Introduction

Since 1975, USDA has utilized marbling (combined with physiological maturity) as the primary determinant in beef quality grades. The decision to include marbling as a primary predictor of eating quality was based on the premise that marbling is positively correlated to tenderness, which was later substantiated by scientists who found this to, in fact, be true (McBee and Wiles, 1967, Jennings et al., 1978, Tatum et al., 1980, Dolezal et al., 1982). Smith et al. (1987) illustrated how marbling could be effectively used to sort carcasses based on palatability, when the population exhibited the entire range of possible quality grades. However, in the current market, over 80% (87.5% as of March 6, 2000), of the beef carcasses produced in the United States grade USDA Select and Choice (USDA, 2000) with a large portion of these carcasses having Slight and Small degrees of marbling. Within this narrow range of marbling scores, marbling does not do an adequate job of sorting beef carcasses into differing palatability groups (Smith et al., 1995). Despite efforts to continually update Official USDA Quality Grade standards to more effectively sort beef carcasses into differing palatability groups, new technologies with the ability to more precisely sort carcass on the basis of beef palatability are necessary.

Instrument Technologies

As a result of the inability to explain variation in beef carcass palatability using marbling scores, the beef industry began investigating the use of instruments to help improve characterization, sorting, and pricing of cattle and beef carcasses nearly three decades ago (Cross and Whittaker, 1992). In 1994, the National Livestock and Meat Board convened a National Beef Instrument Assessment Planning (NBIAP) symposium to assess state-of-the-art capabilities in carcass evaluation tools and to make recommendations as to which technologies new research should focus. The NBIAP symposium resulted in the following conclusions: (1) reliable, accurate tools for instrument assessment hold the promise of more accurately measuring the factors that contribute to consumer satisfaction with beef, while reducing production costs and waste, (2) testing experimental technology under real-world conditions is critical to achieving commercial success, (3) VIA technology was ready for commercial testing and was the most promising technology for short-term implementation, and (4) ToBEC, Tendertec, Swatland's Probe and Real-Time Ultrasound for seedstock evaluation were ranked second through fifth in applied research priority, respectively (NLSMB, 1994).

Opinions as to how instruments should be used in carcass assessment are diverse. They have ranged from those who would eliminate Federal USDA grading altogether and replace it with services provided by a private grading company that may, or may not, incorporate instruments into the system (Helming, 1996), to those who believe that technology should be used to augment the application of USDA grade standards (Belk et al., 1996). Supporters of augmenting USDA grades are excited about the possibility of increasing both the accuracy and repeatability of the current grade standards using instrument technology, however they also realize that (1) privatization of the grading system would not prove to be a credible, third-party conformity assessment system, (2) the current system is voluntary and therefore, if grading were not desired by customers of beef packing companies, it could have already been eliminated, (3) eliminating USDA grades would require a change in the Agricultural Marketing Act of 1946, (4) current USDA grades are extremely important merchandising tools in the international market, and (5) elimination of USDA grades would have an adverse effect on other marketing services currently provided by USDA, such as certification and Process Verification programs (Belk et al., 1999).

Several systems have been developed over the years for use in measuring yield and quality factors on beef carcasses. Those systems that could potentially sort carcasses based on eating quality subsequent to aging and cooking can be di-
vided into two classes, invasive systems and non-invasive systems. An invasive system requires the removal of product from the carcass or penetration of muscles of the carcass, resulting in lost yield to a packer or potential food safety concerns, while a non-invasive system relies on the visual characteristics of the carcass and/or lean tissue, and results in no lost product and minimal food safety concerns. Within the scope of this paper, only instrumentation or augmentation systems designed to predict end product palatability will be discussed.

