



J. Dairy Sci. 101:1–15  
<https://doi.org/10.3168/jds.2017-14232>  
 © American Dairy Science Association®, 2018.

## Relationship between herd size and measures of animal welfare on dairy cattle farms with freestall housing in Germany

Daniel Gieseke,<sup>\*1</sup> Christian Lambertz,<sup>†</sup> and Matthias Gauly<sup>†</sup>

<sup>\*</sup>Department of Animal Sciences, Georg-August-University of Göttingen, 37075 Göttingen, Germany

<sup>†</sup>Faculty of Science and Technology, Free University of Bolzano, 39100 Bolzano, Italy

### ABSTRACT

The objective of this study was to examine the association of herd size with animal welfare in dairy cattle herds. Therefore, 80 conventional dairy cattle farms were classified by the number of cows into 4 herd size classes, C1 (<100 cows), C2 (100–299 cows), C3 (300–499 cows), and C4 ( $\geq 500$  cows), and assessed using multiple animal-based measures of the Welfare Quality Assessment protocol for dairy cattle. Data were recorded from April 2014 to September 2016 by an experienced single assessor in Northern Germany. Each farm was visited 2 times at an interval of 6 mo (summer period and winter period) to avoid seasonal effects on the outcome. The average herd size was  $383 \pm 356$  Holstein-Friesian cows (range 45 to 1,629). Only farms with freestall (cubicle) housing and a maximum of 6 h access to pasture per day were included in the study. Data were statistically analyzed using a generalized linear mixed model. None of the farms reached the highest overall rating of “excellent.” The majority of the farms were classified as “enhanced” (30%) or “acceptable” (66%), and at 6 assessments the farms were rated as “not classified” (4%). Regarding single indicators, mean trough length per cow, percentage of cows with nasal discharge, and vulvar discharge increased with increasing herd size, whereas it was vice versa for displacements of cows. Percentage of lean cows, percentage of dirty lower legs, and duration of the process of lying down showed a curvilinear relationship with the number of cows per farm. Herd size was not associated with any other measures of the Welfare Quality protocol. In conclusion, herd size effects were small, and consequently herd size cannot be considered as a feasible indicator of the on-farm animal welfare level. Housing conditions and management practices seem to have a greater effect on animal welfare than the number of dairy cows per farm.

**Key words:** animal welfare, herd size, dairy cow, welfare quality protocol for cattle

### INTRODUCTION

Over the last decades, milk production has intensified continuously in Europe. The number of dairy cattle farms in the member states of the European Union-10 (Belgium, Denmark, Germany, Ireland, Greece, France, Italy, Luxembourg, the Netherlands, and United Kingdom) decreased from 1,514,441 to 288,600 farms between 1983 and 2013 (–81%). During this period the dairy cattle population declined by 31% from 25,143,828 to 15,460,770 animals, whereas average milk production remained stable (100 million t). This indicates a significant improvement of the milk yield per cow (Eurostat, 2015), which might affect the welfare of dairy cows especially because of increasing health disorders (Coignard et al., 2014). Driven by this development, the average herd size in the stated European Union countries increased from 17 to 54 dairy cows per farm (Eurostat, 2015). A similar herd size development is noted on a global scale. Between 1970 and 2006, herd sizes increased from 19 to 120 animals in the United States (MacDonald et al., 2007), from 93 to 284 in Australia (Dairy Australia, 2015), and from 140 to 413 in New Zealand (Dairy New Zealand, 2014). During the same period, public awareness of animal welfare issues in livestock farming has increased (European Commission, 2016) and many consumers are concerned about the enduring intensification of livestock production (Spooner et al., 2014). The so-called factory farms are perceived as having serious animal health and welfare problems (Prickett et al., 2010; Vanhonacker and Verbeke, 2014). From the consumers’ perspective, natural housing conditions are essential for animal welfare and these would only be provided in small-scale family farms (Krystallis et al., 2009; Spooner et al., 2014). Contrastingly, farmers do not seem to associate herd size with welfare-related issues. Performance and health of the animals are from their point of view more suitable indicators of animal welfare and can be

Received December 1, 2017.

Accepted April 4, 2018.

<sup>1</sup>Corresponding author: [dgiesek1@gwdg.de](mailto:dgiesek1@gwdg.de)

maintained independent of the herd size (Vanhonacker et al., 2008; Sorensen and Fraser, 2010). However, little scientific evidence is available about whether there is a direct correlation between herd size and farm animal welfare status. Larger herds are associated with increasing stock per labor unit, increasing stocking density, and less access to pasture. Furthermore, better-trained staff, routine veterinary herd health visits, and monitoring systems are more likely in large herds (Beggs et al., 2015; Robbins et al., 2016). Numerous studies have considered herd size as an influencing factor for animal welfare issues with inconsistent results. For example, increasing herd size was reported as detrimental factor for lameness (Alban, 1995; de Vries et al., 2014) or mastitis (Archer et al., 2013; Lievaart et al., 2007), whereas in other studies a protective effect of increasing herd sizes was found for lameness (Dippel et al., 2009; Solano et al., 2015) or mastitis (Oleggini et al., 2001; Simensen et al., 2010). Others, however, did not observe any relationship between herd size and lameness (Barker et al., 2010; Fabian et al., 2014) or mastitis (Whitaker et al., 2004; Ivemeyer et al., 2011). Robbins et al. (2016) pointed out in their comprehensive review about farm size and animal welfare that study designs in the existing literature differ considerably and the definition of large and small herd sizes are country specific. Moreover, herd size has only been taken into account as a risk factor for single animal welfare indicators. Due to the multidimensional character of animal welfare, a holistic approach considering various health and behavior parameters is necessary to investigate the overall effect of herd size on animal welfare (Robbins et al., 2016). Therefore, the objective of this study was to examine the relationship between herd size and animal welfare and to analyze whether herd size could be used as an indicator of animal welfare at the herd level. A total of 80 conventional dairy cattle farms with different herd sizes (small: <100; medium: 100–299; large: 300–499; very large:  $\geq$ 500 cows) were assessed using multiple animal-based measures of the Welfare Quality Assessment protocol (WQP) for dairy cattle (Welfare Quality, 2012).

## MATERIALS AND METHODS

### Study Design

The data collection was conducted from April 2014 to September 2016 by a single assessor on 80 conventional dairy farms located in Northern Germany. Each farm was visited 2 times at an interval of 6 mo (summer period and winter period) to avoid seasonal effects on the animal welfare assessment. At both visits the

animal welfare was assessed using the WQP. This is a standardized indicator system for on-farm animal welfare assessment. It focuses mainly on animal-based measures, which directly reflect the actual welfare state of the animals. More than 30 animal welfare indicators covering aspects of feeding, housing, health, and behavior were measured and aggregated to 12 welfare criteria and 4 welfare principles (Welfare Quality, 2012). Finally, farms were assigned to 1 of 4 overall welfare categories, representing an “excellent,” “enhanced,” or “acceptable” animal welfare state. In cases where minimum requirements could not be achieved, the farms were rated as “not classified.” The assessor was trained intensively by a member of the Welfare Quality Network (Christoph Winckler, University of Natural Resources and Life Sciences, Vienna, Austria) to ensure the correct application of the dairy-cattle-specific indicators of the WQP. The official 4-d training course at the University of Natural Resources and Life Sciences in Vienna (Austria) consisted of theoretical exercises with photos and videos as well as practical applications of the WQP on different dairy cattle farms. Data collection of this study was conducted by this assessor only.

### Farm Selection

Farm acquisition was organized with the support of different agricultural stakeholders (e.g., chamber of agriculture, milk recording association, and research facilities). For participation in the study, some specific requirements had to be fulfilled to guarantee the comparability of the housing environment. All lactating dairy cows in the sample had to be kept in loose housing barns with deep bedded or rubber mat-equipped cubicles. The dominant breed was Holstein Friesian so that genetic effects could be excluded. Farms with access to pasture for more than 6 h per day were omitted from the study because this resource-based indicator has a high weighting within the aggregation system of the WQP and an inclusion would lead to a substantial confounding effect. Nine farms provided access to pasture for less than 6 h per day without significant feed intake (“outdoor loafing areas”). There were no other limitations regarding housing conditions, milking techniques, or feeding systems (for further characterizations of the farms see Table 1). The 80 dairy cattle farms were classified based on the herd sizes according to the classification of the federal statistical office in Germany (Destatis, 2017). The first class had <100 dairy cows (C1), the second 100 to 299 dairy cows (C2), the third 300 to 499 dairy cows (C3), and the fourth  $\geq$ 500 cows (C4). Each class consisted of 20 farms. Maximum group sizes were documented, defined as the maximum

**Table 1.** Characteristics (farm data, performance data, and housing conditions) of 80 dairy farms classified by herd size into C1 = <100 cows, C2 = 100–299 cows, C3 = 300–499 cows, or C4 = ≥500 cows

Farm characteristic <sup>1</sup>	Parameter	Herd size class			
		C1 (n = 20)	C2 (n = 20)	C3 (n = 20)	C4 (n = 20)
<b>Farm data</b>					
Herd size	Mean ± SD	74 ± 14	185 ± 59	381 ± 57	892 ± 327
	Range	45–99	102–284	304–493	548–1,629
Group size	Mean ± SD	60 ± 10	85 ± 28	118 ± 39	147 ± 50
	Range	38–77	32–134	48–188	84–267
Hectares/cow	Mean ± SD	1.5 ± 0.9	3.0 ± 3.5	3.0 ± 2.3	2.0 ± 1.5
	Range	0.7–4.6	0.3–14.3	0.4–7.6	0.4–6.2
<b>Performance data</b>					
Fat, %	Mean ± SD	4.04 ± 0.19	4.03 ± 0.16	3.95 ± 0.21	3.92 ± 0.18
	Range	3.60–4.40	3.66–4.46	3.65–4.60	3.50–4.42
Protein, %	Mean ± SD	3.34 ± 0.08	3.38 ± 0.07	3.38 ± 0.07	3.36 ± 0.10
	Range	3.15–3.53	3.21–3.50	3.19–3.58	3.13–3.50
ECM, kg	Mean ± SD	9,464 ± 1,024	9,588 ± 1,158	10,127 ± 716	10,027 ± 771
	Range	7,091–11,046	6,434–11,761	8,637–11,792	8,343–11,736
<b>Housing conditions, % (no.)</b>					
Cubicles	Deep bedded	65 (13)	70 (14)	90 (18)	70 (14)
	Rubber mat	35 (7)	30 (6)	10 (2)	30 (6)
Floors	Plain	5 (1)	60 (12)	75 (15)	75 (15)
	Slatted	95 (19)	40 (8)	25 (5)	25 (5)
Feeding	Feed rack	95 (19)	45 (9)	40 (8)	20 (4)
	Neck tube	5 (1)	55 (11)	60 (12)	80 (16)
Insulation	Insulated	80 (16)	35 (7)	40 (8)	35 (7)
	Not insulated	20 (4)	65 (13)	60 (12)	65 (13)

number of dairy cows within a lactation group. These were considered because associations between single animal welfare indicators and group size, rather than herd size, were assumed (e.g., agonistic interactions).

