Effects of Sample Age on the Suitability of Fresh Milk Samples for Instrumental Composition Analysis

J. P. MELICHERCIK,* C. D. FULTON,* L. SZIJARTO,* and A. R. HILL[1,†]

*Laboratory Services Division and
†Department of Food Science,
University of Guelph, Guelph, ON, Canada N1G 2W1

ABSTRACT

This investigation was undertaken by the Ontario Agriculture and Food Laboratory Services Centre to determine whether the officially acceptable time for testing milk composition could be extended from the currently allowed maximum of 5 d to at least 6 d. Samples of milk from producers were divided into aliquots and stored at 4°C in separate vials. Each day, one vial from each sample was tested on infrared instruments. Daily testing continued until the physical properties of the sample made it impossible to introduce the sample to the instrument. Signals for fat A (5.7 μm), protein (6.5 μm), lactose (9.6 μm), and fat B (3.5 μm) and estimates for fat, protein, lactose, and lactose plus other solids were recorded. The maximum sample age at testing was 16 d, and the total number of observations was about 1220. Small but statistically significant effects of age were found both within and beyond the currently accepted testing period of 2 to 5 d. However, there was no consistent effect of age on instrumental estimates of protein, fat, lactose, or lactose plus other solids in samples less than 10 d old. Signals and estimates for lactose were most affected by age.

(**Key words**: fresh milk samples, storage time, instrumental analyses, composition estimates)

Abbreviation key: **LOS** = lactose plus other solids, **LSM** = least squares mean.

INTRODUCTION

Several reports (1, 5, 6, 7) confirm that sampling fresh milk for composition analysis of producer milk is as accurate as composite sampling. There is evidence that fresh samples produce higher estimates for all three major components (1) but no evidence that fresh milk sampling is less accurate than composite sampling (7 or 8 samples per composite; two composites tested per month) provided that testing frequency is increased to four tests per month. The requirements of systems for sampling fresh milk and an evaluation of two cooling systems for fresh milk samples have been described (9).

The primary requirements for fresh milk sampling relate to storage time and consistent refrigerated temperature. Fresh milk samples should be stored and transported at temperatures less than 4°C and should reach the testing laboratory as soon as possible after sampling, preferably within 24 h (4). However, no investigation on acceptable storage time of fresh milk samples for instrumental analysis of composition has been reported. Biggs et al. (2) compared results obtained from samples that were 0, 3, 4, or 5 d old. Differences between days were significant, but there was no significant trend over time; therefore, the acceptability of samples older than 5 d was not determined.

The acceptable age of samples has become an important financial issue to the dairy industry in Ontario, where a central laboratory tests all producer milks for the entire province. Based partly on logistical considerations and partly on results from an earlier investigation (1), a policy was established that instrumental analysis for payment purposes would use fresh samples between 2 and 5 d old. Because of certain logistical factors related to sample transport and replacement of missing samples, considerable economy is realized by extending the acceptable time for milk composition testing from 5 to 6 d. There was also interest in determining the longest storage time for raw milk samples that would be consistent with accurate estimates of composition.

This study, therefore, was designed to determine whether composition results obtained using 6-d-old samples (d 6) were different from control samples tested on d 2 and to determine the maximum period that samples could be held before the estimates of composition differed significantly from those taken on d 2.

Received January 13, 1997.
Accepted July 30, 1997.
[1]Author to whom correspondence should be addressed.

MATERIALS AND METHODS

Sampling

Two 4.5-L samples of milk were obtained directly from farm bulk tanks. Producer milks were selected to provide a range of fat contents from approximately 3.0 to 5.5%. Bulk tank agitators were operated for 5 min before sampling. Samples were taken by draining milk from the gate at the bottom of the tank rather than by dipping from the surface. All samples were between 0.5 and 4°C at the time of sampling. The samples were refrigerated during transport to the laboratory. The entire 9-L sample was poured into a single small vat and agitated with a magnetic stirrer. Agitation continued while 35-ml aliquots were manually poured into 37-ml capacity vials. The sample vials were stored at 4°C until required for instrumental analysis.

