Prediction of serum metabolic profile biomarkers in early lactation dairy cows using mid-infrared spectroscopy of milk

T.D.W. Luke[1,2], S. Rochfort[1,2], W.J. Wales[2] and J.E. Pryce[1,2]

[1]Department of Economic Development, Jobs, Transport and Resources and La Trobe University, Agribio, 5 Ring Road, Bundoora, VIC 3083, Australia
[2]School of Applied Systems Biology, La Trobe University, Bundoora, Victoria 3083, Australia
[3]Department of Economic Development, Jobs, Transport & Resources, Ellinbank Centre, 1301 Hazeldean Rd., Ellinbank, Victoria, 3820 Australia

Summary

Metabolic diseases in early lactation have significant negative effects on dairy cow health and welfare, and farm profitability. The most commonly described metabolic diseases are ketosis, hypocalcaemia, and hypomagnesaemia. Subclinical metabolic diseases, which are not associated with obvious clinical signs, are of particular interest due to their relatively high prevalence and significant effects on animal welfare and performance. Currently one of the most common methods for monitoring the metabolic health of cows is serum metabolic profiling, which utilises well-established associations between the concentrations of several metabolites in serum, and the presence of both subclinical and clinical metabolic disease.

An emerging technology to evaluate subclinical metabolic disease is mid-infrared spectroscopic analysis (MIR) of milk samples. In this cross-sectional study we investigated the use of MIR spectroscopy of milk for estimating the concentrations a number of serum metabolites commonly employed in metabolic profiling. A single plain/clotted blood sample was taken from 1027 cows from 5 farms in the Gippsland region of south-eastern Victoria, Australia, on the same day as milk recording. All cows had calved within 8 weeks of sampling. Serum samples were analysed for beta hydroxybutyrate (BHB), non-esterified fatty acids (NEFA), calcium, magnesium, blood urea nitrogen, total protein, albumin and globulins.

Milk samples were analysed by MIR spectroscopy using a Bentley Instruments FTS Combi. Calibration models were constructed using partial least square (PLS) regression, and external validation was performed using both a farm exclusion (a calibration equation derived using data from 4 farms was used to predict the outcome on the 5th farm) and a random sampling method (a random sample of 20% of cows was excluded from the calibration dataset and used for validation). The R^2 and root mean square error (RMSE) values for MIR predictions using random external validation were 0.49 (0.19) for BHB, 0.51 (0.24) for NEFA, 0.88 (0.79) for Urea, 0.18 (0.14) for calcium, 0.23 (0.10) for magnesium, 0.28 (2.01) for albumin, and 0.31 (4.76) for globulins.

Corresponding Author: tim.luke@ecodev.vic.gov.au

Our results demonstrate that MIR spectroscopic analysis of milk shows promise for evaluating short-term protein status of animals through accurate estimation of BUN concentration, and reasonable prediction of energy balance by estimation of serum BHB and NEFA concentrations.

Keywords: Mid-infra-red spectral prediction, metabolic profiles, ketosis.

Introduction

Metabolic diseases in early lactation have significant negative effects on dairy cow health and welfare, and farm profitability (McArt *et al.*, 2015, Suthar *et al.*, 2013). The most commonly described metabolic diseases are ketosis, hypocalcaemia and hypomagnesaemia. Subclinical metabolic diseases, which are not associated with obvious clinical signs, are of particular interest due to their relatively high prevalence and significant effects on animal welfare and performance (Compton *et al.*, 2013, Macrae *et al.*, 2006, McArt *et al.*, 2012, Suthar *et al.*, 2013).

Serum metabolic profile testing (MPT) utilises well-established epidemiological associations between the concentrations of several metabolites in serum and the presence of both subclinical and clinical metabolic disease, and provides objective information on the metabolic health and nutritional status of cows (Payne *et al.*, 1970). The metabolites evaluated in MPT vary, but often include beta-hydroxybutyrate (BHB) and non-esterified fatty acids (NEFAs) as indicators of energy balance, albumin, globulins and blood urea nitrogen (BUN) as indicators of protein status, and magnesium (Mg) and calcium (Ca) and as indicators of mineral status (Anderson, 2009, Whitaker, 2004). NEFA and BHB are particularly important, as elevated concentrations of one or both of these metabolites are indicative of mal-adaptation to the period of post-partum negative energy balance and are associated with an increased risk of subsequent negative health and production outcomes (Ospina *et al.*, 2010, Sordillo and Raphael, 2013, Chapinal *et al.*, 2012)