### Data Collection

Animal welfare assessments were conducted following the instructions of the WQP for dairy cattle (Welfare Quality, 2012). Minor modifications were made for practical or statistical reasons (e.g., nonlactating cows were excluded). Modified animal welfare indicators and assessment methods are described in Table 2. The farm visits usually started at 0800 h ( $\pm 1$  h) after morning milking and lasted, depending on herd size, 6 to 13 h. Data collection was performed at each farm in a fixed order. At the beginning, the human-animal-relationship was assessed using the avoidance-distance-test at the feeding rack. Subsequently the Qualitative Behavior Assessment was executed for 25 min to evaluate the emotional state of the herd. Behavioral observations took place in up to 12 different segments of the barn (with a maximum of 25 dairy cows per segment). Lying behavior, agonistic behavior, and coughing were recorded using a continuous behavior sampling (approximately 150–180 min). The clinical scoring of individual dairy cows was carried out in a sample according to

group size, if animals were kept in different groups. All measures were recorded in the same sample of animals (body condition, cleanliness, lameness, integument alterations, discharge, diarrhea, hampered respiration). Depending on the herd size, a sample of 32 to 102 cows were assessed at each farm visit. In total, 10,758 individual animals were observed during the 2-yr data collection period (5,405 in summer and 5,353 in winter season). Finally, resource-based measures (water provision, cleanliness, functionality) were assessed in the barn and management-based measures (mortality, dystocia, downer cows, and dehorning practice) as well as milk-recording data (SCC) were gathered during a farmer interview. Five farms were excluded because herd sizes changed considerably during data collection period and 2 farms were excluded because the housing system changed. Data sets of the 80 dairy cattle farms were complete for all measures of the WQP. In addition to the indicators of the WQP, farm characteristics such as design of barn (insulated, not insulated), cubicles (deep bedded, rubber mat), floor (slatted, plain), or feeding place (neck tube, feed rack) were recorded.

### Statistical Analysis

Single indicators were summarized to criteria, principles, and overall score using decision trees, I-spline-

**Table 2.** Data collected for the assessment of the animal welfare level of dairy cows using the Welfare Quality protocol (modified after Coignard et al., 2013; modifications of the Water Quality protocol are highlighted with bold letters)

Welfare indicator	Source <sup>1</sup>	Frequency calculation <sup>2</sup>	Method for collecting data
Good feeding			
Body condition	ABM	BCS observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = regular body condition, 1 = very lean body condition, 2 = very fat body condition
Water provision and cleanliness	RBM	Length, cleanliness, and functionality of water troughs and bowls	<b>Troughs and bowls: length measured in cm</b> , and test of sufficient water flow. Check of cleanliness: 0 = clean: drinkers and water clean, 1 = partly dirty: drinkers dirty, but water fresh, 2 = dirty: drinkers and water dirty
Good housing			
Duration of the process of lying down	ABM	Mean time to lie down observed on a sample of at least 6 cows	Recording of time needed to lie down in seconds from the moment the carpal joint was bent and lowered (before touching the ground) until hind quarter had fallen down and pulled front leg
Collisions during lying down	ABM	Observed on a sample of at least 6 cows	Animal collides with the housing equipment during the lying down movement: 0 = no collision observed, 2 = collision observed
Lying outside the lying area	ABM	Observed in representative segments of the barn	Percentage of animals lying partly or completely outside the designated lying area, out of all lying animals
Cleanliness of udder and legs	ABM	Cleanliness observed on a sample of cows ( $N_{fin}$ )	Assessment on lower hind legs, hind quarters, and udder. Animal is assessed using the scale: 0 = no dirt or minor splashing, 2 = separate or continuous dirt plaques
Presence of tethering (trestall)	RBM	—	Only farms with cubicle housing systems were chosen for this study.
Good health			
Lameness (loose housed)	ABM	Lameness observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = not lame: timing of steps and weight-bearing equal on all 4 feet, 1 = moderately lame: imperfect temporal rhythm in stride creating a limp, 2 = severely lame: reluctance to bear weight on 1 limb or more than 1 limb affected
Integument alterations	ABM	Integument alterations observed on a sample of cows ( $N_{fin}$ )	Observation of 1 side of the animal. The number of hairless patches and lesions/swellings of a minimum diameter of 2 cm are recorded.
Coughing	ABM	Mean number of coughs expressed per cow and per 15 min	Recording using continuous behavior sampling of sudden and noisy expulsion of air from the lungs during a total period of 120 min
Nasal discharge	ABM	Nasal discharge observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = no evidence of nasal discharge, 2 = evidence of clearly visible discharge from the nostrils
Ocular discharge	ABM	Ocular discharge observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = no evidence of ocular discharge, 2 = evidence of clearly visible discharge (wet/dry); at least 3 cm
Hampered respiration	ABM	Hampered respiration observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = no evidence of hampered respiration, 2 = evidence of deep and labored respiration
Diarrhea	ABM	Diarrhea observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = no evidence of diarrhea, 2 = evidence of watery manure on both sides of the tail; >size of a hand
Vulvar discharge	ABM	Vulvar discharge observed on a sample of cows ( $N_{fin}$ )	Animal is assessed using the scale: 0 = no evidence of vulvar discharge, 2 = evidence of purulent effluent from the vulva
Milk SCC	QUE	Cows with subclinical mastitis within the last 3 mo (milk recording)	Cow milk SCC are assessed using the scale: 0 = SCC below 400,000 cells/mL for the last 3 mo, 2 = SCC above 400,000 cell/mL at least once in the last 3 mo
Mortality	QUE	Annual cumulative incidence of mortality	Defined as the percentage of dairy cows that died or were euthanized by a veterinarian or emergency slaughtered during last 12 mo
Dystocia	QUE	Annual cumulative incidence of dystocia	Defined as percentage of calvings where major assistance was required during the last 12 mo
Downer cows	QUE	Annual cumulative incidence of downer cows	Defined as percentage of severe cases of downer cows during the last 12 mo (>24 h lying)
Disbudding/dehorning	QUE	—	Farmer is asked about management practices for disbudding/dehorning of the calves or dairy cows (procedures, anesthetics/analgesics).
Tail docking	QUE	—	Not practiced in Germany except an emergency case (legislation)

*Continued*



**Table 2 (Continued).** Data collected for the assessment of the animal welfare level of dairy cows using the Welfare Quality protocol (modified after Coignard et al., 2013; modifications of the Water Quality protocol are highlighted with bold letters)

Welfare indicator	Source <sup>1</sup>	Frequency calculation <sup>2</sup>	Method for collecting data
Appropriate behavior	ABM	Observed in representative segments of the barn	Recording using continuous behavior sampling during a total period of 120 min: number of head butts, displacements, chases, and fights
Agonistic behavior	ABM	Days per year with at least 6 h at pasture	Number of days per year on pasture, number of hours per day on pasture
Access to pasture	ABM	Observed on a sample of cows ( $N_{fm}$ )	Starting at 2 m distance at the feed bunk; observer approached with an extended hand. If cow showed withdrawal, avoidance distance was estimated between the hand and the muzzle.
Avoidance distance test	ABM	Observed in representative segments of the barn	Behavioral observation of the herd for being active, relaxed, fearful, agitated, calm, content, indifferent, frustrated, friendly, bored, playful, positive, lively, inquisitive, irritable, uneasy, sociable, apathetic, happy, or distressed (20 min)
Qualitative behavior assessment	ABM		

<sup>1</sup>ABM = animal-based measure; RBM = resource-based measure; QUE = questionnaire.

<sup>2</sup> $N_{fm}$  = sample size according to Welfare Quality (2012).

functions, and Choquet integrals as described in the WQP. Data were checked before analysis for normal distribution using the test of Shapiro-Wilk ( $P < 0.05$ ). If normal distribution was not given, data were log-, sqrt-, or arcsin-transformed. For parameters without normal distribution after transformation,  $P$ -values were calculated assuming a Poisson distribution. All statistical analyses were computed with SAS, version 9.4 (Statistical Analysis Systems, Cary, NC) using the GLIMMIX procedure. The following linear mixed effects model was used to analyze principles, criteria, and indicators:

$$Y_{ijkl} = \mu + a_i + b_j + c_{ij} + d_k + e_l + \varepsilon_{ijklm},$$

with  $Y_{ijkl}$  = observed value of principles/criteria/indicators,  $\mu$  = overall average of sample,  $a_i$  = fixed effect of herd size  $i$  (C1, C2, C3, C4),  $b_j$  = fixed effect of season  $j$  (summer, winter),  $c_{ij}$  = interaction between herd size  $i$  and season  $j$ ,  $d_k$  = covariate of maximum group size  $k$  (continuous),  $e_l$  = random effect of the farm  $l$ , and  $\varepsilon_{ijklm}$  = residual error. Farm was defined as the statistical unit and results were considered as statistically significant at  $P < 0.05$ . Differences between herd size classes were assessed by applying the Tukey-Kramer test. Scores are presented as least squares means and pooled standard errors. Additionally, a chi-squared test of independence (Fisher's exact test) followed by pair-wise comparisons with Bonferroni-Holm correction were performed at the level of the overall classification to test for possible differences between herd size classes.

