An Evaluation of Two Indirect Methods of Estimating Body Weight in Holstein Calves and Heifers

R. T. Dingwell,¹ M. M. Wallace, C. J. McLaren, C. F. Leslie, and K. E. Leslie

Department of Population Medicine, University of Guelph, Guelph, Ontario, Canada N1G 2W1

ABSTRACT

Monitoring the growth of replacement heifers is a useful management tool to assist producers in achieving a reasonable goal for age at first calving. Standard growth curves have been established, and heart girth tapes are widely available to estimate body weight (BW). Probably the easiest, and undoubtedly the most accurate, means of determining the actual BW of heifers is by using a calibrated electronic scale. However, if an electronic scale is not available, indirect methods of BW estimation are required. The hipometer is a new indirect tool that uses the external width between the greater trochanters of the left and right femurs to estimate BW. The purpose of this observational study was to evaluate the hipometer and the heart girth tape to estimate the BW of Holstein heifers, as compared with their actual weight recorded by an electronic scale. A total of 311 Holstein heifers in 4 research herds, ranging in age from 1 wk old to immediately prior to calving (24 mo), were used in this comparison. The mean BW of all heifers was 261 ± 124 kg. The Pearson values of the correlation between the scale and hipometer weights, and the scale and tape weights were 0.92 and 0.94, respectively. The concordance correlations of scale weight with hipometer and tape weights were 0.98 and 0.99, respectively. The agreement among the 3 methods, as assessed by the kappa statistic, was substantial for heifers aged 3 to 15 mo. However, poor to no agreement was observed in heifers younger than 3 mo, as well as at 15 mo of age or greater (kappa 0 to 0.18). This is of particular concern because these groups represent the age when dairy heifers would be weaned (<3 mo) and the age when breeding would normally commence (>15 mo). We concluded that the hipometer is an easy and useful alternative method of estimating the BW of Holstein heifers, particularly in heifers aged 3 to 15 mo.

Key words: body weight, heifer, Holstein, growth

INTRODUCTION

An overall objective of raising dairy heifers is to achieve sufficient weight and stature for calving at approximately 24 mo of age (Radostits et al., 1994), which has resulted in optimal profitability in Holstein replacement programs. A recent study of the current growth of Holstein heifers in US dairy farms revealed that heifers were heavier and taller at the withers than in the standards published 30 to 50 yr ago (Heinrichs and Losinger, 1998). Therefore, it is important to understand and monitor the performance of heifers and compare the results to known standards. Accurate knowledge of the rate of gain in BW in heifers allows for the creation of appropriate nutritional management programs (Donovan and Braun, 1987), assessment of feed efficiency, accurate determination of pharmaceutical doses, and assessment of the overall health status of the animals (Enevoldsen and Kristensen, 1997). Body weight information can also be used to determine the value of culled animals and the efficiency of rearing replacement heifers.

A variety of methods are used to measure or estimate BW accurately in Holstein heifers, the most accurate being to weigh animals individually on a traditional or electronic weigh scale. Surveys have determined that a majority of dairy producers feel that measurement with a weigh scale is too time-consuming and costly to implement (Heinrichs et al., 1992). As a result, indirect methods of estimating BW have been developed.

The most common and cheapest indirect method of estimating the BW of dairy heifers is with the heart girth tape. This method consists of a measuring tape placed around the circumference of the animal just behind the withers. The heart girth circumference corresponds to an estimated BW, which was devised by Heinrichs and Hargrove (1987) based on measurements of 5,723 heifers on commercial farms in Pennsylvania.

Previous studies have investigated the relationships among various measurements of growth in dairy cattle to predict BW. Heinrichs et al. (1992) and Enevoldsen and Kristensen (1997) found, by means of linear regression models, that hip width was one of the skeletal measurements more highly related to BW that was not greatly influenced by body condition. Recently, the hipo-

Received October 6, 2005.

Accepted May 24, 2006.