Despite the advantages of MPT, blood testing animals on a regular basis is invasive, logistically challenging and potentially costly. Given the ready availability of milk, its use as a biofluid for diagnostic purposes is of increasing interest. A milk fat to protein ratio (FTP) of greater than 1.4 has been described as an indicator of negative energy balance and subclinical ketosis in early lactation (Schcolnik, 2016). More recently, mid-infrared spectroscopy (MIRS) of milk has shown promise for predicting the risk of subclinical ketosis and evaluating energy balance in early lactation (McParland *et al.*, 2011, Grelet *et al.*, 2016, van Knegsel *et al.*, 2010, Belay *et al.*, 2017, de Roos *et al.*, 2007). The aim of our study was to determine if MIRS of milk is a better predictor of subclinical ketosis risk and negative energy balance in early lactation dairy cows than milk fat to protein ratio, and whether MIRS can be used to estimate the concentrations of other serum metabolites.

Materials and methods

All procedures were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes (National Health and Medical Research Council 2004). Approval to proceed was obtained from the DEDJTR Agricultural Research and Extension Animal Ethics Committee (Department of Economic Development, Jobs, Transport and Resources, 475 Mickleham Road, Attwood, Victoria 3049, Australia).

Sample collection

Blood samples were taken from 1028 cows on the same day as milk-recording between July and October 2017. The cows were located on five farms in the Gippsland region of south-eastern Australia. All five farms implement a feeding system reliant on grazed pasture plus other forages and more than one tonne grain/concentrates fed in the bail per cow per year. Three of the farms (farms 3, 4 and 5) had rotary milking platforms which allowed blood samples to be collected during milking. Samples were taken immediately after milking on the other two farms. Samples were collected at the morning milking on farms 1 and 3, and during the afternoon milking on farms 2, 4 and 5. Blood was collected from the coccygeal vein into 10ml serum clot activator vacutainer tubes (Becton Dickinson). Samples were allowed to clot for a minimum of one hour at room temperature before centrifugation at 1200 g for ten minutes at 18°C. All samples were processed within six hours of collection. Serum samples were refrigerated at 4°C then transported on ice to Regional Laboratory Services (Benalla, Victoria, Australia). Samples were analysed for BHB, NEFA, total Ca, Mg, total protein and albumin using colourimetric analysis, with reagents supplied by Randox Laboratories. Globulin concentrations were calculated as total protein concentration minus albumin concentration. Milk samples were collected as part of normal milk-recording by the Herd Improvement Co-Operative Australia (Maffra, Victoria, Australia). Samples were preserved with bronopol and analysed fresh using mid-infrared (MIR) spectroscopy (Bentley Instruments NexGen FTS Combi) by TasHerd Pty Ltd (Hadspen, Tasmania, Australia). MIR spectra were expressed in absorbance, with 899 spectral points between 649cm^{-1} and 3998cm^{-1}. Spectral regions associated with the O-H bending and stretching regions of water were excluded prior to analysis (Afseth *et al.*, 2010, Belay *et al.*, 2017). A total of 563 spectral wavelengths between 928cm^{-1} and 3025cm^{-1} were included in the analysis.

Statistical analysis

Descriptive statistics

Descriptive statistics of the animals included in the analysis are summarised in Table 1.

The distribution of serum metabolite concentrations were visually assessed for normality using frequency histograms. NEFA and BHB concentration distributions were both skewed, with lower values over represented. This type of distribution leads to decreased accuracy in predicting high values (Grelet *et al.*, 2016), so a logarithmic (10) transformation was applied.

MIR predictions

All MIR spectral data analysis was performed with MATLAB R2017a (MathWorks, Natick, MA) utilising PLS Toolbox (Eigenvector Research, Manson, WA).

Table 1. Number of cows that had metabolic profiles and milk MIR spectral data by farm, including stage of lactation (days in milk means and ranges) and percentage of animals in their first lactation.

