Short communication: Potential prediction of vitamin B₁₂ concentration based on mid-infrared spectral data using Holstein Dairy Herd Improvement milk samples

M. Duplessis,^{1*} D. Pellerin,² C. L. Girard,¹ D. E. Santschi,³ and H. Soyeurt⁴

¹Centre de Recherche et Développement de Sherbrooke, Sherbrooke, QC, J1M 0C8, Canada

²Département des Sciences Animales, Université Laval, Québec, G1V 0A6, Canada

³Lactanet, Sainte-Anne-de-Bellevue, QC, H9X 3R4, Canada

⁴Agriculture, Bio-Engineering, and Chemistry Department, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium

ABSTRACT

The purpose of this study was (1) to predict the quantitative concentration of vitamin B₁₂ in milk using mid-infrared (MIR) spectrometry, and (2) to evaluate the potential of MIR spectra to discriminate different clusters of records based on their B₁₂ concentration. Milk samples were collected from 4,340 Holstein cows between 3 and 592 d in milk and located in 100 herds. Samples were taken using in-line milk meters and divided into 2 aliquots: one for MIR spectrometry and the other for B₁₂ concentration reference analyses by radioassay. Analyses were performed on 311 selected spectral wavelengths. A partial least squares regression model was built to quantify B₁₂ concentration. Discriminant analysis was used to isolate B₁₂ concentration clusters. A B₁₂ concentration threshold was set at 442 ng/dL, because this represents the cutoff value for a 250-mL glass of milk to fulfill 46% of the daily vitamin B₁₂ recommended dietary allowance for individuals 14 yr or older. For each analysis, records coming from two-thirds of herds were used to calibrate prediction equations, and the remaining records (one-third of herds for validation) were used to assess the prediction performance. In the case of discriminant analysis, validation sets were divided into evaluation sets (one-third of herds) to obtain alternate probability cutoffs and in test sets (two-thirds of herds) to validate equations. Spectral and B₁₂ concentration outliers were identified by calculating standardized Mahalanobis distance and with a residual analysis, respectively ($n = 3,154$). Regarding quantitative B₁₂ concentration, cross-validation and validation coefficients of determination averaged 0.51 and 0.46, respectively, which are relatively low,

which would limit the potential use of the developed quantitative equations. In addition, root mean square errors of prediction of cross validation and validation sets averaged 88.9 and 94.7 ng/dL, respectively. Area under the receiver operating characteristic curve of test sets averaged 0.81 based on the 442 ng/dL threshold, which could be considered to represent good accuracy of classification. However, the false discovery rate averaged 36%. In summary, models predicting quantitative B₁₂ concentration had low cross-validation and validation coefficients of determination, limiting their use, but the proposed discriminant models could be used to identify milk samples with naturally high B₁₂.

Key words: cyanocobalamin, mid-infrared, milk, cow, discriminant analysis

Short Communication

Humans rely on exogenous sources of vitamin B₁₂ to fulfill their daily recommended dietary allowance (RDA). Because this vitamin is solely synthesized by bacteria if the cobalt supply is sufficient (Martens et al., 2002), animal products, especially those from ruminants such as dairy products, are an excellent source of this vitamin. Bacteria dwelling in the rumen synthesize vitamin B₁₂ that can either be used by the animal or naturally secreted in milk. According to the National Academies of Sciences (1998), humans over 13 yr old have an RDA of vitamin B₁₂ of 2.4 µg/d; this recommendation increases by 0.2 and 0.4 µg/d for pregnant and lactating women, respectively. Based on standard pasteurized milk composition (USDA-ARS, 2012), a 250-mL glass of milk should provide 46% of the daily vitamin B₁₂ RDA for an individual aged 14 yr or older. In human nutrition, adequate supply of vitamin B₁₂ prevents neuropathy, megaloblastic anemia, and gastrointestinal symptoms (Baik and Russell, 1999).

Previous studies have shown that a combined supplement of folic acid and vitamin B₁₂ given in early lacta-

Received October 16, 2019.

