Fit or flabby: Can we simplify the body condition scoring method?

Presented in Partial Fulfillment of the Requirements for Graduation with Honors Research Distinction

Elizabeth M. Schafer

Department of Animal Sciences

College of Food, Agricultural, and Environmental Sciences

The Ohio State University

Expected Graduation May 2023

Project Advisor: Dr. Benjamin Wenner

Department of Animal Sciences

College of Food, Agricultural, and Environmental Sciences

The Ohio State University

2023

Abstract

Body condition scoring (BCS) on a 1 - 5 scale is a common method of assessing dairy cow body fat reserves, serving as a management tool to consistently identify patterns and address herd nutrition, reproduction, and health. However, the efficacy of scoring in different BCS point increments has not been well described. Our objective was to evaluate the variation in BCS based on scoring in half- (**HALF**) or quarter-point (**QRT**) increments and to evaluate the accuracy of categorically assessing cows as thin, ideal, or fat (**CAT**) for current lactation stage compared to assigning traditional BCS. Our hypothesis was that HALF scoring would not increase variation of managerial value. We further hypothesized that CAT score would be interpretively similar to traditional BCS. Jersey cows (n = 20) balanced by lactation stage were scored by undergraduate students (n = 14) of varying experience for 6 weeks. Prior to trial initiation, all participants received BCS training and were provided an on-farm scoring guide. Each week, participants assigned QRT, HALF, and CAT scores. Statistical analysis included a SAS mixed-model approach (fixed effects: scoring method, lactation stage, week, scorer experience; random effects: scorer, cow) with repeated week as appropriate; residuals were regressed against predicted BCS from HALF, centered to the mean. CAT method was on a 2 - 4 scale; QRT and HALF were on a 1 - 5 scale. Regression of the residuals indicated a mean bias ($P < 0.01$) where HALF underestimates herd BCS by 0.018 and a slope bias ($P < 0.01$) where HALF increased by 0.110 points compared to QRT for every whole point BCS gain observed. Scoring method influenced mean score ($P < 0.01$) where CAT was lower than QRT or HALF by 0.055. Also of note, novice scorers estimated cows 0.24 points lower than more experienced scorers ($P < 0.01$). Variance of QRT and HALF were not different ($P > 0.10$). These data indicate that HALF could be equally accurate as QRT, while potentially decreasing producer

effort and time to score. Further work is needed to support these results on farms with different breeds and management styles and with scorers of different ages and backgrounds.

Introduction

Body condition scoring (BCS) is a measurement of the leanness or fatness of an animal, determined by assessing the amount of adipose tissue that has accumulated over specified anatomical features. BCS is a central management tool used across many species and production systems, with variations in the scale (i.e., 1 - 9 for beef cattle (Eversole et al., 2009) and horses (Henneke et al., 1984); 1 - 5 for dairy cattle (Wildman et al., 1982)) and combination of techniques used (i.e., visual, palpation) to score different animals. While collecting body weight data on animals is highly useful for determining drug dosages or formulating rations, it is less useful as a measure of obesity. Variation in absolute body weight is confounded by differences in breed, age, frame size, and genetic variation, not only between different populations but also within a group (Wildman et al., 1982). For example, a young animal may be overweight and yet still weigh less than an ideally conditioned mature animal, or even another young animal that is taller. This can lead to inaccurate interpretations of which animal is obese or underweight. Assessing fat cover provides better insight into the animal's leanness because it is not swayed by these misleading factors. While ultrasound technology is useful to measure back fat depth to assess reserves (Domecq et al., 1995), this tool is expensive and therefore not widely available for producer use. Use of digital camera systems to score cows automatically has also been investigated, using both 2-dimensional overhead (Bewley et al., 2008; Halachmi et al., 2008; Azzaro et al., 2011) and 3-dimersional rear (Fischer et al., 2015) views. These automated approaches yield promising results, removing the human variation and scorer bias element from BCS, but have had minimal implementation so far. Therefore, BCS scoring continues to be the accepted method to measure body fat cover.