## RESULTS

### Overall Assessment

Table 3 illustrates that none of the farms reached the highest rating of "excellent." The majority of the farms were classified as "enhanced" (30%) or "acceptable" (66%). At 6 assessments (3 in summer, 3 in winter), the minimum welfare requirements of the WQP were not met and farms were consequently rated as "not classified" (4%). Most of the dairy farms achieved the same overall scores in summer and winter season (68%), whereas 8 farms were rated better in summer (10%) and 18 farms better in winter (22%). Herd size classes C1 and C3 achieved higher overall scores in winter, whereas C2 and C4 were rated better in the summer season. No significant effect of the season was found ( $P > 0.05$ ). Pair-wise comparisons revealed that C2 achieved significantly lower overall scores in winter compared with the classes with larger herds (C3 and C4). In contrast, overall scores of the herd size classes did not differ in the summer season ( $P > 0.05$ ).

**Table 3.** Overall classification of the Welfare Quality assessments in season 1 (summer) and season 2 (winter) for dairy farms with <100 cows (C1), 100–299 cows (C2), 300–499 cows (C3), or  $\geq 500$  cows (C4) and  $\chi^2$  test of independence (Fisher's exact test;  $n = 20$ )<sup>1</sup>

Overall classification	Herd size class								$\chi^2$ test comparing herd size classes		
	C1		C2		C3		C4		S	W	T
	S	W	S	W	S	W	S	W			
Excellent	0	0	0	0	0	0	0	0	0.13	<0.01	<0.01
Enhanced	3	7	3	1	5	12	9	8			
Acceptable	15	13	17	16	14	8	11	12			
Not classified	2	0	0	3	1	0	0	0			
$\chi^2$ test comparing summer to winter assessment within herd size classes	0.17		0.22		0.05		1.00				

<sup>1</sup>S = summer; W = winter; T = total.

### Principles

Results from the level of principles in summer and winter season are illustrated in Table 4. For the principle “good feeding,” an effect of herd size ( $P < 0.01$ ) and herd size and season interaction was observed ( $P < 0.01$ ). Farms of class C2 achieved lower scores in summer compared with the larger farms of classes C3 in winter and C4 in summer. The mean principle score of C2 in winter was significantly lower than the scores of C3 and C4 at both assessments. Additionally, a significant effect of group size was found ( $P < 0.01$ ). The larger the group size, the lower was the principle score for “good feeding.” Scores of the principle “good housing” were not affected by class, season, or the interaction ( $P > 0.05$ ). The principle “good health” ( $22.6 \pm 1.1$ ) was the one with the lowest average scores, compared with “good feeding” ( $37.4 \pm 5.3$ ), “good housing” ( $60.0 \pm 1.8$ ), and “appropriate behavior” ( $38.2 \pm 1.1$ ). Results of this principle were not affected by herd size ( $P > 0.05$ ), but in all classes values were higher in winter ( $24.0 \pm 0.6$ ) than in summer ( $21.1 \pm 0.6$ ;  $P < 0.01$ ). For the principle “appropriate behavior,” neither of the tested factors had an effect ( $P > 0.05$ ).

### Criteria

As presented in Table 4, herd size classes differed in the criterion “absence of prolonged hunger” ( $P < 0.05$ ). Averaged over both seasons, C2 received a significantly lower score ( $38.8 \pm 2.9$ ) compared with C3 ( $51.4 \pm 2.9$ ) and C4 ( $52.0 \pm 3.3$ ), respectively. The scores of C1 ( $41.4 \pm 3.3$ ) did not differ from the scores of the other herd size classes ( $P > 0.05$ ). For the criterion “absence of prolonged thirst,” an effect of herd size ( $P < 0.05$ ) and the interaction of herd size and season was found ( $P < 0.01$ ). Scores increased from summer to winter in C1 and C3, whereas they decreased in both

other classes. The lowest scores were found in C2 and the highest in C4. Additionally, a significant effect of group size was found ( $P < 0.01$ ). The larger the group size, the lower was the score of the criterion “absence of prolonged thirst.” Within the criterion “comfort around resting,” C1 and C3 were rated better in winter compared with summer, whereas no seasonal difference was found for C2 and C4 ( $P < 0.05$ , interaction effect). A significant effect of the season was determined for the criterion “absence of injuries” ( $P < 0.01$ ). In each herd size category, results were better in winter ( $29.7 \pm 1.3$ ) than in summer ( $21.9 \pm 1.3$ ). The prevalence of diseases within the dairy farms was not influenced by the examined effects ( $P > 0.05$ ). Within the criterion “expression of social behaviors,” social interactions were at a comparable level between the 4 herd size classes ( $P > 0.05$ ). For the criterion “human-animal-relationship,” no influence of the herd size, but an effect of the season was found ( $P < 0.05$ ). Averaged over both farm visits, the scores in winter were higher compared with summer, particularly in C3 and C4. The results of the qualitative behavior assessment, more specifically the related criterion “positive emotional state” was influenced by the interaction between herd size and season ( $P < 0.01$ ). Whereas the classes C1 ( $74.5 \pm 3.4$  points), C2 ( $78.5 \pm 3.0$  points), and C4 ( $78.1 \pm 3.4$  points) performed better in the winter season, C3 ( $77.8 \pm 3.0$  points) showed a higher amount of positive emotions in the summer season.

### Indicators

Results at indicator level are described in Table 5. For the indicator “percentage of very lean cows,” effects of herd size ( $P < 0.05$ ) and the herd size  $\times$  season interaction were found ( $P < 0.01$ ). The smaller farms of class C1 ( $13.9 \pm 1.6\%$ ) and C2 ( $16.0 \pm 1.4\%$ ) had a higher amount of lean cows than the larger farms

**Table 4.** Results of the Welfare Quality assessments at the level of principles and criteria in season 1 (summer) and season 2 (winter) for dairy farms with <100 cows (C1), 100–299 cows (C2), 300–499 cows (C3), or ≥500 cows (C4) expressed on a 0 to 100 value scale (0 = poor, 100 = good; n = 20)<sup>1</sup>

Welfare principle/criterion	Herd size class												P-value			
	C1			C2			C3			C4						
	S	W	PSE	S	W	PSE	S	W	PSE	S	W	PSE		Herd size (H)	Season (S)	Interaction (H × S)
Good feeding	30.9	33.5	20.2	24.8	48.5	51.0	45.2	5.7	<0.01	0.10	<0.01	<0.01	<0.01	0.70	0.23	0.20
Absence of prolonged hunger	40.5	42.4	37.4	40.1	53.6	54.9	49.0	3.7	0.02	0.70	0.23	0.20	<0.01	0.53	<0.01	<0.01
Absence of prolonged thirst	45.0	47.8	30.4	33.3	66.6	68.2	61.3	10.2	0.01	0.53	<0.01	<0.01	0.55	0.24	0.06	0.15
Good housing	55.0	58.9	60.1	60.2	64.2	59.9	59.8	2.0	0.11	0.24	0.55	0.06	0.02	NT	NT	NT
Comfort around resting	28.5	34.7	36.6	36.8	43.1	36.4	36.1	3.2	0.08	0.01	0.02	0.15	0.84	0.97	0.19	0.22
Ease of movement	100.0	100.0	100.0	100.0	100.0	100.0	100.0	0.0	0.75	<0.01	0.84	0.97	0.78	<0.01	0.22	0.19
Good health	21.1	23.1	23.3	20.7	25.3	20.9	24.4	1.3	0.31	<0.01	0.78	0.22	0.71	0.21	0.19	0.85
Absence of injuries	21.2	26.3	29.0	18.6	24.9	22.9	31.4	2.7	0.37	0.21	0.71	0.19	<0.01	0.54	0.60	0.20
Absence of diseases	25.1	24.3	22.8	25.1	23.6	20.5	20.7	1.7	0.39	<0.01	<0.01	0.60	0.27	0.73	0.12	0.20
Absence of pain	37.8	32.3	31.6	38.3	37.9	42.6	43.6	3.4	0.08	0.12	0.08	0.20	NT	NT	NT	NT
Appropriate behavior	36.3	37.9	39.6	39.1	36.8	38.1	37.2	1.2	0.46	0.03	0.31	0.41	<0.01	0.76	<0.01	0.53
Expression of social behaviors	77.0	82.2	84.6	85.7	87.8	90.0	86.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Expression of other behaviors	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Human-animal relationship	75.7	75.6	76.1	76.8	71.0	72.9	67.9	2.8	0.46	0.03	0.31	0.41	<0.01	0.76	<0.01	0.53
Positive emotional state	72.6	76.3	79.0	78.0	71.9	77.7	78.5	3.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

<sup>1</sup>S = summer; W = winter. PSE = pooled standard error; NT = not tested.

of class C3 (10.3 ± 1.4%) and C4 (10.5 ± 1.6%). A difference between the herd size classes could only be confirmed for C2 in winter, compared with C3 in winter and C4 in summer ( $P < 0.05$ ). Trough length per animal rose with increasing herd size ( $P < 0.01$ ). Over both seasons, C3 (7.8 ± 0.6 cm) and C4 (8.0 ± 0.7 cm) provided more water surface per animal compared with C1 (4.2 ± 0.6 cm) and C2 (5.0 ± 0.6 cm). Moreover, a significant effect of group size and season was determined. The larger the group size, the lower was the trough length per animal ( $P < 0.01$ ). Dairy cows had more water surface available in summer season (6.4 ± 0.3 cm), compared with the winter months (6.0 ± 0.3 cm;  $P < 0.05$ ). An influence of the herd size was found for the indicator “duration of the process of lying down in a stall” ( $P < 0.05$ ). On average, dairy cows of the smallest class C1 needed more time to lie down (6.0 ± 0.1 s), compared with the animals of the class C3 with 300 to 499 cows (5.5 ± 0.1 s). The indicator “percentage of cows lying outside cubicles” was influenced by season. The scores of all herd size classes were lower in winter than in summer ( $P < 0.05$ ). Furthermore, an influence of group size was observed for this indicator. The larger the group size, the lower was the amount of cows lying outside the cubicles ( $P < 0.01$ ). For the indicator “percentage of dirty animals (legs),” better scores were achieved in summer season, compared with winter season ( $P < 0.01$ ). Group size as well as herd size influenced the results significantly ( $P < 0.05$ ). The larger the group size, the higher was the percentage of cows with dirty lower hind legs. In class C3 (41.3 ± 5.3%), fewer animals with dirty legs compared with C1 (57.2 ± 6.0%) were found. With respect to the cleanliness of the body regions flank and udder, no differences were observed ( $P > 0.05$ ). The percentages of animals with lesions or swellings and severe lameness were different in summer and winter ( $P < 0.01$ ), but not between the herd size classes ( $P > 0.05$ ). Nasal discharge was found more often in summer compared with winter ( $P < 0.01$ ). This symptom occurred less often in C1 (<100 cows) than in the other classes ( $P < 0.05$ ). For the indicator “percentage of cows with ocular discharge,” an effect of season and interaction of season and herd size was found ( $P < 0.01$ ). The larger the group size, the higher the incidence rate of animals with diarrhea ( $P < 0.01$ ). In addition, a seasonal effect on the results of this indicator was determined ( $P < 0.01$ ). The percentage of dairy cows with mastitis and vulvar discharge was influenced by the interaction of herd size and season ( $P < 0.01$ ). The prevalence of vulvar discharge was lower in C1 in summer (1.0 ± 0.6%) compared with C4 in winter (2.9 ± 0.5%) and summer (3.5 ± 0.5%), respectively. For the remaining health indicators “coughing,” “dystocia,” “downer cows,” and “mortality,” no effects of any of

