Comparison of Models to Identify Lame Cows Based on Gait and Lesion Scores, and Limb Movement Variables

P. G. Rajkondawar,*¹ M. Liu,† R. M. Dyer,‡ N. K. Neerchal,† U. Tasch,§ A. M. Lefcourt,§ B. Erez,∥ and M. A. Varner#

*Bou-Matic, LLC, Madison, WI 53708

†Department of Mathematics and Statistics, University of Maryland at Baltimore County (UMBC), Baltimore 21250

‡Department of Animal and Food Sciences, University of Delaware, Newark 19717

§Department of Mechanical Engineering, UMBC, Baltimore 21250

|| Agriculture Experiment Station (Central Maryland Research and Education Center), University of Maryland, Ellicott City 21042

#Department of Animal and Avian Science, University of Maryland, College Park 20742

ABSTRACT

Bovine lameness results in pain and suffering in cattle and economic loss for producers. A system for automatically detecting lame cows was developed recently that measures vertical force components attributable to individual limbs. These measurements can be used to calculate a number of limb movement variables. The objective of this investigation was to explore whether gait scores, lesion scores, or combined gait and lesion scores were more effectively captured by a set of 5 limb movement variables. A set of 700 hind limb examinations was used to create gait-based, lesion-based, and combined (gait- and lesion-based) models. Logistic regression models were constructed using 1, 2, or 3 d of measurements. Resulting models were tested on cows not used in modeling. The accuracy of lesion-score models was superior to that of gait-score models; lesionbased models generated greater values of areas under the receiving operating characteristic curves (range 0.75 to 0.84) and lower mean-squared errors (0.13 to 0.16) compared with corresponding values for the gaitbased models (0.63 to 0.73 and 0.26 to 0.31 for receiving operating characteristic and mean-squared errors, respectively). These results indicate that further model development and investigation could generate automated and objective methods of lameness detection in dairy cattle.

Key words: dairy cow lameness, automated lameness detection, gait score, lesion score

INTRODUCTION

Lameness in dairy cattle is associated with economic losses and cow welfare concerns in the dairy industry (Greenough et al, 1997; Whay and Main, 1999). Problems such as inaccurate diagnosis and producer desensitization may exacerbate the problem (Wells et al., 1993; Mill and Ward, 1994; Whay et al., 2003). Diagnostic accuracy has been hindered by subjectivity and by low between- and within-observer reproducibility of diagnosis (Mill and Ward, 1994). It has been reported that producers correctly identify only 40 to 45% of lame cattle (Wells et al., 1993). Others have found that the probability of observer and expert agreement on exact estimates of 9 classes of gait scores (GS) ranged between 25 and 47% and rose to 80% if one class difference was permitted between observer and expert (Engel et al., 2003). Observer training inconsistently improved the agreement between observer and expert estimates, but the probability of agreement on exact score estimates remained low, at 47%.

Methods of objective, automated scoring may enable earlier, more accurate detection of lameness. Early, accurate detection could lead to more efficacious practices that reduce lameness in commercial dairy units. Cattle caretakers, claw trimmers, and veterinarians normally detect lame cows by visual observation. One popular method for detecting lameness is to score cows from 1 (sound) to 5 (severely lame) based on subjective observations of postural and stride characteristics (Sprecher et al., 1997). To begin introducing automated, continuous, and objective methodology into lameness detection, the authors recently developed an automated system for detecting lameness based on vertical ground reaction force (GRF) measurements of individual limbs (Rajkondawar et al., 2002b; Tasch et al., 2004). The system is called the reaction force detection (RFD) system and consists of 2 parallel-force plates with leveling platforms before and after the plates. Side railings guide the cows so that the left limbs contact the left plate and right limbs the right plate. Vertical forces measured over time for each plate can be used to calculate a number of limb movement variables (LMV).

Received February 11, 2005. Accepted June 20, 2006.

¹Corresponding author: parimal@boumatic.com

4268

Table 1. Number of weeks in which cows were enrolled in the study

No. of cows	No. of weeks
	— (Modeling data) —
2	4
1	6
2	9
1	11
2	12
1	13
2	15
2	20
1	21
1	22
1	23
1	26
4	27
	—— (Test data) ——
10	6

The step of cattle is divided into a weight-bearing (stance) and non-weight-bearing (swing) phase. The weight-bearing phase is initiated at the time the foot lands on the ground and is terminated after push-off when the foot is lifted off the ground. During this phase, vertical and horizontal GRF are transmitted by the foot to the ground. The vertical GRF increases until midstance and then decreases until the foot is lifted off the ground (Scott, 1988, 1989; Van der Tol et al., 2003). Lameness changes the maximal force applied by lame limbs, relative times of peak loading, and deacceleratory and acceleratory longitudinal forces in the lame as well as the nonlame contralateral limbs (Scott, 1989). Conceivably, these changes could provide the basis for the development of objective methods for lameness detection. In this regard, a pilot study demonstrated that 5 LMV generated from the RFD system can be used to detect lameness (Rajkondawar et al., 2002a).

The objective of the current study was to determine how best to use the 5 LMV to detect lameness. An additional objective was to develop statistical models that related measures of foot and hind limb health to LMV derived from the RFD system. The effect of including more days in model development also was investigated. The models were then tested using a previously unevaluated set of cows from the same herd.

