

Screening for ketosis using multiple logistic regression based on milk yield and composition

Mitsunori KAYANO¹* and Tomoko KATAOKA²

¹Department of Animal and Food Hygiene, Obihiro University of Agriculture and Veterinary Medicine, Inada-Cho, Obihiro, Hokkaido 080-8555, Japan

²Faculty of Animal Husbandry, Obihiro University of Agriculture and Veterinary Medicine, Inada-Cho, Obihiro, Hokkaido 080-8555, Japan

(Received 27 December 2014/Accepted 2 June 2015/Published online in J-STAGE 14 June 2015)

ABSTRACT. Multiple logistic regression was applied to milk yield and composition data for 632 records of healthy cows and 61 records of ketotic cows in Hokkaido, Japan. The purpose was to diagnose ketosis based on milk yield and composition, simultaneously. The cows were divided into two groups: (1) multiparous, including 314 healthy cows and 45 ketotic cows and (2) primiparous, including 318 healthy cows and 16 ketotic cows, since nutritional status, milk yield and composition are affected by parity. Multiple logistic regression was applied to these groups separately. For multiparous cows, milk yield (kg/day/cow) and protein-to-fat (P/F) ratio in milk were significant factors ($P < 0.05$) for the diagnosis of ketosis. For primiparous cows, lactose content (%), solid not fat (SNF) content (%) and milk urea nitrogen (MUN) content (mg/dl) were significantly associated with ketosis ($P < 0.01$). A diagnostic rule was constructed for each group of cows: (1) $9.978 \times \text{P/F ratio} + 0.085 \times \text{milk yield} < 10$ and (2) $2.327 \times \text{SNF} - 2.703 \times \text{lactose} + 0.225 \times \text{MUN} < 10$. The sensitivity, specificity and the area under the curve (AUC) of the diagnostic rules were (1) 0.800, 0.729 and 0.811; (2) 0.813, 0.730 and 0.787, respectively. The P/F ratio, which is a widely used measure of ketosis, provided the sensitivity, specificity and AUC values of (1) 0.711, 0.726 and 0.781; and (2) 0.678, 0.767 and 0.738, respectively.

KEY WORDS: herd test, metabolic disorder, perinatal disease, P/F ratio, regression analysis

doi: 10.1292/jvms.14-0691; *J. Vet. Med. Sci.* 77(11): 1473–1478, 2015

Ketosis is a metabolic disorder of postpartum cows. The clinical signs include poor feeding, decreased milk production, weight loss, hypoglycemia and hyperketonemia [1]. The level of ketone bodies (acetoacetic acid, β -hydroxybutyric (BHB) acid and acetone) in the blood, urine and milk of cows is used as a measure of ketosis. Even if there are no clinical signs of ketosis, a high concentration of ketone bodies (particularly BHB in the blood) can indicate the early stages of metabolic and infectious disorders, such as metritis, mastitis and displaced abomasum (subclinical ketosis [3, 9, 10]). To make early and comprehensive screening of ketotic cows possible, the development of rapid and accurate diagnostic methods is necessary. Such screening is required particularly for modern high-yielding dairy cows, which have a high risk of metabolic disorders, such as ketosis.

Milk yield and composition typically reflect the nutritional status and condition of dairy cows [3, 6, 7, 10, 13, 14]. In Japan, a dairy herd performance test (herd test) is conducted monthly on approximately half of the farms to assess these factors as well as somatic cell count and to gather feeding and reproduction information (Livestock Improvement Associa-

tion of Japan). Specifically, as of October 2014, the herd test is carried out on 68% of the farms in Hokkaido. These farmers therefore have monthly information about milk yield and composition for all of the cows on their farms. This presents a cost-effective opportunity to screen for disorders in cows without necessitating further testing.

