Implementation of Instrument Grading for Beef Carcasses

57th Florida Beef Cattle Short Course
Gainesville, FL

Kerry R. Smith, Ph.D.
United States Department of Agriculture
Agricultural Marketing Service
Livestock and Seed Program
Standards, Analysis and Technology Branch
Government, Industry and Academia Working Together

- USDA, AMS
  - J. Wise, M. Feil, J. Ransom, K. Smith, L. Yates, M. O’Connor
- USDA, MARC
  - M. Koohmaraie, S. Shackelford, T. Wheeler, A. King
- NCBA
  - James “Bo” Reagan, Bucky Gwartney
- Colorado State University
  - K. Belk, C. Moore
- Technology Providers
  - E + V, Oranienburg, Germany
  - RMS, Fort Collins, CO
Image Based Assessment

- Obtain image
- Determine numerous metrics
- Calculate factors using approved equations
Accuracy and Precision

- The average value might be an *accurate* estimate of the true value, but the individual values may be *inaccurate*.

- It is not possible to reliably achieve accuracy in individual measurements without precision.
Accuracy
the degree a measured or calculated value conforms to an actual (true) value or to an accepted standard
Precision

the degree to which two or more measurements or calculations show the same or similar results
Precision

the degree to which two or more measurements or calculations show the same or similar results
Beef Grades

A Reflection of not only the Quality but the Quantity

● Quality Grades
  - Prime
  - Choice
  - Select

● Factors
  - Marbling
    ● Traces or less
    ● Slight
    ● Small
    ● Modest
    ● Moderate or more
  - Maturity
    ● Skeletal
    ● Lean
  - Lean Firmness

● Yield Grades
  - Adjusted Preliminary Yield Grade (Fat Thickness)
  - Kidney, Pelvic & Heart Fat Percent; KPH
  - Ribeye Area; REA
  - Hot Carcass Weight
Marbling
Beef Grading

26,034,000 Carcasses were graded in 43 plants across the U.S. in FY 2007
# Beef Grading Assignments

<table>
<thead>
<tr>
<th>Plants</th>
<th>Graders</th>
<th>Graded Per Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>One</td>
<td>56,250</td>
</tr>
<tr>
<td>6</td>
<td>Two</td>
<td>60,000</td>
</tr>
<tr>
<td>5</td>
<td>Three</td>
<td>75,000</td>
</tr>
<tr>
<td>4</td>
<td>Four</td>
<td>80,000</td>
</tr>
<tr>
<td>11</td>
<td>Six</td>
<td>330,000</td>
</tr>
<tr>
<td>2</td>
<td>Eight</td>
<td>66,000</td>
</tr>
</tbody>
</table>
Just as there is an inherent variation among animals and carcasses ...

There is inherent variation in assessing beef quality and yield!
Impact of Two or More Graders within Plants
Impact of Two or More Graders within Plants
Impact of Two or More Graders within Plants
Impact of Two or More Graders within Plants

Grader 1
Grader 2
Grader 3
Plant Average
Impact of Two or More Plants within an Area
Impact of Two or More Areas within a Region
Impact of Two or More Regions Across the Nation
Rationale for Instrument Grading

- Reduce variation
  - Within plants
  - Between plants, areas and regions
- Increase accuracy, precision and resolution of grade factor assessment
  - Official yield grades in tenths
  - Refinement of marbling degree categories
    - Instead of a Select Grade
      - “Low Select” and “Select Plus”
    - Instead of a Choice Grade
      - “Low, Mid and High” Choice
History of Instrument Grading

- **1970’s**
  - USDA and NASA
  - Kansas State University

- **1980’s**
  - USDA/MARC
  - KSU, Whole carcass cutout test

- **1990’s**
  - Beef Industry Value-based Marketing Task Force
  - National Beef Instrument Assessment Symposium
History of Instrument Grading

● Dawn of a New Millennium
  – 2003
    ● Ribeye Area
  – 2005
    ● Yield Grade
  – 2006
    ● Marbling
Instrument Grading Task Force

- Composed of Representatives from:
  - Beef producer groups
  - Equipment manufacturers
  - Academia
  - Beef processing companies
  - USDA

*Instrument Performance Standards and In-Plant Operational Procedures*
Instrument Assessment and Approval

- Methodology for assessing instrument systems was adapted from the methods used for assessing Supervisory Meat Graders

- Grading Metrics
  - Accuracy
  - Precision
  - Bias
  - Comparative referencing

NOTICE TO THE TRADE

USDA PROPOSES STANDARD FOR APPROVAL AND USE OF INSTRUMENT GRADING SYSTEMS FOR MARBLING SCORE DETERMINATION OF BEEF CARCASSES

The Department of Agriculture’s (USDA) Agricultural Marketing Service is proposing a standard for the accuracy and repeatability of instrument grading systems for assessing marbling in beef carcasses. Once approved under this standard, an instrument grading system may be used to assess marbling in the official USDA grading process.