Instrumental Analysis

Each day, one aliquot from each sample was tested on one of three infrared milk analyzers (Multispec Micronul M; Foss Electric A/S, Hillerød, Denmark). A brief description of sample preparation and instrument calibration is given herein [Hill et al. (3) report more detail]. Samples were removed from cold storage, manually mixed with eight inversion cycles through an arc of 8 cm in height, placed in a water bath maintained at 40 to 42°C for 15 to 25 min, mixed again with eight inversion cycles, and introduced to the instrument via an automatic sampler. Equations for calibration of the instruments were determined by multiple linear regression of instrument signals against chemical reference values using the model: fat A (5.7 μm) + protein (6.5 μm) + lactose (9.6 μm) + fat B (3.5 μm). Five reference samples were selected each day from each of three instruments and were analyzed for fat, protein (N × 6.38), and total solids by standard methods described previously (1). Reference values for lactose plus other solids (**LOS**) were determined by subtracting the combined values for fat and protein from the values for total solids. New calibration equations were made weekly using the most recent 50 reference results. Adjustments for bias were made daily as required.

Daily testing continued until changes in physical properties made it impossible to introduce the sample to the instrument. Estimates were recorded for fat A, fat B, protein, and lactose signals and for fat, protein, lactose, and LOS.

TABLE 1. Summary of ANOVA statistics for effect of sample age on component estimate differences (d N estimates minus d 2 estimates).

Variable	df	$P > F$	R^2
Fat differences (n = 1221)			
Age (A)	10	0.0119	
Block (B)	8	1×10^{-16}	
Fat (F)	1	0.0035	
F × A	10	0.0142	
Model	29	3×10^{-19}	0.1212
Protein differences (n = 1231)			
A	10	3×10^{-09}	
B	8	5×10^{-06}	
Instrument	2	0.0489	
Model	20	5×10^{-16}	0.0927
Lactose differences (n = 1231)			
A	10	0.0005	
B	8	7×10^{-23}	
F	1	0.0006	
F × A	10	0.0382	
Model	29	2×10^{-62}	0.2673
Lactose plus other solids differences (n = 1231)			
A	10	0.0019	
B	8	3×10^{-46}	
F	1	0.0349	
F × A	10	0.0583	
Model	29	1×10^{-100}	0.3723

Statistical Design and Analysis

The prescribed sampling period for fresh milk samples at the Ontario milk testing laboratory was 2 to 5 d; d 0 was the day the sample was taken. For the purpose of this investigation, d 2 was chosen as the reference point, and experimental results were reported as the value of estimates or signals on d N minus the corresponding value for the same milk on d 2. The relevant statistical question is whether differences for d N were significantly different ($P < 0.05$) from 0.

The experiment was an unbalanced factorial design. Sample age (3, 5, 6, 7, 8, 9, 10, 12, 13, 14, and 16 d), block (nine producer milks), instruments (1, 2, and 3), and fat content were continuous quantitative variables. The entire experiment was replicated 17 times for a total of about 1220 observations.

Sample age, blocks, and instrument were treated as qualitative or class variables (ANOVA tables in Tables 1 and 2). The design was unbalanced because some older samples were unavailable for testing because of physical instability, and logistics prevented other samples from being tested on certain days (holidays and weekends). For the data on fat, 10 outliers were removed to normalize the distribution of data. The data for the other components did not meet assumptions with respect to normality (i.e., a univariate test for nonnormal distribution was highly signifi-

cant). However, the frequency plots (not shown) appeared normal. Nonnormality is frequently observed for large data files because univariate tests for normality become more conservative as the number of observations increases. Tables 3 and 4 indicate the number of observations on each day.

Sample age could be treated as a quantitative variable to evaluate overall trends (e.g., by regression analysis). However, we were interested in specific comparisons as defined in the introduction, so it was prudent to test for significant differences between mean estimates for specific days. Therefore, sample age was treated as a class variable. Equation [1] is the full model. Models for individual component estimates and signals (Tables 1 and 2) are simplified to exclude statistically unimportant variables.

$$Y = A + B + I + F + (A \times I) + (F \times A) + (F \times I) + (F \times A \times I) \quad [1]$$

where

Y = component or signal difference (d N – d 2),
A = sample age (days),
B = producer milks,
I = instrument number, and
F = fat estimate.

