J. Dairy Sci. 105:4565–4580
https://doi.org/10.3168/jds.2021-21518

Phenotyping metabolic status of dairy cows using clustering of time profiles of energy balance peripartum

F. Vossebeld,[1,2] A. T. M. van Knegsel,[2] and E. Saccenti[1]*
[1]Laboratory of Systems and Synthetic Biology, Wageningen University and Research, Stippeneng 4, 6708 WE, Wageningen, the Netherlands
[2]Adaptation Physiology group, Department of Animal Sciences, Wageningen University and Research, PO Box 338, 6700 AH, Wageningen, the Netherlands

ABSTRACT

Due to a combination of a relatively low energy intake and a high demand of energy required for milk production, dairy cows experience a negative energy balance (EB) at the start of lactation. This energy deficit causes body weight reduction and an increased risk for metabolic diseases. Severity and length of negative EB can differ among cows. Peripartum time profiles of EB for dairy cows are not described yet in the literature. Creating EB-derived time profiles with corresponding metabolic status and disease treatments could improve understanding the relationship between EB and metabolic status, as well as enhance identification of cows at risk for compromised metabolic status. In this research we propose a novel method to cluster EB time series and examine associated metabolic status and disease treatments of dairy cows in the peripartum period. In this study, data of 3 earlier experiments were merged and examined. Four dairy cow clusters for time profiles of EB from wk −3 until +7 relative to calving were generated by the global alignment kernel algorithm. For each cluster, mean of body weight prepartum was distinguishable, indicating this might be a possible on-farm biomarker for the peripartum EB profile. Moreover, cows with severe EB drop postpartum were more treated for milk fever and had high plasma nonesterified fatty acids and β-hydroxybutyrate concentration, and low IGF-1, insulin, and glucose concentration in the first 7 wk of lactation. Overall, this study demonstrated that cows can be clustered based on EB time profiles and that characteristics such as prepartum body weight, and postpartum nonesterified fatty acids and glucose concentration are promising biomarkers to identify the time profile of EB and potentially the risk for metabolic diseases.

Key words: negative energy balance, time trends, metabolites, body condition, disease treatments

INTRODUCTION

Due to extensive energy requirement for milk production in combination with a relatively low energy intake from feed, dairy cows experience a negative energy balance (**NEB**) in the early stage of lactation. To compensate for this shortage in energy, dairy cows mobilize fats originating from body reserves (Tamminga et al., 1997). Because of this body fat mobilization, the availability of lipogenic to glucogenic compound ratio shifts to more lipogenic compounds, which is associated with an increase of plasma metabolites such as nonesterified fatty acids (**NEFA**) and BHB, which in turn could increase the risk for metabolic diseases such as fatty liver and ketosis (van Knegsel et al., 2005; Klein et al., 2010; Sun et al., 2014).

Metabolic impairment and diseases can lead to a reduced productive lifespan, compromised animal welfare, and increased veterinary costs (Seifi et al., 2011; Probo et al., 2018) and for these reasons understanding the variability of energy balance (**EB**) levels of cows in early lactation is pivotal to facilitate management of animal health and wellbeing. Traditionally, EB levels in dairy cows are measured and examined independently over the pre- and postpartum period or studied as an average per time period. In most existing studies EB is grouped based on experimental treatments, such as dietary (e.g., Odens et al., 2007; Garnsworthy et al., 2009) or management measures (e.g., Rastani et al., 2005).

To our knowledge limited studies have grouped cows based on EB level [e.g., by dividing cows in 4 quartile groups based on EB (De Vries and Veerkamp, 2000)], which can be considered a relatively arbitrary approach. It is known, however, that the severity and length of NEB period can vary among cows (Mellouk et al., 2019). Moreover, little is known about the relationship between the temporal dynamics of EB in the peripartum period and the dynamics of blood metabolites and other cow characteristics. However, it can be expected that understanding the relationship between the dynamics of EB during the peripartum period and

Received November 3, 2021.
Accepted February 3, 2022.
*Corresponding author: edoardo.saccenti@wur.nl

Table 1. Characteristics of studies that were used for this research

Item	Study Ia: van Knegsel et al. (2014)	Study Ib: Chen et al. (2015)	Study II: van Hoeij et al. (2017)	This study
Characteristic				
Number of cows	168	130	127	350
Start week (prepartum)	−8	−8	−3	−3
End week (postpartum)	14	9	44	7
Variable				
Milk	Fat, FPCM, lactose, milk yield, protein, SCC			
Plasma	BHB, glucose, IGF-1, insulin, NEFA			
Cow	BW, energy balance			
Independent	DPL, feed			

[1]FPCM = fat- and protein-corrected milk yield; NEFA = nonesterified fatty acids; DPL = dry period length.

associated metabolic status during early lactation can enhance identification of cows that are at risk for compromised metabolic status or even metabolic diseases and support customized management decisions.

The aim of this study is to cluster EB time series to stratify dairy cows into different groups characterized by EB time patterns measured over a 10-wk period before calving and during early lactation, making use of data from previously published studies (van Knegsel et al., 2014; Chen et al., 2015; van Hoeij et al., 2017). In addition, the EB time series clusters were related to plasma metabolites, parity, dry period length, BW, and disease treatments in the same period.

MATERIALS AND METHODS

Experimental Data

This study uses data collected during 3 previously published studies: study Ia (van Knegsel et al., 2014), study Ib (Chen et al., 2015), and study II (van Hoeij et al., 2017). The experimental protocols of all 3 studies were approved by the Institutional Animal Care and Use Committee of Wageningen University & Research (the Netherlands) and comply with the Dutch law on Animal Experimentation (study Ia and Ib: protocol number 2010026; study II: protocol number 2014125). A summary of the characteristics of these 3 studies is presented in Table 1. In brief, all 3 studies included Holstein Friesian dairy cows at the Dairy Campus research herd (WUR Livestock Research, Lelystad, the Netherlands). In study Ia, 168 cows were assigned randomly to 3 groups with a 0-d, 30-d, or 60-d dry period length (**DPL**). After lactation of study Ia, 130 of the 168 cows were examined during the next lactation in study Ib. In study Ib, 19 cows originally assigned to 0-d DPL were attributed to 0→67 d DPL (actual days dry: 67 ± 8 d). Cows were allocated to this new group and dried off when they had a milk yield of <4 kg/d at least 30 d before the expected calving date (Chen et al., 2015). In study II, 127 cows were assigned randomly to 1 of the 2 groups with either a 0-d or 30-d DPL.

Data from the 3 considered studies were merged considering that study II does not include data before wk 3 prepartum; thus, data before wk 3 prepartum from study Ia and Ib were excluded (Figure 1). In addition, because cows in study II obtained different rations after wk 7 postpartum, data from subsequent weeks were removed from all 3 studies. This left data from wk −3 to +7 relative to calving for 425 animals, of which 7 were removed because of dropout. This left 418 cows and associated data available for analysis in the present study.

Housing and Feeding

In all studies, cows were housed in a freestall, consisting of cubicles and slatted floors. Cows were milked twice a day (~0600 h and ~1800 h). Ration composition and feeding strategies were described earlier in detail (study Ia/Ib in van Knegsel et al., 2014; study II in van Hoeij et al., 2017). In short, in study Ia/Ib lactating cows were fed a basal mixed ration matching energy requirement for 25 kg of milk production and consisting of grass silage, corn silage, wheat straw, and rapeseed meal or soybean meal (51:34:2:13, DM basis). Dry cows were fed a basal mixed ration of grass silage, corn silage, wheat straw, soybean meal, and rapeseed meal (39:25:25:4:8, DM basis). From d 10 before the expected calving date, both lactating and dry cows were fed either glucogenic or lipogenic concentrate. After calving, concentrate supply was increased stepwise by 0.5 kg/d until 8.5 kg/d. In study II, lactating cows were fed a basal mixed ration matching energy requirement for 25 kg of milk production and consisting of grass silage, corn silage, wheat straw, soybean meal, and sugar beet pulp (44:34:2:10:8, DM basis). Dry cows were fed a basal mixed ration of grass silage, corn silage, wheat straw, and rapeseed meal (47:18:24:8, DM basis). From d 10 before the expected calving date, both lactating

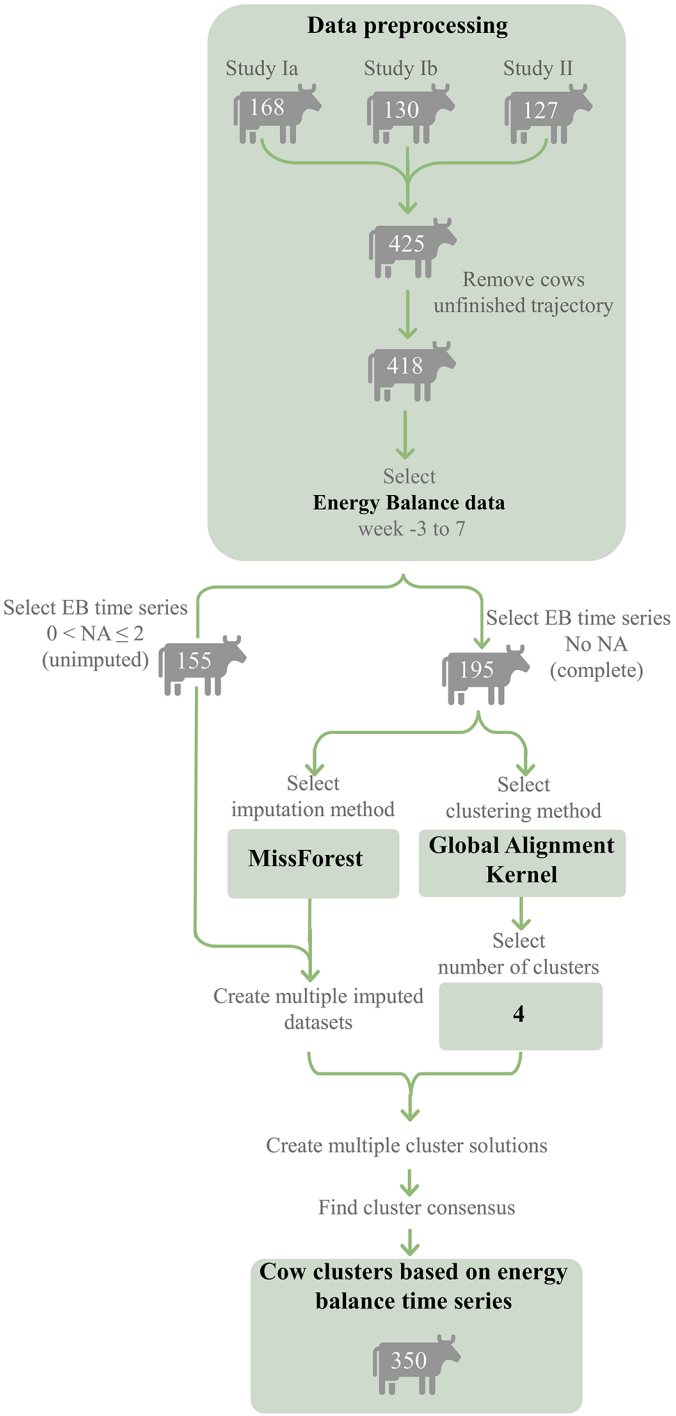

Figure 1. Workflow of the clustering of energy balance (EB) time profiles. First, data originating from 3 studies were merged and filtered. Two EB time series data sets were selected, one with complete EB time series and one with at most 2 missing EB time points for each cow. These 2 data sets were used to select the most optimal imputation method, number of clusters, and clustering method. Because a nondeterministic algorithm was chosen to impute missing EB values, multiple imputations were averaged to obtain a justifiable consensus. NA = not available.