USDA’s beef grading service provides third-party beef carcass evaluation to segregate a large and diverse population into smaller, more homogeneous groups (grades) that are utilized in the price/value discovery process by both buyers and sellers of beef. It is essential that USDA take all reasonable steps – including the adoption of new technology – to ensure accurate and consistent grading decisions on a nationwide basis.

The proposed standard for marbling assessment was established in consultation with an Industry working group consisting of members from the National Cattlemen’s Beef Association, meat packing firms, producers, technology manufacturers, and academia. A performance element trial assessing various technologies was conducted earlier this year. USDA has previously established standards for...

Performance Standard

BACKGROUND:
The Department of Agriculture (USDA), Agricultural Marketing Service (AMS), Livestock and Seed (LS) Program will accept beef carcass marbling scores made by approved instruments to assist in determining the official grade of beef carcasses. To be approved by the LS Program for such a purpose, the instrument must meet certain performance requirements for accuracy and repeatability in the prediction of the marbling scores for beef carcasses. The performance requirements outlined in this document were established after consultation with an Industry Working Group that was convened by the LS Program and the National Cattlemen’s Beef Association (NCBA). The Industry Working Group was comprised of representatives of USDA, NCBA, packing companies, producers, instrument manufacturers, and academia.

PURPOSE:
The purpose of this document is to provide the beef industry with a performance requirement for instrument grading systems for marbling score determination of beef carcasses.
Marbling Standard Development Trials

- Two Facilities
- Every carcass imaged
- Official USDA Grade Data collected on 1,000+ carcasses at each facility
- Data analyzed multiple ways by AMS
  - In-depth discussion with USDA and Industry Grading Working Group
- Outcome – *Heteroscedastic Variation*
  - degree of variation increased with increasing marbling score
Official Standard for Instrument Marbling Evaluation

- Published June 2006

“Performance Requirements for Instrument Marbling Evaluation (PRIME)”

- Two Sets of Performance Requirements
  - Demonstration of Repeatability, Accuracy, and Precision
  - Operational Procedures for Implementation and Verification
PRIME I – Demonstration of Repeatability, Accuracy, and Precision

- Use of multiple plants (four) to test instrument performance
- 1,000+ carcasses per facility
- Data collected created 2 data sets
  - Calibration
  - Test
- Five expert panel members
  - Higher level of confidence
- Two Phases
  - Repeatability
  - Accuracy and precision
How Do the Instrument Systems Compare?

<table>
<thead>
<tr>
<th>Location</th>
<th>Correlation N</th>
<th>N</th>
<th>Average Residual Marbling Score</th>
<th>Residual Standard Deviation</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Philadelphia</td>
<td>-5.5</td>
<td>280</td>
<td>44.5</td>
<td>-0.0809</td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>-1.4</td>
<td>300</td>
<td>42.0</td>
<td>-0.0479</td>
<td></td>
</tr>
<tr>
<td>Greeley</td>
<td>-4.8</td>
<td>760</td>
<td>34.9</td>
<td>-0.0753</td>
<td></td>
</tr>
<tr>
<td>Fresno</td>
<td>-8.4</td>
<td>1500</td>
<td>33.2</td>
<td>-0.0856</td>
<td></td>
</tr>
<tr>
<td>Plainview</td>
<td>-1.7</td>
<td>960</td>
<td>36.2</td>
<td>-0.0554</td>
<td></td>
</tr>
<tr>
<td>Garden City</td>
<td>-4.5</td>
<td>1760</td>
<td>44.1</td>
<td>-0.0360</td>
<td></td>
</tr>
<tr>
<td>Greeley</td>
<td>5.0</td>
<td>499</td>
<td>37.8</td>
<td>0.0444</td>
<td></td>
</tr>
<tr>
<td>Amarillo</td>
<td>-2.0</td>
<td>2140</td>
<td>30.7</td>
<td>0.0111</td>
<td></td>
</tr>
<tr>
<td>Greeley</td>
<td>2.9</td>
<td>1680</td>
<td>38.2</td>
<td>0.0182</td>
<td></td>
</tr>
<tr>
<td>ALL</td>
<td>-2.5</td>
<td>9879</td>
<td>37.2</td>
<td>-0.0271</td>
<td></td>
</tr>
<tr>
<td>Instrument</td>
<td>2.1</td>
<td>2061</td>
<td>30.5</td>
<td>-.050</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>0 ± 10</td>
<td></td>
<td>≤ 35</td>
<td>0 ± 0.0750</td>
<td></td>
</tr>
</tbody>
</table>
Instrument Marbling Assessment

