An observational study investigating uniformity of manual body condition scoring in dairy cows

Abhishek Paul1, Champak Bhakat2, Santu Mondal1, and Ajoy Mandal3

Received: 29 October 2019 / Accepted: 28 January 2020 / Published online: 27 February 2020
© Indian Dairy Association (India) 2020

Abstract: Body condition scores (BCS) are very useful for dairy herd management; however, its reliability and consistency of recordings made by observers have been questioned. Moreover, regular high-quality manual body condition scoring of an individual cow is difficult as a routine practice on the farm. This study aimed to find out the agreement in BCS within- and between the observers. An observational study was conducted in which two observers independently scored BCS of 43 crossbred animals from one dairy herd. Cohen’s kappa (κ) and spearman’s rho (ρ) was computed to quantify the agreement between observer for overall BCS measurement and only kappa test for different BCS groups (high, medium and low). For overall BCS score ê of 0.62 to 0.71 (p<0.001) and 0.68 to 0.74 (p<0.001) were obtained for intra observer 1 and 2 respectively, such values would be interpreted as moderate agreement. Whereas, the inter-observer agreement (ê) was found to be low i.e. between 0.52 to 0.60 (p<0.001). However, the spearman’s rho value was higher in all the cases, indicating a good correlation among the individual observers. Besides, kappa value (κ) for different BCS groups revealed the lowest agreement between observers (0.31 to 0.37, p<0.001) and within observers (0.34 to 0.59, p<0.001) for medium BCS group as compared to the other BCS groups. These findings suggest multiple observers should perform manual Body Condition Scoring for better accuracy in the outcome.

Keywords: BCS, Dairy cattle, Observation variation

Introduction

Initially, BCS evaluation chart was developed by assessing the eight anatomical locations of the rear half of the cow within the areas of the loin, pelvis and tail head and prepared a scale ranges from 1 to 5, using 0.25-unit increments (Wildman et al. 1982, Edmonson et al. 1989). Several researchers from different countries had developed various scales by observing and palpating the animal body. Bewley and Schutz (2008) reviewed international BCS systems, in the UK and Ireland widely used scale ranges from 1-5 with 0.50 and 0.25 intervals, respectively. Likewise, in Australia and New Zealand most common and widely used scale ranges from 1-8 and 1-10 with an increment of 0.5 scale intervals. In India, Prasad (1994) has developed a modified Body Condition Scoring scale ranges from 1 to 6, with an increment of 0.5 for better accuracy. Evans, (1978) and Nicoll, (1981) have studied the factors causing variation in BCS and found 60 to 70 % was due to animal variation, 5 % from the evaluator and 10 % happened animal-evaluator variation. The quality of manual Body Condition Scoring depends on the observers and scoring protocol (Kristensen et al. 2006), where a trained person had consistency up to 58 to 67 % accuracy and 27 % precision by an untrained person (Ferguson et al. 1994). Inter-observer agreement was more reliable than the single observer conducted all BCS evaluation (Morin et al. 2017). Also, BCS (1-5 point scale) changes of 0.25 cannot realistically be detected, even with trained observers (Bewley et al. 2008; Bewley and Schutz, 2008). When comparing the methods for monitoring the BCS, Mazeris (2015) found an automated BCS system to be highly accurate to a human scorer, with 98% of scores being within a quarter-point. Previously, intra and inter observational agreements were done in 1 to 5 BCS scale with 0.25 unit increments (Kristensen et al. 2006; Morin et al. 2017) based on the scale of Ferguson et al. (1994). Our hypothesis of the present study was observational agreement within and between professionals on BCS in 1-6 scale with 0.5 unit increments.

Materials and Methods

Data collection

Two veterinarians were independently assessed BCS of a total of 43 cows. BCS was measured using a visual plus palpation
technique in 1 to 6 scales with 0.5 increments, which was established by Prasad (1994). The data were collected two consecutive days in a week for three weeks. All the crossbred Jersey cattle were divided into three BCS groups i.e. high (> 4.5/above 4.5, N= 12), Medium (3 to 4.5, N= 22) and Low (< 3/ below 3, N= 9) by an expert. Both the veterinarians were well trained for manual body condition scoring of dairy cattle. Prasad (1994) have mentioned specific regions for body condition scoring in dairy cattle, which needs sound anatomical knowledge of the observer for better accuracy. The anatomical regions i.e. vertebral column, spinous processes, transverse processes, pin and hook bone, tail head region and ribs were considered, based on that scoring were done.