Accepted March 4, 2020.

*Corresponding author: melissa.duplessis@canada.ca

tion, despite having no effect on DMI, has a positive effect on energy status (Preynat et al., 2009; Duplessis et al., 2017). However, it is not possible under the current state of knowledge to foresee which cows would benefit from this supplement in early lactation, but the plasma vitamin concentration of the animal could be an important factor determining it (Duplessis et al., 2017). Although Duplessis et al. (2018) concluded that plasma and milk vitamin B₁₂ concentrations were not strongly correlated in Holstein cows, assessing vitamin B₁₂ concentrations in milk could permit an estimation of the extent of its secretion in milk and, therefore, how much is not available for metabolism in extra-mammary tissues.

The reference method used to quantify vitamin B₁₂ concentration in milk requires preparation of milk with several reagents followed by incubation with an enzyme before analysis by radioassay (Duplessis et al., 2015). This analysis is not suitable for routine determination of vitamin B₁₂ concentration in milk because it is time consuming and costly and uses radioactivity. It has been showed that mid-infrared (MIR) spectrometry could be used to predict, among other things, the milk concentrations of citrate, BHB (Grelet et al., 2016), protein fraction (McDermott et al., 2016), Ca and P (Soyeurt et al., 2009). The advantages of MIR analysis are its speed as well as its low cost, and it is widely used worldwide by DHI agencies for analyses of milk fat, protein, and lactose concentrations. Given the aforementioned roles of vitamin B₁₂ in human health and cow metabolism and the advantages of MIR, it could be interesting to predict this vitamin by using MIR. The objectives of this experiment were (1) to predict the concentration of vitamin B₁₂ in milk of individual cows by using MIR spectrometry, and (2) to determine whether a discriminant analysis could be used to distinguish different clusters of vitamin B₁₂ concentration in milk of individual cows using MIR spectrometry.

All procedures of this experiment were approved by the animal care committee from Université Laval, Québec, Canada, following the guidelines of the Canadian Council on Animal Care (2009). Participating herds and cows were previously described by Duplessis et al. (2019). Briefly, 100 dairy herds located in the province of Québec, Canada, voluntarily participated in this experiment. Cows were milked twice daily and housed in tiestall ($n = 98$) and freestall ($n = 2$) barns. Herd size ranged from 16 to 113 Holstein lactating cows. To participate, herds were required to record milk production using the regular a.m.-p.m. milk testing scheme through the DHI agency, have Holstein as the main breed, and milk cows twice daily. Dairy herds that fulfilled these requirements were contacted and par-

ticipation was voluntary. From October 2014 to June 2015, herds were visited once for 2 consecutive milkings during routine milk recording. Milk yield was recorded at each milking. Separate milk samples using in-line milk meters from evening and morning milkings were taken on 4,340 Holstein cows (1,484 primiparous and 2,856 multiparous cows) between 3 and 592 DIM. Milk samples were divided in 2 aliquots: a 60-mL aliquot was preserved with bronopol and sent immediately to the Lactanet laboratory (Dairy Production Center of Expertise, Québec and Atlantic Provinces, Sainte-Anne-de-Bellevue, QC, Canada) and a 15-mL aliquot, without bronopol, was stored at -20°C until analysis.