![Graph showing the relationship between Official Marbling Score and Predicted Marbling Score. The graph demonstrates a strong positive correlation, with data points clustering closely around the diagonal line.]
Instrument Marbling Assessment
PRIME II – Implementation and Verification of Operational Procedures

- Instrument Implementation
  - Plant-by-plant basis
  - Quality Management System
  - Additional USDA Requirements
- ASTM Standard F2341-03
  “Procedure for user requirements for Livestock, Meat and Poultry Evaluation Devices or Systems”
  - Installation
  - Training
  - Calibration
  - Testing
  - Maintenance
  - Traceability
Additional USDA Requirements

- Identity of Hardware Components
- Maintain Carcass Identity
- Maintain Transactional Integrity and Security
  - Image
  - Data
- Digital Security
  - Strong firewall consisting of multiple layers of security
  - Electronic points of entry management - invisible or unavailable
  - Patch security management
Reducing Risk

FAA worries on-board Net opens jets to cyberattack

WASHINGTON — The government is enabling rules to open a new kind of computer door: A cyberattack first could discover the in-flight Internet service that connects passengers to the outside world.

The FAA, which regulates airline safety, said in a notice that it will consider allowing in-flight computer networks to be connected to the Internet. The agency is concerned about the potential for such networks to be vulnerable to attacks.

The agency said in a notice that it will consider allowing in-flight computer networks to be connected to the Internet. The agency is concerned about the potential for such networks to be vulnerable to attacks.
Defining Requirements

ISO/IEC Standard 15408

A flexible catalogue of standardized IT security requirements (features and assurances)

Protection Profiles

✓ Operating Systems
✓ Database Systems
✓ Firewalls
✓ Smart Cards
✓ Applications
✓ Biometrics
✓ Routers
✓ VPNs

Access Control
Identification
Authentication
Audit
Cryptography

Consumer-driven security requirements in specific information technology areas
Industry Responds

**Protection Profile**

**Firewall Security Requirements**

**Consumer statement of IT security requirements to industry in a specific information technology area**

**Security Targets**

- CISCO Firewall
- Lucent Firewall
- Checkpoint Firewall
- Network Assoc. FW

**Vendor statements of security claims for their IT products**

**Security Features and Assurances**
Demonstrating Conformance

Vendors bring IT products to independent, impartial testing facilities for security evaluation.

Private sector, accredited security testing laboratories conduct evaluations.

Security Features and Assurances

Common Criteria Testing Labs

Test results submitted for post-evaluation validation.
Validating Test Results

Validation Body validates laboratory’s test results

Laboratory submits test report to Validation Body

Common Criteria Validation Body

NIAP issues Validation Report and Common Criteria Certificate
In plant monitoring procedures

- Expertise
  - Meat Grading and Certification Branch
- Monitor instrument operation
- Continuous review of on-line operation
  - 100% monitoring
    - Carcass-to-carcass
    - Day-to-day
- Data and grade integrity
Accurate Instrument Assessment

![Graph showing the comparison between reality and instrument readings. The graph has a bell curve with a peak at zero, indicating a high degree of accuracy. The x-axis represents the scale difference, ranging from -150 to 250, and the y-axis shows the corresponding readings.]
Biased Instrument Assessment

[Graph showing a comparison between Reality, Instrument, and Difference]
Implementing Instrument Grading

- Packers are establishing a broad instrument base
- USDA instrument monitoring procedures have been distributed to industry
- Final draft of USDA’s Digital Security Protection Profile delivered
  - Strong firewall consisting of multiple layers of security
  - Electronic points of entry management
    - invisible or unavailable
  - Transactional Integrity and Security
    - Image
    - Data

However ...
Industry Identified a Difference Between Instruments & Graders

- Disparity between USDA field quality grades and instrument assessed grades
Basis of the Divergence

- Instrument performance?
- USDA field graders?
- Plant procedures?
- Or some possible combination of factors?
Industry-Government Task Force Identifies Objectives

- Quantify any divergence between USDA field and instrument quality grades among nine plants
- Identify the source of any divergence among USDA field graders and instruments through the use of a four-member expert panel
Important Design Elements

- USDA blinded to identity of all plants and grader identifications
- USDA graders should be unaware that data is being collected
  - Hawthorne effect - If graders perceived that they were being observed, then their grading pattern might not reflect their normal day-to-day patterns

Study 1: Is There a Divergence?

