



Technical note: Evaluation of a sonographic overbagging edema scoring system for show cows: Comparison with visual inspection

M. Balmer,* M. Alsaad,* M. Boesiger,* R. O'Brien,† and A. Steiner*¹

*Clinic for Ruminants, Vetsuisse-Faculty, University of Bern, 3001 Bern, Switzerland

†BitRads, Nobleboro, ME 04555

ABSTRACT

Overbagged udders are commonly seen at dairy cow shows due to prolonged milking intervals. The aims of this study were to describe the prevalence of udder edema in dairy cows at shows as determined by sonography and to suggest a sonographic scoring system, which was evaluated for reproducibility and repeatability. At 4 highly competitive Swiss dairy cow shows, 319 cows of various breeds were examined. Sonographic scans—1 from each fore quarter and 1 from the rear, representing both hind quarters—were collected by 3 experienced veterinarians at defined positions predisposed to overbagging edema. Sonographic scans were scored with the aid of a newly suggested scoring system (score 0 = no edema, grade 1 = slight edema, grade 2 = moderate edema, grade 3 = severe edema). Further, 139 video sequences from behind while the cows were walking and the same number of photographs of the udder each from the left, right, and behind were selected. The photographs and video sequences were visually scored for different parameters as 0 = not present or 1 = present except for abduction, which was scored on a visual analog scale (0 = no abduction; 100 = maximal imaginable abduction). Visual scores were then compared with the sonographic scoring (Pearson's chi-squared). The prevalence of udder edema scores 1, 2, and 3 was 14.0, 6.5, and 2.3%, respectively. Interobserver reliability for objective sonographic scoring ($\kappa = 0.815$) and intraobserver agreement ($\kappa = 0.90$ and 0.85 for 2 different observers) was estimated as “almost perfect.” Interobserver agreements for visual parameters were $\kappa = 0.40$ or less except for teat shape ($\kappa = 0.52$). Classification of the videos resulted in better interobserver agreement; when rating hind limb abduction, Spearman correlation coefficient was 0.61, whereas $\kappa = 0.61$ for absence of udder movement. Udder movement and teat shape were significantly associated with the edema score. The higher inter- and intraobserver

reliability of the sonographic edema scoring compared with the visual scoring system indicates that the newly suggested scoring system might be used in the future to objectively identify udder edema in cows at dairy shows.

Key words: udder edema, sonographic scoring system, overbagging, dairy show

Technical Note

Overbagged udders are commonly seen at dairy cow shows due to prolonged duration of not milking, which occasionally leads to milking intervals of as much as 24 to 36 h (Waller et al., 2007; O'Brien, 2017). This practice is a common part of show preparation, which aims to make an udder appear what is considered to be the ideal udder for competition. The adverse effects of prolonged milking intervals (PMI) on a cow's well-being have recently been shown (Kohler et al., 2016). They include behavioral changes such as decreased eating time and increased hind limb abduction as well as non-physiological status such as increased udder firmness and the development of udder edema. Similar results were described by Bertulat et al. (2017) at dry-off. This study group demonstrated higher udder pressure and a greater increase in stress levels after dry-off in high-yielding cows (yielding ≥ 16 kg/d at the time of dry-off) compared with a control group treated with cabergoline (to stop milk secretion). The effect of a sudden dry-off on low-yielding cows was negligible. After PMI, edema develops at defined udder locations, and the differentiation of physiological postpartum edema, injection-induced edema, and edema caused by overbagging is possible, mainly according to the sites of occurrence (O'Brien et al., 2002; Waller et al., 2007). Although the exact mechanism leading to edema formation is not known, Melendez et al. (2006) stated that “the highly vascular nature of the bovine mammary gland makes the tissue more prone to develop localized edema due to an increase in blood and lymphatic flow.” This refers to periparturient cows. The opposite might occur in the case of overbagging edema, as increased pressure in the udder ought to impede venous and lymphatic

Received January 17, 2018.

Accepted March 30, 2018.