and dry cows were fed 1 kg/d of concentrate. After calving, concentrate supply was increased stepwise by 0.3 kg/d from 4 DIM until 8.5 kg/d at 28 DIM for cows receiving standard concentrate level, or 0.3 kg/d from 4 DIM until 6.7 kg/d at 22 DIM for cows receiving a low concentrate level. In all 3 studies cows had free access to water and the basal mixed ration.

Measurements

Milk Yield and Milk Composition. Milk samples were collected 4 times per week and analyzed as a pooled sample for fat, protein, lactose, and SCC (ISO 9622; ISO, 2013; Qlip). Fat- and protein-corrected milk (**FPCM**) was calculated as (CVB, 2018)

$$\text{FPCM (kg)} = [0.337 + 0.116 \times \text{fat (\%)} + 0.06 \times \text{protein (\%)}] \times \text{milk yield (kg)}. \quad [1]$$

Feed Intake, BW, BCS, and EB. Energy balance was calculated according to the Dutch NE system for lactation (VEM; Van Es, 1975; CVB, 2018) as the difference between VEM supplied with feed and VEM required for maintenance and milk production. Animal maintenance requirements are 42.4 VEM/$kg^{0.75}$·d (1,000 VEM = 6.9 MJ of net energy). The VEM required for milk production is 442 VEM/kg of FPCM (Van Es, 1975). In calculating the maintenance and milk energy requirements, a correction factor to scale requirements to an average cow was applied as described in Van Es (1975). The energy intake and EB are expressed in kJ/$kg^{0.75}$ per day, where $kg^{0.75}$ indicates metabolic BW (Van Es, 1975).

Blood Collection and Analysis. Blood samples were collected once a week from 3 wk prepartum until 7 wk postpartum. Blood samples were collected after the morning milking and between 3 and 1 h before the morning feeding in 10-mL EDTA tubes (Vacuette, Greiner BioOne) from the coccygeal vein. Blood was centrifuged (3,000 × *g* for 15 min, 4°C) and plasma was isolated and stored at −20°C until further analysis. Concentrations of NEFA and BHB were measured enzymatically using Wako Chemicals kit no. 994-75409 and Randox Laboratories kit no. RB1007, respectively (Graber et al., 2012). The plasma glucose concentration was measured using BioMerieux kit no. 61269 (Graber et al., 2012). The plasma insulin concentration was measured using EMD Millipore Corporation kit no. PI-12K. The plasma IGF-1 concentration was measured using the Beckman Coulter kit no. A15729. For more details, we refer the reader to the original publications (van Knegsel et al., 2014; Chen et al., 2015; van Hoeij et al., 2017).

Disease Treatments. All treatments of disease during the experiment were recorded by farm staff according to farm protocols (Dairy Campus, Lelystad, the Netherlands). Disease treatments were registered daily based on treatments for disease, including milk fever, retained placenta, vaginal discharge, endometritis, cystic ovaries, mastitis, claw disorders, fever, ketosis, diarrhea, displaced abomasum, peritonitis, pneumonia, and other diseases, and then summed per EB cluster.

Table 2. Overview of energy balance (EB) cut-off values to divide cows into groups traditionally based on Mellouk et al. (2019)

Group	EB[1]
Great positive EB	EB > 77
Positive EB	0 < EB < 77
Negative EB	−77 < EB < 0
Severe negative EB	EB < −77

[1]In kJ/kg$^{0.75}$ per day.

Data Analysis

In this section we present a strategy to cluster cows based on EB time profiles. The rationale behind this strategy is to prevent cows migrating to other clusters over time, which is possible for the traditional clustering approach since each time point is treated separately.

Traditional Clustering of EB. To be able to evaluate our novel clustering method, we compared it with the traditional approach. For each time point cows were assigned to different EB groups based on cut-off values (great positive EB, positive EB, NEB, severe NEB) derived from literature (Mellouk et al., 2019) and data are presented in Table 2.

Handling of Missing Data. Because of missing data at one or more time points (weeks) it was not possible to calculate complete EB time profiles (over the 10-wk period) for all animals in the study. Complete EB profiles were available for 195 animals (45%). A total of 107 (25%) animals missed one time point, 48 (11%) missed 2, and 84 (19%) missed 3 or more time points and needed to be imputed. We took a conservative approach and decided to discard animals with 3 or more missing time points: this left 155 animals (36%) for which the EB times series needed to be imputed.

Data Imputation. Missing data were imputed using an approach based on Random Forest (Breiman, 2001) as implemented in the missForest method (Stekhoven, 2011; Stekhoven and Bühlmann, 2012). The method is based on Random Forest regression and works by averaging over many unpruned regression trees. Because there is intrinsic randomness in the Random Forest algorithm, the missForest imputation is nondeterministic: this means that if the imputation is performed multiple times, results can (slightly) change because if the randomness of the imputation procedure. This naturally leads to multiple imputation framework that must be considered when performing clustering of the imputed time series (Van Buuren, 2018).

Clustering of Time Profiles. We aimed to cluster animals on the basis of the similarity of the EB times profiles; to this scope we used a clustering approach based on global alignment kernel (**GAK**; Cuturi et al., 2007), which is well suited for clustering time series on the bases of similarity, since similarities are based on kernels that consider costs over all optional alignment distances (Marques et al., 2018).

The optimal number of clusters was determined by looking at consensus of 2 different approaches: the elbow approach (Thorndike, 1953) and the silhouette method (Rousseeuw, 1987). The elbow method is a heuristic criterion in which the total within sum of squares (i.e., a measure of the compactness of the cluster solution) is plotted against the number of clusters: the elbow of the curve corresponds to the optimal number of clusters for the given problem (Kaufman and Rousseeuw, 2009). The silhouette method measures how similar an object is to its own cluster compared with other clusters and ranges from −1 to +1. A high silhouette value indicates that a time series is well matched to a given cluster and poorly matched to other clusters. If a set of time series have all large silhouette value (thus high average silhouette value), then the clustering configuration (i.e., the number of clusters) is appropriate. The optimal number of clusters is the one that maximizes the average silhouette over a range of possible number of clusters. When applied on the 51 complete (nonimputed) time series, we found the optimal number of clusters to be between 4 and 5 and we chose the most economical solution for sake of interpretability; $k = 4$ was then also used to cluster the imputed time series as described below, which gave consistent results as in the case of the complete time series.

Definition of the Consensus Clustering. We started by generating 1,000 imputed solutions using the missForest approach, on which the GAK clustering algorithm was applied to obtain 4 clusters of times series as previously described obtaining 1,000 clustering solutions. Because cluster labeling is arbitrary (there is no fixed relationship between a partition and any class label), there is the problem of defining a consensus of the labeling. Basically, the problem is to make sure that clusters with the same labels are the same among the 1,000 solutions. The problem of finding correspondence among the labeling of different clusters can be recast as an assignment problem and solved using the Hungarian method (Kuhn, 1955), which is a standard tool to attack this type of problem (Burkard et al., 2012). We

considered the first clustering solution as a reference and found correspondence of the cluster labels of the remaining 999 cluster solutions with respect to the first solution.

Once correspondence among cluster labels of the 1,000 cluster solutions was achieved, we were left with the problem of determining a consensus clustering (i.e., pooling the clustering results obtained from the multiple imputed data sets). A consensus solution was obtained by majority voting, which is widely used when dealing with ensembles of clustering because of its simplicity, robustness, and stability (Wang et al., 2013); basically the class of a time series is the cluster label, which was selected most often across the 1,000 imputed data sets.

Statistical Analysis

Univariate Analysis. The Kruskall-Wallis test (Kruskal and Wallis, 1952) was used to compare plasma metabolites, milk variables, body condition, BW, and calf birth weight among clusters. Dunn's post hoc test was used for multiple comparison (Dunn, 1961).

Receiver Operating Characteristic Analysis. Analysis of receiver operating characteristic (**ROC**) curves (Søreide, 2009) was used to assess the discriminative power of BW in discriminating between cow clusters based on time profiles of EB; it does not require a predetermined cut-off point and is not dependent on the group size. The area under the ROC was obtained with the corresponding accuracy, specificity, and sensitivity and the estimated best concentration threshold.

Association of Disease Treatment and Cluster. Association of disease treatment and EB clusters was assessed by means of a Pearson's χ^2 (Pearson, 1900) test applied on 4×2 contingency table (clusters $\times$ number disease/nondiseased animals) for each one of the 14 different disease treatments recorded in the studies.

Software. All analysis was performed in R (The R Project; www.r-project.org/). Imputation was performed using the missForest function (options ntree = 1,000, maxiter = 10) from the missForest R package (Stekhoven, 2011). Clustering of time profiles was performed using the tsclust function (distance = 'gak') from the dtwclust package (Sarda-Espinosa, 2017). Silhouette values were calculated with the silhouette function from the cluster package (Brigo et al., 2002). Correspondence of the labels via the Hungarian method was obtained using the R function available at https://www.r-bloggers.com/2012/11/matching-clustering-solutions-using-the-hungarian-method/; clustering consensus was obtained using the majority voting function from the diceR package (Chiu and Talhouk, 2018). The ROC analysis was performed using the R package pROC (Robin et al., 2011). The χ^2 analysis was performed using the R built-in function chisq.test.