The concentration of vitamin B₁₂ of evening and morning milk samples without bronopol was separately analyzed in duplicate by radioassay (Duplessis et al., 2015) using a commercial kit (SimulTrac B₁₂/Folate-S, MP Biomedicals, Solon, OH). The latter was considered the reference method to analyze vitamin B₁₂ concentration in milk in the current study because Brito et al. (2015) reported a correlation coefficient of 0.97 between the radioassay and liquid chromatography-MS methods. The interassay coefficient of variation was 3.1%. Milk samples with bronopol were analyzed by Fourier transform MIR reflectance spectrometry for milk fat, protein, and lactose concentrations using MilkoScanTM FT 6000 spectrometers (Foss, Hillerød, Denmark) in the laboratory of Lactanet. Out of the 4,340 sampled cows, 4,082 spectra from evening samples and 4,112 spectra from morning samples were obtained and contained 1,060 data points located from 5,000 to 925 cm^{-1} . Spectral data were obtained in transmittance and transformed to absorbance using the following formula: Absorbance = $[\log_{10}(1/\text{transmittance})]$. To obtain representative 24-h spectra, evening and morning spectra and vitamin B₁₂ data were weighted according to milk yield recorded at each of the milkings. Spectral data from both evening and morning milkings were missing for 6 herds, and spectral data from either the evening or morning milking were missing for 7 other herds. Hence, the data set used for further analysis contained 3,712 cow records. A first-derivative transformation using a gap of 5 consecutive wavenumbers was applied to the recorded spectral data (Soyeurt et al., 2011). As recommended by Soyeurt et al. (2011) and based on coefficients of variation analysis of each spectral point, 311 spectral points were used and located between 1,600 and 926, 1,809 and 1,712, and 2,989 and 2,561 cm^{-1} . Then, using the FactoMineR package (Lê et al., 2008) of R software (version 1.1.419; <https://www.R-project.org/>), principal component analyses were performed on treated spectral data, and a standardized Mahalanobis (GH) distance was calculated based on 15

principal components explaining 99.1% of the spectral variability. A GH threshold of 3 was used to conclude that a spectrum was an outlier (Whitfield et al., 1987). Hence, 61 samples were suppressed. This data set ($n = 3,651$) is hereafter referred to as the raw data set.

The detection of reference outliers of vitamin B₁₂ concentration in milk was performed a posteriori based on a residual analysis. A first partial least squares (PLS) regression predicting the concentration of vitamin B₁₂ in milk using the selected 311 spectral data was built using the PLS package (Mevik and Wehrens, 2007) of R software on the raw data set. The number of PLS components was estimated from the root mean square error of prediction of cross-validation set (**RMSEP_{cv}**) obtained after a 10-fold cross-validation. If a record had an absolute residual (i.e., observed minus predicted vitamin B₁₂ concentration in milk) >2 times the square root of the residual variance, it was deleted (Rousseeuw et al., 2006). This operation was iterated until the enhancement of the cross-validation R^2 (**R^2_{cv}**) was less than 1%. This data set is hereafter called the cleaned data set ($n = 3,154$).

Calibration and validation sets were created from the raw and cleaned data sets. The splitting was based on herds to ensure independency between calibration and validation sets. All herds were randomly divided into 3 groups using Proc SURVEYSELECT of SAS software (version 9.4; 2012; SAS Institute Inc., Cary, NC). Then, 3 different calibration (records from two-thirds of herds) and validation (records from one-third of herds) sets were created to validate records from each herd once. Proc UNIVARIATE of SAS was used to determine descriptive statistics of vitamin B₁₂ concentration in milk of each of the calibration and validation sets.

Partial least squares regressions were built from each calibration set to predict the concentration of vitamin B₁₂ in milk using the PLS package of R software. A 10-fold cross-validation was performed on calibration data sets, but a maximum of 16 components was set to avoid overfitting. To assess the robustness and the accuracy of the predictions obtained by the developed regressions, different statistical parameters were estimated: R^2 of calibration (**R^2_c**) and R^2_{cv} , root mean square error of prediction of calibration set (**RMSEP_c**) and RMSEP_{cv}. As stated by Soyeurt et al. (2009), new MIR equations are useful only if the studied component cannot be estimated by existing equations used for well-known milk components. Pearson correlation coefficients between major daily milk components (i.e., fat, protein, and lactose) predicted by MIR using manufacturer equations and daily milk concentration of vitamin B₁₂ from reference analysis were calculated with R software. If cross-validation correlation coefficients

are higher than the Pearson correlation coefficients between milk components and reference daily milk concentration of vitamin B₁₂, the developed equations give additional information and the component cannot be estimated based on existing equations. For each developed prediction equation, a validation was done on the remaining records from the one-third of herds not used for calibration. Predicted values of the validation data set were computed from intercept and PLS regression coefficients previously determined by calibration sets. Coefficient of determination of validation set (**R^2_v**) and root mean square error of prediction of validation set (**RMSEP_v**) were calculated. Moreover, the ratio of SD from reference vitamin B₁₂ concentration in milk values to RMSEP_v (RPD; Williams and Sobering, 1993) was calculated to evaluate the potential application of the developed prediction equations. A ratio >3 indicates that the prediction equations can be considered reliable and used in practice for screening purposes (Soyeurt et al., 2011).