The use of milk composition for the diagnosis of conditions, such as ketosis, has been tested [2, 3, 5, 7, 8, 14]. The protein-to-fat ratio of milk (P/F ratio) is widely used to diagnose ketosis, with a cutoff value of approximately 0.70 (e.g. [8]), but this measure is not very accurate. Indeed, even direct diagnosis based on concentrations of ketone bodies in milk and urine is not always sufficiently accurate [8].

In this study, we used multiple logistic regression analysis to investigate whether data collected as part of the herd test, specifically milk yield and composition data, could be used in combination to diagnose ketosis. We then used several components of the herd test data to construct a scheme of diagnosis.

MATERIALS AND METHODS

Samples: Data were collected from June 2011 to February 2014 from 50 farms in the Kawanishi and Taisho subareas of the Tokachi area of Hokkaido, Japan. The average number of lactating cows per year was less than 50 for 19 farms, between 50 and 100 for 26 farms and between 100 and 200 for 5 farms. The herd test datasets and medical records were obtained from the Obihiro Husbandry Center (Obihiro Chikusan Center) and Tokachi Agricultural Insurance Associa-

*CORRESPONDENCE TO: KAYANO, M., Department of Animal and Food Hygiene, Obihiro University of Agriculture and Veterinary Medicine, Inada-Cho, Obihiro, Hokkaido 080-8555, Japan.
e-mail: kayano@obihiro.ac.jp

©2015 The Japanese Society of Veterinary Science

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License <<http://creativecommons.org/licenses/by-nc-nd/3.0/>>.

Table 1. The result of the *t*-tests for each group of cows, separated on the basis of parity

	P-value		
	All parity	Primiparity	Multiparity
Days in milk	0.052	0.380	0.008 **
Parity	1.93×10^{-5} ***	-	7.25×10^{-4} ***
Milk yield	0.091	0.633	3.19×10^{-4} ***
Fat	2.50×10^{-7} ***	0.196	4.38×10^{-7} ***
SNF	1.00×10^{-6} ***	2.07×10^{-3} **	8.29×10^{-4} ***
Protein	0.048 *	0.002 **	0.633
Lactose	4.56×10^{-4}	0.642	2.39×10^{-4} ***
MUN	0.007 **	0.017 *	0.124
P/F	1.03×10^{-10} ***	0.014 *	2.56×10^{-9} ***

***' 0.001, '**' 0.01, '*' 0.05.

tion (Tokachi NOSAI), respectively.

The focus of the analysis was cows (both healthy and ketotic) that had ≤ 30 days in milk (DIM), since in this dataset, 82.26% of ketosis cases occurred in these cows. After cows that had recovered from ketosis were removed from the dataset, it included 632 records for healthy cows and 61 for ketotic cows. The ketotic cows were identified by veterinarians based on clinical signs (for all cows) and the level of BHB in their milk (for approximately 70% of cows); i.e., diagnosis was based on clinical signs only for 30% of the ketotic cows and on both clinical signs and the level of BHB in the milk for the remaining 70%. We defined healthy cows as those that did not have any records of disorders during the study period. The data were analyzed as described below, using 9 variables: DIM (7–30 days), parity (1–12), milk yield (5.7–55.8 kg/day/cow) and fat content (2.51–7.37%), protein content (2.30–4.42%), solid not fat (SNF) content (7.34–9.76%), lactose content (3.85–5.00%), milk urea nitrogen (MUN) content (1.00–18.95 mg/dl) and P/F ratio (0.36–1.32).

Statistical analysis (t-test, multiple logistic regression and ROC analysis): T-tests can be used to investigate whether some components of the herd test data are significantly different between healthy and ketotic cows. In contrast, multiple logistic regression analysis simultaneously takes into account interactions between the components of the herd test data and identifies those that are significantly associated with ketosis. Using the formula for multiple logistic regression:

$$\log\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + \dots + b_px_p,$$

the probability p that a cow is ketotic is calculated by the following reformulation:

$$p = \frac{\exp(b_0 + b_1x_1 + \dots + b_qx_q)}{1 + \exp(b_0 + b_1x_1 + \dots + b_qx_q)},$$

