined in the study conducted by Grelet et al. (2015). The mid-infrared (**MIR**) spectra were pretreated by a first derivative with a gap of 5 wavenumbers. The selected spectral area consisted of 212 wavenumbers from 968.1 to 1,577.5 cm^{-1}, 1,731.8 to 1,762.6 cm^{-1}, 1,781.9 to 1,808.9 cm^{-1}, and 2,831.0 to 2,966.0 cm^{-1} to exclude spectral areas not reproducible between instruments (Grelet et al., 2021) and the absorbance values were mean-centered before being modeled. Each sample was analyzed in triplicate on the spectrometer; only the third replicate was taken for building the different model to avoid carryover.

Model Development

Three datasets were employed to build 3 classes of FFA models: spontaneous lipolysis model, induced lipolysis model, and entire lipolysis model. For each dataset, every FFA model was computed using partial least squares (**PLS**) regressions using Python language (Python, 2024) and the Scikit-learn package (Pedregosa et al., 2011) to construct the most accurate predictive models. This algorithm amalgamates features from predictors and responses to construct latent variables. It maximizes covariance between predictors and responses iteratively, generating independent components that capture shared variance (Fornell and Cha, 1994). The use of PLS is particularly advantageous for reducing the dimensionality of high-dimensional, multicollinear datasets, thereby improving predictive accuracy while limiting overfitting. It is especially effective in situations where the quantity of predictors (i.e., wavenumber) surpasses the number of observations, such as in spectroscopy, as its first step is a dimension reduction that adeptly manages multicollinearity and noise while facilitating predictive analysis (Fornell and Cha, 1994). To prevent overfitting, the

maximum number of PLS latent variables was limited to 20. The optimal number of factors was determined by selecting the minimum number of PLS latent variables that yielded the highest R^2 and the lowest root mean square error (**RMSE**) during 10-fold cross-validation. Each model was evaluated through cross-validation and the same external validation. Before calibration, the entire dataset, including samples with mechanical treatment, was therefore randomly divided into 2 sets: the calibration set (80% of milk) and the validation set (20% of milk). The external validation is an independent prediction of the test set and the metrics are calculated with this independent set of prediction to approach real performances. The 10-fold cross-validation was performed with the calibration set, which is randomly divided into 10 symmetrical sets. The model was calibrated with 9 of the 10 sets and the last is predicted; this was done 10 times and the R^2 of cross-validation (**R^2cv**) and root mean square error of cross-validation (**RMSEcv**) were the average of these 10 results.

To prevent the inclusion of anomalous spectral data, an initial phase of data cleaning was conducted. On the calibration set, spectra that displayed a standardized Mahalanobis distance (**GH**) surpassing 30 were discarded from the dataset. This high threshold for GH was purposefully established to enhance the model's robustness by allowing a great spectral variability, while concurrently ensuring the exclusion of any abnormal spectra.

A second round of data cleaning related to the reference sample was performed during model development. Indeed, the samples with an error of prediction 2.5 times higher than the standard deviation of the error predictions were discarded during the cross-validation step because they were considered outliers. The "final calibration" model was calibrated excluding these samples, whereas no samples were removed from the validation dataset. This selection methodology was applied for each FFA in the calibration process. The samples removed through these 2 selection rules were quantified as a percentage of outliers.

Pasteurization Test

The goal of the study was to predict FFA content in milk, and to do so several mechanical treatments were performed on milk. Those mechanical treatments affected the physical structure of milk (e.g., breaking down the fat globule size). The FT-MIR spectra of milk are not only affected by chemical composition, but also by physical structure, such as fat globule size interacting with light diffraction and so-called Christiansen effects (Goulden, 1961; Di Marzo and Barbano, 2016). Therefore, a potential artifact of inducing lipolysis by mechanical treatment is that both effects could be confused by FT-MIR models and therefore predict time after homogenization rather than the concentration of FFA. To evaluate such a potential effect, milks were first pasteurized, to inactivate LPL enzymes, before applying mechanical treatment. This will enable observation of the effect of mechanical treatment only on FFA contents. For this a total of 10 milk samples sourced from various farms in Belgium were selected and subjected to pasteurization for 10 min at 72°C, with each sample divided into 2 aliquots (A and B). When the temperature reached 40°C following the pasteurization, aliquot A from each sample was analyzed using MIR and GC-MS/MS techniques. For the other aliquot B, a homogenization step was performed, followed by a storage period of 4 h at 4°C. Aliquots B from each sample were then also analyzed by using MIR and GC-MS/MS techniques after heating at 40°C. The analysis comparing predictions with reference data for both groups may reveal the effectiveness of the model to predict FFA and not homogenization.

RESULTS AND DISCUSSION

Data Distribution and Availability

The complete dataset comprises 792 samples, with 314 samples coming from spontaneous lipolysis and 478 samples derived from induced lipolysis. Table 1 presents the descriptive statistics of this dataset. The minimum values correspond to their intrinsic limit of quantification (**LOQ**), and FFA availability refers to the percentage of samples with values above this minimum, that is, the proportion of quantifiable samples. The coefficient of variation for SCFFA demonstrates superior variability within the induced lipolysis dataset. The mean and median values for SCFFA are notably close to the minimum, indicating a limited availability of data for the spontaneous lipolysis dataset.

Figure 1 illustrates the data distribution for each FFA. In the spontaneous lipolysis dataset, the figure reveals a skewed (exponential) distribution for SCFFA and mid-chain FFA (**MCFFA**), whereas LCFFA exhibits a logarithmic distribution. Conversely, the induced lipolysis dataset also shows a skewed distribution close to a gamma distribution, with the maximum value shifted toward higher levels for both SCFFA and MCFFA. Although the distribution for LCFFA in the induced dataset remains unchanged compared with the LCFFA distribution obtained from spontaneous lipolysis, it appears to be flattened.

Mechanically Induced Lipolysis of Milk

Because the variability of FFA obtained from the spontaneous lipolysis was not sufficient to build a

Table 1. Descriptive statistics of the different datasets for each FFA[1]

