THESIS

AN UPDATE OF BEEFCAM™ TENDERNESS PREDICTION ABILITIES OF 14 DAY AGED LONGISSIMUS MUSCLES

Submitted by

Melissa Dianne Green

Animal Sciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2010

COLORADO STATE UNIVERSITY

April 1st, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY MELISSA DIANNE GREEN ENTITLED AN UPDATE OF BEEFCAM™ TENDERNESS PREDICTION ABILITIES OF 14 DAY AGED LONGISSIMUS MUSCLES BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate work

Joseph D. Tatum

Phillip L. Chapman

Advisor: Keith E. Belk

Department Head: William R. Wales Jr.

ABSTRACT OF THESIS

AN UPDATE OF BEEFCAM^{TM} IMAGE ANALYSIS OUTPUT TO PREDICT BEEF LONGISSIMUS TENDERNESS

The research presented herein was conducted to update BeefCam^{TM} tenderness predictive abilities of 14 day aged longissimus muscle samples by creating new regression equations. In this investigation, image data were collected from 670 carcasses at four beef packing plants using a video image analysis system, BeefCam^{TM}, and those data were used to predict the tenderness of aged (14 d), fresh beef Longissimus muscle (LM). Portions of the LM were removed from the striploin subprimal (NAMP #180) on both sides of each carcass. All LM samples remained fresh, were aged at 2°C for 14 d, and were cooked to a target internal temperature of 71°C. The LM samples collected from the right side of each carcass were assessed for tenderness by means of Warner-Bratzler shear force (WBSF) analysis, whereas samples collected from the left side of each carcass were evaluated by means of slice shear force (SSF) analysis. Data were sorted by SSF values and half the carcasses from each day of collection were utilized as a sequestered validation dataset (N = 334), while the remaining 336 carcasses constituted an instrument calibration dataset. BeefCam^{TM} output measures were used in regression analyses to predict beef LM tenderness following aging. A regression equation was developed using the calibration dataset that correctly classified 280 carcasses out of 336 (83.3%) as tough or tender based on LM tenderness. When the same equation was

applied to the sequestered validation dataset, it correctly classified 266 out of 334 (79.6%) carcasses as tough or tender. The developed regression equation was very successful in classifying tender carcasses, although BeefCam™ had difficulty properly identifying the tough carcasses. The root mean square error (RMSE), predicted residual sum of squares (PRESS) and R^2 statistics for the regression model were 0.1239, 2.418 and 0.3300, respectively. BeefCam™ repeatability has previously been verified and approved by USDA-AMS, but in this study, repeatability was determined to be 92.6% for the calibration dataset (N = 314) when a novice operated the instrument.

Melissa Dianne Green
Animal Sciences
Colorado State University
Fort Collins, CO 80523
Spring 2010

ACKNOWLEDGMENTS

The completion of my Master's thesis project could not have been possible without the hard work and dedication of the people around me. First, I would like to thank Dr. Belk for his unwavering support and continued patience. Your expert guidance is surely the foundation to my successes. To my family: mom, dad, Nate, Sara, and Joe, for never ceasing to express your pride in me and for standing alongside of me throughout my many endeavors. Thank you to my California friends and adopted family for their constant praise, laughter, and love. Finally, my successes throughout life could not have been possible if not through the grace of God.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER I

OBJECTIVE OF THESIS

Objective: To update BeefCam™ tenderness predictive abilities of 14 day aged longissimus muscle samples by creating new regression equations.

CHAPTER II

REVIEW OF LITERATURE

Consumers' Ability to Identify Tender Steaks and Their Willingness to Pay

In the late 1990's, United States beef consumption hit an all time low due to inconsistencies in tenderness, flavor, quality, nutrition, safety and price (NCBA, 2002). According to the National Beef Tenderness Survey, variation among fed cattle in the slaughter population was responsible for the wide range of tenderness in retail beef cuts (Morgan et al., 1991). Savell and Shackelford (1992) reported a 20% occurrence of dissatisfaction with the beef eating experience due to variation in palatability. To address tenderness, the main attribute that consumers associate with a positive beef eating experience, research efforts focused on identifying and eliminating tough carcasses (NCBA, 2002).

While determining protocols for classifying tough and tender carcasses can be beneficial to the industry, it is important to ensure that consumers will purchase verified tender products. Fortunately, foodservice companies and consumers are willing to pay premiums for cuts that will assure a positive eating experience (Savell and Shackelford, 1992). Wheeler, Shackelford, and Koohmaraie (1999) determined that the proper allocation of retail labels classifying steaks as tough or tender resulted in more satisfied consumers. Furthermore, the ability to categorize tender and tough steaks provides valuable information for proper cooking techniques in order to facilitate a positive eating experience (Wheeler et al., 1999). Research reports have indicated that consumers can easily identify which beef products are tender and they are willing to pay more to purchase the tender product, regardless of USDA quality grade (Boleman et al., 1997; Miller, Carr, Ramsey, Crockett, and Hoover, 2001). In fact, in a study that polled consumers, 72% were willing to pay extra for beef products that were guaranteed tender (NCBA, 2002). Results of a study of the influence of USDA quality grades on beef muscle tenderness indicated that product quality grade (assigned by USDA graders) had little effect on the consumer or the Warner-Bratzler shear force (WBSF) tenderness evaluations (Brooks et al., 2000), while, WBSF values were determined to be highly correlated with consumer panel tenderness ratings (Platter, Tatum, Belk, Chapman, Scanga, and Smith, 2003). On the other hand, Platter, Tatum, Belk, Koontz, Chapman, and Smith (2005) found that both marbling and WBSF had significant influence in consumers' beef strip loin purchasing decisions. Furthermore, consumers were more likely to purchase products with marbling scores greater than Modest^{50} and with WBSF values less than 3.9 kg (Platter et al., 2005).

Additional Subjective Methods for Determining Beef Carcass Quality

The use of LM marbling scores (based upon amount, size, and distribution of intramuscular fat deposits) remain useful for sorting carcasses into groups that are more consistent with regard to potential eating quality (Smith et al., 1984). However, since USDA quality grade accounts for less than 1/3 of the variation in palatability traits, various studies have been conducted to determine what beef carcass characteristics could identify and segregate tender from tough carcasses (Wulf, O'Connor, Tatum, and Smith, 1997; Smith et al., 1986). Due to the moderate relationship between marbling scores and tenderness ratings by sensory panels, Wulf et al. (1997) analyzed correlations between muscle color and beef longissimus tenderness. Results indicated that, in their study population, color measurements such as L*, a*, and b* and ultimate muscle pH were more highly correlated to muscle tenderness than the range of marbling scores, Slight^{00} to Moderately Abundant^{80} (Wulf et al., 1997).

Color measurements (a* and b* values) and pH values of the LM surface have been used in previous studies to segregate carcasses for palatability (Vote, Belk, Tatum, Scanga, and Smith, 2003, Wyle et al., 2003). Wulf and Page (2000) noted a trend in LM tenderness which indicated that the higher the b*, the better. In addition, beef carcass L* values were observed to have a relationship with the prediction of LM tenderness (Vote et al., 2009). Furthermore, muscle pH values of 5.45 to 6.65 were associated with less tender steaks (Wulf and Page, 2000). Swatland, Brooks, and Miller (1998) utilized an optical-electromechanical probe to measure ultraviolet reflectance of muscle surfaces in beef striploins. Observations concluded that the scattering of light in meat was affected by myoglobin content, animal age, and pH, therefore, making use of color as the sole tenderness predictor problematic (Swatland et al., 1998). While evaluation of muscle color and pH may not be the most effective method for segregating the most tender carcasses, it can be utilized to eliminate tough beef carcasses (Wulf and Page, 2000).