Body condition scoring on a 1 - 5 scale is the accepted method for dairy cows in the US, where a score of 1 represents an emaciated cow and a score of 5 represents an obese one. BCS scores are assigned by observation and palpation of adipose tissue over the bony prominences of specific body regions, namely the back (spinous and transverse processes of vertebrae), pelvis (including the tuber coxae or hooks, tuber ischii or pins, and thurl area), and tail head (including the coccygeal vertebrae, ischiorectal fossa, and ileo-sacral and ischial-coccygeal ligaments) regions (Wildman et al., 1982; Edmonson et al., 1989; Ferguson et al., 1994). This method is utilized on-farm by some producers, but is widely used by nutritionists, veterinarians, and other consultants.

Regular assessment of body energy reserves is especially important within the dairy industry because the condition of these animals fluctuates cyclically over each lactation period, in response to changes in the cow's net energy balance at levels of high or low production (Roche et al., 2007). This means that the BCS interpreted as "ideal" from a manager standpoint will also fluctuate depending on what stage of lactation the cow is currently in. From parturition through peak lactation, the average U.S. dairy cow is physically unable to consume enough energy in the diet to keep up with the energy demand of her current level of milk production. This deficit is referred to as a negative energy balance – when the body's energetic demands are greater than the energy taken in through feed.

To meet this excess in energy demand, the cow's body begins to mobilize adipose tissue, breaking it down from its storage form so that it can be utilized to make milk (Roche et al., 2009). Maintaining this high level of production is desirable for farm profitability, but excessive mobilization of adipose produces a buildup of ketones in the liver, leading to ketosis, which has negative impacts on cow health, fertility, and milk yield (Wathes et al., 2007), meaning this disease risk has a high economic impact on the dairy (Cainzos et al., 2022). As well, ketosis is often connected to an increased risk for other diseases, including displaced abomasum and metritis (Duffield et al., 2009).

It is important not only to watch for cows becoming too thin from excessive mobilization of fat, but also to identify cows becoming too fat. As the lactation progresses, the cow's production begins to decrease as she partitions more energy towards growing a new fetus and building up body reserves for the next lactation (Roche et al., 2007). This means that for every pound of feed eaten, the cow is less efficient at the goal of producing milk, a clear economic loss. However, becoming obese in late lactation can also lead to increased health risks including dystocia as well as metabolic disorders in early lactation as the cow pulls energy from an excessively large reserve of adipose tissue (Schroeder & Staufenbiel, 2006).

Ketosis is a complex disease related to many on-farm factors, but its overall risk and severity can be addressed through nutrition and management strategies in efforts to limit the amount and speed of fat mobilization (Roche et al., 2013). This requires a way to measure condition consistently, so that fat cover on different dates can be compared and changes in cow energy reserves can be observed. Using a defined and systematic method to assign and record BCS scores fills this need, allowing producers and consultants to discuss cow fat cover in common terms, watch for patterns or warning signs that might indicate an issue, and assess if their cows are producing efficiently.

Despite the usefulness of performing BCS on a regular basis, it is not always practiced by herd managers; rather it is often left to the nutritionist or other consultant to complete. It is recommended that BCS should be performed weekly, or at very least biweekly to allow the manager to make informed decisions, and scoring time is expected to take an estimated average of 45 seconds per cow (Bewley & Schutz, 2008). Scoring 25% of the herd, this requires an additional 16 labor hours or more per year for just a 100-cow dairy to complete, or over 160 hours for a 1000-cow dairy. This labor and time required to complete scoring and maintain training for staff performing it (Bewley & Schutz, 2008), as well as the expectation for each employee to care for more animals due to increasing farm size and difficulty in finding skilled labor (Bell et al., 2018) are considered barriers to universal adoption of BCS practices at the producer level.