Fatty acid (mg/L)	Spontaneous lipolysis dataset						Induced lipolysis dataset						Entire lipolysis dataset					
	Min	Median	Max	Mean	SD	CV (%)	Min	Median	Max	Mean	SD	CV (%)	Min	Median	Max	Mean	SD	CV (%)
C4	2.0	2.2	40.2	4.2	5.0	119	2.0	26.2	638.68	42.0	51.5	122	2.0	12.9	638.68	27.1	44.2	163
C6	3.5	3.5	30.7	4.4	3.1	70	3.5	20.7	395.75	27.9	30.2	108	3.5	9.9	395.75	18.6	26.2	141
C8	2.5	2.5	19.5	3.3	2.1	64	2.5	14.7	225.28	19.2	20.5	107	2.5	7.6	225.28	12.9	17.7	138
C10	3.5	3.6	28.1	5.1	3.1	62	3.5	13.6	325.9	19.4	23.7	122	3.5	7.7	325.9	13.7	19.8	144
C12	5.5	5.5	21.2	6.1	1.8	29	5.5	8.4	162.56	13.8	15.2	110	5.5	5.6	162.56	10.8	12.4	116
C14	3.5	7.1	47.8	8.1	5.3	66	3.5	10.7	203.7	18.3	21.3	117	3.5	8.5	203.7	14.2	17.6	124
C16	7.5	25.6	105.8	27.4	15.6	57	9.1	34.3	445.79	49.9	48.2	97	7.5	31.1	445.79	41.0	40.2	98
C18	7.5	25.5	93	29.6	15.9	54	8.2	21.8	140	29.7	21.4	72	7.5	23.2	140	29.7	19.4	65
C18:1	4.5	7.2	211.5	12.7	18.1	142	4.5	9.1	374.18	17.8	28.7	161	4.5	8.4	374.18	15.8	25.1	159
C18:2	3.0	3.0	22.9	3.6	2.0	55	3.0	3.0	59.27	4.4	4.5	103	3.0	3.0	59.27	4.1	3.7	92
SCFFA	8.0	8.4	90.4	12.0	10.0	83	8.0	64.3	1,259.71	89.1	96.5	108	8.0	31.1	1,259.71	58.5	84.1	144
MCFFA	12.5	16.5	87.3	19.2	9.5	49	12.5	33.6	692.16	51.5	57.6	112	12.5	22.1	692.16	38.7	47.8	124
LCFFA	22.5	66.3	354.3	73.3	44.6	61	25.1	72.5	771.09	101.8	92.2	91	22.5	68.8	771.09	90.5	78.2	86

[1]Min = minimum; Max = maximum.

model, adding induced lipolysis samples in the dataset was therefore necessary. Table 2 presents the availability and number of samples from the spontaneous lipolysis and from the different mechanical treatments applied to induce lipolysis. Because the availability of SCFFA was extremely bad (Table 2) and the bad taste of the milk is more linked to this kind of FFA, the way of doing the mechanical induction was focused on these FFA after a literature review. So, 3 approaches were tested in this study: (1) vortexing and freezing samples, (2) bubbling samples, and (3) impact of homogenization of milk fat globules in milk samples associated with different resting times before FFA analysis.

The first approach to induce lipolysis in milk was to vortex fresh and frozen milk samples. The data availability from this process is described in Table 2. The difference in FFA concentrations found following the vortex time and the kind of milk samples used are depicted in Figure 2. From this figure, we observed that the duration of vortexing directly influences the concentration of SCFFA. The fresh sample exhibited approximately a 100% increase in concentration, whereas the frozen sample showed a more substantial increase of 400%. The increase of FFA observed with fresh vortexing milk samples reached a stagnant state after 120 s. The inclusion of a freezing step enhanced the SCFFA concentration when the sample was shaken. The difference observed between fresh and frozen milk samples is explained by the damage sustained by the MFGM as a result of ice crystal formation during the freezing process. In both cases, with fresh and frozen milk samples, the concentration of SCFFA reached a plateau, indicating that this method is not suitable for increasing variability over 400%. Despite this acceptable increase, it is important to note that only fresh milk samples can be used for modeling purposes, as frozen milk samples decrease modeling performance with infrared analysis (Coppa et al., 2014). Indeed, the processes of freezing alter the milk's structure and then change the spectral signature compared with fresh milk. Therefore, the model would not be suitable to be applied on fresh milk. Because the freezing and vortexing were not suitable to increase the concentration of FFA, the samples from this vortexing test were not included in the final model.

The second approach was based on adding air in milk leading to a bubbling milk sample. Unfortunately, the results based on the 25 samples presented in Table 2 were highly unfavorable, as there was hardly any noticeable improvement in the increase of the concentration of FFA and the availability of data for modeling. Indeed, the availability of the bubble-induced milk was similar to milk coming from spontaneous lipolysis (Table 2). Although this methodology would be compatible with MIR, it can be concluded that this method is not suitable for inducing lipolysis effectively.

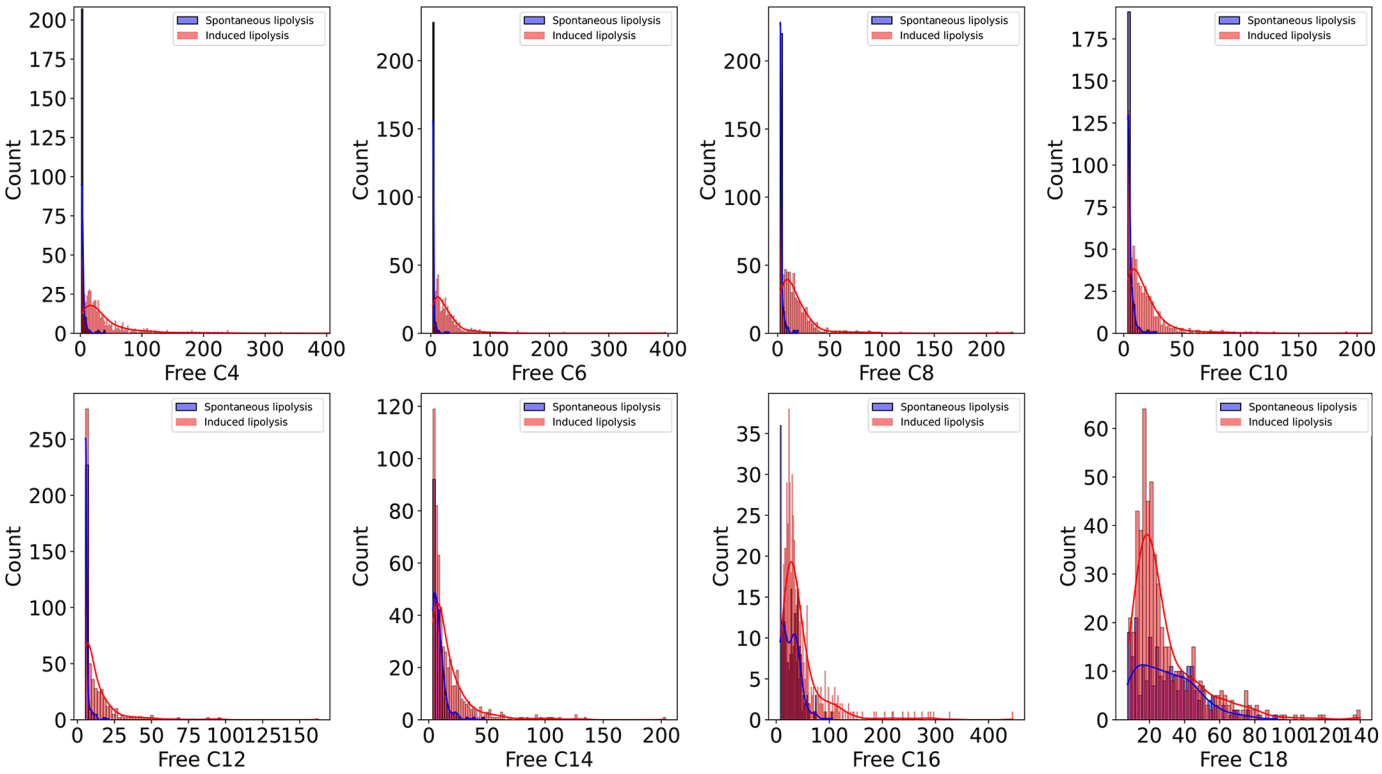

Figure 1. Data distribution count: difference between spontaneous lipolysis and induced lipolysis in free fatty acids (mg/L).

