

# United States beef quality as chronicled by the National Beef Quality Audits, Beef Consumer Satisfaction Projects, and National Beef Tenderness Surveys — A review

John Michael Gonzalez<sup>1,a,\*</sup> and Kelsey Jean Phelps<sup>1,a</sup>

\* **Corresponding Author:** John Michael Gonzalez  
**Tel:** +1-785-532-3448, **Fax:** +1-785-532-7059,  
**E-mail:** [johngonz@k-state.edu](mailto:johngonz@k-state.edu)

<sup>1</sup> Department of Animal Sciences and Industry, Kansas State University, Manhattan, KS 66506, USA

<sup>a</sup> These authors contributed equally to this work.

## ORCID

John Michael Gonzalez  
<https://orcid.org/0000-0002-1905-9294>  
Kelsey Jean Phelps  
<https://orcid.org/0000-0003-4559-1484>

Submitted Mar 14, 2018; Revised May 11, 2018;  
Accepted May 19, 2018

**Abstract:** Meat quality is a very difficult term to define because it means different things to different people. When purchasing beef, consumers in the United States are likely to consider color, price, marbling level, subcutaneous fat trim, or cut thickness when determining the quality of beef. Once consumers have consumed the product, meat quality becomes exponentially more difficult to define due to the subjective nature of this term. Traditionally, tenderness, juiciness, and flavor have been considered the three most important factors that determine the palatability of beef. Therefore, American meat science beef research and industry focus has turned to measuring and quantifying these 3 attributes objectively and subjectively, and to determining what influences them. In reviewing the scientific literature, attempting to meaningfully summarize the findings of the thousands of studies on beef meat quality is impossible due to the inherent differences in the objective and methodology of studies. Fortunately, the United States beef industry and their national trade association, the National Cattlemen's Beef Association (NCBA), have conducted numerous surveys and audits to characterize the quality of the products being produced and marketed by their cattlemen and the palatability perceptions of their consumers. The data produced by these studies is quite large and impossible to summarize in entirety in this review. Therefore, this review concentrates on the most important attributes that determine the value of a beef carcass and objectively measured and consumer-assessed palatability characteristics of fresh meat from these carcasses from 1987 through 2010.

**Keywords:** Beef; Palatability; Quality; Surveys; Tenderness

## INTRODUCTION

The term “meat quality” for the beef producer, packer, and retailer differs quite drastically when compared to the American beef consumer. These segments of the industry supply chain determine what constitutes meat quality of carcasses and, ultimately, their value based on a multitude of factors including lean maturity, bone maturity, lean color, marbling, loin eye area, and subcutaneous fat thickness. Along with hot carcass weight, the two most important measures of beef carcass and meat quality in the United States are yield grade and quality grade. Yield grade estimates the percentage of boneless, closely trimmed retail cuts from the chuck, rib, loin, and round. Carcass measures such as carcass weight, subcutaneous fat depth opposite the ribeye, ribeye area, and percentage of kidney, heart, and pelvic fat contribute to yield grade. The American yield grade system is based on a scale of 1 to 5, with 1 expected to yield the greatest percentage of hot carcass weight as closely-trimmed beef and 5 the least. Quality grade predicts expected palatability, and carcass measures such as maturity, amount and distribution of marbling, and lean firmness texture, and color all determine

quality grade. In young cattle harvested in the United States, 4 major quality grades exist, with Prime having the best expected palatability, followed by choice, select, and standard [1]. Today, these measures are primarily measured by instrumentation and USDA graders, and they are utilized to set the value of carcasses at the packer (abattoir); however, the packer also makes subtle adjustments to the value based on other measures including carcass weight, bruising, and breed influences, and chronological age of the animal. At the retailer, quality grade is mainly utilized to set the price of cuts when beef is merchandized. As Asia becomes a larger market for the export of American beef products, maintaining and improving the quality and value of these products will be instrumental in expanding the presence of American beef in the Asian diet.