With muscle color and USDA quality grades accounting for approximately 30% of tenderness variation (Smith et al., 1986), additional beef carcass characteristics have been explored to increase palatability predictability. Li, Tan, Martz, and Heymann (1999) studied the ability of image texture features, along with color and marbling, to determine beef carcass tenderness. Prediction models showed that the inclusion of muscle texture measurements significantly increased R^2 values in principal component regression and partial least squares, from 0.30 to 0.72 and from 0.35 to 0.70, respectively (Li et al., 1999). Muscle images aided Li, Tan, and Shatadal (2000) in the development of methods that correctly identified samples as tough or tender with accuracy rates of 74.4% to 83.3%. However, results also indicated that the sole use of muscle texture features was not likely adequate for properly identifying tender carcasses (Li et al., 2000).

Review of Technologies Capable of Predicting Beef Carcass Tenderness

Many forms of technology have been developed and tested to correctly analyze muscle characteristics (color, pH, texture, etc.) in attempts to determine beef carcass tenderness. There were five main technologies that seemed, in the early 1990's, to have potential for beef quality and tenderness estimations: nuclear magnetic resonance (NMR), near-infrared reflectance (NIR), ultrasound, video image analysis (VIA), and computerized axial tomography (CAT-Scan) (Cross and Belk, 1994). Energy differences that occur between atomic nuclei in the muscle are measured by NMR, but it is also expensive and a slow process (Cross and Belk, 1994). The measurement of infrared wavelengths between fat and lean is measured by NIR (Cross and Belk, 1994). Ultrasound probes measure the reflectance of fat and muscle components and can be utilized on live animals or hide-on carcasses (Cross and Belk, 1994). To record images for VIA, a camera is placed over the surface of the muscle to be analyzed and signals are transmitted to a computer for electronic readings (Cross and Belk, 1994). Finally, density differences across muscle and fat tissues are mapped by CAT-Scans; however, the process is costly and inappropriate for on-line speeds (Cross and Belk, 1994).

Due to the variation in technology available, the National Beef Instrument Assessment Plan II (NCBA, 2002) set requirements for instruments to be accurate, fast, durable, cost effective, and operate in indirect, non-invasive methods to predict beef carcass tenderness. In addition, according to Cross and Whittaker (1992), the development of instruments to classify carcasses eases cattle producers concerns that the current grading system is subjective and does not provide producers with ample confidence. Furthermore, Cross and Whittaker (1992) stated that instruments must be able to predict percentage of lean, marbling, skeletal maturity, and be able to operate at on-line speeds. Therefore, according to the requirements set by NCBA, the most applicable methods for carcass evaluation were NIR and VIA (NCBA, 2002). Ferguson (2004) specifically endorsed the use of VIA systems by stating that VIA essentially emulates what a trained assessor does. Many systems, such as UV probes, Tendertec, NIR, and BeefCam^{TM} have been developed and tested in order to accomplish the goal of correctly classifying carcasses as tender or tough.

Ultraviolet Probe and Tendertec

Swatland, Gullet, Hore, and Buttenham (1995) conducted a study utilizing an ultraviolet (UV) fiber-optic probe to measure the presence of connective tissue in beef semitendinosus and LM and its correlation to taste panel chewiness evaluations. Dransfield, (1992) has speculated that the excessive presence of connective tissues is negatively associated with tenderness. Taste panelists classified steaks as chewy and tough when the UV probe identified samples with connective tissues located close to one another (Swatland et al., 1995). Furthermore, the presence of connective tissue and its effects on product tenderness was evaluated by Tendertec (George et al., 1997). The Tendertec probe penetrates 8 cm perpendicularly between thoracic and lumbar vertebra in order to measure tenderness by the force required for penetration (George et al., 1997). Results indicated that Tendertec variables were correlated ($P \leq 0.05$) with sensory panel connective tissue amounts, overall tenderness, and juiciness (George et al., 1997). However, there were no significant correlations between Tendertec readings and WBSF (George et al., 1997). Furthermore, Tendertec was a very weak predictor of overall tenderness, having R^2 values no greater than 0.017 (George et al., 1997). Similar results were reported by Belk et al. (2001), where Tendertec had low and non-significant correlations with WBSF values and sensory panel ratings. Therefore, Tendertec has limited ability to sort beef carcasses by predicting tenderness (Belk et al., 2001).

Near-Infrared Reflectance

The analysis of NIR's tenderness predictability was assessed on bull and cow carcasses by Hildrum, Nilsen, Mielnik, and Naes (1994). When separate analyses were utilized for only the bull samples, results showed higher correlations with tenderness and greater variance explained by tenderness, 0.86 and 76%, respectively, (Hildrum et al., 1994). Furthermore, Hildrum et al. (1994) suggested that the reason for low overall performance was likely due to the limited wavelength range, 850-1050 nm. Thus, Byrne, Downey, Troy, and Buckley (1998) analyzed NIR wavelengths of 750-1098 nm and their ability to predict quality attributes of beef carcasses. Utilizing principal component regression analysis, the overall correlation between NIR readings and WBSF tenderness evaluation was 0.61 (Byrne et al., 1998). However, the overall correlation coefficient

(R^2) of NIR to predict tenderness was only 0.53 (Byrne et al., 1998). Both Hildrum et al. (1994) and Byrne et al. (1998) conducted their testing on sample sizes of 70 carcasses or less; so, it was concluded that additional studies must be conducted on larger sample populations to fully analyze the ability of NIR to predict beef carcass tenderness.

Park, Chen, Hruschka, Shackelford, and Koohmaraie (1998) evaluated 119 carcasses to determine the capability of NIR readings over a wavelength range of 1100-2498 nm to predict LM tenderness. Wavelength absorption rates were highest in steaks determined by WBSF evaluation to be tender (Park et al., 1998). A prediction model was developed using six partial least squares factors that, when tested on the validation population, showed synonymous performance, $R^2 = 0.67$ and 0.63, respectively (Park et al., 1998). In that study, NIR was able to accurately predict 48.7, 87.7, and 97.4% of samples to be within ranges of 1.0, 2.0, and 3.0 kg, respectively, of observed WBSF values of (Park et al., 1998). However, the NIR data collection method utilized in that study did not comply with the criteria set by Cross and Whittaker (1992) to operate on-line.

Shackelford, Wheeler, and Koohmaraie (2005) evaluated NIR on-line with 145 carcasses in each of the calibration and the validation datasets. In that study, classification of tenderness was determined by being less than or greater than the median predicted SSF value (Shackelford et al., 2005). A regression equation was developed for the calibration population using ten variables accounting for 38% of the variation in SSF; furthermore, only 22% of the variance in SSF was explained when the same algorithm was applied to the validation population (Shackelford et al., 2005). The ten variables utilized were derived from the wide range of 350-2500 nm wavelengths; however, all variables included were associated with only wavelengths between 552 and 930 nm (Shackelford et al., 2005). Thus, Shackelford et al. (2005) surmised that equal prediction levels could be obtained by using a less complex machine, encompassing a smaller wavelength range.

Price, Hilton, VanOverbeke, and Morgan (2007) recently conducted a study to determine the ability of NIR to classify beef carcasses as tough or tender. While NIR was unable to predict specific tenderness values with appropriate accuracy, it was able to correctly classify tough carcasses with 92.9% accuracy (Price et al., 2007). Furthermore, NIR was able to classify 20 of the 39 tough carcasses correctly with an error rate of 3.7% (Rust et al., 2008). All results from NIR studies illustrate that classification of beef carcass tenderness, and not actual prediction of tenderness values, is possible.

