On-line classification of US Select beef carcasses for longissimus tenderness using visible and near-infrared reflectance spectroscopy

S.D. Shackelford *, T.L. Wheeler, M. Koohmaraie

Roman L. Hruska US Meat Animal Research Center, USDA, ARS, Clay Center, NE 68933-0166, USA

Received 6 January 2004; received in revised form 24 August 2004; accepted 24 August 2004

Abstract

The current experiment was conducted to evaluate the on-line application of visible and near-infrared spectroscopy (VISNIR) to US Select carcasses during commercial beef carcass grading procedures to predict tenderness of longissimus steaks after 14 days of refrigerated storage. A regression model was calibrated using 146 carcasses and tested against an additional 146 carcasses. Carcasses were segregated into VISNIR-based tenderness classes based on whether their VISNIR-predicted slice shear force value was less than (tender) or greater than (tough) the median predicted slice shear force value. Carcasses classified as tender by VISNIR had a lower mean SSF value, were less likely to have slice shear force values greater than 245 N, had higher trained sensory panel tenderness ratings, and were less likely to have trained sensory panel tenderness ratings below slightly tender than were carcasses classified as tough \((P < 0.001)\). This technology might be useful for identification of US Select carcasses that excel in longissimus tenderness.

Published by Elsevier Ltd.

Keywords: Beef; Near-infrared spectroscopy; Tenderness

1. Introduction

US Select cuts are currently marketed at a discount relative to US Choice cuts despite the fact that many cuts from US Select carcasses are very tender. Consumers have shown a strong willingness to pay a premium for “Tender Select” cuts that combine superior tenderness with the leanness of Select (Shackelford et al., 2001). Therefore, meat retailers have expressed strong interest in marketing a “Tender Select” product line. However, beef packing companies have not been able to meet the needs of the retailers because non-invasive methods to predict tenderness, such as Beefcam (Belk et al., 2000) and colorimeter (Wulf & Page, 2000), have been ineffective for classifying Select carcasses (Wheeler et al., 2002). Thus, there is a need to develop a non-invasive method to accurately identify US Select carcasses that excel in meat tenderness. Several studies have shown that near-infrared reflectance spectroscopy can be used to predict beef tenderness (Byrne, Downey, Troy, & Buckley, 1998; Hildrum, Nilsen, Mielenik, & Naes, 1994; Hildrum et al., 1995; Mitsumoto, Maeda, Mitsuhashi, & Ozawa, 1991; Naes & Hildrum, 1997; Park, Chen, Hruschka, Shackelford, & Koohmaraie, 1998). However, the procedures used in those studies
were either destructive in that they required excision of a muscle sample for spectroscopy (Hildrum et al., 1994, 1995; Naes & Hildrum, 1997; Park et al., 1998) or they were limited to sampling a very small area (4 cm²) and thus, would be highly subject to error induced by non-representative sampling of the target muscle (Byrne et al., 1998; Mitsumoto et al., 1991). We (Shackelford, Wheeler, & Koohmaraie, 2004) developed a highly repeatable method for on-line spectroscopic evaluation of *longissimus* quality traits of ribbed beef carcasses using a high-intensity reflectance probe that allowed for sampling of a representative portion of the exposed *longissimus* cross-section of ribbed beef carcasses. Therefore, the current experiment was conducted to evaluate the on-line application of visible and near-infrared spectroscopy (VISNIR) to US Select carcasses during commercial beef carcass grading procedures to predict tenderness of *longissimus* steaks after 14 days of refrigerated storage.

### 2. Materials and methods

Spectroscopy was conducted on-line at two large-scale commercial fed-beef processing facilities. At 24 h postmortem, carcasses were ribbed conventionally between the 12th and 13th ribs for determination of USDA beef quality and yield grades and spectroscopy was conducted on the *longissimus* cross-section. Spectroscopy was conducted on the beef grading bloom chain approximately 2 min after the carcasses were ribbed. Spectroscopy only was conducted on carcasses that were likely to be graded USDA Select and USDA yield grade 1, 2, or 3. After grading, any carcasses that did not meet these specifications were excluded from the experiment. The carcasses that were included in the experiment had a slight amount of marbling in the *longissimus* which suggests that the *longissimus* intramuscular fat content was approximately 3.4% (Savell, Cross, & Smith, 1986).