The third method to induce lipolysis was the homogenization of milk fat globules associated with different resting times before analysis, to enable LPL action. Figure 3 presents 2 images of milk from an optical microscope, illustrating its state before and after the homogenization process. Both images are scaled identically and reveal distinct particle size distributions. Specifically, the milk posthomogenization displays smaller and more aggregated particles, in contrast to the larger particles observed in the pre-homogenization sample. Additionally, a well-defined MFGM surrounding the fat particles is distinctly visible in the milk before homogenization, whereas this membrane is not clearly discernible in the homogenized milk. Figure 2 confirms the effect of homogenization on fat globule size and number suggesting that the triglycerides present in the fat globule are more available to be hydrolyzed by the LPL. By this explanation, the control of the time after homogenization is a crucial factor to monitor as the catalytic reaction could constantly increase through the time of storage. So, in this work, different times at 4°C after homogenization were applied on different milks. Figure 4 shows the

Table 2. Number and availability of FFA data depending on the type of lipolysis

Fatty acid (%)	Spontaneous lipolysis	Vortex		Bubbling	Storage after homogenization						
		Fresh	Frozen		0 h	0.5 h	1 h	2 h	4 h	12 h	24 h
N	314	10	10	20	15	223	125	51	16	15	14
C4	53	75	100	35	100	100	100	100	100	100	100
C6	24	70	100	15	100	99	99	100	100	100	100
C8	38	78	100	35	80	100	100	100	100	100	100
C10	51	80	84	55	73	97	85	100	100	100	100
C12	23	50	95	10	67	78	53	80	88	90	100
C14	80	60	100	65	100	93	96	100	100	94	100
C16	93	100	100	100	100	100	100	100	100	100	100
C18	100	100	100	100	100	100	100	100	100	100	100
C18:1	75	50	75	45	60	80	82	88	81	90	100
C18:2	30	18	56	0	13	37	7	49	6	78	50

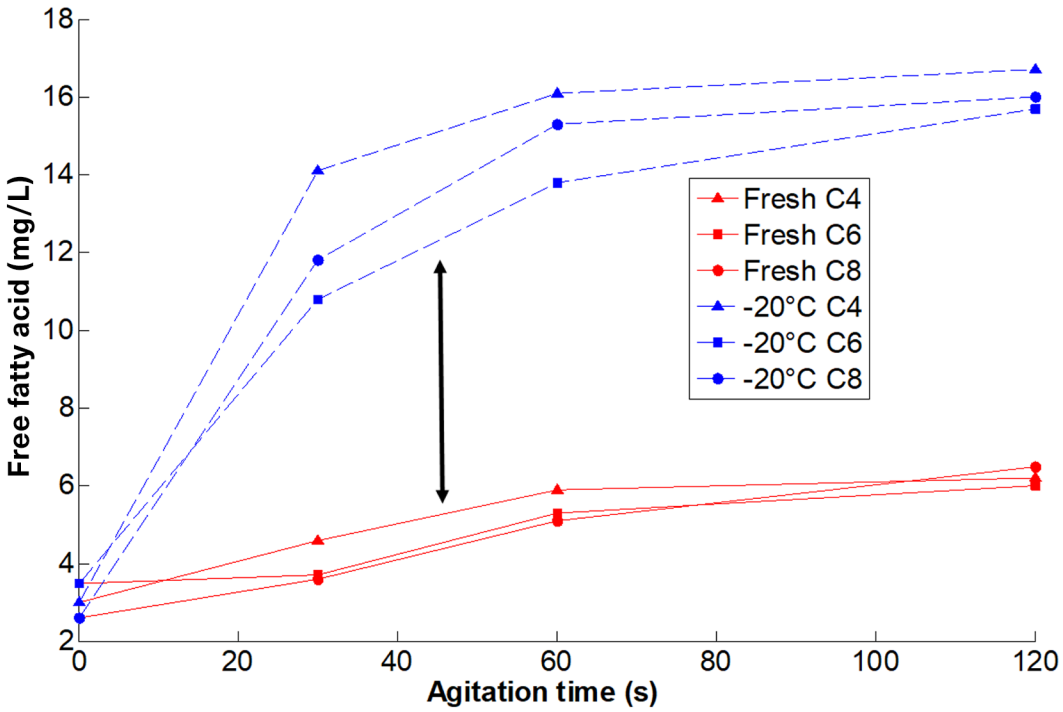

Figure 2. Change in the mean concentration of free fatty acids in fresh and frozen bulk milk. Arrow shows the difference between fresh and thawed samples.

change in SCFFA concentration following the resting time after homogenization for 5 different milks. The initial SCFFA concentration was under the LOQ. Figure 4 showed a clear and almost constant SCFFA increase according to the time of storage until 48 h of storage. Through the homogenization step, the concentration increase is mainly related to SCFFA. These findings are in accordance with the literature, indicating that LPL exhibits fatty acid positional specificity on the triglycerides rather than specificity for the fatty acid structure (Morley and Kuksis, 1977). The positional specificity of LPL is observed at the sn-1 and sn-3 positions (Somerharju et al., 1978). Given that SCFFA predominantly reside at the sn-3 position (Ouattara et al., 2004), LPL may appear to preferentially release these fatty acids.

In contrast to other methods of mechanical induction, homogenization can enhance lipolysis by up to 100-fold compared with nonhomogenized lipolysis. Depending on the time after homogenization, a wide range of values could be obtained. This technique is clearly better than

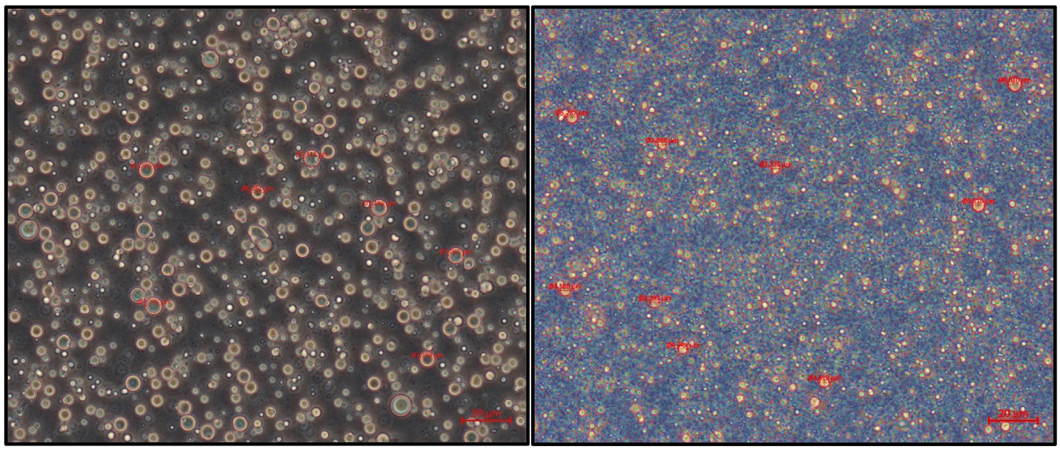

Figure 3. Milk fat globule micrographs before and after the homogenization step. Micrographs taken by an optical microscope.