The threshold used to discriminate classes of vitamin B₁₂ concentration in milk was based on the fact that a 250-mL glass of milk should provide 46% of daily vitamin B₁₂ RDA for an individual over 13 yr of age. Based on that assumption, a concentration of vitamin B₁₂ of 441.6 ng/dL in milk should be reached to meet this requirement; hence, this value was used as the threshold to categorize samples based on their ability to meet the RDA. The Caret package (Kuhn, 2008) of R software was used to develop models discriminating vitamin B₁₂ concentration in milk classes based on PLS regressions from the cleaned data set. A 10-fold cross-validation was also used to determine the number of components based on the receiver operating characteristic (**ROC**) curve, but a maximum number of 16 was set to avoid overfitting. Calibration sets from the cleaned data set were the same as those used to build the quantitative PLS prediction models. To overcome class imbalance, alternate probability cutoffs needed to be determined (Kuhn and Johnson, 2013). Each validation set of the cleaned data set was further divided into evaluation (one-third of herds) and test (two-thirds of herds) sets. The evaluation set was used to obtain alternate probability cutoffs, as described by Kuhn and Johnson (2013). Alternate probability cutoffs were used to predict classifications of calibration and test sets based on prediction equations. A confusion matrix was obtained and the statistics were expressed in terms of ROC area under the curve (**AUC**), sensitivity, specificity, and accuracy.

Milk yield on a 305-d basis averaged $9,814 \pm 1,948$ kg/cow. Milk fat, protein, and lactose percentages averaged 4.13 ± 0.62 , 3.34 ± 0.36 , and $4.57 \pm 0.19\%$,

Table 1. Descriptive statistics regarding vitamin B₁₂ concentration in milk (ng/dL) analyzed by radioassay (the reference method) for calibration and validation sets¹

Data set	Calibration set						Validation set					
	Herds (no.)	Samples (no.)	Mean	SD	Minimum	Maximum	Herds (no.)	Samples (no.)	Mean	SD	Minimum	Maximum
1	58	2,145	404.4	129.0	98.6	981.5	29	1,009	401.3	122.3	125.7	867.5
2	58	2,105	403.8	127.4	100.9	981.5	29	1,049	402.6	126.1	98.6	888.6
3	58	2,058	402.0	124.4	98.6	888.6	29	1,096	406.1	131.8	100.9	981.5

¹Only data from cleaned data sets shown. They include data from the raw data set in which all absolute residuals >2 times the square root of the residual variance after a first partial least squares regression were deleted.

respectively. Daily vitamin B₁₂ concentration in milk per cow from raw data set averaged 420.9 ± 152.2 ng/dL and ranged from 68.4 to 1,471.6 ng/dL. This wide range is similar to that reported by Rutten et al. (2013). Table 1 shows descriptive statistics for daily vitamin B₁₂ concentration in milk for all calibration and validation data sets based on cleaned data sets. The numbers of samples between calibration and validation sets were different because the splitting of data was based on herds and not on records. The repeatability of the reference method to quantify milk vitamin B₁₂ concentration could be considered good according to the obtained interassay coefficient of variation.