Spectroscopy was conducted using the optimal protocol developed by Shackelford et al. (2004). Spectroscopy was conducted using a Model A108310 LabSpec Pro portable spectrophotometer (ASD; Analytical Spectral Devices, Inc., Boulder, CO) which was equipped to collect spectra from 350 to 2500 nm. Spectra were collected via a ASD Model 135090 2-m long fiber optic jumper cable attached to a ASD Model A122000 high-intensity reflectance probe, that served as an external light source (2900 K color temperature quafiz halogen light) to illuminate the object of interest. The field of view was restricted to 50 mm in diameter using a modified ASD Model A122040 field of view limiter and plate. This field of view size was such that a large area of the cross-section of the *longissimus* was sampled.

#### 2.1. Slice shear force

Following spectroscopy and grading, carcasses were fabricated and the NAMP 180 strip loin (NAMP, 1997) was obtained from the left side of each carcass (n = 292). Strip loins were obtained from 143 and 149 carcasses at the two respective processing facilities. Strip loins were transported to the MARC abattoir and aged (1 °C) until 14 d postmortem. At 14 d postmortem, a 2.54-cm thick *longissimus* steak was removed from the anterior end of each strip loin, cooked using a belt grill as described by Wheeler, Shackelford, and Koohmaraie (1998), and slice shear force (SSF) was measured as described by Shackelford, Wheeler, and Koohmaraie (1999). The remainder of each strip loin was frozen (−20 °C). Subsequently, two steaks (2.54 cm thick) were obtained from the anterior end of each frozen strip loin for trained descriptive attribute sensory panel evaluation.

<table>
<thead>
<tr>
<th>Trait</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>% &gt;245 N</th>
<th>% &lt;5²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calibration data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slice shear force (N)</td>
<td>146</td>
<td>186.4</td>
<td>67.7</td>
<td>92.2</td>
<td>455.2</td>
<td>17.8</td>
<td>–</td>
</tr>
<tr>
<td>Tenderness*</td>
<td>146</td>
<td>5.3</td>
<td>0.8</td>
<td>2.4</td>
<td>6.6</td>
<td>–</td>
<td>28.8</td>
</tr>
<tr>
<td>Juiciness*</td>
<td>146</td>
<td>5.4</td>
<td>0.3</td>
<td>4.7</td>
<td>6.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Beef flavor intensity*</td>
<td>146</td>
<td>5.1</td>
<td>0.3</td>
<td>4.5</td>
<td>5.8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Prediction data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slice shear force (N)</td>
<td>146</td>
<td>186.4</td>
<td>66.7</td>
<td>91.2</td>
<td>437.6</td>
<td>17.8</td>
<td>–</td>
</tr>
<tr>
<td>Tenderness*</td>
<td>146</td>
<td>5.4</td>
<td>0.8</td>
<td>2.3</td>
<td>7.0</td>
<td>–</td>
<td>23.3</td>
</tr>
<tr>
<td>Juiciness*</td>
<td>146</td>
<td>5.4</td>
<td>0.3</td>
<td>4.6</td>
<td>6.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Beef flavor intensity*</td>
<td>146</td>
<td>5.1</td>
<td>0.3</td>
<td>4.2</td>
<td>5.6</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

* Tenderness, juiciness, and beef flavor intensity rated on 8-point scales (1 = extremely tough, dry, and bland; 8 = extremely tender, juicy, and intense).

