Improvements and validation of milk fatty acid predictions using mid-infrared spectrometry

H Soyeurt[1,2], S McParland[4], D Berry[4], E Wall[5], N Gengler[1,2], F Dehareng[3], P Dardenne[3]
[1]Gembloux Agro-Bio Tech - University of Liège, Animal Science Unit, Gembloux, Belgium, [2]National Funds for Scientific Research, Brussels, Belgium, [3]Walloon Agricultural Research Centre, Valorisation of Agricultural Products, Gembloux, Belgium, [4]Teagasc, Moorepark Dairy Production Research Centre, Fermoy, Cork, Ireland, [5]Scottish Agricultural College, Sustainable Livestock Systems, Penicuik, Midlothian, United Kingdom *Email: hsoyeurt@ulg.ac.be*

Introduction The requirement for gas chromatography (GC) to quantify milk FA concentration is expensive to be undertaken on large number of samples. The recent development of equations based on mid-infrared spectrometry (MIR) for the prediction of milk FA content (Soyeurt *et al.*, 2006) offers a solution. The first objective was to improve the predictions of FA by using different approaches. The second objective was to validate these new equations using an independent sample set.

Materials and methods The calibration set contained 239 Belgian milk samples collected between March 2005 and December 2007 from several cows and breeds. These samples were selected based on their spectral variability. The MIR spectrum from each sample was obtained (Foss MilkoScan FT6000) and the FA content of each sample was quantified by GC. The equations were built by Foss WINISI software using partial least squares (PLS) and/or first derivation and/or repeatability file. Using a derivative applied to the spectra permits to normalize the spectral data. The repeatability file contained spectra generated from the same samples but from five different spectrometers. A cross-validation using 20 groups from the calibration set was used to estimate the accuracy of the FA predictions. The methods used were 1) just PLS, 2) PLS and first derivative, 3) PLS and repeatability and 4) PLS, first derivative and repeatability (*). These methods were compared using the ratio of the standard deviation of GC results (SD) to the standard error of cross-validation (RPD). An external validation was done using 362 GC independent milk samples collected between April 2008 and August 2009 from several breeds and cows in Belgium, Ireland, and Scotland to confirm the results obtained by cross-validation. The validation coefficient of determination (R^2v) was calculated.

Results If RPD is superior to 3, the predictions given by the equation can be considered as good. Table 1 presents results for equations showing RPD superior or equal to 3 and summarises the descriptive statistics of the GC results. As expected, the equations with the higher values of RPD were globally from groups of FA, rather than the individual FA. The different approaches used to develop the equations showed generally different RPD values. All equations were not better using a calibration equation built from first derivative and repeatability file even if the results were generally better with this approach. These results suggest adapting the methodology used to develop the equation in function of the studied FA. R^2v shown in Table 1 confirms it. The highest R^2v were observed for the same equations, which showed the highest RPD values except for C18:0.

Table 1 Descriptive statistics of the calibration set, RPD values, and R^2v obtained from the developed equations using the 4 methods.

	N=239		RPD (N=239)				R^2v (N=362)			
Constituent	Mean	SD	1 (*)	2 (*)	3 (*)	4 (*)	1 (*)	2 (*)	3 (*)	4 (*)
C6:0	0.08	0.02	3.95	4.02	3.89	3.95	0.88	0.90	0.87	0.90
C8:0	0.05	0.02	3.21	3.27	3.21	3.33	0.84	0.88	0.86	0.81
C10:0	0.12	0.04	3.03	2.99	3.07	3.07	0.82	0.87	0.80	0.73
C14:0	0.48	0.14	3.51	3.62	3.66	3.70	0.90	0.90	0.90	0.90
C16:0	1.29	0.42	3.07	3.12	3.17	3.16	0.91	0.90	0.90	0.90
C18:0	0.49	0.23	2.89	2.93	2.90	3.01	0.73	0.62	0.74	0.72
Total C18:1 trans	0.15	0.09	3.16	3.09	3.05	3.09	0.46	0.46	0.52	0.49
C18:1 cis-9	0.89	0.36	4.61	4.68	4.35	4.60	0.86	0.92	0.93	0.91
Total C18:1 cis	0.96	0.37	4.62	4.71	4.50	4.73	0.85	0.93	0.94	0.93
Saturated	2.98	0.85	9.34	10.01	9.55	9.95	0.98	0.98	0.98	0.98
Monounsaturated	1.26	0.43	5.47	5.85	5.41	5.88	0.93	0.95	0.95	0.95
Unsaturated	1.46	0.48	5.82	6.24	5.77	6.26	0.93	0.95	0.96	0.96
Short chain (C4-C10)	0.39	0.11	3.85	3.96	3.90	3.97	0.89	0.91	0.91	0.93
Medium chain (C12-C16)	2.19	0.64	4.10	4.19	4.14	4.27	0.92	0.94	0.92	0.94
Long chain (C17-C22)	1.86	0.69	4.56	4.86	4.64	4.93	0.94	0.92	0.95	0.95

Conclusions MIR is a good technology to predict the contents of major FA in milk especially saturated fatty acids. Results presented here are superior to those presented by Soyeurt *et al.* (2006). The 3rd and 4th proposed methodologies give globally the best results.

Acknowledgement The authors gratefully acknowledge funding from the FP7 KBBE-2007-1 project, RobustMilk. The National Fund for Scientific Research is also acknowledged for its financial support through the FRFC 2.4.623.08.F.

References

Soyeurt, H., P. Dardenne, G. Lognay, D. Veselko, P. Mayeres, and N. Gengler. 2006. Journal of Dairy Science 89, 3690-3695.