where x_1, \dots, x_q are explanatory variables that explain the probability of ketosis (e.g. components of the herd test data, such as milk yield and composition), and b_1, \dots, b_q are regression coefficients. If the probability of ketosis thus calculated is greater than a fixed threshold (e.g. 0.5), the corresponding cows are classified as ketotic. The performance of multiple logistic regression for the diagnosis of ketosis can be evaluated by receiver operating characteristic (ROC) analysis, which plots sensitivities and specificities for many cutoff values. In this study, the inputs used to draw the ROC curves were (i) the probability of ketosis based on a logistic regression including some components of the herd test data and (ii) a binary variable indicating whether the cows were ketotic or not. In ROC analysis, the measure of diagnostic accuracy is the area under the curve (AUC), which ranges between 0 and 1. An AUC of 1 means perfect diagnosis, and an AUC of 0.5 indicates random diagnosis, implying that the components of the herd test data provide no useful information on which to base diagnosis. A large value (between 0.5 and 1) of AUC indicates an increasingly accurate diagnosis.

RESULTS

All analyses were conducted using the statistical software R (ver. 3.0.1 for WIN).

T-tests: First of all, *t*-tests were applied to the dataset containing the following herd test components: DIM, parity, milk yield (kg/day/cow) and composition (fat (%), protein (%), SNF (%), lactose (%), MUN (mg/dl) and P/F ratio) (Table 1). In addition, *t*-tests were conducted on datasets in which the cows were separated into polyparous cows (A), which included 314 healthy cows and 45 ketotic cows, and primiparous cows (B), which included 318 healthy cows and 16 ketotic cows (Table 1). These tests showed that the herd test components significantly associated with ketosis were different between primiparous and multiparous cows. Interestingly, SNF and milk protein were the two main components associated with ketosis in primiparous cows ($P < 0.001$). Boxplots of those components are shown in Fig. 1.

Logistic regression for multiparous cows: Multiple logistic regression was carried out on the dataset containing mul-

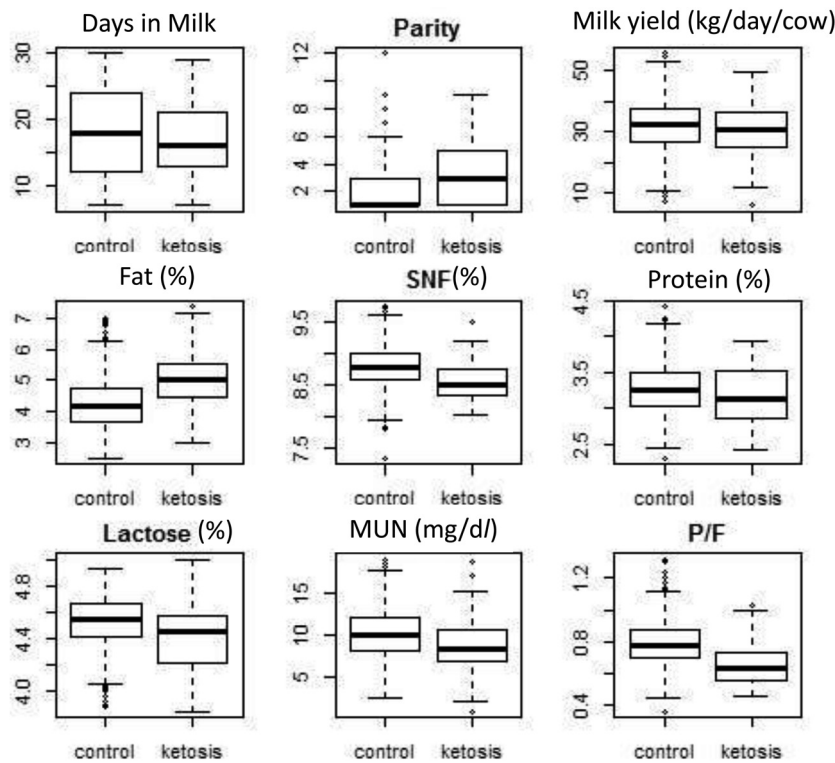


Fig. 1. Box plots of the herd test variables that were used in analysis of records for 632 healthy (control) and 61 ketotic cows. SNF represents solid not fat, MUN represents milk urea nitrogen, and P/F is the protein-to-fat ratio.