RESULTS

Clustering of Cows Per Week

Distributions of observations of the EB values for all weeks are presented in Figure 2A. We observed a shift toward more negative EB values directly after calving (wk +1), followed by a gradual shift toward positive EB values in the following weeks. Distributions of weekly EB values were more dispersed in the prepartum weeks compared with postpartum weeks. Next, we explored the clustering of animals based on EB values for each week separately (traditional clustering), based on earlier published cut-off values (Table 2). In the prepartum weeks, most of the cows were in the positive EB cluster (Figure 2B). The number of cows in the severe NEB cluster strongly increased postpartum and slowly decreased after wk 2 postpartum.

Clustering of Cows Based on EB Time Profiles

The 10 wk (−3 until +7 wk relative to calving) (imputed) EB time profiles for 350 cows corresponding to the 1,000 imputed data sets were clustered using the GAK algorithm in combination with consensus clustering. Clustering was based on the similarity of the cow-specific EB dynamics (i.e., on the similarity of the time progression to a status of EB unbalance after calving and following recovery toward a normal EB status). No other information (such as metabolite or hormone concentration) was used to inform the clustering: the animals were clustered solely on the basis of the patterns of calculated EB variation over the 10 wk. We found an optimal solution for $k = 4$ clusters (Figure 3A). We assigned descriptive names to each cluster based on the characteristics of the EB dynamics in postpartum weeks, with no reference to metabolite or hormone concentration:

(1) **SP** cluster: includes cows with a stable positive EB time profile over the 10-wk period. The animals do not experience NEB (EB > 0).
(2) **MN** cluster: includes cows with a moderate NEB (-100 kJ/kg$^{0.75}$ $<$ EB < 0) after calving.
(3) **IN** cluster: includes cows with an average NEB (-400 kJ/kg$^{0.75}$ $<$ EB < -200 kJ/kg$^{0.75}$) after calving.
(4) **SN** cluster: includes cows with severe NEB (-600 kJ/kg$^{0.75}$ $<$ EB < -200 kJ/kg$^{0.75}$) after calving.

The MN, IN, and SN clusters included cows with NEB patterns postpartum, each to a different extent; cluster SN did also differ from IN by EB levels prepartum, which were higher for the IN cluster. The number of animals in each cluster is shown in Figure 3B. The mean EB time profiles of each cluster with the associated 95% confidence interval are shown in Figure 3C. The clusters are well resolved, with minimal overlap in the prepartum weeks and distinct EB time profiles during the 7 wk postpartum. Four clusters were also obtained when the analysis was performed on the complete time profiles (i.e., excluding animal with no missing time points) as shown in Supplemental Figure S1 (https://zenodo.org/record/6337900#.Yids5JYo9PY).

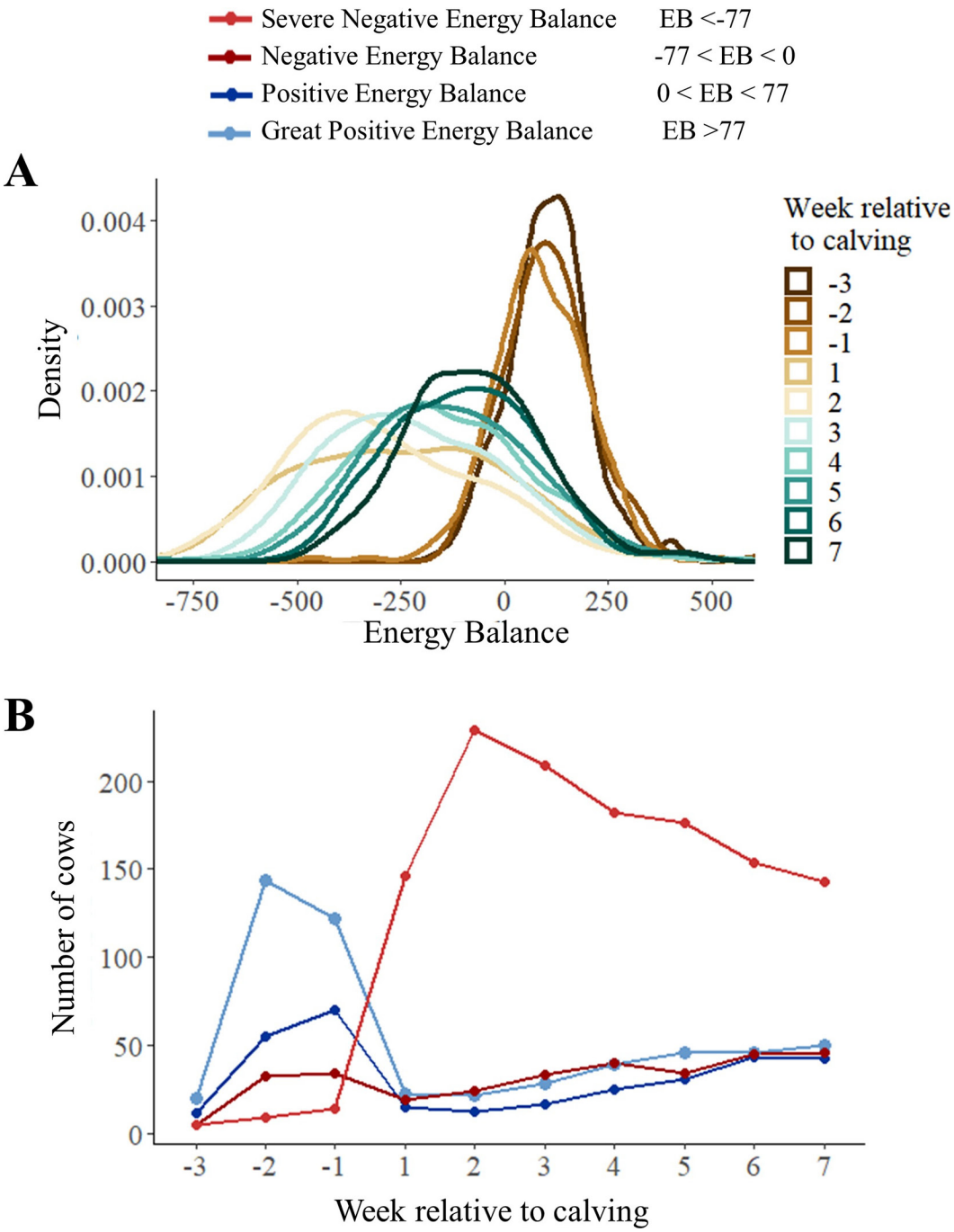

Figure 2. (A) Distribution of weekly energy balance (EB) observations from wk −3 until +7 relative to calving. Energy balance is expressed in kJ/kg$^{0.75}$ per day. (B) Group sizes of the 4 clusters for EB time profiles from wk −3 until +7 relative to calving

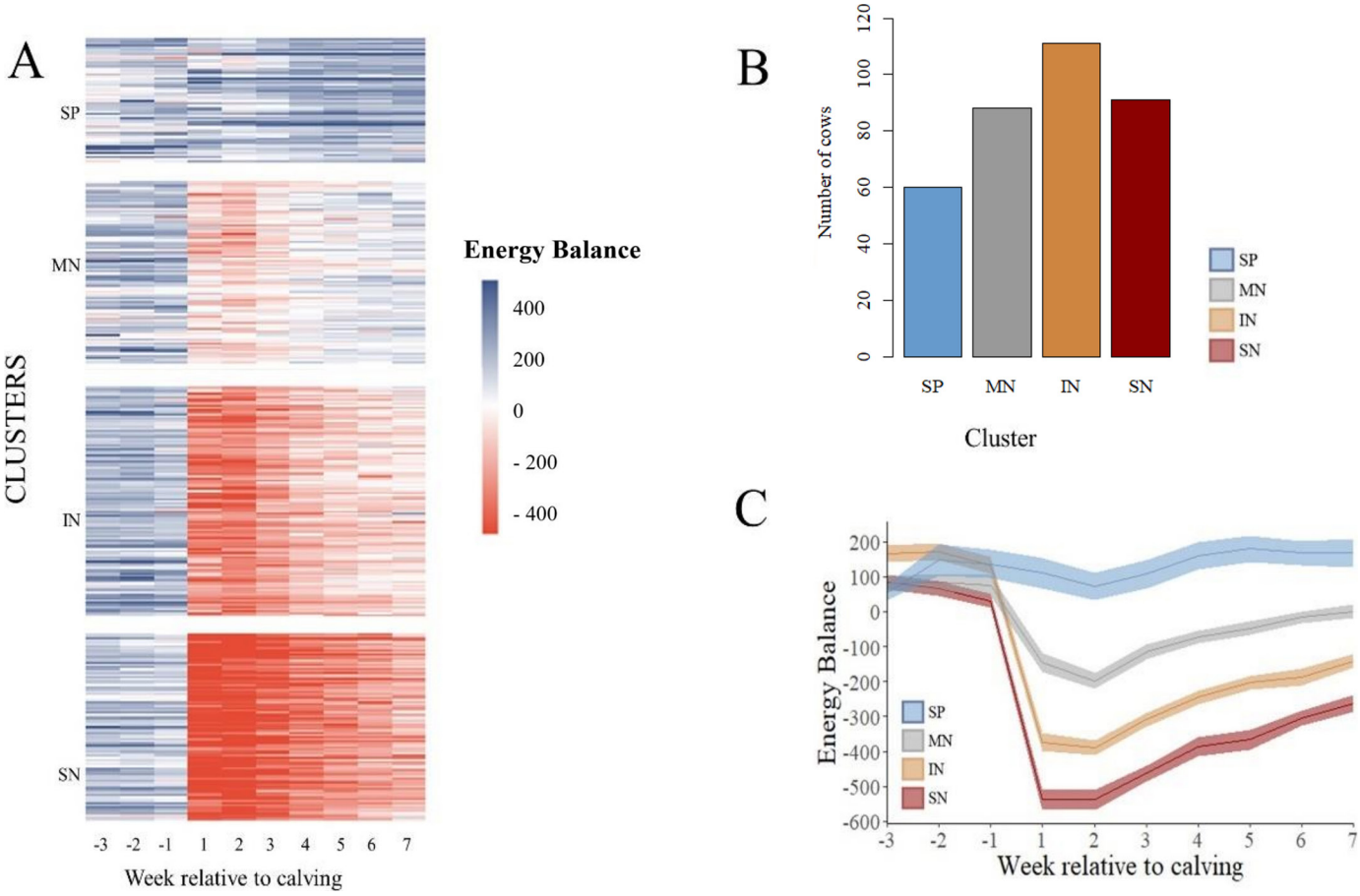

Figure 3. (A) Consensus clustering of energy balance time profiles. Each line represents time series of energy balance of an individual cow. Energy balance is expressed in kJ/kg$^{0.75}$ per day. SP = stable positive cow cluster; MN = mild negative cow cluster; IN = intermediate negative cow cluster, SN = severe negative cow cluster. (B) Number of cows in each cluster of time series of energy balance using imputed data. (C) 95% CI of means of energy balance time series for each cow cluster.