Established body condition scoring methodology is performed in quarter-point increments, requiring highly detailed observations of small physiological features to arrive at a final score. However, with such small nuances to distinguish quarter-point differences, analysis of BCS repeatability shows that scorers agree with each other substantially within 0.5 points, but only moderately agree on exact scores (Vasseur et al., 2013). I proposed scoring cows to the nearest half-point value, requiring less scrutiny to decide a cow's score. Comparisons of the accuracy and variation of assessing BCS in different increments has not been well described in literature. After scoring cows, a manager or consultant must then consider the stage of lactation to determine whether the cows are under, ideally, or over conditioned in that context before addressing the situation. I proposed creation of a scale to directly assess idealness of condition rather than taking this indirect route, which has not been well explored in the past.

Based on the information that BCS scoring is most highly repeatable within half point ranges (Vasseur et al., 2013), I hypothesized that scoring in half point increments (HALF) would not increase variation of managerial value compared to scoring in quarter point increments (QRT). I further hypothesized that categorically assessing cows as thin, ideal, or fat for current stage of lactation (CAT) would be interpretively similar to traditional BCS scores.

Therefore, the objectives of this study were to evaluate the variation in BCS, across several levels of scorer experience, based on scoring with the HALF or QRT methods, and to evaluate the accuracy of assigning a CAT score compared to interpretation of a traditional BCS score. If a scoring method yields similar results while requiring less inputs to complete, this could be a beneficial procedure for implementation.

Materials and Methods

Lactating Jersey cows (n = 20), balanced for stage of lactation (average DIM = 127 at trial initiation, 162 at trial conclusion) and housed at The Ohio State University's Waterman Dairy Farm, were selected to be scored each week by undergraduate student participants (n = 14 for 3 weeks, n = 13 for 2 weeks, n = 12 for 1 week). The average days in milk (DIM) of the cows was 127 at trial initiation and 162 at trial conclusion. Cows were classified as fresh (0 - 28 DIM), early lactation (29 - 100 DIM), mid lactation (101 - 200 DIM), or late lactation (greater than 200 DIM). Because BCS is a regular management tool employed at the dairy, no approval from the IACUC was required or sought for this study. Prior to trial initiation, participants received BCS training from study personnel on how to perform QRT, HALF, and CAT scoring, as well as an explanation of how cow body condition fluctuates over a lactation and what score ranges are desired at different stages. At the farm, an illustrated scoring guide (Elanco Animal Health, 1997) was provided for reference (available at all times), as well as a data collection sheet listing the ear tag numbers of cows to be scored and the current days in milk for each selected cow. For six weeks, starting in the last week of January and finishing in the first week of March, participants assessed each cow to assign QRT, HALF, and CAT scores, weekly. While they were encouraged to assist each other in locating and moving cows, especially for the safety of novice scorers, all participants were instructed to complete their scores without discussing or consulting with one another.

Participants had varying skill levels when entering the study, and each was classified as a Novice (learned about Dairy BCS scoring method for first time with study training), Beginner (had prior exposure to Dairy BCS scoring method before entering study, but had little to no experience utilizing it, i.e. no more than 1 - 2 times within a class setting), or Experienced (familiar with Dairy BCS scoring method before entering study, had previously performed scoring in multiple settings on-farm due to a prior job, judging team, or other field experience) scorer.

Upon completion of data collection, data analysis was generated using SAS software, Version 9.4. (SAS Institute Inc., Cary, NC, USA). This analysis included a mixed-model approach with fixed effects of scoring method, lactation stage, week, and scorer experience, and random effects of scorer and cow, and utilized repeated effect of week as appropriate. Further analysis of QRT versus HALF included residuals regressed against predicted BCS from HALF, centered to the mean (St-Pierre, 2003). To align the center values, the CAT method was plotted on a 2 - 4 scale; QRT and HALF were on a 1 - 5 scale.