Video Image Analysis, Computer Vision System, and BeefCam™

In 1978, the U.S. Government determined that the USDA needed to conduct more research in order to develop instruments that could accurately measure beef carcass characteristics (Woerner and Belk, 2008). Thus, the Food Safety and Quality Service (now AMS and FSIS) worked jointly with NASA's Office of Technology to ultimately identify ultrasound and VIA as the two technologies with the greatest potential of applying technology to beef grading (Cross and Whittaker, 1992). Cross, Gilliland, Durland, and Seideman (1983) began some of initial research on the ability of VIA systems to correctly predict beef carcass yield grades. Two decades later, VIA was thought to be the most useful technology in determining cutability, USDA yield grade, marbling score and tenderness prediction (Woerner and Belk, 2008).

Researchers began testing the abilities of VIA systems in the 1990's in packing plants to determine its capabilities. Video image analysis was determined to have considerable potential for determining beef yields by measuring total lean area, total fat area and fat thickness (Cross et al., 1983). Furthermore, the potential of VIA systems led to uses in the lamb industry (Brady et al., 2003; Cunha, Belk, Scanga, LeValley, Tatum, and Smith, 2004). A modified VIA system, called a lamb vision system (LVS), was utilized to create regression algorithms to accurately predict lamb carcass fabrication yields (Brady et al., 2003; Cunha et al., 2004). In addition, Colorado State University researchers worked with Hunter Associates Laboratory to develop a VIA system that was capable of reading lean and fat color as L*, a*, and b* values (Woerner and Belk, 2008). Belk, Cannell, Tatum, and Smith (1997) determined that the VIA system was capable of assessing Hunter color values in muscle surfaces of beef carcasses. The ability of the new system to relate lean and fat color of the LM area surface to cooked sensory characteristics of beef carcasses led to the development of a VIA prototype, BeefCam™ (Woerner and Belk, 2008).

After the ability of VIA systems to accurately measure beef carcass USDA yield grades was substantiated, researchers began studies relating VIA's ability to correctly classify beef carcasses according to USDA marbling scores (Woerner and Belk, 2008). Moore (2007) determined through a series of research studies that VIA systems were able to predict USDA marbling scores at accuracy rates greater than 89%. In order to achieve precision rates higher than that obtained for any other instrument, Moore (2007) utilized the Computer Vision System (CVS) variables that accounted for amount, size and distribution of fat, as well as fat and lean color of the muscle surface exposed at the ribeye. Results of this research led to approval by USDA to utilize VIA systems for the determination of beef carcass marbling scores (Woerner and Belk, 2008).

With demonstration of the ability of VIA systems to correctly ascertain beef carcass USDA yield and marbling scores, research began on the potential for BeefCam™ to predict beef carcass tenderness. Wheeler et al. (2002) assessed three different objective systems, BeefCam™, colorimeter, and SSF, to determine their accuracy in identifying tender beef. Results indicated that BeefCam™ was less accurate at sorting beef carcasses for tenderness than SSF (Wheeler et al., 2002). However, SSF does not provide the industry with the required non-invasive technique in order to properly classify beef carcasses by tenderness. Compared to previous VIA studies, Wyle et al. (2003) utilized BeefCam™ variables that measured lean and fat color and ribeye area; the inclusion of all lean and fat color measurements produced highly significant results in classifying beef carcasses according to tenderness (Wyle et al., 2003). However, conclusions indicated that further development of BeefCam™ was needed in order to increase the accuracy in identifying tough beef carcasses (Wyle et al., 2003).

Further development of BeefCam™ allowed Vote et al. (2003) to accurately classify 80% of carcasses as certified tender. In beef carcass populations that ranged from Traces^{50} to Slightly Abundant^{80}, BeefCam™ variable outputs for ribeye area, marbling, and lean L*, a*, and b* were all highly correlated to WBSF (Vote et al., 2003). As had been reported in other studies (Wulf et al., 1997; Wulf and Page, 2000), Vote et al. (2003) determined that the relationship found between lean color measurements and WBSF suggested that tough steaks are associated with darker colored muscles. However, no fat color measurements were correlated with WBSF tenderness values, indicating that white fat does not add to beef product tenderness in carcasses of youthful cattle (Vote et al., 2003). The most precise and accurate regression equation developed to classify beef carcasses by tenderness included BeefCam^{TM} measurements for lean a*, lean b*, and marbling, to produce an R^2 value of 0.30, a higher correlation than for any equation utilizing USDA grade factors (Vote et al., 2003).

BeefCam^{TM} was assessed in Uruguay to further evaluate its ability to predict beef carcass tenderness (Vote et al., 2009). In that study, the sample population included approximately equal number of carcasses from youthful and mature cattle. Data from previous research studies had quantified the relationship between dark muscle colors (high lean L*, a*, and b* values) and tough steaks. Therefore, Vote et al. (2009) eliminated dark-cutting carcasses from the calibration population and thereby improved correlations between lean L* and a* values with WBSF. Furthermore, BeefCam^{TM} lean color measurements were more highly correlated with WBSF values when images were obtained after a 50 min bloom time instead of obtaining immediately after beef carcass ribbing (Vote et al., 2009). Contrary to the magnitude of correlations with youthful carcasses reported by Vote et al. (2003), the study conducted utilizing more mature Uruguayan cattle showed no relationship between ribeye area and WBSF tenderness values (Vote et al., 2009). Furthermore, results illustrated that steaks from carcasses with high fat L* values were more tender (based upon WBSF values) (Vote et al., 2009).

Because numerous studies have been conducted to determine the ability of VIA systems to predict USDA yield grades, marbling scores, and carcass tenderness, it is important to ensure that the instrument's results are consistent when the measurement protocol is repeated. Shackelford, Wheeler, and Koohmaraie (1998) determined that image analysis was able to accurately predict longissimus area ($R^2 = 0.88$). Steiner et al. (2003) evaluated the repeatability of CVS, by (1) placing the camera over stationary carcass LMs and collecting three sequential images without moving the camera head; (2) placing the camera over the LM and collecting three images by removing and repositioning the camera head after each image acquisition; and (3) placing the camera head over the LM and collecting three sequential images at a chain speed of 360 carcasses per hour. All three of the CVS image acquisition procedures were compared with measurements obtained by use of a plastic grid and of acetate paper tracing (Steiner et al., 2003). Results indicated that all three CVS procedures were more accurate and consistent than both of the human-obtained ribeye area measurements ($R^2 = 0.92$, 0.90, 0.84, and 0.94, 0.93, 0.86 respectively) (Steiner et al., 2003). Thus, Steiner et al. (2003) illustrated that CVS can operate in both a stationary and on-line setting while still acquiring accurate images with high levels of repeatability.

Slice and Warner-Bratzler Shear Force Tenderness Thresholds

The ability of instrument systems to predict beef carcass tenderness relies upon their capability to correctly assign classifications based on WBSF or SSF values. The tender/tough threshold varies widely throughout the scientific literature. Thresholds for WBSF evaluation of tough LM steaks have been set at 9.0 kg (Shackelford et al., 1997), 5.0 kg (Shackelford et al., 1999b; Wheeler et al., 1999; Wulf, Emnett, Leheska, and Moeller, 2002), and 4.6 kg (Miller et al., 2001; Rust et al., 2008). Platter, et al. (2003) determined that 50% of consumers would accept beef steaks with WBSF values of 4.4 kg, while 68% of consumers would find WBSF values of 3.7 acceptable. Furthermore, food service establishments have a threshold acceptability value of 3.9 kg for WBSF (Voisinet, Grandin, O'Connor, Tatum, and Deesing, 1997), coinciding with findings from Platter, et al. (2005) that consumers would pay more for steaks with WBSF values less than 3.9 kg. In addition, beef strip loin steaks had a value decrease of $1.02/kg for every 1 kg increase in WBSF value (Platter, et al., 2005).