## UNITED STATES NATIONAL BEEF QUALITY AUDITS

During 1973 and 1974, the USDA Market Consist Report [2] documented the carcass characteristics of cattle harvested in the United States. The industry did not conduct a follow-up audit for 17 years until the 1991 National Beef Quality Audit (NBQA). During this latter audit, researchers surveyed large packing plants encompassing 80% of inspected slaughters to characterize the carcasses. Serving as the baseline for future audits, this audit found genetic changes to the cattle population, combined with management changes, produced carcass with lower yield grades, less subcutaneous fat, larger ribeye areas, and greater hot carcass weights. Unfortunately, these practices also reduced the amount the high quality grades of carcasses presented for grading, which was a concern for eating quality, and many carcasses were still considered too fat as indicated by elevated yield grades [3]. Because of the 1991 audit, the authors suggested subsequent audits should be conducted every 4 to 5 years to track the progress of the industry. Therefore, an NBQA was conducted in 1995, which showed a reduction of excess fat had been achieved, as evidenced by a reduction in average carcass yield grade, but average carcass marbling score and quality grade were worse than in 1991. The authors speculated that because genetics control carcass characteristics, and genetics is slow to produce change, progress in these measures would not be immediate [4].

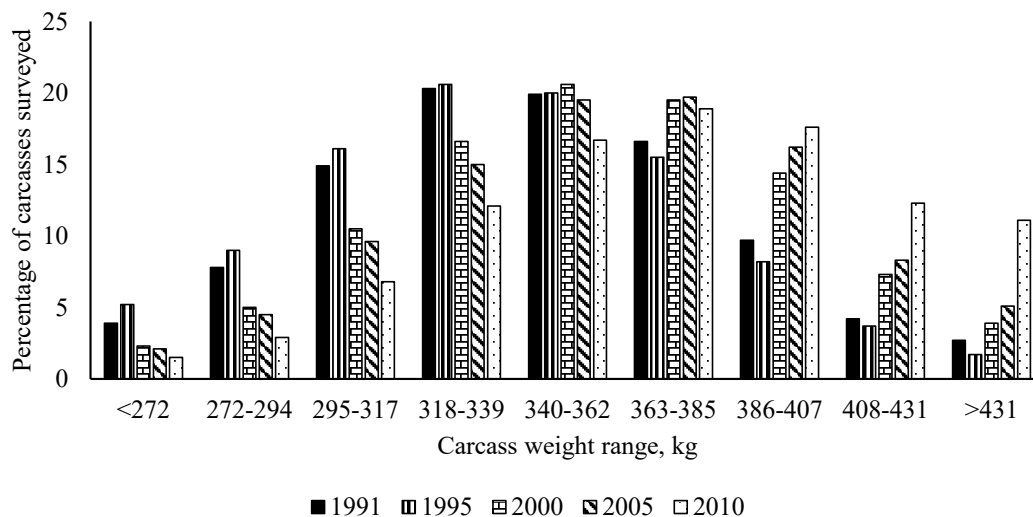
As the industry progressed into the next century, the 2000 NBQA found average carcass marbling scores and quality grades had returned to 1990 levels, but there were also large increases in average hot carcass weight and muscling as indicated by larger longissimus muscle areas. This resulted in carcasses maintaining similar yield grades as in previous years [5]. In the 2005 NBQA, researchers observed the same trends of heavier carcasses, larger longissimus muscle areas, greater marbling scores, increased quality grade, and constant yield grade. The authors noted this continued improvement in car-

cass quality may have been a function of genetics as packers harvested more black-hided cattle during this time period [6]. In the final peer-reviewed, published NBQA in 2011, Moore et al [7] found continued improvements in the same carcass attributes as the previous 2 surveys. Most surprisingly, average hot carcass weight and longissimus muscle area continued to increase, while quality grade, including the percentage of carcasses grading prime and choice, and marbling score also increased to the greatest values reported over the 20 years covered by the audits. Over this period, carcasses that would receive discounts because they had too elevated of a yield grade or too low a quality grade decreased, which suggested improvement in the consistency of carcasses being produced and marketed.