¹Corresponding author: dingwell@uoguelph.ca

Figure 1. Diagrammatic representation of the placement of the hipometer on the greater trochanters of the left and right femurs to estimate the BW of replacement heifers.

meter (Dairy Innovations, Alexander, NY) has been developed to estimate BW from the width between the hip joints at the point of the greater trochanters of the femurs (Skidmore, 2001). The hipometer consists of 2 arms that cup over the greater trochanters of the femurs (Figure 1). Based on the width of the caliper arms, BW can be estimated. Although the heart girth tape has been extensively validated with published results (Heinrichs and Hargrove, 1987), there are no validation reports of the hipometer method or a direct comparison of these 2 methods of assessing BW.

The objective of this study was to evaluate the hipometer and the standard heart girth tape for measurement of BW in Holstein heifers as compared with the actual weight, as measured by a scale, as well as to study the hipometer and heart girth tape over a range of heifer weights for sources of error.

MATERIALS AND METHODS

Study Population

The study animals included Holstein heifers from 3 research herds associated with the University of Guelph (Elora, Kemptville, and Ponsonby Research Stations), and the Shur Gain research herd in Burford, Ontario. These calves and heifers ranged in age from 1 wk to 24 mo of age, and were all raised in calf build-ings, calf hutches, or confinement barns with self-lock- ing head gates, depending on their age. The calves were fed whole milk and a commercial calf starter. The older

heifers were fed either a TMR or a component-based feeding program, which was formulated according to recommended nutritional guidelines from the National Research Council. No routine management, including feeding and housing, was altered for purposes of this trial.

Data Recording

Data were collected during one or more visits to each farm to maximize the data that were collected. A repeat visit was made to only a single farm, where the possibility existed of having reweighed heifers in an older age group from their original age group. A conscious effort was made not to have this happen. The heifers were sent through a chute and weigh scale system, and one individual would record the actual BW of the animal. The electronic scales were calibrated annually by each research herd as part of their standard operating procedures. Subsequently, the animal was released from the weigh scale and put in a locking head gate. The heifer was then weighed with the Nasco Holstein Dairy Heart Girth Tape (Nasco, Ft. Atkinson, WI) and the hipometer (Dairy Innovations). The hipometer measurement was performed by a different person, who was unaware of any previous measurements, to avoid investigator bias. With the heart girth tape method, the tape was placed around the heart girth of the animal and pulled tautly. The weight was read directly off the tape according to the circumference of the animal. With the hipometer, the forcep-like arms were cupped over the greater trochanters of the left and right femurs of the heifer (Figure 1). The weight was then recorded according to the width of the opening of the caliper arms by reading directly from the sliding scale of the hipometer (Skidmore, 2001). Periodically during each weighing session, the investigator used the calibration rod to ensure that consistent and appropriate tension was placed on the calipers.

Statistical Analysis

For means of comparison, the 311 observations were grouped into 1 of 8 age categories: <3 mo old, 3 to <6 mo, 6 to <9 mo, 9 to <12 mo, 12 to <15 mo, 15 to <18 mo, 18 to <21 mo, and >21 mo old. The mean weight values and associated measures of variation for each of the 3 methods were calculated using SAS (PROC UNIVARIATE, PROC MEANS, SAS, v. 8.1; SAS Institute, 1999–2000) overall and for each age category. A 2-sided *t*-test was used to detect any significant difference in the mean weight of animals as measured with the 2 methods at any age interval, with significance declared at $P < 0.05$. Pearson correlation coefficients (*r*)

Table 1. Comparison of BW measurements using an electronic scale, heart girth tape, and hipometer for different age groups of 311 Holstein heifers¹