Results

Results for influence of scoring method on mean score are summarized in Figure 1. Scoring method influenced mean score ($P < 0.01$) where CAT was lower than QRT or HALF by 0.055, while QRT and HALF were not different from each other ($P > 0.10$). Mean score did not differ by week or parity. However, Novice scorers estimated cows 0.24 points lower than scorers with more experience ($P < 0.01$). Variances of QRT and HALF (summarized in Figure 2) were not different ($P = 0.73$). Due to the difference in its scale of measurement, variance of CAT was not compared to that of QRT and HALF.

The residuals (observed - predicted, or in this case QRT - HALF) were regressed against the predicted (HALF) score to further investigate variation of QRT vs. HALF. This regression of residuals (summarized in Figure 3) indicated a mean bias ($P < 0.01$) where HALF underestimates herd BCS by 0.018 and a slope bias ($P < 0.01$) where HALF increased by 0.110 points greater than QRT for every whole point BCS.

Discussion

Several automated image technologies are being investigated for use in the dairy industry, with potential to decrease the labor and time necessary to obtain important information on cow body condition and to increase the accuracy of this information by removing the subjective human element (Bewley et al., 2008; Halachmi et al., 2008; Azzaro et al., 2011; Fischer et al., 2015). However, until this technology is widely accepted and made available, producers and consultants must implement a scoring methodology on-farm that is feasible to complete and gives them an accurate and repeatable assessment of the cows. Further, those systems represent extensive financial investment that may not be economically feasible in many small- to mid-size operations. The aim of this study was to propose a scoring method requiring less inputs to complete and to assess its validity to inform management decisions.

The mean scores obtained from the HALF and QRT methods were not statistically different, meaning that the two methods are equally accurate ways to assess body condition. The variances of HALF and QRT were also not statistically different, meaning both methods yielded a similar spread of results. While the regression of residuals shows HALF increases by a tenth of a point more than expected for every whole point, management decisions on the farm would not be impacted by this small of a BCS change, and the nature of the scoring system with 5 as the extreme limits this error. This fits well with previous research stating that scorers agree strongly on BCS to the nearest 0.5 points but have less agreement on exact QRT scores (Vasseur et al., 2013). Further, it is likely that excessive statistical power from the number of observers, cows, and weeks contributed to detection of a managerially insignificant, yet statistically significant, result.

The CAT score was statistically lower than HALF or QRT, meaning it was less accurate in comparison. A mean difference of 0.055 is still not large enough to be of management value, but CAT was not as effective at matching the standard QRT compared to HALF. I believe that more training could be required before scorers feel confident using this method, as it requires simultaneously assessing the cow's appearance while also considering how many days in milk she is and how much fat cover is ideal for the associated stage of lactation. As this would require additional training input and employee skill level, and possibly the creation of training materials, this may not be an ideal solution for producers and could incorporate more subjectivity of the scorer.

While an average scoring time of 45 seconds per cow has been estimated for QRT in previous literature (Bewley & Schutz, 2008), the time saved by the HALF method is unknown. Predicting a time savings of 15 seconds per cow, HALF has the potential to decrease the hours of labor necessary to perform scoring, saving a 100-cow dairy around $200 per year and allowing that labor to focus elsewhere, or an estimated $550 of consultant labor if that individual was the scorer, and these savings multiply for larger farms (see Table 1). In addition, the ability to complete HALF scoring more quickly could prevent the need to disturb the cow's routine in order to score, allowing this task to be added to another time cows are already being assessed, or to be recorded when simply walking pens rather than needing to have the cow in headlock so she cannot avoid the scorer. More investigation into the average time per cow needed to record HALF scores would be interesting and allow full assessment of the method.

Previous research has identified a lack of study cows scoring on either extreme of the BCS scale as a limitation (Bewley et al., 2008), and this study also observed this, with only a minimum score of 1.75 and maximum score of 4 being assigned throughout the experimental period. This lack of extreme cows BCS scores in the study prevents confident application of the HALF scale's usage in that range. While it is desired to not observe either extreme on-farm, such cows do exist and should be able to be scored accordingly.