Vitamin B₁₂ concentrations in milk >900 ng/dL were not frequent and could be related to unhealthy cows (Duplessis et al., 2018). As mentioned in the previous section, a PLS regression was first carried out to remove records with too-high absolute residuals from the raw data set to obtain the cleaned data set. This procedure was done until the increase in R²_{cv} was <1%; 7 iterations were necessary to achieve this. Using 30 components in the PLS in the model to remove outliers, the final R²_{cv} was 0.54 for the cleaned data set. This procedure allowed us to keep 86.4% of the initial raw data set. This cleaning procedure removed most of the extremely high values of vitamin B₁₂ concentration in milk.

The number of components used in all final PLS models ranged from 15 to 16. The R²_c and R²_{cv} ranged, respectively, from 0.37 to 0.41 and from 0.34 to 0.38 for the 3 calibration raw data sets, which are low. Moreover, all RPD were <3, suggesting that the developed calibration equations were inadequate to estimate quantitative vitamin B₁₂ concentration in milk (Soyeurt et al., 2011). Regarding validation of the raw data sets, as expected, R²_v were lower and RMSEP_v were higher than those of the cross validation (not shown). Prediction of vitamin B₁₂ concentration in milk improved when using the cleaned data set, as suggested by increases in R²_c and R²_{cv} (Table 2). Soyeurt et al. (2011) obtained similar results after deleting residuals. The equations developed could be considered robust because the differences between R²_c and R²_{cv} were low (i.e., 0.03). However, RPD values were still <3, limiting potential use of the developed equations. The differences between R²_{cv} and R²_v were low except for the first calibration and validation sets.

Vitamins, including vitamin B₁₂, make up about 0.1% of the composition of cow milk (Pereira, 2014). It has been shown that potential MIR prediction was greater for fatty acids present at a higher concentration than for minor fatty acids in milk (<0.1%; De Marchi et al., 2011). Vitamin B₁₂ is a minor milk constituent, which could explain why R²_{cv} of prediction of its

Table 2. Statistics of calibration and validation sets while predicting quantitative vitamin B₁₂ concentration in milk (ng/dL) using spectral data^{1,2}

Data set	Calibration set							Validation set				
	Mean	SD	ncomp ³	R ² _c	RMSEP _c	R ² _{cv}	RMSEP _{cv}	Mean	SD	R ² _v	RMSEP _v	RPD
1	404.4	96.4	16	0.56	85.7	0.53	88.5	394.3	88.0	0.38	96.7	1.3
2	403.8	92.6	16	0.53	87.5	0.50	89.7	402.7	99.1	0.48	91.2	1.4
3	402.0	89.7	16	0.52	86.1	0.49	88.5	406.8	90.8	0.47	96.1	1.3

¹Only data from cleaned data sets shown. They include data from raw data set in which all absolute residuals >2 times the square root of the residual variance after a first partial least squares regression were deleted.

²R²_c, R²_{cv}, and R²_v = coefficients of determination of calibration, cross validation, and validation sets, respectively; RMSEP_c, RMSEP_{cv}, and RMSEP_v = root mean square errors of prediction of calibration, cross validation, and validation sets, respectively; RPD = SD of vitamin B₁₂ concentration in milk by radioassay analysis/RMSEP_v.

³ncomp = number of components.

concentration in milk were low in the current study. However, Grelet et al. (2016) showed that citrate and BHB, which are also minor milk constituents, could be accurately predicted by MIR. When developing a new MIR equation, it is important to determine whether the equation brings new information compared with those already available. Pearson correlation coefficients between fat, protein, and lactose concentrations predicted by MIR using manufacturer equations and vitamin B₁₂ concentration in milk measured by radioassay were 0.37, 0.44, and −0.16, respectively. Averaged cross-validation correlations (0.60 and 0.72 for raw and cleaned data sets, respectively) were higher than absolute Pearson correlation coefficients between major milk components and vitamin B₁₂ concentration in milk, suggesting that the developed equations did add new information.