The NBQA audit has been a tool with which the beef industry recorded the changes in carcass composition and value from 1991 through 2011. When observing the carcass characteristics that are important in determining the value of a carcass and the quality of the meat produced, a couple of trends are apparent from these data. First, over the 20 years covered by the audits, there was a shift in the distribution of cattle carcass weights to indicate a larger percentage of carcasses being produced became heavier as time progressed (Figure 1). Initially, the shift toward heavier carcasses was due to changes in the genetics of cattle and nutritional programs; however, Moore et al [7] noted the large increases in hot carcass weights in the 2011 audit may have been due to use of growth promoting technologies and increased utilization of breeds that are larger framed and heavier muscled. Second, as depicted in Figure 2, there has also been an increase in muscling as indicated by larger average carcass longissimus muscle areas over time. At the same time, average carcass quality grade was not adversely affected, and actually increased numerically to on average almost meet the choice category. In contrast, average carcass yield grade had not changed since the 1995 audit and has stayed at around 2.9. Both Garcia et al [6] and Moore et al [7] hypothesized that the trend of increased average carcass weights and muscling while maintaining quality and yield grade may be due to the increase in branded beef programs and producers striving to produce carcasses that fit the requirements of these programs.

## CONSUMER SURVEYS OF BEEF PALATABILITY

In the early 1980s, the United States beef industry recognized consumer eating satisfaction was likely highly variable, but there were few studies that had evaluated consumer perceptions of palatability. The first National Consumer Retail Beef Study was conducted in 1985 [8,9], with further surveys occurring in 1993 [10], 1998 [11], 2006 [12], and 2010 [13]. Each was designed to examine the impact of quality grade, which



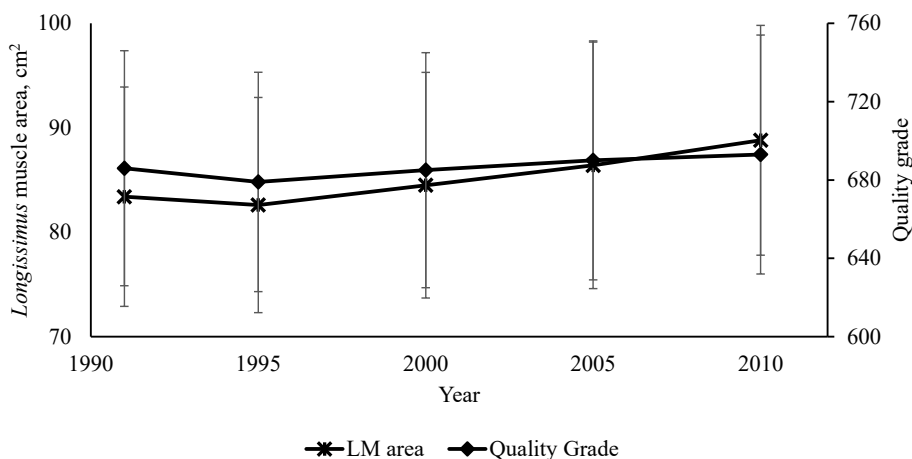
**Figure 1.** Distribution shift of carcass weights associated with the 5 National Beef Quality Audits. Adapted from Lorenzen et al [3], Boleman et al [4], McKenna et al [5], Garcia et al [6], and Moore et al [7].

is used to sort carcasses based on potential eating quality [14], and to examine the palatability of different steaks that are available at retail or foodservice for purchase by consumers.

The first National Consumer Retail Beef Survey was conducted in 2 phases. Phase I of the survey found two influencers of palatability, degree of doneness and marbling level, differed among consumers based on geographical region [8]. Phase II of the survey also detected regional differences in the influencers of a consumer’s palatability like or dislike of Choice and Select steaks [9]. Both these studies associated with the National Consumer Retail Beef Study demonstrated there were differences in palatability perceptions based on region. More importantly, an opportunity for the beef industry to market their products differently based on the location of their customers was identified.