Statistical analysis was performed using SAS for Personal Computers (8).

RESULTS AND DISCUSSION

Significant effects for sample age were indicated for all component estimates (Table 1), and specific statistical comparisons were necessary to determine which daily means were significantly different from those on d 2. Significant effects of age on component estimates and instrument signals are summarized in Tables 3 and 4, respectively. Daily mean differences, standard deviations of differences, least squares means (**LSM**), and corresponding standard errors of estimates are tabulated in Tables 3 and 4. Bar charts showing LSM and standard errors of estimates for all components and signals are shown in Figure 1, A through H. The LSM are adjusted by a statistical averaging procedure of SAS that considers unequal numbers of observations, unequal variance within class levels (e.g., within days), and interaction effects among independent variables. Some LSM for signals (Table 4) could not be estimated because of missing observations and significant three-way interactions.

All of the component and signal differences show some days for which the differences between d N and d 2 vary significantly from 0 (Tables 5 and 6). Low R^2 values for all components (Tables 1 and 2) and apparently inconsistent age effects indicate that most of the variation in differences (d N – d 2) was not explained by the models. Because of numerous changes associated with aging, deterioration with age would likely have different effects on infrared signals for different samples. For example, bacterial metabolism of lactose would decrease lactose signals, but associated bacterial metabolites could absorb at any wavelength, which is another reason to determine the earliest day after which component estimates are consistently different from d 2 estimates, rather than only looking for trends in component estimates.

A consistent effect is that all component differences were significant for d 13, 14, and 16. The variance in fat estimates between d 2 and d 3 is largely due to variance in the fat B signal rather than to variance in the fat A signal (Figure 1, F and E, respectively).

The occurrence of significant differences during d 3 to d 10 appears to be due to parameters other than age. Numerous sources of variation exist, but none is

TABLE 2. Summary of ANOVA statistics on instrument (Instr.) signal differences (d N – d 2)

Variable	df	*P* > *F*	R^2
Fat A differences (n = 1263)			
Age (A)	10	0.0014	
Block (B)	8	2×10^{-16}	
Instr.	2	0.0031	
Fat (F)	1	0.1771	
A × Instr.	17	0.055	
F × A	10	0.0003	
F × Instr.	2	0.0021	
F × A × Instr.	17	0.0134	
Model	67	3×10^{-19}	0.2208
Fat B differences (n = 1263)			
A	10	0.0295	
B	8	1×10^{-16}	
Instr.	2	0.0835	
F	1	0.0855	
A × Instr.	17	0.0002	
F × A	10	0.0313	
F × A × Instr.	17	0.0002	
Model	67	5×10^{-24}	0.1890
Protein differences (n = 1263)			
A	10	.3577	
B	8	3×10^{-22}	
Instr.	2	0.8196	
F	1	0.4068	
F × A	10	0.0510	
Model	31	3×10^{-39}	0.1958
Lactose plus other solids differences (n = 1263)			
A	10	7×10^{-56}	
B	8	2×10^{-44}	
Model	18	4×10^{-88}	0.3141

an adequate explanation of the significant differences between days. In particular, significant differences between d 3 and 2 for fat, lactose, and LOS estimates appear inexplicable. On the plus side, this difference is evidence that the inconsistent but significant effects between days for samples fewer than 10 d are not due to spoilage. Lipolysis occurs during aging, but significant effects because of lipolysis would have a greater effect on fat A (fat A measures the ester linkage) than on fat B. Some of the observed variation is certainly due to incomplete blending of fat globules before the sample is introduced to the instrument (1, 2), but this effect should occur equally at all ages or perhaps increase with age.

TABLE 3. Effect of sample age on component estimate mean differences (diff.) (d N – d 2) and least squares mean (LSM) differences.