This study was also limited in that only undergraduate students were recruited to participate, and that even those labeled as experienced scorers had only a few years of practice at most. In reality, farm employees, producers, and consultants are all a diverse population in regard to experience, age, background, and other factors, and future work is needed to determine how these factors influence the scores obtained. The observed tendency of novice scorers to assign scores 0.24 points lower is especially interesting. This could be explained by these individuals lacking familiarity with the lean appearance of dairy cows, which have been genetically selected to put energy towards milk production rather than adding body mass. Experienced scorers are not exempt from error, but can develop bias or lose accuracy if training is not maintained. Therefore, the ideal timing to assess employees and consider providing a refresher of training would also benefit from further investigation.

Conclusion

In conclusion, the HALF scoring method was an equally valid system of assessing body condition compared to the accepted QRT method. The CAT score was less accurate, and more extensive training of scorers may be necessary for it to succeed – not only in cow appearance, but also in understanding the physiological changes a cow undergoes throughout the lactation cycle. The HALF scoring method has the potential to decrease cow-side time and training time required to record BCS scores on-farm. Gaining interpretively similar data while reducing these inputs could encourage more producers to implement regular body condition scoring, allowing management teams to more informed management decisions to the benefit of the cows and people connected to the dairy.

References

Azzaro, G., M. Caccamo, J. D. Ferguson, S. Battiato, G. M. Farinella, G. C. Guarnera, G. Puglisi, R. Petriglieri, and G. Licitra. 2011. Objective estimation of body condition score by modeling cow body shape from digital images. J. Dairy Sci. 94:2126-2137.

Bell, M. J., M. Maak, M. Sorley, and R. Proud. 2018. Comparison of methods for monitoring the body condition of dairy cows. Front. Sustain. Food Syst. 2:80.

Bewley, J. M. and M. M. Schutz. 2008. An interdisciplinary review of body condition scoring for dairy cattle. The Prof. Anim. Scientist. 24:507-529.

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

Cainzos, J. M., C. Andreu-Vasquez, M. Guadagnini, A. Rijpert-Duvivier, and T. Duffield. 2022. A systematic review of the cost of ketosis in dairy cattle. J. Dairy Sci. 105:6175-6195.

Duffield, T. F., K. D. Lissemore, B. W. McBride, and K. E. Leslie. 2009. Impact of hpyerketonemia in early lactation dairy cows on health and production. J. Dairy Sci. 92:571-580.

Edmonson, A. J., I. J. Lean, L. D. Weaver, T. Farver, and G. Webster. 1989. A body condition scoring chart for Holstein dairy cows. J. Dairy Sci. 72:68-78.

Elanco Animal Health. 1997. Body condition scoring in dairy cattle. Accessed 12 January 2023. https://www.vet.cornell.edu/sites/default/files/1e_Elanco%20Cow%20Body_condition_scoring_V3.pdf.

Eversole, D. E., M. F. Browne, J. B. Hall, and R. E. Dietz. 2009. Body condition scoring beef cows. Virginia Cooperative Extension Publication 400-791.

Ferguson, J. D., D. T. Galligan, and N. Thomsen. 1994. Principal descriptors of body condition score in Holstein cows. J. Dairy Sci. 77:2695-2703.

Fischer, A., T. Luginbuhl, L. Delattre, J. M. Delouard, and P. Faverdin. 2015. Rear shape in 3 dimensions summarized by principal component analysis is a good predictor of body condition score in Holstein dairy cows. J. Dairy Sci. 98:4465-4476.

Halachmi, I., P. Polak, D. J. Roberts, and M. Klopcic. 2008. Cow body shape and automation of condition scoring. J. Dairy Sci. 91:4444-4451.

Henneke, D. R., G. D. Potter, and J. L. Kreider. 1984. Body condition during pregnancy and lactation and reproductive efficiency of mares. Theriogenology. 6:897-909.

N.R. St-Pierre. 2003. Reassessment of Biases in Predicted Nitrogen Flows to the Duodenum by NRC 2001. J. Dairy Sci. 86:344-350.