The second national consumer research project was conduct-

ed in 1993 when its name was changed to The Beef Customer Satisfaction Project. In this survey, consumers in San Francisco, Houston, Chicago, and Philadelphia evaluated top loin, top sirloin, and top round steaks from five different quality grades during in-home consumption. Similar to the findings of the National Consumer Retail Beef Study, geographical location had a large influence on how consumers perceived the palatability of their steaks. Consumers in Chicago and Philadelphia rated Top Choice steaks more desirable for tenderness, juiciness, and beef flavor than the other quality grades; however, Houston and San Francisco consumers were not able to perceive differences in the palatability attributes between grades. The authors hypothesized the reason consumers in Chicago and Philadelphia could detect differences in the palatability characteristics of the different grades, while consumers in Houston and San Francisco could not, may have been due to



**Figure 2.** Increases in *Longissimus* muscle cross-sectional area and USDA quality grade over the 20 years associated with the National Beef Quality Audits. Adapted from Lorenzen et al [3], Boleman et al [4], McKenna et al [5], Garcia et al [6], and Moore et al [7].

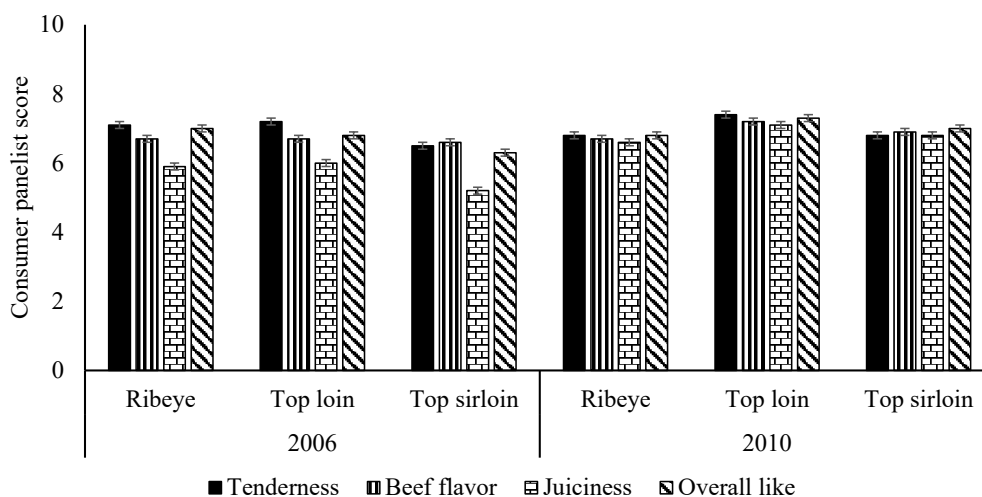
consumers in Chicago and Philadelphia being used to eating Choice beef and consumers in the other markets predominantly eating Select beef [10]. These findings provided further evidence to the beef industry that consumers in different locations of the United States perceive the palatability of beef products differently and marketing of products should consider this.

These early market studies provided an excellent platform for future national surveys. Instead of conducting separate “market basket” surveys in subsequent years, consumer palatability evaluations were conducted in accordance with the National Beef Tenderness Surveys (NBTS) [11-13]. While these surveys were designed to evaluate tenderness of beef steaks (as further discussed below), consumers were also asked to evaluate juiciness, flavor, and overall liking at university controlled-panel sessions. This allowed for better control of influencers of palatability including degree of doneness and method of cooking. Because different muscles were examined during the various studies, the following discussion will focus on consumer perceptions of ribeye steaks and include quality grade where data is available.