Sample age	N	Mean diff.	SED[1]	LSM	SE LSM	$P > t$
			Fat estimates			
3	154	–0.01	0.05	–0.02	0.0039	2×10^{-06}
5	102	–0.01	0.04	0.00	0.0048	0.8339
6	192	–0.00	0.03	–0.01	0.0034	0.1023
7	195	0.00	0.06	–0.00	0.0034	0.6067
8	196	0.00	0.04	0.00	0.0034	0.5298
9	145	0.00	0.04	–0.00	0.004	0.2950
10	57	0.00	0.04	0.00	0.0065	0.5069
12	29	–0.00	0.06	0.00	0.0092	0.6718
13	87	0.01	0.05	0.01	0.0052	0.0116
14	40	0.02	0.05	0.02	0.0074	0.0385
16	24	0.02	0.06	0.03	0.0096	0.0008
			Protein estimates			
3	154	0.00	0.02	0.00	0.0045	0.4560
5	104	0.01	0.03	0.01	0.0055	0.1208
6	192	0.02	0.06	0.02	0.004	5×10^{-06}
7	198	–0.01	0.08	–0.01	0.004	0.0350
8	199	–0.01	0.03	–0.01	0.0039	0.0490
9	145	0.01	0.07	0.01	0.0046	0.1401
10	58	–0.01	0.02	–0.01	0.0074	0.0590
12	29	0.04	0.04	0.03	0.0105	0.0061
13	88	–0.01	0.05	–0.01	0.006	0.0215
14	40	–0.02	0.04	–0.03	0.0086	0.0018
16	24	–0.02	0.05	–0.03	0.011	0.0113
			Lactose estimates			
3	154	0.03	0.05	0.04	0.0039	3×10^{-19}
5	104	–0.00	0.03	0.00	0.0047	0.785
6	192	0.01	0.03	0.02	0.0034	5×10^{-06}
7	198	0.00	0.04	0.01	0.0033	0.0025
8	199	0.00	0.05	0.01	0.0033	0.0037
9	145	–0.02	0.05	–0.01	0.0039	0.0158
10	58	–0.00	0.07	0.00	0.0063	0.4947
12	29	–0.03	0.02	–0.02	0.009	0.0339
13	88	–0.03	0.06	–0.02	0.0051	1×10^{-05}
14	40	–0.05	0.04	–0.04	0.0073	2×10^{-09}
16	24	–0.06	0.09	–0.04	0.0094	3×10^{-06}
			Lactose plus other solids estimates			
3	154	0.04	0.05	0.05	0.0035	2×10^{-42}
5	104	–0.02	0.04	–0.02	0.0043	9×10^{-06}
6	192	0.01	0.04	0.01	0.0031	0.0083
7	198	0.01	0.05	0.01	0.0031	0.0004
8	199	0.01	0.05	0.01	0.003	0.0014
9	145	0.00	0.04	0.01	0.0036	0.0005
10	58	0.01	0.03	0.01	0.0058	0.1337
12	29	–0.06	0.03	–0.03	0.0083	0.0001
13	88	–0.03	0.05	–0.02	0.0047	4×10^{-07}
14	40	–0.05	0.03	–0.05	0.0067	5×10^{-12}
16	24	–0.05	0.05	–0.04	0.0086	2×10^{-06}

[1]Standard error of the difference.

Changes after 10 d of storage for estimates and signals for both protein and lactose are consistent with proteolysis and fermentation, respectively. However, fat estimates increased from 10 to 16 d. Our planned comparisons did not include an evaluation of estimates versus age by regression analysis; however, there is a consistent upward trend in fat estimates (Figure 1, A). This pattern is more or less supported by increases in signals of both fat A and fat B (Figures 1, E and F, respectively), although data are incomplete because LSM values could not be computed. This effect cannot be explained by lipolysis

TABLE 4. Effect of sample age on mean instrument signal differences (diff.) (d N – d 2) and least squares mean (LSM) differences.