Tough LM steaks have been determined by SSF values of 40.0 kg (Shackelford et al., 1999b), 25.0 kg (Price et al., 2007), and 23.0 kg (Wheeler et al., 1999). Shackelford, Wheeler, and Koohmaraie (1999a) conducted a study to evaluate the correlations between SSF and WBSF with trained sensory panel tenderness ratings. Results indicated that sensory panel tenderness was more highly correlated with SSF than WBSF ($r = -0.82$ and -0.77, respectively), although there was no significant difference between the two correlations (Shackelford et al., 1999a). However, Platter et al. (2003) found that WBSF had moderately high correlations ($r = 0.63$) with consumer tenderness ratings. While it is important to identify tenderness thresholds in a trained sensory panel setting, in reality, the determination of tough versus tender product is ultimately determined by consumers.

Research has been conducted to determine the effects of cooking methods on the final tenderness determination of beef longissimus. Obuz, Kikeman, and Loughin (2003) found that cooking steaks with a convection method resulted in higher WBSF values. Coinciding with these results, Berth, Blair-Kerth, and Jones (2003) determined that clam-shell grills generated slightly more tender steaks than those cooked in an oven. When utilizing conduction type cooking methods, Wheeler, Shackelford, and Koohmaraie (1998) found that belt grills produced steaks with significantly higher tenderness values than those cooked on an electric broiler. Wharton, Apple, Yancy, Sawyer, and Lee

(2008) conducted a study to determine the effects of five cooking apparatus on beef longissimus tenderness; air-impingement oven, clam-shell griddle, convection oven, electric griddle, and gas-fired, char-grill. Results indicated that convection cookery methods produced steaks that were more tender ($P < 0.05$) than conduction methods, contradicting previous research (Wharton et al., 2008).

CHAPTER III

INTRODUCTION

Beef producers can receive premiums for cattle that qualify for certain certified and branded programs. In addition, the majority of consumers are willing to pay higher prices for certified and branded beef products that guarantee a good eating experience (Platter, Tatum, Belk, Koontz, Chapman, and Smith, 2005; Savell and Shackelford, 1992). Therefore, the ability to designate beef carcasses as tough or tender in a packing plant can have implications which may be beneficial for consumers, packers and producers alike.

Traditionally, marbling scores based upon amount, size and distribution of intramuscular fat within the cross-section of the LM have been useful for sorting carcasses into groups that are more consistent with regard to potential eating quality (Smith et al., 1984). Wulf, O'Connor, Tatum, and Smith (1997) determined that objective color measurements of the longissimus muscle (LM), such as b*, were also useful in the prediction of beef tenderness. Color measurements (L*, a*, and b* values) of the LM surface have been used in previous studies to categorize carcasses according to differences in tenderness (Vote, Belk, Tatum, Scanga, and Smith 2003; Wyle et al., 2003). Beef carcass L* values were observed to have a relationship with the prediction of LM tenderness (Vote et al., 2009). Muscle color, along with marbling in, and texture of the LM, has been utilized to predict beef carcass tenderness (Li, Tan, Martz, and Heymann, 1999; Li, Tan, and Shatadal, 2000). Some video imaging systems have had difficulty accurately sorting beef carcasses for tenderness (Wheeler et al., 2002). While the BeefCam™ video imaging system may be able to identify carcasses that are tender after appropriate aging and cooking, it does so with limited accuracy. Therefore, the objective of this study was to update BeefCam™ tenderness predictive abilities of 14 day aged longissimus muscle samples by creating new regression equations.

MATERIALS AND METHODS

Data Collection

Trained Colorado State University personnel collected video images from carcasses of 670 commercially raised, fed cattle. Images were collected in retail coolers at four commercial slaughter facilities in the Midwestern US. Carcasses were selected to ensure that at least 10% of each of the following USDA quality grades were represented: upper 2/3 Choice (Modest^{00} – Moderate^{100}), lower 1/3 Choice (Small$^{00\text{-}100}$), and Select (Slight $^{0\text{-}100}$). To obtain a balance of tender and tough carcasses, those selected represented certain breeds/species or from geographic origins known to generate tough beef. Cattle fed zilpaterol hydrochloride were excluded from carcass selection because it significantly affects tenderness while it inconsistently alters carcass tenderness indicators, such as lean and fat color (Hilton et al., 2008). Carcasses were selected in approximately 50 hd increments, quartered between the 12th and 13th ribs, and allowed to bloom for 20 to 90 minutes.

Image Acquisition

The BeefCam™ (Research Management Systems U.S.A., Fort Collins, CO) video image analysis (VIA) system was calibrated using standardized color cards at the beginning of each sampling day. Colorado State University personnel were trained in proper operating procedures and collected the required images. A video image was obtained at the 12^{th} and 13^{th} rib interface from the lead side of each carcass. The camera shroud was carefully placed on the LM surface to acquire proper images and to prevent image distortions. Digital images obtained were processed by proprietary software and repeatability was established by placing the camera head on each 12^{th} rib surface and pressing the trigger three consecutive times without removing the camera head, according to the methods of Steiner et al. (2003).

Warner-Bratzler and Slice Shear Force Determination

After all instrument readings were obtained, a single 4.25 cm section of LM from the boneless striploin (NAMP #180) was collected, on the rail, from each carcass side. The LM samples remained fresh (unfrozen) and were allowed to age 14 d postmortem before shear force evaluation. Tenderness of LM samples from right carcass sides was assessed at Texas A&M University using Warner-Bratzler shear force (WBSF) evaluation, whereas LM samples from left carcass sides were evaluated at the University of Missouri by slice shear force (SSF) evaluation. The handling of LM samples was conducted differently than in previous research for tenderness evaluation. Specifically, LM samples were removed from refrigerated conditions, retained at room temperature, and were handled excessively during additional instrument imaging following carcass evaluation with BeefCam™. After image acquisition and LM sample collection, LM samples were shipped in coolers containing frozen ice packs to the appropriate universities during the 14 d aging period. Upon arrival, LM samples were refrigerated at approximately 2°C until they were evaluated for shear force. A deli-slicer (model 3750, Globe Food Equipment Co., Dayton, OH or Hobart Deli Slicer, Hobart, Troy, OH) was used to cut the LM samples into 2.54 cm steaks for shear force evaluation. Striploin samples were cooked using an impingement conveyor oven (XLT Oven model 1832-EL, BOFI Inc., Wichita, KS) at 180°C to a target internal temperature of 71°C. Steaks to be analyzed by WBSF were placed on trays, covered with Saran wrap and chilled for 12 to 18 h at 2°C, then cored (3-6 cores per steak) using a 1.27 cm diameter coring device, parallel to the orientation of the muscle fibers. All samples evaluated by WBSF were determined to be tender if the mean compression force for a given steak was less than 5.0 kg (Wheeler et al., 1999, Shackelford et al., 1999b). Steaks analyzed for SSF were evaluated immediately after internal temperatures had peaked. A 1 cm slice was excised from the distal end of each LM sample parallel to the muscle fibers and sheared perpendicular to the muscle fibers. The threshold for determining if LM samples were tender was set at 25.0 kg per steak when analyzed by SSF.

Statistical Analysis

All statistical analyses, including descriptive statistics, simple correlations and regression analyses, were performed using the SAS statistical analysis software package (SAS Inst. Inc., Cary, NC). The sample population (N = 670) was divided into a calibration dataset (N =336) and a sequestered validation dataset (N = 334). The allocation of LM samples to the sequestered validation dataset was conducted by selecting every other sample by tenderness performance ranking, thus ensuring equal numbers of tender and tough LM samples in each of the datasets.