**MARC Tenderness Classification**

Researchers at the Meat Animal Research Center (MARC) in Clay Center, NE, have developed a system that utilizes slice shear force measurements (a variation of Warner-Bratzler shear force designed to be completed in less time) on samples collected from carcasses after chilling (generally 36-48 hours post-mortem) to sort beef carcasses into groups described as “tender,” “intermediate” or “tough.” A recent research project that utilized this system determined that 47% of the carcasses were sorted into the “tender” group, 48% were sorted into the “intermediate” group and 5% were sorted into the “tough” group (Shackelford et al., 1999). When compared against trained taste panel responses for steaks that were aged 14 days, 100% of the “tender” group were actually tender, 91% of the “intermediate” group were actually tender, and 28% of the “tough” group were actually tender (Shackelford et al., 1999). Samples from carcasses in the study were obtained in a commercial packing plant and transported to the MARC facility for slice shear force testing. Therefore, these results represent the best-case scenario for the system and not real-world conditions as the carcasses were not sorted with a commercial, online system. Furthermore, the MARC tenderness classification system has been met with opposition from packers due to the invasive nature of the technology. The removal of one steak from a carcass becomes costly when considering that most large packing facilities are capable of processing in excess of 5000 carcasses per day.

**Armour Tenderometer**

Many researchers have attempted to develop probe systems that are moderately-invasive, believing that the industry would much more readily accept a system of this type. The first system of this type was the Armour Tenderometer (AT). This system utilized a group of probes that were inserted into the *longissimus* following carcass chilling and that measured the force required to penetrate the muscle and used this information to predict cooked meat tenderness. Carpenter et al. (1972) concluded that the AT effectively categorized USDA Choice beef carcasses into tenderness desirability groups, however Huffman (1974) reported that AT reading were related (R² = .22) to WBS values, but there was no relationship (P > .05) between AT readings and trained taste panel scores for tenderness taken from 192 carcasses ranging in USDA Quality Grade from Prime to Standard. Parrish et al. (1973) reported low correlations between AT values and both WBS and organoleptic tenderness ratings from *longissimus* steaks aged for 7 d (R² = .07 and .12, respectively). More recently, Harris et al. (1992) evaluated usefulness of Armour Tenderometer readings from 384 “A” maturity beef carcasses described in the study of Smith et al. (1984) and reported simple correlation coefficients of .10 (P < .05) and -.13 (P < .01) between AT readings and sensory panel tenderness ratings and WBS force values, respectively. Because of the low correlation to ultimate meat tenderness and palatability, and the apparent ineffectiveness of this technology, it has since been abandoned as a tenderness-predicting tool.

**Swatland’s Probe**

The second moderately-invasive system developed was the connective tissue probe (CT probe) which uses an optical fiber probe and measures the reflectance of initially polarized light to predict the palatability of beef, predominantly by characterizing the connective tissue properties of the muscle (Swatland, 1991). Despite initial laboratory success, further improvements to the prototype system were needed to obtain more reliable results in commercial use. Later reports found that measurements (reflectance at 460 nm, fluorescence peak 3 and mean length disorder) taken using an optical-electrochemical probe accounted for 34% of the variation in perceived tenderness of 21 d aged *longissimus* steaks, but further work and improvements are needed to obtain more reliable predictions (Swatland et al., 1998).

**Tendertec**

The third moderately-invasive system evaluated was the Tendertec Mark III Beef Grading Instrument, an Australian probe developed to measure the amount of connective tissue and other factors that contribute to the toughening of meat. In research conducted by George et al. (1997a), no statistical significance was found between the Tendertec outputs and Warner-Bratzler shear force values. Tendertec output variables were significantly correlated with sensory panel ratings for connective tissue amount and overall tenderness, but the coefficients were very low (George et al., 1997a). These results were similar to previous findings by Belk et al. (1996) for the Tendertec instrument’s ability to predict beef carcass palatability. Working with Lester Jeremiah (Agriculture Canada, Lacombe), George et al. (1997b) compared efficacy of Tendertec and Swatland’s CT probe (on carcasses) and the Meat Industry Research in New Zealand (MIRINZ) Tenderometer (on raw muscle tissue) as predictors of WBS force values for beef loin steaks. Simple correlation coefficients for Tendertec, C.T. probe and MIRINZ Tenderometer with WBS force values for samples from more than 400 carcasses and/or muscles were not statistically different from zero (.19, .17, .00 to .36, respectively).

