

# Facility and management characteristics of large Upper Midwest dairy herds clustered by Dairy Herd Improvement records

**R.L. Brotzman, DVM; N.B. Cook, BSc BVSc Cert. CHP DBR Dip. ECBHM MRCVS; M.R. Foy; J.P. Hess; K.V. Nordlund, DVM; T.B. Bennett; A. Gomez, Lic. Vet., MSc; D.D. Döpfer, Dr. med. vet. DD, MSc, PhD, Dipl. ECBHM**  
*University of Wisconsin-Madison School of Veterinary Medicine, Madison, WI 53706*

## Introduction

Principal component analysis (PCA) is used on datasets with vast numbers of numeric variables and limited observations (e.g., dairy herd improvement [DHI] data) for unbiased selection of uncorrelated variables that describe the largest amount of variance. Cluster analysis (CA) divides objects of interest (e.g., dairy herds) into groups on the basis of similarity in multiple characteristics simultaneously. The aims of this project were to develop a novel method for discovering important DHI variables by use of PCA and then grouping herds by those variables via CA, and to survey herds to determine herd management characteristics of each group.

## Materials and Methods

Complete year 2011 DHI data was obtained for 557 Upper Midwest dairy herds with a test-day mean herd size of  $\geq 200$  cows. Twenty-two DHI variables were selected and reduced to 16 through PCA. Farms were divided into 6 groups by CA on the basis of PCA-selected variables. All herds were contacted by phone to survey management and housing practices. (201 completed surveys [36.1%]).

## Results

Of the 557 herds evaluated, 201 (36.1%) herd managers agreed to complete the management survey. Survey results indicated trends in herd management consistent with mean DHI performance of each group. Select responses that highlighted management practices were: mean herd size (COWS); mean energy-corrected milk (ECM); mean linear somatic cell score (LSCS); mean proportion of herd with new subclinical mastitis infection at test (NINF); mean dry-cow subclinical mastitis cure rate (DRYCR); mean transition cow index (TCI); mean death rate (DIED); mean proportion of herd as primiparous cows (L1); proportion of herds that used recombinant bovine somatotropin (BST); proportion of herds that used monensin sodium (MON); mean milking frequency (MF); proportion of herds that used a reproductive synchronization program for the first breeding in cows (SYNCH); proportion of herds that used sand

bedding for mature lactating cows (SAND); and mean cows per 50-hr full-time equivalent employee (CFTE).

Group 1 herds were characterized as 'high input–high health–high production' (mean  $\pm$  SD or proportion of group; COWS,  $581 \pm 377$  cows; ECM,  $91.7 \pm 7.85$  lb ( $41.7 \pm 3.57$  kg); LSCS,  $2.2 \pm 0.31$ ; NINF,  $9 \pm 1.9\%$ ; DRYCR,  $75 \pm 10.0\%$ ; TCI,  $457.6 \pm 772.2$  lb ( $208 \pm 351$  kg); DIED,  $5.7 \pm 2.6\%$ ; L1,  $38.4 \pm 4.93\%$ ; BST,  $73.0\%$ ; MON,  $88.9\%$ ; MF,  $3.0 \pm 0.1$ ; SYNCH,  $87.3\%$ ; SAND,  $71.4\%$ ; CFTE,  $51.92 \pm 15.82$  cows/FTE).

Group 2 herds had somewhat similar DHI parameters as group 1 herds, but managed cows differently and were characterized as 'lower input–high health–high production' (COWS,  $300 \pm 186$  cows; ECM,  $86.6 \pm 8.3$  lb [ $39.4 \pm 3.76$  kg]; LSCS,  $2.3 \pm 0.31$ ; NINF,  $9 \pm 2.3\%$ ; DRYCR,  $66 \pm 11.6\%$ ; TCI,  $519.2 \pm 664.4$  lb [ $236 \pm 302$  kg]; DIED,  $5.7 \pm 2.7\%$ ; L1,  $38.1 \pm 5.94\%$ ; BST,  $33.3\%$ ; MON,  $75.8\%$ ; MF,  $2.0 \pm 0.0$ ; SYNCH,  $60.6\%$ ; SAND,  $63.6\%$ ; CFTE,  $58.95 \pm 16.03$  cows/FTE).