Various regression analyses were conducted on the LM data from the calibration dataset in order to accurately predict beef carcass tenderness. Some relationships between the dependent variable (SSF) and the independent variables (BeefCam™ output) were determined to be non-linear. Transformations by polynomial functions were conducted on independent variables while the dependent variable was log transformed to improve linearity and homogeneity of error variance. Once a final prediction regression equation was developed, the predictions for the dependent variable were back transformed to the original scale. In the attempt to increase consumer consumption and acceptance of beef, it is more acceptable to incorrectly classify tender carcasses as tough instead of allowing tough carcasses to be predicted as tender. Therefore, the tough/tender threshold was reduced to 24.0 kg in order to accurately classify more observations as tough. Furthermore, the reduction of the tough/tender threshold helps to adjust for bias due to the back transformation of predicted SSF values from the log scale.

The distribution of LM samples by SSF was approximately a bell curve. However, for the purposes of prediction, the excessive representation of samples (N = 236) with intermediate tenderness scores (15 to 25 kg) created weight on those sample characteristics. Since the objective of this study was to identify and predict the tough extremes, it was important that such extremes were well represented in the data. Therefore, a random subset of the intermediate samples were selected in order to create an equal SSF distribution of tender N = 45 (<15 kg), intermediate N = 50 (15 to 24.9 kg), and tough N = 55($\geq$25 kg). Forward, backward, and stepwise selection methods were used to identify independent variables that were significant for regression analysis. A selection criterion for independent variable inclusion into the model was $\alpha = 0.10$. Root mean square error (RMSE) and predicted residual sum of squares (PRESS) were evaluated to assess accuracy and precision.

RESULTS AND DISCUSSION

Simple statistics for the calibration dataset prior to and after equalizing the distribution is illustrated in Table 1. Simple statistics representing carcass characteristics for the calibration and the sequestered validation carcass datasets are displayed in Table 2. Tenderness LM evaluations by WBSF were excluded in the prediction equations of this study due to the limited number of samples designated as tough (N = 3) by the threshold of 5.0 kg in the calibration dataset. Therefore, all algorithms utilized SSF only in order to predict LM tenderness. Previous studies indicated that consumers can identify tough steaks at around a SSF of 23.0 kg (Wheeler et al., 1999; Shackelford et al., 1999b). All carcasses with LM steaks that had a SSF value $\geq$ 25.0 kg were designated as tough. The threshold for determining tough or tender SSF values was set by the principle investigators with assistance from the National Beef Instrument Assessment Plan II (NCBA, 2002; Price et al., 2007; Rust et al., 2008).

Numerous regression equations were developed in order to produce the greatest percentage of correct carcass classifications based on LM tenderness category. Correlations between BeefCam^{TM} variables that are associated with beef carcass tenderness and SSF are presented in Table 3. While Wulf et al. (1997) noted that all lean and fat color measurements were correlated with LM tenderness, in this study b* values for lean and fat were not significantly correlated with SSF. Such results contradict findings that specifically identify lean and fat b* measurements as being significant predictors of beef carcass tenderness (Wulf and Page, 2000; Wulf et al., 1997). Furthermore, the correlation computed in this study between ribeye area and tenderness coincides with findings from the Uruguayan study, which included more mature carcasses (Vote et al., 2009), but not with the study of Vote et al. (2003) that included a more youthful carcass population.

Initial algorithms were established utilizing combinations of the 203 characteristics captured by BeefCam™ during image acquisition. Upon the elimination of unusable or illogical variables, viable regression equations were developed using a combination of 165 of the BeefCam™ recorded characteristics, with accurate classification percentages, indicating the proportion of the population that was correctly classified, ranging from 92.6 to 39.0. To improve predictive capabilities, regression equations were developed for individual plants based on data-collection locations. Within each data-collection location, samples under the intermediate classification were randomly removed in order to recreate an equal distribution. This method increased the prediction accuracy in some plants (to 94.7%) while drastically affecting the predictability of subsequent plants (to 54.6%). The regression equations developed to predict beef carcass tenderness, along with appropriate statistics, are presented in Table 4. Additional attempts to establish higher predictive potential included logistic regression, interaction principles, principle components, forcing appropriate variables, and discriminate analysis. Furthermore, combinations of equations, by plant, by day, and overall, were assessed.

The final regression equation developed using the BeefCam^{TM} output correctly classified 280 out of 336 carcasses (83.3%) in the calibration dataset where it accurately predicted that 257 carcasses out of 281, as tender (91.5%) and 23 carcasses out of 55, as tough (41.8%). The final prediction equation was comprised of seven variables from the BeefCam^{TM} output, including descriptors for marbling, color, and lean area:

$$\hat{y} = 2.482 + a\,(0.176) + b\,(-0.242) + c\,(-0.014) + d\,(-0.015)$$
$$+ f\,(-0.099) + g\,(-0.004) + h\,(0.149) + e$$

Variables included in the final model are proprietary and therefore not specified. The root mean square error (RMSE), predicted residual sum of squares (PRESS) and R^2 statistics were 0.1239, 2.418 and 0.3300, respectively. Since 186 observations were randomly removed from the intermediate category, observations for the selected fifty were re-randomized twice and included in subsequent evaluations. Equations developed with the re-randomized observations had classification responses similar to the original equation, indicating that equalizing the distribution of the calibration dataset did not affect predictive accuracy.

The tough/tender threshold was reduced to 24.0 kg for the predicted SSF values in order to achieve the highest overall accuracy prediction rates while also correctly classifying more tough carcasses. Furthermore, a threshold (20.4 kg) was utilized to correctly capture the highest number of tough carcasses along with a high proportion of the tender carcasses. Utilizing this lower threshold, BeefCam^{TM} was able to predict 39 of the 55 carcasses as tough (70.9%) and 195 out of 281 carcasses as tender (69.4%). The ability to account for and identify more tough carcasses, thereby preventing them from reaching consumers, is more important that solely classifying a greater number of tender carcasses.

Consistent with previous studies (Wulf and Page, 2000; Wulf et al., 1997), neither b* values for lean nor for fat were utilized in the present study for the tenderness prediction algorithm. However, the inclusion of lean color measurements such as L* and a* coincided with results from Vote et al. (2003, 2009) and Wyle et al. (2003). On the other hand, Vote et al. (2003) concluded that fat color measurements were not associated with tenderness values, while this study has included such measurements in the prediction equation. The ability of VIA systems to predict USDA marbling scores (Moore, 2007) ensures that the inclusion of VIA marbling variables does not detract from the accuracy of prediction capabilities.

The same algorithm was applied to the sequestered validation dataset and predicted a total of 33 tough carcasses, out of 334. Final predictions were sent to the University of Missouri for validation and results are shown in Table 5. Comparison between predicted SSF and actual SSF values for the sequestered validation dataset resulted in accurately identifying 257 out of 281 carcasses (91.5%) as tender and 9 out of 53 carcasses (17.0%) as tough, with an overall accuracy of 79.6%. The association between the calibrated dataset SSF values and sequestered validation dataset SSF values is illustrated in Figure 1. The graph shows that BeefCam™ has a tremendous ability to properly classify tender carcasses while its performance had difficulty correctly identifying tough carcasses.

In the present study, the difference in handling methods of LM samples during additional instrument assessment could have had an adverse effect on the overall tenderness. In previous studies, standard error rates for tenderness evaluation ranged from 0.59 to 0.80 (Belk et al., 2001; George et al., 1997; Price et al., 2007) while the standard error associated with tenderness evaluation for this work was 0.28. This shows that there was less variance associated with SSF and, therefore, variance among the tenderness observations. It could thus indicate that the improper handling of the LM samples led to increased tenderness and reduced variation in tenderness, and thereby drastically influenced the presented results.

