Prediction of marbling score in ribeye quartered at the 5th- 6th rib of French beef using the Q-FOM™ beef assessment camera

Nathalia da Silva Rodrigues Mendes, Mette Christensen, Moïse Kombolo-Ngah, Pascal Faure, Laure Thoumy, Alix Neveu, Amanda Gobeti Barro, Jingjing Liu, Tatianne Ferreira de Oliveira, Marie-Pierre Ellies-Oury, et al.

Meat Science 222 (2025) 109759

Contents lists available at ScienceDirect

Meat Science

journal homepage: www.elsevier.com/locate/meatsci

Prediction of marbling score in ribeye quartered at the 5th- 6th rib of French beef using the Q-FOM™ beef assessment camera

Nathalia da Silva Rodrigues Mendes [a,b], Mette Christensen [c], Moïse Kombolo-Ngah [a], Pascal Faure [a], Laure Thoumy [d], Alix Neveu [e], Amanda Gobeti Barro [a], Jingjing Liu [a], Tatianne Ferreira de Oliveira [b], Marie-Pierre Ellies-Oury [a,f], Sghaier Chriki [a,g], Jean-François Hocquette [a,*]

[a] Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), Université Clermont Auvergne, VetAgro Sup, UMR1213, Recherches sur les Herbivores, Theix, 63122 Saint-Genès Champanelle, France

[b] School of Agronomy, Federal University of Goiás-UFG, Campus Samambaia, Rodovia Goiânia-Nova Veneza Km-0, Caixa Postal 131, CEP 74690-900 Goiânia, Brazil

[c] Frontmatec A/S, Hassellunden 9, Smoerum 2765, Denmark

[d] Institut de l'Elevage, Maison Régionale de l'Agriculture du Limousin, Boulevard des Arcades, 87060 Limoges, France

[e] IMR3GF, ul. Smulikowskiego 4/217, 00-389 Warszawa, Poland

[f] Bordeaux Sciences Agro, CS 40201, 33175 Gradignan, France

[g] ISARA, Agro School for Life, 23 rue Jean Baldassini, CEDEX 07, 69364 Lyon, France

A R T I C L E I N F O

Keywords:
Marbling
Beef eating quality
MSA
Non-destructive method
Camera

A B S T R A C T

In the Meat Standards Australia (MSA) and Guaranteed Global Grading (3G) grading schemes, beef marbling is scored visually in the chiller by accredited graders from 100 to 1190 marble score points in increments of 10. This study aimed to evaluate a hand-held camera (Q-FOM™ Beef) for determining MSA marbling scores of carcasses quartered between the 5th and 6th rib. The carcasses were scored by two accredited graders, including an expert grader (i.e. a more experienced grader). The R^2 of correlation between scores of the two graders for 377 carcasses was 0.78 with a RMSE of 47.9. The R^2 of correlation between the scores of the expert grader and the Q-FOM™ for 285 carcasses was 0.75 with a RMSE of 44.9. For the grader-to-grader comparison, 75.9 %, 97.1 % and 100 % of the values were within 50, 100 and 200 marbling points, respectively. For the comparison between Q-FOM™ predictions and the expert grader, 78.6 %, 96.8 % and 99.7 % of the values were within 50, 100 and 200 marbling points, respectively. Both between visual graders and the Q-FOM™ against expert grader showed acceptable accuracy performance and fulfilled the accreditation criteria defined by AUS-MEAT in Australia. Additionally, 124 Q-FOM™ images were assessed on-screen by the expert grader. The R^2 of correlation between the in-chiller and on-screen MSA marbling scores was 0.78 with a RMSEP of 48.7. Thus, on-screen assessing met requirements for accreditation, and both in-chiller and on-screen visual assessments of MSA marbling score are acceptable inputs for developing a Q-FOM™ Beef marbling model.

1. Introduction

Intramuscular fat (IMF%) is found in muscle tissue between skeletal muscle fibers. Marbling is the visible white flecks and streaks of IMF (Hocquette et al., 2010; Lee & Choi, 2019; Lee, Yoon, & Choi, 2019). Both IMF% and marbling score have a significant effect on a number of beef quality traits such as juiciness, color, tenderness, and taste (Chen, Li, Du, & Cao, 2019; Hocquette et al., 2010; Stewart, Gardner, et al., 2021). Both Dikeman (1996) and Hocquette et al. (2011) found that variation of 10 to 15 % of tenderness evaluation, and 2 to 56 % of variation in flavor, can be accounted for by IMF%, respectively. IMF% and marbling score varies between species, between breeds, and even more between muscle types in the same animal (Chriki et al., 2013). Other factors affecting IMF% and marbling score include gender, age, and feeding regimen (Chambaz, Scheeder, Kreuzer, & Dufey, 2003; Chriki et al., 2013; Hocquette et al., 2011; Thompson, 2004).

Approximately 75 % of the variation in IMF% can be explained by variation in visual scores by skilled assessors (reviewed by Ferguson,

* Corresponding author.
E-mail address: jean-francois.hocquette@inra.fr (J.-F. Hocquette).

https://doi.org/10.1016/j.meatsci.2025.109759
Received 30 September 2024; Received in revised form 10 December 2024; Accepted 16 January 2025
Available online 22 January 2025

2004). In addition, Stewart, Gardner, et al. (2021), showed that visual marbling scores and data from chemical IMF% analyses give similar results ($R^2 = 0.32$ and $R^2 = 0.28$ respectively) in predicting beef eating quality. According to Stewart, Toft, et al. (2024) whether direct MSA marbling assessment or IMF% derived models for MSA marbling are used, the results are largely the same. The shortcomings of chemical methods for determining intramuscular fat percentage (IMF%) include that they are time-consuming, labor-intensive, and destructive to samples. These limitations hinder their practical application in an industrial environment. To address these issues and to reliably predict beef eating quality in an industrial setting, grading systems relying upon carefully trained and accredited graders were developed internationally and especially in Europe (reviewed by Hocquette et al., 2020).

However, in the EU, beef carcasses are evaluated using the EUROP grid, which describes the carcass conformation and fatness scores, but this grid does not include marbling indicators despite previous recommendations (Monteils et al., 2017). Studies in France (Bonny et al., 2018; Legrand, Hocquette, Polkinghorne, & Pethick, 2013; Liu et al., 2021) and in other European countries (Farmer & Farrell, 2018; Pogorzelski, Woźniak, Polkinghorne, Półtorak, & Wierzbicka, 2020 have shown the importance of marbling for beef eating quality. Thus, the French National Food Conference recently recommended that the meat sector, represented by INTERBEV, introduce marbling into the French beef grading system (Etats Generaux de l"Alimentation EGA, 2018).

The MSA was developed in the 1990s to assess beef palatability under commercial conditions (Watson, Polkinghorne, & Thompson, 2008). Since then, it has come to be widely recognized as an innovative method for predicting beef eating quality (review by Bonny et al., 2018; Mendes, Briceno et al., 2024). The United Nations Economic Commission for Europe has recommended following the guidelines of the Guaranteed Global Grading (3G) protocols (Guaranteed Global Grading, 2023) which are similar to the MSA ones, since both grading systems (MSA and 3G) are based on the same control points and protocols, of which one of the most important is marbling (Hocquette et al., 2020).

In both the MSA and 3G protocols, marbling is evaluated in the chiller by accredited graders, who make their assessment at one quartering site between the 5th to 13th rib depending on the country (Meat, Livestock Australia, Meat Standards Australia, 2024). Historically, two marbling scores were designed: 1) the AUS-MEAT marbling score on a scale of 0 (no visual marbling) to 9 (extensive visual marbling) which provides an indication of the amount of marbling in beef and 2) the MSA marbling score on a scale from 100 to 1190 in increments of 10, which provides indications of amount and distribution of marbling. MSA accreditation is expensive due to the initial training, and the additional requirement of regularly renewing certification every six months adds to this expense. In addition, the visual assessment of the ribeye in this way varies from grader to grader, and even assessments by the same grader can vary over the course of a workday (Stewart, Toft, et al., 2024). For these reasons, there is a need for more consistent, reliable, and cost-effective methods for assessing carcasses for marbling instead of MSA-accredited graders.

Implementations of several technologies have been studied to address this problem of quantitatively and qualitatively predicting carcass composition and meat quality (Mendes, Briceno et al., 2024; Sanchez, Arogancia, Boyles, Pontillo, & Ali, 2022). These include Hyperspectral Imaging (HSI) (Stewart, Lauridsen, et al., 2021) which serves as the foundational technology for the Q-FOM™ camera (Drachmann et al., 2024; Stewart, Lauridsen, et al., 2021). Unlike traditional HSI systems, the Q-FOM™ integrates specific algorithms for beef quality assessment, setting it apart as a unique tool among current sensing and imaging methods (Drachmann et al., 2024). Additional methods include Magnetic Resonance Imaging (Lee et al., 2015), X-ray Computed Tomography (Anderson, Cook, Williams, & Gardner, 2018), Computerized Tomography Imaging (Prieto et al., 2010), Near-infrared Spectroscopy (Kombolo-Ngah et al., 2023) including smartphone NIRS sensor (Coombs, Fajardo, & González, 2021), Raman spectroscopy (Cama-Moncunill et al., 2020) and ultrasound (Beriain et al., 2021). However, these techniques were mostly developed for laboratory use, with only a few studies evaluating their industrial implementation. An additional disadvantage of all these technologies is the need for a large number of reference carcasses that have been evaluated according to the grading system of the country in which the device(s) is to be used. Algorithms must be developed for each implementation, as well, in order to be accurately calibrated (Cheng, Cheng, Sun, & Pu, 2015).