¹Corresponding author: adrian.steiner@vetsuisse.unibe.ch

drainage. This might explain the different locations of periparturient versus overbagging edema, as described by Waller et al. (2007). As demonstrated by Kohler et al. (2016) in Holstein-Friesian cows at 90 d postpartum, udder edema occurred after a PMI of 24 h in 10 out of 15 cows (first sign of edema seen after 18 h), and it was not found in any of the experimental cows after a milking interval of 12 h. Somatic cell count was increased at 12 h after PMI (from 66.3 cells/ μ L before PMI to 216.3 cells/ μ L 12 h after PMI; Kohler et al., 2016). Subjective (visual) scoring systems have been developed for describing edema in periparturient cows. The criteria include subjective scoring done by herd personnel without specific instructions (Dentine and McDaniel, 1983) and the use of visual and tactile criteria such as waxy appearance, abnormal udder shape, and recognition of the pitting, which represents a sign of edema, by applying pressure with a fingertip (Tucker et al., 1992). As these scoring systems do not allow an examiner to objectively determine and rate edema formation during overbagging, objective methods for assessing edema associated with PMI are needed. Sonography is a sensitive method for identifying edema (O'Brien, 2017). The sonographic characteristics of edema are best described as "subcutaneous alternating bands of hyperechoic subcutaneous fascial planes and homogenous hypoechoic fluid" (Waller et al., 2007). A reliable system of objectively measuring overbagging at cow shows is therefore required to recognize udder edema and educate the cow's owner about the adverse effect of PMI.

In addition to estimating the prevalence of overbagging udder edema in show cows using sonography, the objectives of this study were to suggest a sonographic scoring system and to evaluate this system for reproducibility and repeatability. Furthermore, the study aimed at adapting the visual udder edema score as described by Dentine and McDaniel (1983) and evaluating its accuracy, defining the sonographic score as the gold standard. We hypothesized that the newly suggested sonographic scoring system provides high inter- and intraobserver agreement and good correlation with the visual udder edema score.

Sonographic scans, video recordings, and photographs of 319 cows were collected at 4 highly competitive dairy shows in Switzerland between September 2016 and April 2017. Of these, 28 cows were examined twice. Sonographic scans of both examinations were used to check inter- and intraobserver reliability of the suggested scoring system. In case of 2 examinations, only the second examination was used to estimate prevalence and provide a comparison with the visual scoring. Twelve cows were excluded from prevalence estimation and comparison with visual scoring (3 cows because

they were examined after milking and 9 cows because they calved within the 21 d before the respective show). Therefore, 307 cows were included in these analyses. The breeds involved were Brown Swiss ($n = 97$), Holstein-Friesian ($n = 61$), Jersey ($n = 20$), Montbéliard ($n = 23$), Original Brown ($n = 22$), Red Holstein ($n = 44$), Simmental ($n = 19$), and Swiss Fleckvieh ($n = 21$). The age of the cows (mean \pm SD) was 4.55 ± 2.21 yr (range: 2.07–14.03 yr), and they were at 115.7 ± 74.2 DIM (range: 22–500 DIM). Based on official milking data of the respective breeder associations, cows yielded 33.4 ± 7.7 kg/d within the month before evaluation (range: 16.1–60.3 kg). Sonographic examination was performed by 3 veterinarians who were experienced in soft tissue scanning (Boesiger, O'Brien, and Steiner). Sonographic images were taken before milking using MyLab One (Esaote, Genova, Italy) scanners and a 10-MHz linear probe. As settings, a depth of 5 cm with focus on 1 cm of depth and 70% gain were chosen. Both front quarters and the hind quarters in the area of the central ligament were examined as shown in Figure 1. Sonographic still pictures were blinded as to shows, identity of the cows and owners, and the udder location. Three sets of 50 still pictures were randomly chosen to evaluate unambiguous assignment. The scoring system was modified twice until a final scoring system was obtained, as described in this article. Thereafter, the complete set of sonographic still pictures ($n = 1,038$) was scored. Grade 0 was defined as no sonographic signs of edema. In the fore and rear quarters, a different definition was given for each sonographic edema score (grade 1 = slight edema; grade 2 = moderate edema; grade 3 = severe edema). The characteristics of each score at different sites are described in the legend of Figure 2. The sonographic still pictures of all cows were independently scored within 2 wk after the last show by 3 experienced scientists (Balmer, O'Brien, and Steiner), all of whom were familiar with evaluating sonographic scans. The interobserver reliability (test for Kendall's coefficient of concordance for ordinal response) of these 3 observers was determined using SAS (SAS Institute Inc., Cary, NC). The intraobserver reliability was determined by 2 experienced scientists (Balmer and Steiner) using NCSS 10 (NCSS LLC, Kaysville, UT) after re-scoring the sonographic still pictures 3 mo after their first evaluation. To determine edema prevalence at cow level, the maximal score at quarter level determined the score at cow level, as the score of the quarters within many cows ($n = 70$) were different.