Percentages of cows having milk concentrations of vitamin B₁₂ >441.6 ng/dL were 38.3, 38.4, and 36.9% in calibration sets 1, 2, and 3 from the cleaned data set, respectively. Evaluation (between 36.0 and 44.7%) and test (between 33.2 and 39.4%) sets had similar values. This means that fewer than 40% of samples fulfilled the RDA (wherein a 250-mL glass of milk provides 46% of daily vitamin B₁₂ RDA for an individual aged 14 yr or older). Alternative probability cutoffs to overcome class imbalance determined using evaluation sets averaged 0.465. The number of components used in all final PLS discriminant analysis models ranged from 15 to 16. Sensitivity, specificity, accuracy, and ROC AUC were similar among the 3 calibration and test sets (Table 3), which implies a good robustness of models. As previously observed with quantitative prediction models, indicators to evaluate PLS discriminant analysis models were smaller for the test sets than for the calibration sets. According to Zhu et al. (2010), a ROC AUC ranging between 0.8 and 0.9 reflects good accuracy of classification, as currently observed when using 441.6 ng/dL as a threshold for vitamin B₁₂ concentration in milk. As measured by the percentage accuracy, from 77 to 79% of samples were correctly categorized regarding calibration sets. Table 3 also shows a confusion matrix for the 3 test sets. These results indicated the potential of using MIR to identify milk samples that fulfill (or not) the RDA for vitamin B₁₂ concentration in milk when a threshold of 441.6 ng/dL is used. This threshold could be of interest to certify the high nutritional value of milk for consumers; that is, that the milk vitamin B₁₂ concentration fulfills the USDA recommendation, hence, promoting this product as a reliable healthy choice. Moreover, this threshold could be used to market a specialty milk with a naturally high vitamin B₁₂ concentration. However, it is important to note that the false discovery rate was between 35 and 39% depending on test sets. This could represent a limitation

Table 3. Receiver operating characteristic area under the curve (ROC AUC), sensitivity (Se), specificity (Sp), accuracy, and confusion matrix from partial least squares analysis to discriminate vitamin B₁₂ concentration in milk classes from the cleaned data set according to a threshold of 441.6 ng/dL¹

Data set	ncomp ²	Probability cutoffs ³		Cross-validation		Calibration set			
				ROC AUC		ROC AUC	Se (%)	Sp (%)	Accuracy (%)
1	15	0.487		0.836		0.854	68	86	79
2	18	0.445		0.828		0.852	77	74	75
3	18	0.464		0.826		0.852	72	80	77
Test set ⁴									
Confusion matrix of test set									
1	Predicted above the threshold		133	73		0.810	60	84	76
	Predicted at or below the threshold		90	375					
2	Predicted above the threshold		206	133		0.824	75	71	73
	Predicted at or below the threshold		69	333					
3	Predicted above the threshold		187	101		0.802	66	77	73
	Predicted at or below the threshold		95	333					

in discriminating milk samples according to vitamin B₁₂ concentration.

One limitation of the current data set is that only Holstein cows were involved. Differences among breeds regarding vitamin B₁₂ concentration in milk have been reported (Duplessis et al., 2018). Including other breeds in this study would have increased the variability of reference values that could help to improve model performance. However, as the data sets included 87 dairy herds, a wide variety of feeding management was used (Duplessis et al., 2019), thus increasing the variability of reference values.

In summary, we showed that predicting quantitative vitamin B₁₂ concentration in milk using spectral data led to low accuracy, limiting use of the developed model. Discriminant analysis models using 441.6 ng/dL as the threshold for classifying samples had good accuracy of classification. However, the false discovery rate averaged 36%, which could limit potential use for discriminating samples according to their vitamin B₁₂ concentration in milk.