Age group, mo	N	Median ²	Technique, kg of BW		
			Electronic scale	Hipometer	Heart girth tape
All ages	311	245	261 ± 124	258 ± 118	254 ± 123
<3.0	32	67	74 ± 25 ^a	77 ± 15 ^a	58 ± 18 ^b
3.0 to 5.9	18	130	131 ± 22	128 ± 22	129 ± 25
6.0 to 8.9	74	201	205 ± 36	203 ± 33	199 ± 33
9.0 to 11.9	100	258	266 ± 47	264 ± 47	261 ± 47
12.0 to 14.9	65	335	343 ± 55	339 ± 51	339 ± 51
15.0 to 17.9	8	506	484 ± 56	474 ± 53	477 ± 47
18.0 to 20.9	5	591	589 ± 22 ^a	544 ± 39 ^b	554 ± 29 ^b
≥21.0	9	600	612 ± 68	591 ± 47	598 ± 49

^{a,b}Values within the same row for a given group with different superscripts are different, $P \leq 0.05$.

¹Hipometer (Dairy Innovations, Alexander, NY); heart girth tape (Nasco, Ft. Atkinson, WI).

²Median value reported from the electronic scale weight recorded for each weight group.

were calculated to assess the linear association between the weights of animals measured by each indirect method and the weight as recorded by the scale. Agreement among the 3 methods was assessed by calculating kappa and the concordance correlation (PROC FREQ, SAS, v. 8.1; SAS Institute, 1999–2000). Kappa measures the amount of agreement, beyond chance, observed by each method. To categorize the data to calculate kappa, each weight measurement recorded by each weighing device was plotted on a standard normal growth curve for Holstein heifers (Heifer Growth Curve, Penn State University, University Park, PA). Pairwise agreement was met if both scales recorded a heifer to be of normal, light, or heavy weight for her age. All other disparate results were considered nonagreement. A simple kappa, as opposed to a weighted kappa, was calculated because the differences between the ordinal groupings (normal, light, heavy) were considered to be weighted evenly. The Pearson correlation measures the linear relationship between 2 variables but fails to detect any departure of data points from a line of perfect agreement (45°; Shoukri and Pause, 1999). Therefore, Pearson correlation coefficient often is not the most appropriate measure to use for comparison. To avoid this error, a concordance correlation was calculated. This concordance correlation evaluates the degree to which pairs of observations fall on a perfect 45° line through the origin. Any departure from this line will result in a concordance of less than one, even if the correlation coefficient is equal to one (Shoukri and Pause, 1999). As an additional means of evaluating the differences between the methods of measuring BW, Bland–Altman graphs of the difference versus the mean of pairs of weights collected on each heifer were constructed and visually checked (Christley and Reid, 2003).

To formally estimate rho (ρ), a 2-way random effects model was used (Shoukri and Pause, 1999). This model estimates the reliability of the different measurements recorded, but characterizes the effect of a randomly selected rater and can be represented by

$$y_{ij} = \mu + g_i + c_j + e_{ij}$$

where y_{ij} denotes the j th rating on the i th subject, μ is the overall population mean of the subjects, g_i is the true measure of the i th subject (variance σ_g^2), c_j is the additive effect of a randomly selected rater (variance σ_c^2), and e_{ij} is the error (variance σ_e^2). The variance of y_{ij} is $\text{var}(y_{ij}) = \sigma_g^2 + \sigma_c^2 + \sigma_e^2$; hence, the intraclass correlation used as an appropriate measure of reliability under this model is

$$R = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_c^2 + \sigma_e^2}$$

RESULTS

Mean weight data results are summarized in Table 1. The mean overall weight for all animals, as recorded by the scale, was 261 ± 124 kg. There was no significant difference between this weight and the overall mean weights as measured by the heart girth tape or the hipometer (254 ± 123 kg and 258 ± 118 kg, respectively). Examined across different age groups, the only differences among recorded weights occurred in the youngest heifers and in the heifers between 18 and 21 mo (Table 1). For heifers less than 3 mo old, the heart girth tape BW estimation was significantly different from the weight recorded by the scale (58 and 74 kg, respectively). For heifers between 18 and 21 mo, the average

Table 2. Agreement and correlation of BW estimations determined by the indirect methods of the hipometer and heart girth tape, relative to the actual BW measured by an electronic scale¹