The most promising avenue of research is the development of a portable handheld device for use in slaughterhouses as a replacement for certified graders or as a toolkit to assist certified graders. The Q-FOM™ Beef handheld camera developed by Frontmatec Smoerum A/S is certified by MSA/AUS-MEAT for use in Australia in carcasses quartered caudal to the 10th to 13th rib (AUS-MEAT, 2023). In addition, the Q-FOM™ camera stands out for its incorporation of three-dimensional sensors and advanced image segmentation algorithms that enhance its precision and repeatability across diverse carcass phenotypes. Unlike conventional HSI systems, which often face challenges related to carcass presentation and operator bias, the Q-FOM™ utilizes a real-time guidance system to standardize image acquisition, thereby minimizing the impact of environmental and operational variability (Stewart, Toft, et al., 2024). This advancement makes the Q-FOM™ not merely an HSI-based device, but an integrated tool tailored for robust and objective beef quality assessment in commercial settings. Marbling, IMF%, Eye Muscle Area, Meat Color, Fat Color, and Rib Fat Thickness are all measured within a few seconds after capturing an image of the ribeye cut surface in the chiller, allowing for adequate blooming. The Q-FOM™ system is characterized by high precision and accuracy in predicting continuous traits such as marbling score and/or chemical IMF% (Drachmann et al., 2024; Stewart, Toft, et al., 2024).

Consequently, this study evaluated the potential of the Q-FOM™ Beef camera to predict MSA marbling scores of French beef cattle in one slaughterhouse in France, especially the practice of quartering at the 5th–6th rib. Additionally, the feasibility of using on-screen images to perform MSA marbling measurements was also examined. The study further compared the MSA marbling and AUS-MEAT marbling scores provided by 3G-accredited graders.

2. Materials and methods

2.1. Experimental design

All the beef carcasses for this study were provided by a commercial slaughterhouse located in Limoges, France (Plainemaison Aquitaine-Beauvallet, Limoges), and were obtained from animals reared and slaughtered in France. A total of 420 carcasses were randomly selected for variability in terms of breed (54.5 % Limousin, 2.9 % Charolais, 11.9 % Aubrac, 19.3 % Primholstein, 4.3 % Montbéliarde, 0.2 % Salers, etc), category (4.3 % Bull, 77.9 % Cow, 11.4 % Heifer, 1.9 % Ox, 4.5 % Steers), age (from 10 to 220 months) and weight (from 101.6 to 730.6 kg). This ensured that the studied animals were representative of the commercial variability of the company's beef carcasses.

Animals were transported the day before slaughter and slaughtered early in the morning at the slaughterhouse in compliance with ethical guidelines for animal care. Transport time from the farm was about 1 h to 1.4 h and animals had free access to water until slaughter. The exsanguination from the jugular vein was performed after penetrating stunning using a captive-bolt pistol. Slaughtering was performed in compliance with French welfare and EU regulations (Council Regulation (EC) No. 1099/2009). The carcasses were dressed according to standard commercial practices and between 30 and 50 min post exsanguination the carcasses were split in half then chilled for 24 h at 2–4 °C.

These beef carcasses were hung by the *Achilles tendon* and were quartered between the 5th and 6th rib. Carcasses with cut surface damage and presence of residues (blood, fat, bone or meat) on the ribeye were discarded prior to evaluation of the marbling score either by the graders and by the Q-FOM™ camera for comparison of both measurements.

2.2. MSA marbling assessment

After 24 h of *post-mortem* chilling in a cold room maintained at an average temperature between 0 and 4 °C, the processed carcasses were assessed. Marbling assessments were performed in accordance with the ABCAS reference standards (Meat, Livestock Australia, Meat Standards Australia, 2024), following a benchmark established by the UNECE Bovine Language Standards, using MSA marbling scores. The marbling standards have been tailored to European cattle and consumers through extensive collaborative research in Europe, with data storage and use facilitated by the IMR3GF (International Meat Research 3G Foundation).

The MSA marbling score, which provides a detailed scale (ranging from 100 to 1190 in increments of 10), was used to indicate the amount, size, fineness and distribution of fat inclusions in beef carcasses. Evaluation of MSA marbling scores was performed according to the MSA methodology (AUS-MEAT, 2022a) described by many authors, for example, Liu et al. (2021) and Stewart, Toft, et al. (2024). However, in this specific study, evaluation of MSA marbling scores was performed on the 5th rib of the carcass (Liu et al., 2021).

Two 3G accredited graders, including one expert grader, performed the evaluations. Both graders have undergone uniform training and received consistent updates in accordance with the Australian Beef Chiller Assessment System (ABCAS) standards to minimize technical variability in their assessments. An expert grader, while also accredited, possesses extensive experience and is recognized for superior expertise and accuracy in marbling assessments. They are often used as benchmarks for comparisons and are recognized as experts for research projects. Thus, the expert grader is considered to provide the best reference data. The expert grader independently assessed all 420 carcasses in the experiment. Of these, 377 carcasses were also evaluated by both 3G graders, including the expert grader, providing a set of common evaluations for comparative analysis. For model calibration, 285 carcasses were selected based on the best segmentation quality of the ribeye area, ensuring the reliability of the data used for accurate model training. After cutting, the ribeye was exposed to air for a minimum of 20 min and up to 3 h to allow the meat to bloom prior to the MSA marbling assessment. Blooming refers to the color change resulting from oxygenation of myoglobin when the meat surface is exposed to oxygen (Jacob, 2020). Assessing was performed using standard visual cards provided by ABCAS to assess AUS-MEAT and MSA scores (AUS-MEAT, 2022a).

2.3. Image analysis for MSA marbling

Among the 420 carcasses included in the trial, a subsample consisting of 285 carcasses was visually assessed by the expert grader in the chiller. A total of 779 images (i.e. between 2 and 3 images per carcass) were acquired using the hand-held Q-FOM™ Beef camera. Prior to image acquisition, excessive bone dust and fat smears present on the cut surface were removed using either the blade of a knife, a traditional fabric commonly used in French abattoirs to cover carcasses or simply paper. These cleaning methods were selected for their effectiveness in removing surface contaminants while maintaining the presentation and quality of the ribeye area for imaging. The camera was positioned over the assessing surface using a real-time built-in camera operator guiding tool appearing on the screen during image capture.

Each morning, a self-test was performed by placing the camera in a self-test stand with a NIST traceable chessboard and triggering image capture. The self-test verifies that color and geometry measures established during calibration are within acceptable limits. In a first analysis with the 285 carcasses, a Q-FOM™ Beef calibration model using the expert grader MSA Marbling score as reference was developed. The construction of the Q-FOM™ Beef calibration model was a multi-step process involving specific hardware configurations, experimental design, and advanced image processing techniques, as outlined in Stewart, Toft, et al. (2024).

Two months later, a PDF file containing the Q-FOM™ images from 100 of those carcasses was presented to the expert grader. The carcasses were assessed again by evaluating the marbling score of the captured images. Triplicate images of 12 different carcasses were included in the PDF file to determine the expert grader's repeatability via a blind test. The expert grader was not informed that triplicate images were included, and all images (a total of 124 pages) were presented in a randomized way using a consecutive ID number.

2.4. Statistical analysis

The datasets utilized in this study were sourced from three distinct origins: (a) a dataset from 377 carcasses with marbling scores provided by the expert grader and a 3G accredited grader. This dataset was used to compare the consistency of marbling scores between the two graders; (b) a subset consisting of 285 carcasses with marbling scores provided by the expert grader and images captured by the Q-FOM™ beef camera. This dataset was used to develop a MSA marbling score prediction model for the Q-FOM™ beef camera and (c) additionally 124 images were assessed for marbling score by the expert grader on-screen. This dataset was used to compare the consistency of marbling scores between on-site in-chiller and on-screen assessing by the expert grader only.

2.4.1. Model assessment

The dataset was analyzed using R software (version 4.3.0 - R Core Team, 2023). The assessment of model precision and goodness of fit on test datasets between the Q-FOM™ Beef camera predicted data and the expert grader MSA Marbling scores was conducted using coefficients of determination (R^2), root mean squared error (RMSE), and bias, as described by Tedeschi (2006).

For the relationships between actual MSA marbling score of the independent validation, the R^2 and root mean squared error of the prediction (RMSEP) were used to report the precision, whereas the slope of the relationship between actual and predicted values and bias were used to report accuracy. Bias is defined as the difference between the actual and the predicted value at the mean of the predicted trait, as described by Stewart, Toft, et al. (2024).

In this study, we employed confusion matrices to assess agreement between evaluations performed by the expert grader and the secondary grader for AUS-MEAT marbling scores. Precision was calculated as the average proportion of scores assigned by the secondary grader that fell within a ± 1 class range of those assigned by the expert grader. Accuracy was measured as the proportion of carcasses for which the secondary grader's scores matched exactly with the median scores assigned by the expert grader. Furthermore, bias was analyzed by comparing the differences between the expert grader's scores and those predicted by the secondary grader across all class levels (James, Witten, Hastie, & Tibshirani, 2013; Kuhn & Johnson, 2013).

2.4.2. Discriminant analysis

The marbling scores in the different datasets were classified into three homogeneous MSA marbling classes: Low, Medium, and High. The thresholds for these classes were: Low (≤ 290), Medium ($290 <$ scores ≤ 360), and High (> 360). These classes are derived from the 25th to the 75th percentile, with whiskers indicating minimum and maximum values. The use of these divisions is justified by the need to more accurately reflect the intrinsic variability within the dataset, ensuring that subsequent analyses are based on a stratification that is representative of the observed data. This approach allows for a categorization that is more aligned with the actual distribution of MSA Marbling scores in the present study, thereby ensuring that the statistical analyses capture the differences in sensitivity within each identified class with greater precision.

The confusion matrix, which describes the classification performance of the discrimination model, was evaluated based on three metrics: sensitivity, specificity, and overall accuracy for each class.

$$Sensitivity\ (\%) = \frac{TP}{TP + FN} \times 100$$

$$Specificity\ (\%) = \frac{TN}{FP + TN} \times 100$$

$$Accuracy\ (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

where FP, FN, TP, and TN represent the number of false positives, false negatives, true positives, and true negatives predicted during external validation, respectively. Sensitivity and specificity are used to calculate the true rates: sensitivity represents the model's ability to correctly assign the sample to the actual class it belongs to, whereas specificity indicates the model's ability to correctly identify the class to which the sample does not belong (Almeida, Fidelis, Barata, & Poppi, 2013). Accuracy is the overall proportion of correctly classified samples.