Roche, J. R., D. P. Berry, J. M. Lee, K. A. Macdonald, and R. C. Boston. 2007. Describing the body condition score change between successive calvings: A novel strategy generalizable to diverse cohorts. J. Dairy Sci. 90:4378-4396.

Roche, J. R., N. C. Friggens, J. K. Kay, M. W. Fisher, K. J. Stafford, and D. P. Berry. 2009. Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. J. Dairy Sci. 92:5769-5801.

Roche, J. R., J. K. Kay, N. C. Friggens, J. J. Loor, and D. P. Berry. 2013. Assessing and managing body condition score for the prevention of metabolic disease in dairy cows. Vet. Clinics of N. Amer.: Food Anim. Prac. 29:323-336.

SAS Software, Version 9.4. 2017. SAS Institute Inc., Cary, NC, USA.

Schroeder, U. J. and Staufenbiel R. 2006. Invited review: Methods to determine body fat reserves in the dairy cow with special regard to ultrasonographic measurement of backfat thickness. J. Dairy Sci. 89:1-14.

Vasseur, E., J. Gibbons, J. Rushen, and A. M. de Passille. 2013. Development and implementation of a training program to ensure high repeatability of body condition scoring of dairy cows. J. Dairy Sci. 96:4725-4737.

Wathes, D. C., M. Fenwick, Z. Cheng, N. Bourne, S. Llewellyn, D. G. Morris, D. Kenny, J. Murphy, and R. Fitzpatrick. 2007. Influence of negative energy balance on cyclicity and fertility in the high producing dairy cow. Theriogenology. 68S1:S232-S241.

Wildman, E. E., G. M. Jones, P. E. Wagner, R. L. Boman, H. F. Troutt Jr., and T. N. Lesch. 1982. A dairy cow body condition scoring system and its relationship to selected production characteristics. J. Dairy Sci. 65:495-501.

Figure 1. Mean BCS scores obtained from three methods. Animals were scored on a 1 - 5 scale, where 1 is emaciated and 5 is obese. Scores were assessed to nearest quarter-point increment (QRT) and half-point increment (HALF), and then a categorical score (CAT) of thin, ideal, or fat for current stage of lactation was assigned. QRT and HALF were not statistically different ($P > 0.10$), while CAT was significantly lower ($P < 0.01$) than QRT or HALF by 0.055. SE = 0.071.