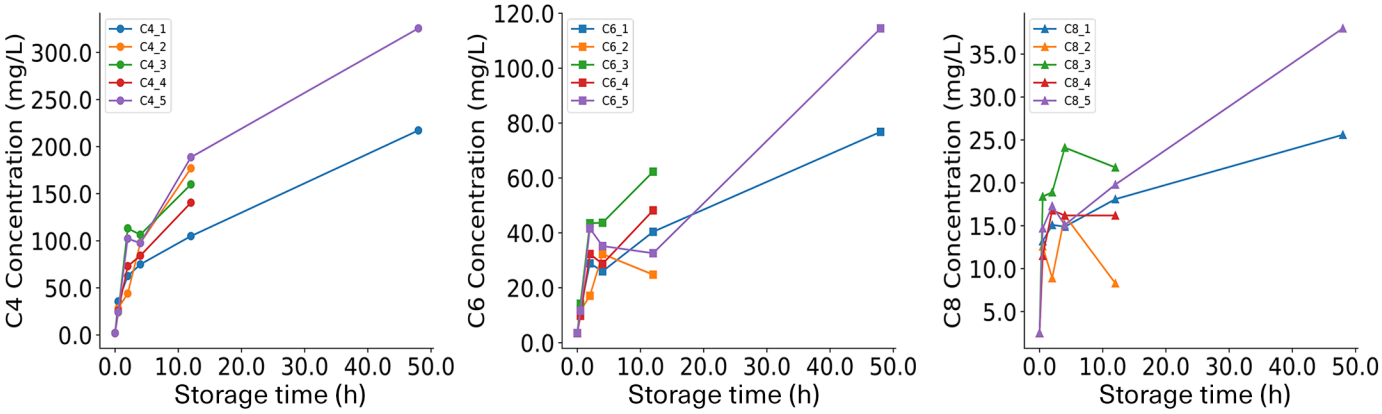

Figure 4. Changes in short-chain free fatty acids through storage after high-pressure homogenization.

the other approaches assessed to increase lipolysis as extreme values of SCFFA could be obtained. The other advantage is that this technique does not prevent performance of MIR spectral analysis. However, the largest increase is not in line with reality where such a high lipolysis rate does not take place. Therefore, to be in the same order of magnitude as the real lipolysis level, only a few samples with longer storage times will be included in models including 15 samples with 12 h and 14 samples with 24 h of resting time. Table 2 summarizes the number of samples and mechanical induction.

Free Fatty Acid Modeling

To build predictive models, the PLS algorithm was employed. A 10-fold cross-validation and a random 20% external validation set (160 samples) was used to evaluate the model performance. Table 3 shows the performance of the different models, and Figure 5 shows the cross-validation model performance and data distribution of summing models. The same external validation is applied on the different models to compare them. To assess the interest of adding induced lipolysis samples, PLS models were built from the dataset having (1) only spontaneous lipolysis samples, (2) only induced samples, and (3) both induced and spontaneous samples. The developed models demonstrate relatively heterogeneous performances.

For the spontaneous dataset, SCFFA models developed show low predictive performances, as they displayed a range of R^2cv varying from 0.01 to 0.09, indicating a poor fit of the model. The lack of variability in the dataset complicates the process of accurate modeling as the R^2 is directly affected by the range of reference values (Grelet et al., 2021). The MCFFA model demonstrates limited predictive capability for C10 and C12, yielding R^2cv values of 0.21 and 0.01, respectively. In contrast, C14 shows a moderate level of predictive performance, achieving a R^2cv of 0.43. Long-chain free fatty acids, specifically C16 and C18, demonstrate a moderate level of performance, as indicated by their higher availability, which yields R^2cv values of 0.53 and 0.69, respectively. The external validation results, as anticipated, demonstrate severely inadequate predictive performance. The calibration lacks any high values, leading to a situation where all high values are predicted as low, resulting in a graph that aligns closely with the x-axis leading to negative R^2 of validation. A negative R^2 indicates that the regression model performs worse in fitting the data compared with merely using the mean of the observed values. The random selection of samples, which encompasses variability in factors such as breed, country, DIM, and parity, exposes certain limitations within the modeling process. The difficulties associated with detecting spontaneous variability in FFA indicate that the current sampling methodology is inadequate.

For the induced dataset, SCFFA models developed show R^2cv ranged from 0.50 to 0.72. As the availability of the data is almost perfect (Table 2), the performance of the models has improved, reflecting the high variability present in the dataset. In return, the performance of the LCFFA has decreased for R^2cv from 0.53 to 0.38 for C16 and from 0.69 to 0.6 for C18. This decline can be attributed to the limited impact of homogenization on the enhancement of these specific FFA. The external validation demonstrates superior performance compared with the model developed solely with spontaneous lipolysis. This outcome was anticipated, as the dataset includes greater variability in the calibration. However, a significant drop of performance is noted for all FFA. The most pronounced reductions are evident for C6 and C18, with performance metrics decreasing from 0.62 to 0.2 for C6 and from 0.56 to 0.18 for C18 when comparing cross-validation to validation results.