Data	No. of cows	Percentage primiparous	Days in milk		
			Mean	Min	Max
Farm 1	480	17.7	23.15	4	53
Farm 2	137	27.0	19.94	3	39
Farm 3	81	29.6	27.48	5	56
Farm 4	149	12.1	21.01	4	39
Farm 5	180	17.8	29.51	6	52
Total	1027	19.1	23.86	3	56

Pre-processing of spectra

Preliminary analysis of the spectral data was conducting using principal component analysis (PCA). A single outlier spectrum with a zero absorbance was identified from the PCA plot and removed from the data set, leaving 1027 samples. MIR spectra were pre-processed with Savitzy-Golay 2nd derivative transformation and smoothing, removal of linear trend and auto scaling.

Calibration and validation

The relationship between blood metabolite concentrations and milk MIR spectra was investigated using partial least square (PLS) regression analysis.

Cross-validation (CV) was performed using a venetian blinds method, which split the data into 20 subsets and performed CV on two samples per subset. Each model was assessed for over-fitting using a permutation test with 50 iterations, the statistical significance ($P<0.05$) of which was tested using a Wilcoxon sign test.

The number of latent variables (LV) included in each model was based on maximising the percentage of variance captured while minimising the root mean square error of cross-validation ($RMSE_{CV}$). The optimum number of LVs was determined by examining a plot $RMSE_{CV}$ as a function of number of LVs

External validation (EV) was performed using two methods. The first method involved randomly sorting the data then splitting it into calibration (n=822) and validation (n=205) sets. Each dataset was designed to have a representative number of samples from each farm and parity category (primiparous or multiparous), and was balanced for days in milk. The second method involved using data from one farm to validate calibration equations constructed using data from the remaining four farms.

Comparison with milk fat to protein ratio

The usefulness of milk MIR predictions of serum NEFA and BHB for estimating the degree of fat mobilisation and the risk of ketosis respectively, was determined by comparing the accuracy of MIR prediction outputs with that of milk fat to protein ratio.

Multiple linear regression models were constructed to predict serum BHB and NEFA concentrations using combinations of MIR prediction of BHB (MIR_{BHB}), MIR prediction of NEFA (MIR_{NEFA}) and milk fat to protein ratio. Farm identification, parity category and weeks in milk (WIM) were included as fixed effects in each model. The significance ($P<0.001$) of each predictor variable within each model was determined using ANOVA F-tests. The overall accuracy of the models was assessed by comparing the adjusted R^2 values.

Results and discussion

Descriptive statistics

Of the 1027 animals included in the analysis, 480 (46.74%) were from one farm. The remaining 547 animals were evenly distributed between the remaining four farms. Of the animals sampled, 757 (73.7%) had been calved 30 days or less, which is the period of highest risk for development of metabolic disease (LeBlanc *et al.*, 2006). The overall percentage of primiparous animals in the dataset was 19.1%, with a range of 12.1% to 29.6% between farms.

MIR calibration and validation

The results of PLS regression models investigating the relationships between blood metabolite concentrations and MIR spectra from milk samples are shown in Table 2. The coefficient of determination of CV (R^2_{CV}) and EV (R^2_{EV}) for the model predicting serum BHB in the random calibration and validation analysis were 0.49 and 0.49 respectively. The results for the NEFA model were similar with an R^2_{CV} of 0.46 and an R^2_{EV} of 0.51. The most promising results were for the model predicting serum urea concentration, which had an R^2_{CV} of 0.89 and an R^2_{EV} of 0.88. The prediction models for both Ca and Mg had comparatively low accuracies with R^2_{CV} and R^2_{EV} of 0.18 and 0.18, and 0.17 and 0.23 respectively. Serum albumin and globulin concentration prediction models performed slightly better with R^2_{CV} and R^2_{EV} values of 0.26 and 0.28, and 0.20 and 0.30 respectively.

The results of the second external validation method, which involved using prediction equations derived from the data from four farms to predict the outcome on the excluded farm, were consistently lower than the results of the random external validation. This is important from a practical point of view, as it suggests that if MIRS is to be employed commercially to predict the metabolic health of cows from multiple herds, the results may be more accurate if data from a subset of animals in each herd being tested are included in the calibration/reference dataset. The relative accuracies of prediction models using farm exclusion external validation were similar to the results of random external validation, with urea, BHB and NEFA models having the highest R^2_{EV} values.

Table 2. Results of partial least square regression models for the prediction of serum metabolite concentrations from milk MIR spectra, using a random external validation method.