**Statistical analysis**

All statistical analyses were performed using IBM SPSS statistics software (version 22.0). However, we considered BCS as an ordinal variable because it classifies cows into different categories of body condition. The quality of manual body condition scoring of the two observers was quantified by the scoring agreement Cohen’s kappa (κ) and the scoring correlation coefficient Spearman’s rho (ρ). Cohen’s kappa (κ) measures the pairwise agreement of two observers or one observer across 2 days, including the possibility of agreement occurring by chance. Spearman’s rho (ρ) is a pairwise rank correlation and is defined as the statistical dependence between the rankings of 2 manual scores without interference from the systematic scoring difference. The ρ value ranges from 1 (i.e. a pair set of scores with identical ranks) to –1 (i.e. a pair set of scores with opposite ranks). Both ê and ñ values were calculated for intra-observer agreement between the consecutive days and for inter-observer comparison of the 2-d average scores between observers 1 and 2.

**Results and Discussion**

The descriptive statistics of BCS observation by two veterinarians presented in Table 1, whereas intra and inter-observer agreement (ê) and Spearman correlation coefficients (ñ) of the overall manual BCS of two observers shown in Table 2. In the present study, two experienced veterinarians were engaged in BCS evaluation to increase homogeneity in inter-class variations. Both the observers assessed overall BCS with good intra observational agreement i.e., κ values were 0.62 to 0.71 (p<0.001) and 0.68 to 0.74 (p<0.001) respectively. However, inter-observer agreement was moderate and kappa value (κ) was significantly (p<0.001) varied in the range of 0.52 to 0.60. Additionally, intra and inter-observer correlation coefficient (Spearman rho, ρ) were calculated and it was significantly higher within observers (ρ>0.95) than between observers (ρ > 0.88) during the study period. However, Ferguson et al. (1994) also reported a high correlation among observers i.e. 0.76 to 0.85. The kappa value (κ) is widely used as an indicator of agreement and greater than 0.80 is representing excellent agreement in clinical observation (Ersboll et al. 2004). Kristensen et al. (2006) have found low to moderate Intra and inter-observer agreement on BCS measurement among veterinarians i.e. κ, ranges between 0.22 to 0.75 and 0.17 to 0.78 respectively. Whereas, Song and coworker (2019) have observed moderate intra-observer agreement (κ, 0.40 to 0.59 and 0.60 to 0.79) with lower inter-observer agreement (ê, 0.20 to 0.79). On the contrary, excellent intra (κ, 0.86) and inter-observer (κ, 0.76 to 0.85) agreement was observed for highly trained instructors by Kristensen et al. (2006) and Morin et al. (2017) respectively. In this present study, intra and inter observational variations were comparably narrower than Kristensen et al. (2006). However, the comparison was not error free with other studies as it uses different BCS scale i.e. 6 point scale with 0.5 increments. Furthermore, in our study, BCS observation values ranged from 1 to 6 due to wide variation in the BCS groups. On the contrary, Morin et al. (2017) found all BCS values ranged between 2.5 to 3.5, resulting in higher inter-observer agreement (κ) between observers than the present findings.

| Time  | Observer | N  | Minimum | Maximum | Mean | Std. Error | Std. Deviation | Variance |
|-------|----------|----|---------|---------|------|------------|----------------|----------|
| 1st week | 1        | 43 | 1.5     | 5.5     | 4.05 | 0.178      | 1.17           | 1.37     |
|        |          |    | 43      | 1.5     | 6    | 3.95       | 0.191          | 1.25     |
|        |          |    | 43      | 1.5     | 6    | 3.67       | 0.201          | 1.32     |
|        |          |    | 43      | 1.5     | 5.5  | 3.59       | 0.189          | 1.24     |
| 2nd week | 1        | 43 | 1.5     | 5.5     | 4.03 | 0.174      | 1.15           | 1.31     |
|        |          |    | 43      | 1.5     | 6    | 3.93       | 0.185          | 1.20     |
|        |          |    | 43      | 1.5     | 6    | 3.88       | 0.199          | 1.30     |
|        |          |    | 43      | 1.5     | 5.5  | 3.74       | 0.188          | 1.24     |
| 3rd week | 1        | 43 | 1.5     | 5.5     | 4.03 | 0.174      | 1.15           | 1.31     |
|        |          |    | 43      | 1.5     | 6    | 3.95       | 0.194          | 1.27     |
|        |          |    | 43      | 1.5     | 6    | 3.85       | 0.179          | 1.17     |

Further, the intra and inter-observer agreement (κ) were measured for three different BCS groups and represented in Table 3. In the high BCS group, Kappa (κ) values were ranges between 0.62 to
Indian J Dairy Sci 73(1): 77-80