Additionally, 3 digital photographs of the udders (1 each from the left and right sides and 1 from behind) and a video sequence of the cow walking were taken from behind before milking and digitally stored. Photographs and video sequences were taken by different

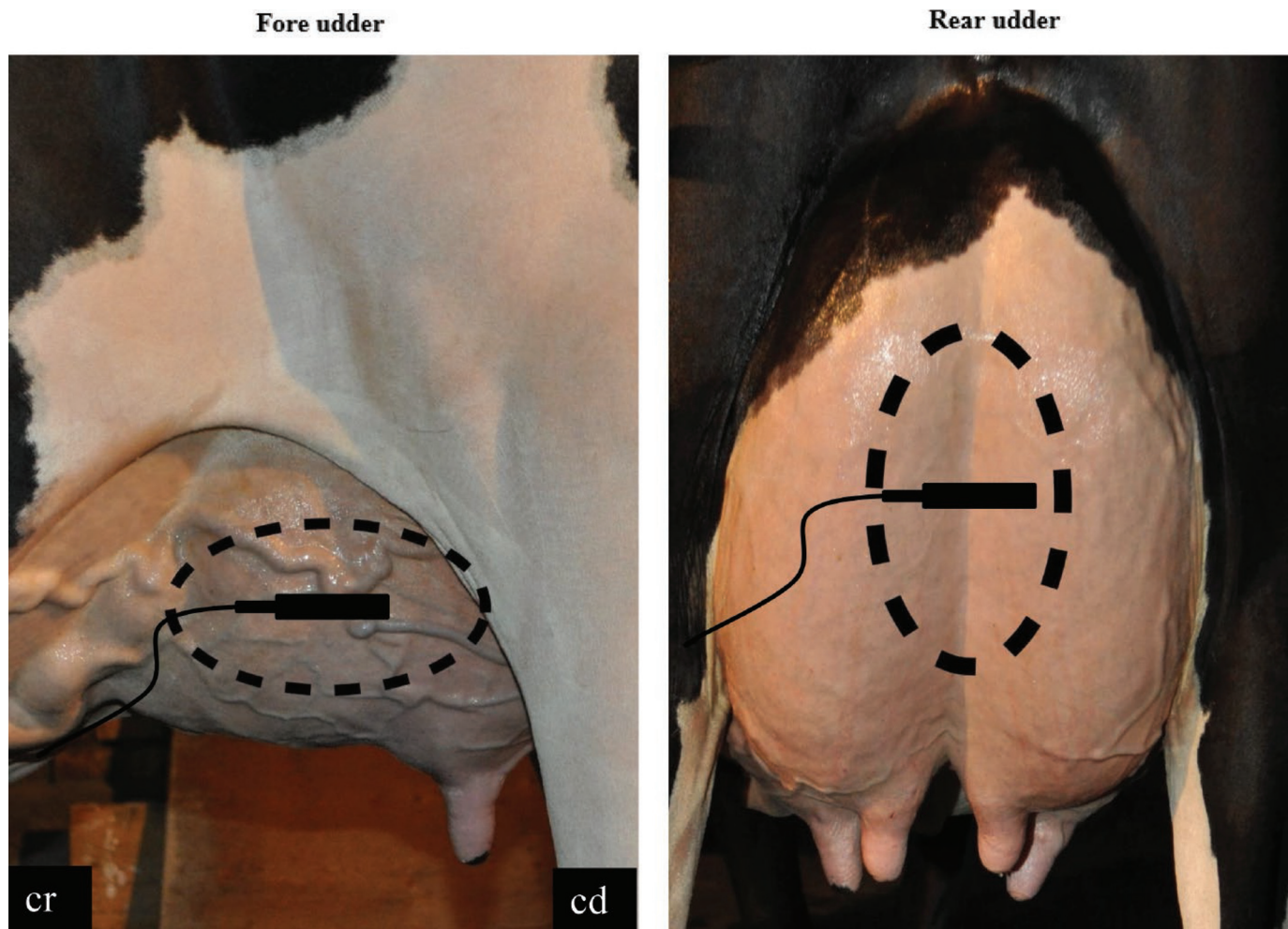


Figure 1. Positions of the sonographic linear probe: on both front quarters and on the cleft between rear quarters. cr = cranial; cd = caudal. The dotted lines represent the area where edema typically is located. Color version available online.