- e+v served as data repository
- Nine plants
- Instrument marbling and USDA grade data
- 1.2 Million Carcasses
  - approximately one-half of the national beef slaughter under Federal inspection during the four-week period
- e+v estimated grade line offsets
- Grade line divergence was reviewed by USDA and the data was reanalyzed for comparison
Quality Grade Distribution by Instrument Marbling Score

![Quality Grade Distribution by Instrument Marbling Score](image_url)
In a perfect world ...
In a perfect world ... Only Select
In a perfect world ... Only Choice
In a perfect world ... Only Prime
In Reality ...

Marbling Score
Percent Quality Grade
No Roll Select Choice Prime
In Reality ... with Respect to Grades, Probabilities are not All or None
Grade Probability as a Function of Instrument Marbling Score

![Graph showing grade probability as a function of instrument marbling score. The graph includes data points for %Se, %Ch, and %Pr across different marbling scores. The x-axis represents the e+v marbling score, ranging from 100 to 1100, and the y-axis represents the percent grade at each marbling score, ranging from 0% to 100%. The graph highlights the distribution and probability of grades at various marbling scores.]
Industry Suggested Adjustment

Percent Grade at Each Marbling Score

- %Se
- %Ch
- %Pr

e+v Marbling Score

100 200 300 400 500 600 700 800 900 1000 1100
USDA Assessment of Field Grade Lines

Percent Grade at Each Marbling Score

%Se %Ch %Pr
## Quality Grade Line Divergence Estimates

<table>
<thead>
<tr>
<th>Grade Line</th>
<th>Approved e+v Instrument</th>
<th>e+v Suggested Adjustment</th>
<th>Assessed USDA Field Grade Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select</td>
<td>295</td>
<td>292</td>
<td>283</td>
</tr>
<tr>
<td>Choice</td>
<td>395</td>
<td>376</td>
<td>376</td>
</tr>
<tr>
<td>Prime</td>
<td>695</td>
<td>630</td>
<td>656</td>
</tr>
</tbody>
</table>

### Marbling Degree Based Quality Grade

<table>
<thead>
<tr>
<th>% No Roll</th>
<th>Marbling Degree Based Quality Grade</th>
<th>Actual USDA Quality Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>% Se</td>
<td>51.9%</td>
<td>42.0%</td>
</tr>
<tr>
<td>% Ch</td>
<td>47.5%</td>
<td>57.0%</td>
</tr>
<tr>
<td>% Pr</td>
<td>0.5%</td>
<td>1.1%</td>
</tr>
</tbody>
</table>
Study 1 Summary

● A significant divergence exists for all grade lines
● The question now becomes,

What is the source of the difference?
Study 2: Source of the Divergence?

- **Image Collection**
  - Raw and processed images
  - Nine plants
  - e+v served as image repository
  - Instrument marbling and USDA grade data

- **Image Selection**
  - Critical marbling lines
  - Randomly selected to adequately represent
    - Grader diversity
    - Shifts
    - Plants

- **One image per carcass**
  - Not necessarily the side graded
Methods – *Panel and Facilitators*

- **Academicians**
  - Dr. Dan Hale, Texas A&M University
  - Dr. Daryl Tatum, Colorado State University

- **USDA**
  - Joel Gottlieb, Audit, Review & Compliance Branch
  - Martin O’Connor, Standards, Analysis & Technology Branch

- **Process oversight**
  - Dr. Bo Reagan, National Cattlemen’s Beef Association

- **Process facilitation**
  - Dr. Lawrence Yates, USDA
Methods – *Image Review*

- **Raw images**
  - Marbling degree
  - Carcass presentation
  - Image capture process

- **Processed images**
  - Image processing suitability

- **Sample size**
  - 5,377 carcasses
  - Three day period
  - 300 image pairs per hour
Result Summary

- Seven plants, instrument = panel
  - Five plants, instrument = panel < USDA
  - Two plants, instrument = panel = USDA
- Two plants, instrument < panel < USDA
- Panel-instrument agreement rates were greater than USDA field grade-instrument agreement rates
# Quality Grade Divergence Estimates