The overall accuracy of the confusion matrix was also evaluated using receiver operating characteristic (ROC) curves, where AUC-ROC (Area Under the ROC Curve) corresponds to the area under a ROC curve and provides a single value that measures the overall performance of a binary classifier (Hanley & McNeil, 1982). The AUC ranges from 0.5 to 1, where the lowest value represents a random classifier, and the highest value represents a perfect classifier (Hanley & McNeil, 1982). Three AUC-ROC curves were measured to evaluate the model's ability to discriminate between the different classes, and this was performed for both models.

3. Results

3.1. Variability in carcass data

The descriptive statistics (mean, standard deviation (SD), coefficient of variation (CV), minimum, and maximum) of animal and carcass traits along with MSA traits measured by certified graders and predicted by the Q-FOM™ beef camera between the 5th and 6th rib are shown in Table 1.

Table 1
Descriptive statistics for the cattle population.

	Mean	SD	CV (%)	Minimum	Maximum
Animal maturity (months)	74	52.93	71.76	10	220
Cold Carcass Weight (kg)	372.0	95.30	25.62	101.6	730.6
EU Conformation score[1]	6.84	3.58	52.3	1	14
EU Fat score[2]	3.87	3.30	85.60	1	9
MSA Marbling score (Visual)	313.08	91.29	29.16	100	650
MSA Marbling score (Q-FOM™)[3]	313.78	77.92	24.83	111	486.3
AUS-MEAT Marbling score (Visual)	0.95	0.83	87.71	0	4

[1] Following Hickey et al. (2007), European conformation scores were converted from P (− /=/+), O (− /=/+), R (− /=/+), U (− /=/+), and E (− /=/+) to classes from 1 (P-) to 15 (E+).

[2] Following Hickey et al. (2007), European fat scores were converted from 1 (− /=/+), 2 (− /=/+), 3 (− /=/+), 4 (− /=/+), and 5 (− /=/+) to classes from 1 (1-) to 15 (5+).

[3] Predictions of MSA marbling scores for 285 carcasses by Q-FOM™ Beef's current algorithms.

3.2. Comparison of MSA marbling scores between visual assessments performed by an expert grader and the 3G accredited grader

Using 377 carcasses, the R^2 of prediction between scores from the expert grader and the 3G accredited grader was 0.78 with a RMSE of 47.9 MSA marbling points (Fig. 1).

To be accredited for MSA marbling visual assessing in Australia, ≥49 % of the samples must be within 50 MSA-MS marbling points from the score assigned by an expert grader, ≥79 % of the samples must be within 100 MSA marbling points and ≥ 97 % must be within 200 MSA Marbling points. Comparing the 3G accredited grader's assessment of MSA marbling scores to the expert grader scores revealed that the accreditation requirements were met, since 75.9, 97.1 and 100.0, respectively, were within the requirements shown above (Fig. 2).

When carcasses are grouped in three classes based on MSA marbling score, the highest sensitivity was achieved for samples correctly predicted to have low MSA scores (≤ 290), with a sensitivity of 83.3 % in the analyzed dataset. Lower sensitivities were observed for samples in the medium (290 < scores ≤ 360) and high (> 360) MSA classes, with sensitivities of 53.3 % and 79.6 %, respectively. On the other hand, specificity ranged from 82.2 % to 93.3 % for all the classes. The confusion matrix for the dataset had an overall accuracy of 73.7 % (Table 2).

In Fig. 3, the ROC curves demonstrate an AUC of 0.840 for MSA marbling scores provided by the expert grader compared to the other 3G accredited grader. The ROC curves for the different classes show high sensitivity at almost all points, suggesting that the 3G accredited grader has a high probability of correctly classifying marbling scores when compared to the expert grader. More generally, we observed a good discriminatory ability of the graders in distinguishing between different marbling classes.

The interpretation of the AUC-ROC curves was guided by the criteria outlined by Yang and Berdine (2017). According to their study, an AUC value of 0.5 indicates no discriminatory power, values between 0.5 and 0.6 indicate poor discrimination, values between 0.6 and 0.7 represent acceptable discrimination, values between 0.7 and 0.8 indicate excellent

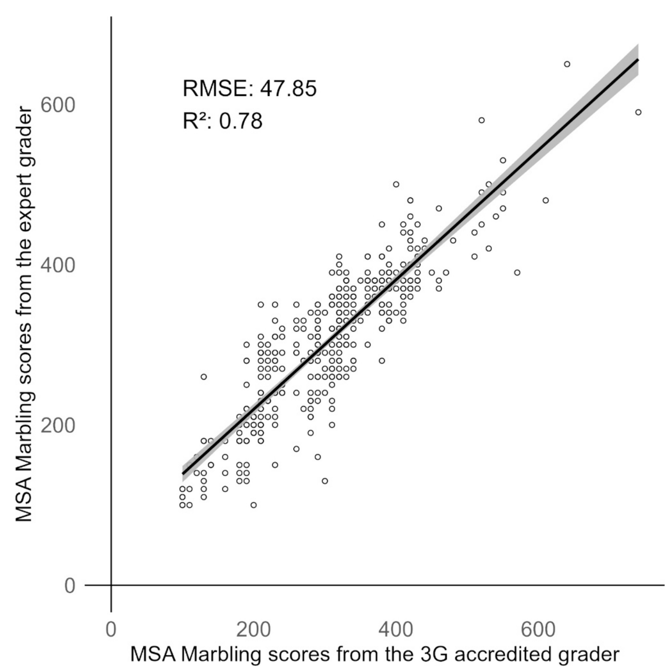

Fig. 1. Relationship between marbling scores measured by the expert grader (Y axis) and the 3G accredited grader (X axis) for 377 carcasses. The regression line is shown with a shaded 95 % confidence interval, showing the precision of the relationship between the two graders.

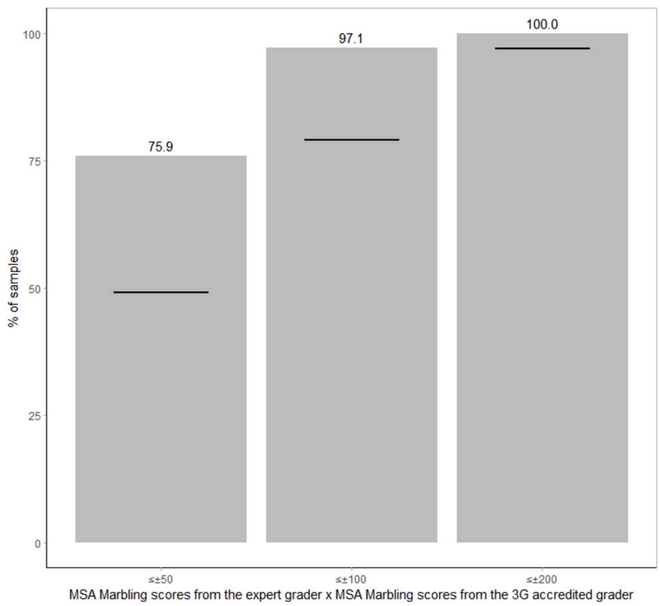

Fig. 2. Histograms showing the percentage of samples (grey bars) assigned a specific MSA marbling score by the expert grader and the 3G accredited grader, within the acceptance thresholds indicated on the x-axes. The black lines represent the AMILSC approved minimum requirements of accuracy standards for MSA marbling assessment (N = 377 carcasses) (AUS-MEAT, 2022b).

Table 2
Confusion matrix for MSA marbling classes using a dataset containing MSA marbling scores provided by the expert grader and a 3G accredited grader, comprising 377 carcasses.

MSA class	Low marbling score (MSA ≤ 290)	Medium marbling score (290 < MSA ≤ 360)	High marbling score (MSA > 360)	Total
Low predicted	135	33	0	168
Medium predicted	26	57	22	105
High predicted	1	17	86	104
Model performance, %				
Sensitivity	83.3	53.3	79.6	
Specificity	84.7	82.2	93.3	
Overall Accuracy				73.7

discrimination, and values above 0.9 signify outstanding discrimination. Based on these thresholds, our model demonstrated the ability to predict both the low and high classes effectively, as the AUC values exceeded 0.8 for our models.

The Low MSA Marbling Class (≤ 290) achieved an AUC of 0.840, indicating a high degree of consistency between the evaluations of the expert grader and the 3G accredited grader for carcasses with lower marbling. The Medium MSA Marbling Class (290 < scores ≤ 360) recorded an AUC of 0.677, suggesting greater variability in the consistency of evaluations for carcasses with medium marbling between the two graders. The High MSA Marbling Class (> 360) attained the highest AUC of 0.865, reflecting strong concordance between the expert grader and the 3G accredited grader for carcasses with MSA Marbling scores exceeding 360. This highlights the second grader's ability to closely correspond their assessments with those of the expert grader for high-quality carcasses.