Sample age	N	Mean diff.	SED[1]	LSM	SE LSM	*t* test *P* > *t*
			Fat A signals			
3	147	0.01	0.05	0.00	0.0058	0.8244
5	108	–0.01	0.05	–0.00	0.0073	0.8370
6	184	–0.00	0.05	–0.00	0.0052	0.6541
7	207	–0.00	0.07	–0.01	0.0048	0.0734
8	204	0.00	0.08	–0.01	0.0050	0.2722
9	164	–0.00	0.05	–0.00	0.0056	0.6390
10	59	0.06	0.04	0.03	0.0096	0.0071
12	29	–0.02	0.10	NE[2]		
13	90	0.04	0.08	0.01	0.0083	0.4656
14	43	0.02	0.06	NE		
16	28	0.05	0.13	NE		
			Fat B signals			
3	147	–0.01	0.07	–0.01	0.0058	0.0380
5	108	0.00	0.08	0.00	0.0074	0.6120
6	184	–0.00	0.06	–0.00	0.0052	0.8678
7	207	0.00	0.08	0.00	0.0048	0.5752
8	204	0.00	0.07	0.00	0.0050	0.7681
9	164	–0.00	0.06	–0.00	0.0056	0.3713
10	59	0.01	0.05	0.00	0.0096	0.8652
12	29	0.02	0.06	NE		
13	90	0.01	0.06	0.01	0.0083	0.2004
14	43	0.00	0.07	NE		
16	28	–0.02	0.10	NE		
			Protein signals			
3	147	–0.01	0.04	–0.01	0.0057	0.1598
5	108	0.01	0.04	0.02	0.0066	0.0114
6	184	–0.00	0.04	0.01	0.0051	0.3165
7	207	–0.00	0.06	–0.00	0.0047	0.5167
8	204	–0.02	0.06	–0.02	0.0048	0.0003
9	164	–0.04	0.11	–0.04	0.0053	0.0000
10	59	–0.00	0.07	0.00	0.0090	0.9508
12	29	–0.01	0.08	–0.01	0.0130	0.4333
13	90	–0.01	0.08	–0.02	0.0073	0.0179
14	43	–0.01	0.07	–0.04	0.0102	0.0005
16	28	–0.04	0.12	–0.06	0.0126	0.0000
			Lactose signals			
3	147	0.04	0.05	0.04	0.0036	0.0000
5	108	0.00	0.04	0.01	0.0041	0.0038
6	184	0.02	0.04	0.02	0.0031	0.0000
7	207	0.03	0.05	0.03	0.0029	0.0000
8	204	0.03	0.04	0.03	0.0029	0.0000
9	164	0.02	0.03	0.03	0.0033	0.0000
10	59	0.02	0.05	0.02	0.0057	0.0003
12	29	–0.04	0.03	–0.01	0.0081	0.1550
13	90	–0.01	0.05	–0.00	0.0045	0.8939
14	43	–0.03	0.04	–0.01	0.0064	0.0698
16	28	–0.04	0.08	–0.02	0.0079	0.0155

[1]Standard error of the difference.

[2]Not estimated.