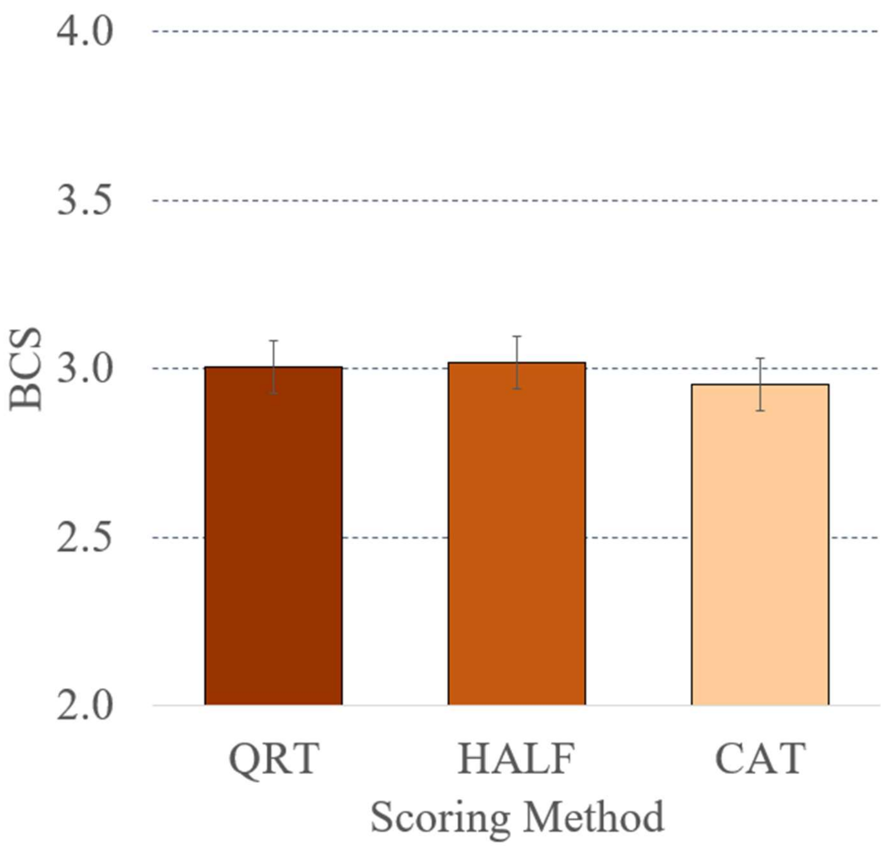

Figure 2. Variances of QRT and HALF scores. QRT and HALF were not statistically different ($P > 0.10$). CAT variance was not compared to QRT and HALF due to its different scale. SE = 0.0096.

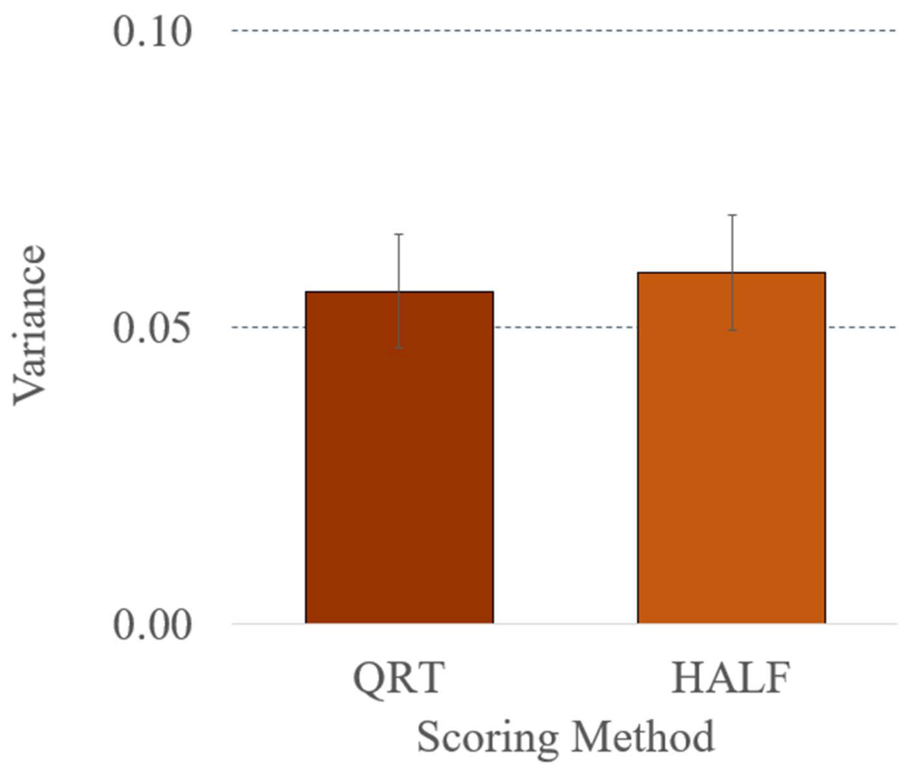

Figure 3. Regression of residuals (observed - predicted, or in this case QRT - HALF) against the predicted (HALF) score. The dots represent residual values, and the solid line represents a significant ($P < 0.01$) linear bias with a slope of -0.110 (SE = 0.012). The dashed line represents a significant ($P < 0.01$) mean bias (intercept) of -0.018 (SE = 0.0047).