Fig. 4 displays a confusion matrix heatmap of AUS-MEAT marbling

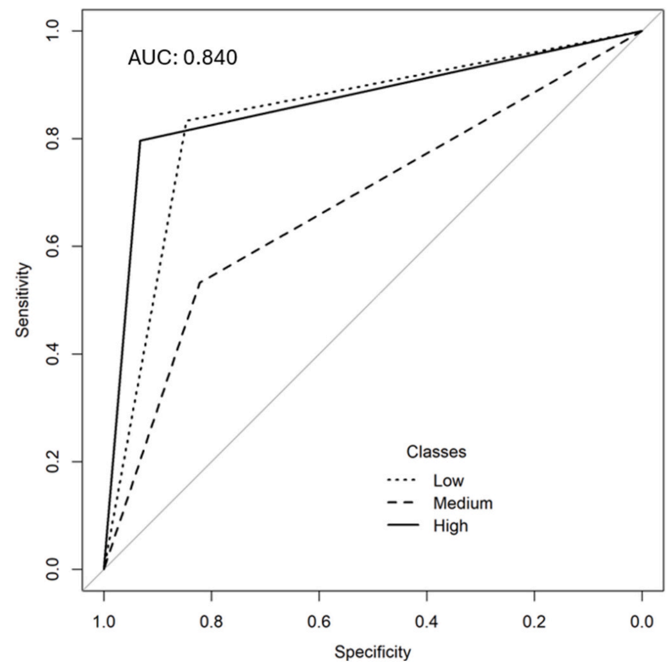

Fig. 3. ROC Curves for MSA marbling scores provided by the expert grader compared to the 3G accredited grader for 377 carcasses. The dotted, dashed, and solid lines represent the ROC curves for the Low (≤ 290), Medium (290 < scores ≤ 360), and High (scores >360) marbling classes, respectively. The diagonal line represents predictions no better than random guessing, and the closer a curve is to the upper-left corner, the better the model's predictive ability. The AUC value of 0.840 indicates excellent discrimination for the model.

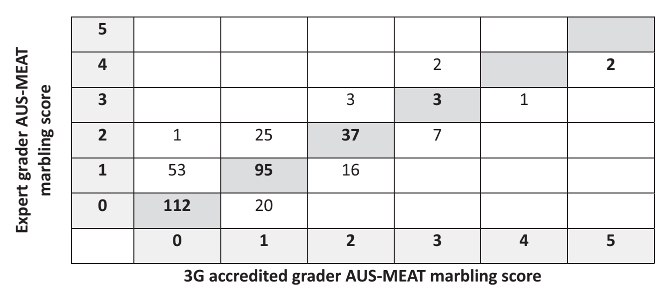

Fig. 4. Confusion matrix heatmap for AUS-MEAT marbling scores measured by the expert grader and the 3G accredited grader for 377 carcasses. 65.5 % of the carcasses (diagonal) were classified similarly by the expert grader and the 3G accredited grader. Allowing ±1 class deviation between the expert grader and the 3G accredited grader resulted in 99.7 % agreement.

scores. Most carcasses were correctly classified or fell within ±1 of the score assigned by the 3G accredited grader. This validates the reliability of the AUS-MEAT marbling assessments conducted by the 3G accredited graders.

3.3. *Q-FOM™ beef MSA marbling score performance*

A Q-FOM™ Beef calibration model using the expert grader MSA marbling scores (MSA-MB) as reference was developed on a dataset including 285 carcasses (Fig. 5).

For the comparison between predictions from the Q-FOM™ Beef camera and data from the expert grader, 78.6 %, 96.8 %, and 99.6 % of the values were within 50, 100, and 200 marbling points, respectively (Fig. 6). This shows that the Q-FOM™ Beef camera can predict MSA marbling scores with good accuracy and precision in carcasses quartered

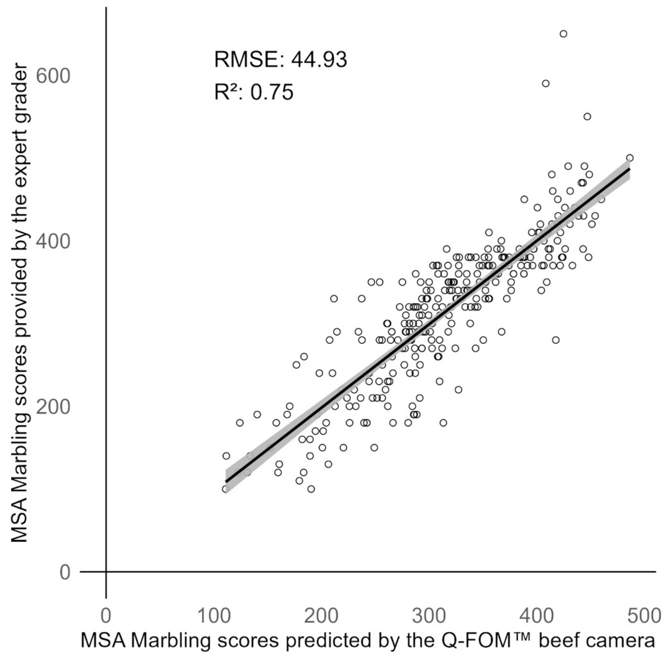

Fig. 5. Relationship between MSA marbling scores predicted by the Q-FOM™ beef camera and the expert grader (N = 285 carcasses). The regression line includes a shaded 95 % confidence interval, showing the variability and precision of the predictions relative to the expert assessments.

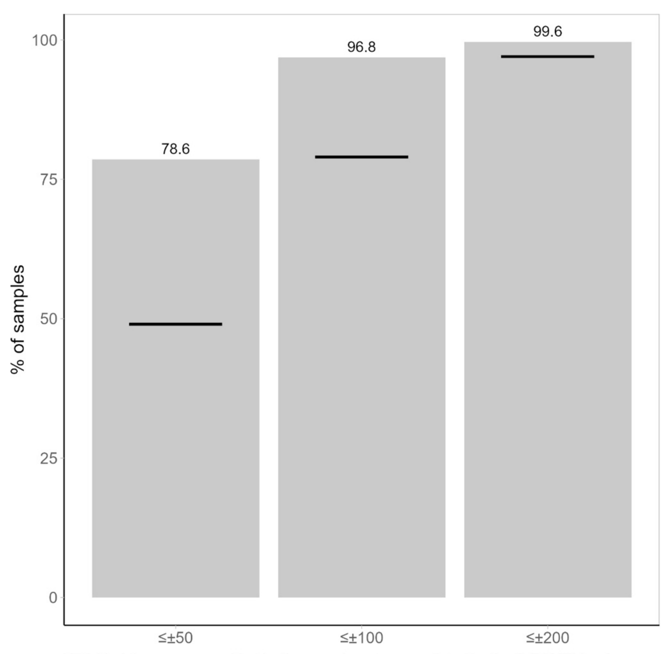

Fig. 6. Histograms showing the percentage of samples (grey bars) assigned a specific MSA marbling score by the expert grader and the Q-FOM™ Beef camera, within the acceptance thresholds indicated on the x-axes. The black lines represent the AMILSC approved minimum requirements of accuracy standards for MSA marbling assessment (N = 285 carcasses) (AUS-MEAT, 2022b).

at the 5th to 6th rib. These results are consistent with previous studies, who demonstrated the accuracy of the Q-FOM™ Beef camera in assessing ribeye traits (including marbling) in beef carcasses quartered at the 10th to 13th (Stewart, Toft, et al., 2024).

The Q-FOM™ MSA Marbling calibration model maintained a high level of precision and accuracy (Figs. 5 and 6). The Q-FOM™ calibration and validation models for MSA marbling also demonstrated high precision and accuracy, as reported by Stewart, Toft, et al. (2024). The Q-FOM™ calibration model tended to score marbling lower than the expert MSA grader when MSA marbling scores were high (+500). This may result from biases in the data used for calibration, or it may be from limitations in algorithms for image analysis. Further analysis and studies are needed to determine the importance of this phenomenon for high marbled carcasses. Contrary to our findings, Stewart, Toft, et al. (2024) reported an increase in bias with higher marbling scores, where the Q-FOM™ camera tended to assign higher marbling scores than the mean scores provided by MSA graders. However, it is important to note that the range of marbling scores differed significantly between the studies, with a broader range observed in the study by Stewart, Toft, et al. (2024), which reported scores ranging from 120 to 1160 compared to our study's range from 100 to 650 reflecting the phenotypic characteristics of European breeds. Despite these differences in range, the consistent upward bias with higher marbling scores underscores the need for further optimization ensuring that the Q-FOM™ system accurately predicts carcasses with extreme marbling levels.

When the carcasses were divided into three groups using MSA marbling scores, our analysis showed that the highest sensitivity was achieved for samples predicted to have low MSA scores, reaching 78.7 % in the dataset. Conversely, sensitivity was lower for samples in the medium (65.1 %) and high (78.0 %) MSA classes. Specificity, however, ranged from 79.4 % to 92.3 % across all classes. The confusion matrix for the dataset demonstrated an overall accuracy of 74.4 % (Table 3).

The ROC curves show an AUC of 0.845. This value is slightly higher than that obtained in the comparison with the second grader, indicating that the Q-FOM™ beef camera has a slightly better or comparable discrimination capability compared to the human second grader. Additionally, the ROC curves suggest that the Q-FOM™ beef camera is able to classify the different marbling classes with a high level of accuracy (Fig. 7).

The Low MSA Marbling Class (≤ 290) with an AUC of 0.845 demonstrates the Q-FOM™ beef MSA marbling score model's power to accurately identify samples with low MSA Marbling scores with minimal false positives. The Medium MSA Marbling Class (290 < scores ≤ 360) recorded an AUC of 0.723, suggesting adequate discrimination but with a higher likelihood of classification errors, especially within the transition zone between the Low and High classes. The High MSA Marbling Class (> 360) achieved the highest AUC, with a value of 0.851 (Fig. 7). This outcome highlights the model's robustness in accurately classifying carcasses with MSA marbling scores exceeding 360, demonstrating excellent precision and a low rate of false negatives.

Table 3
Confusion matrix for MSA marbling classes with model performance using a dataset containing MSA marbling scores provided by the expert grader and scores predicted by the Q-FOM™ beef camera, comprising 285 carcasses.