## Comparison of the Two Studies

<table>
<thead>
<tr>
<th>Grade Line</th>
<th>1.2 M hd Data Set</th>
<th>Panel Data set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Approved e+v Instrument</td>
<td>e+v Suggested Adjustment</td>
</tr>
<tr>
<td>St - Se</td>
<td>295</td>
<td>292</td>
</tr>
<tr>
<td>Se - Ch</td>
<td>395</td>
<td>376</td>
</tr>
<tr>
<td>Ch - Pr</td>
<td>695</td>
<td>630</td>
</tr>
</tbody>
</table>
Overall Conclusion and Outcome

- The difference between instrument performance and that of USDA field graders was confirmed.
- The source of the divergence resulted in a large part from the performance of USDA field graders.
- A pilot project is being conducted using a prediction equation developed from the 1.2 million carcass data set.
Pilot Project

- One plant per company
- One week acclimation period followed by a 4 week trial
- Minimal grading supervision
- Limited scope ensures appropriate controls and evaluations
Pilot Project

- Data includes instrument marbling score, final quality grade, instrument yield grade, final yield grade
- Grader and instrument performance will be analyzed by shift, by day and by plant
- Decision logic between graders and instruments will provide an accurate and timely appraisal of any bias or influence
- Results will provide beta test data for evaluation as USDA migrates to fuller implementation
Percent Over-rides and Carcass Numbers by Quality Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>Over-rides</th>
<th>Inst. Cx</th>
<th>USDA Cx</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Roll</td>
<td>28%</td>
<td>1,973</td>
<td>12,065</td>
</tr>
<tr>
<td>Select</td>
<td>14%</td>
<td>53,313</td>
<td>63,947</td>
</tr>
<tr>
<td>Choice</td>
<td>15%</td>
<td>169,729</td>
<td>150,296</td>
</tr>
<tr>
<td>Prime</td>
<td>32%</td>
<td>8,319</td>
<td>7,026</td>
</tr>
</tbody>
</table>

86% of the Carcasses Quality Graded had Both Sides Imaged
## Comparing Instrument and USDA Quality Grade Assessments

<table>
<thead>
<tr>
<th>USDA Assessed Quality Grade</th>
<th>No Roll</th>
<th>Se</th>
<th>Ch</th>
<th>Pr</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Roll</td>
<td>0.6%</td>
<td>1.6%</td>
<td>2.8%</td>
<td>0.2%</td>
<td>5.2%</td>
</tr>
<tr>
<td>Se</td>
<td>0.2%</td>
<td>19.7%</td>
<td>7.4%</td>
<td>0.0%</td>
<td>27.4%</td>
</tr>
<tr>
<td>Ch</td>
<td>0.0%</td>
<td>1.5%</td>
<td>61.9%</td>
<td>0.9%</td>
<td>64.4%</td>
</tr>
<tr>
<td>Pr</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.6%</td>
<td>2.4%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Total</td>
<td>0.8%</td>
<td>22.8%</td>
<td>72.7%</td>
<td>3.6%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

- **Agree**: 84.7%
- **+ Bias**: 2.3%
- **- Bias**: 13.0%
- **Overrides**: 15.3%

### Instruments vs. USDA

<table>
<thead>
<tr>
<th>Instrument Assessed Quality Grade</th>
<th>Agree</th>
<th>+ Bias</th>
<th>- Bias</th>
<th>Overrides</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>84.7%</td>
<td>2.3%</td>
<td>13.0%</td>
<td>15.3%</td>
</tr>
</tbody>
</table>
## Original Approved Equation

<table>
<thead>
<tr>
<th></th>
<th>Select Line</th>
<th>Choice Line</th>
<th>Prime Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Line</td>
<td>280</td>
<td>385</td>
<td>640</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>23</td>
<td>22</td>
<td>33</td>
</tr>
<tr>
<td>Grade Basis</td>
<td>96.1%</td>
<td>94.1%</td>
<td>86.1%</td>
</tr>
</tbody>
</table>

![Graph showing % of No Roll, Se, Ch, Pr, pNR, pSe, pCh, pPr for different values of 200, 300, 400, 500, 600, 700, 800]
# Adjusted Equation

<table>
<thead>
<tr>
<th></th>
<th>Select Line</th>
<th>Choice Line</th>
<th>Prime Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Line</td>
<td>293</td>
<td>406</td>
<td>704</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>24</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Grade Basis</td>
<td>96.1%</td>
<td>94.0%</td>
<td>87.4%</td>
</tr>
</tbody>
</table>

![Graph showing the adjusted equation with data points and line segments for No Roll (Se), Select (Ch), Choice (Pr), and Prime (pNR, pSe, pCh, pPr).]
Future Opportunities

- Continuous improvement
  - Check Standards
  - Quality control attributes
- New technologies for tenderness
- Lamb Instrument Grading
  - Yield
  - Quality