Data originated from 3 previously published studies with slightly different animal management and diet interventions. We did not observe confounding of the EB clusters with diet intervention (Figure 4B, dietary treatment; i.e., glucogenic versus lipogenic diets for study Ia and Ib were equally represented in the 4 EB time profile clusters and not associated with a particular EB cluster) or with the study (Figure 4C, study samples for each study are spread over the 4 clusters), indicating that the existence of (groups of) different EB time profiles in dairy cows is due to inherent physiological differences and not to animal management or external interventions.

Characteristics of EB Time Profile Clusters

Cow Characteristics. We examined the characteristics of the clusters with respect to dietary treatment, DPL, and parity (Figure 4). Mean parity of the cows varied among cow clusters, with average parity increasing with the severity of the NEB after calving. Average parity was 2.8 for the SP cluster, 2.9 for MN cluster, 3.3 for IN cluster, and 3.8 for SN cluster and differed among clusters (Kruskal-Wallis test: P-value < 0.001, chi-squared $= 26.91$, df $= 3$, post hoc Dunn test: SP-MN P-value $= 0.73$; SP-IN P-value $= 0.014$; SP-SN P-value < 0.001; MN-IN P-value $= 0.011$; MN-SN P-value < 0.001; IN-SN P-value $= 0.032$). Dry period length also varied among clusters: cows with a short DPL had a more positive EB profile postpartum compared with cows with longer DPL. The SN cluster (i.e., cows with severe NEB) had the largest number of cows who dried off by themselves, although they were initially planned to have a 0-d DP.

Milk, Metabolic Profiles, and Disease Treatments. Next, we explored the dynamics of milk, milk composition, and blood metabolites for the 4 cow clusters based on EB time profiles (Figure 5). Milk-related variables had distinctive patterns among clusters postpartum. The SN cluster had the highest milk

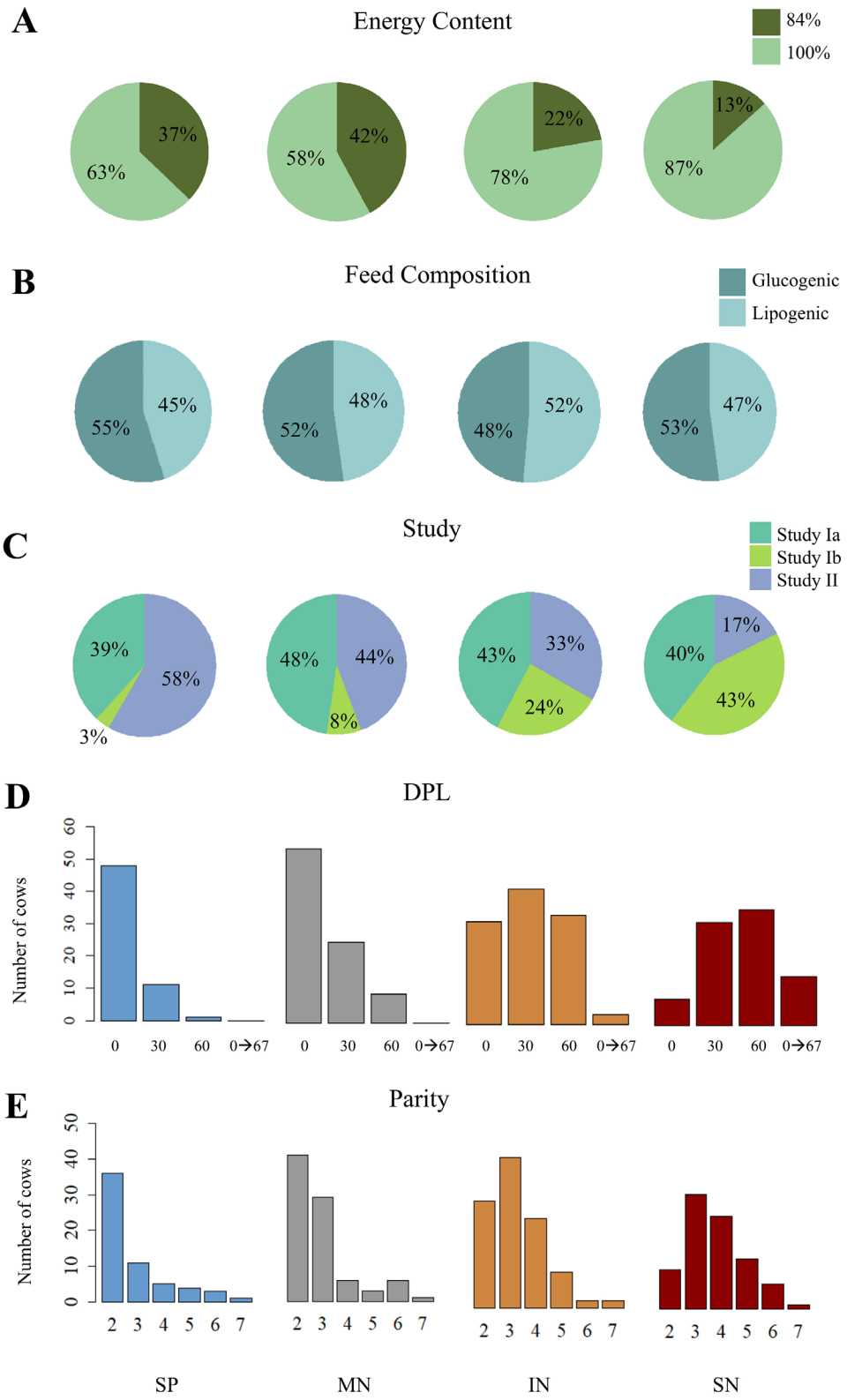

Figure 4. Relationship between energy balance (EB) time profile clusters and study characteristics. Energy content (A), feed composition (B), study (C), parity (D), and dry period length (DPL; E) distributions are shown. Feed composition is only applicable for study Ia and Ib. SP = stable positive cow cluster, MN = mild negative cow cluster, IN = intermediate negative cow cluster, SN = severe negative cow cluster. The DPL classes are expressed in days except for DPL = 0→67, which are cows that were initially classified as DPL = 0 but were dried off due to low milk yield (<4 kg/d) at least 30 d before calving.

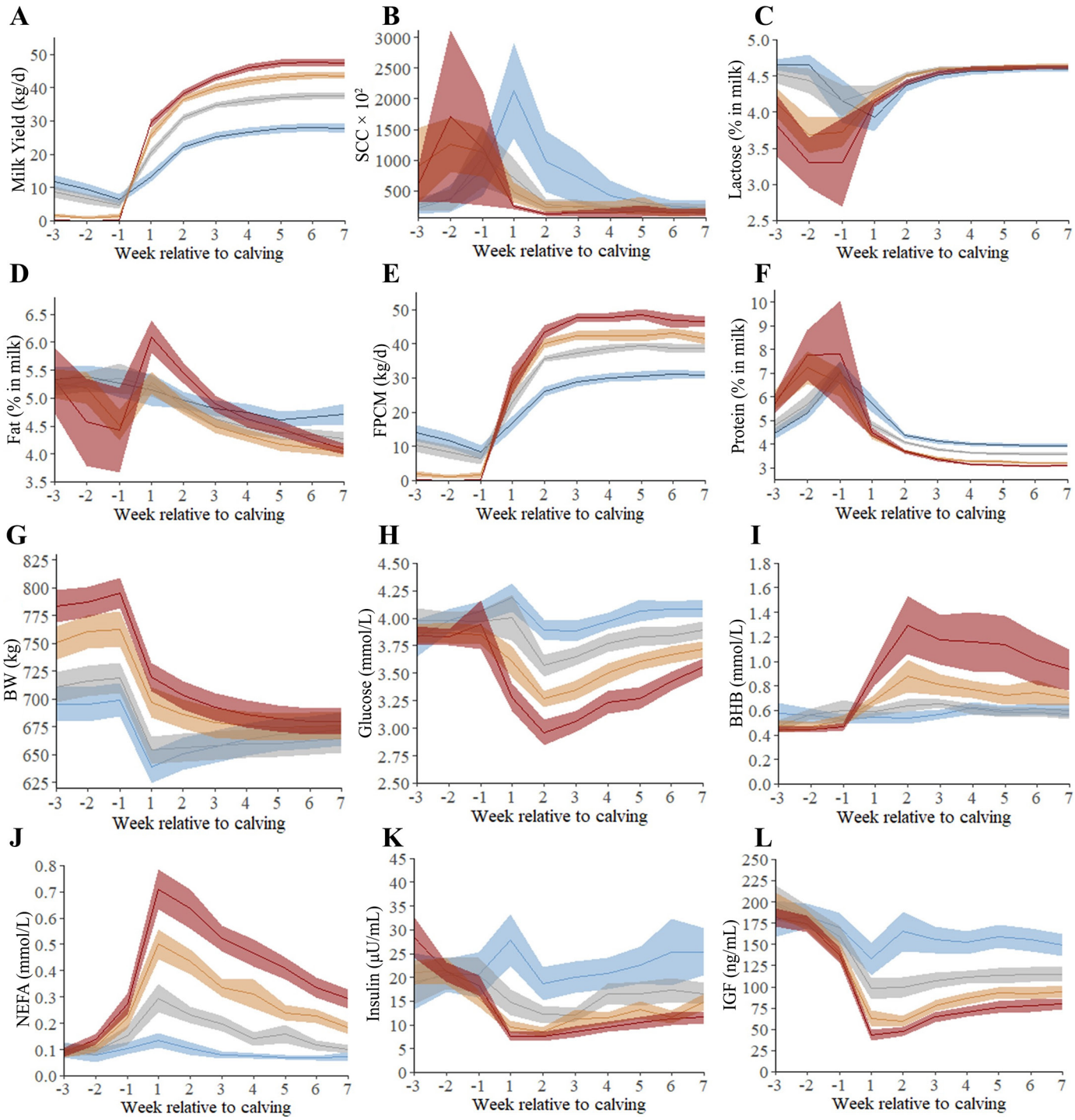

Figure 5. Metabolic profiles for each cow cluster. (A) Milk yield (in kg/d); (B) SCC (in 10^2 cells/mL), (C) lactose (in % in milk); (D) fat (in % in milk); (E) fat- and protein-corrected milk yield (FPCM; in kg/d); (F) protein (in % in milk); (G) BW (in kg); (H) plasma glucose concentration (in mmol/L); (I) plasma BHB concentration (in mmol/L); (J) plasma nonesterified fatty acid concentration (NEFA, in mmol/L); (K) plasma insulin concentration; and (L) plasma IGF-1 concentration from wk −3 until +7 relative to calving for each cluster of time series of energy balance. Mean and confidence intervals of the mean are visualized. SP = stable positive, MN = mild negative, IN = intermediate negative, SN = severe negative.