Due to the limited success of probes and industry opposition to invasive systems, researchers also have investigated the use of lean color as a palatability predictor. Hodgson et al. (1992) and Hilton et al. (1997) found that lean and fat color scores for mature cow carcasses were related to subsequent cooked beef palatability. The lean and fat color scores used in the Hodgson et al. (1992) and Hilton et al. (1997) studies were determined by personnel trained to evaluate such...
carcass traits, and did not represent the use of instruments to sort beef carcasses into specific palatability classes.

Belk et al. (1999) reported that the lean and fat color of beef carcasses can be used to measure several traits that are related to beef carcass palatability, including: (1) presence/absence of marbling, (2) physiological maturity of the lean, (3) muscle pH, (4) production and feeding management history, and (5) ultrastructural status of sarcomeres and connective tissue within the muscle. In addition, lean color has been shown to be related to calpastatin activity of postmortem muscle (Tatum et al., 1997). As a result, researchers began searching for objective ways to measure lean and fat color of beef carcasses.

Wulf et al. (1997) utilized a Minolta Colorimeter (a portable colorimeter) to measure the Commission Internationale de l'Eclairage (International Commission on Illumination; CIE) values for L* (lightness; dark = 0, white = 100), a* (red = + values, green = - values), and b* (yellow = + values, blue = - values). Wulf et al. (1997) found that L*, a*, and b* values measured on the exposed longissimus muscle of beef carcasses were related to beef carcass palatability. Similarly, Tatum et al. (1997), found that L*, a*, and b* values, measured using the HunterLab MiniScan portable spectrophotometer, could be used to decrease the variation that occurs in beef carcass palatability. Both of these studies used color measurement instruments with small aperture sizes to measure the lean and fat color of longissimus muscle cross sections. Therefore, information concerning lean color was only generated for a small portion of the exposed longissimus muscle, and was not representative of the variation in muscle color that occurs across the cross-sectional face of the longissimus muscle surface at the 12th rib.

Recent work resulted in the use of video image analysis (VIA) to make color measurements on the entire exposed surface of the longissimus muscle at the 12th rib. Early work using VIA technology to measure beef muscle color was marginally successful (Li et al., 1997). The early VIA systems used the computer compatible RGB color measurements computed from the video images to determine the lean color of beef longissimus muscle. While RGB colors were correlated to tenderness, attempts to sort beef carcasses into palatability classes using these color measurements were unsuccessful (Li et al., 1997). Early VIA research did prove that computer software could be written that would accurately segment a video image of a ribeye—via image processing techniques—into fat, lean and connective tissue components and conduct analysis of color and other attributes generated by color measurements on each of these components, independently.

**BeefCam™**

In 1996, Colorado State University initiated pilot work with Hunter Associates Laboratory (manufacturers of the HunterLab MiniScan portable spectrophotometer) to develop a VIA system that could measure beef carcass lean and fat color using the L*, a*, and b* color scale. The bench-top VIA system developed by Colorado State University and Hunter Associates Laboratory used off-the-shelf technology that was not specifically designed to analyze beef muscle color. The off-the-shelf technology was used to obtain images of beef longissimus steaks for the purpose of objective color analysis. When the VIA-derived color measurements were used, in conjunction with expert quality grade factors, the probability of encountering a tough (WBS > 4.5 kg) steak after 14 to 21 d of aging was reduced from .18 to .25 and .15 to .02 for USDA Choice and USDA Select steaks, respectively (Belk et al., 1997). Furthermore, Belk et al. (1997) reported that the pilot study data confirmed that (1) color is related to subsequent cooked palatability of beef carcasses, independent of differences in marbling or carcass maturity, and (2) VIA technology is capable of ascertaining color attributes of beef ribeyes, using the color information to augment USDA quality grades, and thereby improve the accuracy of quality grades in sorting carcasses based on expected eating palatability across narrow ranges of marbling scores.

Based on the results of the pilot study, Colorado State University and Hunter Associates Laboratory began development of a prototype portable video imaging system (BeefCam™) which contained hardware and software that were specifically designed for the analysis of beef carcass lean and fat color in a packing plant environment. Researchers at Colorado State University tested the BeefCam™ system for its ability to sort beef carcasses based on the expected eating quality of subsequent cooked product. A study conducted by Wyle et al. (1998) used the BeefCam™ system, either alone (Model I), or in conjunction with USDA Quality Grade (Model II) to certify carcasses as being tender (WBS < 4.5 kg) or tough (WBS > 4.5 kg). In this study, use of Model I resulted in 51.9% of the carcasses evaluated (n = 500) being certified as being tender and 92.2% of those that were certified were actually tender. Using Model II, 53.4% of the carcasses evaluated (n = 500) were certified as being tender and 94.4% of those certified were actually tender (Table 1).