Consumers that participated in the 1998 NBTS rated Prime retail steaks better than all other grades for overall like and overall flavor, but there were no differences in ratings between the other grades. Consumers did not detect differences between quality grades for the other palatability characteristics (tenderness, juiciness, and beef flavor), and did not detect differences between the grades for the overall like rating of foodservice ribeye steaks [11]. In the 2006 NBTS, consumers did not compare retail ribeye steaks evaluated across grade, but ribeye steaks consistently had some of the greatest like scores for all palatability attributes including overall like, tenderness, beef flavor, and juiciness. Consumers in the 2006 NBTS did

evaluate foodservice ribeye steaks for the same attributes and determined that Select ribeye steaks had the greatest flavor like scores above all other grades which did not differ [12]. This was a surprising finding considering the opposite was found in retail ribeye steaks by Brooks et al [11], and that steaks with less intramuscular fat are expected to be less flavorful. In the 2010 NBTS, consumers continued to rate retail ribeye steaks as one of the best steaks in the beef carcass across all palatability characteristics. In contrast to Voges et al [12], consumers in the 2010 NBTS rated foodservice ribeye steaks and were able to find grade differences in juiciness, tenderness, and overall like. For the most part, Prime ribeye steaks received greater like scores than Low Choice and Ungraded steaks for overall like and juiciness. Interestingly, Select ribeye steaks received similar scores to Prime and Top Choice steaks for all palatability characteristics evaluated [13].

Consumer ratings of the 3 most important determinants of beef palatability, namely tenderness, juiciness and beef flavor of foodservice ribeye, top loin, and top sirloin steaks during the 2006 and 2010 NBTS are depicted in Figure 3. Over the 4 years between the studies, ribeye steaks only increased in like ratings of juiciness, but there was no effect on overall like of these steaks. In contrast, top loin and top sirloin steaks consumer like scores numerically increased between 0.3 and 0.5 of a point for tenderness, beef flavor and juiciness. This resulted in overall like scores increasing approximately 0.50 and 0.75 of a point, respectively. It is impossible to track the progress from 1991 and 1998 due to use of different like scales during both years; however, it appears when only 1 palatability attribute is improved, there was no effect on overall like scores, but when all three influencers of overall palatability are improved, overall like is modestly improved.



**Figure 3.** Consumer panel evaluations of tenderness, beef flavor, juiciness, and overall like for rib and loin steaks evaluated during the 2006 and 2010 National Beef Tenderness Surveys. 10 = very tender, extreme amount, very juicy, and like extremely; 1 = not at all tender, none at all, not at all juicy, and dislike extremely. Adapted from Voges et al [12] and Guelker et al [13].

## FOCUS ON TENDERNESS

Tenderness has been identified as the single most important factor affecting consumers' palatability perceptions, despite its influence not being strongly demonstrated in the original National Consumer Retail Beef Study [8]. Therefore, the NCBA commissioned the series of NBTS to track objective and subjective measures of tenderness. The first study, completed in 1990, was designed to provide baseline data from which industry leaders and researchers could determine the magnitude of the beef tenderness problem and set research priorities to improve beef tenderness. In the study conducted by Morgan et al [15], Warner-Bratzler shear force (WBSF) and trained taste panel palatability scores for chuck, rib, loin, and round roasts and steaks purchased at retail stores in 14 metropolitan cities were determined. While researchers identified cuts from the chuck and round as the most likely to be rated tough as reflected by WBSF and trained sensory panel scores, 23% of rib cuts and 17.5% of loin cuts had WBSF greater than 4.6 kg. In a study conducted in 1996 to gauge progress by industry, George et al [16] concluded the industry had made little progress in improving the tenderness of strip loin steaks, as the mean WBSF value was 3.46 kg and the chances of receiving a tough sensory panel rating was 24%. The authors also concluded insufficient postmortem aging and carcasses grading as Select and "commodity" Choice were the two issues negatively affecting tenderness.