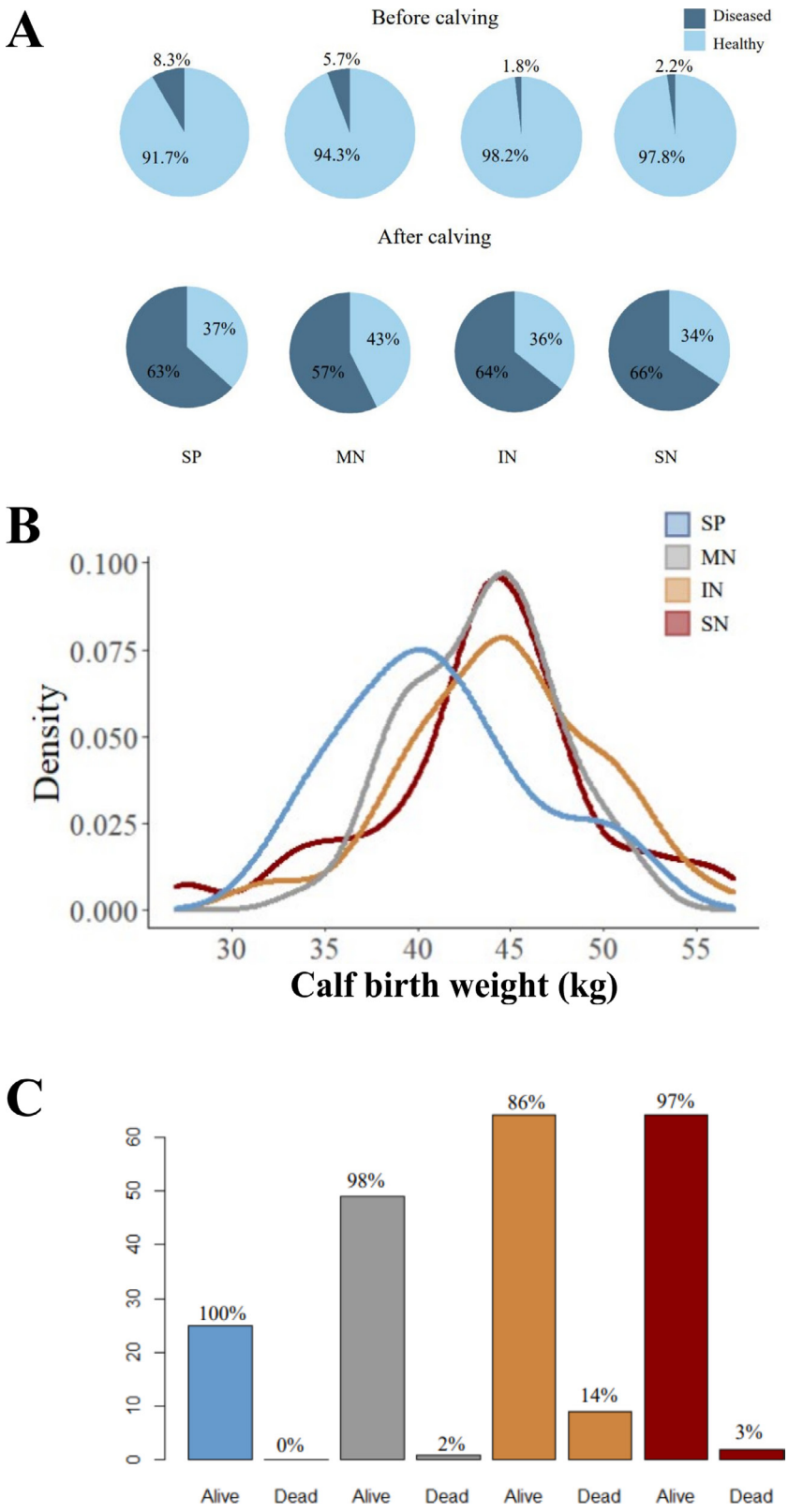

Figure 6. (A) Distributions of disease treatments before and after calving among clusters. (B) Distribution of observations of calf birth weight among clusters for time series of energy balance. (C) Distributions of observations for liveborn (+) or death calves for each cluster for time series of energy balance. SP = stable positive cow cluster, MN = mild negative cow cluster, IN = intermediate negative cow cluster, SN = severe negative cow cluster. Percentages show the percentage dead or alive for each cluster.

Table 3. Overview of disease treatments per cluster for wk 1 until 7 postpartum[1]

Item	Cluster				Chi-squared test[2]
	SP	MN	IN	SN	
Cluster size	60	88	111	91	
Milk fever	0 (0%)	4 (4.6%)	7 (6.3%)	16 (18%)	χ^2 (3, n = 350) = 19, $P < 0.001$
Retained placenta	6 (10%)	15 (17%)	13 (12%)	9 (9.9%)	χ^2 (3, n = 350) = 2.7, $P = 0.45$
Vaginal discharge	10 (17%)	13 (15%)	9 (8.1%)	8 (8.8%)	χ^2 (3, n = 350) = 4.4, $P = 0.22$
Endometritis	7 (12%)	9 (10%)	14 (13%)	9 (9.9%)	χ^2 (3, n = 350) = 0.47, $P = 0.92$
Cystic ovaries	4 (6.7%)	7 (8.0%)	19 (17%)	16 (18%)	χ^2 (3, n = 350) = 7.4, $P = 0.061$
Mastitis	14 (23%)	11 (13%)	10 (9.0%)	18 (20%)	χ^2 (3, n = 350) = 4.9, $P = 0.18$
Claw disorders	4 (6.7%)	12 (14%)	12 (11%)	13 (14%)	χ^2 (3, n = 350) = 2.5, $P = 0.48$
Fever	4 (6.7%)	7 (8.0%)	9 (8.1%)	7 (7.7%)	χ^2 (3, n = 350) = 0.12, $P = 0.99$
Ketosis	1 (1.7%)	0 (0%)	2 (1.8%)	0 (0%)	χ^2 (3, n = 350) = 3.2, $P = 0.37$
Diarrhea	0 (0%)	1 (1.1%)	2 (1.8%)	4 (4.4%)	χ^2 (3, n = 350) = 4.3, $P = 0.24$
Displaced abomasum	2 (3.3%)	0 (0%)	3 (2.7%)	1 (1.1%)	χ^2 (3, n = 350) = 3.3, $P = 0.35$
Peritonitis	0 (0%)	0 (0%)	0 (0%)	1 (1.1%)	χ^2 (3, n = 350) = 2.9, $P = 0.41$
Pneumonia	0 (0%)	1 (1.1%)	0 (0%)	0 (0%)	χ^2 (3, n = 350) = 3.0, $P = 0.39$
Other	1 (1.7%)	3 (3.4%)	6 (5.4%)	4 (4.4%)	χ^2 (3, n = 350) = 1.5, $P = 0.67$
Total no. of disease treatments	53	83	106	106	
Total no. of cows with disease treatments	38	51	72	61	

[1]The numbers of cows within each cluster treated for a disease are shown. The percentages of cows from the cluster that were treated for the disease are shown in parentheses.
[2]χ^2 (df, n = sample size).

yield and SP cluster had the lowest milk yield. Milk fat percentage was more stable over time for the SP cluster when compared with the other clusters. Somatic cell count for the SP cluster rose postpartum compared with other clusters. The SN, IN, and MN clusters were characterized by reduced plasma IGF-1, insulin, and glucose concentrations postpartum compared with the SP cluster. Plasma NEFA and BHB concentrations were highest postpartum for the SN cluster. Disease treatment rates are similar among clusters prepartum (Figure 6A). After calving, an increase of disease treatments is visible for all clusters. More cows were treated for milk fever ($P < 0.001$) and there was a tendency ($P = 0.061$) for more cows treated for cystic ovaries in the SN cluster (Table 3).

Table 4. Statistical comparison of BW in the 4 clusters for 3 wk prepartum[1]

Item	Prepartum week relative to calving		
	−3	−2	−1
Kruskal-Wallis test			
	$H^2 = 63.96$ df = 3 $P < 0.001$	H = 69.25 df = 3 $P < 0.001$	H = 73.51 df = 3 $P < 0.001$
Dunn test		*P*-value	
SP-MN	0.20	0.099	0.10
SP-IN	<0.001	<0.001	<0.001
SP-SN	<0.001	<0.001	<0.001
MN-IN	<0.001	<0.001	<0.001
MN-SN	<0.001	<0.001	<0.001
IN-SN	0.0036	0.015	0.0043

[1]SP = stable positive cow cluster; MN = mild negative cow cluster; IN = intermediate negative cow cluster, SN = severe negative cow cluster.
[2]H = H statistic.