Age group, mo	N	Hipometer				Heart girth tape			
		Kappa	95% CI	Concordance	R ²	Kappa	95% CI	Concordance	R ²
All ages	311	0.61	(0.5, 0.7)	0.98	0.92	0.63	(0.5, 0.7)	0.99	0.94
<3.0	32	0.13	(−0.1, 0.4)	0.01	0.92	0.00	(0.3, 0.3)	0.72	0.97
3.0 to 5.9	18	1.00	(0.5, 1.5)	0.91	0.97	0.45	(0.2, 0.7)	0.96	0.97
6.0 to 8.9	74	0.68	(0.5, 0.8)	0.92	0.92	0.87	(0.6, 0.9)	0.95	0.96
9.0 to 11.9	100	0.74	(0.6, 0.8)	0.96	0.95	0.64	(0.5, 0.8)	0.95	0.95
12.0 to 14.9	65	0.60	(0.4, 0.7)	0.87	0.87	0.53	(0.4, 0.7)	0.93	0.93
15.0 to 17.9	8	0.14	(−0.5, 0.8)	0.88	0.97	0.60	(0.1, 0.9)	0.87	0.97
18.0 to 20.9	5	0.00	(0.7, 0.5)	0.31	0.65	0.00	(0.6, 0.6)	0.00	0.00
>21.0	9	0.18	(0.1, 0.5)	0.71	0.82	0.18	(0.1, 0.5)	0.75	0.81

¹Hipometer (Dairy Innovations, Alexander, NY); heart girth tape (Nasco, Ft. Atkinson, WI). CI = confidence interval.

weights, as estimated by the hipometer (544 kg) and by the heart girth tape (554 kg), were both significantly different from the average scale weight (598 kg).

The agreement between the hipometer and the heart girth tape as compared with the weigh scale measurements is shown in Table 2. For the total population, the kappa statistics for estimations of BW were 0.61 and 0.63 for the hipometer and the tape, respectively. Pearson's correlation coefficients between the hipometer and scale weights, and the tape and scale weights were both 0.99. When age was controlled for, the correlation coefficients were 0.92 and 0.94 between the hipometer and scale, and the tape and scale, respectively. Simple scatter plots for both the hipometer versus the scale, and the heart girth tape versus the scale are shown in Figures 2 and 3, respectively. A general linear relationship is observed in each plot, with widening variation noted for each measurement above 500 kg. The concordance correlation values between the hipometer and scale, and the tape and scale were 0.98 and 0.99, respectively.

The kappa statistic was also used to compare agreement between the 2 indirect tests and the scale weight (Table 2). For the hipometer, the kappa values ranged from 0.00 to 1.00. The lowest kappa value was for the 5 heifers between 18 and 21 mo, and the highest was observed in the 18 heifers aged 3 to 6 mo. Similarly, the heart girth tape had kappa values ranging from 0.00 to 0.87 in the 18- to 21-mo and the 6- to 9-mo age groups, respectively. The lowest kappa values were in the less-than-3-mo and the 18- to 21-mo age groups. The highest kappa value was in heifers 6 to 9 mo of age. The Bland–Altman graphs of the differences and the mean difference of pairs of weights collected on each heifer again demonstrated that the agreement between methods of measuring weight is not perfect, especially for animals greater than 300 kg (Figures 4 and 5). For

Figure 2. Scatter plot of BW estimations as determined by the hipometer and the electronic scale. **Figure 3.** Scatter plot of BW estimations as determined by the heart girth tape and the electronic scale.

Figure 4. Difference versus mean for each pair of BW recorded by the electronic scale and the hipometer.

many of the heifers, there were reasonably large variations between the methods of measurement.

A 2-way random mixed ANOVA of reliability was performed. The intraclass correlation was calculated to be high, at 0.84, demonstrating that the measuring device used was not a significant factor. An age group $\times$ measuring device interaction was tested and found not to be significant ($P > 0.05$). There was a significant effect of farm ($P < 0.05$).