² 5 = slightly tender.
2.2. Trained sensory panel

Frozen steaks were thawed (24 h at 5 °C) and cooked using a belt grill as described by Wheeler et al. (1998). Immediately after post-cooking temperature rise was completed, steaks were sliced and served. Each panelist received three random cubes (1.3 cm × 1.3 cm × cooked steak thickness) from each sample. Sensory panelists scored steaks for tenderness on an eight-point scale (1 = extremely tough and 8 = extremely tender). The eight-member sensory panel was selected and trained according to Cross, Moen, and Stanfield (1978) and was highly experienced. With the protocol used in this experiment, the eight-member sensory panel has been reported (Wheeler et al., 1998) to measure tenderness with a high (0.87) level of repeatability.

![Calibration dataset (n=146)](calibration_dataset.png)

![Prediction dataset (n=146)](prediction_dataset.png)

Fig. 1. Scatter plots for predicted vs observed slice shear force for the calibration (a) and prediction (b) data sets. The regression model was developed using the calibration data set and tested on the prediction data set. Open and closed circles correspond to the carcasses represented by white and black bars, respectively, in Fig. 2.
2.3. Statistical analysis

To facilitate computations and to reduce spectral noise, spectra were reduced by averaging groups of nine consecutive wavelengths using Unscrambler® (Version 7.5; Camo, Inc., Corvallis, OR).

Carcasses were blocked by plant and observed slice shear force and one-half of the carcasses were assigned to a calibration data set, which was used to develop regression equations, and one-half of the carcasses were assigned to a prediction data set, which was used to validate the regression equations (Neter, Wasserman, & Kutner, 1989).

Forwards stepwise regression (SAS Inst., Inc., Cary, NC) was used to construct 1 to 20-variable regression equation for the prediction of SSF. The 10-variable regression equation was tested against the prediction data set.

The equations were validated as follows: carcasses were classified as “Tender” if their VISNIR-predicted slice shear force value was less than the median predicted slice shear force value. Carcasses were classified as “Tough” if their VISNIR-predicted slice shear force value was greater than the median predicted slice shear force value. The median predicted slice shear force value was 184 and 182 N for the calibration and prediction data sets, respectively. Analysis of variance was used to determine the effect of VISNIR classification on slice shear force and trained sensory panel traits. Analysis of variance was conducted using the GLM procedure of SAS. The frequency of carcasses with slice shear force values >245 N and the frequency of carcasses with trained sensory panel tenderness ratings <5 was calculated for each VISNIR class. Differences in these frequencies among VISNIR classes was compared using the DIFFER program of PEPI (Version 2; USD, Inc., Stone Mountain, GA).

Table 2
Effect of VISNIR classification on slice shear force and trained sensory panel traits for the calibration and prediction data sets

<table>
<thead>
<tr>
<th>VISNIR classification a</th>
<th>Slice shear force (N)</th>
<th>Percentage of slice shear force values &gt;245 N</th>
<th>Tenderness rating b</th>
<th>Percentage of tenderness ratings &lt;5 c</th>
<th>Juiciness rating b</th>
<th>Beef flavor intensity rating b</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calibration data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tender</td>
<td>157</td>
<td>5.5</td>
<td>5.6</td>
<td>17.8</td>
<td>5.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Tough</td>
<td>216</td>
<td>30.1</td>
<td>5.0</td>
<td>39.7</td>
<td>5.4</td>
<td>5.1</td>
</tr>
<tr>
<td>SEM</td>
<td>7.2</td>
<td>4.3</td>
<td>0.1</td>
<td>5.2</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.01</td>
<td>0.34</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Prediction data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tender</td>
<td>160</td>
<td>5.5</td>
<td>5.6</td>
<td>11.0</td>
<td>5.5</td>
<td>5.1</td>
</tr>
<tr>
<td>Tough</td>
<td>212</td>
<td>30.1</td>
<td>5.1</td>
<td>35.6</td>
<td>5.4</td>
<td>5.1</td>
</tr>
<tr>
<td>SEM</td>
<td>7.1</td>
<td>4.3</td>
<td>0.1</td>
<td>4.8</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.20</td>
<td>0.99</td>
</tr>
</tbody>
</table>

a Carcasses were classified as “Tender” if their VISNIR-predicted slice shear force value was less than the median predicted slice shear force value. Carcasses were classified as “Tough” if their VISNIR-predicted slice shear force value was greater than the median predicted slice shear force value. The median predicted slice shear force value was 184 and 182 N for the calibration and prediction data sets, respectively.