MATERIALS AND METHODS

Lesion and gait scores were measured weekly in dairy cows at the University of Maryland Research Farm in Clarksville, Maryland (Table 1). For 3 d before the gait and lesion examination, vertical GRF was measured daily to allow calculation of the 5 LMV depicted in Table 2. To assess detection capabilities, the 5 LMV calculated for individual hind limbs were related to corresponding GS or lesion scores (**LS**) by statistical models.

Because both GS and LS were ordinal measures, logistic regression was used to develop the statistical models. The dependent variables were ordinal scores, and the independent variables were the 5 LMV. Data for the 5 LMV for up to 3 d before gait and lesion observation were used for modeling. The resulting models were evaluated using separate data from the modeling data. The validity of the models was determined by comparing 1) differences between GS and LS and model-derived lameness indices; and 2) areas under the receiving operating characteristic (ROC) curves (Hosmer and Lemeshow, 2000).

Experimental Design and Procedures

Cows were fed a balanced TMR consisting primarily of corn and alfalfa silage; grass hay; soybeans; cotton-seed; ground, shelled corn; and vitamins and minerals. They were housed in a free-stall facility with stalls containing rubber mats and sawdust, and hooves were trimmed by the herdsmen following an undefined procedure at the time of dry off. Floors of the alleyways were deeply grooved cement with crosshatched 6.4-cm squares. As cows were assigned to the study, the hooves were trimmed by the veterinarian according to the Dutch technique guidelines (van Amstel et al., 2000) and any lesions were noted and scored. Gaits were scored for all 4 limbs. Regardless of the number, score, and type of lesion, when all 4 limbs received a score of 1, cows were not considered lame.

Data were collected from 10 to 15 cows at a time. Cows were allowed to walk through an RFD system located in the return alley from the parlor during 2 to 3 h before milking. Each cow in the study was examined once weekly by the same veterinarian. The veterinarian assigned GS and LS, and treated cows as necessary. Cows with normal GS in all 4 limbs (score = 1) for 2 consecutive weeks were released from the study. The herdsman randomly selected replacement cows without regard for gait status from the low-production group (<16 kg of milk/d) or the dry cow lot. The modeling data were constructed during a 6-mo period using 21 Holstein cows of mixed age. A test data set was acquired using 10 cows during a subsequent 6-wk period (Table 1).

LMV

Vertical GRF measured over time for individual hind and forelimbs were used to calculate 5 LMV (Rajkondawar et al., 2002b). The maximum of the sum of instantaneous GRF for a cow was used as an estimate of her BW. To enable between-cow comparisons, the LMV were scaled to dynamic BW.

Table 2. Limb movement variables (LMV) used in the study

LMV	Units	Description
$\begin{array}{c} \text{PGRF} \\ \text{AGRF} \\ \text{ST} \\ \text{I} \\ \text{GRF}_{\omega} \end{array}$	s s s 1/s	Peak ground reaction force (GRF) of a tested limb Average GRF of a tested limb Stance time = time during which a limb is in contact with the floor Impulse = integral of the GRF with respect to time Area under the Fourier transformed curve of a GRF signature

GS

Once weekly, cows were first observed in a stance and then while walking in a straight line. To facilitate identification of lame limbs, cows also were circled to the right and to the left. All examinations were performed in an alleyway on concrete flooring. Gait observations were performed by observing the cow perpendicular, parallel, posterior, and anterior to the line of travel. During gait evaluation, cows and limbs were observed for freedom of motion, left- and right-sided stride length, length of the anterior and posterior swing phases, symmetry and arc of the foot flight, foot placement relative to body position and limb axis, foot rotation during weight bearing, symmetry of weight distribution at the walk and stance, and position of the top line at a walk and in a stance. A score of 1 to 5 was assigned for each limb, as described in Table 3. The scoring procedure is an adaptation of other methods (Wells et al., 1993; Sprecher et al., 1997). To score 2 limbs on the same cow, the lamer of the 2 limbs was assigned the GS (Table 3) normally assigned to the cow. The unscored second limb was assigned the same score or a lower score from Table 3, depending on the relative level of gait changes displayed in the second limb.

LS

The interdigital area was brushed with water and soap and then examined by vision and palpation. The digits and bulbs of the heel were manually separated and examined. The sole was cleaned, trimmed if appropriate, and examined for puncture wounds, lacerations, and hemorrhage within the white line or epidermis of the sole. Localization of pain, when necessary, was established with a hoof tester. Pressure was applied from the dorsal wall to the sole across regions 3, 4, and 5 (Shearer et al., 2004) and then along the axial—abaxial axis in regions 4, 5, and 6. Claw compression was manually increased until the sole or wall displaced under the arm of the hoof tester. A positive pain response was recorded when cows reacted to 3 repeated compression tests along the same axis.

Lesions were diagnosed as described elsewhere (Bergstem, 1997; Collick et al., 1997; Leach et al., 1998). All lesions were scored as summarized in Table 4. A lesion score of 1 signified no visible lesion.