ACKNOWLEDGMENTS

The authors acknowledge the invaluable participation of dairy producers in this experiment. We appreciate the help of Roxane Robichaud, Liliana Fadul-Pacheco, Valérie Audet, and Isabelle Duval (Département des sciences animales, Université Laval, Québec, QC, Canada) for data collection; and Valérie Beaudet, Jasmin Brochu, and Liette Veilleux (Agriculture et Agroalimentaire Canada, Sherbrooke, QC, Canada) and Lactanet (formerly Valacta, Sainte-Anne-de-Bellevue, QC, Canada) for laboratory analyses. This work was supported by the Dairy Research Cluster Initiative (Dairy Farmers of Canada, Agriculture and Agri-Food Canada, the Canadian Dairy Network, and the Canadian Dairy Commission; Ottawa, ON, Canada). Sample collection was funded through the Programme de recherche en partenariat sur la préservation et l'amélioration de la valeur nutritive des aliments en lien avec la santé Fondation des maladies du cœur et de l'AVC et Visez santé (FMC), Ministère de l'Agriculture, des Pêcheries et de l'Alimentation (MAPAQ), Ministère de la Santé et des Services Sociaux (MSSS), Fonds de recherche du Québec – Nature et technologies (FRQ-NT) et le Fonds de recherche du Québec – Santé (FRQ-S). The authors have not stated any conflicts of interest.

REFERENCES

- Baik, H. W., and R. M. Russell. 1999. Vitamin B₁₂ deficiency in the elderly. *Annu. Rev. Nutr.* 19:357–377. <https://doi.org/10.1146/annurev.nutr.19.1.357>.
- Brito, A., J. Chiquette, S. P. Stabler, R. H. Allen, and C. L. Girard. 2015. Supplementing lactating dairy cows with a vitamin B₁₂ precursor, 5,6-dimethylbenzimidazole, increases the apparent ruminal synthesis of vitamin B₁₂. *Animal* 9:67–75. <https://doi.org/10.1017/S1751731114002201>.
- Canadian Council on Animal Care. 2009. Guide to the Care and Use of Experimental Animals. 2nd ed. Vol. 1. E. D. Rolfert, B. M. Cross, and A. A. McWilliam ed. Can. Coun. Anim. Care, Ottawa, ON, Canada.
- De Marchi, M., M. Penasa, A. Cecchinato, M. Mele, P. Secchiari, and G. Bittante. 2011. Effectiveness of mid-infrared spectroscopy to predict fatty acid composition of Brown Swiss bovine milk. *Animal* 5:1653–1658. <https://doi.org/10.1017/S175173111000747>.
- Duplessis, M., R. I. Cue, D. E. Santschi, D. M. Lefebvre, and C. L. Girard. 2018. Short communication: Relationships among plasma and milk vitamin B₁₂, plasma free fatty acids, and blood β-hydroxybutyrate concentrations in early lactation dairy cows. *J. Dairy Sci.* 101:8559–8565. <https://doi.org/10.3168/jds.2018-14477>.
- Duplessis, M., H. Lapierre, D. Pellerin, J. P. Laforest, and C. L. Girard. 2017. Effects of intramuscular injections of folic acid, vitamin B₁₂, or both, on lactational performance and energy status of multiparous dairy cows. *J. Dairy Sci.* 100:4051–4064. <https://doi.org/10.3168/jds.2016-12381>.
- Duplessis, M., S. Mann, D. V. Nydam, C. L. Girard, D. Pellerin, and T. R. Overton. 2015. Short communication: Folate and vitamin B₁₂ in colostrum and milk from dairy cows fed different energy levels during the dry period. *J. Dairy Sci.* 98:5454–5459. <https://doi.org/10.3168/jds.2015-9507>.
- Duplessis, M., D. Pellerin, R. Robichaud, L. Fadul-Pacheco, and C. L. Girard. 2019. Impact of diet management and composition on vitamin B₁₂ concentration in milk of Holstein cows. *Animal* 13:2101–2109. <https://doi.org/10.1017/S1751731119000211>.
- Grete, 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>.
- Kuhn, M. 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28:1–26. <https://doi.org/10.18637/jss.v028.i05>.
- Kuhn, M., and K. Johnson. 2013. Remedies for severe class imbalance. Pages 419–443 in *Applied Predictive Modeling*. 1st ed. Springer-Verlag, New York, NY.
- Lê, S., J. Josse, and F. Husson. 2008. FactoMineR: An R package for multivariate analysis. *J. Stat. Softw.* 25:1–18. <https://doi.org/10.18637/jss.v025.i01>.
- Martens, J. H., H. Barg, M. Warren, and D. Jahn. 2002. Microbial production of vitamin B₁₂. *Appl. Microbiol. Biotechnol.* 58:275–285. <https://doi.org/10.1007/s00253-001-0902-7>.
- McDermott, A., G. Visentin, M. De Marchi, D. P. Berry, M. A. Fanelon, P. M. O'Connor, O. A. Kenny, and S. McParland. 2016. Prediction of individual milk proteins including free amino acids in bovine milk using mid-infrared spectroscopy and their correlations with milk processing characteristics. *J. Dairy Sci.* 99:3171–3182. <https://doi.org/10.3168/jds.2015-9747>.
- Mevik, B. H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. *J. Stat. Softw.* 18:1–23. <https://doi.org/10.18637/jss.v018.i02>.
- National Academy of Sciences. 1998. Vitamin B₁₂. Pages 306–356 in *Dietary reference intakes for thiamin, riboflavin, niacin, vitamin B₆, folate, vitamin B₁₂, pantothenic acid, biotin and choline*. National Academy Press, Washington, DC. <https://doi.org/10.17226/6015>.
- Pereira, P. C. 2014. Milk nutritional composition and its role in human health. *Nutrition* 30:619–627. <https://doi.org/10.1016/j.nut.2013.10.011>.
- Preynat, A., H. Lapierre, M. C. Thivierge, M. F. Palin, J. J. Matte, A. Desrochers, and C. L. Girard. 2009. Effects of supplements of folic acid, vitamin B₁₂, and rumen-protected methionine on whole-body metabolism of methionine and glucose in lactating dairy cows. *J. Dairy Sci.* 92:677–689. <https://doi.org/10.3168/jds.2008-1525>.