To validate the effectiveness of the BeefCam™ Model 1 algorithm, researchers at Colorado State University selected 292 beef carcasses from a commercial Colorado packing plant (Cannell et al., 1999; unpublished data), a different plant from those sampled in Wyle et al. (1999). The sample population evaluated contained carcasses that were assigned USDA quality grades ranging from U.S. Standard to U.S. Prime, with the greatest proportion of carcasses falling into the U.S. Select and U.S. Choice grades (mimicking the U.S. beef population). Sample carcasses were assigned USDA yield grades ranging between 1 and 5, and all carcasses were selected to reflect the normal variability in composition, dressing defects and quality attributes encountered by the facility on a daily basis. Of the 292 carcasses evaluated, 47.3% (138 carcasses) were certified by BeefCam™ Model 1 as being palatable (WBS < 4.5 kg). The BeefCam™ Model I certified palatable carcasses produced steaks with a mean WBS value and overall tenderness panel rating of 3.6 kg and 5.0, respectively. Of those carcasses certified as palatable using BeefCam™ Model I, only 1.4% (2 carcasses) generated steaks that had WBS greater than 4.5 kg, 6.5% (9 carcasses) generated steaks that had WBS greater than 4.0 kg, and only 20.3% (28 carcasses) generated steaks that were assigned trained taste panel ratings of less...
than 5.0 for overall tenderness (on an 8-point scale where: 1 = extremely tough). Of those carcasses rejected for certification using BeefCam™ Model I, 21 carcasses (13.6%) generated steaks that had WBS values greater than 4.5 kg, 48 carcasses (31.2%) generated steaks with WBS values greater than 4.0 kg, and 70 carcasses (45.5%) generated steaks that were assigned trained taste panel ratings of less than 5.0 for overall tenderness. From this validation, when tested on a separate and unique beef carcass sample population, relative to WBS force and trained taste panel ratings, BeefCam™ performed similarly (if not better) in accuracy to its performance on the initial population from which the sorting algorithms were developed.

USDA Quality Grade Augmentation

A similar system was developed by Wulf and Page (2000) who augmented the current USDA Quality Grading Standards with colorimeter readings (L*, a*, and b*), pH and hump height (maximum dorsal protrusion of the rhomboideus muscle; measured as the distance from the dorsal edge of the ligamentum nuchae to the dorsal edge of the rhomboideus, not counting subcutaneous fat). When evaluated under “carefully-controlled” bloom times, it was reported that this augmentation system could predict longissimus WBS force (R² = .36) measures and a carcass palatability (R² = .46) index (additive measure of longissimus, gluteus medius and semimembranosus WBS shear force values and sensory panel attributes) following 7 days of postmortem aging. From this research, two proposed augmentation schemes were outlined for USDA Choice and Select beef carcasses (Table 2).

The effectiveness of these two proposed systems to segregate cattle into palatability groups are demonstrated in Figure 1. Furthermore, augmentation of the current USDA grade standards with these two proposed systems could reduce the chance of an unpleasant eating experience from 14% to 1%.

### TABLE 1. Percentage of carcasses certified by use of BeefCam™ alone (Model I) or in conjunction with USDA Quality Grade (Model II) and percentages of unacceptable carcasses based on mean Warner-Bratzler shear force values of cooked longissimus steaks (n = 769).