the tested effects were revealed ( $P > 0.05$ ). “Hampered respiration” was not observed at any of the assessments. Whereas head butts per animal and hour were not influenced by the herd size ( $P > 0.05$ ), the amount of displacements per animal and hour was influenced by herd size ( $P < 0.01$ ). Compared with the other herd size classes, more displacements were counted in C1. The number of animals with an avoidance distance of 0 or 1 to 50 cm was influenced by the interaction of herd size and season ( $P < 0.05$ ). In contrast, the number of animals with an avoidance distance of 50 to 100 cm was only affected by the season ( $P < 0.01$ ). More cows were observed in this category during winter ( $7.2 \pm 0.7\%$ ) compared with the summer season ( $5.3 \pm 0.7\%$ ).

## DISCUSSION

### *Limitations of the Study*

The objective of the present study was to examine the relationship between herd size and animal welfare. Therefore, 80 loose housing dairy cattle farms with a maximum of 6 h access to pasture per day were selected with different herd sizes, and animal welfare level was assessed using the WQP. This indicator system is used in several European working groups for animal welfare assessment in dairy cattle farms, but recent evaluations of the WQP revealed some methodological problems (de Graaf et al., 2017b). For example, de Vries et al. (2013) stated that single welfare indicators such as water provision or lean cows have a disproportionate effect on the overall classification, whereas other welfare-related health indicators (e.g., lameness or mastitis) were less important for the overall score. Challenges concerning the aggregation process were also described by other authors (Heath et al., 2014a,b; de Graaf et al., 2017b; Sandoe et al., 2017).

Furthermore, the current strategy of sampling dairy cows for clinical examinations (e.g., lameness, lean cows, and cleanliness) might influence the results and is recommended to be modified (Van Os et al., 2018). Seasonal effects on several animal welfare indicators were observed in the present study and should also be taken into account when comparing results of different studies (Heath et al., 2014a; de Graaf et al., 2017a). Finally, insufficient inter-observer reliabilities of single indicators of the WQP such as qualitative behavior assessment (Bokkers et al., 2012), avoidance distance test (de Graaf et al., 2017a), or cleanliness and integument alterations (Heath et al., 2014b) were determined. However, under consideration of the described limitations, the WQP can be a useful instrument for on-farm animal welfare assessment in research settings (Gieseke et al., 2014; Heath et al., 2014b).

Mean herd size of the sample (383 cows) was much larger than average herd size (61 cows) in Germany (Destatis, 2017). Nevertheless, under consideration of the herd size development in recent years, it is likely that the intensification of the milk production continues with further increasing herd sizes. It has to be mentioned that the thresholds of the herd size classes were selected in accordance with the farm structures in Germany. Definitions of small or large dairy cattle farms might be different in other countries. Therefore, results of the present study should be interpreted under consideration of the selected herd size classification. Moreover, ECM yields of the 2 lower herd size classes was about 500 kg lower compared with the 2 upper classes. An effect of the breed can, however, be excluded given that all farms raised Holstein Friesian cows. Other factors related to the feeding regimens or environmental conditions may have caused these differences (Bewley and Schutz, 2008; Gauly et al., 2013), but its effect could not be further specified in this study. Even though Coignard et al. (2014) did not find an association between milk yield and overall welfare as measured with the WQP in commercial French dairy herds, further studies on this relationship are warranted and the difference in milk yield might have influenced some of the welfare parameters in the present study. Several farm types (e.g., tiestalls, straw-bedded barns, farms with pasture access) were excluded in the present study. The distribution of specific housing conditions (e.g., cubicle types, flooring types, and stall climate) was not considered during farm acquisition because no statistical data on their proportions among German dairy cattle farms were available. In general, the housing characteristics of the dairy farms in the present study (see Table 1) can be considered as typical of herds in Central Europe of the sizes assessed. Different housing conditions were considered as random farm effects in the statistical model. Additionally, known effects of the housing system (barn, floor, and cubicle design) on different animal welfare indicators were integrated in the Discussion section.

### *Overall Assessment*

The WQP overall classification widely agrees with a recent study by de Graaf et al. (2017b), who analyzed a large data set of 491 dairy cattle farms from different European countries (e.g., Macedonia, Scotland, Denmark, and Austria). The authors found 0% “excellent,” 35% “enhanced,” 63% “acceptable,” and 2% “not classified” dairy cattle farms, which showed broad variations in terms of housing and management conditions (de Graaf et al., 2017b). Similar results were published by other working groups. Heath et al. (2014a)



**Table 5.** Results of the Welfare Quality assessments at the level of indicators in season 1 (summer) and season 2 (winter) for dairy farms with <100 cows (C1), 100–299 cows (C2), 300–499 cows (C3), or ≥500 cows (C4; n = 20)<sup>1</sup>

Welfare indicator	Herd size class												P-value				
	C1			C2			C3			C4							
	S	W	PSE	S	W	PSE	S	W	PSE	S	W	PSE		Herd size (H)	Season (S)	Interaction (H × S)	Group size (G)
<b>Good feeding</b>																	
Percentage of very lean animals	14.4	13.4	14.2	17.7	11.0	9.6	9.0	11.9	1.8	<0.05	0.10	<0.01	0.54				
Length of troughs/bowls per animal (cm)	4.2	4.2	5.3	4.6	8.1	7.5	8.2	7.9	0.6	<0.01	0.02	0.22	<0.01				
Mean cleanliness water points (0–2)	0.4	0.6	0.6	0.4	0.5	0.6	0.4	0.4	0.1	0.82	0.80	0.73	0.85				
<b>Good housing</b>																	
Duration of the process of lying down (s)	5.9	6.0	5.8	5.9	5.3	5.6	5.7	6.0	0.2	0.03	0.07	0.83	0.37				
Percentage of collisions with cubicles	25.2	14.4	18.4	21.8	20.0	9.6	17.5	19.0	3.7	0.26	<0.01	<0.01	0.15				
Percentage of cows lying outside cubicles	5.6	3.5	3.8	2.7	6.5	4.1	3.6	2.9	1.4	0.35	0.04	0.77	<0.01				
Percentage of dirty animals (legs)	51.4	62.9	48.9	60.0	36.9	45.6	48.9	54.5	6.2	0.03	<0.01	0.24	0.02				
Percentage of dirty animals (flanks)	62.2	63.4	68.2	70.4	62.0	63.9	64.0	68.8	5.0	0.69	0.14	0.83	0.73				
Percentage of dirty animals (udder)	45.4	47.3	50.9	51.0	34.1	35.7	39.8	44.2	5.3	0.09	0.67	0.89	0.38				
<b>Good health</b>																	
Percentage of cows with lesions/swellings	31.3	28.4	40.0	30.1	32.8	28.4	34.9	26.5	3.5	0.50	<0.01	0.62	0.80				
Percentage of cows with severe lameness	39.4	19.8	33.1	18.1	21.5	12.1	22.9	11.3	3.5	0.10	<0.01	0.89	0.16				
Percentage of cows with nasal discharge	17.1	15.4	24.6	18.4	26.3	19.5	27.0	19.4	2.0	0.02	<0.01	0.41	0.20				
Percentage of cows with ocular discharge	6.8	0.9	2.7	0.9	4.1	2.0	4.9	1.5	0.9	0.14	<0.01	<0.01	0.91				
Frequency of coughing per cow/15 min	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.1	0.0	0.14	0.62	0.37	0.60				
Percentage of cows with diarrhea	42.8	36.2	40.1	33.6	33.2	29.5	29.4	25.3	3.6	0.16	<0.01	0.92	<0.01				
Percentage of cows with mastitis	20.6	16.8	19.7	21.4	21.1	22.4	22.5	24.7	2.0	0.31	0.78	0.01	0.67				
Percentage of cows with vulvar discharge	1.0	2.2	1.8	2.3	1.7	2.5	3.5	2.9	0.5	<0.05	<0.01	0.01	0.80				
Percentage of dystocia (12 mo)	6.1	7.6	6.6	6.3	5.3	6.0	7.8	6.4	1.7	0.41	0.72	0.91	0.39				
Percentage of downer cows (12 mo)	4.5	4.4	3.1	4.3	3.5	3.6	2.7	3.0	0.7	0.72	0.05	0.91	1.00				
Percentage of mortality (12 mo)	3.6	3.3	4.5	4.5	5.2	5.5	5.8	6.3	0.8	0.38	0.73	0.26	0.33				
<b>Appropriate behavior</b>																	
Head butts per animal/60 min	0.5	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.1	0.69	0.97	0.57	0.62				
Displacements per animal/60 min	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.0	<0.01	0.02	0.31	0.07				
Avoidance distance test 0 cm (%)	53.0	56.6	54.7	54.7	47.9	43.5	45.2	38.3	4.5	0.16	0.21	0.04	0.35				
Avoidance distance test 1 to 49 cm (%)	40.9	35.6	40.0	38.9	47.6	48.6	47.6	51.2	3.6	0.07	0.39	0.02	0.20				
Avoidance distance test 50 to 100 cm (%)	5.8	7.4	5.1	5.6	3.9	6.9	6.1	9.0	1.5	0.55	<0.01	0.11	0.86				

<sup>1</sup>S = summer; W = winter. PSE = pooled standard error.

observed 62% “enhanced” and 38% “acceptable” dairy cattle farms in a British study, whereas De Boyer de Roches et al. (2014) reported 37% “enhanced,” 58% “acceptable,” and 5% “not classified” farms in France. Popescu et al. (2014) assessed loose houses and tiestalls in Romania and classified 43% of the farms as “enhanced” and 47% as “acceptable.” Every 10th farm was “not classified.” As in the present study, none of the aforementioned studies found a farm being classified as “excellent.” Comparisons to other studies regarding the effect of herd size on the overall WQP classification are not possible because the herd size was never considered as influencing factor.