BW of Cows and Calves. Body weight was different in the prepartum weeks for all 4 clusters (Table 4). The ROC analysis showed that discrimination between SP and SN clusters is feasible and that it is possible to assign a cow to different clusters based on BW prepartum (Table 5). Since before calving the total cow BW is the sum of the weight of the calf and that of the mother, we also examined the distribution of calf weight across the 4 clusters (Figure 6B).

We observed smaller calves for the cows in the SP cluster, compared with calves of the cows in the IN, MN, and SN clusters (Kruskall-Wallis test: *P*-value = 0.02, df = 3, $\chi^2 = 9.60$, post hoc Dunn test: SN vs. IN *P*-value = 0.04; SN vs. MN *P*-value = 0.002; SN vs. SP *P*-value = 0.016).

We also explored the association between the number of calves born dead and the EB clusters (Figure 6C). Chi-squared analysis of the 4 × 2 contingency table corresponding to Figure 6C indicated that the number of calves born dead is different among EB clusters (*P*-value = 0.0206; $\chi^2 = 9.77$, df = 3).

DISCUSSION

In this study we obtained well-resolved cow clusters based on EB time profiles in the peripartum period. The EB-based time profiles have not been developed earlier, although cluster analysis has been applied before on early-lactation dairy cow data (Tremblay et al.,

Table 5. Receiver operating characteristic analysis of BW (kg) for 3 wk prepartum[1]

	Prepartum week														
	−3					−2					−1				
Item	AUC	Acc	Spec	Sens	Threshold (kg)	AUC	Acc	Spec	Sens	Threshold (kg)	AUC	Acc	Spec	Sens	Threshold (kg)
SN-IN	0.62	0.62	0.68	0.57	770	0.60	0.61	0.69	0.54	762	0.62	0.63	0.62	0.63	767
SN-MN	0.80	0.73	0.70	0.76	773	0.78	0.72	0.69	0.75	762	0.78	0.72	0.61	0.83	764
SN-SP	0.87	0.82	0.81	0.83	733	0.84	0.78	0.74	0.83	744	0.83	0.78	0.77	0.79	741
IN-MN	0.67	0.65	0.79	0.48	699	0.66	0.65	0.80	0.47	705	0.65	0.64	0.77	0.49	702
IN-SP	0.75	0.69	0.61	0.83	724	0.75	0.69	0.63	0.80	735	0.72	0.69	0.65	0.76	742
MN-SP	0.58	0.58	0.40	0.83	729	0.59	0.58	0.43	0.80	735	0.57	0.57	0.43	0.78	743

[1]AUC = area under the curve; Acc = accuracy; Spec = specificity; Sens = sensitivity. SP = stable positive cow cluster; MN = mild negative cow cluster; IN = intermediate negative cow cluster; SN = severe negative cow cluster.

2018; De Koster et al., 2019; Xu et al., 2019). These studies focused on clustering of orthogonal components, milk, or plasma variables, respectively, to unravel metabolic status during early lactation. Our study focused on the same topic with a different point of view by now clustering EB time profiles, not permitting cows to change cluster from week to week as in more traditional clustering approaches and allowing us to gain new insights in the dynamics of metabolic status peripartum. This novel way of examining dynamics of EB levels of cows could improve understanding of the relationship between EB and metabolic status and would facilitate customized management strategies based on individual time profiles of EB.

Cow clusters were associated with fixed variables such as parity. Cows within the SN cluster had on average a higher parity than cows within the SP, IN, and MN cluster. Cows of greater parity produce more milk (Lee and Kim, 2006) and are therefore more likely to experience a severe EB drop postpartum. This finding is consistent with earlier studies (Friggens et al., 2007; Macrae et al., 2019; Van et al., 2020), where positive correlations between EB and parity were identified.

In addition, cows with a severe EB drop postpartum had a longer dry period. The SN cluster contained relatively more cows which dried off themselves (DPL = 0→67) than other clusters: from this we can derive that cows within this group (SN) are likely to have a severe NEB postpartum. This may be related to the fact at the previous lactation, this group of cows had a lower milk yield, which resulted in spontaneous drying off and was associated with an increase in BCS prepartum (Chen et al., 2015). It is well known that an elevated BCS at calving is related to a decrease in energy intake, a more severe NEB and increased risk for metabolic disorders after calving (Morrow, 1976). Indeed, the specific group of cows that dried themselves off in the current data (DPL = 0→67) had a lower energy intake and more negative EB after calving (Chen et al., 2015).

Somatic cell count levels in milk of cows in the SP cluster were higher postpartum, compared with the other clusters. This agrees with the observation that cows with shorter or no DPL have a relatively high SCC, but a more positive EB (van Knegsel et al., 2014). An explanation for this observation could be that cows with a short or no DP have a reduced generation of mammary epithelial cells in the weeks before calving. Less renewal of mammary epithelial cells could result in a lower concentration of young and active cells at calving, resulting in a lower milk yield (Capuco et al., 1997). Because of the lower milk yield after calving, cows in the SP cluster have a better EB, but the somatic cells in milk are less diluted and increase in concentration (Steeneveld et al., 2013).

Metabolic status corresponding to each EB cluster was examined and showed distinguishable patterns postpartum. The cluster of cows with severe EB drop postpartum had high NEFA and BHB concentration and low IGF-1, insulin, and glucose concentration, which is in line with earlier studies (Reist et al., 2002; Friggens et al., 2007; Van et al., 2020) and earlier explained by van Hoeij et al. (2017). Cows experiencing a severe NEB have a low glucose concentration in plasma, and therefore insulin production is not enhanced, causing increased levels of NEFA. Increased plasma NEFA are known to be related to metabolic diseases (Adewuyi et al., 2005; Ospina et al., 2010). Cows in the SN clusters had higher NEFA concentration postpartum and NEFA signatures postpartum could be used to indicate whether a cow goes into a severe NEB state postpartum. Although average metabolic status of cows differed among EB clusters in the current study, it could be possible that cows in the same EB cluster differ in metabolic status. It can be hypothesized that including information on metabolic profiles in the EB clustering results in differentiation of metabolic status within and between EB clusters.

Although levels for milk and blood variables among the different EB clusters tend to overlap in the prepartum weeks, complicating cluster discrimination, we observed different average BW in the week prepartum for the cows in the different clusters. Prepartum BW and prepartum EB had a negative correlation, which is in line with literature (Heuer et al., 2000, 2001). An explanation for this could be that cows experiencing severe NEB are able to mobilize more fat postpartum from a larger body mass. Mäntysaari and Mäntysaari (2010) reported that EB and BW postpartum are negatively correlated, which is in line with our results. Increased body reserves were associated with increased levels of body fat content, causing in turn increased leptin, which has a negative effect on feed intake (Liefers et al., 2003). Cows with a relatively high body fat will eat less, which results in a severe NEB. Likewise, more fat mobilization will lead to a disturbed availability in C2:C3 ratio. This causes increased plasma NEFA concentration, therefore increasing the chance to obtain metabolic diseases (van Hoeij et al., 2017). In the current study cows experiencing a severe EB drop postpartum had higher BW prepartum compared with cows experiencing an intermediate, mild, or no EB drop postpartum.

Body weight prepartum could be an interesting prepartum biomarker to detect which cows go into a severe NEB postpartum. For this reason, we examined whether BW prepartum could be used to assign cows to different EB groups. We applied ROC analysis to investigate the power of BW to discriminate between different postpartum cow clusters. The ROC analysis did show that it is possible to discriminate between SP and SN clusters and that it is possible to assign a cow to different clusters based on weight. It should be noted that BW is also used to determine EB values and thus association between BW and EB is expected. In addition, clusters have also been determined using the dynamics of the EB time profiles, indicating that not only the actual EB are related to BW but also the temporal dynamics.

The number of disease treatments, especially for milk fever and cystic ovaries, during wk 1 until 7 differed among EB clusters (Table 3). It is well known that disease incidence is high in dairy cows in early lactation, which is related to the calving process and the increase in milk yield directly after calving (Friggens et al., 2004; Koeck et al., 2012). Most existing studies, however, differentiated to type of disease only to a limited extent (Collard et al., 2000; Koeck et al., 2012).

In this study, cows were more often treated for milk fever in the clusters with a more negative EB, compared with the clusters with a better EB. First, this could be related to the detrimental characteristics of milk fever causing a poor start of the new lactation, with milk fever in itself also being a risk factor for secondary diseases during the early-lactation period (DeGaris and Lean, 2008). Second, cows in the clusters with a more negative EB also had a higher milk yield, whereas high-producing cows are more at risk for milk fever (Fleischer et al., 2001). Also, for cystic ovarian disease it is known that its occurrence is positively related to milk yield potential of dairy cows (Koeck et al., 2012), as also reviewed in Vanholder et al. (2006). The etiology of cystic ovarian disease might be related to alterations in metabolic status in high-producing dairy cows in early lactation (Vanholder et al., 2006). Moreover, it can be hypothesized that cystic ovaries occur in the cascade of multiple disease events of dairy cows that have a very poor start of their lactation, also related to a severe NEB in these cows.

In the cluster with cows with a stable EB, more calves were born with a lower birth weight (Figure 6b). In the positive EB cluster there were more cows with a 0-d dry period, compared with the other clusters. In an earlier study (Mayasari et al., 2015), we reported that cows with a 0-d dry period had calves with on average a lower birth weight, which was partially explained by also a shorter pregnancy length for cows with a 0-d dry period (278 vs. 280 or 1,281 d for 0-d vs. 30-d or 60-d dry period).

Calf mortality was negatively associated with EB postpartum (Figure 6c). It is well known that stillbirth of calves is a risk factor for cows and associated, for example, with an increase in cow mortality (Shahid

et al., 2015). Moreover, stillbirth is associated with an increased risk for dystocia and retained placenta in dairy cows (Ghavi Hossein-Zadeh and Ardalan, 2011). As a consequence, recovery from parturition is more complex after a stillbirth, which could be associated with limited feed intake and poor adaptation to a new lactation resulting in a more negative EB, as observed in the current study.