Independence of Measurements

The experimental design included repeated measurements of the 5 LMV over time for every cow. Procedure GENMOD (SAS Inst. Inc., Cary, NC) was used to test whether these repeated measurements violated the independence assumption of the logistic regression analysis used in this study. Using procedure GENMOD, a scale parameter was estimated that represented the extent of lack of independence attributable to the repeated measures taken on individual limbs for the same

Table 3. Method for assigning a gait score (GS)¹

Classification	GS	Gait characteristics
Sound	1	Cows walk freely with unrestricted motion and symmetry in the stride. No postural abnormalities are exhibited.
Mildly lame	2	No visible gait abnormalities, but an arched top line develops with the walk. Top line may be normal or arched and stride length may be slightly shortened on the inside limb during circling.
Moderately lame	3	Gait is tentative and the top line is arched at a walk and a stance. Abnormalities in gait are apparent at a straight walk and do not require circling to become visible.
Lame	4	Cows are reluctant to rise and do so with difficulty. Weight bearing at a stance is distributed asymmetrically with reluctance to move and bear weight on the affected limbs. The top line is profoundly arched at a stance and walk and the gait is tentative and slow.
Severely lame	5	Cows are unable or unwilling to rise, move, or bear weight on the affected limb(s). The limb is held off the ground, gently rested on the ground, or carried in a fixed position.

¹Adapted from Sprecher et al. (1997), Murray et al. (1996), and Wells et al. (1993).

Table 4. Assignment of lesion score (LS)

Lesion	Description
Sole ulcer	Circumscribed areas of red hemorrhagic sole sometimes lying beneath or within a variable zone of yellow, soft horn tissue (score 2). Ulcerations with corium infection, exudation, and granulation tissue formation or infection of deeper structures, or both (score 3).
White line disease	Blackened, separated axial or abaxial junction between the wall and sole with tracts penetrating through the crest and terminal horn into the underlying epidermis and dermis of the crest and terminal papilla (score 2). Septic white line disease associated with drainage from the coronet of regions 7 and 8 or the heel bulb in region 6 (score 3).
Vertical crack and separation of the wall	Fissures in the periople, coronary, and wall segment of the claw (score 2). Fissures associated with infection of the periople cushion, coronary cushion, and the corium (score 3).
Puncture wound of the sole	Punctures appeared as punctate or elongated disruptions in the sole horn with or without subsolar abscess formation (score 2). Draining tracts may emerge from the skin–periople junction at the base of the bulb (score 3).
Sole bruising	A small (score 2) or large (score 3) well-circumscribed, discretely demarcated, dark red or maroon entrapment of blood in or under the sole.
Interdigital dermatitis	Superficial dermatitis, foul odor, and erythema of the interdigital skin with no heel horn erosion (score 2). Progression to ulceration; 1- to 2-mm-deep fissures; exudation, hyperplasia, or both of the interdigital skin with heel horn destruction in the bulb (score 3).
Interdigital pododermatitis (interdigital phlegmon)	Edema, erythema of the interdigital area extending to the periople and coronette, and moderate claw separation (score 2). Edema, fissures, and exudation of the interdigital space; edema of the pastern and fetlock; foul odor and marked claw separation (score 3).
Digital dermatitis	A 2- to 6-cm-diameter circumscribed circular or oval, hyperemic erosion or a concave bed of red or blackened granulation tissue surrounded by an elevated border of thickened epithelium (score 2). A 3- to 6-cm-diameter wartlike lesion with variable numbers of tentacle-like projections on the plantar or palmar skin bordering the interdigital space (score 3).
Interdigital fibroma	Small- to moderate-sized (1- to 2-cm-wide, 1-to 2-cm-deep) hyperplastic areas of the interdigital integument (score 2). Larger than 2-cm-wide and 2-cm-deep hyperplastic interdigital integument (score 3).
Hemorrhage	A diffuse, poorly demarcated brushlike pink or small red or yellow discoloration of the sole horn, or axial and abaxial white line (score 2). A diffuse, poorly demarcated brushlike, moderate to severe bright red or large bright red or purple discoloration of the sole horn, or axial and abaxial white line (score 3).

¹Adapted from Dyer et al. (2004).

cows; an estimate of 1 corresponded to perfect independence. The reader is referred to Morel et al. (2003) for a description of the estimation methodology used in procedure GENMOD.

Modeling Approach

Analyses were based on data collected for individual hind limbs, and all weekly sets of measurements for individual limbs were considered to be statistically independent experimental trials. Several models of lameness were developed using GS, LS, and combined scores (CS). The CS was defined as the maximum of contemporaneous GS and LS, or max {GS, LS}, for a single limb. Gait and lesion scores for individual limbs were related to the 5 LMV for corresponding limbs using logistic regression (Hosmer and Lemeshow, 2000). Because of a very limited number of GS and LS greater than 3, raw

scores of 4 and 5 were not included in the development of the lameness models. Therefore, the GS and LS used in the analysis had ordinal values of 1, 2, or 3. Logistic regression models were developed to generate predictive equations for the probability that a limb had been assigned a specific score as an S-shaped function of the 5 LMV. [*Note*: The mathematical form of an S-shaped function is $f(x) = e^{\beta_1 x}/1 + e^{\beta_2 x}$]. The general form of these models is as follows:

$$P(\text{score} = 1) = \frac{e^{\alpha_1 + f(\text{LMV})}}{1 + e^{\alpha_1 + f(\text{LMV})}}$$

$$P(\text{score} = 2) = \frac{e^{\alpha_2 + f(\text{LMV})}}{1 + e^{\alpha_2 + f(\text{LMV})}}$$

$$P(\text{score} = 3) = 1 - P(\text{score} = 1) - P(\text{score} = 2)$$

where P(A) is the probability that A is true, and f represents a weighted linear combination of the 5 LMV. The coefficients for these models were estimated using the maximum likelihood methods applied to the underlying multinomial distribution. For details, the reader is referred to Hosmer and Lemeshaw (2000) and Dyer et al. (2004). These cumulative logistic regression models are used widely in the statistical literature to model ordinal scores (Hosmer and Lemeshaw, 2000). Five LMV were calculated for each limb every day. Models were constructed based on 1 d (1D), 2 d (2D), and 3 d (3D) of LMV observation. Therefore, the 1D, 2D, and 3D models used 5, 10, and 15 LMV, respectively. Based on the above probabilities, a lameness index (LI) was calculated as a weighted sum:

$$LI = \sum_{i=1}^{3} P(score = i) \times score,$$

where "score" was GS, LS, or CS.

Model Validation

Model validity was evaluated using sums of squared errors and ROC (Hosmer and Lemeshow, 2000; Appendix A). The sum of squared error of the gait score ($E_{\rm GS}$) was calculated as

$$E_{GS} = \sum_{N} (GS - LI_{GS})^{2},$$

where GS was the gait score and $LI_{\rm GS}$ was the lameness index estimated by the model being tested. The summation extended over all N data points, and $E_{\rm LS}$ and $E_{\rm CS}$ were similarly defined.

The area under the ROC curve (**AUC**) represents the level of discrimination achieved by a model. The AUC values can be used to compare the performance of various models much as correlation coefficients are used to compare results of linear regression models. Interpreta-

tions of AUC values were as described in Hosmer and Lemeshow (2000): 1) $0.7 \le AUC < 0.8$ (acceptable discrimination); 2) $0.8 \le AUC < 0.9$ (very good discrimination); and 3) $0.9 \le AUC$ (excellent discrimination). A detailed description of ROC in terms of true and false positives, and true and false negatives is provided in Appendix A.

RESULTS

Means and standard deviations of the 5 LMV used for computation of the lameness models are listed in Tables 5 and 6. Note the elevated standard deviation values of the stance time and the small number of limbs with score values of 4 and 5. Therefore, data from these limbs were not included in the development of the lameness models. Table 7 shows cross-correlations among the 5 LMV for the entire modeling data set. The scale parameter estimates, which represent the extent of lack of independence attributable to repeated measures collected from the same cow, ranged between 1.04 to 1.69, and were always ≥1. A scale parameter estimate of 1 corresponds to perfect independence. For GS, LS, and CS, the smallest values (1.27, 1.04, and 1.25, respectively) occurred for the 3D model. Thus, correlations among observations from the same limb were best accounted for by the inclusion of the LMV from the previous days, demonstrating an advantage of the 3D model. These overall small values provided justification for the assumption that all weekly sets of measurements for individual limbs could be considered to be statistically independent experimental trials.

Mean-squared errors and AUC values for the various models are listed in Table 8. Note that the GS models possessed greater mean-squared errors and smaller AUC values than the LS models (i.e., the LS models predicted LS more effectively than the GS models predicted GS). In particular, notice the unacceptable AUC values of the 3D GS models (AUC < 0.7). The AUC

Table 5. Means (standard deviations) of limb movement variables (LMV) for individual hind limbs in the modeling data set by lesion score¹

	Lesion score				
LMV	1	2	3	4	5
Observations, no.	922	1,436	526	38	2
PGRF	0.41(0.07)	0.36(0.07)	0.34(0.08)	0.31(0.08)	0.23 (0.03)
AGRF	0.30(0.05)	0.25(0.06)	0.25(0.16)	0.24(0.07)	0.15 (0.01)
ST	1.29(0.85)	1.46 (1.10)	1.23 (0.91)	1.19(0.78)	1.32 (0.10)
I	0.38(0.22)	0.35(0.22)	0.29(0.20)	0.26(0.15)	0.20 (0.01)
GRF_{ω}	0.39 (0.08)	0.33 (0.09)	$0.32\ (0.09)$	0.32(0.11)	0.22 (0.01)

 $^{^{1}}$ PGRF = Peak ground reaction force (GRF); AGRF = average GRF; ST = time during which a limb is in contact with the floor; I = integral of GRF with respect to time; and GRF $_{\omega}$ = area under the Fourier transformed curve of a GRF signature.