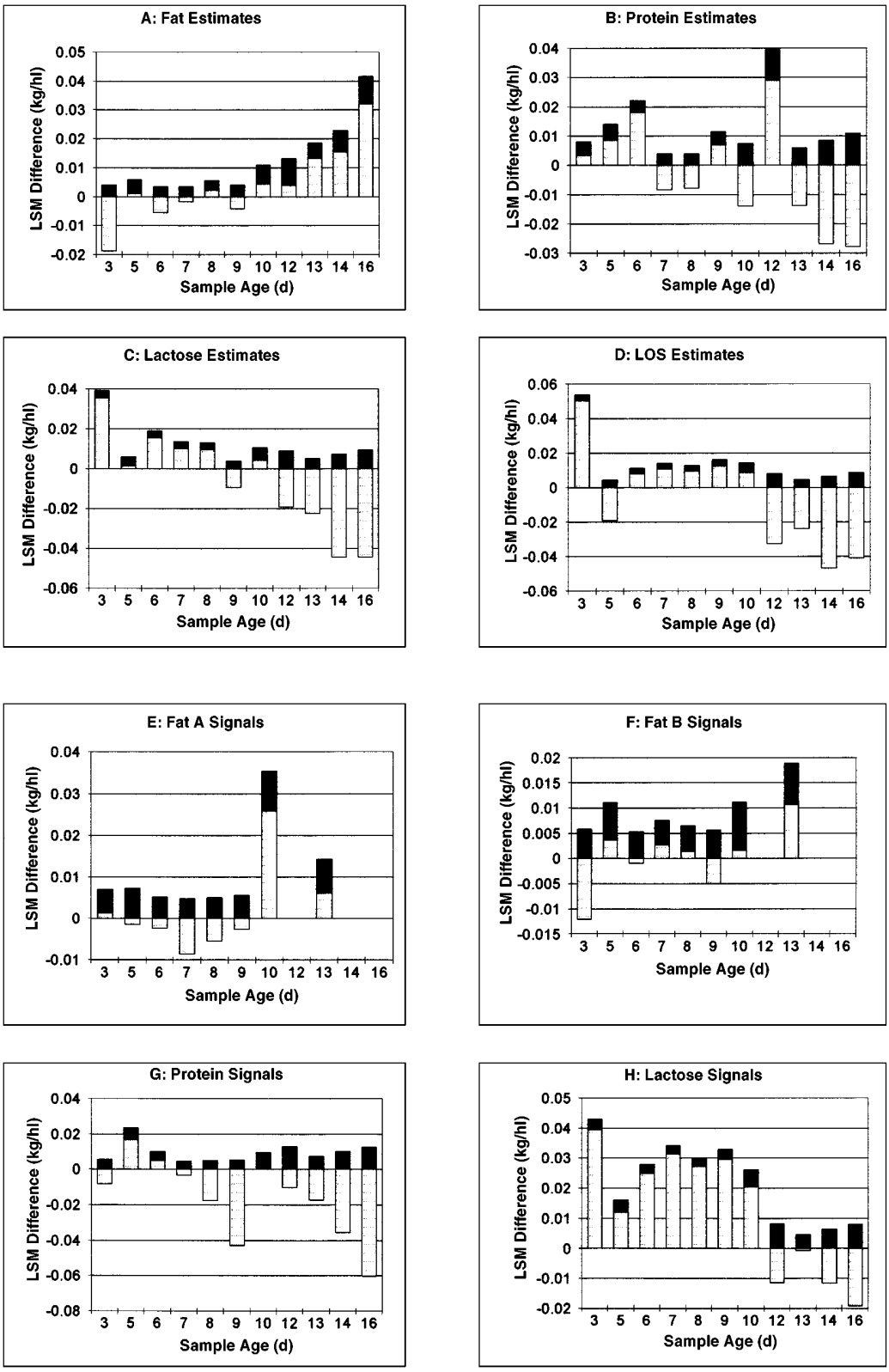

Figure 1. Least squares mean (LSM) differences (□) and standard errors (■) by sample age for estimates of fat (A), protein (B), lactose (C), lactose plus other solids (LOS) (D), fat A signals (5.7 μm) (E), fat B signals (3.5 μm) (F), protein signals (6.5 μm) (G), and lactose signals (9.6 μm) (H).

TABLE 5. Summary of significant effects ($P < 0.05$) of milk sample age on component estimates.[1]

Significant differences	Test days										
	3	5	6	7	8	9	10	12	13	14	16
Fat	—	.	.	.	.	.	.	.	+	+	+
Protein	.	.	+	—	—	.	.	+	–	–	–
Lactose	+	.	+	+	+	–	.	–	–	–	–
LOS[2]	+	–	+	+	+	+	.	–	–	–	–

[1]Negative (–), positive (+), or no effect (.).

[2]Lactose plus other solids.

TABLE 6. Summary of significant effects of milk sample age on instrument signals.[1]

Significant differences	Sample age										
	3 d	5 d	6 d	7 d	8 d	9 d	10 d	12 d	13 d	14 d	16 d
Fat A	.	.	.	.	.	.	+	?	.	?	?
Fat B	–	.	.	.	.	.	.	?	.	?	?
Protein	.	+	.	.	–	–	.	.	–	–	–
Lactose	+	+	+	+	+	+	+	.	.	–	–

[1]Negative (–), positive (+), no effect (.), or not estimated (?).

because a release of free fatty acids would decrease fat A signals and would have no effect on fat B signals.