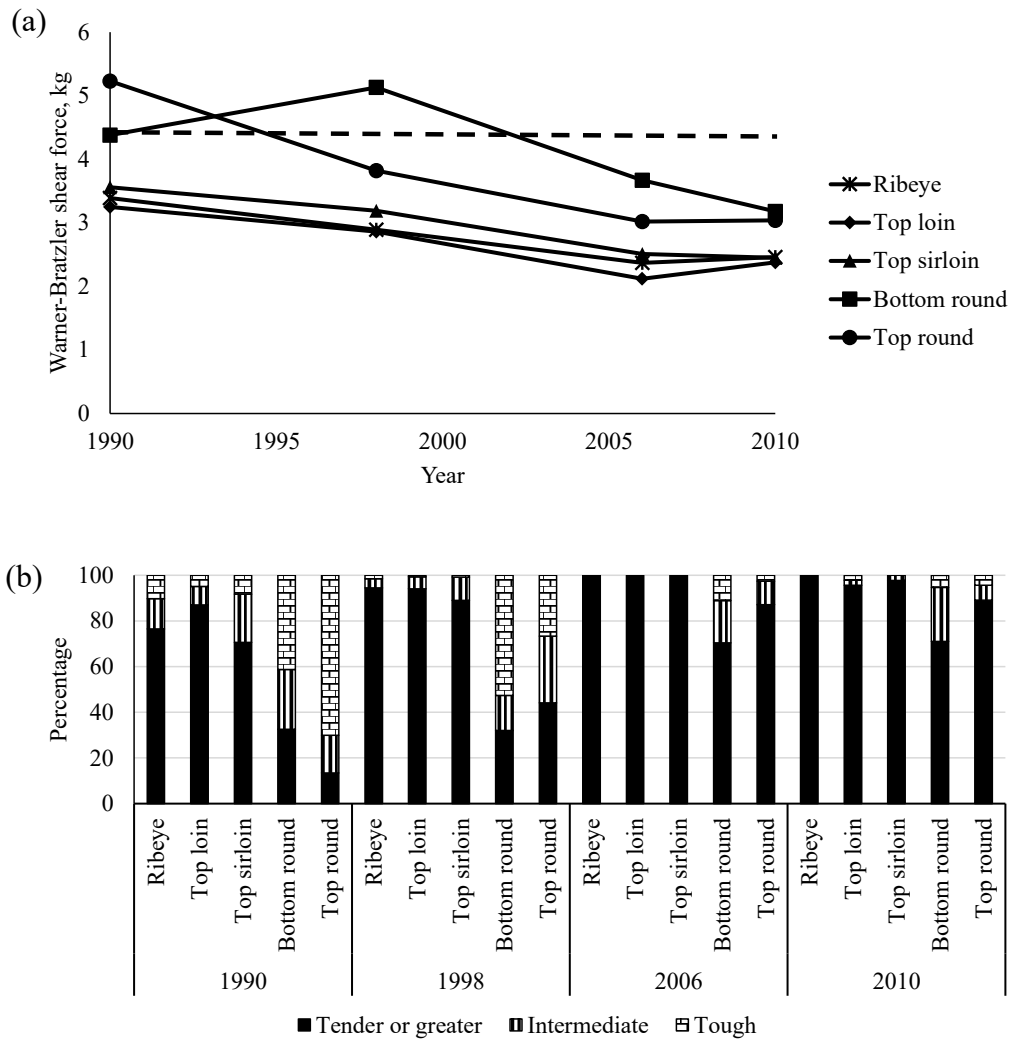
In the 1998 second NBTS, researchers narrowed their sample collection to retail outlets in 8 major metropolitan cities, but also examined steaks of the rib and loin available in the foodservice market [11]. Vast differences were noted in the aging period for sub-primals when comparing retail (19 d) and foodservice (32 d), although mean postmortem aging times for retail cuts for both groups were greater than the minimum 14 d believed to be needed to maximize tenderness [17]. Overall, improvements were observed in the tenderness of beef steaks, which the authors attributed to an increased focus on postmortem aging and longer chilling; however, the authors identified a need to continue to improve processing and preparation methods associated with retail cuts from the round to improve tenderness. The 2006 study conducted by [12] found continued improvement in overall steak tenderness, which was attributed to continued increases in postmortem aging times, slower carcass chill rates, and the prevalence of retail and packer programs that attempted to control the major contributors of tenderness. The most recent peer-reviewed publication involving the 2010 National Beef Tenderness Survey, found most steaks continued to be rated tender by objective and subjective measures; however, not all cuts decreased in tenderness, which led the authors to conclude a tenderness plateau may have been reached. The authors also concluded the increase in retail and packer programs focusing

on tenderness from the 2006 survey were an on-going contributor to the maintenance of acceptable tenderness levels across cuts [13].