Instrument repeatability was established by analyzing the predicted SSF for the multiple images obtained from each individual carcass included in the assessment. The evaluation procedure established by Steiner et al. (2003) was utilized by placing the camera over stationary carcass LMs and collecting three sequential images without moving the camera head. Due to instrument irregularities, 22 carcasses from the calibrated dataset were removed (N = 314) and instrument repeatability was established at 92.6% (with a novice operator). Repeatability results coincided with the conclusions of Steiner et al. (2003) where R^2 levels ranged from 0.94 to 0.84.

IMPLICATIONS

The capability of the prediction equation to accurately predict LM tenderness was challenged due to the limited number of carcasses in the calibration dataset that were classified as tough by SSF ($\geq$ 25.0 kg, N = 55). BeefCam™ was able to predict tender carcasses with a high level of accuracy; however, BeefCam™ was very limited in its ability to segregate tough carcasses from the population. Further research should be conducted on a more equally distributed population and using a more strict tenderness threshold to accurately predict tenderness and provide consumers with a consistent and desirable product.

Table 1
Simple statistics for slice shear force (SSF) tenderness evaluation before and after equal distribution.

SSF classification	N	Mean kg	SD
<15 kg	45	13.64	1.03
15 to 25 kg without equal distribution	236	19.13	2.67
15 to 25 kg with equal distribution	50	18.05	2.25
≥25 kg	55	29.72	4.11

Table 2
Simple statistics for hot carcass weights, output variables from BeefCam™, and slice shear force (SSF) instrument by dataset.

Trait	N	Mean	SD	Minimum	Maximum
Calibration dataset					
Marbling score[1]	336	417.0	78.8	271.0	746.0[2]
Hot carcass wt, kg	328	375.352	33.71	257.47	356.11
Fat thickness, cm	336	10.667	5.561	0.4136	27.427
Longissimus muscle area, cm^2	336	91.76	11.57	61.40	127.43
SSF, kg	336	20.13	5.41	11.0	43.3
WBSF, kg	336	2.67	0.62	1.65	5.56
Sequestered validation dataset					
Marbling score[1]	334	408.0	70.4	238.0	654.0
Hot carcass wt, kg	325	383.25	33.55	314.48	415.84
Fat thickness, cm	334	10.652	5.273	0.765	30.711
Longissimus muscle area, cm^2	334	93.49	12.67	63.70	133.37

[1] Marbling: 200=Traces^{00}, 300=Slight^{00}, 400=Small^{00}, 500=Modest^{00}, 600=Moderate^{00}, 700=Slightly Abundant^{00}.

[2] Marbling score value is greater than the marbling range declared for selection, however, this marbling score is assigned by BeefCam™ and not the USDA grader.

Table 3
Simple correlations between BeefCam™ output variables and Slice Shear Force (SSF) tenderness.

BeefCam™ variable	Pearson correlation	P value
Marbling degree	-0.1666	0.0022
Ribeye Area cm^2	-0.0019	0.9710
Lean L*	-0.2387	<0.0001
Lean a*	-0.2759	<0.0001
Lean b*	0.0477	0.3833
Fat L*	-0.1960	0.0003
Fat a*	-0.2341	<0.0001
Fat b*	-0.0649	0.2352

Table 4
Regression equations developed for predicting beef carcass tenderness with corresponding variable numbers, R^2, percent and classification accuracy, root mean square error (RMSE), and predicted residual sum of squares (PRESS) values.

Equation	Variables included	Adjusted R^2	Percent accuracy[a]	Classification accuracy[b]	RMSE	PRESS
1	7	0.3248	83.94%	263/281, 19/55	0.1219	2.36
1:I[c]	7	0.2969	83.04%	263/281, 16/55	0.1239	2.36
1:II[c]	6	0.2719	83.00%	261/281, 18/55	0.1262	2.50
2[d]	11	0.2156	85.42%	274/281, 12/55	0.1314	2.79
3[d]	9	0.2244	84.82%	274/281, 11/55	0.1307	2.70
4[d]	11	0.2258	84.82%	274/281, 11/55	0.1305	2.73
5[d]	8	0.2078	84.23%	272/281, 11/55	0.1321	2.75
6	5	0.3214	83.04%	262/281, 17/55	0.1222	2.60
7[e]	8	0.3569	83.33%	260/281, 20/55	0.1190	2.27
8[e]	10	0.3623	82.44%	257/281, 20/55	0.1185	2.33
9[e]	5	0.3122	80.36%	255/281, 15/55	0.1230	2.36
10[e]	10	0.3624	82.70%	258/281, 20/55	0.1185	2.37
Interaction 1[f]	8	0.3717	83.93%	257/281, 25/55	0.1176	2.20
Interaction 2[f]	9	0.3759	83.29%	264/281, 15/55	0.1172	2.18
Interaction 3[f]	8	0.3766	81.55%	252/281, 22/55	0.1171	2.17
Plant A1[e]	6	0.3384	93.68%	86/86, 3/9	0.0839	0.75
Plant B1	1	0.2513	88.00%	44/44, 0/6	0.1166	0.31
Plant B2[e]	1	0.2541	88.00%	44/44, 0/6	0.1164	0.30
Plant C1	4	0.2159	85.89%	134/134, 0/22	0.0956	1.16
Plant C2[e]	4	0.1675	86.54%	134/134, 1/19	0.0985	1.60
Plant D1	6	0.4491	80.77%	90/103, 15/27	0.1085	0.83
Plant D2	3	0.3248	83.08%	97/103, 11/27	0.1201	0.96
Plant D3[e]	6	0.4459	81.54%	91/103, 15/27	0.1088	0.86
Plant D4[e]	6	0.4461	58.46%	51/103, 24/27	0.1088	0.83

[a] Accuracy percentages are the proportion of carcasses correctly classified as tough or as tender by the BeefCam™ derived algorithm at the SSF tough threshold of 25.0 kg.
[b] Numbers given as ratio of number correctly classified as tender/number of tender, number of correctly classified as tough/number tough.
[c] Equation developed by re-randomizing the intermediate portion of the dataset.
[d] Variables within the equation were highly correlated with one another.
[e] Equation contains exponential variables and/or combinations of exponential variables.
[f] Regression equation that utilized interactions between BeefCam™ variables as a single variable.

Table 5
Results of regression equation when applied to the sequestered validation cattle dataset.

	Predicted Tender	Predicted Tough
Day 1		
Actual Tender	18	2
Actual Tough	0	0
Day 2		
Actual Tender	63	3
Actual Tough	7	2
Day 3		
Actual Tender	38	6
Actual Tough	6	0
Day 4		
Actual Tender	42	5
Actual Tough	10	2
Day 5		
Actual Tender	77	7
Actual Tough	7	1
Day 6		
Actual Tender	19	1
Actual Tough	14	4
Percent of carcasses correctly classified	91.5%	17.0%