Table 2 Intra and inter-observer agreements and correlation coefficient for manual BCS (p< 0.001)

| Groups | 1st week | 2nd week | 3rd week |
|--------|----------|----------|----------|
| Intra observer 1 | Cohen’s kappa (κ) | 0.62 | 0.68 | 0.71 |
| | Spearman’s rho (ρ) | 0.95 | 0.96 | 0.98 |
| Intra observer 2 | Cohen’s kappa (κ) | 0.68 | 0.71 | 0.74 |
| | Spearman’s rho (ρ) | 0.97 | 0.96 | 0.98 |
| Inter observer | Cohen’s kappa (κ) | 0.52 | 0.55 | 0.60 |
| | Spearman’s rho (ρ) | 0.88 | 0.90 | 0.91 |

Table 3 Description of Cohen’s kappa (κ) values among Low BCS (L-BCS), Medium BCS (M-BCS) and High BCS (H-BCS) groups

| Groups | Scoring agreement, Cohen’s kappa (κ) |
|--------|-------------------------------------|
| 1st week | 2nd week | 3rd week |
| Intra observer r 1 | L-BCS | 0.64 | 0.82 | 0.66 |
| | M-BCS | 0.34 | 0.42 | 0.54 |
| | H-BCS | 0.70 | 0.71 | 0.88 |
| Intra observer 2 | L-BCS | 0.82 | 0.83 | 0.85 |
| | M-BCS | 0.54 | 0.56 | 0.59 |
| | H-BCS | 0.62 | 0.75 | 0.74 |
| Inter observer | L-BCS | 0.53 | 0.52 | 0.68 |
| | M-BCS | 0.31 | 0.33 | 0.37 |
| | H-BCS | 0.57 | 0.60 | 0.58 |

Kappa ranges from 0 to 1

** 0 to 0.20 Slight agreement** 0.4 to 0.6 Moderate agreement** 0.8 to 1.0 Perfect agreement

Acknowledgments

Authors are grateful to both the Director, NDRI, and Head, ERS-NDRI, Kalyani for allowing the use of animals to complete this research.

References

Bercovich A, Edan Y, Alchanatis V, Moallem U, Parment Y, Honig H, Maltz E, Antler A, Halachmi I (2013) Development of an automatic cow body condition scoring using body shape signature and fourier descriptors. J Dairy Sci 96: 8047-8059

Bewley J, Schutz M (2008) Review: An interdisciplinary review of body condition scoring for dairy cattle. Professional Anim Scientist 24: 507-529

Bewley J, Peacock A, Lewis O, Boyce R, Roberts D, Coffey M, Kenyon S, Schutz M (2008) Potential for estimation of body condition scores in dairy cattle from digital images. J Dairy Sci 91: 3439-3453

Edmonson AJ, Lean IJ, Weaver LD, Farver T, Webster G (1989) A body condition scoring chart for Holstein dairy cows. J Dairy Sci 72: 68-78

Ersbøll AK, Bruun J, Toft N (2004) Data analysis. Chap.13 in Introduction to Veterinary Epidemiology. H. Houe, A. K. Ersbøll, and N. Toft, ed. Biofalia, Frederiksberg, Denmark.

Evans DG (1978) The interpretation and analysis of subjective body condition scores. Anim Prod 26: 119-125

Ferguson JD, Galligan DT, Thomsen N (1994) Principal descriptors of body condition score in Holstein cows. J Dairy Sci 77: 2695-2703

Kristensen E, Dueholm L, Vink D, Andersen JE, Jakobsen EB, Illum-Nielsen S, Petersen FA, Enevoldsen C (2006) Within- and across-
person uniformity of body condition scoring in Danish Holstein cattle. J Dairy Sci 89: 3721-3728
Mazeris F (2015) DeLaval body condition scoring BCS: daily, automatic & consistent scoring of cows. Pages 47-50 in Precision Dairy Conference
Morin PA, Chorfi Y, Dubuc J, Roy JP, Santschi D, Dufour S (2017) Short communication: An observational study investigating inter-observer agreement for variation over time of body condition score in dairy cows. J Dairy Sci 100: 3086-3090
Nicoll G (1981) Sources of variation in the condition scoring of cows. Ir J Agric Res 20: 27-33
Prasad, S (1994) Studies on body condition scoring and feeding management in relation to production performance of crossbred dairy cattle. Ph. D. thesis, NDRI Deemed University, Karnal, India
Song X, Bokkers EAM, van Mourik S, Groot Koerkamp PWG, van der Tol PPJ (2019) Automated body condition scoring of dairy cows using 3-dimensional feature extraction from multiple body regions. J Dairy Sci 102: 4294-4308
Wildman EE, Jones GM, Wagner PE, Boman RL, Troutt HF, Lesch TN (1982) A dairy cow body condition scoring system and its relationship to selected production characteristics. J Dairy Sci 65: 495-497