Table 3. Performance of the different PLS models predicting FFA content in milk for spontaneous lipolysis including 313 samples, induced lipolysis including 479 samples, and for the complete dataset including 634 samples in the calibration set and 160 samples in the external validation set[1]

Dataset	Fatty acid	N PLS	Outlier	R^2cv	R^2val	RMSEcv	RMSEval	RPDcv	RPDval
Spontaneous lipolysis	C4	7	26	0.02	−0.28	1.78	40.03	1.01	0.88
	C6	4	21	0.05	−0.51	0.63	24.59	1.03	0.81
	C8	6	26	0.18	−0.45	0.62	17.92	1.11	0.83
	C10	5	20	0.23	−0.37	1.30	19.31	1.14	0.86
	C12	1	22	0.00	−0.20	0.42	12.25	1.00	0.91
	C14	6	18	0.43	−0.24	2.47	15.69	1.33	0.90
	C16	11	9	0.53	0.00	9.54	33.45	1.47	1.00
	C18	11	6	0.69	−0.34	10.09	17.28	1.81	0.86
	C18:1	1	21	−0.01	−0.12	5.77	14.93	1.00	0.95
	C18:2	7	17	0.09	−0.14	0.44	2.41	1.05	0.93
	SCFFA	6	20	0.07	−0.48	3.41	81.22	1.04	0.82
	MCFFA	6	21	0.15	−0.29	4.23	45.91	1.08	0.88
	LCFFA	12	14	0.44	−0.05	23.04	59.09	1.33	0.98
Induced lipolysis	C4	17	8	0.72	0.71	20.90	19.00	1.89	1.87
	C6	17	16	0.62	0.20	13.21	17.96	1.62	1.12
	C8	14	15	0.60	0.39	7.98	11.59	1.57	1.29
	C10	15	15	0.55	0.39	9.29	12.95	1.49	1.28
	C12	8	28	0.26	0.26	5.95	9.59	1.16	1.17
	C14	6	23	0.31	0.30	10.03	11.76	1.20	1.20
	C16	7	28	0.38	0.24	21.34	29.23	1.27	1.15
	C18	20	15	0.56	0.18	10.63	13.54	1.51	1.11
	C18:1	6	29	0.08	0.14	10.85	13.12	1.04	1.08
	C18:2	8	33	0.10	0.05	1.15	2.20	1.05	1.03
	SCFFA	17	9	0.70	0.41	41.02	51.93	1.84	1.30
	MCFFA	9	24	0.37	0.32	23.70	33.29	1.25	1.21
	LCFFA	7	27	0.32	0.13	43.53	53.96	1.21	1.07
Full	C4	19	10	0.76	0.84	16.80	14.04	2.03	2.53
	C6	16	16	0.68	0.60	10.95	12.71	1.77	1.57
	C8	14	23	0.65	0.53	6.37	10.20	1.69	1.46
	C10	9	29	0.51	0.36	6.88	13.24	1.43	1.25
	C12	9	43	0.33	0.24	4.18	9.78	1.22	1.14
	C14	9	43	0.32	0.28	6.70	11.97	1.21	1.18
	C16	9	39	0.38	0.27	17.45	28.51	1.27	1.17
	C18	20	27	0.52	0.33	11.07	12.19	1.45	1.23
	C18:1	4	57	0.08	0.07	7.91	13.64	1.04	1.04
	C18:2	7	58	0.05	−0.03	0.68	2.29	1.03	0.98
	SCFFA	19	11	0.73	0.77	33.31	31.97	1.94	2.10
	MCFFA	14	36	0.45	0.44	17.08	17.58	1.35	1.33
	LCFFA	18	50	0.32	0.30	35.91	42.56	1.21	1.20

[1]N PLS = number of components in the partial least squares regression; Outlier = number of samples excluded from model building; R^2val = R^2 of validation; RMSEval = root mean square error of validation; RPDcv = ratio per deviation of cross-validation; RPDval = ratio per deviation of validation.

For the complete dataset combining samples with induced and spontaneous lipolysis, the ratio per deviation (Williams, 2014) of validation ranged from 0.98 to 2.53, whereas validation R^2 ranged between 0.05 and 0.84. All models developed with the complete dataset outperformed the models created using partial datasets. Considering the performances, some of these models could be suitable for rough screening, especially considering the SCFFA and the C4 both showing a validation R^2 above 0.7. Both C6 and C8 show relatively good cross-validation performances, with R^2 around 0.7. However, in validation, the R^2 drops to 0.57 and 0.56, respectively, which might indicate that the model lacks robustness and that additional variability and samples should be integrated in the calibration dataset. Other models show relatively poor performances and are not likely to be used to better monitor lipolysis. The performance metrics of the C16 and C18 models exhibited a decline upon the introduction of samples characterized by induced lipolysis, with ratio per deviation of cross-validation values decreasing from 1.41 and 1.86 to 1.28 and 1.5, respectively. This reduction in performance may lend credence to the hypothesis that LCFFA are not significantly affected by induced lipolysis. Notably, these fatty acids predominantly reside in the sn-2 position of triglycerides, as demonstrated by the findings of Somerharju et al. (1978). Furthermore, Figure 1 supports this hypothesis, indicating that the distribution remains unchanged despite the inclusion of samples derived from mechanical induction.

However, this randomly segmented validation approach is suboptimal because it fails to accurately reflect real-world conditions. Specifically, samples devoid of