- Rousseeuw, P. J., M. Debruyne, S. Engelen, and M. Hubert. 2006. Robustness and outlier detection in chemometrics. *Crit. Rev. Anal. Chem.* 36:221–242. <https://doi.org/10.1080/10408340600969403>.
- Rutten, M. J. M., A. C. Bouwman, R. C. Sprong, J. A. M. van Arendonk, and M. H. P. W. Visker. 2013. Genetic variation in vitamin B₁₂ content of bovine milk and its association with SNP along the bovine genome. *PLoS One* 8:e62382. <https://doi.org/10.1371/journal.pone.0062382>.
- Soyeurt, H., D. Bruwier, J. M. Romnee, N. Gengler, C. Bertozzi, D. Veselko, and P. Dardenne. 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454. <https://doi.org/10.3168/jds.2008-1734>.
- 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>.
- USDA-ARS (Agricultural Research Service). 2012. National Nutrient Database for Standard Reference Release 28, Basic Report: 01077, Milk, whole, 3.25% milkfat, with added vitamin D Accessed Oct. 3, 2017. <http://ndb.nal.usda.gov/ndb/foods/list>.
- Whitfield, R. G., M. E. Gerger, and R. L. Sharp. 1987. Near-infrared spectrum qualification via Mahalanobis distance determination. *Appl. Spectrosc.* 41:1204–1213. <https://doi.org/10.1366/0003702874447572>.
- Williams, P. C., and D. C. Sobering. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spectrosc.* 1:25–32. <https://doi.org/10.1255/jnirs.3>.
- Zhu, W., N. Zeng, and N. Wang. 2010. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. In NESUG (Northeast SAS Users Group) Proceedings: Health Care and Life Sciences, Baltimore, MD. https://www.lexjansen.com/cgi-bin/xsl_transform.php?x=shl&c=nesug.

ORCIDS

- M. Duplessis <https://orcid.org/0000-0003-3503-6612>
D. Pellerin <https://orcid.org/0000-0002-4310-1077>
C. L. Girard <https://orcid.org/0000-0002-5470-0841>
D. E. Santschi <https://orcid.org/0000-0002-3938-6073>
H. Soyeurt <https://orcid.org/0000-0001-9883-9047>