MSA class	Low marbling score (MSA ≤ 290)	Medium marbling score (290 < MSA ≤ 360)	High marbling score (MSA > 360)	Total
Low predicted	85	17	0	102
Medium predicted	21	56	20	97
High predicted	2	13	71	86
Model performance, %				
Sensitivity	78.7	65.1	78.0	
Specificity	90.4	79.4	92.3	
Overall Accuracy				74.4

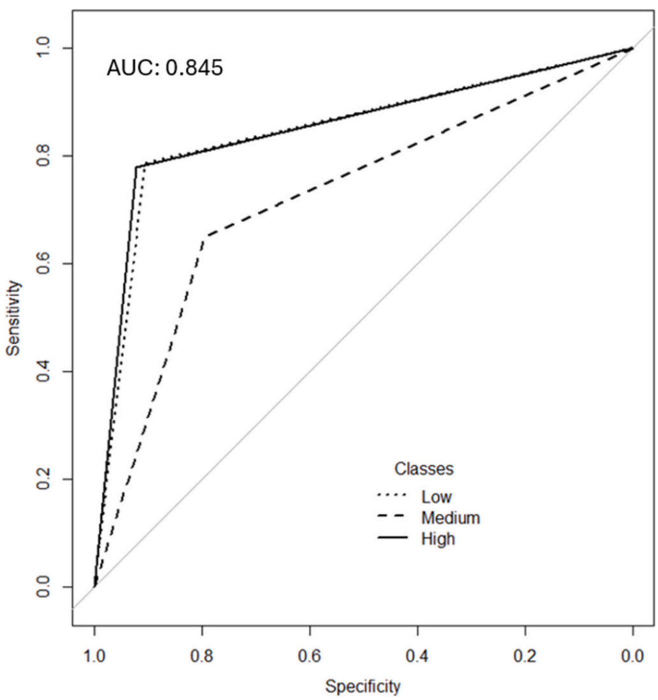

Fig. 7. ROC Curves for MSA Marbling Classes provided by the expert grader compared to those predicted by the Q-FOM™ beef camera for 285 carcasses. The dotted, dashed, and solid lines represent the ROC curves for the Low (≤ 290), Medium (290 < scores ≤ 360), and High (scores >360) marbling classes, respectively. The diagonal line represents predictions no better than random guessing, and the closer a curve is to the upper-left corner, the better the model's predictive ability. The AUC value of 0.845 indicates excellent discrimination for the model.

3.4. Comparison of in-chiller assessment to on-screen image assessing by the same expert grader

In another analysis, images of a subset ($N = 100$ carcasses, 124 images) of the 285 carcasses were assessed on-screen by the expert grader. The R^2 of prediction between the in-chiller and on-screen images was 0.78 with a RMSE of 48.7 (Fig. 8). The scatter plot in Fig. 8 highlights two notable outliers, which correspond to the carcasses shown in Fig. 9. The difference in scores between in-chiller assessment and on-screen image assessment was within 50 MSA-MS for 79.8 %, within 100 MSA-MS for 96.8 %, and within 200 MSA-MS for 100 % of the carcasses analyzed. This suggests that on-screen assessment of MSA marbling scores would be an acceptable method to develop a marbling calibration model, although we acknowledge that the dataset could be expanded to include more carcasses. (See Fig. 10.)

Our research, which utilized 100 unique carcasses and 12 of these carcasses with triplicate images, is one of the first to explore this approach with a specific focus on the variability of MSA Marbling score (100 to 650) at the 5th–6th ribbing site. While these findings are promising and provide a solid foundation, further investigation is required to confirm these results across a broader range of marbling scores and carcass types. Expanding the dataset to include a wider range of images and phenotypes would enhance the robustness of the calibration model, ensuring its applicability across different conditions.

The outliers shown in Fig. 9 revealed significant differences between the MSA marbling scores obtained from on-screen assessments and traditional in-chiller assessments. In Fig. 9a, the MSA marbling score was 480 in- chiller, but 610 when assessed on-screen. This discrepancy can be attributed to several factors, including errors during in-chiller evaluations due to time constraints, as well as typographical mistakes that can occur during assessment. Similarly, in Fig. 9b, the other sample

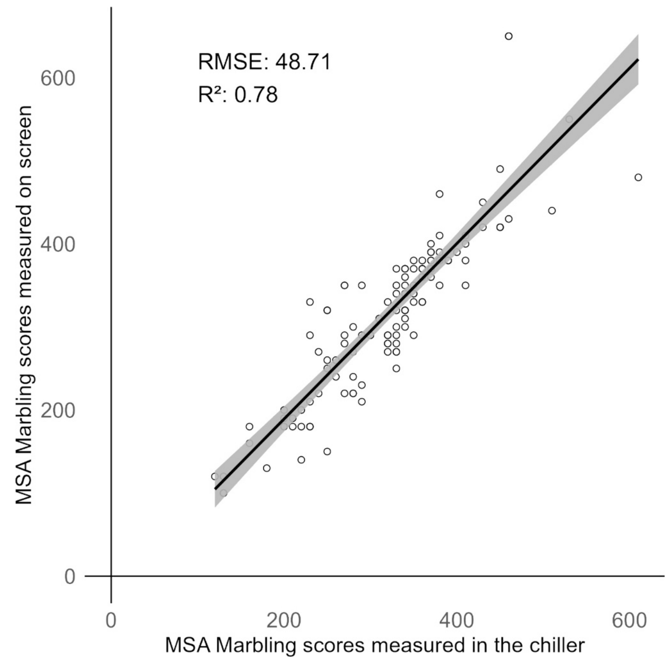

Fig. 8. Relationship between MSA Marbling scores measured in the chiller to on-screen image assessing by the same accredited expert grader (N = 100 carcasses, 124 images). The regression line with a shaded 95 % confidence interval shows the precision of the on-screen assessments compared to the chiller measurements.

displayed a marbling score of 650 with traditional assessment and 460 on screen.

When carcasses are divided into three classes by their MSA marbling score, the greatest sensitivity was achieved for samples correctly predicted to have high MSA scores, with 75.5 % in the analyzed dataset. Lower sensitivity was obtained for samples belonging to medium (68.6 %) and low MSA classes (77.5 %). On the other hand, specificity ranged from 76.4 % to 96.4 % for all classes. The confusion matrix for the dataset had an overall accuracy of 74.2 % (Table 4).

The ROC curves show an AUC of 0.824. This value is slightly lower than that obtained from the comparison of the traditional assessments by the expert grader with assessments using the Q-FOM™ beef camera. The ROC curves also suggest that on-screen assessments using images provided by the Q-FOM™ beef camera can be used to classify different marbling classes with a high degree of accuracy (Fig. 11).

The Low MSA Marbling Class (≤ 290) has an AUC of 0.82, indicating that the expert grader's on-screen assessment of Q-FOM™ acquired images can accurately identify samples with in-chiller low MSA Marbling scores. This suggests that the on-screen assessment is effective in correctly identifying carcasses with lower marbling scores, giving a low number of false positives. The Medium MSA Marbling Class (290 < Scores ≤360) showed an AUC of 0.72, implying good discrimination but with a higher tendency for classification errors. The High MSA Marbling Class (> 360) exhibited the highest AUC, with a value of 0.87 (Fig. 11). This shows the robustness of the model for correctly identifying carcasses with MSA marbling scores greater than 360, and consequently a low number of false negatives.

4. Discussion

4.1. Range and variability of marbling scores

The European conformation and fat scores in our study are consistent with Liu et al. (2021) and Drachmann et al. (2024). In addition, the average MSA marbling score obtained in our study was 313, which is consistent with the values of 293 to 329 reported by Liu et al. (2020) for

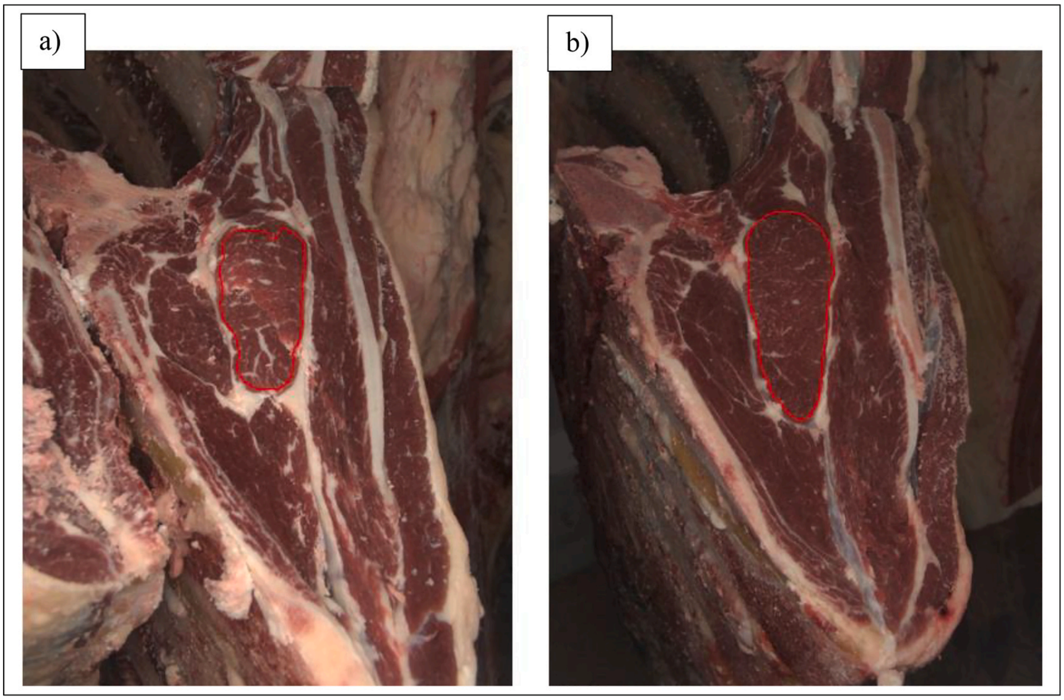

Fig. 9. Images captured using the Q-FOM™ camera of the *Longissimus thoracis* muscle surfaces from bovine carcasses identified as outliers in Fig. 8, representing cases with significant discrepancies between the marbling scores assigned by the expert grader in the chiller and those from on-screen assessments: (a) 480 MSA points (chiller) and 610 MSA points (on-screen image); (b) 650 MSA points (chiller) and 460 MSA points (on-screen image).