Overall, we demonstrated that cows have distinct EB time profiles and that characteristics such as prepartum BW, and postpartum NEFA, glucose, insulin, and IGF-1 concentrations identify the time profile of EB and potentially the risk for metabolic diseases. Under practical circumstances, the prediction EB profile of individual cows could be used to adjust cow management strategies. In the current study, however, EB time profiles are to some extent simplified by a limited number of clusters. With a more advanced approach, it might be interesting to examine gradient-wise EB profiling instead of dividing cows into a limited number of clusters. In addition, metabolic time profiles could be used to discriminate EB state of cows, it is known that for example NEFA and insulin are related to EB (Reist et al., 2002; Fenwick et al., 2008; Van et al., 2020). Moreover, plenty of studies are available that illustrate the value of metabolites in milk (Gustafsson and Emanuelson, 1996; van Knegsel et al., 2010) or milk fatty acid profiles, either determined by GC or estimated by milk infrared spectra (McParland and Berry, 2016; Bach et al., 2019; Churakov et al., 2021) to indicate the energy status of dairy cows in early lactation. Because of the noninvasive characteristics of these milk measures, exploration of the value of these measures with EB time profiles would be highly relevant from a practical perspective.

CONCLUSIONS

This study demonstrated that dairy cows can be phenotyped based on their EB time profiles (i.e., on the patterns of transition to and recovery from NEB in the weeks before and after calving). We could obtain 4 well-resolved clusters. Moreover, BW predicted EB profile peripartum, although ROC analysis indicated it is difficult to discriminate between neighboring clusters due to overlapping BW distributions. Furthermore, the EB time series clusters were mirrored by specific metabolic profiles as well as dynamics in milk yield, composition, and BW. Cows with severe EB drop (SN cluster) postpartum had high NEFA and BHB, low IGF-1, insulin, and glucose concentration and were treated more for milk fever and cystic ovaries. Since NEFA is known to be related to metabolic diseases, SN cluster might be related to it as well. Being able to distinguish this cluster from the other clusters could pave the way to personalized animal treatment related to metabolic diseases. In addition, more analysis is required to indicate if BW is a possible on-farm biomarker to detect energy profile of a cow in an early stage, which might help to prevent severe metabolic states or diseases.

ACKNOWLEDGMENTS

The authors thank the staff of Dairy Campus (Lelystad, the Netherlands) for their technical support during the 3 studies. The authors thank the Dutch Dairy Board (PZ, the Netherlands), Product Board Animal Feed (PDV, the Netherlands), and CRV (the Netherlands) for financing study Ia and Ib. The authors thank DairyNL (ZuivelNL, the Netherlands) and Ministry of Economic Affairs (EZ, the Netherlands) for financing study II, which was also part of the Public-Private Partnership "Sustainable Dairy Chain" (Duurzame Zuivelketen, the Netherlands). The authors have not stated any conflicts of interest.

REFERENCES

Adewuyi, A., E. Gruys, and F. Van Eerdenburg. 2005. Non esterified fatty acids (NEFA) in dairy cattle. A review. Vet. Q. 27:117–126. https://doi.org/10.1080/01652176.2005.9695192.

Bach, K., D. Barbano, and J. McArt. 2019. Association of mid-infrared-predicted milk and blood constituents with early-lactation disease, removal, and production outcomes in Holstein cows. J. Dairy Sci. 102:10129–10139. https://doi.org/10.3168/jds.2019-16926.

Breiman, L. 2001. Random forests. Mach. Learn. 45:5–32. https://doi.org/10.1023/A:1010933404324.

Brigo, D., F. Mercurio, and F. Rapisarda. 2002. Parametrizing correlations: A geometric interpretation. Milan: Banca IMI Working Paper.

Burkard, R., M. Dell'Amico, and S. Martello. 2012. Assignment problems: Revised reprint. SIAM.

Capuco, A., R. Akers, and J. Smith. 1997. Mammary growth in Holstein cows during the dry period: Quantification of nucleic acids and histology. J. Dairy Sci. 80:477–487. https://doi.org/10.3168/jds.S0022-0302(97)75960-5.

Chen, J., J. J. Gross, H. van Dorland, G. Remmelink, R. Bruckmaier, B. Kemp, and A. van Knegsel. 2015. Effects of dry period length and dietary energy source on metabolic status and hepatic gene expression of dairy cows in early lactation. J. Dairy Sci. 98:1033–1045. https://doi.org/10.3168/jds.2014-8612.

Chiu, D. S., and A. Talhouk. 2018. diceR: An R package for class discovery using an ensemble driven approach. BMC Bioinformatics 19:11. https://doi.org/10.1186/s12859-017-1996-y.

Churakov, M., J. Karlsson, A. E. Rasmussen, and K. Holtenius. 2021. Milk fatty acids as indicators of negative energy balance of dairy cows in early lactation. Animal 15:100253. https://doi.org/10.1016/j.animal.2021.100253.

Collard, B. L., P. J. Boettcher, J. C. Dekkers, D. Petitclerc, and L. R. Schaeffer. 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. J. Dairy Sci. 83:2683–2690. https://doi.org/10.3168/jds.S0022-0302(00)75162-9.

Cuturi, M., J.-P. Vert, O. Birkenes, and T. Matsui. 2007. A kernel for time series based on global alignments. Pages II-413–II-416 in Proc. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07. IEEE.

CVB. 2018. Chemische samenstellingen en nutritionele waarden van voedermiddelen. Federatie Nederlandse Diervoederketen. Centraal Veevoederbureau.

De Koster, J., M. Salavati, C. Grelet, M. Crowe, E. Matthews, R. O'Flaherty, G. Opsomer, L. Foldager, and M. Hostens. 2019. Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers. J. Dairy Sci. 102:2631–2644. https://doi.org/10.3168/jds.2018-15533.

De Vries, M., and R. Veerkamp. 2000. Energy balance of dairy cattle in relation to milk production variables and fertility. J. Dairy Sci. 83:62–69. https://doi.org/10.3168/jds.S0022-0302(00)74856-9.

DeGaris, P. J., and I. J. Lean. 2008. Milk fever in dairy cows: A review of pathophysiology and control principles. Vet. J. 176:58–69. https://doi.org/10.1016/j.tvjl.2007.12.029.

Dunn, O. J. 1961. Multiple comparisons among means. J. Am. Stat. Assoc. 56:52–64. https://doi.org/10.1080/01621459.1961.10482090.

Fenwick, M., S. Llewellyn, R. Fitzpatrick, D. Kenny, J. Murphy, J. Patton, and D. Wathes. 2008. Negative energy balance in dairy cows is associated with specific changes in IGF-binding protein expression in the oviduct. Reproduction 135:63–75. https://doi.org/10.1530/REP-07-0243.

Fleischer, P., M. Metzner, M. Beyerbach, M. Hoedemaker, and W. Klee. 2001. The relationship between milk yield and the incidence of some diseases in dairy cows. J. Dairy Sci. 84:2025–2035. https://doi.org/10.3168/jds.S0022-0302(01)74646-2.

Friggens, N. C., J. B. Andersen, T. Larsen, O. Aaes, and R. J. Dewhurst. 2004. Priming the dairy cow for lactation: A review of dry cow feeding strategies. Anim. Res. 53:453–473. https://doi.org/10.1051/animres:2004037.

Friggens, N. C., P. Berg, P. Theilgaard, I. Korsgaard, K. Ingvartsen, P. Løvendahl, and J. Jensen. 2007. Breed and parity effects on energy balance profiles through lactation: Evidence of genetically driven body energy change. J. Dairy Sci. 90:5291–5305. https://doi.org/10.3168/jds.2007-0173.

Garnsworthy, P. C., A. A. Fouladi-Nashta, G. E. Mann, K. D. Sinclair, and R. Webb. 2009. Effect of dietary-induced changes in plasma insulin concentrations during the early post partum period on pregnancy rate in dairy cows. Reproduction 137:759–768. https://doi.org/10.1530/REP-08-0488.

Ghavi Hossein-Zadeh, N., and M. Ardalan. 2011. Cow-specific risk factors for retained placenta, metritis and clinical mastitis in Holstein cows. Vet. Res. Commun. 35:345–354. https://doi.org/10.1007/s11259-011-9479-5.

Graber, M., S. Kohler, A. Müller, K. Burgermeister, T. Kaufmann, R. Bruckmaier, and H. Van Dorland. 2012. Identification of plasma and hepatic parameters related to metabolic robustness in dairy cows. J. Anim. Physiol. Anim. Nutr. (Berl.) 96:75–84. https://doi.org/10.1111/j.1439-0396.2010.01124.x.

Gustafsson, A., and U. Emanuelson. 1996. Milk acetone concentration as an indicator of hyperketonaemia in dairy cows: The critical value revised. Anim. Sci. 63:183–188. https://doi.org/10.1017/S1357729800014739.

Heuer, C., W. Van Straalen, Y. Schukken, A. Dirkzwager, and J. Noordhuizen. 2000. Prediction of energy balance in a high yielding dairy herd in early lactation: Model development and precision. Livest. Prod. Sci. 65:91–105. https://doi.org/10.1016/S0301-6226(99)00177-3.

Heuer, C., W. Van Straalen, Y. Schukken, A. Dirkzwager, and J. Noordhuizen. 2001. Prediction of energy balance in high yielding dairy cows with test-day information. J. Dairy Sci. 84:471–481. https://doi.org/10.3168/jds.S0022-0302(01)74497-9.

ISO. 2013. ISO 9622:2013. Milk and liquid milk products—Guidelines for the application of mid-infrared spectrometry. International Standardization Organization.

Kaufman, L., and P. J. Rousseeuw. 2009. Finding Groups in Data: An Introduction to Cluster Analysis. Vol. 344. Wiley.

Klein, M. S., M. F. Almstetter, G. Schlamberger, N. Nürnberger, K. Dettmer, P. J. Oefner, H. Meyer, S. Wiedemann, and W. Gronwald. 2010. Nuclear magnetic resonance and mass spectrometry-based milk metabolomics in dairy cows during early and late lactation. J. Dairy Sci. 93:1539–1550. https://doi.org/10.3168/jds.2009-2563.

Koeck, A., F. Miglior, D. Kelton, and F. Schenkel. 2012. Health recording in Canadian Holsteins: Data and genetic parameters. J. Dairy Sci. 95:4099–4108. https://doi.org/10.3168/jds.2011-5127.

Kruskal, W. H., and W. A. Wallis. 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47:583–621. https://doi.org/10.1080/01621459.1952.10483441.

Kuhn, H. W. 1955. The Hungarian method for the assignment problem. Nav. Res. Logist. Q. 2:83–97. https://doi.org/10.1002/nav.3800020109.