It is very difficult to earnestly compare the four National Beef Tenderness Surveys discussed above because of minor method adjustments by the research teams including the number of cities sampled, the cities sampled, cooking methods, and types of sensory panels utilized. Additionally, the focus of the surveys were not consistent which makes comparison difficult; however, this portion of the review aims to follow muscles of the rib, loin, and round during the 20 years of surveys, as shown in Figure 4a. Over time, the beef industry and USDA have established thresholds for tenderness, with a WBSF value less than 4.4 kg being used today for products to enter a "Guaranteed Tender" program [18]. Surveyed cuts from the loin were never tough according to this threshold, but over time the ribeye, top loin, and top sirloin steaks improved in tenderness by 31%, 27%, and 27% respectively. Cuts of the round were initially considered tough and in need of tenderness improvement. It took until 2006 for the top round and bottom round to both reach the "Guaranteed Tender" threshold and during that time both muscles improved in tenderness, by 42 and 27% respectively. Aside from overall WBSF values, consistency of tenderness has been a major issue for the industry. Improvements in tenderness of these same cuts discussed above can be seen by the shift in the percentage of cuts falling into "Tender" ( $\leq 3.9$  kg), "Intermediate" (3.9 to 4.5 kg), and "Tough" ( $> 4.5$  kg) tenderness categories as measured by WBSF (Figure 4b). In 1990, an average of 22% of rib and loin cuts had WBSF values outside the "Tender" category, while 77% of round cuts were outside this category. In the subsequent surveys, the most noticeable improvement in reduced tenderness variation was in the rib and loin cuts, with 2006 being the highlight when 100% of steaks tested had WBSF in the "Tender" category. Progress in the round cuts was much slower, but vast improvement in variations in tenderness have occurred with 11% of top round and 29% of bottom round steaks not having WBSF in the "Tender" category. However, as stated earlier, a tenderness threshold may have been met, and tenderness of the round muscles may not improve further due to the physiology of these muscles of locomotion. Overall, the National Beef Tenderness Surveys demonstrate that beef industry has made quantifiable improvements in tenderness consistency.

## CONCLUSION

It is unknown what level of knowledge Asian beef consumers as a whole have on the US quality grading system. It is assumed Australia has the greatest knowledge given their current system is based off the US system. Ngapo et al [19] reported consumers in more developed Asian countries such as Japan,



**Figure 4.** Rib, loin, and round objective tenderness values and categorization over the 20 years covered by the National Beef Tenderness Surveys. (a) Warner-Bratzler shear force (WBSF) values of steaks. Blue dashed line represents a WBSF value of 4.4 which is the threshold at which tested beef products can enter guaranteed tender beef programs [18] (b) Percentage of steaks grouped into tender or greater (WBSF≤3.9), intermediate (3.8 to 4.5), and tough (>4.5 kg) tenderness categories. Adapted from Morgan et al [15], Brooks et al [11], Voges et al [12], and Guelker et al [13].

Korea, and Taiwan prefer meat with increased levels of marbling. Additionally, Frank et al [20] found the importance of flavor and juiciness in the consumers’ eating experience vary by country within Asia. In the US, more focus has been placed on improving these two important palatability influencers because tenderness has become less variable and improved as demonstrated above. Therefore, there is a great opportunity for US beef producers to improve the flavor and juiciness of their products to satisfy not only their consumers, but the rapidly growing Asian market.

**CONFLICT OF INTEREST**

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

**ACKNOWLEDGMENTS**

Contribution no. 18-171-J of the Kansas Agricultural Experiment Station, Manhattan 66506, USA.