Group 3 herd data did not reveal any clear pattern for characterization (COWS,  $370 \pm 143$  cows; ECM,  $88 \pm 8.16$  lb [ $40.0 \pm 3.71$  kg]; LSCS,  $2.6 \pm 0.32$ ; NINF,  $12 \pm 2.7\%$ ; DRYCR,  $64 \pm 9.7\%$ ; TCI,  $-24.2 \pm 805.2$  lb [ $-11 \pm 366$  kg]; DIED,  $6.3 \pm 3.2\%$ ; L1,  $38.6 \pm 5.72\%$ ; BST,  $70.4\%$ ; MON,  $77.8\%$ ; MF,  $2.9 \pm 0.3$ ; SYNCH,  $59.3\%$ ; SAND,  $63.0\%$ ; CFTE,  $53.09 \pm 22.99$  cows/FTE).

Group 4 herds were characterized as 'low input–low production–low milk quality' (COWS,  $294 \pm 104$  cows; ECM,  $74.58 \pm 9.98$  lb [ $33.9 \pm 4.54$  kg]; LSCS,  $3.0 \pm 0.39$ ; NINF,  $15 \pm 4.0\%$ ; DRYCR,  $56 \pm 14.1\%$ ; TCI,  $-378.4 \pm 721.6$  lb [ $-172 \pm 328$  kg]; DIED,  $4.9 \pm 2.6\%$ ; L1,  $38.0 \pm 7.89\%$ ; BST,  $26.1\%$ ; MON,  $60.9\%$ ; MF,  $2.2 \pm 0.4$ ; SYNCH,  $39.1\%$ ; SAND,  $69.6\%$ ; CFTE,  $58.61 \pm 16.86$  cows/FTE).

Group 5 herds were characterized as 'high death loss–low milk quality–low production' (COWS,  $480 \pm 330$  cows; ECM,  $81.18 \pm 8.8$  lb [ $36.9 \pm 4.02$  kg]; LSCS,  $2.8 \pm 0.41$ ; NINF,  $14 \pm 4.2\%$ ; DRYCR,  $64 \pm 10.0\%$ ; TCI,  $-468.6 \pm 574.2$  lb [ $-213 \pm 261$  kg]; DIED,  $12.4 \pm 3.6\%$ ; L1,  $43.8 \pm 5.11\%$ ; BST,  $60.9\%$ ; MON,  $91.3\%$ ; MF,  $2.8 \pm 0.4$ ; SYNCH,  $69.6\%$ ; SAND,  $52.2\%$ ; CFTE,  $51.42 \pm 14.45$  cows/FTE).

Group 6 herds were characterized as 'larger, technology-adopting' (COWS,  $1,478 \pm 1,356$  cows; ECM,  $88.4 \pm 6.5$  lb [ $40.2 \pm 2.96$  kg]; LSCS,  $2.7 \pm 0.34$ ; NINF,  $13$

$\pm 2.7\%$ ; DRYCR,  $71 \pm 8.4\%$ ; TCI,  $-30.8 \pm 563.2$  lb [ $-14 \pm 256$  kg]; DIED,  $7.6 \pm 2.6\%$ ; L1,  $43.8 \pm 5.11\%$ ; BST,  $84.4\%$ ; MON,  $84.4\%$ ; MF,  $3.0 \pm 0.1$ ; SYNCH,  $93.8\%$ ; SAND,  $68.8\%$ ; CFTE,  $66.57 \pm 19.40$  cows/FTE).

Further statistical investigation into the potential relationship between management practices and CA results will be conducted.

### Significance

PCA and CA were useful for selection of variables and the meaningful grouping of herds into several plausible patterns of management by the simultaneous evaluation of multiple DHI variables.