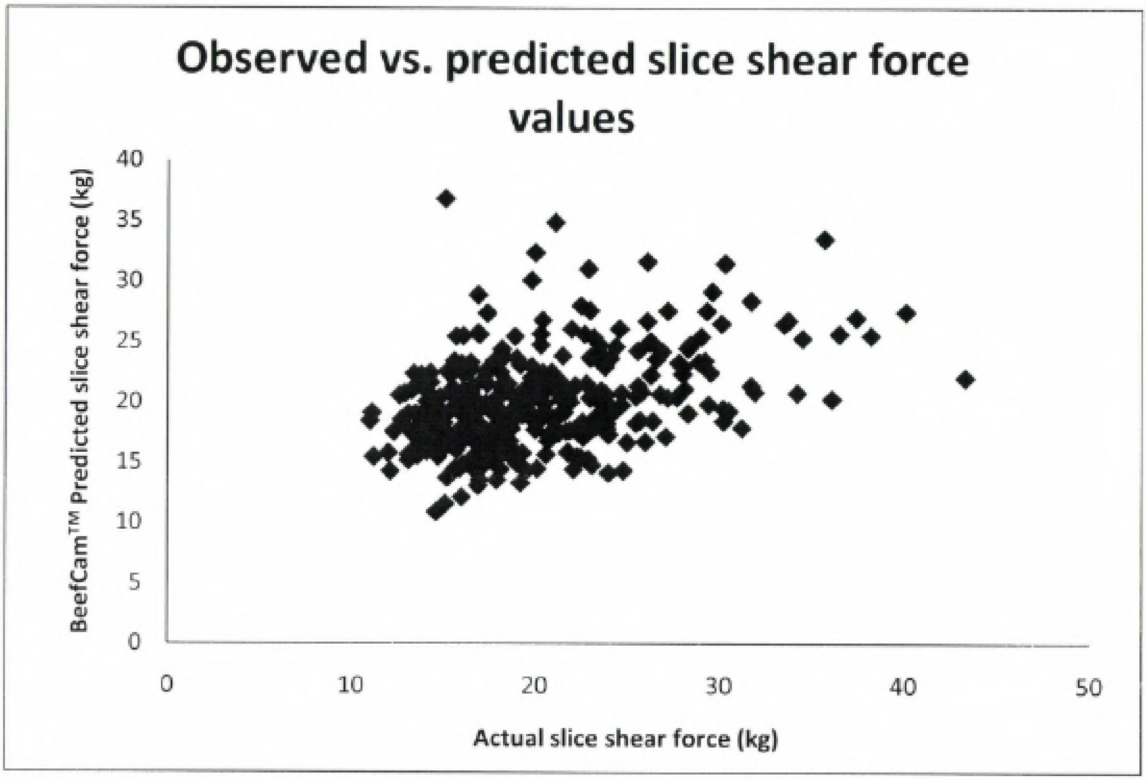

Figure 1. Relationship between the actual slice shear force values for steaks from individual carcasses and the predicted slice shear force values based on BeefCam™ output.

LITERATURE CITED

Belk, K. E., Cannell, R. C., Tatum, J. D., and Smith, G. C. (1997). Video imaging systems for composition and quality. Paper presented at the *Meat Industry Research Conference,* 28 October, 1997, Chicago, Il.

Belk, K. E., George, M. H., Tatum, J. D., Hilton, G. G., Miller, R. K., Koohmaraie, M., Reagan, J. O., and Smith, G. C. (2001). Evaluation of the Tendertec beef grading instrument to predict the tenderness of steaks from beef carcasses. *Journal of Animal Science*, *79,* 688-697.

Boleman, S. J., Boleman, S. L., Miller, R. K. , Taylor, J. F., Cross, H. R., Wheeler, T. L., Koohmaraie, M., Shackelford, S. D., Miller, M. F., West, R. L., Johnson, D. D., and Savell, J. W. (1997). Consumer evaluation of beef of known categories of tenderness. *Journal of Animal Science*. *75,* 1521-1524.

Brady, A. S., Belk, K. E., LeValley, S. B., Dalsted, N. L., Scanga, J. A., Tatum, J. D., and Smith, G. C. (2003). An evaluation of the lamb vision system as a predictor of lamb carcass red meat yield percentage. *Journal of Animal Science, 81*, 1488-1498.

Brooks, J. C., Belew, J. B., Griffin, D. B., Gwartney, B. L., Hale, D. S., Henning, W. R., Johnson, D. D., Morang, J. B., Parrish Jr., F. C., Reagan, J. O., and Savell, J. W. (2000). National Beef Tenderness Survey – 1998. *Journal of Animal Science*, *78,* 1852-1860.

Byrne, C. E., Downey, G., Troy, D. J., and Buckley, D. J. (1998). Non-destructive prediction of selected quality attributes of beef by near-infrared reflectance spectroscopy between 750 and 1098 nm. *Meat Science, 49*, 399-409.

Cross, H. R., and Whittaker, A. D. (1992). The role of instrument grading in beef value-based marketing system. *Journal of Animal Science, 70,* 984-989.

Cross, H. R., and Belk, K. E. (1994). Objective measurements of carcass and meat quality. *Meat Science*, *36*, 191-202.

Cross, H. R., Gilliland, D. A., Durland, P. R., and Seideman, S. (1983). Beef carcass evaluation by use of a video image analysis system. *Journal of Animal Science, 57*, 908-917.

Cunha, B. C. N., Belk, K. E., Scanga, J. A., LeValley, S. B., Tatum, J. D., and Smith, G. C. (2004). Development and validation of equations utilizing lamb vision system output to predict lamb carcass fabrications yields. *Journal of Animal Science, 82,* 2069-2076.

Dransfield, E. (1992). A unified model for meat tenderness. In Swatland, H. J., Gullett, E., Hore, T., and Buttenham, S. (1995). UV fiber-optic probe measurements of connective tissue in beef correlated with taste panel scores for chewiness. *Food Research International, 28,* 23-20.

Ferguson, D. M. (2004). Objective on-line assessment of marbling: A brief review. *Australian Journal of Experimental Agriculture, 44,* 681-685.

George, M. H., Tatum, J. D., Dolezal, H. G., Morgan, J. B., Wise, J. W., Calkins, C. R., Gordon, T., Reagan, J. O., and Smith, G. C. (1997). Comparison of USDA quality grade with Tendertec for the assessment of beef palatability. *Journal of Animal Science*, *75,* 1538-1546.

Hildrum, K. I., Nilsen, B. N., Mielnik, M., and Naes, T. (1994). Prediction of sensory characteristics of beef by near-infrared spectroscopy. *Meat Science, 38,* 67-80.

Hilton, G. G., Montgomery, J. L., Krehbiel, C. R., Yates, D. A., Hutcheson, J. P., Nichols, W. T., Streeter, M. N., Blanton Jr., J. R., and Miller, M. F. (2008). Effects of feeding zilpaterol hydrochloride with and without monensin and tylosin on carcass cutability and meat palatability of beef steers. *Journal of Animal Science, 87*, 1394-1406.

Kerth, C. R., Blair-Kerth, L. D., and Jones, W. R. (2003). Warner-Bratzler shear force repeatability in beef longissimus steaks cooked with a convection oven, broiler, or clam-shell grill. *Journal of Food Science, 68*, 668-670.

Li, J, Tan, J., Martz, F. A., and Heymann, H. (1999). Image texture features as indicators of beef tenderness. *Meat Science*, *53,* 17-22.

Li, J., Tan, J., and Shatadal, P. (2000). Classification of tough and tender beef by image texture analysis. *Meat Science, 57,* 341-346.

Miller, M. F., Carr, M. A., Ramsey, C. B., Crockett, K. L., and Hoover, L.C. (2001). Consumer thresholds for establishing the value of beef tenderness. *Journal of Animal Science, 79,* 3062-3068.

Moore, C. B. (2007). Establishing an appropriate gold standard for measuring performance of system derived beef carcass marbling scores. M.S. Thesis. *Center for Red Meat Safety and Quality, Department of Animal Sciences, Colorado State University, Fort Collins, CO.*

Morgan, J. B., Savell, J. W., Hale, D. S., Miller, R. K., Griffin, D. B., Cross, H. R., and Shackelford, S. D. (1991). National beef tenderness survey. *Journal of Animal Science, 69,* 3274-3283.

NCBA. (2002). The National Beef Instrument Assessment Plan – II. Focus on Tenderness. *Natl. Cattlemen's Beef Assoc., Centennial, CO.*

Obuz, E., Dikeman, M. E., and Loughin, T. M. (2003). Effects of cooking method, reheating, holding time, and holding temperature on beef *longissimus lumborum* and *biceps femoris* tenderness. *Meat Science, 65*, 841-851.