**REFERENCES**

1. Aberle ED, Forrest JC, Gerrard DE, Mills EW. Principles of meat science. 4th ed. Dubuque, IA, USA: Kendall/Hunt Publishing Co; 2003.
2. USDA. Grades of fed beef carcasses. Agricultural Marketing Service. Marketing Res. Rep. Washington, DC, USA: United States Department of Agriculture; 1974. Report No.: 1073.
3. Lorenzen CL, Hale DS, Griffin DB, et al. National Beef Quality Audit: survey of producer-related defects and carcass quality and quantity attributes. J Anim Sci 1993;71:1495-502.

4. Boleman SL, Boleman SJ, Morgan WW, et al. National Beef Quality Audit–1995: survey of producer-related defects and carcass quality and quantity attributes. *J Anim Sci* 1998;76:96-103.
5. McKenna DR, Roebert DL, Bates PK, et al. National Beef Quality Audit–2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *J Anim Sci* 2002;80:1212-22.
6. Garcia LG, Nicholson KL, Hoffman TW, et al. National Beef Quality Audit–2005: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *J Anim Sci* 2008;86:3533-43.
7. Moore MC, Gray GD, Hale DS, et al. National Beef Quality Audit–2011: in-plant survey of targeted carcass characteristics related to quality, quantity, value, and marketing of fed steers and heifers. *J Anim Sci* 2012;90:5143-51.
8. Savell JW, Branson RE, Cross HR, et al. National Consumer Beef Retail Beef Study: palatability evaluations of beef loin steaks that differed in marbling. *J Food Sci* 1987;52:517-9.
9. Savell JW, Cross HR, Francis JJ, et al. National Consumer Retail Survey: interaction of trim level, price and grade on consumer acceptance of beef steaks and roasts. *J Food Qual* 1989;12:251-74.
10. Neely TR, Lorenzen CL, Miller RK, et al. Beef customer satisfaction: role of cut, USDA quality grade, and city on in-home consumer ratings. *J Anim Sci* 1998;76:1027-32.
11. Brooks JC, Belew JB, Griffin DB, et al. National beef tenderness survey-1998. *J Anim Sci* 2000;78:1852-60.
12. Voges KL, Mason CL, Brooks JC, et al. National Beef Tenderness Survey – 2006: assessment of Warner-Bratzler shear and sensory panel ratings for beef from US retail and food service establishments. *Meat Sci* 2007;77:357-64.
13. Guelker MR, Haneklaus AN, Brooks JC, et al. National Beef Tenderness Survey – 2010: Warner-Bratzler shear force values and sensory panel ratings for beef steaks from United States and food service establishments. *J Anim Sci* 2013;91:1005-14.
14. USDA. Official United States standards for grades of carcass beef. Washington, DC, USA: Livestock Seed Program Agricultural Marketing Service; 1997.
15. Morgan JB, Savell JW, Hale DS, et al. National beef tenderness survey. *J Anim Sci* 1991;69:3274-83.
16. George MH, Tatum JD, Belk KE, Smith GC. An audit of retail beef loin steak tenderness conducted in eight U.S. cities. *J Anim Sci* 1999;77:1735-41.
17. Lorenzen CL, Weatherly BH, Savell JW. Determination of an aging index. College Station, TX, USA: A final report to the Texas Beef Council, Austin, from the Meat Science Section, Department of Animal Science, Texas A&M University; 1998.
18. Standard specification for tenderness marketing claims associated with meat cuts derived from beef [Internet]. Washington, DC, USA: ASTM International; 2011 [cited 2017 Sept 16]. Available from: [https://compass.astm.org/EDIT/html\\_annot.cgi?F2925+11](https://compass.astm.org/EDIT/html_annot.cgi?F2925+11)
19. Ngapo TM, Martin JE, Dransfield E. International preferences for pork appearance: II. Factors influencing consumer choice. *Food Qual Prefer* 2007;18:139-51.
20. Frank D, Joo ST, Warner R. Consumer acceptability of intramuscular fat. *Korean J Food Sci Anim Resour* 2016;36:699-708.