### **Principle of “Good Feeding”**

Results of the present study with a mean of 12.7% lean cows (9.0 to 17.0%) are in accordance with other studies. For example, Popescu et al. (2014) observed on average 13.1% cows with a low BCS in loose houses and 10.2% in tiestalls. Zuliani et al. (2017) observed 18.3% in Italian mountain farms. Heath et al. (2014a) found on average 5.7% in a British study, whereas de Graaf et al. (2017a) observed 5.5% in Belgium. The highest mean percentages of lean cows in the present study were found in C2, which increased from summer to winter, whereas the lowest percentages were found in C3 and C4 with lower seasonal differences. Because feeding management (TMR) did not differ in the farms with limited access to pasture between seasons and seasonal calving was not practiced both can be excluded as explanatory factors. Herd size effects were also stated by other studies. Adams et al. (2017) compared small (<100 cows), medium (100–499 cows), and large ( $\geq 500$  cows) dairy farms. The highest within-herd prevalence of cows with a BCS  $\leq 2.5$  was observed in small (9.1%) compared with medium (3.0%) and large (2.0%) farms (Adams et al., 2017). An association of lower percentages of lean cows with increasing herd size was also observed in a Dutch study (de Vries et al., 2016). The advantage of larger herds might depend on the establishment of feeding groups. Different rations may be provided, which are adjusted for specific feeding requirements in particular lactation stages (Bewley and Schutz, 2008; Adams et al., 2017).

The mean length of troughs per cow in the present study was significantly higher in larger herds (>300 cows). Both classes C3 and C4 provided a 7.5 cm trough length per cow, which is sufficient according to the WQP (threshold:  $\geq 6$  cm). Contrastingly, classes with smaller herd sizes C1 and C2 ranged below this threshold (4 to 6 cm; partly sufficient). In other studies, comparable scores for sufficient (43%), partly sufficient (35%), and not sufficient (22%) water provision were

found (de Vries et al., 2013; Heath et al., 2014b; De Boyer de Roches et al., 2014). In summer, most farms were scored better than in winter. This discrepancy may be explained by installation of additional troughs in the barns during summer and a noticeable number of troughs that were broken due to frost in winter. Dairy cows in a Belgian study had on average higher scores in the criterion “absence of prolonged thirst” at the end (80 points), compared with the beginning of the indoor period (60 points). Due to the high variability within the groups, no statistical effect was found (de Graaf et al., 2017a).

### **Principle of “Good Housing”**

The average time needed to lie down ranged in between 5.3 and 6.0 s, depending on class and season. Comparable values were measured, for example, in France with 5.9 s (De Boyer de Roches et al., 2014), United Kingdom with 5.2 s (Heath et al., 2014a), and Denmark with 6.0 s (Andreasen et al., 2014), with a wide range between the farms (3.1 to 10.7 s). Variances of the duration of the process of lying down might be caused by different housing conditions (Plesch et al., 2010). Farms of C3 achieved lower mean durations than farms of C1. These findings could be partly explained by the higher amount of 90% deep-bedded cubicles in C3 compared with 65% in C1. Deep-bedded ones are more comfortable for the dairy cows and might reduce the time needed to lie down (Wechsler et al., 2000). Inadequate cubicle dimensions could also be detrimental for the lying down process because dairy cows might be disturbed by the cubicle partitions (Veisier et al., 2004). Furthermore, painful conditions can influence the time to lie down. For example, Popescu et al. (2013) determined strong correlations between the duration of the process of lying down and cows with lameness, lesions, or mastitis ( $P < 0.05$ ). However, these animal welfare indicators did not differ between herd size classes in the present study.

In most studies, alarm thresholds of the WQP for dirtiness of lower hind legs (50%) and hindquarters as well as udders (20%) were widely exceeded. For example, Heath et al. (2014a) found more than 50% dirty udders (2 to 98%). Comparable results for soiling of the hindquarters and lower hind legs were reported by other studies (De Boyer de Roches et al., 2014; Heath et al., 2014a; Zuliani et al., 2017). The percentage of dirty lower legs was influenced by herd size in the present study, but not the percentages of dirty hindquarters or udders. The discrepancy of 57.2% dirty lower legs in C1 and 41.2% in C3 was unexpected. Farms of C3 had a higher proportion of plain floors (75%), whereas dairy farms of C1 had a higher proportion of slatted floors

(95%). Plain floors are usually associated with severe contaminations of the lower legs due to accumulated manure in front of the automatic scraper (Cramer et al., 2009). Probably, higher scraping frequencies in the farms of the present study have reduced the amount of manure in the alleys in comparison to other studies (DeVries et al., 2012). The observed results might also be explained by the higher percentage of deep bedded cubicles in C3 (90%) compared with C1 (65%) because these have a higher absorptive capacity of the bedding material (de Vries et al., 2015; Cook et al., 2016). In every herd size class a higher prevalence of dirty lower legs was found in winter compared with summer. Walking areas are usually more humid in winter because the floors do not dry off. Therefore, dairy cows are at a higher risk of being dirty at the lower legs during winter (Ruud et al., 2010).

### **Principle of “Good Health”**

In the present study severe integument alterations (lesions/swellings) were assessed at 31.6% of the animals, whereas herd sizes did not differ but alterations were observed more often in summer than in winter. These results comply with the mean prevalence of 39.2% in French (Coignard et al., 2013), 37.6% in Dutch (de Vries et al., 2013), and 29.8% in British (Heath et al., 2014a) dairy cattle farms. Risk factors for lesions and swellings are cubicles with mattresses, inadequate dimensions of feeding racks or cubicles, low BCS, and high milk yields (Coignard et al., 2013; Zaffino Heyerhoff et al., 2014). Summer pasturing positively affects the prevalence of integument alterations (Haskell et al., 2006; de Graaf et al., 2017a) because hairless patches, lesions, and swellings could recover on the pasture due to the softer underground (Winckler et al., 2007). In contrast, farms of the present study with a maximum of 6 h access to pasture per day achieved constantly higher integument scores in summer (34.8%) than in winter (28.3%). Conceivably, cleanliness of the dairy cows might have contributed to the observed differences in the assessment of lesions and swellings. Dairy cows of the present study were less dirty in summer and could therefore be examined more easily compared with soiled animals in winter. Smaller lesions might have been invisible beneath large-scale plaques of dirt.

The mean prevalence of severely lame cows (22.3%) in the present study was higher compared with other studies using the WQP. A British study detected 4.9% (0.0 to 47.6%) severely lame cows (Heath et al., 2014a), whereas de Vries et al. (2013) found 5.0% (0.0 to 65.9%) in the Netherlands. The discrepancy may be explained by the specific combination of risk factors in the study

design (Dippel et al., 2009; Solano et al., 2015). Lameness is observed more often in cubicle housing systems compared with straw barns and tiestalls, which might be provoked by permanent contact with soiled alleys (Haskell et al., 2006; Coignard et al., 2013; Popescu et al., 2014). Methodological challenges of lameness detection might also have contributed to lower lameness prevalences in tiestalls (Palacio et al., 2017). However, the most important effect for preventing lameness is access to pasture during summer season (Cook et al., 2016; de Graaf et al., 2017a). Contrastingly, higher percentages of severely lame cows were observed in summer (29.7%) compared with the winter season (15.3%). Under heat stress conditions, dairy cows are at a higher risk of developing claw disorders due to prolonged standing times in soiled alleys (Cook et al., 2004; Sanders et al., 2009). No association of herd size and lameness was found in the present study, which is in accordance with other authors (Barker et al., 2010; Fabian et al., 2014). In contrast, Alban (1995) and de Vries et al. (2014) determined a positive relationship between herd size and lameness in dairy cattle. Increasing herd size was associated with intensive mechanization and less attention for the single cow (Alban, 1995). Other authors stated a negative relationship because professionalized management (trained staff, regular footbaths) is provided more frequently in larger than in smaller herds (Dippel et al., 2009; Solano et al., 2015).

Different information on prevalence of nasal discharge was found in the literature. Some authors detected only 3.7% (Heath et al., 2014a) or 4.2% (Zuliani et al., 2017) cows with nasal discharge. Other authors described higher proportions of cows with this symptom in France (16.4%; Coignard et al., 2013) and Belgium (17.1%; de Graaf et al., 2017a), which were at a comparable level to the findings of the present study (21.0%). Increasing herd size was associated with an increased frequency of nasal discharge. Conceivably, this result is attributed to a higher infection risk in larger herds due to intensified contact to other animals and several regroupings during lactation (Torres-Cardona et al., 2014; Beggs et al., 2015). All herd size classes had higher proportions of cows with nasal discharge in summer (23.7%) compared with winter (18.2%). Dairy cows have an impaired immune status under heat stress conditions (Kadzere et al., 2002), which might have increased the risk of nasal discharge due to viral or bacterial infections in summer (Canali et al., 2009). Nasal discharge as an unspecific symptom of respiratory disorders in cattle can also be caused by environmental factors such as the dust concentration in the barn (Brscic et al., 2012). However, a previous study did not determine different dust exposures in dairy cattle farms between the sum-

mer and winter season (Takei et al., 1998). Therefore, the seasonal effect on nasal discharge observed in the present study still remains unclear.