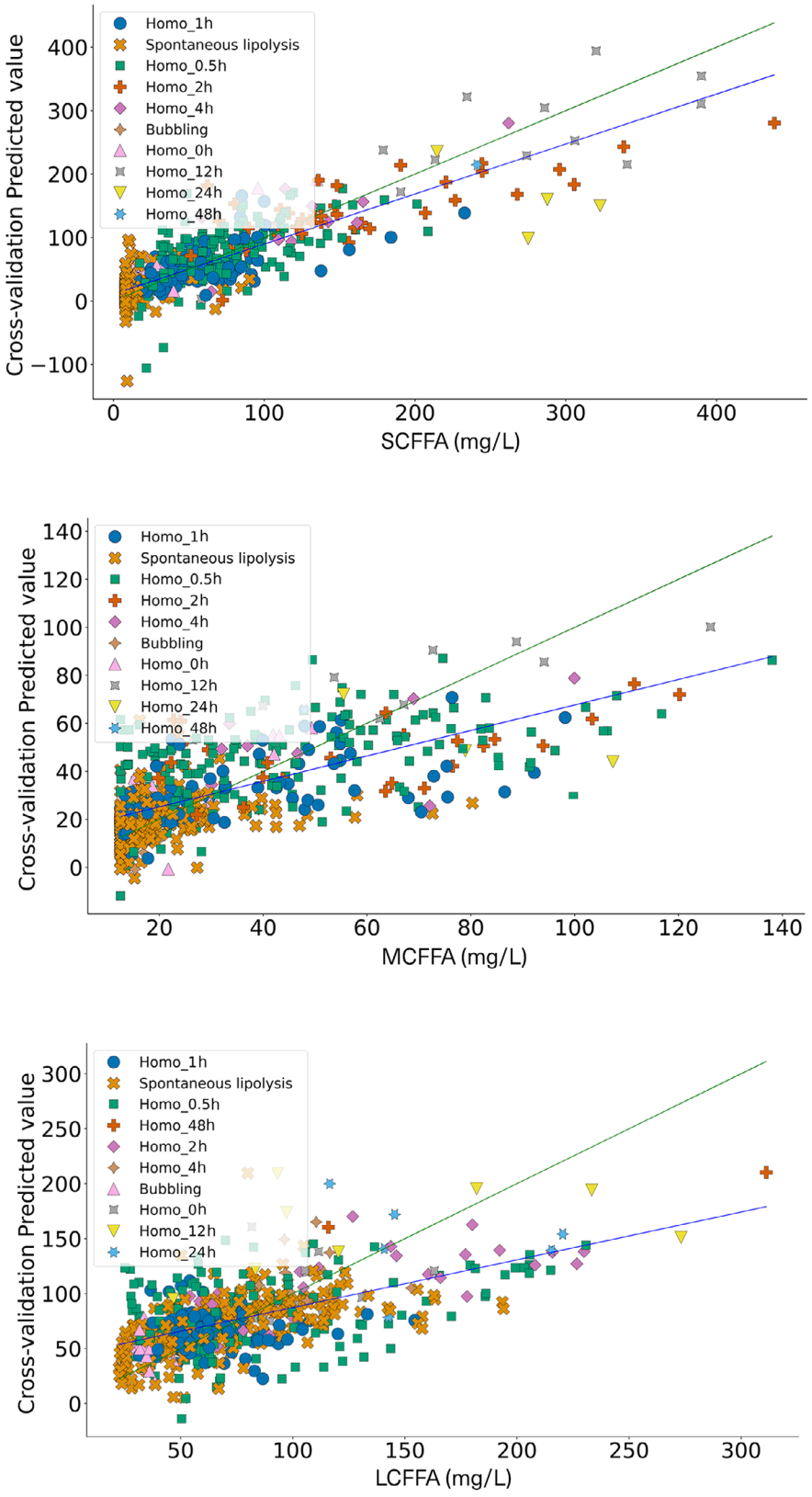

Figure 5. Data distribution of the partial least squares (PLS) models for the sum of the free fatty acids (FFA): short-chain FFA (SCFFA), mid-chain FFA (MCFFA), and long-chain FFA (LCFFA). Different symbols indicate the time of storage after the homogenization and before the analysis in hours (e.g., Homo_1h = 1 h between homogenization and analysis).

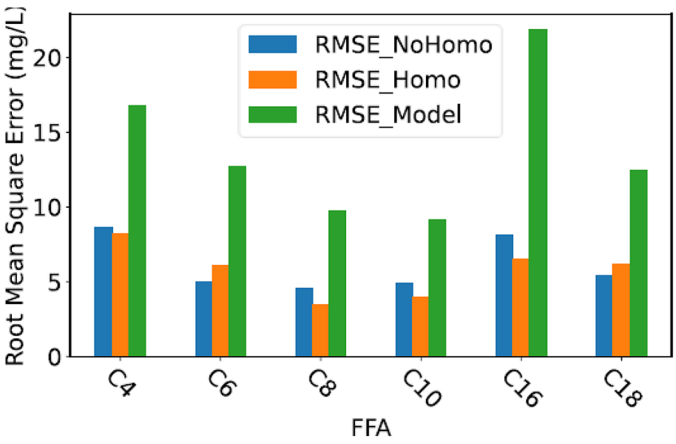

Figure 6. Root mean square model of the pasteurization test. Differences of pasteurized samples with and without homogenization. RMSE_NoHomo = RMSE with no homogenization; RMSE_Homo = RMSE = with homogenization; RMSE_Model = intrinsic RMSE of the models.

mechanical induction should be employed to effectively assess the performance of our models. Unfortunately, the scarcity of high-value samples without mechanical induction hinders the implementation of this approach.

An alternative methodology to modeling using a continuous quantitative target would be the development of classification models, discriminating low versus high FFA contents. Discriminant models are expected to outperform models present in Table 3, but it would be necessary to establish a threshold to determine whether a milk sample exhibits a milk taste defect or not. Thresholds have been established in previous studies for global FFA (Pillay et al., 1980). However, to align with the current research, it is essential to define new thresholds for each FFA using an updated quantification method. To the best of our knowledge, no published articles have addressed the prediction of lipolysis in milk, and only models from instrument manufacturers are currently accessible. These models rely on various reference methods, but above all they predict the global FFA and not the individual ones, making the comparison with our results precarious.

Pasteurization Test

To validate the appropriate functioning of our model, a pasteurization test was conducted. This process effectively inhibits all lipases typically found in milk, allowing us to assess whether the model accurately predicts the concentration of FFA or the homogenization or duration following homogenization. The results are illustrated in Figure 6. Comparable RMSE are observed for the pasteurized test set, with and without homogenization. The outcomes of this test affirm the model's correct behavior. Specifically, if the homogenization or the time postho-mogenization were predicted, a higher RMSE would be expected for the test set corresponding to a homogenization and a duration of 4 h after homogenization. This test validates that the model predicts the FFA content well, but not the structural and chemical changes induced by homogenization and resting time. Both RMSE are under the RMSEval of the building model, confirming the right behavior of our predictive model.

CONCLUSIONS

The development of predictive equations for FFA content is a significant advancement in the field of milk quality control. This study successfully demonstrates the potential of using MIR analysis of milk with machine learning techniques to predict FFA levels with heterogeneous accuracy. Although SCFFA could be used for rough screening, especially C4 and total SCFFA with validation R^2 above 0.7, the MCFFA and LCFFA models showed poor predictive performance. These results are therefore positive as the negative tastes mainly originate from SCFFA. The diverse equations of prediction were produced through 2 kinds of methodologies with spontaneous lipolysis samples and induced lipolysis samples. Various forms of induced lipolysis were examined, including vortexing, freezing bubbling, and storage duration following the homogenization process. Among these methods, the storage after homogenization emerged as the most effective approach to increase the range of FFA contents and therefore model performances. Additionally, a pasteurization test was conducted to assess the accuracy of our equation and to validate the reliability of our models. Finally, the proposed equations could offer a reliable alternative to traditional, time-consuming laboratory methods, providing quick insights into the organoleptic quality of samples.

NOTES

This research was conducted as part of the Improving Taste and Flavor of Milk Products (ITF-milk; Eurpean Innovation Partnership for Agricultural Productivity and Sustainability) initiative. The authors express their gratitude to the Ministry of Agriculture, Viticulture, and Rural Development of the Grand Duchy of Luxembourg (Luxembourg City, Luxembourg) for funding this project. The authors gratefully thank Tom Engel from Administration des services techniques de l'agriculture (ASTA; Luxembourg City, Luxembourg) and Matthieu Dubuisson, Mohamed El Morabit, Petimat Kitaeva, and Quentin Vigneron from the Walloon Agricultural Research Center (CRA-W, Gembloux, Belgium) for their invaluable

work. Supplemental material for this article is available at https://doi.org/10.6084/m9.figshare.30138346. All samples were collected following the guidelines edited by the International Committee for Animal Recording (ICAR, 2022) and with ICAR-approved milk samplers. The authors have not stated any conflicts of interest.