DISCUSSION

For the majority of weight measurements, the methods of indirect measurement did not differ from the

direct measurement. In the heifers aged less than 3 mo, the heart girth tape measurement was significantly lower than the scale measurements. This would suggest that further study of the heart girth tape for measuring the weight of very young Holstein calves may be necessary. For heifers in the 18- to 21-mo age interval, both the hipometer and the heart girth tape yielded lower estimated weights. It is noteworthy that a limited number of measurements were collected in some of the age intervals. Thus, to make an accurate comparison of the methods of measurement within specific age categories, those groups with low numbers of observations should be studied using larger numbers of heifers. Overall, there was no significant difference between the indirect methods of measurement and the electronic scale, suggesting that the hipometer may be used as a method of estimating the BW of Holstein calves and heifers.

Pearson's correlation coefficients for the hipometer and the heart girth tape, when compared with the scale, were very high. When controlling for age, the correlation coefficients for the hipometer and the tape remained high as well. This suggests that even when age differences were taken into account, the hipometer and the heart girth tape measurements were still closely related with the electronic scale measurement.

The high concordance in correlation values for the indirect methods of the hipometer and the tape, when compared with the scale, suggest that there was very little departure in observations from the perfect 45° line through the origin on the linear regression plot. In the 18- to 21-mo age group, the concordance values for the hipometer and the tape dropped drastically. These low values would suggest that there was much variation and departure of weights from what would be perfect. It is noteworthy that the majority of the heifers in this age category were from one research station, indicating that specific herd factors relating to nutritional management, body condition, or genetics of conformation may have played a predominant role in these differences.

Kappa was used to measure the amount of agreement beyond chance between the 2 indirect methods of measurement and the scale method. Kappa values of 0.4 to 0.5 indicated moderate levels of agreement (Martin et al., 1987). Therefore, the agreement beyond chance between the hipometer and the scale, and between the tape and the scale, would be considered substantial. However, it is important to note that although the overall kappa was moderate, there was no agreement to poor agreement in heifers older than 15 mo and under 3 mo. The low kappa values that were observed may be a result of the small sample size in those specific age intervals. More weight measurements in these age

Figure 5. Difference versus mean for each pair of BW recorded by the electronic scale and the heart girth tape.

intervals, perhaps as well as using more herds of origin, should be collected for an accurate comparison of methods in this age range.

In the random mixed model, it was clear that age was a significant predictor of weight, as was the farm of origin. This coincides with the previous results and discussion on the influence that feeding and genetics may have had on a particular farm. A measurement of BCS was not assigned to each animal during the visits to each farm. Undoubtedly, BCS is directly related to the feeding management, and consequently the body weight, of heifers. This may be of particular interest with older heifers, because differences in weight and size would be due mainly to BCS. Further examination of these variables would be required to explain any differences seen between farms. Had the purpose of this study been to describe the variability in BW, then BCS and withers height would have been recorded, as would have fatness and rumen fill, which are the major determinants of BW (Enevoldsen and Kristensen, 1997). Instead, this study evaluated the agreement of an indirect measurement of BW that was least likely to be influenced by BCS. Because skeletal development is relatively slow and progressive, hip width, as measured by the hipometer, should not have been greatly influenced by the BCS of the heifers (Enevoldsen and Kristensen, 1997).

Overall, the results obtained from the hipometer and the heart girth tape appear to correspond closely to the results obtained from the electronic weigh scale. The heart girth tape method has already been validated and is widely used (Heinrichs and Hargrove, 1987). Because the hipometer results also correspond closely to the scale measurements, it is apparent that the hipometer may be another useful method to estimate the BW of Holstein heifers indirectly. One can acknowledge that the heart girth tape is much more readily available on dairy farms than a calibrated electronic weigh scale. In addition, it is most likely the least expensive of all available methods, and can be performed cow side. However, the tape method still involves having very close contact with the animal to reach under and wrap the tape around its circumference. In some situations, this process can be inconvenient, time-consuming, and somewhat dangerous. In addition, the heart girth measurements may be difficult to perform consistently, in that a change in the position of the animal during measurement or the tightness with which the tape is pulled can affect the results. The hipometer requires much less close contact with the animal. One advantage of the hipometer method is its ease of use. The operator can walk down tie-stall aisles or behind animals in head gates and measure the BW relatively quickly and easily.