On average, 21% of the dairy cows showed signs of mastitis, determined as SCC >400,000 cells at least once in the last 3 mo (Welfare Quality, 2012). The alarm threshold of the WQP was exceeded by most of the farms, which widely agrees with previous studies. Andreasen et al. (2014) described that 11 farms have surpassed the warning (8.8%) and 29 farms the alarm threshold (17.5%) out of 44 farms in Denmark. Coignard et al. (2013) observed on average 20.6% dairy cows with mastitis (2.0 to 46.6%). Lower incidence rates were found with 11.1% affected cows in the Netherlands (de Vries et al., 2013) and 15.5% in the United Kingdom (Heath et al., 2014a). No significant effect of herd size on the mastitis incidence was examined in the present study. These findings were supported by other studies (Ivemeyer et al., 2011; Schewe et al., 2015). Contrastingly, some authors observed a higher risk of mastitis with increasing herd size (Lievaart et al., 2007; Archer et al., 2013), whereas other authors reported a lower risk (Oleggini et al., 2001; Ingham et al., 2011). The contradictory results found in the literature might be caused by differences within the study design. Regarding herd size, milk yield, or dominant breed, only the US studies were comparable to our own study population. Oleggini et al. (2001) examined the effect of several herd size classes (20 to 49, 50 to 99, 100 to 149, 150 to 249, 250 to 449,  $\geq 450$  cows) on different dairy herd performance parameters. Ingham et al. (2011) compared SCC in 3 herd size categories ( $\leq 118$ ; 119 to 713;  $\geq 714$  cows). In both publications, lower SCC with increasing herd size were determined.

In several European countries mean percentages of 0.4 to 1.0% dairy cows with vulvar discharge were found (Coignard et al., 2013; Popescu et al., 2014; Heath et al., 2014a). Results of the present study were slightly higher (2.3%) and increased continuously with increasing herd size. Average percentages in herd size classes C1 to C3 were below the warning threshold (2.3%), whereas the mean value of C4 was above (Welfare Quality, 2012). The larger farms of C4 ( $\geq 500$  cows) had frequently implemented a separate group for cows in the early postpartum period. Most of the dairy cows in these groups showed signs of vulvar discharge due to physiological cleaning processes of the uterus in the first weeks of lactation. The probability to include a cow with vulvar discharge in the sample was therefore higher in larger compared with smaller herds without different lactation groups. Fourichon et al. (2001) also examined the effect of herd size on the incidence of metritis and detected a higher percentage of affected cows in larger herds. In contrast, no relationship be-

tween herd size and metritis incidence was observed in Denmark (Bruun et al., 2002).

### ***Principle of “Appropriate Behavior”***

Cows in larger herds are usually confronted with regular regrouping because larger herds are often divided into smaller subgroups (Boe and Faerevik, 2003). Frequent regrouping might lead to increased agonistic interactions in the groups, due to increasing conflicts between unfamiliar cows (Boe and Faerevik, 2003; Estevez et al., 2007). The results of the present study did not support this hypothesis. A significant effect of herd size was determined, but most frequent conflicts were observed in the smallest class C1. The highest amount of displacements was associated with less space per cow in small walking alleys and conflicts concerning limited resources like water troughs, concentrate feeders, or cow brushes. Similarly, de Vries et al. (2015) found a positive relationship between the presence of cow brushes and the number of displacements. However, in all herd size classes displacements were at a relatively low level (0.1 to 0.2 displacements), compared with the value of 0.4 displacements found in the Netherlands (de Vries et al., 2015), in Belgium (de Graaf et al., 2017a), and in the United Kingdom (Heath et al., 2014a).

In larger herds different stock persons take care of a higher amount of animals and the human-animal-relationship might be less pronounced (Raussi, 2003). However, no negative association of herd size and avoidance distance test at the feed rack was found in the present study. Mattiello et al. (2009) and Verkerk and Hemsworth (2010) determined higher flight distances with increasing herd sizes, whereas other working groups could not verify such a relationship (Waiblinger and Menke, 1999; Waiblinger et al., 2003). The scores of the criterion “human-animal-relationship” in the present study (74 points) were comparable to 68.0 points in the Netherlands (de Vries et al., 2013) and 78.7 points in Italy (Zuliani et al., 2017). Lower scores were determined with 42.0 points in France (De Boyer de Roches et al., 2014). The observed seasonal effect with higher flight distances in winter (especially for C3 and C4) could not be explained by the collected data. Battini et al. (2011) did also find varying human-animal-relationships throughout the year, but this was attributed to management changes due to alpine summer pasturing. Contrastingly, de Graaf et al. (2017a) determined comparable human-animal-relationship scores at the beginning (35.7 points) and end (36.4 points) of the indoor period. Results of the avoidance distance tests are primarily influenced by quantity and quality of individual interactions between dairy cows and farmers (Waiblinger et al., 2003). Seasonal differ-



ences in the human-animal-relationship are warranted to be investigated in more detail in further studies.

## CONCLUSIONS

Findings of the present study did not indicate a linear relationship between herd size and animal welfare in dairy cattle. Single animal welfare indicators (lean cows, trough length, duration of the process of lying down in a stall, dirtiness of lower legs, nasal discharge, vulvar discharge, and displacements) were associated with the number of dairy cows per farm. But the observed results were not consistent, as some welfare indicators worsened with increasing herd size, whereas others improved or showed a curvilinear relationship. Large variations of welfare levels between farms were observed in each herd size class. Therefore, herd size is not a valid indicator of animal welfare at the herd level. Housing conditions and management practices seem to have a greater effect on animal welfare than the number of dairy cows per farm.

## ACKNOWLEDGMENTS

This study was funded by the German Federal Ministry of Education and Research (Berlin) within the project “Sustainable Land Management in North German Lowland” (NaLaMa-nt) and financially supported by the H. Wilhelm Schaumann Foundation (Hamburg, Germany). The authors thank the different stakeholders for their support during farm acquisition and the participating farmers for their time and effort.

## REFERENCES

- Adams, A. E., J. E. Lombard, C. P. Fossler, I. N. Román-Muñoz, and C. A. Koprál. 2017. Associations between housing and management practices and the prevalence of lameness, hock lesions, and thin cows on US dairy operations. *J. Dairy Sci.* 100:2119–2136. <https://doi.org/10.3168/jds.2016-11517>.
- Alban, L. 1995. Lameness in Danish dairy cows: Frequency and possible risk factors. *Prev. Vet. Med.* 22:213–225. [https://doi.org/10.1016/0167-5877\(94\)00411-B](https://doi.org/10.1016/0167-5877(94)00411-B).
- Andreasen, S. N., P. Sandoe, and B. Forkman. 2014. Can animal-based welfare assessment be simplified? A comparison of the Welfare Quality<sup>®</sup> protocol for dairy cattle and the simpler and less time-consuming protocol developed by the Danish Cattle Federation. *Anim. Welf.* 23:81–94. <https://doi.org/10.7120/09627286.23.1.081>.
- Archer, S. C., F. Mc Coy, W. Wapenaar, and M. J. Green. 2013. Association of season and herd size with somatic cell count for cows in Irish, English, and Welsh dairy herds. *Vet. J.* 196:515–521. <https://doi.org/10.1016/j.tvjl.2012.12.004>.
- Barker, Z. E., K. A. Leach, H. R. Whay, N. J. Bell, and D. C. J. Main. 2010. Assessment of lameness prevalence and associated risk factors in dairy herds in England and Wales. *J. Dairy Sci.* 93:932–941. <https://doi.org/10.3168/jds.2009-2309>.
- Battini, M., E. Andreoli, S. Barbieri, and S. Mattiello. 2011. Long-term stability of avoidance distance tests for on-farm assessment of dairy cow relationship to humans in alpine traditional husbandry systems. *Appl. Anim. Behav. Sci.* 135:267–270. <https://doi.org/10.1016/j.applanim.2011.10.013>.
- Beggs, D. S., A. D. Fisher, E. C. Jongman, and P. H. Hemsworth. 2015. A survey of Australian dairy farmers to investigate animal welfare risks associated with increasing scale of production. *J. Dairy Sci.* 98:5330–5338. <https://doi.org/10.3168/jds.2014-9239>.
- Bewley, J. M., and M. M. Schutz. 2008. Review: An interdisciplinary review of body condition scoring for dairy cattle. *Prof. Anim. Sci.* 24:507–529. [https://doi.org/10.15232/S1080-7446\(15\)30901-3](https://doi.org/10.15232/S1080-7446(15)30901-3).
- Boe, K. E., and G. Faerevik. 2003. Grouping and social preferences in calves, heifers and cows. *Appl. Anim. Behav. Sci.* 80:175–190. [https://doi.org/10.1016/S0168-1591\(02\)00217-4](https://doi.org/10.1016/S0168-1591(02)00217-4).
- Bokkers, E. A. M., M. de Vries, I. C. M. A. Antonissen, and I. J. M. de Boer. 2012. Inter- and intra-observer reliability of experienced and inexperienced observers for the Qualitative Behaviour Assessment in dairy cattle. *Anim. Welf.* 21:307–318. <https://doi.org/10.7120/09627286.21.3.307>.
- Brcsic, M., H. Leruste, L. F. M. Heutinck, E. A. M. Bokkers, M. Wolthuis-Fillerup, N. Stockhofe, F. Gottardo, B. J. Lensink, G. Cozzi, and C. G. Van Reenen. 2012. Prevalence of respiratory disorders in veal calves and potential risk factors. *J. Dairy Sci.* 95:2753–2764. <https://doi.org/10.3168/jds.2011-4699>.
- Bruun, J., A. K. Ersbøll, and L. Alban. 2002. Risk factors for metritis in Danish dairy cows. *Prev. Vet. Med.* 54:179–190. [https://doi.org/10.1016/S0167-5877\(02\)00026-0](https://doi.org/10.1016/S0167-5877(02)00026-0).
- Canali, E., H. R. Whay, and K. A. Leach. 2009. Cattle health status. Pages 77–88 in *Assessment of Animal Welfare Measures for Dairy Cattle, Beef Bulls and Veal Calves*. B. Forkman, and L. Keeling, ed. Welfare Quality Reports No. 11.
- Coignard, M., R. Guatteo, I. Veissier, A. de Boyer des Roches, L. Mounier, A. Lehébel, and N. Bareille. 2013. Description and factors of variation of the overall health score in French dairy cattle herds using the Welfare Quality<sup>®</sup> Assessment protocol. *Prev. Vet. Med.* 112:296–308. <https://doi.org/10.1016/j.prevetmed.2013.07.018>.
- Coignard, M., R. Guatteo, I. Veissier, A. Lehébel, C. Hoogveld, L. Mounier, and N. Bareille. 2014. Does milk yield reflect the level of welfare in dairy herds? *Vet. J.* 199:184–187. <https://doi.org/10.1016/j.tvjl.2013.10.011>.
- Cook, N. B., T. B. Bennett, and K. V. Nordlund. 2004. Effect of free stall surface on daily activity patterns in dairy cows with relevance to lameness prevalence. *J. Dairy Sci.* 87:2912–2922. [https://doi.org/10.3168/jds.S0022-0302\(04\)73422-0](https://doi.org/10.3168/jds.S0022-0302(04)73422-0).
- Cook, N. B., J. P. Hess, M. R. Foy, T. B. Bennett, and R. L. Brotzman. 2016. Management characteristics, lameness, and body injuries of dairy cattle housed in high-performance dairy herds in Wisconsin. *J. Dairy Sci.* 99:5879–5891. <https://doi.org/10.3168/jds.2016-10956>.
- Cramer, G., K. D. Lissemore, C. L. Guard, K. E. Leslie, and D. F. Kelton. 2009. Herd-level risk factors for seven different foot lesions in Ontario Holstein cattle housed in tie stalls or free stalls. *J. Dairy Sci.* 92:1404–1411. <https://doi.org/10.3168/jds.2008-1134>.
- Dairy Australia. 2015. Cows and farms. Accessed Feb. 10, 2017. [www.dairyaustralia.com.au/Markets-and-statistics/Farm-facts/Cows-and-Farms.aspx](http://www.dairyaustralia.com.au/Markets-and-statistics/Farm-facts/Cows-and-Farms.aspx).
- Dairy New Zealand. 2014. New Zealand Dairy Statistics 2013–2014. Accessed Feb. 10, 2017. [www.dairynz.co.nz/media/1327583/nz-dairy-statistics-2013-2014-web.pdf](http://www.dairynz.co.nz/media/1327583/nz-dairy-statistics-2013-2014-web.pdf).
- De Boyer des Roches, A., I. Veissier, M. Coignard, N. Bareille, R. Guatteo, J. Capdeville, E. Gilot-Fromont, and L. Mounier. 2014. The major welfare problems of dairy cows in French commercial farms: An epidemiological approach. *Anim. Welf.* 23:467–478. <https://doi.org/10.7120/09627286.23.4.467>.
- de Graaf, S., B. Ampe, and F. A. M. Tuytens. 2017a. Assessing dairy cow welfare at the beginning and end of the indoor period using the Welfare Quality<sup>®</sup> protocol. *Anim. Welf.* 26:213–221. <https://doi.org/10.7120/09627286.26.2.213>.
- de Graaf, S., B. Ampe, C. Winckler, M. Radeski, L. Mounier, M. K. Kirchner, M. J. Haskell, F. J. C. M. van Eerdenburg, A. de Boyer des Roches, S. N. Andreasen, J. Bijttebier, L. Lauwers, W. Verbeke, and F. A. M. Tuytens. 2017b. Trained-user opinion about