Age and block were the only variables that had a consistent and highly significant effect on component differences. For all components except protein, the block effect was stronger than the age effects. Instrument effect was marginally significant for protein differences, and fat content was marginally significant for differences in fat, lactose, and LOS. In general, significant age effects (i.e., significant daily mean differences) ranged from ± 0.01 to 0.04 kg/hl. These differences, although statistically significant, were small and similar in magnitude to the error associated with calibration of the instruments. Accumulated mean differences for chemical reference samples, minus corresponding estimates during the period of this experiment, were 0.03 to 0.04, 0.048 to 0.058 and 0.059 to 0.074, respectively.

CONCLUSIONS

Statistical analysis indicates variable effects of age on instrumental estimates of milk composition. Effects on individual components and signals are complicated by interactions among components, unequal sample sizes, and unequal variance among responses. However, by our interpretation, the data support the following conclusions:

1. The effect of sample age up to 10 d is probably not significant for instrumental estimates of fat, protein, lactose, and LOS. Statistically significant differences occur at fewer than 10 d, but no trend is evident, and the magnitude of differences is similar to effects of instrument and block and is within the range of calibration error.
2. The accuracy of estimation of all components is compromised for milk samples that are more than 10 d old, but a considerable margin of safety is provided over the target age of 6 d.
3. Fat estimates show no consistent age effect up to d 12 when fat estimates begin to increase.
4. Protein estimates show no consistent age effect up to d 12 when estimates begin to decrease.
5. Neither LOS nor lactose estimates show a consistent age effect up to d 10 when estimates begin to decrease.
6. Except for protein, all component estimates for d 3 were significantly different than for d 2. This effect, which is apparently not an age effect, deserves further investigation.

ACKNOWLEDGMENTS

This work was supported by the Ontario Dairy Council, the Dairy Farmers of Ontario, and the Ontario Ministry of Agriculture and Food and Rural Affairs. Thanks are expressed to W. Mathes-Sears, Department of Mathematics and Statistics, University of Guelph, for assistance with statistical analysis.

REFERENCES

1 Biggs, D. A., L. F. Szijarto, and F. R. van de Voort. 1984. Fresh milk sampling for centralized testing. J. Dairy Sci. 67: 3085–3092.

2 Biggs, D. A., L. F. Szijarto, and F. R. van de Voort. 1984. The Biggs' study on fresh milk sampling. Ontario Agriculture and Food Laboratory Services Centre, Guelph, ON, Canada.
3 Hill, A. R., M. J. Makarchuck, and L. F. Szijarto. 1991. Comparison of alternative calibration procedures for infra-red milk analyzers. Food Res. Int. 24:228–232.
4 International Dairy Federation. 1985. Milk and milk products: methods of sampling. IDF Stand. 50B. Int. Dairy Fed., Brussels, Belgium.
5 Johnston, K. A. 1983. A comparison of daily and composite analysis for fat, protein and lactose in raw milk. N.Z. J. Dairy Sci. Technol. 18:79–82.
6 O'Conner, F., and J. Conolly. 1981. Random or composite milk sampling for estimating fat content of milk supplies. Farm Food Res. 12(4):108–109.
7 Pelsue, N. H., Jr. 1973. Minimum number of fresh milk samples: test of the theory. J. Dairy Sci. 56:968–972.
8 SAS® for Personal Computers, Version 6. 1989. SAS Inst., Cary, NC.
9 Szijarto, L. F., F. Harding, A. R. Hill, and J. Melichercik, 1990. Cooling systems for transport of unpreserved milk samples. J. Dairy Sci. 73:2299–2308.
10 Williams, C. J., and R. G. Peterson. 1978. Variation in estimates of milk fat, protein, and lactose content associated with various bulk milk sampling programs. J. Dairy Sci. 61: 1093–1102.