tiparous cows with DIM, parity, milk yield (kg/day/cow) and composition (SNF (%), lactose (%), MUN (mg/dl) and P/F ratio) as explanatory variables. Multiple logistic regression analysis showed that P/F ratio and milk yield were significantly associated with ketosis ($P < 0.05$; Table 2). The AUC value was 0.811, with a sensitivity of 80.0% and specificity of 72.9% (Table 3, Fig. 2). The diagnostic rule for ketotic multiparous cows was given by

$$9.978 \times \text{P/F ratio} + 0.085 \times \text{milk yield (kg/day/cow)} < 10. \dots(1)$$

This diagnosis rule had a sensitivity of 80.0% and specificity of 72.9% based on the dataset used in this study (Fig. 3). When only the P/F ratio was used, with a threshold of 0.70, the AUC value was 0.781, with a sensitivity of 71.1% and specificity of 72.6% (Table 3 and Fig. 2). Even if all seven variables were included, AUC, sensitivity and specificity were 0.852, 0.778 and 0.803, respectively (Fig. 2). Milk yield improved the accuracy of diagnosis of ketotic multiparous cows.

Logistic regression for primiparous cows: Multiple logistic regression was applied to the dataset containing primiparous cows with DIM, milk yield (kg/day/cow) and composition (SNF (%), lactose (%), MUN (mg/dl) and P/F ratio) as explanatory variables. Three components, SNF, lactose and MUN, were significantly associated with ketosis

($P < 0.01$, Table 2). The AUC value was 0.787, with a sensitivity of 81.3% and specificity of 73.0% (Table 3, Fig. 4). The equation for diagnosis of ketosis for primiparous cows was given by

$$2.327 \times \text{SNF (%) - 2.703} \times \text{lactose (%) + 0.225} \times \text{MUN (mg/dl)} < 10. \dots(2)$$

This diagnosis rule had a sensitivity of 81.3% and specificity of 73.0% based on the dataset used in this study (Fig. 5). When only the P/F ratio was used, with a threshold of 0.70, the AUC value was 0.738, with a sensitivity of 76.7% and specificity of 68.7% (Table 3 and Fig. 4). Even if all 6 variables were included, AUC, sensitivity and specificity were 0.812, 0.688 and 0.906, respectively (Fig. 4). Three milk composition factors (SNF, lactose and MUN) also provided a high accuracy of diagnosis of ketosis for primiparous cows.

DISCUSSION

An important clinical sign for diagnosing ketotic cows is hyperketonemia, in which the concentration of ketone bodies (acetoacetic acid, BHB acid and acetone) in the cow's blood increases. Ketone bodies are also excreted into the urine and milk of ketotic cows. Ketosis can therefore be diagnosed by measuring the concentrations of ketone bodies in the cow's blood, urine and milk. Krogh *et al.* [8] reported

Table 2. Results of the multiple logistic regression analysis of two datasets, one containing multiparous cows and the other containing primiparous cows

Multiparity			
Variable	Estimate	Std.Error	P-value
(Intercept)	6.654		
Milk yield	-0.073	0.023	0.0012 **
P/F ratio	-8.573	1.627	1.37×10^{-7} ***
Primiparity			
Variable	Estimate	Std.Error	P-value
(Intercept)	10.027		
Lactose	3.525	1.743	0.0431 *
SNF	-3.034	0.996	0.0023 **
MUN	-0.293	0.107	0.0061 **