Nonstandard abbreviations used: AMS = automatic milking system; FFA = free fatty acids; FT-MIR = Fourier transform mid-infrared; GC-MS/MS = GC coupled with tandem mass spectrometry; GH = Mahalanobis distance; LCFFA = long-chain free fatty acids; LOQ = limit of quantification; LPL = lipoprotein lipase; MCFFA = mid-chain free fatty acids; MFGM = milk fat globule membrane; MIR = mid-infrared; PLS = partial least squares; R^2cv = R^2 of cross-validation; R^2val = R^2 of validation; RMSE = root mean square error; RMSEcv = RMSE of cross-validation; RMSE_Homo = RMSE = with homogenization; RMSE_Model = intrinsic RMSE of the models; RMSE_NoHomo = RMSE with no homogenization; RMSEval = root mean square error of validation; RPDcv = ratio per deviation of cross-validation; RPDval = ratio per deviation of validation; SCFFA = short-chain free fatty acids; SPE = solid phase extraction.

REFERENCES

Abeni, F., L. Degano, F. Calza, R. Giangiacomo, and G. Pirlo. 2005. Milk quality and automatic milking: Fat globule size, natural creaming, and lipolysis. J. Dairy Sci. 88:3519–3529. https://doi.org/10.3168/jds.S0022-0302(05)73037-X.

Astrup, H., L. Vik-Mo, O. Skrøvseth, and A. Ekern. 1980. Milk lipolysis when feeding saturated fatty acids to the cow. Milchwissenschaft 35:1–4.

Christophe, O. S., C. Grelet, C. Bertozzi, D. Veselko, C. Lecomte, P. Höeckels, A. Werner, F.-J. Auer, N. Gengler, F. Dehareng, and H. Soyeurt. 2021. Multiple breeds and countries' predictions of mineral contents in milk from milk mid-infrared spectrometry. Foods 10:2235. https://doi.org/10.3390/foods10092235.

Collins, Y. F., P. L. H. McSweeney, and M. G. Wilkinson. 2003. Lipolysis and free fatty acid catabolism in cheese: A review of current knowledge. Int. Dairy J. 13:841–866. https://doi.org/10.1016/S0958-6946(03)00109-2.

Connolly, J. F., J. J. Murphy, C. B. O'Connor, and D. R. Headon. 1979. Relationship between lipolysed flavour and free fatty acid levels in milk and butter. Int. J. Food Sci. Technol. 3:79–92.

Coppa, M., A. Revello-Chion, D. Giaccone, A. Ferlay, E. Tabacco, and G. Borreani. 2014. Comparison of near and medium infrared spectroscopy to predict fatty acid composition on fresh and thawed milk. Food Chem. 150:49–57. https://doi.org/10.1016/j.foodchem.2013.10.087.

Deeth, H. C. 2006. Lipoprotein lipase and lipolysis in milk. Int. Dairy J. 16:555–562. https://doi.org/10.1016/j.idairyj.2005.08.011.

Deeth, H. C. 2019. Milk lipids: Lipolysis and hydrolytic rancidity. Reference Module in Food Science. Elsevier. https://doi.org/10.1016/B978-0-08-100596-5.00925-2.

Deeth, H. C. 2022. Lipolysis and hydrolytic rancidity. Pages 827–834 in Encyclopedia of Dairy Sciences. 3rd ed. P. L. H. McSweeney and J. P. McNamara, ed. Academic Press, Oxford.

Dehareng, F., C. Delfosse, E. Froidmont, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde, and P. Dardenne. 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. Animal 6:1694–1701. https://doi.org/10.1017/S1751731112000456.

Delosière, M., L. Bernard, C. Hurtaud, M. Guilleton, D. Viala, A. Rau, M. Bonnet, and C. Cebo. 2023. Protein signatures of spontaneous lipolysis and lipoprotein lipase activity in cow's milk. J. Proteomics 285:104951. https://doi.org/10.1016/j.jprot.2023.104951.

Di Marzo, L., and D. M. Barbano. 2016. Effect of homogenizer performance on accuracy and repeatability of mid-infrared predicted values for major milk components. J. Dairy Sci. 99:9471–9482. https://doi.org/10.3168/jds.2016-11618.

Downey, W. K. 1980. Flavour impairment from pre- and post-manufacture lipolysis in milk and dairy products. J. Dairy Res. 47:237–252. https://doi.org/10.1017/S0022029900021117.

Duncan, S. E., G. L. Christen, and M. P. Penfield. 1991. Rancid flavor of milk: Relationship of acid degree value, free fatty acids, and sensory perception. J. Food Sci. 56:394–397. https://doi.org/10.1111/j.1365-2621.1991.tb05288.x.

Evers, J. M. 2003. Determination of free fatty acids in milk using the BDI method—Some practical and theoretical aspects. Int. Dairy J. 13:111–121.

Fornell, C., and J. Cha. 1994. Partial least squares. Adv. Methods Mark. Res. 407:52–78.

Gholson, J. H., R. H. Schexnailder, and L. L. Rusoff. 1966. Influence of a poor-quality low-energy ration on lipolytic activity in milk. J. Dairy Sci. 49:1136–1139. https://doi.org/10.3168/jds.S0022-0302(66)88033-5.

González-Córdova, A. F., and B. Vallejo-Cordoba. 2003. Detection and prediction of hydrolytic rancidity in milk by multiple regression analysis of short-chain free fatty acids determined by solid phase microextraction gas chromatography and quantitative flavor intensity assessment. J. Agric. Food Chem. 51:7127–7131. https://doi.org/10.1021/jf030347w.

Goulden, J. D. S. 1961. Quantitative analysis of milk and other emulsions by infra-red absorption. Nature 191:905–906. https://doi.org/10.1038/191905a0.

Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. A. Fernandez Pierna, F. G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825. https://doi.org/10.3168/jds.2015-10477.

Grelet, C., P. Dardenne, H. Soyeurt, J. A. Fernandez, A. Vanlierde, F. Stevens, N. Gengler, and F. 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., E. Froidmont, L. Foldager, M. Salavati, M. Hostens, C. P. Ferris, K. L. Ingvartsen, M. A. Crowe, M. T. Sorensen, J. A. Fernandez Pierna, A. Vanlierde, N. Gengler, and F. Dehareng. 2020. Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. J. Dairy Sci. 103:4435–4445. https://doi.org/10.3168/jds.2019-17910.

ICAR (International Committee for Animal Recording). 2022. Section 2-Guidelines for Dairy Cattle Milk Recording. ICAR.

Ivanov, G. Y., E. Bilgücü, T. G. Balabanova, and I. V. Ivanova. 2021. Effect of somatic cell count of cow's milk on the lipolysis and fatty acid profile of farmer cheese. Int. Food Res. J. 28:1171–1178. https://doi.org/10.47836/ifrj.28.6.08.

Jellema, A., and C. Schipper. 1975. Influence of physiological factors on the lipolytic susceptibility of milk. Annual Bulletin, International Dairy Federation. 86:2–6.

Johnson, P. E., and R. Von Gunten. 1962. A study of factors involved in the development of rancid flavor in milk. Oklahoma Agricultural Experiment Station, Bulletin no. 593, February 1962.