b Tenderness, juiciness, and beef flavor intensity rated on 8-point scales (1 = extremely tough, dry, and bland; 8 = extremely tender, juicy, and intense).

c 5 = slightly tender.
pay a premium for a product that is both lean and tender. A processing company could use this technology to select carcasses with low predicted SSF values and increase the likelihood of achieving a high level of consumer satisfaction. For example, if the processor that produced the carcasses used in this experiment had a premium product line that used 50% of its US Select carcasses, the processor could have assigned the carcasses with predicted SSF values less than the median to the premium line and reduced the frequency of *longissimus* SSF values greater than 245 N from 17.8% (Table 1) to 5.5% (Table 2) for the premium line.

When carcasses in the calibration data set were segregated into VISNIR-based tenderness classes based on whether their VISNIR-predicted slice shear force value was less than (tender) or greater than (tough) the median predicted slice shear force value, it was determined that the tender class had a lower mean SSF value, a lower

![Calibration data set](image1)

![Prediction data set](image2)

Fig. 2. Frequency distributions for observed slice shear force for the calibration (a) and prediction (b) data sets.
percentage of carcasses with slice shear force values greater than 245 N, higher trained sensory panel tenderness ratings, and a lower percentage of carcasses with trained sensory panel tenderness ratings below slightly tender (Table 2). Likewise, when the equation developed in the calibration data set was tested on the prediction data set, the tender class had a lower mean SSF value, a lower percentage of carcasses with slice shear force values greater than 245 N, higher trained sensory panel tenderness ratings, and a lower percentage of carcasses with trained sensory panel tenderness ratings below slightly tender (Table 2). For the prediction data set, 22 of the 73 (30.1%) carcasses with VISNIR-predicted SSF values greater than the median (182 N) had SSF values above 245 N. In contrast, only 4 of the 73 (5.5%) carcasses with VISNIR-predicted SSF values less than the median had SSF values above 245 N. Likewise, whereas 26 of the 73 (35.6%) carcasses with VISNIR-predicted SSF values greater than the median had sensory panel tenderness ratings below slightly tender, only 8 of the 73 (11.0%) carcasses with VISNIR-predicted SSF values less than the median had sensory panel tenderness ratings below slightly tender. Further inspection of the frequency distributions of the slice shear force values of the VISNIR classes revealed that the frequency of carcasses in each 49 N increment from 196 to 491 N was numerically lower for the tender class (Fig. 2).

This technology was evaluated on US Select carcasses because (1) US Select cuts are currently marketed at a discount relative to US Choice cuts despite the fact that many cuts from US Select carcasses are very tender, (2) consumers have shown a strong willingness to pay a premium for “Tender Select” cuts that combine superior tenderness with the leanness of Select (Shackelford et al., 2001), (3) it is likely that higher degrees of marbling would interfere with the ability of VISNIR to predict tenderness and (4) identification of tender longissimus from US Select would increase the value of carcasses usually discounted because of assumed lower eating quality. To our knowledge, this is the first report of a system that is capable of non-invasively classifying US Select carcasses for longissimus tenderness with a useful level of accuracy. However, further testing is needed to determine whether the effectiveness of the VISNIR prediction achieved in the current experiment can be achieved in other populations.

4. Conclusions

The present experiment indicates that US Select carcasses can be non-invasively classified for longissimus tenderness using visible and near-infrared spectroscopy. This technology might be useful for identification of US Select carcasses that excel in longissimus tenderness for use in branded beef programs.

Acknowledgments

We are grateful to Patty Beska, Kathy Mihm, Pat Tammen, In-Ho Hwang, and Min-Suk Rhee for technical assistance and we are grateful to Marilyn Bierman and Carol Grummert for secretarial assistance.

References


systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science, 80*, 3315–3327.