collaborators according to a standardized protocol. Photographs were always taken at a 90° angle to the quarter(s) of interest with the cow in standing position, putting equal weight on both hind limbs. Pictures were standardized to ensure that each fore quarter and both rear quarters, respectively, were in the center of the photograph. For photographs of the rear quarters, the tail was gently pulled aside. Video sequences recording walking for at least 5 m were taken from behind the cow, also at a 90° angle to the caudal surface of the hind quarters. A Nikon D5000 (Nikon Corporation, Tokyo, Japan) digital camera and a Sony FDR-AX33 (Sony Corporation, Tokyo, Japan) video camera were used for collecting photographs and video sequences, respectively. At 3 mo after the last show, the photographs and video sequences were also blinded as to shows and the identity of the cows and owners and were scored by 2 independent observers (Balmer and Alsaad) within a

1-wk period. Photographs were scored as 0 (not present) or 1 (present) relating to the following parameters: cleft impression of the median suspensory ligament, teat shape, udder shape, bulging of the subcutaneous veins, and udder shape abnormalities (Table 1). The video sequences were scored according to Kohler et al. (2016) for hind limb abduction on a visual analog scale (0 = no abduction and 100 = maximal imaginable abduction) and for absence of udder movement (0 = sufficient space and unrestricted movement possible; 1 = no space present between udder and legs; Table 1). These parameters were defined according to the literature (Tucker et al., 1992) and completed by adding information discovered (1) by studying photographs and videos of 25 cows out of the study population, including all cows with score 3 ($n = 7$) and 18 randomly selected cows with udder edema scores 0 to 2 and (2) by including the expert opinion of 3 licensed Swiss show