**** 0.001, *** 0.01, ** 0.05.

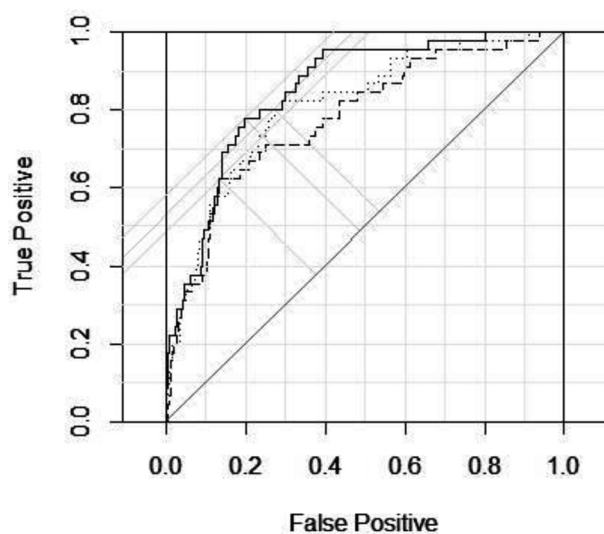


Fig. 2. ROC curve based on a logistic regression for multiparous cows developed using records for 314 healthy and 45 ketotic cows. The inputs used to draw the ROC curves were (i) the probability of ketosis based on a logistic regression including some components of the herd test data and (ii) a binary variable indicating whether the cows were ketotic or not. Solid line (AUC=0.852): 7 components of the herd test data, namely DIM, parity, milk yield (kg/day/cow), SNF (%), lactose (%), MUN (mg/dl) and P/F ratio; dotted line (AUC=0.811): milk yield + P/F ratio; broken line (AUC=0.781): P/F ratio.

the sensitivity (Sen) and specificity (Spe) of diagnosis based on BHB in milk, acetoacetic acid in urine and F/P ratio (as opposed to P/F ratio): Sen=0.78 (95% Bayesian confidence interval of 0.55–0.98) and Spe=0.99 (0.97–0.99) for BHB in milk; Sen=0.58 (0.39–0.93) and Spe=0.99 (0.97–0.99) for acetoacetic acid in urine; and Sen=0.63 (0.58–0.71) and 0.79 (0.77–0.81) for F/P ratio. In contrast, our diagnostic rule, based on a multiple logistic regression including variables describing milk yield and composition (P/F ratio and milk yield for multiparous cows, and SNF, lactose and MUN for primiparous cows), had Sen \geq 0.80 and Spe \approx 0.73. Such a

Table 3. Performance of the diagnosis of ketosis through a few components

Multiparity			
Components	AUC	Sensitivity	Specificity
Milk yield +P/F ratio	0.811	0.800	0.729
P/F ratio (0.7)	0.781	0.711	0.726
Primiparity			
Components	AUC	Sensitivity	Specificity
SNF + Lactose +MUN	0.787	0.813	0.730
P/F ratio (0.7)	0.738	0.687	0.767