- Welfare Quality measures and integrated scoring of dairy cattle welfare. *J. Dairy Sci.* 100:6376–6388. <https://doi.org/10.3168/jds.2016-12255>.
- Destatis. 2017. Land- und Forstwirtschaft, Fischerei Rinder- und Schweinebestand Statistisches Bundesamt. Fachserie 3 Reihe 4.1. Accessed Jun. 7, 2017. [www.destatis.de/DE/Publikationen/Thematisch/LandForstwirtschaft/ViehbestandTierischeErzeugung/Viehhaltung2030213169004.pdf?\\_blob=publicationFile](http://www.destatis.de/DE/Publikationen/Thematisch/LandForstwirtschaft/ViehbestandTierischeErzeugung/Viehhaltung2030213169004.pdf?_blob=publicationFile).
- de Vries, M., E. A. M. Bokkers, C. G. van Reenen, B. Engel, G. van Schaik, T. Dijkstra, and I. J. M. de Boer. 2015. Housing and management factors associated with indicators of dairy cattle welfare. *Prev. Vet. Med.* 118:80–92. <https://doi.org/10.1016/j.prevetmed.2014.11.016>.
- de Vries, M., E. A. M. Bokkers, G. van Schaik, R. Botreau, B. Engel, T. Dijkstra, and I. J. M. de Boer. 2013. Evaluating results of the Welfare Quality multi-criteria evaluation model for classification of dairy cattle welfare at the herd level. *J. Dairy Sci.* 96:6264–6273. <https://doi.org/10.3168/jds.2012-6129>.
- de Vries, M., E. A. M. Bokkers, G. van Schaik, B. Engel, T. Dijkstra, and I. J. M. de Boer. 2016. Exploring the value of routinely collected herd data for estimating dairy cattle welfare. *J. Dairy Sci.* 97:715–730. <https://doi.org/10.3168/jds.2013-6585>.
- de Vries, M., E. A. M. Bokkers, G. van Schaik, B. Engel, T. Dijkstra, and I. J. M. de Boer. 2016. Improving the time efficiency of identifying dairy herds with poorer welfare in a population. *J. Dairy Sci.* 99:8282–8296. <https://doi.org/10.3168/jds.2015-9979>.
- DeVries, T. J., M. G. Aarnoudse, H. W. Barkema, K. E. Leslie, and M. A. G. von Keyserlingk. 2012. Associations of dairy cow behavior, barn hygiene, cow hygiene, and risk of elevated somatic cell count. *J. Dairy Sci.* 95:5730–5739. <https://doi.org/10.3168/jds.2012-5375>.
- Dippel, S., M. Dolezal, C. Brenninkmeyer, J. Brinkmann, S. March, U. Knierim, and C. Winckler. 2009. Risk factors for lameness in cubicle housed Austrian Simmental dairy cows. *Prev. Vet. Med.* 90:102–112. <https://doi.org/10.1016/j.prevetmed.2009.03.014>.
- Estevez, I., I. L. Andersen, and E. Naevdal. 2007. Group size, density and social dynamics in farm animals. *Appl. Anim. Behav. Sci.* 103:185–204.
- European Commission. 2016. Attitudes of Europeans towards animal welfare. Special Eurobarometer 442. Accessed May 5, 2017. <http://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/2096>.
- Eurostat. 2015. Milk and milk products—30 years of quotas. Historical data on the milk sector (1983–2013). Accessed May 10, 2017. [http://ec.europa.eu/eurostat/statistics-explained/index.php/Milk\\_and\\_milk\\_products\\_-\\_30\\_years\\_of\\_quotas](http://ec.europa.eu/eurostat/statistics-explained/index.php/Milk_and_milk_products_-_30_years_of_quotas).
- Fabian, J., R. A. Laven, and H. R. Why. 2014. The prevalence of lameness on New Zealand dairy farms: A comparison of farmer estimate and locomotion scoring. *Vet. J.* 201:31–38. <https://doi.org/10.1016/j.tvjl.2014.05.011>.
- Fourichon, C., F. Beaudeau, N. Bareille, and H. Seegers. 2001. Incidence of health disorders in dairy farming systems in western France. *Livest. Prod. Sci.* 68:157–170. [https://doi.org/10.1016/S0301-6226\(00\)00249-9](https://doi.org/10.1016/S0301-6226(00)00249-9).
- Gauly, M., H. Bollwein, G. Breves, K. Brüggeman, S. Dänicke, G. Das, J. Demeler, H. Hansen, J. Isselstein, S. König, M. Lohölter, M. Martinsohn, U. Meyer, M. Potthoff, C. Sanker, B. Schröder, N. Wrage, B. Meibaum, G. von Samson-Himmelstjerna, H. Stinshoff, and C. Wrenzycki. 2013. Future consequences and challenges for dairy cow production systems arising from climate change in Central Europe—A review. *Animal* 7:843–859. <https://doi.org/10.1017/S1751731112002352>.
- Gieseke, D., C. Lambertz, I. Traulsen, J. Krieter, and M. Gauly. 2014. Beurteilung von Tiergerechtigkeit in der Milchviehhaltung – Evaluierung des Welfare Quality® Protokolls. *Züchtungskunde* 86:58–70.
- Haskell, M. J., L. J. Rennie, V. A. Howell, M. J. Bell, and A. B. Lawrence. 2006. Housing system, milk production, and zero-grazing effects on lameness and leg injury in dairy cows. *J. Dairy Sci.* 89:4259–4266. [https://doi.org/10.3168/jds.S0022-0302\(06\)72472-9](https://doi.org/10.3168/jds.S0022-0302(06)72472-9).
- Heath, C. A. E., W. J. Browne, S. Mullan, and D. C. J. Main. 2014b. Navigating the iceberg: Reducing the number of parameters within the Welfare Quality® Assessment protocol for dairy cows. *Animal* 8:1978–1986. <https://doi.org/10.1017/S1751731114002018>.
- Heath, C. A. E., Y. Lin, S. Mullan, W. J. Browne, and D. C. J. Main. 2014a. Implementing Welfare Quality® in UK assurance schemes: Evaluating the challenges. *Anim. Welf.* 23:95–107. <https://doi.org/10.7120/09627286.23.1.095>.
- Ingham, S. C., Y. Hu, and C. Ané. 2011. Comparison of bulk-tank standard plate count and somatic cell count for Wisconsin dairy farms in three size categories. *J. Dairy Sci.* 94:4237–4241. <https://doi.org/10.3168/jds.2011-4310>.
- Ivemeyer, S., U. Knierim, and S. Waiblinger. 2011. Effect of human-animal relationship and management on udder health in Swiss dairy herds. *J. Dairy Sci.* 94:5890–5902. <https://doi.org/10.3168/jds.2010-4048>.
- Kadzere, C. T., M. R. Murphy, N. Silanikove, and E. Maltz. 2002. Heat stress in lactating dairy cows: A review. *Livest. Prod. Sci.* 77:59–91. [https://doi.org/10.1016/S0301-6226\(01\)00330-X](https://doi.org/10.1016/S0301-6226(01)00330-X).
- Krystallis, A., M. Dutra de Barcellos, J. O. Kügler, W. Verbeke, and K. G. Grunert. 2009. Attitudes of European citizens towards pig production systems. *Livest. Sci.* 126:46–56. <https://doi.org/10.1016/j.livsci.2009.05.016>.
- Lievaart, J. J., H. W. Barkema, W. D. J. Kremer, J. van den Broek, J. H. M. Verheijden, and J. A. P. Heesterbeek. 2007. Effect of herd characteristics, management practices, and season on different categories of the herd somatic cell count. *J. Dairy Sci.* 90:4137–4144. <https://doi.org/10.3168/jds.2006-847>.
- MacDonald, J. M., E. J. O'Donoghue, W. D. McBride, R. F. Nehring, C. L. Sandretto, and R. Mosheim. 2007. Profits, costs, and the changing structure of dairy farming. Accessed Apr. 10, 2017. [https://www.ers.usda.gov/webdocs/publications/45868/11138\\_err47\\_1\\_.pdf?v=41746](https://www.ers.usda.gov/webdocs/publications/45868/11138_err47_1_.pdf?v=41746).
- Mattiello, S., C. Klotz, D. Baroli, M. Minero, V. Ferrante, and E. Canali. 2009. Welfare problems in alpine dairy cattle farms in Alto Adige (Eastern Italian Alps). *Ital. J. Anim. Sci.* 8:628–630. <https://doi.org/10.4081/ijas.2009.s2.628>.
- Oleggini, G. H., L. O. Ely, and J. W. Smith. 2001. Effect of region and herd size on dairy herd performance parameters. *J. Dairy Sci.* 84:1044–1050.
- Palacio, S., L. Peignier, C. Pachoud, C. Nash, S. Adam, R. Bergeron, D. Pellerin, A. M. de Passillé, J. Rushen, D. Haley, T. J. DeVries, and E. Vasseur. 2017. Technical Note: Assessing lameness in tie-stalls using live stall lameness scoring. *J. Dairy Sci.* 100:6577–6582. <https://doi.org/10.3168/jds.2016-12171>.
- Plesch, G., N. Broerkens, S. Laister, C. Winckler, and U. Knierim. 2010. Reliability and feasibility of selected measures concerning resting behaviour for the on-farm welfare assessment in dairy cows. *Appl. Anim. Behav. Sci.* 126:19–26. <https://doi.org/10.1016/j.applanim.2010.05.003>.
- Popescu, S., C. Borda, E. A. Diugan, M. Niculae, R. Stefan, and C. D. Sandru. 2014. The effect of the housing system on the welfare quality of dairy cows. *Ital. J. Anim. Sci.* 13:2940. <https://doi.org/10.4081/ijas.2014.2940>.
- Popescu, S., C. Borda, E. A. Diugan, M. Spinu, I. S. Groza, and C. D. Sandru. 2013. Dairy cows welfare quality in tie-stall housing system with or without access to exercise. *Acta Vet. Scand.* 55:43–54. <https://doi.org/10.1186/1751-0147-55-43>.
- Prickett, R. W., F. B. Norwood, and J. L. Lusk. 2010. Consumer preferences for farm animal welfare: Results from a telephone survey of US households. *Anim. Welf.* 19: 335–347.
- Raussi, S. 2003. Human-cattle interactions in group housing. *Appl. Anim. Behav. Sci.* 80:245–262. [https://doi.org/10.1016/S0168-1591\(02\)00213-7](https://doi.org/10.1016/S0168-1591(02)00213-7).
- Robbins, J. A., M. A. G. von Keyserlingk, D. Fraser, and D. M. Weary. 2016. Invited review: Farm size and animal welfare. *J. Anim. Sci.* 94:5439–5455. <https://doi.org/10.2527/jas.2016-0805>.
- Ruud, L. E., K. E. Boe, and O. Osteras. 2010. Risk factors for dirty dairy cows in Norwegian freestall systems. *J. Dairy Sci.* 93:5216–5224. <https://doi.org/10.3168/jds.2010-3321>.