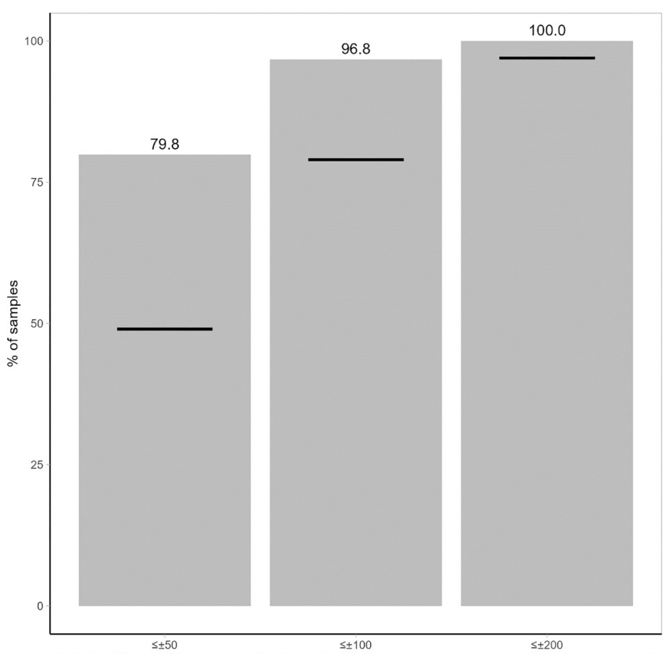

Fig. 10. Histograms showing the percentage of Q-FOM™ acquired images (grey bars) assigned a specific MSA marbling score by the expert grader in-chiller and on-screen, within the acceptance thresholds indicated on the x-axes. The black lines represent the AMILSC approved minimum requirements of accuracy standards for MSA marbling assessment ($N = 124$ images) (AUS-MEAT, 2022b).

a variety of European cattle and with the value of 310.8 reported by Mendes, Silva, et al. (2024) for Limousin cows, but it is slightly higher than that of 288 obtained by Liu et al. (2021) on the 5th rib from 208 bovines mainly French Limousine cows.

Table 4
Confusion matrix for MSA marbling classes with model performance using a dataset containing MSA marbling scores provided by the expert grader in the chiller and scores provided by the same grader using on-screen images ($N = 100$ carcasses, 124 images).

MSA class	Low marbling score (MSA ≤ 290)	Medium marbling score (290 < MSA ≤ 360)	High marbling score (MSA > 360)	Total
Low predicted	37	8	0	45
Medium predicted	12	24	9	45
High predicted	0	3	31	34
Model performance %				
Sensitivity	75.5	68.6	77.5	
Specificity	89.3	76.4	96.4	
Overall Accuracy				74.2

Our average MSA Marbling scores were also lower than those reported in Australian studies by Stewart, Toft, et al. (2024), who examined a diverse range of carcass phenotypes, and Stewart, Gardner, and Tarr (2024), who focused on Angus, Hereford, and *Bos indicus* cattle breeds, with scores of 495 and 445, respectively.

Additionally, the mean AUS-MEAT Marbling score in our study was 0.95, which is slightly higher than the score reported by Liu et al. (2021), but lower than the mean values observed by Santinello et al. (2024), and both Stewart, Toft, et al. (2024) and Stewart, Gardner, and Tarr (2024), who reported scores of 2.11 and 2.0, respectively.

The differences observed in our study compared to the Australian results are likely to be explained by differences in cattle breeds (Stewart, Gardner, and Tarr, 2024; Stewart, Toft, et al., 2024). Indeed, higher MSA Marbling scores are typically observed in early maturity breeds such as Angus and Hereford. By contrast, late-maturity European breeds, such as Limousin, Charolais, and Blonde d'Aquitaine, typically exhibit lower

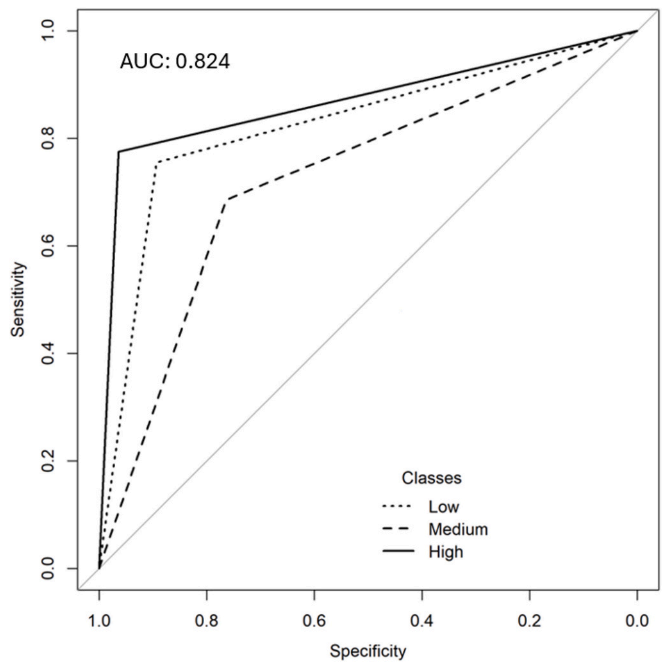

Fig. 11. ROC Curves for MSA Marbling measured on-screen image assessing by the expert grader ($N = 124$ images). The dotted, dashed, and solid lines represent the ROC curves for the Low (≤ 290), Medium ($290 <$ scores ≤ 360), and High (scores >360) marbling classes, respectively. The diagonal line represents predictions no better than random guessing, and the closer a curve is to the upper-left corner, the better the model's predictive ability. The AUC value of 0.824 indicates excellent discrimination for the model.

IMF content, even at similar maturity levels. These breeds face lower ability in depositing IMF% before intermuscular and subcutaneous fat, resulting in reduced marbling and, consequently, lower meat quality compared to early maturity breeds (Pethick, Harper, & Oddy, 2004). The characteristics of the production systems also play a critical role. European systems tend to focus on producing heavier carcasses from late-maturing beef breeds with less marbling despite using diets rich in maize or forages, which contrasts with the Australian systems where cattle from more early-maturing breeds are reared in feedlots using cereals in the finishing diet (Hocquette et al., 2018).

The differences observed in our study compared to those reported in Italian research can be partially attributed to the intensive rearing conditions on Italian feedlots, where young bulls are fed a diet supplemented with concentrates for six months prior to slaughter (Santinello et al., 2022; Santinello et al., 2024). In contrast, our rearing conditions are more extensive than the Italian ones.

The average marbling score and also its range of the studied carcass population are important for developing prediction models of marbling score. Consequently, there was a possibility that the calibration models of the Q-FOM™ camera initially designed for more marbled beef were not adapted to the European cattle due to lower and less variable marbling levels. Our results showed it is not the case since human graders can provide reliable data for marbling assessment. Indeed, despite that the MSA Marbling scores in our study ranged from 100 to 650, compared to the broader ranges (120 to 1160 and 160 to 1190 respectively) observed in the studies by Stewart, Gardner, and Tarr (2024) and Stewart, Toft, et al. (2024), our results demonstrated the capability to develop a robust model to predict marbling score with high accuracy. The broader range of MSA Marbling scores in these Australian studies may have contributed to their ability to create models using devices with enhanced predictive accuracy due to the greater variability in the data. However, our study also achieved strong predictive performance, suggesting that even with a more constrained range of marbling scores, the model can effectively capture the relevant factors influencing marbling and produce reliable predictions.

While the variability for MSA marbling between scores from accredited human graders is low enough, the presence of bias between the Q-FOM™ Beef camera calibration model and the expert grader underscores a critical issue associated with using human-derived scores for calibrating and validating measurement technologies (Jang, Ishdorj, Anderson, Purevjav, & Dahlke, 2017; Stewart, Toft, et al., 2024). As new technologies develop, there is a risk that new devices might not be calibrated and/or accredited using compatible reference data. This issue is particularly significant because it hampers the comparison of the performance of technologies which were calibrated with data from different graders (Stewart, Toft, et al., 2024). In our study, to minimize this problem, we have developed a calibration model based on the expert grader reference data. More generally, to mitigate this problem, the Australian Meat Industry Language and Standards Committee has recently endorsed chemical IMF% as an industry trait, highlighting the necessity of objective gold standard traits for validating technologies (AUS-MEAT, 2022c).

4.2. Prediction of marbling score and IMF%

Previous studies by Stewart, Toft, et al. (2024) and Drachmann et al. (2024) established the efficacy of the Q-FOM™ system in predicting IMF %. Further, Stewart, Gardner, and Tarr (2024) developed and validated models to use chemical IMF% to find equivalent MSA marbling scores and AUS-MEAT marbling score. Consequently, IMF% can be used as a proxy for MSA marbling score in comparison of our results with those of other studies. Additionally, our research focuses on the feasibility of the Q-FOM™ beef camera to predict MSA marbling score in French cattle breeds, which exhibit a limited marbling range when compared to the cattle breeds found in Australia.

The sensitivity, specificity, and overall accuracy results indicate that the models are particularly effective in identifying samples with low MSA scores, while their performance in predicting medium and high MSA classes is slightly lower. The high specificity values indicate that the models are reliable in correctly identifying samples that do not belong to each class, reducing the rate of false positives. These findings suggest that the models used in this study demonstrate a higher effectiveness compared to the results reported by Kombolo-Ngah et al. (2023), who also evaluated the implementation of handheld near-infrared spectrometers for the on-line prediction of beef marbling score. This superior performance underlines the quality of the device, and the robustness of the models used in this study in distinguishing between different marbling classes, demonstrating the potential usefulness of the Q-FOM™ beef camera in practical applications.