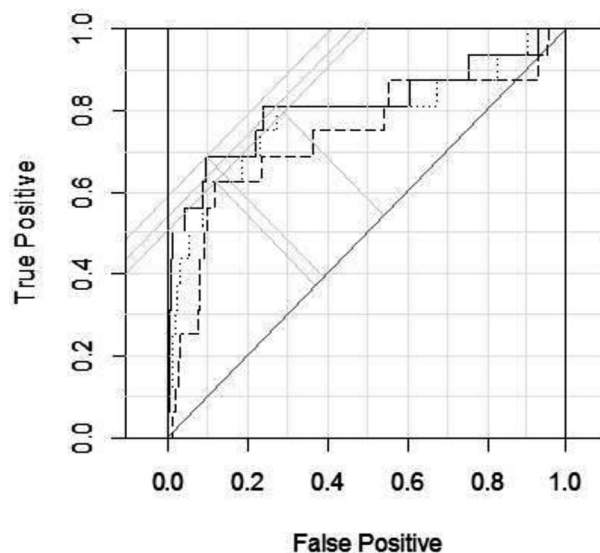


Fig. 4. ROC curve based on a logistic regression for primiparous cows developed using records for 318 healthy and 16 ketotic cows. The inputs used to draw the ROC curves were (i) the probability of ketosis based on a logistic regression including some components of the herd test data and (ii) a binary variable indicating whether the cows were ketotic or not. Solid line (AUC=0.812): 6 components of the herd test data, namely DIM, milk yield (kg/day/cow), SNF (%), lactose (%), MUN (mg/dl) and P/F ratio; dotted line (AUC=0.787): SNF + lactose + MUN; broken line (AUC=0.738): P/F ratio.

diagnosis and screening scheme is comparatively cost-effective and straightforward, since it can utilize the information collected in herd tests that are routinely conducted on a monthly basis to check milk yield and composition for all cows in a herd. Although milk yield can also decrease as a result of many other causes, such as puerperal metritis, lymphoma and abomasal displacement, ketotic cows can be reliably diagnosed based on both milk yield and composition.

Our screening rule for multiparous cows is given by equa-

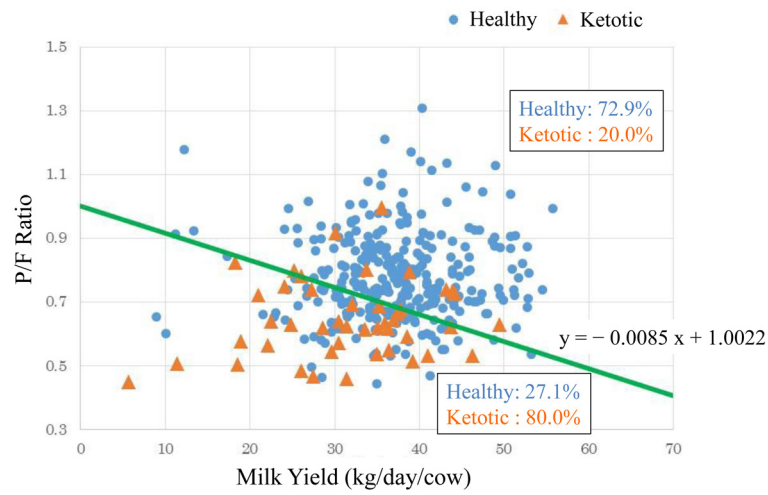


Fig. 3. Scatter plot of milk yield (kg/day/cow) and P/F ratio for multiparous cows. The line indicates the diagnosis boundary for ketotic cows: $9.978 \times \text{P/F ratio} + 0.085 \times \text{milk yield (kg/day/cow)} = 10$. Red triangles and blue circles represent ketotic (n=45) and healthy (n=314) cows, respectively.