 $\textbf{Table 6.} \ \ Means \ (standard \ deviations) \ of \ limb \ movement \ variables \ (LMV) \ for \ individual \ hind \ limbs \ in \ the \ modeling \ data \ set \ by \ gait \ score^1$

	Gait score				
LMV	1	2	3	4	5
Observations, no.	1,479	1,262	175	18	9
PGRF	0.39(0.08)	0.35(0.07)	0.31(0.08)	0.30(0.08)	0.28(0.13)
AGRF	$0.28\ (0.06)$	0.26(0.11)	0.23(0.06)	0.21(0.07)	0.21(0.10)
ST	1.37(0.85)	1.39(1.16)	1.13(0.79)	1.29(1.09)	0.82(0.20)
I	0.38(0.21)	0.33(0.24)	0.25(0.16)	0.24(0.11)	0.17 (0.08)
GRF_{ω}	0.37(0.09)	0.33(0.09)	0.31 (0.10)	0.28(0.09)	0.29(0.14)

 1 PGRF = Peak ground reaction force (GRF); AGRF = average GRF; ST = time during which a limb is in contact with the floor; I = integral of GRF with respect to time; and GRF $_{\omega}$ = area under the Fourier transformed curve of a GRF signature.

values for the CS model were the same as those for the LS model.

DISCUSSION

The original intent of this project was to determine the sensitivity of the RFD system by monitoring the recovery of lame cows following treatment. Unfortunately, environmental conditions on the test farm produced a situation in which cows often acquired a new, unrelated lameness before they had fully recovered from their prior lameness. This situation resulted in a number of repeated trials using the same cows, albeit for different lameness conditions. The repeated trials using a limited number of 21 cows and the environmental conditions may have created limitations for broad application of the derived statistical models.

Other confounding factors may have had a potentially negative impact on the robustness of the lameness models reported herein. The nature of the lameness problem in the experimental population may have been problematic. Lameness was associated with single or multiple lesions within claws, single or multiple lesions between claws, and lameness in single or multiple limbs. It is unclear how single or multiple lesions between and within claws, or single and multiple limb lameness in-

Table 7. Correlation matrix of limb movement variables (LMV) for the modeling data

		LMV				
LMV^1	PGRF	AGRF	I	ST	$\mathrm{GRF}\omega$	
$\begin{array}{c} \operatorname{PGRF} \\ \operatorname{AGRF} \\ \operatorname{I} \\ \operatorname{ST} \\ \operatorname{GRF}_{\omega} \end{array}$	1	0.611 1	0.326 0.060 1	0.013 0.196 0.912 1	0.774 0.565 0.067 0.205	

 $^{1}PGRF$ = Peak ground reaction force (GRF); AGRF = average GRF; ST = time during which a limb is in contact with the floor; I = integral of GRF with respect to time; and GRF_ ω = area under the Fourier transformed curve of a GRF signature.

fluenced the process of modeling. In addition, the models were generated with data restricted to one type of herd with one type of management, housing, ration, and flooring. Despite these potential limitations, the data were sufficiently robust that the general relations in terms of the relative importance and value of the 5 LMV, and the relative ability to associate the 5 LMV to GS and LS were of value for refining lameness detection models. Evidence of the robustness of the data was the finding that the effect of lack of independence attributable to repeated measures diminished when LMV from additional days were included in the models.

Cost is an important factor to consider when developing systems for agricultural use. To reduce the potential cost of a commercial system to detect lameness, the RFD system used in this study measured only vertical reaction forces. An earlier study demonstrated that 5 LMV derived from vertical GRF could be used to identify lame cows (Rajkondawar et al., 2002a). However, a variety of changes in the GRF in both the vertical and longitudinal components are associated with lameness (Scott, 1989; Clayton et al., 2000). Thus, measurement of all components of GRF could potentially provide a more robust approach to lameness modeling. The tradeoff would be a greatly increased cost of the measurement system.

Table 8. Squared error (E), number of observations (N), and area under the curve (AUC) values for models including 1, 2, or 3 d of limb movement variable data (1D, 2D, and 3D, respectively) using gait, lesion, or combined scores as the dependent variable

Model	E	N	E/N	AUC
1D gait	29.80	105	0.284	0.708
2D gait	26.93	105	0.256	0.734
3D gait	32.33	105	0.308	0.632
1D lesion	14.30	100	0.143	0.752
2D lesion	13.21	100	0.132	0.809
3D lesion	16.23	100	0.162	0.840
1D combined	16.03	104	0.154	0.750
2D combined	15.23	104	0.146	0.794
3D combined	17.49	104	0.168	0.842

The mean of the 5 LMV (Tables 5 and 6) tended to decrease with increasing LS and GS. However, changes within any single LMV were not sufficient to distinguish sound from lame cows within either GS or LS. These findings are in accord with Scott (1989), who showed that bovine lameness was associated with a variety of inconsistent changes in GRF. Inconsistent GRF changes between and within lame cows would increase the variability of the 5 LMV and therefore decrease the discriminatory value of any single LMV in lameness detection. For equines, normal (nonlame) animals showed 5 to 8% variability in vertical GRF and 2 to 6% variability in stance time (Hodson et al., 2001; Khumsap et al., 2001). Lameness increased the variability 3- to 4-fold (Clayton et al., 2000). In our study, variability tended to increase with lameness, but did not approach the magnitude reported by Clayton et al. (2000).