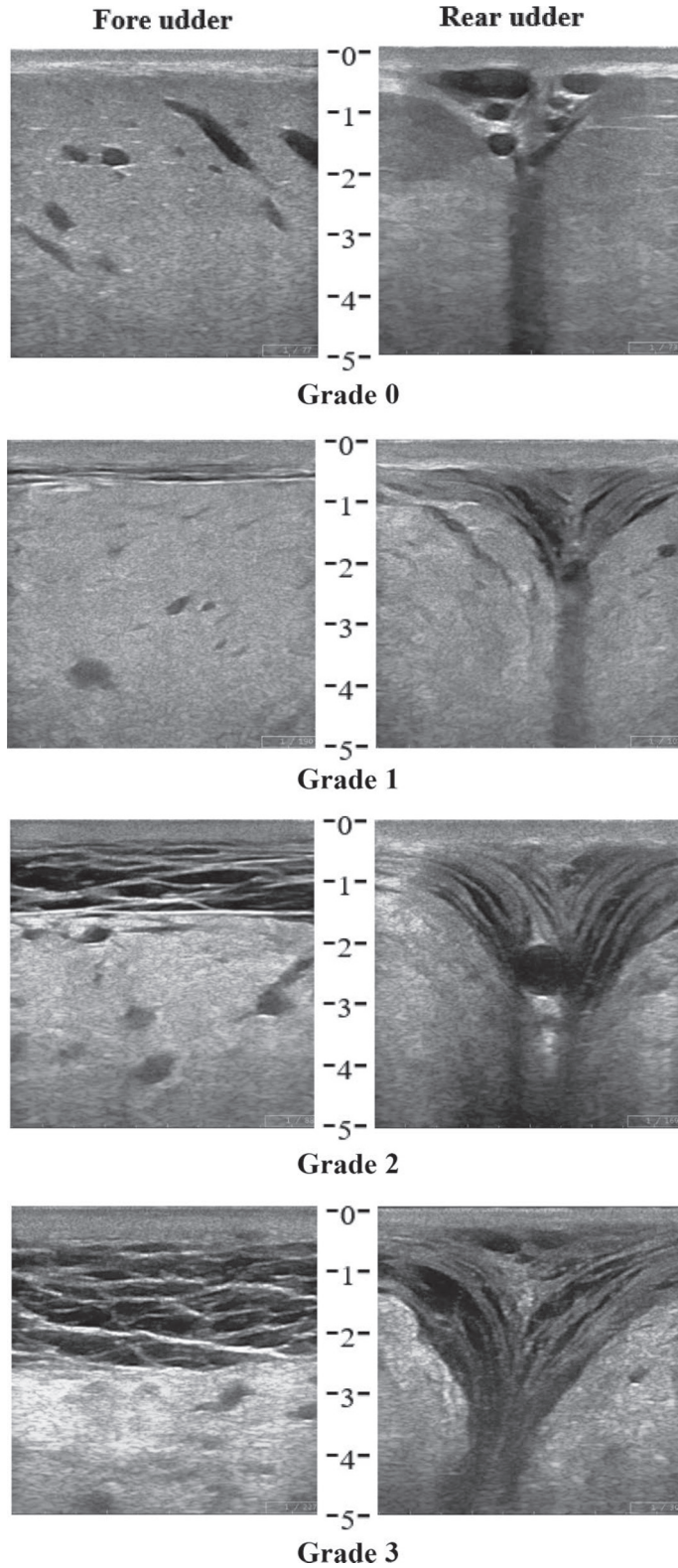


Figure 2. Grade 0 = normal. No signs of edema present. Grade 1 = slight edema. Fore udder: 2 to 4 distinct edema lines or <1 cm deep or both; rear udder: 2 to 5 distinct edema lines or <2 cm deep or both. Grade 2 = moderate edema. Fore udder: 5 to 8 distinct edema lines or >1 but ≤2 cm deep or both; rear udder: 6 to 8 distinct edema lines or >2 but ≤3 cm deep or both. Grade 3 = severe edema. Fore udder: >8 edema lines or >2 cm deep or both; rear udder: >8 distinct edema lines or >3 cm deep or both. Depth is measured from the probe-skin interface (0–5 cm).

Table 1. Parameters for scoring the photographs (parameters 1–5) and video sequences (parameters 6 and 7)¹

Parameter	Definition
1. Cleft of the median suspensory ligament	Cleft caused by the median suspensory ligament is less prominent due to edema formation, which fills the slot
2. Teat shape	Teats are funnel formed
3. Udder shape	Udder is spheroid shaped
4. Bulging of the subcutaneous veins	Bulging of the subcutaneous veins is reduced
5. Udder shape abnormalities	Udder shape abnormalities (e.g., sagging formation)
6. Hindlimb abduction	Degree of hindlimb abduction at walking on a visual analog scale
7. Absence of udder movement	Absence of free udder movement

¹Parameters were scored as either 0 (not present) or 1 (present) except parameter 6, which was based on a visual analog scale (0 = no abduction, 100 = maximal imaginable abduction) according to Kohler et al. (2016).

judges. The photographs and videos were studied by 2 independent veterinary scientists comparing photographs and video sequences of knowingly overbagged with non-overbagged cows. These scientists were not involved in scoring the pictures. Expert opinion from licensed Swiss show judges was collected with the aid of standardized personal interviews.

All photographs and videos of cows with a sonographic score of 1, 2, or 3 ($n = 70$) were selected. Then, 70 randomly selected cows with a sonographic score of 0 were chosen as negative controls. One of the 70 cows with score 0 was excluded from comparison due to the poor quality of photographs and video sequences. Therefore, 139 cows were included in the analysis. The interobserver reliability of the 2 observers was determined (Pearson's chi-squared for categorical data, correlation for continuous data). Kappa values (κ) were classified as follows: <0.00 = poor; 0.00 to 0.20 = slight; 0.21 to 0.40 = fair; 0.41 to 0.60 = moderate; 0.61 to 0.80 = substantial; 0.81 to 1.00 = almost perfect agreement (Landis and Koch, 1977). Whenever agreement was at least moderate ($\kappa \geq 0.41$) for categorical data (udder movement and teat shape), the 2 observers rescored diverging data together to match the results. Furthermore, defining the sonographic score as a binominal result (grade 0 = negative; grades 1, 2, 3 = positive) as gold standard, sonographic score and (matched) visual scores were compared by Pearson's chi-squared.