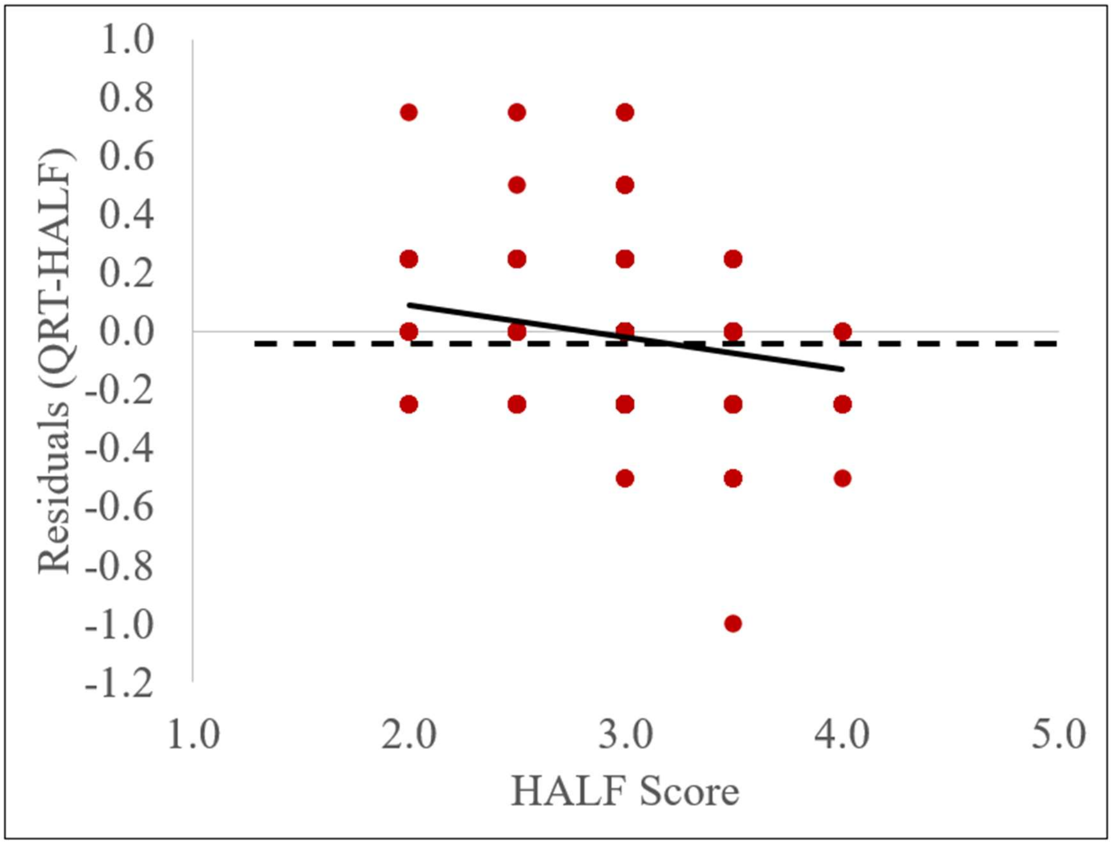

Table 1. Projection of labor savings achieved by implementation of HALF method to replace QRT method on a 100- or 1000-cow dairy. Calculations were performed considering 25% of the herd would be scored on a weekly basis. Hourly cost of manager or consultant labor were estimated at $35 and $100, respectively.

Inputs to Body Condition Scoring	Labor for 100 Cow Herd (hours/year)	Labor for 1,000 Cow Herd (hours/year)
45 seconds per cow*	16.25 hours	162.5 hours
30 seconds per cow^	10.83 hours	108.3 hours
Hours Saved	5.42 hours	54.2 hours
Manager Savings	@ $35/hour = **$190/year**	@ $35/hour = **$1,900/year**
Consultant Savings	@ $100/hour = **$550/year**	@ $100/hour = **$5,500/year**

*estimated time to score with QRT method, according to Bewley & Schutz, 2008

^projected time to score with HALF method, saving 15 seconds per cow