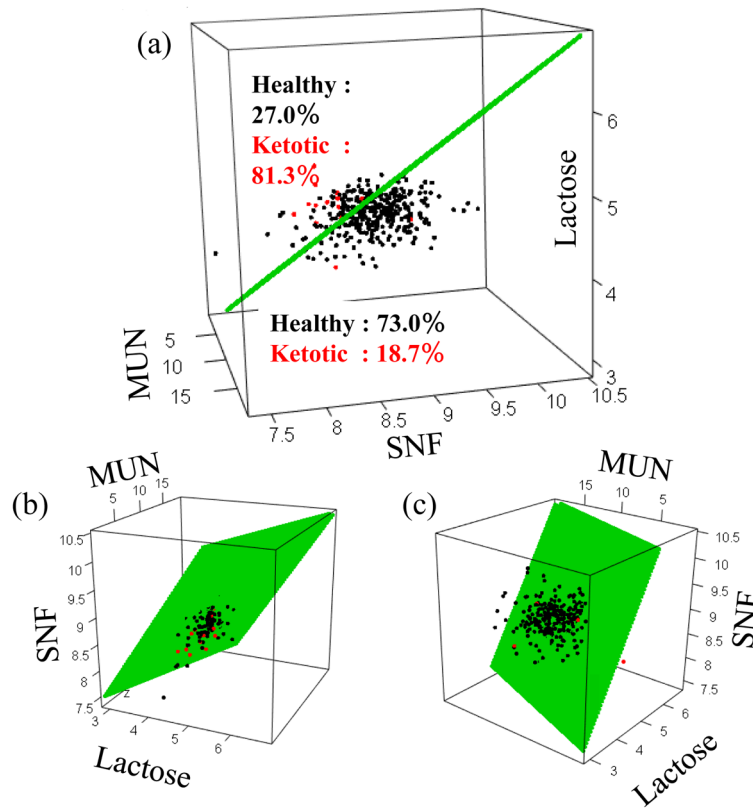


Fig. 5. Three-dimensional plots of SNF (%), lactose (%) and MUN (mg/d) for primiparous cows, viewed from three different angles (a-c). The surface is the diagnosis boundary for ketotic cows: $2.327 \times \text{SNF (\%)} - 2.703 \times \text{lactose (\%)} + 0.225 \times \text{MUN (mg/d)} = 10$. Red and black points indicate ketotic (n=16) and healthy (n=318) cows, respectively.

tion (1), which is based on milk yield and P/F ratio. Although decreased milk production is a sign of ketosis, in the current dataset, milk yield was affected significantly only in multiparous cows and hence appears only in that screening rule. Nevertheless, this variable, together with P/F ratio, was found to be highly valuable for screening of ketotic multiparous cows. In addition, for multiparous cows, fat content was significantly increased in ketotic cows' milk ($P=4.4 \times 10^{-7}$). This implies a drastic decrease in the P/F ratio for ketotic cows ($P=2.6 \times 10^{-9}$), even though the protein content was not significantly affected. A possible reason for the increased fat content of ketotic multiparous cows' milk is active fat mobilization.

For primiparous ketotic cows, the screening rule is given by equation (2), which is based on the SNF, lactose and MUN content of the milk. However, there were relatively few primiparous ketotic cows ($n=16$) upon which to base this rule. The P/F ratio does not appear in this screening rule, because other composition factors, such as SNF ($P=0.002$), were more significantly related to ketosis than the P/F ratio ($P=0.014$). The small difference in P/F ratio between healthy and ketotic cows was the result of a relatively small difference in protein content ($P=0.002$) and no difference in fat content ($P>0.05$). However, protein appeared in the screening rule through SNF content, since SNF consists of protein (37.4%), lactose (54.9%) and ash (7.7%). Replacing the P/F ratio and SNF content with protein and fat content in the multiple logistic regression resulted in protein and MUN content becoming significant ($P<0.05$). A possible reason for the decreased protein in the milk produced by ketotic primiparous cows ($P=0.002$) is decreased microbial synthesis, while a negative energy balance is present. The decrease in MUN content may be the result of poor feeding by ketotic cows. Further analyses based on records for many more primiparous ketotic cows are required to strengthen and clarify these results.

The screening rules developed in this study facilitate detection of subclinical ketosis. However, subclinical ketosis cases tend to show lower blood BHB concentrations than do clinical cases (e.g. [4, 11, 12]). Weak subclinical ketosis cases with low BHB levels might pass our screening rules as healthy cows. In fact, subclinical ketosis is defined by a high BHB level (>1.2 – 1.4 mmol/l), and clear clinical signs tend to appear only at BHB levels of >3.0 mmol/l (e.g. [4, 11]). Sun *et al.* [12] reported mean \pm SD BHB levels of 2.49 ± 0.60 , 1.22 ± 0.17 and 0.82 ± 0.12 mmol/l for 24 clinical, 33 subclinical and 24 healthy cows in China, respectively. A difference in BHB levels between subclinical and clinical ketosis might affect the resulting screening rule. Further analyses of subclinical ketosis cases with data on BHB levels are thus required to construct highly accurate screening rules for subclinical cases.