Variability of the 5 LMV measured in our study generally exceeded that observed in equine studies for a few plausible reasons. It is presently unclear how much of the variability can be attributed to the RFD, the cows, or the methodology. Normal limbs in this study included all individual hind limbs scored sound as normal gait or normal gait with the absence of lesions. Socalled normal limbs included limbs on completely sound cows or the single sound limb on a cow with unilateral lameness. This approach could be problematic particularly in the latter case. Compensatory changes in vertical forces applied to the contralateral sound limb of a unilaterally lame cow could be expected to generate vertical GRF that differ from GRF generated by normal limbs of a bilaterally sound cow with symmetrical weight bearing (Scott, 1989). Differences could be expected to increase LMV variability within the pool of normal limbs used in our modeling. Similar arguments could be raised about LMV for any limb contralateral to a lame limb. Second, LMV data used in the equine studies were selected from a larger pool of force signatures generated by multiple repeated passes. This selection process likely resulted in an underestimation of the variability in LMV in many equine studies. In this bovine study, all 5 LMV were derived from single observations. This method of data assimilation is considerably different from that used in equine studies and may present a more accurate assessment of the variability in the 5 LMV of sound and lame limbs.

Results showed that the lesion models identified cows with lesions better than the gait models identified them with abnormal gaits. The reason for this difference is not clear. One possibility may be that identification of lesions was less subjective than identifying changes in gait characteristics. Earlier studies showed that ob-

server agreement tended to be good when gaits of moderately lame cows or limbs were scored. However, observer agreement was much poorer when cows and limbs with subtle lameness were scored (Wells et al., 1993; Mill and Ward, 1994; Keegan et al., 1998). In the current study, subtle lameness dominated the experimental group in that 34, 46, and 18% of the limbs in the investigation scored 1, 2 and 3, respectively. Although we approached the difficulties of scoring subtle lameness with a meticulously detailed lameness examination, the risk of subjectivity and low reproducibility in gait scoring may have remained problematic because of the large prevalence of mild lameness.

Accuracy of the CS model was no greater than the LS model because LS were generally \geq GS for most limbs. The magnitude of the maximum lesion severity score was \geq GS in 92.4% of the 656 limbs with LS and GS. Thus, modeling lameness by selecting the greater of the 2 scores for the CS model produced results that were nearly identical to that utilized in the LS model.

Another confounding problem is the temporal nature of the data collection. For each data record, the 5 LMV were measured daily and the corresponding GS and LS were obtained on the day after the last measurement of the 5 LMV. Conceivably, gait adjustment to accommodate pain, changes in lesion sensitivity, or movement of the nociceptive threshold (Whay et al., 1998) during the 24- to 96-h period of LI computation could have generated disparities between LI scores and GS or LS, or both. The AUC values for the 1D, 2D, and 3D lesion and gait models may support this contention. The AUC values for the GS models worsened with additional days of data. This could reflect the very fluid nature of paindriven changes in locomotion with time. In contrast, AUC values for the LS models improved as additional days were incorporated into the model (Table 8). This may reflect the inability of lesions (compared with gait) to change appearance during the same time period.

Opportunities also exist to improve the lesion scoring system. An improved lesion scoring system might allow construction of more accurate detection models. The scoring scheme used in this investigation (Table 4) generated a degree of graded severity within a lesion, but did not present an accurate assessment of lesion severity across different types of lesions. For example, sole ulcers with a score of 2 and sole hemorrhages with a score of 2 were equivalent in terms of modeling, but could have very different effects on pain perception. Hemorrhage in horn tissue is regarded as historical evidence of corium damage (Greenough and Vermunt, 1991; Ossent et al., 1997), and therefore little correlation exists between lesions and gait manifestations in limbs bearing hemorrhagic stains of horn tissue. Tem-

poral disassociation between the time of hemorrhage in the corium and the appearance of hemorrhage on weight-bearing surfaces likely created inaccuracies in the lesion model and underscores the need for weighting lesion severity scores for pain. This problem might be addressed in the future by the development of objective measures of pain.

Spatial variability of force distributions within and between medial and lateral claws may also have complicated lesion-based modeling. It has been demonstrated that the maximal forces are normally borne by the lateral claw rather than the medial claw of the hind limb (Van der Tol et al., 2002, 2003, 2004). Within regions of the claw, forces tended to be maximally applied in regions 3, 5, and 6 (Van der Tol et al., 2002). All things being equal, one could predict that lesions with similar severity scores appearing in regions 3, 5, and 6 of the lateral hind claw might generate greater changes in the 5 LMV than similar score lesions appearing in the medial claw or in regions 1, 2, and 4 of the lateral claws. Thus, modeling lameness by lesion scoring without regard for either pain assessment or regional distribution runs the risk of inappropriate lesion score assignment during modeling. Again, this problem might be addressed by the development of objective measures of pain.

The lack of objectivity in scoring lesions and locomotion presents a fundamental problem for modeling lameness with objective LMV data. One solution would be to develop objective measures of the dependent variables. In this regard, it might be appropriate to develop an objective, graduated measure of claw pain and then relate the 5 LMV to this measure. Another possibility would be to incorporate an exact measurement of lesion size and location, as described by Leach et al. (1998). With independent, continuous measures of lesions and locomotion, it may be possible to model lameness more accurately and relate lameness problems to milk and reproductive losses.