Park, B., Chen, Y.R., Hruschka, W. R., Shackelford, S. D., and Koohmaraie, M. (1998). Near-infrared reflectance analysis for predicting beef longissimus tenderness. *Journal of Animal Science, 76,* 2115-2120.

Platter, W. J., Tatum, J. D., Belk, K. E., Chapman, P. L., Scanga, J. A., and Smith, G. C. (2003). Relationships of consumer sensory ratings, marbling score, and shear force value to consumer acceptance of beef strip loin steaks. *Journal of Animal Science, 81,* 2741-2750.

Platter, W. J., Tatum, J. D., Belk, K. E., Koontz, S. R., and Smith, G. C. (2005). Effects of marbling and shear force on sonsumers' willingness to pay for beef strip loin steaks. *Journal of Animal Science, 83,* 890-899.

Price, D. M, Hilton, G. G., VanOverbeke, D. L., and Morgan, J. B. (2007). Using the near-infrared (VIS-NIR) system to sort various beef "middle and end muscle cuts" into tenderness categories. *Journal of Animal Science, 86,* 413-418.

Rust, S. R., Price, D. M., Subbiah, J., Kranzler, G., Hilton, G. G., VanOverbeke, D. L., and Morgan, J. B. (2008). Predicting beef tenderness using near-infrared spectroscopy. *Journal of Animal Science, 86,* 211-219.

Savell, J. W., and Shackelford, S. D. (1992). Significance of tenderness to the meat industry. *Reciprocal Meat Conference Proceedings, 45*, 43-46.

Shackelford, S. D., Wheeler, T. L., and Koohmaraie, M. (1997). Tenderness classification of beef: I. Evaluation of beef longissimus shear force at 1 or 2 days postmortem as a predictor of aged beef tenderness. *Journal of Animal Science, 75,* 2417-2422.

Shackelford, S. D., Wheeler, T. L., and Koohmaraie, M. (1998). Coupling of image analysis and tenderness classification to simultaneously evaluate carcass cutability, longissimus area, subprimal cut weights, and tenderness of beef. *Journal of Animal Science, 76,* 2631-2640.

Shackelford, S. D., Wheeler, T. L., and Koohmaraie, M. (1999a). Evaluation of slice shear force as an objective method of assessing beef longissimus tenderness. *Journal of Animal Science, 77,* 2693-2699.

Shackelford, S. D., Wheeler, T. L., and Koohmaraie, M. (1999b). Tenderness classification of beef: II. Design and analysis of a system to measure beef longissimus shear force under commercial processing conditions. *Journal of Animal Science, 77,* 1474-1481.

Shackelford, S. D., Wheeler, T. L., and Koohmaraie, M. (2005). On-line classification of US Select beef carcasses for longissimus tenderness using visible and near-infrared reflectance spectroscopy. *Meat Science*, *69,* 409-415.

Smith, G. C., Savell, J. W., Cross, H. R., Carpenter, Z. L., Murphey, C. E., Davis, G. W., Abraham, H. C., Parrish Jr., F. C., and Berry, B.W. (1986). Relationship of USDA quality grades to palatability of cooked beef. *Journal of Food Quality, 10,* 269-286.

Smith, G. C., Carpenter, Z. L., Cross, H. R., Murphey, C. E., Abraham, H. C., Savell, J. W., Davis, G. W., Berry, B. W., and Parrish Jr., F. C. (1984). Relationship of USDA marbling groups to palatability of cooked beef. *Journal of Food Quality, 7,* 289-308.

Steiner, R., Vote, D. J., Belk, K. E., Scanga, J. A., Wise, J. W., Tatum, J. D., and Smith, G. C. (2003). Accuracy and repeatability of beef carcass longissimus muscle area measurements. *Journal of Animal Science, 81,* 1980-1988.

Swatland, H. J., Gullett, E., Hore, T., and Buttenham, S. (1995). UV fiber-optic probe measurements of connective tissue in beef correlated with taste panel scores for chewiness. *Food Research International, 28,* 23-20.

Swatland, H. J., Brooks, J. C., and Miller M. F. (1998). Possibilities for predicting taste and tenderness of broiled beef stakes using an optical-electromechanical probe. *Meat Science, 50*, 1-12.

Voisinet, B. D., Grandin, T., O'Connor, S. F., Tatum, J. D., and Deesing, M. J. (1997). *Bos Indicus*-cross feedlot cattle with excitable temperaments have tougher meat and a higher incidence of borderline dark cutters. *Meat Science, 46*, 367-377.

Vote, D. J., Belk, K. E., Tatum, J. D., Scanga, J. A., and Smith, G. C. (2003). Online prediction of beef tenderness using a computer vision system equipped with BeefCam module. *Journal of Animal Science, 81,* 457-465.

Vote, D. J., Bowling, M. B., Cunha, B. C. N., Belk, K. E., Tatum, J. D., Montossi, F., and Smith, G. C. (2009). Video image analysis as a potential grading system for Uruguayan beef carcasses. *Journal of Animal Science*, *87,* 2376-2390.

Wharton, M. D., Apple, J. K., Yancey, J. W. S., Sawyer, J. T., and Lee, M. S. (2008). Internal color and tenderness of the *longissimus thoracis* are affected by cooking methods and degree of doneness. *Arkansas Animal Science Department Report 2008*. 105-108. http://arkansasagnews.uark.edu/563-30.pdf Accessed February 17, 2010.

Wheeler, T. L., Vote, D. J., Leheska, J. M., Shackelford, S. D., Belk, K. E., Wulf, D. M., Gwartney, B. L., and Koohmaraie, M. (2002). The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science*, *80*,3315-3327.

Wheeler, T. L., Shackelford, S. D., and Koohmaraie, M. (1998). Cooking and palatability traits of beef longissimus steaks cooked with a belt grill or an open hearth electric broiler. *Journal of Animal Science, 76,* 2805-2810.

Wheeler, T. L., Shackelford, S. D., and Koohmaraie, M. (1999). Tenderness classification of beef: III. Effect of interaction between end point temperature and tenderness on Warner-Bratzler shear force of beef longissimus. *Journal of Animal Science , 77,* 400-407.

Woerner, D. R., and Belk, K. E. (2008). The history of instrument assessment of beef – A focus on the last ten years. Final Report submitted to the National Cattlemen's Beef Association, Centennial, CO by the *Center for Red Meat Safety and Quality, Department of Animal Sciences, Colorado State University, Fort Collins, CO.*

Wulf, D. M., and Page, J. K. (2000). Using measurements of muscle color, pH, and electrical impedance to augment the current USDA beef quality grading standards to improve the accuracy and precision of sorting carcasses into palatability groups. *Journal of Animal Science, 78,* 2595-2607.

Wulf, D. M., Emnett, R. S., Leheska, J. M., and Moeller, S. J. (2002). Relationships among glycolytic potential, dark cutting (dark, firm, and dry) beef, and cooked beef palatability. *Journal of Animal Science, 80*, 1895-1903.

Wulf, D. M., O'Connor, S. F., Tatum, J. D., and Smith, G. C. (1997). Using objective measures of muscle color to predict beef longissimus tenderness. *Journal of Animal Science, 75*, 684-692.

Wyle, A. M., Vote, D. J., Roeber, D. L., Cannell, R. C., Belk, K. E., Scanga, J. A., Goldberg, M., Tatum, J. D., and Smith, G. C. (2003). Effectiveness of the SmartMV prototype BeefCam System to sort beef carcasses into expected palatability groups. *Journal of Animal Science*, *81,* 441-448.