Screening for metabolic disorders, such as clinical and subclinical ketosis, in cows using herd test results, including information on milk yield and composition, is comparatively effective. The further development of such approaches should proceed by utilizing comprehensive datasets from individual farm.

ACKNOWLEDGMENTS. The datasets from the herd tests and medical records were obtained from the Obihiro Husbandry Center (Obihiro Chikusan Center) and Tokachi Agricultural Insurance Association (Tokachi NOSAI), respectively. This study was supported in part by Japanese Society for the Promotion of Science (JSPS) KAKENHI Grant 24700290.

REFERENCES

- Andrews, A. H. 2000. *The Health of Dairy Cattle*. Blackwell Science Ltd., Oxford.
- Duffield, T. F., Kelton, D. F., Leslie, K. E., Lissemore, K. E. and Lumsden, J. H. 1997. Use of test day milk fat and milk protein to detect subclinical ketosis in dairy cattle in Ontario. *Can. Vet. J.* **38**: 713–718. [Medline]
- Duffield, T. F., Lissemore, K. D., McBride, B. W. and Leslie, K. E. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. *J. Dairy Sci.* **92**: 571–580. [Medline] [CrossRef]
- Ehret, A., Hochstuhl, D., Krattenmacher, N., Tetens, J., Klein, M. S., Gronwald, W. and Thaller, G. 2015. Use of genomic and metabolic information as well as milk performance records for prediction of subclinical ketosis risk via artificial neural networks. *J. Dairy Sci.* **98**: 322–329. [Medline] [CrossRef]
- Geishauser, T. D., Leslie, K. E., Duffield, T. F. B. and Edge, V. L. 1998. An evaluation of protein/fat ratio in first DHI test milk for prediction of subsequent displaced abomasum in dairy cows. *Can. J. Vet. Res.* **62**: 144–147. [Medline]
- Hamann, J. and Kromker, V. 1997. Potential of specific milk composition variables for cow health management. *Livest. Prod. Sci.* **48**: 201–208. [CrossRef]
- Heuer, C., Schukken, Y. H. and Dobbelaar, P. 1999. Postpartum body condition score and results from the first test day milk as predictors of disease, fertility, yield, and culling in commercial dairy herds. *J. Dairy Sci.* **82**: 295–304. [Medline] [CrossRef]
- Krogh, M. A., Toft, N. and Enevoldsen, C. 2011. Latent class evaluation of a milk test, a urine test, and the fat-to-protein percentage ratio in milk to diagnose ketosis in dairy cows. *J. Dairy Sci.* **94**: 2360–2367. [Medline] [CrossRef]
- Ospina, P. A., Nydam, D. V., Stokol, T. and 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**: 3595–3601. [Medline] [CrossRef]
- Rajala-Schultz, P. J., Grohn, Y. T. and McCulloch, C. E. 1999. Effects of milk fever, ketosis, and lameness on milk yield in dairy cows. *J. Dairy Sci.* **82**: 288–294. [Medline] [CrossRef]
- Suthar, V. S., Canelas-Raposo, J., Deniz, A. and Heuwieser, W. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. *J. Dairy Sci.* **96**: 2925–2938. [Medline] [CrossRef]
- Sun, L. W., Zhang, H. Y., Wu, L., Shu, S., Xia, C., Xu, C. and Zheng, J. S. 2014. ¹H-Nuclear magnetic resonance-based plasma metabolic profiling of dairy cows with clinical and subclinical ketosis. *J. Dairy Sci.* **97**: 1552–1562. [Medline] [CrossRef]
- Sutton, J. D. 1989. Altering milk composition by feeding. *J. Dairy Sci.* **72**: 2801–2814. [CrossRef]
- Toni, F., Vincenti, L., Grigoletto, L., Ricci, A. and Schukken, Y. H. 2011. Early lactation ration of fat and protein percentage in milk is associated with health, milk production, and survival. *J. Dairy Sci.* **94**: 1772–1783. [Medline] [CrossRef]