Lee, J.-Y., and I.-H. Kim. 2006. Advancing parity is associated with high milk production at the cost of body condition and increased periparturient disorders in dairy herds. J. Vet. Sci. 7:161–166. https://doi.org/10.4142/jvs.2006.7.2.161.

Liefers, S., R. Veerkamp, M. Te Pas, C. Delavaud, Y. Chilliard, and T. Van der Lende. 2003. Leptin concentrations in relation to energy balance, milk yield, intake, live weight, and estrus in dairy cows. J. Dairy Sci. 86:799–807. https://doi.org/10.3168/jds.S0022-0302(03)73662-5.

Macrae, A., E. Burrough, J. Forrest, A. Corbishley, G. Russell, and D. Shaw. 2019. Risk factors associated with excessive negative energy balance in commercial United Kingdom dairy herds. Vet. J. 250:15–23. https://doi.org/10.1016/j.tvjl.2019.06.001.

Mäntysaari, P., and E. A. Mäntysaari. 2010. Predicting early lactation energy balance in primiparous Red Dairy Cattle using milk and body traits. Acta Agric. Scand. A Anim. Sci. 60:79–87. https://doi.org/10.1080/09064702.2010.496002.

Marques, A. R., H. Forde, and C. W. Revie. 2018. Time-series clustering of cage-level sea lice data. PLoS One 13:e0204319. https://doi.org/10.1371/journal.pone.0204319.

Mayasari, N., G. de Vries Reilingh, M. Nieuwland, G. Remmelink, H. Parmentier, B. Kemp, and A. Van Knegsel. 2015. Effect of maternal dry period length on colostrum immunoglobulin content and on natural and specific antibody titers in calves. J. Dairy Sci. 98:3969–3979. https://doi.org/10.3168/jds.2014-8753.

McParland, S., and D. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. J. Dairy Sci. 99:4056–4070. https://doi.org/10.3168/jds.2015-10051.

Mellouk, N., C. Rame, D. Naquin, Y. Jaszczyszyn, J.-L. Touzé, E. Briant, D. Guillaume, T. Ntallaris, P. Humblot, and J. Dupont. 2019. Impact of the severity of negative energy balance on gene expression in the subcutaneous adipose tissue of periparturient primiparous Holstein dairy cows: Identification of potential novel metabolic signals for the reproductive system. PLoS One 14:e0222954. https://doi.org/10.1371/journal.pone.0222954.

Morrow, D. A. 1976. Fat cow syndrome. J. Dairy Sci. 59:1625–1629. https://doi.org/10.3168/jds.S0022-0302(76)84415-3.

Odens, L. J., R. Burgos, M. Innocenti, M. J. VanBaale, and L. H. Baumgard. 2007. Effects of varying doses of supplemental conjugated linoleic acid on production and energetic variables during the transition period. J. Dairy Sci. 90:293–305. https://doi.org/10.3168/jds.S0022-0302(07)72630-9.

Ospina, P., D. Nydam, T. Stokol, and T. Overton. 2010. Evaluation of nonesterified fatty acids and β-hydroxybutyrate in transition dairy cattle in the northeastern United States: Critical thresholds for prediction of clinical diseases. J. Dairy Sci. 93:546–554. https://doi.org/10.3168/jds.2009-2277.

Pearson, K. 1900. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Lond. Edinb. Dublin Philos. Mag. J. Sci. 50:157–175. https://doi.org/10.1080/14786440009463897.

Probo, M., O. B. Pascottini, S. LeBlanc, G. Opsomer, and M. Hostens. 2018. Association between metabolic diseases and the culling risk of high-yielding dairy cows in a transition management facility using survival and decision tree analysis. J. Dairy Sci. 101:9419–9429. https://doi.org/10.3168/jds.2018-14422.

Rastani, R., R. Grummer, S. Bertics, A. Gümen, M. Wiltbank, D. Mashek, and M. Schwab. 2005. Reducing dry period length to simplify feeding transition cows: Milk production, energy balance, and metabolic profiles. J. Dairy Sci. 88:1004–1014. https://doi.org/10.3168/jds.S0022-0302(05)72768-5.

Reist, M., D. Erdin, D. Von Euw, K. Tschuemperlin, H. Leuenberger, Y. Chilliard, H. Hammon, C. Morel, C. Philipona, Y. Zbinden, N. Kuenzi, and J. W. Blum. 2002. Estimation of energy balance at the individual and herd level using blood and milk traits in high-yielding dairy cows. J. Dairy Sci. 85:3314–3327. https://doi.org/10.3168/jds.S0022-0302(02)74420-2.

Robin, X., N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, and M. Müller. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12:77. https://doi.org/10.1186/1471-2105-12-77.

Rousseeuw, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

Sarda-Espinosa, A. 2017. Package 'dtwclust': Time series clustering along with optimizations for the dynamic time warping distance. https://cran.r-project.org/web/packages/dtwclust/dtwclust.pdf.

Seifi, H. A., S. J. LeBlanc, K. E. Leslie, and T. F. Duffield. 2011. Metabolic predictors of post-partum disease and culling risk in dairy cattle. Vet. J. 188:216–220. https://doi.org/10.1016/j.tvjl.2010.04.007.

Shahid, M. Q., J. K. Reneau, H. Chester-Jones, R. C. Chebel, and M. I. Endres. 2015. Cow- and herd-level risk factors for on-farm mortality in Midwest US dairy herds. J. Dairy Sci. 98:4401–4413. https://doi.org/10.3168/jds.2014-8513.

Søreide, K. 2009. Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research. J. Clin. Pathol. 62:1–5. https://doi.org/10.1136/jcp.2008.061010.

Steeneveld, W., Y. Schukken, A. Van Knegsel, and H. Hogeveen. 2013. Effect of different dry period lengths on milk production and somatic cell count in subsequent lactations in commercial Dutch dairy herds. J. Dairy Sci. 96:2988–3001. https://doi.org/10.3168/jds.2012-6297.

Stekhoven, D. J. 2011. Using the missForest package. R package:1–11. https://stat.ethz.ch/education/semesters/ss2012/ams/paper/missForest_1.2.pdf.

Stekhoven, D. J., and P. Bühlmann. 2012. MissForest—Non-parametric missing value imputation for mixed-type data. Bioinformatics 28:112–118. https://doi.org/10.1093/bioinformatics/btr597.

Sun, L., H. Zhang, L. Wu, S. Shu, C. Xia, C. Xu, and J. Zheng. 2014. 1H-Nuclear magnetic resonance-based plasma metabolic profiling of dairy cows with clinical and subclinical ketosis. J. Dairy Sci. 97:1552–1562. https://doi.org/10.3168/jds.2013-6757.

Tamminga, S., P. Luteijn, and R. Meijer. 1997. Changes in composition and energy content of liveweight loss in dairy cows with time after parturition. Livest. Prod. Sci. 52:31–38. https://doi.org/10.1016/S0301-6226(97)00115-2.

Thorndike, R. L. 1953. Who belongs in the family? Psychometrika 18:267–276. https://doi.org/10.1007/BF02289263.

Tremblay, M., M. Kammer, H. Lange, S. Plattner, C. Baumgartner, J. Stegeman, J. Duda, R. Mansfeld, and D. Döpfer. 2018. Identifying poor metabolic adaptation during early lactation in dairy cows using cluster analysis. J. Dairy Sci. 101:7311–7321. https://doi.org/10.3168/jds.2017-13582.

Van, Q. C. D., E. Knapp, J.-L. Hornick, and I. Dufrasne. 2020. Influence of days in milk and parity on milk and blood fatty acid concentrations, blood metabolites and hormones in early lactation Holstein cows. Animals (Basel) 10:2081. https://doi.org/10.3390/ani10112081.

Van Buuren, S. 2018. Flexible Imputation of Missing Data. CRC Press.

Van Es, A. 1975. Feed evaluation for dairy cows. Livest. Prod. Sci. 2:95–107. https://doi.org/10.1016/0301-6226(75)90029-9.

van Hoeij, R., J. Dijkstra, R. Bruckmaier, J. J. Gross, T. Lam, G. Remmelink, B. Kemp, and A. van Knegsel. 2017. The effect of dry period length and postpartum level of concentrate on milk production, energy balance, and plasma metabolites of dairy cows across the dry period and in early lactation. J. Dairy Sci. 100:5863–5879. https://doi.org/10.3168/jds.2016-11703.

van Knegsel, A. T., G. Remmelink, S. Jorjong, V. Fievez, and B. Kemp. 2014. Effect of dry period length and dietary energy source on energy balance, milk yield, and milk composition of dairy cows. J. Dairy Sci. 97:1499–1512. https://doi.org/10.3168/jds.2013-7391.

van Knegsel, A. T., S. Van Der Drift, M. Horneman, A. De Roos, B. Kemp, and E. Graat. 2010. Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. J. Dairy Sci. 93:3065–3069. https://doi.org/10.3168/jds.2009-2847.

van Knegsel, A. T., H. Van den Brand, J. Dijkstra, S. Tamminga, and B. Kemp. 2005. Effect of dietary energy source on energy balance, production, metabolic disorders and reproduction in lactating dairy cattle. Reprod. Nutr. Dev. 45:665–688. https://doi.org/10.1051/rnd:2005059.

Vanholder, T., G. Opsomer, and A. De Kruif. 2006. Aetiology and pathogenesis of cystic ovarian follicles in dairy cattle: A review. Reprod. Nutr. Dev. 46:105–119. https://doi.org/10.1051/rnd:2006003.

Wang, H., Y. Yang, H. Wang, and D. Chen. 2013. Soft-voting clustering ensemble. Pages 307–318 in Multiple Classifier Systems. Lecture Notes in Computer Science. Vol. 7872. Z. H. Zhou, F. Roli, and J. Kittler, ed. Springer-Verlag.

Xu, W., A. T. M. van Knegsel, J. J. M. Vervoort, R. M. Bruckmaier, R. J. van Hoeij, B. Kemp, and E. Saccenti. 2019. Prediction of metabolic status of dairy cows in early lactation with on-farm cow data and machine learning algorithms. J. Dairy Sci. 102:10186–10201. https://doi.org/10.3168/jds.2018-15791.

ORCIDS

F. Vossebeld https://orcid.org/0000-0003-4275-3863
A. T. M. van Knegsel https://orcid.org/0000-0003-1959-3363
E. Saccenti https://orcid.org/0000-0001-8284-4829