Similar to the heart girth tape, a concern when using the hipometer is the variance associated with different strengths and pressures when applying the device. It is important to emphasize that the hipometer must be used with an appropriate amount of pressure on the caliper arms, which is determined by repeated periodic use of the calibration rod. It is also obviously important that the cups on the hipometer arms be placed in the correct position over the greater trochanter of the left and right femurs. In the current study, the technicians were trained in placement and calibration of the rod. A formal interreader analysis was not performed because on any given day of data collection, a single technician recorded all weights by both means of indirect measurement. Another limitation of the hipometer is that it has not been calibrated or manufactured to measure larger animals, such as lactating cows. The caliper scale does not extend beyond 700 kg.

In conclusion, weighing heifers on a calibrated electronic scale remains the easiest and most accurate method of obtaining BW. However, when scales are not available, indirect methods are required. The hipometer performed very similarly to both the previously validated heart girth tape and an actual electronic scale measurement of weight for heifers of various ages. The correlation of measurements among all 3 measuring methods was high, and the agreement was substantial in heifers aged 3 to 15 mo. The hipometer may be a useful alternative method of estimating the BW of Holstein heifers. Further studies to validate this method are encouraged, particularly in heifers aged less than 3 mo and those at or above the breeding age.

ACKNOWLEDGMENTS

The authors acknowledge the managers and staff of the Elora, Ponsonby, and Kemptville Research Stations of the University of Guelph, and the Shur Gain Research Station in Burford, Ontario. Thanks to Dairy Innovations for supplying the hipometers and to the Ontario Ministry of Agriculture and Food for financial support.

REFERENCES

- Donovan, G. A., and R. K. Braun. 1987. Evaluation of dairy heifer replacement rearing programs. *Comp. Cont. Educ. Pract. Vet.* 9:F133–F139.
- Christley, R. M., and S. W. J. Reid. 2003. No significant difference: Use of statistical methods for testing equivalence in clinical veterinary literature. *J. Am. Vet. Med. Assoc.* 222:433–437.
- Enevoldsen, C., and T. Kristensen. 1997. Estimation of body weight from body size measurements and body condition scores in dairy cows. *J. Dairy Sci.* 80:1988–1995.

- Heinrichs, A. J., and G. L. Hargrove. 1987. Standards of weight and height for Holstein heifers. *J. Dairy Sci.* 70:653–660.
- Heinrichs, A. J., G. W. Rogers, and J. B. Cooper. 1992. Predicting body weight and wither height in Holstein heifers using body measurements. *J. Dairy Sci.* 75:3576–3581.
- Heinrichs, A. J., and W. C. Losinger. 1998. Growth of Holstein dairy heifers in the United States. *J. Anim. Sci.* 76:1254–1260.
- Martin, S. W., A. H. Meek, and P. Willeberg. 1987. *Veterinary Epidemiology: Principles and Methods*. Iowa State University Press, Ames, IA.
- Radostits, O. M., K. E. Leslie, and J. Fetrow. 1994. *Herd Health: Food Animal Production Medicine*. 2nd ed. W. B. Saunders, Philadelphia, PA.
- SAS Institute. 1999–2000. *SAS User's Guide: Statistics*, Release 8.1. SAS Inst., Inc., Cary, NC.
- Shoukri, M. M., and C. A. Pause. 1999. *Statistical Methods for Health Sciences*. 2nd ed. CRC Press, Boca Raton, FL.
- Skidmore, A. 2001. Hipometer for estimation of body weight using the hip width measurement [Package insert]. Dairy Innovations, Alexander, NY.