<table>
<thead>
<tr>
<th></th>
<th>Percent of Population Certified</th>
<th>Percent of Population Tough</th>
<th>Percent of Cattle That Were Tough</th>
<th>Percent of Non-Certified Cattle That Were Tough</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Carcasses</td>
<td>51.9</td>
<td>13.8</td>
<td>7.8</td>
<td>20.3</td>
</tr>
<tr>
<td>Top Choice</td>
<td>57.3</td>
<td>7.9</td>
<td>4.3</td>
<td>12.6</td>
</tr>
<tr>
<td>Low Choice</td>
<td>58.5</td>
<td>10.3</td>
<td>6.3</td>
<td>16.0</td>
</tr>
<tr>
<td>Select</td>
<td>37.5</td>
<td>24.7</td>
<td>16.5</td>
<td>29.6</td>
</tr>
<tr>
<td><strong>Model II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Carcasses</td>
<td>53.4</td>
<td>13.8</td>
<td>5.6</td>
<td>23.2</td>
</tr>
<tr>
<td>Top Choice</td>
<td>78.0</td>
<td>7.9</td>
<td>4.8</td>
<td>18.9</td>
</tr>
<tr>
<td>Low Choice</td>
<td>59.1</td>
<td>10.3</td>
<td>6.7</td>
<td>15.4</td>
</tr>
<tr>
<td>Select</td>
<td>19.8</td>
<td>24.7</td>
<td>4.4</td>
<td>23.8</td>
</tr>
</tbody>
</table>

### TABLE 2. Proposed USDA Quality Grade augmentation models (Wulf and Paige, 2000).

<table>
<thead>
<tr>
<th>System #1</th>
<th>System #2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>USDA Choice</strong></td>
<td><strong>USDA Choice</strong></td>
</tr>
<tr>
<td>Must be “A” or “B” overall maturity.</td>
<td>Must be “A” or “B” overall maturity.</td>
</tr>
<tr>
<td>Must have a minimum marbling score of Small100.</td>
<td>If L* is from 36.0 to 40.0 then must have a minimum marbling score of Small100.</td>
</tr>
<tr>
<td>Must have a minimum L* value of 36.0.</td>
<td>If L* is &gt; 40.0, then must have a minimum marbling score of Slight100.</td>
</tr>
<tr>
<td>Must have a hump height &lt; 8.9 cm.</td>
<td>Must have a hump height &lt; 8.9 cm.</td>
</tr>
<tr>
<td><strong>USDA Select</strong></td>
<td><strong>USDA Select</strong></td>
</tr>
<tr>
<td>Must be “A” or “B” overall maturity.</td>
<td>Must be “A” or “B” overall maturity.</td>
</tr>
<tr>
<td>Must have a minimum marbling score of Slight100.</td>
<td>Must have a minimum marbling score of Slight100.</td>
</tr>
<tr>
<td>Must have a minimum L* value of 38.0.</td>
<td>Must have a minimum L* value of 36.0.</td>
</tr>
<tr>
<td>Must have a hump height &lt; 8.9 cm.</td>
<td>Must have a hump height &lt; 8.9 cm.</td>
</tr>
</tbody>
</table>
and 36% to 7% for USDA Choice and Select carcasses, respectively. It is evident that augmentation systems can improve the accuracy and precision of sorting beef carcasses into palatability groups.

**FIGURE 1.**

Response-surface analysis ($R^2 = .40$) showing the combined effect of marbling score (300 = Slight$^0$, 400 = Small$^0$, etc.) and $L^*$ value on overall palatability index, overlaid with three different quality grading systems. Palatability index: Very High = 45.00 and higher, High = 40.00 to 44.99, Above Average = 35.00 to 39.99, Below Average = 30.00 to 34.99, Low = 25.00 to 29.99, Very Low = less than 25.00.

**Conclusions and Implications**

Video imaging systems and Wulf’s system have been shown to perform at current chain speeds (over 300 hd/h) and accurately (over 90%) segment the cattle population into tender versus not tender categories, while doing so in a non-invasive fashion. In an industry where consumers are becoming more demanding of the end product and are willing to purchase “branded” or “certified” products in search of a consistently good eating experience, instrument technologies will be an integral part of identifying potentially tender carcasses and more effectively sorting and marketing beef products; these technologies show overwhelming potential as the next phase in USDA Quality Grade Standard improvement.

**References**


Helming, B. 1996. The time has come to shift the responsibility of beef quality grading from the government to the private sector (Parts 1 and 2). Report No. 442, 443. Bill Helming Consulting Services, Olathe, KS.