CONCLUSIONS

The results of this investigation indicated that the presence or absence of lameness in cattle is amenable to modeling with GRF signatures. Statistical modeling was performed without regard for the causality of lameness. Analyses of the results indicated that LS models were better descriptors of lameness than were GS models. The findings suggest that model performance could be improved with the development of a more objective lesion scoring system. In conclusion, our findings indicate that methods for automated, objective detection of bovine lameness are feasible and offer the possibility

of automated, timely, and accurate detection of lameness in commercial dairy farms.

ACKNOWLEDGMENTS

The authors acknowledge support of the University of Maryland Experiment Station and the USDA through award no. NRICGP 2000-02017. The authors further acknowledge Robert Miller for his constructive comments and thank the 3 anonymous referees for their thorough review.

APPENDIX A: ROC CURVE

An ROC curve is the plot of sensitivity on the y-axis and 1 – specificity on the x-axis, and is a method widely used for evaluating the quality and performance of diagnostic tests. Sensitivity and specificity are terms commonly used in the epidemiological literature. These terms are defined in terms of false positives (FP), false negatives (FN), true positives (TP), true negatives (TN), true diagnoses (TD), and false diagnoses (FD), as indicated below (see Table A1):

sensitivity =
$$\frac{TP}{TD}$$
; specificity = $\frac{TN}{FD}$.

An ROC is plotted for a binary classification model by applying different classification thresholds, determining the values listed above, and plotting the result as shown in Figure A1.

The ROC AUC can be used to summarize overall performance relative to an ROC. It is desirable to have a very large AUC, indicating that the sensitivity is very high, even for smaller values of specificity. Note that the AUC cannot exceed 1, because the ROC is always contained in the unit square, as shown above. As recommended by Hosmer and Lemeshow (2000), an AUC value of below 0.70 is unsatisfactory, between 0.70 to 0.80 is acceptable, and above 0.80 is excellent.

Table A1. Model and clinical cow classification with definitions of true positives (TP), false positives (FP), false negatives (FN), and true negatives $(TN)^1$

	Clinical classification		
Model classification	Lame	Sound	
Lame	TP	FP	
Sound	FN	TN	
Total	TD	FD	

 $^{1}\mathrm{TP}=$ Lameness identified by clinical observation and model classification; TN = normal gait identified by clinical observation and model classification; FP = gait classified as normal by clinical observer and lame by model classification; FN = gait classified as lame by clinical observer and normal by model classification.

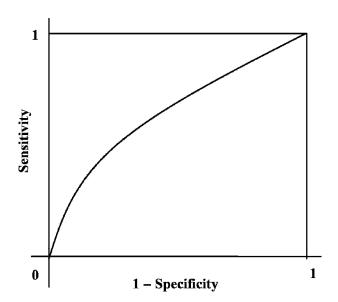


Figure A1. Schematic representation of a receiver operating characteristic (ROC) curve.

REFERENCES

- Bergstem, C. 1997. Infectious diseases of the digit. Pages 89–100 in Lameness in Cattle. P. R. Greenough. and A. D. Weaver, ed. W. B. Saunders, Philadelphia, PA.
- Clayton, H. M., H. C. Schamhardt, M. A. Willemen, J. I. Lanovaz, and R. Colborne. 2000. Kinematics and ground reaction forces in horses with superficial digital flexor tendonitis. Am. J. Vet. Res. 61:191–196.
- Collick, D. W., A. D. Weaver, and R. R. Greenough. 1997. Interdigital space and claw. Pages 101–122 in Lameness in Cattle. P. R. Greenough. and A. D. Weaver, ed. W. B. Saunders, Philadelphia. PA.
- Dyer, R. M., P. G. Rajkondawar, N. Neerchal, U. Tasch, A. M. Lefcourt, and M. A. Varner. 2004. The role of lesion and gait analyses in models of bovine lameness. Pages 76–77 in Proc. 13th Int. Conf. on Ruminant Lameness, Moribor, Slovenia. Ungula Zemljic and Co., Slovenia.
- Engel, B., G. Bruin, G. Andre, and W. Buist. 2003. Assessment of observer performance in a subjective scoring system: Visual classification of the gait of cows. J. Agric. Sci. 140:317–333.
- Greenough, P. R., and J. J. Vermunt. 1991. Evaluation of subclinical laminitis in a dairy herd and observations on associated nutritional and management factors. Vet. Rec. 128:11–17.
- Greenough, P. R., A. D. Weaver, R. J. Esselmont, and F. A. Galindo. 1997. Basic concepts of bovine lameness. Pages 3–13 in Lameness in Cattle. P. R. Greenough and A. D. Weaver, ed. W. B. Saunders, Philadelphia, PA.
- Hodson, E., H. M. Clayton, and J. L. Lanovaz. 2001. The hind limb in walking horses: 1. Kinematics and ground reaction forces. Equine Vet. J. 33:38–43.
- Hosmer, D. W., and S. Lemeshow. 2000. Applied Logistic Regression. John Wiley & Sons, New York, NY.
- Keegan, K. G., D. A. Wilson, D. J. Wilson, M. S. Bryan-Smith, E. M. Gaughan, R. S. Pleasant, J. D. Lillich, J. Kramer, R. D. Howard, C. Bacon-Miller, E. G. Davis, K. A. May, H. S. Cheramie, W. L. Valentino, and P. D. Harreveld. 1998. Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessments with kinematic gait analysis. Am. J. Vet. Res. 59:1370–1377.