		Cross-validation (n=822)			External Validation (n=205)	
Metabolite	No. LVs[1]	R^2_{cv}[2]	$RMSE_{cv}$[3]	p-value	R^2_{ev}[4]	$RMSE_{ev}$[5]
BHB	12	0.49	0.16	< 0.05	0.49	0.19
NEFA	12	0.46	0.25	< 0.05	0.51	0.24
Urea	20	0.89	0.77	< 0.05	0.88	0.79
Calcium	16	0.18	0.14	< 0.05	0.18	0.13
Magnesium	16	0.17	0.11	< 0.05	0.23	0.10
Albumin	16	0.26	2.07	< 0.05	0.28	2.01
Globulin	16	0.20	5.56	< 0.05	0.31	4.76
Alb:Glob	14	0.23	0.15	< 0.05	0.27	0.13

[1] Number of latent variables included in the model
[2] Coefficient of determination for cross-validation
[3] Root mean square error of cross-validation
[4] Coefficient of determination for external validation
[5] Root mean square error of external validation

Comparison of MIR with milk fat:protein

The results of fixed effect models constructed to compare the value of milk MIR predictions of serum BHB and NEFA concentrations with milk fat to protein ratio for predicting the risk of subclinical ketosis and negative energy balance respectively, are shown in Table 3.

MIR_{BHB} was considerably better at estimating serum BHB, and therefore the risk of ketosis, than milk fat to protein ratio. Similarly, MIR_{NEFA} was significantly better at estimating serum NEFA concentrations, and therefore the degree of negative energy balance, than milk fat to protein ratio. As mentioned previously, the accuracies of models using farm exclusion external validation were much lower than the accuracies of models using random external validation. However, even when the accuracies of

Table 3. The accuracy (R^2) of predicting SCK risk and fat mobilisation using milk fat:protein (FTP), and MIRS predictions of serum BHB (MIR_{BHB}) and NEFA (MIR_{NEFA}) concentrations.

	Risk of subclinical ketosis (BHB)		Negative energy balance (NEFA)	
Data	FTP	MIR_{BHB}	FTP	MIR_{NEFA}
Random	0.21	0.47	0.44	0.59
Farm 1	0.19	0.32	0.34	0.36
Farm 2	0.21	0.40	0.43	0.61
Farm 3	0.00	0.35	0.00	0.03
Farm 4	0.17	0.27	0.31	0.41
Farm 5	0.06	0.29	0.20	0.29

prediction models were low, MIRS out-performed milk fat to protein ratio. For example, the R^2_{EV} of the MIR_{NEFA} model for farm 3 was only 0.08, yet it was still a better indicator of serum NEFA concentration than milk fat to protein ratio.

ANOVA testing of the fixed effect models demonstrated that MIR_{BHB} were a poor predictor of serum NEFA concentrations, and therefore negative energy balance. Similarly, the usefulness of MIR_{NEFA} for the prediction of serum BHB concentration, and therefore SCK risk, was also poor. This suggests that the PLS prediction models for NEFA and BHB are significantly different, and are likely utilising different parts of the MIR spectra. This highlights the need for further studies to better understand the association between biomarkers in serum and milk, and the potential use of milk MIR spectral data for further investigation of the genetic parameters of metabolic diseases.

Conclusion

In this study we assessed the accuracy of MIRS performed as part of routine milk-recording for predicting the metabolic health and nutritional status of early lactation dairy cows. We found that MIRS of milk provided accurate estimation of BUN concentration, and good prediction of energy balance by reasonable estimation of serum BHB and NEFA concentrations. The accuracy of MIRS of milk for predicting serum albumin and globulin concentrations, and mineral concentrations (as estimated by serum Ca and Mg concentrations), was poor. We also noted that the accuracy of external prediction was much higher when animals from all farms were represented in both the calibration and validation data sets. This has important implications for the commercial application of MIRS, and implies that it is important to include data from as many herds as possible in the calibration data set.

Acknowledgements

We thank DairyBio, jointly funded by Dairy Australia and Agriculture Victoria for funding this project and T. Luke's PhD project. This work was part of a project "MIRprofit: integrating very large genomic and milk midinfrared data to improve profitability of dairy cows", supported by funding from the Australian Government Department of Agriculture as part of the its Rural R&D for Profit programme.

List of references

AFSETH, N. K., MARTENS, H., RANDBY, A., GIDSKEHAUG, L., NARUM, B., JORGENSEN, K., LIEN, S. & KOHLER, A. 2010. Predicting the fatty acid composition of milk: a comparison of two Fourier transform infrared sampling techniques. Appl Spectrosc, 64, 700-7.