The analysis of the ROC curves and AUC values suggests that both models (expert grader vs. second grader and expert grader vs. Q-FOM™ beef camera) have a good discriminatory ability between different MSA marbling classes. The AUC close to 0.85 for both models indicates robust performance in correctly classifying the marbling classes.

Comparing the two models, the Q-FOM™ beef camera showed a slightly higher AUC, suggesting that it may be a viable alternative to the 3G accredited grader for evaluating marbling scores. However, the difference is not substantial, and both methods are effective for classifying MSA marbling.

4.3. Prediction of the marbling scores using the camera on-site and from images

The prediction accuracy and precision of in-chiller and on-screen evaluations are nearly equivalent to those of expert and 3G graders. This suggests that when in-chiller trials are not feasible, on-screen assessing by an expert grader could serve as a viable alternative. It remains to be elucidated if on-screen assessment by accredited graders would also be a viable alternative in the future. Moreover, this approach would allow for the inclusion of a diverse range of cattle, sexes, and feeding, thereby enhancing the robustness of predictive models for MSA marbling score. While on-screen assessing may not be applicable to all quality traits (e.g., tenderness, color), it appears to perform well for MSA marbling, demonstrating comparable accuracy and precision to visually performed in-chiller assessments.

Discrepancies were observed in only two carcasses (Fig. 9a and b), suggesting a possible need to calibrate and standardize monitor settings to account for variations in luminosity, contrast, and color, but it is likely to result from simple errors. Additionally, it is possible that these discrepancies stem from errors made by the expert grader in predicting the marbling scores for these two samples. In the chiller, assessments are conducted under controlled conditions, with specific lighting and angles as prescribed by the MSA methodology. These conditions are carefully chosen to enhance the accuracy of visual assessments by minimizing glare and shadows. In contrast, on-screen evaluation lacks such controlled lighting, relying instead on monitor settings and on the picture taken by the Q-FOM™ camera, which can introduce variability in the perception of marbling and potentially lead to overestimation of marbling scores.

The rapid assessment conducted in the chiller, which is required to minimize production interruptions, may contribute some variability. However, in this study, an expert grader performed the assessments, accustomed to the fast-paced and challenging conditions of industrial environments, thereby reducing the likelihood of error. Consequently, the controlled environment and lighting in the chiller offer a consistent and potentially reliable assessment as on-screen evaluations.

These findings emphasize the critical importance of maintaining consistent assessment conditions across different evaluation methods. The discrepancies noted in outlier cases suggest that even minor variations in environmental factors, such as lighting and observation angle, can significantly impact marbling assessments. Further research is required to explore the significance of these factors and to develop screen calibration techniques that can harmonize results across various platforms used for marbling assessment.

5. Conclusion

This study demonstrated for the first time that the Q-FOM™ Beef assessing camera can accurately predict MSA marbling scores using the MSA grading scheme for carcasses of breeds typical for European cattle quartered between the 5th and 6th rib, which is common practice in Europe. The Q-FOM™ Beef camera proved to be effective in predicting marbling score in the context of French beef production, where marbling scores are generally lower than beef produced in other parts of the world, particularly in Australia where the MSA grading scheme originated. Both in-chiller and on-screen visual assessments done by an expert grader were found to be reliable methods for developing a Q-FOM™ Beef marbling calibration model, making it a valuable tool for enhancing the consistency and accuracy of marbling evaluations within the French livestock industry.

Funding declaration

The authors wish to acknowledge the generous support provided by the FAPEG/CAPES (Fundação de Amparo à Pesquisa do Estado de Goiás / Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) under grant number No. 202110267000231, the French Embassy in Brazil and the Beauvallet Company for the support of the scholarship of Nathalia da Silva Rodrigues Mendes. The authors are also grateful to the INTAQT project for funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 101000250.

CRediT authorship contribution statement

Nathalia da Silva Rodrigues Mendes: Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Mette Christensen:** Writing – review & editing, Writing – original draft, Validation, Formal analysis. **Moïse Kombolo-Ngah:** Formal analysis, Data curation. **Pascal Faure:** Formal analysis, Data curation. **Laure Thoumy:** Formal analysis, Data curation. **Alix Neveu:** Writing – review & editing, Formal analysis, Data curation. **Amanda Gobeti Barro:** Formal analysis, Data curation. **Jingjing Liu:** Formal analysis, Data curation. **Tatianne Ferreira de Oliveira:** Validation, Supervision, Funding acquisition. **Marie-Pierre Ellies-Oury:** Writing – review & editing, Validation, Supervision. **Sghaier Chriki:** Writing – review & editing, Validation, Supervision. **Jean-François Hocquette:** Conceptualization, Funding acquisition, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare to have no conflicts of interest.

Data availability

Data will be made available on request.

References

Almeida, M. R., Fidelis, C. H. V., Barata, L. E. S., & Poppi, R. J. (2013). Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation. *Talanta, 117*, 305–311. https://doi.org/10.1016/j.talanta.2013.09.025

Anderson, F., Cook, J., Williams, A., & Gardner, G. E. (2018). Computed tomography has improved precision for prediction of intramuscular fat percent in the M. *Longissimus thoracis* et lumborum in cattle compared to manual grading. *Meat Science, 145*, 425–430. https://doi.org/10.1016/j.meatsci.2018.07.025

AUS-MEAT. (2022a). *Handbook of Australian Meat international red Meat manual* (8th ed.). Sydney, Australia: AUS-MEAT Limited.

AUS-MEAT. (2022b). *AUS-MEAT AMILSC approved guidelines for experiments to achieve accreditation of technologies for predicting chemical IMF% in beef (EAI007)*. Brisbane, Australia: AUS-MEAT Limited.

AUS-MEAT. (2022c). *AUS-MEAT Australian Meat industry language and standards committee (AMILSC) approved minimum requirements of accuracy standards for cut surface cameras version 5.0 (EAI002)*. Brisbane, Australia: AUS-MEAT Limited.

AUS-MEAT. (2023). AUS-MEAT list of approved equipment suppliers (Version 08). https://www.ausmeat.com.au/WebDocuments/AUSMEAT_List_of_Approved_Equipment_Suppliers.pdf Accessed on 06.15.24.

Beriain, M. J., Insausti, K., Valera, M., Indurain, G., Purroy, A., Carr, T. R., & Horcada, A. (2021). Effectiveness of using ultrasound readings to predict carcass traits and sensory quality in young bulls. *Computers and Electronics in Agriculture, 183*, Article 106060. https://doi.org/10.1016/j.compag.2021.106060

Bonny, S. P. F., O'Reilly, R. A., Pethick, D. W., Gardner, G. E., Hocquette, J. F., & Pannier, L. (2018). Update of meat standards Australia and the cuts based grading scheme for beef and sheepmeat. *Journal of Integrative Agriculture, 17*(7), 1641–1654. https://doi.org/10.1016/S2095-3119(18)61924-0

Cama-Moncunill, R., Cafferky, J., Augier, C., Sweeney, T., Allen, P., Ferragina, A., … Hamill, R. M. (2020). Prediction of Warner-Bratzler shear force, intramuscular fat, drip-loss and cook-loss in beef via Raman spectroscopy and chemometrics. *Meat Science, 167*. https://doi.org/10.1016/j.meatsci.2020.108157

Chambaz, A., Scheeder, M. R. L., Kreuzer, M., & Dufey, P. A. (2003). Meat quality of Angus, Simmental, Charolais and Limousin steers compared at the same intramuscular fat content. *Meat Science, 63*(4), 491–500. https://doi.org/10.1016/S0309-1740(02)00109-2

Chen, D., Li, W., Du, M., & Cao, B. (2019). Adipogenesis, fibrogenesis and myogenesis related gene expression in longissimus muscle of high and low marbling beef cattle. *Livestock Science, 229*, 188–193. https://doi.org/10.1016/j.livsci.2019.09.032

Cheng, W., Cheng, J. H., Sun, D. W., & Pu, H. (2015). Marbling analysis for evaluating meat quality: Methods and techniques. *Comprehensive Reviews in Food Science and Food Safety, 14*, 523–535. https://doi.org/10.1111/1541-4337.12149

Chriki, S., Picard, B., Faulconnier, Y., Micol, D., Brun, J. P., Reichstadt, M., … Hocquette, J. F. (2013). A data warehouse of muscle characteristics and beef quality in France and a demonstration of potential applications. *Italian Journal of Animal Science, 12*(2), Article e41. https://doi.org/10.4081/ijas.2013.e41

Coombs, C. E. O., Fajardo, M., & González, L. A. (2021). Comparison of smartphone and lab-grade NIR spectrometers to measure the chemical composition of lamb and beef. *Animal Production Science, 61*(16), 1723. https://doi.org/10.1071/AN21069

Dikeman, M. E. (1996). The relationship of animal leanness to meat tenderness. In *, vol. 49. Reciprocal Meat Conference* (pp. 87–101).

Drachmann, F. F., Christensen, M., Esberg, J., Lauridsen, T., Fogh, A., Young, J. F., & Therkildsen, M. (2024). Beef-on-dairy: Meat quality of veal and prediction of intramuscular fat using the Q-FOM™ beef camera at the 5th-6th thoracic vertebra. *Meat Science, 213*, Article 109503. https://doi.org/10.1016/j.meatsci.2024.109503

Etats Generaux de l''Alimentation EGA. (2018). https://agriculture.gouv.fr/alimagri-les-etats-generaux-de-lalimentation Accessed on 11.30.2023.