Kiełczewska, K., K. Ambroziak, D. Krzykowska, and M. Aljewicz. 2021. The effect of high-pressure homogenisation on the size of milk fat

globules and MFGM composition in sweet buttermilk and milk. Int. Dairy J. 113:104898. https://doi.org/10.1016/j.idairyj.2020.104898.
Kintner, J. A., and E. A. Day. 1965. Major free fatty acids in milk. J. Dairy Sci. 48:1575–1581. https://doi.org/10.3168/jds.S0022-0302(65)88529-0.
Kodgev, A., and R. Rachev. 1970. The influence of some factors on the acidity of milk fat. No. 1E, 200 in XVIII International Dairy Congress, Sydney.
Luhtala, A., and M. Antila. 1968. Über die Lipasen und die Lipolyse der Milch. Fette Seifen Anstrichm. 70:280–288. https://doi.org/10.1002/lipi.19680700413.
Macedo Mota, L. F., V. Bisutti, A. Vanzin, S. Pegolo, A. Toscano, S. Schiavon, F. Tagliapietra, L. Gallo, P. Ajmone Marsan, and A. Cecchinato. 2023. Predicting milk protein fractions using infrared spectroscopy and a gradient boosting machine for breeding purposes in Holstein cattle. J. Dairy Sci. 106:1853–1873. https://doi.org/10.3168/jds.2022-22119.
Mcdaniel, M. R., L. A. Sather, and R. C. Lindsay. 1969. Influence of free fatty acids on sweet cream butter flavor. J. Food Sci. 34:251–254. https://doi.org/10.1111/j.1365-2621.1969.tb10334.x.
Morley, N., and A. Kuksis. 1977. Lack of fatty acid specificity in the lipolysis of oligo- and polyunsaturated triacylglycerols by milk lipoprotein lipase. Biochim. Biophys. Acta BBA Lipids Lipid Metab. 487:332–342. https://doi.org/10.1016/0005-2760(77)90009-1.
Murphy, S. C., K. Cranker, G. F. Senyk, D. M. Barbano, A. I. Saeman, and D. M. Galton. 1989. Influence of bovine mastitis on lipolysis and proteolysis in milk. J. Dairy Sci. 72:620–626. https://doi.org/10.3168/jds.S0022-0302(89)79152-9.
Ortiz, M. J., E. M. Kesler, G. H. Watrous Jr., and W. H. Cloninger. 1970. Effect of the cow's body condition and stage of lactation on development of milk rancidity. J. Milk Food Technol. 33:339–342. https://doi.org/10.4315/0022-2747-33.8.339.
Ouattara, G., I. Jeon, R. Hart-Thakur, and K. Schmidt. 2004. Fatty acids released from milk fat by lipoprotein lipase and lipolytic psychrotrophs. J. Food Sci. 69:C659–C664. https://doi.org/10.1111/j.1750-3841.2004.tb18014.x.
Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12:2825–2830.
Python Software Foundation. 2024. Python Language Reference, version 3.2. Accessed Oct. 2, 2025. https://www.python.org.
Pillay, V. T., A. N. Myhr, and J. I. Gray. 1980. Lipolysis in milk. I. Determination of free fatty acid and threshold value for lipolyzed flavor detection. J. Dairy Sci. 63:1213–1218. https://doi.org/10.3168/jds.S0022-0302(80)83070-0.
Ray, P. R., K. Chatterjee, C. Chakraborty, and P. K. Ghatak. 2013. Lipolysis of milk: A review. Int. J. Agric. Sci. Vet. Med. 1:58–74.
Santos, M. V., Y. Ma, and D. M. Barbano. 2003. Effect of somatic cell count on proteolysis and lipolysis in pasteurized fluid milk during shelf-life storage. J. Dairy Sci. 86:2491–2503. https://doi.org/10.3168/jds.S0022-0302(03)73843-0.
Scanlan, R. A., L. A. Sather, and E. A. Day. 1965. Contribution of free fatty acids to the flavor of rancid milk. J. Dairy Sci. 48:1582–1584. https://doi.org/10.3168/jds.S0022-0302(65)88530-7.
Shimizu, M., H. Miyaji, and K. Yamauchi. 1982. Inhibition of lipolysis by milk fat globule membrane materials in model milk fat emulsion. Agric. Biol. Chem. 46:795–799. https://doi.org/10.1080/00021369.1982.10865129.
Somerharju, P., T. Kuusi, F. Paltauf, and P. K. J. Kinnunen. 1978. Stereospecificity of lipoprotein lipase is an intrinsic property of the active site of the enzyme protein. FEBS Lett. 96:170–172. https://doi.org/10.1016/0014-5793(78)81086-2.
Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J. Dairy Sci. 89:3690–3695. https://doi.org/10.3168/jds.S0022-0302(06)72409-2.
Tedde, A., P. N. Grelet, J. E. Ho, D. Pryce, Z. Hailemariam, G. Wang, N. Plastow, E. Gengler, F. Froidmont, C. Dehareng, M. A. Bertozzi, H. Crowe, and H. Soyeurt. 2021. Multiple country approach to improve the test-day prediction of dairy cows' dry matter intake. Animals (Basel) 11:1316. https://doi.org/10.3390/ani11051316.
Vanbergue, E., L. Delaby, J. L. Peyraud, S. Colette, Y. Gallard, and C. Hurtaud. 2017. Effects of breed, feeding system, and lactation stage on milk fat characteristics and spontaneous lipolysis in dairy cows. J. Dairy Sci. 100:4623–4636. https://doi.org/10.3168/jds.2016-12094.
Vanlierde, A., F. Dehareng, N. Gengler, E. Froidmont, S. McParland, M. Kreuzer, M. Bell, P. Lund, C. Martin, B. Kuhla, and H. Soyeurt. 2021. Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra. J. Sci. Food Agric. 101:3394–3403. https://doi.org/10.1002/jsfa.10969.
Wells, M. E., O. P. Pryor, D. M. Haggerty, H. C. Pickett, and J. B. Mickle. 1969. Effect of estrous cycle and lactation on lipase activity in bovine milk and blood. J. Dairy Sci. 52:1110–1113. https://doi.org/10.3168/jds.S0022-0302(69)86705-6.
Wiking, L. 2005. Milk fat globule stability: Lipolysis with special reference to automatic milking systems. Ph.D. diss. Swedish University of Agricultural Sciences.
Williams, P. 2014. The RPD statistic: A tutorial note. NIR News 25:22–26. https://doi.org/10.1255/nirn.1419.

ORCIDS

Octave Christophe, https://orcid.org/0000-0002-9037-8079
Romain Reding, https://orcid.org/0009-0008-5946-9459
Julie Leblois, https://orcid.org/0000-0001-5112-0262
Clément Grelet, https://orcid.org/0000-0003-3313-485X
Hélène Soyeurt, https://orcid.org/0000-0001-9883-9047
Frédéric Dehareng https://orcid.org/0000-0002-6733-4334