- Sanders, A. H., J. K. Shearer, and A. de Vries. 2009. Seasonal incidence of lameness and risk factors associated with thin soles, white line disease, ulcers, and sole punctures in dairy cattle. *J. Dairy Sci.* 92:3165–3174. <https://doi.org/10.3168/jds.2008-1799>.
- Sandoe, P., B. Forkman, F. Hakansson, S. N. Andreassen, R. Nohr, M. Denwood, and T. B. Lund. 2017. Should the contribution of one additional lame cow depend on how many other cows on the farm are lame? *Animals (Basel)* 7:96–108. <https://doi.org/10.3390/ani7120096>.
- Schewe, R. L., J. Kayitsinga, G. A. Contreras, C. Odom, W. A. Coats, P. Durst, E. P. Hovingh, R. O. Martinez, R. Mobley, S. Moore, and R. J. Erskine. 2015. Herd management and social variables associated with bulk tank somatic cell count in dairy herds in the eastern United States. *J. Dairy Sci.* 98:7650–7665. <https://doi.org/10.3168/jds.2014-8840>.
- Simensen, E., O. Osteras, K. E. Boe, C. Kielland, L. E. Ruud, and G. Naess. 2010. Housing system and herd size interactions in Norwegian dairy herds; associations with performance and disease incidence. *Acta Vet. Scand.* 52:14.
- Solano, L., H. W. Barkema, E. A. Pajor, S. Mason, S. J. LeBlanc, J. C. Zaffino Heyerhoff, C. G. R. Nash, D. B. Haley, E. Vasseur, D. Pellerin, J. Rushen, A. M. de Passillé, and K. Orsel. 2015. Prevalence of lameness and associated risk factors in Canadian Holstein-Friesian cows housed in freestall barns. *J. Dairy Sci.* 98:6978–6991. <https://doi.org/10.3168/jds.2015-9652>.
- Sorensen, J. T., and D. Fraser. 2010. On-farm welfare assessment for regulatory purposes: Issues and possible solutions. *Livest. Sci.* 131:1–7. <https://doi.org/10.1016/j.livsci.2010.02.025>.
- Spooner, J. M., C. A. Schuppli, and D. Fraser. 2014. Attitudes of Canadian citizens toward farm animal welfare: A qualitative study. *Livest. Sci.* 163:150–158.
- Takai, H., S. Pedersen, J. O. Johnsen, J. H. M. Metz, P. W. G. Groot Koerkamp, G. H. Uenk, V. R. Phillips, M. R. Holden, R. W. Sneath, J. L. Short, R. P. White, J. Hartung, J. Seedorf, M. Schröder, K. H. Linkert, and C. M. Wathes. 1998. Concentrations and emissions of airborne dust in livestock buildings in Northern Europe. *J. Agric. Eng. Res.* 70:59–77. <https://doi.org/10.1006/jaer.1997.0280>.
- Torres-Cardona, M. G., M. E. Ortega-Cerrilla, J. I. Alejos-de la Fuente, J. Herrera-Haro, and J. G. Peralta Ortiz. 2014. Effect of regrouping Holstein cows on milk production and physical activity. *J. Anim. Plant Sci.* 22:3433–3438.
- Van Os, J. M. C., C. Winckler, J. Trieb, S. V. Matarazzo, T. W. Lehenbauer, J. D. Champagne, and C. B. Tucker. 2018. Reliability of sampling strategies for measuring dairy cattle welfare on commercial farms. *J. Dairy Sci.* 101:1495–1504. <https://doi.org/10.3168/jds.2017-13611>.
- Vanhonacker, F., and W. Verbeke. 2014. Public and consumer policies for higher welfare food products: Challenges and opportunities. *J. Agric. Environ. Ethics* 27:153–171. <https://doi.org/10.1007/s10806-013-9479-2>.
- Vanhonacker, F., W. Verbeke, E. Van Poucke, and F. A. M. Tuytens. 2008. Do citizens and farmers interpret the concept of farm animal welfare differently? *Livest. Sci.* 116:126–136. <https://doi.org/10.1016/j.livsci.2007.09.017>.
- Veissier, I., J. Capdeville, and E. Delval. 2004. Cubicle housing systems for cattle: Comfort of dairy cows depends on cubicle adjustment. *J. Anim. Sci.* 82:3321–3337. <https://doi.org/10.2527/2004.82113321x>.
- Verker, G. A., and P. H. Hemsworth. 2010. Managing cow welfare in large dairy herds. Pages 436–443 in *Proc. 4th Australasian Dairy Science Symposium*. Lincoln University, Christchurch, New Zealand.
- Waiblinger, S., and C. Menke. 1999. Influence of herd size on human-cow relationships. *Anthrozoos* 12:240–247. <https://doi.org/10.2752/089279399787000156>.
- Waiblinger, S., C. Menke, and D. W. Fölsch. 2003. Influences on the avoidance and approach behavior of dairy cows towards humans on 35 farms. *Appl. Anim. Behav. Sci.* 84:23–39. [https://doi.org/10.1016/S0168-1591\(03\)00148-5](https://doi.org/10.1016/S0168-1591(03)00148-5).
- Wechsler, B., J. Schaub, K. Friedli, and R. Hauser. 2000. Behaviour and leg injuries in dairy cows kept in cubicle systems with straw bedding or soft lying mats. *Appl. Anim. Behav. Sci.* 69:189–197.
- Welfare Quality. 2012. *Welfare Quality® Assessment protocol for cattle*. Welfare Quality® Consortium, Lelystad, Netherlands. Accessed Aug. 21, 2017. <http://www.welfarequalitynetwork.net/en-us/news/assessment-protocols/>.
- Whitaker, D. A., A. I. Macrae, and E. Burrough. 2004. Disposal and disease rates in British dairy herds between April 1998 and March 2002. *Vet. Rec.* 155:43–47. <https://doi.org/10.1136/vr.155.2.43>.
- Winckler, C., J. Brinkmann, and J. Glatz. 2007. Long-term consistency of selected animal-related welfare parameters in dairy farms. *Anim. Welf.* 16:197–199.
- Zaffino Heyerhoff, J. C., S. J. LeBlanc, T. J. DeVries, C. G. R. Nash, J. Gibbons, K. Orsel, H. W. Barkema, L. Solano, J. Rushen, A. M. de Passillé, and D. B. Haley. 2014. Prevalence of and factors associated with hock, knee, and neck injuries on dairy cows in freestall housing in Canada. *J. Dairy Sci.* 97:173–184. <https://doi.org/10.3168/jds.2012-6367>.
- Zuliani, A., A. Romanzin, M. Corazzin, S. Salvador, J. C. Abrahantes, and S. Bovolenta. 2017. Welfare assessment in traditional mountain dairy farms: Above and beyond resource-based measures. *Anim. Welf.* 26:203–211. <https://doi.org/10.7120/09627286.26.2.203>.