- Khumsap, S., H. M. Clayton, and J. L. Lanovaz. 2001. Effect of walking velocity on ground reaction force variables in the hind limb of clinically normal horses. Am. J. Vet. Res. 62:901–906.
- Leach, K. A., D. N. Lougue, J. M. Randall, and S. A. Kempson. 1998. Claw lesions in dairy cattle: Methods for assessment of sole and white line lesions. Vet. J. 155:91–102.
- Mill, J. M., and W. R. Ward. 1994. Lameness in dairy cows and farmers' knowledge, training and awareness. Vet. Rec. 134:162–164.
- Morel, J. G., M. C. Bokossa, and N. K. Neerchal. 2003. Small sample correction for the variance of GEE estimator. Biometr. J. 45:1–15.
- Murray, R. D., D. Y. Downham, M. J. Clarkson, W. B. Faull, J. W. Hughes, F. J. Manson, J. B. Merritt, W. B. Russell, J. E. Sutherst, and W. R. Ward. 1996. Epidemiology of lameness in dairy cattle: Description and analysis of foot lesions. Vet. Rec. 138:586–591.
- Ossent, P., P. R. Greenough, and J. J. Vermunt. 1997. Laminitis in bovine lameness. Pages 284–287 in Lameness in Cattle. P. R. Greenough and A. D. Weaver ed. W. B. Saunders Co., Philadelphia, PA.
- Rajkondawar, P. G., A. M. Lefcourt, N. K. Neerchal, R. M. Dyer, M. A. Varner, B. Erez, and U. Tasch. 2002a. The development of an objective lameness scoring system for dairy herds-Pilot study. Trans. ASAE 45:1123–1125.
- Rajkondawar, P. G., U. Tasch, A. M. Lefcourt, B. Erez, R. M. Dyer, and M. A. Varner. 2002b. A system for identifying lameness in dairy cattle. Appl. Eng. Agric. 18:87–96.
- Scott, G. B. 1989. Changes in limb loading with lameness for a number of Friesian cattle. Br. Vet. J. 145:22–38.
- Scott, G. B. 1988. Studies of the gait of Friesian heifer cattle. Vet. Rec. 123:245–248.
- Shearer, J., D. Anderson, W. Ayars, E. Belknap, S. Berry, C. Guard, K. Haoblet, E. Hovingh, G. Kirksey, A. Langill, A. Mills, D. Miskimins, J. Osterstock, R. Price, D. Prigel, A. Roussel, S. van Amstel, R. Wallace, J. Wasson, N. Cook, E. Garret, D. G. Hosteller, and L. Schugel. 2004. A record-keeping system for capture of lameness and foot-care information in cattle. Bovine Pract. 38:83–92.
- Sprecher, D. J., D. E. Hosteler, and J. B. Kaneene. 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. Theriogenology 47:1179–1187.
- Tasch, U., B. Erez, A. M. Lefcourt, and M. A. Varner. 2004. A method and apparatus for detecting lameness in animals. UMBC, assignee. US Pat. No. 6,699,207.
- van Amstel, S. R., J. K. Shearer, and D. K. Haines. 2000. Maintenance claw trimming in cattle with special emphasis on the Dutch method. Bovine Pract. 33:13–20.
- Van der Tol, P. P., J. H. Metz, E. N. Noordhuizen-Stassen, W. Back, C. R. Braam, and W. A. Weijss. 2002. The pressure distribution under the bovine claw during square standing on a flat substrate. J. Dairy Sci. 85:1476–1481.
- Van der Tol, P. P., J. H. Metz, E. N. Noordhuizen-Stassen, W. Back, C. R. Braam, and W. A. Weijss. 2003. The vertical ground reaction force and the pressure distribution on the claws of dairy cows while walking on a flat substrate. J. Dairy Sci. 86:2875–2883.
- Van der Tol, P. P., S. S. van der Beek, J. H. Metz, E. N. Noordhuizen-Stassen, W. Back, C. R. Braam, and W. A. Weijs. 2004. The effect of preventive trimming on weight bearing and force balance on the claws of dairy cattle. J. Dairy Sci. 87:1732–1738.
- Wells, S. J., A. M. Trent, W. E. Marsh, and R. A. Robinson. 1993. Prevalence and severity of lameness in lactating dairy cows in a sample of Minnesota and Wisconsin herds. J. Am. Vet. Med. Assoc. 202-78–82
- Whay, H. R., and D. C. J. Main. 1999. The way cattle walk: Steps towards lameness management. Cattle Pract. 7:357–363.
- Whay, H. R., D. C. J. Main, L. E. Green, and A. J. F. Webster. 2003. Assessment of the welfare of dairy cattle using animal-based measurements: Direct observations and investigation of farm records. Vet. Rec. 153:197–202.
- Whay, H. R., A. J. F. Watterman, and J. K. Obrien. 1998. The influence of lesion type on the duration of hyperalgesia associated with hind limb lameness in dairy cattle. Vet. J. 156:23–29.