A total of 236 cows at dairy shows had a sonographic score of 0 (76.9%), 43 cows had a score of 1 (14.0%), 20 cows had a score of 2 (6.5%), and 7 cows had a score of 3 (2.3%). From 1 cow, no scan was available (0.3%). Kappa value for interobserver reliability for 3 observers concerning the sonographic score was 0.815, representing an almost perfect agreement (Landis and Koch, 1977). Intraobserver agreement was $\kappa = 0.90$ and 0.85 for 2 different observers. Interobserver agreement for visual parameters showed a substantial agreement for absence of udder movement ($\kappa = 0.61$) and the hindlimb abduction (Spearman correlation coefficient

= 0.61), respectively, and a moderate agreement for teat shape ($\kappa = 0.52$). There was only a fair interobserver agreement for the parameters udder shape ($\kappa = 0.40$), bulging of the subcutaneous veins ($\kappa = 0.33$), cleft of the median suspensory ligament ($\kappa = 0.23$), and udder shape abnormalities ($\kappa = 0.22$). The (matched) values of 2 observers concerning udder movement ($P < 0.01$) and teat shape ($P = 0.04$) revealed significant associations with sonographic edema score.

The sonographic scoring system as suggested in this study proved to be a suitable tool for objectively determining the degree of udder edema present in dairy show cows. Inter- and intraobserver agreements of the proposed sonographic score are high. The sonographic scoring system could easily be applied during cow shows, not significantly increasing the udder evaluation time if performed by an experienced evaluator (<1 min; data not reported). Other methods such as measuring udder firmness and milk leaking do not represent appropriate techniques. Measuring udder firmness requires multiple firm contacts of the dynamometer with the udder (Bertulat et al., 2012), and no reference values associated with PMI have been defined. Milk leakage is not evident at many dairy shows, as the outer teat orifice is routinely sealed with collodion (currently allowed in Switzerland) or superglue (not allowed in Switzerland) or both.

The sonographic scoring system requires special technical equipment and might replace the currently practiced visual scoring systems proposed by Dentine and McDaniel (1983) and Tucker et al. (1992). The latter are very subjective, not suitable for use at dairy shows (palpation of the udder not accepted by exhibitors), or reveal insufficient agreement with the described sonographic score. The sonographic scoring system described in this study can be used to identify and grade the severity of udder edema at dairy cow shows with a high reproducibility and repeatability. As a high prevalence of udder edema was found, further steps to counteract this practice are required.

ACKNOWLEDGMENTS

This study was generously supported by the Swiss Federal Food Safety and Veterinary Office (Bern, Switzerland) and the Association of Swiss Cattle Breeders (Zollikofen, Switzerland). We thank Gertraud Schuepbach-Regula for statistical support. The scanner was generously provided for the study by Henry Schein Animal Health (Lyssach, Switzerland).

REFERENCES

- Bertulat, S., C. Fischer-Tenhagen, A. Werner, and W. Heuwieser. 2012. Technical note: Validating a dynamometer for noninvasive measuring of udder firmness in dairy cows. *J. Dairy Sci.* 95:6550–6556.
- Bertulat, S., N. Isaka, A. de Prado, A. Lopez, T. Hetreau, and W. Heuwieser. 2017. Effect of a single injection of cabergoline at dry off on udder characteristics in high-yielding dairy cows. *J. Dairy Sci.* 100:3220–3232.
- Dentine, M. R., and B. T. McDaniel. 1983. Variation of edema scores from herd-year, age, calving month, and sire. *J. Dairy Sci.* 66:2391–2399.
- Kohler, P., M. Alsaood, G. Dolf, R. O'Brien, G. Beer, and A. Steiner. 2016. A single prolonged milking interval of 24h compromises the well-being and health of dairy Holstein cows. *J. Dairy Sci.* 99:9080–9093.
- Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159–174.
- Melendez, P., C. C. Hofer, and G. A. Donovan. 2006. Risk factors for udder edema and its association with lactation performance on primiparous Holstein cows in a large Florida herd, USA. *Prev. Vet. Med.* 76:211–221.
- O'Brien, R. T. 2017. Over bagging in dairy show cows: An ethical crisis. *J. Am. Vet. Med. Assoc.* 251:271–272.
- O'Brien, R. T., K. R. Waller, and J. S. Matheson. 2002. Ultrasonographic appearance of edema caused by injections in the mammary gland attachments of dairy cows. *J. Am. Vet. Med. Assoc.* 221:408–410.
- Tucker, W. B., G. D. Adams, M. Lema, M. Aslam, I. S. Shin, P. Leruyet, and D. L. Weeks. 1992. Evaluation of a system for rating edema in dairy cattle. *J. Dairy Sci.* 75:2382–2387.
- Waller, K. R., R. T. O'Brien, and S. M. McGuirk. 2007. Ultrasonographic distribution and duration of udder edema in post-partum and over-bagged dairy cows. *Bovine Pract.* 41:129–133.