Farmer, L. J., & Farrell, D. T. (2018). Review: Beef-eating quality: A European journey. *Animal, 12*(11), 2424–2433. https://doi.org/10.1017/S1751731118001672
Ferguson, D. M. (2004). Objective on-line assessment of marbling: A brief review. *Australian Journal of Experimental Agriculture, 44*(7), 681–685. https://doi.org/10.1071/EA02161
Guaranteed Global Grading. (2023). 3G guaranteed global grading. https://imr3g.org/about-foundation/ Accessed on 15.06.23.
Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747
Hickey, J. M., Keane, M. G., Kenny, D. A., Cromie, A. R., & Veerkamp, R. F. (2007). Genetic parameters for EUROP carcass traits within different groups of cattle in Ireland. *Journal of Animal Science, 85*, 314–321. https://doi.org/10.2527/jas.2006-263
Hocquette, J. F., Ellies-Oury, M. P., Legrand, I., Pethick, D., Gardner, G., Wierzbicki, J., & Polkinghorne, R. (2020). Research in beef tenderness and palatability in the era of big data. *Meat and Muscle Biology, 4*(2). https://doi.org/10.22175/mmb.9488
Hocquette, J. F., Ellies-Oury, M. P., Lherm, M., Pineau, C., Deblitz, C., & Farmer, L. (2018). Current situation and future prospects for beef production in Europe—A review. *Asian-Australasian Journal of Animal Sciences, 31*(7), 1017–1035. https://doi.org/10.5713/ajas.18.0196
Hocquette, J. F., Gondret, F., Baéza, E., Médale, F., Jurie, C., & Pethick, D. W. (2010). Intramuscular fat content in meat-producing animals: Development, genetic and nutritional control, and identification of putative markers. *Animal, 4*(2), 303–319. https://doi.org/10.1017/S1751731109991091
Hocquette, J. F., Meurice, P., Brun, J. P., Jurie, C., Denoyelle, C., Bauchart, D., … Picard, B. (2011). The challenge and limitations of combining data: A case study examining the relationship between intramuscular fat content and flavour intensity based on the BIF-BEEF database. *Animal Production Science, 51*(11), 975–981. https://doi.org/10.1071/AN10044
Jacob, R. (2020). Implications of the variation in bloom properties of red meat: A review. *Meat Science, 162*, Article 108040. https://doi.org/10.1016/j.meatsci.2019.108040
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning with applications in R. *Springer*. https://doi.org/10.1007/978-1-4614-7138-7
Jang, J. W., Ishdorj, A., Anderson, D. P., Purevjav, T., & Dahlke, G. (2017). Exploring the existence of grader bias in beef grading. *Journal of Agricultural and Applied Economics, 49*(3), 467–489. https://doi.org/10.1017/aae.2017.9
Kombolo-Ngah, M., Goi, A., Santinello, M., Rampado, N., Atanassova, S., Liu, J., … Hocquette, J. F. (2023). Across countries implementation of handheld near-infrared spectrometer for the on-line prediction of beef marbling in slaughterhouse. *Meat Science, 200*. https://doi.org/10.1016/j.meatsci.2023.109169
Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3
Lee, B., & Choi, Y. M. (2019). Correlation of marbling characteristics with meat quality and histochemical characteristics in longissimus thoracis muscle from Hanwoo steers. *Food Science of Animal Resources, 39*, 151–161. https://doi.org/10.5851/kosfa.2019.e12
Lee, B., Yoon, S., & Choi, Y. M. (2019). Comparison of marbling fleck characteristics between beef marbling grades and its effect on sensory quality characteristics in high-marbled Hanwoo steer. *Meat Science, 152*, 109–115. https://doi.org/10.1016/j.meatsci.2019.02.019
Lee, S., Lohumi, S., Lim, H. S., Gotoh, T., Cho, B. K., & Jung, S. (2015). Determination of intramuscular fat content in beef using magnetic resonance imaging. *Faculty of Agriculture Kyushu University, 60*(1), 157–162. https://doi.org/10.5109/1526313
Legrand, I., Hocquette, J. F., Polkinghorne, R. J., & Pethick, D. W. (2013). Prediction of beef eating quality in France using the Meat standards Australia system. *Animal, 7*(3), 524–529. https://doi.org/10.1017/S1751731112001553
Liu, J., Chriki, S., Ellies-Oury, M.-P., Legrand, I., Pogorzelski, G., Wierzbicki, J., Farmer, L., Troy, D., Polkinghorne, R., & Hocquette, J.-F. (2020). European conformation and fat scores of bovine carcasses are not good indicators of marbling. *Meat Science, 170*. https://doi.org/10.1016/j. meatsci.2020.108233
Liu, J., Pogorzelski, G., Neveu, A., Legrand, I., Pethick, D., Ellies-Oury, M. P., & Hocquette, J. F. (2021). Are marbling and the prediction of beef eating quality affected by different grading sites? *Frontiers in Veterinary Science, 8*. https://doi.org/10.3389/fvets.2021.611153
Meat, Livestock Australia, Meat Standards Australia. (2024). *Meat standards Australia information kit*. North Sydney: Meat & Livestock Australia. https://www.mla.com.au/globalassets/mla-corporate/marketing-beef-and-lamb/documents/meat-standards-australia/msa-mla-tips-and-tools-msa00-complete-booklet_v03.pdf.
Mendes, N. S. R., Briceno, J. C. C., Mársico, E. T., Ellies-Oury, M. P., Chriki, S., Hocquette, J. F., & e Oliveira, T. F.. (2024). Recent technological developments and future trends in the evaluation and prediction of beef sensory quality in Brazil and France. *Livestock Science*. , Article 105550. https://doi.org/10.1016/j.livsci.2024.105550
Mendes, N. S. R, Silva, R. R., Kombolo-Ngah, M., Rivet, P.-H., Tondusson, J., Oliveira, … Hocquette, J.-F. (2024). Does transport affect the eating quality potential of beef from Limousin cows in France? - A case study. *Veterinary and Animal Science, 26*, 100411. https://doi.org/10.1016/j.vas.2024.100411
Monteils, V., Sibra, C., Ellies-Oury, M. P., Botreau, R., De la Torre, A., & Laurent, C. (2017). A set of indicators to better characterize beef carcasses at the slaughterhouse level in addition to the EUROP system. *Livestock Science, 202*, 44–51.
Pethick, D. W., Harper, G. S., & Oddy, V. H. (2004). Growth, development and nutritional manipulation of marbling in cattle: A review. *Australian Journal of Experimental Agriculture, 44*(7), 705–715. https://doi.org/10.1071/EA02165
Pogorzelski, G., Woźniak, K., Polkinghorne, R., Połtorak, R., & Wierzbicka, A. (2020). Polish consumer categorisation of grilled beef at 6mm and 25mm thickness into quality grades, based on Meat Standards Australia methodology. *Meat Science, 161*, 107953. https://doi.org/10.1016/j.meatsci.2019.107953
Prieto, N., Navajas, E. A., Richardson, R. I., Ross, D. W., Hyslop, J. J., Simm, G., & Roehe, R. (2010). Predicting beef cuts composition, fatty acids and meat quality characteristics by spiral computed tomography. *Meat Science, 86*(3), 770–779.
Sanchez, P. D. C., Arogancia, H. B. T., Boyles, K. M., Pontillo, A. J. B., & Ali, M. M. (2022). Emerging nondestructive techniques for the quality and safety evaluation of pork and beef: Recent advances, challenges, and future perspectives. *Applied Food Research, 2*(2), Article 100147.
Santinello, M., De Marchi, M., Diana, A., Rampado, N., Hocquette, J. F., & Penasa, M. (2022). Effect of commingling animals at sorting facilities on performances and antibiotic use in beef cattle. *Italian Journal of Animal Science, 21*, 771–781. https://doi.org/10.1080/1828051X.2022.2063766
Santinello, M., Rampado, N., Penasa, M., Hocquette, J. F., Pethick, D., & De Marchi, M. (2024). The Meat standards Australia carcass grading site affects assessment of marbling and prediction of meat-eating quality in growing European beef cattle. *Meat Science, 213*, Article 109501. https://doi.org/10.1016/j.meatsci.2024.109501
Stewart, S. M., Gardner, G. E., McGilchrist, P., Pethick, D. W., Polkinghorne, R., Thompson, J. M., & Tarr, G. (2021). Prediction of consumer palatability in beef using visual marbling scores and chemical intramuscular fat percentage. *Meat Science, 181*, Article 108322. https://doi.org/10.1016/j.meatsci.2020.108322
Stewart, S. M., Gardner, G. E., & Tarr, G. (2024). Using chemical intramuscular fat percentage to predict visual marbling scores in Australian beef carcasses. *Meat Science*. , Article 109573. https://doi.org/10.1016/j.meatsci.2024.109573
Stewart, S. M., Lauridsen, T., Toft, H., Pethick, D. W., Gardner, G. E., McGilchrist, P., & Christensen, M. (2021). Objective grading of eye muscle area, intramuscular fat and marbling in Australian beef and lamb. *Meat Science, 181*, Article 108358. https://doi.org/10.1016/j.meatsci.2020.108358
Stewart, S. M., Toft, H., O'Reilly, R. A., Lauridsen, T., Esberg, J., Jørgensen, T. B., … Christensen, M. (2024). Objective grading of rib eye traits using the Q-FOM™ camera in Australian beef carcasses. *Meat Science, 213*, 109500. https://doi.org/10.1016/j.meatsci.2024.109500
Tedeschi, L. O. (2006). Assessment of the adequacy of mathematical models. *Agricultural Systems, 89*, 225–247. https://doi.org/10.1016/j.agsy.2005.11.004
Thompson, J. M. (2004). The effects of marbling on flavour and juiciness scores of cooked beef, after adjusting to a constant tenderness. *Australian Journal of Experimental Agriculture, 44*(7), 645–652. https://doi.org/10.1071/EA02171
Watson, R., Polkinghorne, R., & Thompson, J. M. (2008). Development of the Meat Standards Australia (MSA) prediction model for beef palatability. *Australian Journal of Experimental Agriculture, 48*(11), 1368–1379. https://doi.org/10.1071/EA07184
Yang, S., & Berdine, G. (2017). The receiver operating characteristic (ROC) curve. *The Southwest Respiratory and Critical Care Chronicles, 5*(19), 19. https://doi.org/10.12746/swrccc.v5i19.391