ANDERSON, D. E. 2009. Current Veterinary Therapy: Food Animal Practice, St. Louis, MO, Saunders Elsevier.

BELAY, T. K., DAGNACHEW, B. S., KOWALSKI, Z. M. & ADNOY, T. 2017. An attempt at predicting blood beta-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. Journal of Dairy Science, 100, 6312-6326.

CHAPINAL, N., LEBLANC, S. J., CARSON, M. E., LESLIE, K. E., GODDEN, S., CAPEL, M., SANTOS, J. E., OVERTON, M. W. & DUFFIELD, T. F. 2012. Herd-level association of serum metabolites in the transition period with disease, milk production, and early lactation reproductive performance. J Dairy Sci, 95, 5676-82.

COMPTON, C., MCDOUGALL, S., YOUNG, L. & BRYAN, M. 2013. Prevalence of subclinical ketosis in mainly pasture-grazed dairy cows in New Zealand in early lactation. New Zealand Veterinary Journal, 1-8.

DE ROOS, A. P. W., VAN DEN BIJGAART, H. J. C. M., HØRLYK, J. & DE JONG, G. 2007. Screening for Subclinical Ketosis in Dairy Cattle by Fourier Transform Infrared Spectrometry. Journal of Dairy Science, 90, 1761-1766.

GRELET, C., BASTIN, C., GELÉ, M., DAVIÈRE, J. B., JOHAN, M., WERNER, A., READING, R., FERNANDEZ PIERNA, J. A., COLINET, F. G., DARDENNE, P., GENGLER, N., SOYEURT, H. & DEHARENG, F. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, beta-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. Journal of Dairy Science, 99, 4816-4825.

LEBLANC, S. J., LISSEMORE, K. D., KELTON, D. F., DUFFIELD, T. F. & LESLIE, K. E. 2006. Major Advances in Disease Prevention in Dairy Cattle. Journal of Dairy Science, 89, 1267-1279.

MACRAE, A. I., WHITAKER, D. A., BURROUGH, E., DOWELL, A. & KELLY, J. M. 2006. Use of metabolic profiles for the assessment of dietary adequacy in UK dairy herds. Veterinary Record, 159, 655.

MCART, J. A. A., NYDAM, D. V. & OETZEL, G. R. 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. Journal of Dairy Science, 95, 5056-5066.

MCART, J. A. A., NYDAM, D. V. & OVERTON, M. W. 2015. Hyperketonemia in early lactation dairy cattle: A deterministic estimate of component and total cost per case. Journal of Dairy Science, 98, 2043-2054.

MCPARLAND, S., BANOS, G., WALL, E., COFFEY, M. P., SOYEURT, H., VEERKAMP, R. F. & BERRY, D. P. 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. Journal of Dairy Science, 94, 3651-3661.

OSPINA, P. A., NYDAM, D. V., STOKOL, T. & OVERTON, T. R. 2010. Association between the proportion of sampled transition cows with increased nonesterified fatty acids and beta-hydroxybutyrate and disease incidence, pregnancy rate, and milk production at the herd level. J Dairy Sci, 93, 3595-601.

PAYNE, J., DEW, S. M., MANSTON, R. & FAULKS, M. 1970. The use of a metabolic profile test in dairy herds. Veterinary Record, 87, 150.

SCHCOLNIK, T. 2016. Using milk fat-to-protein ratio to evaluate dairy cows energy balance status.(Report)(Author abstract). Journal of Animal Science, 94, 54.

SORDILLO, L. M. & RAPHAEL, W. 2013. Significance of Metabolic Stress, Lipid Mobilization, and Inflammation on Transition Cow Disorders. Veterinary Clinics of North America: Food Animal Practice, 29, 267-278.

SUTHAR, V. S., CANELAS-RAPOSO, J., DENIZ, A. & HEUWIESER, W. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. Journal of Dairy Science, 96, 2925-2938.

VAN KNEGSEL, A. T. M., VAN DER DRIFT, S. G. A., HORNEMAN, M., DE ROOS, A. P. W., KEMP, B. & GRAAT, E. A. M. 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. Journal of Dairy Science, 93, 3065-3069.

WHITAKER, D. A. 2004. Metabolic profiles. In Bovine Medicine: Diseases and Husbandry of Cattle, Oxford, Blackwell Science.