Decision tree, a learning tool for the prediction of beef tenderness using rearing factors and carcass characteristics

Mohammed Gagaoua, Valérie Monteils, Brigitte Picard

Decision tree, a learning tool for the prediction of beef tenderness using rearing factors and carcass characteristics

Mohammed Gagaoua*, Valérie Monteils, Brigitte Picard

Université Clermont Auvergne, INRA, VetAgro Sup, UMR Herbivores, F-63122 Saint-Genès-Champanelle, France

Correspondence
*Dr. Mohammed GAGAOUA;
mohammed.gagaoua@inra.fr
Phone: +33 473 624 239. Fax: +33 473 624 639.

Abstract:

Background: This study explored the potential use of decision trees on rearing factors (q = 10) and carcass characteristics (q = 12) for the development of prediction model rules of beef tenderness prediction/categorization. For that, 308 young bulls were used to evaluate by a sensory panel the tenderness potential of ribeye steaks grilled at 55°C. The regression and classification C&RT method was implemented and allowed the prediction of tenderness using i) rearing factors, ii) carcass characteristics or iii) both.

Results: The resultant tree models yielded predictive accuracies of 70.78% (with 4 rearing factors: concentrate per cent; fattening duration; initial body weight and dry matter intake); 67.21% (with 4 carcass characteristics: fatness carcass score; carcass weight; dressing per cent and muscle carcass per cent) and 84.41% (with 6 rearing factors and carcass characteristics) compared to the *k*-means clustering of tenderness. In the final and robust regression tree, from the 22 attribute information 2 carcass characteristics (fatness carcass score and muscle carcass per cent) and 4 rearing factors (fattening duration; concentrate per cent; dry matter intake and initial body weight) were retained as predictors. The first splitter of the 308 ribeye steaks according to their tenderness scores was fatness carcass score, followed by fattening duration and concentrate per cent.

Conclusion: This trial highlighted the importance of thresholding approach to efficiently classify ribeye steaks according to their tenderness potential. The overall prediction model rule was as follow: IF (fatness carcass score ≥ 2.88) AND (concentrate ≥ 82%) [AND (muscle carcass ≥ 71%)] THEN meat was [very] tender.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jsfa.9301

Key words: Meat tenderness; Young bulls; Farm–to–fork continuum; Clustering; Decision tree

Introduction

Tenderness is one of the most important traits for assessing beef quality and it is among the main drivers of beef palatability that consumers consider when they are making (re)purchasing decision[1]. Important classical research on tenderness in order to satisfy the consumers was gaining attention over more than 50-yr period. These studies reported that beef tenderness depends on numerous factors that are related or not to the animal[2-5]. Most of each factor including both pre- and post-slaughter events were mostly investigated objectively. They included animal type, age at slaughter, muscle-type, management factors such as feeding regimens or rearing factors, slaughtering conditions and pre-harvest stress or post-harvest conditions and cooking with many interaction among these factors. During the last years, there was considerable progress in the understanding of the mechanisms underlying meat tenderness determinism[6-8], however, there is still an irregularity in tenderness that challenges both researchers and stakeholders[9].

On another hand, consumers are willing to pay for guaranteed-tender meat[10]. During the last three decades and in order to better meet consumers' expectations, especially in terms of perceived sensory eating quality, many different strategies have been developed. Thus, several research groups tried to identify indicators explaining the uncontrolled variability in order to predict the tenderness potential of carcasses soon after slaughter[6, 11, 12]. This allowed the proposition of numerous tenderness prediction tools that were based on subjective or objective evaluations. Several grading systems of carcasses [EUROP classification system, Australian Meat Standards (MSA), Canadian beef grading system, Japanese grading standards, and United States Department of Agriculture (USDA) grading system] were also developed over the world[13], to describe the quality and yield of carcasses to ensure consistent meat quality and consumer satisfaction. However, some of these systems are not completely accurate as they only reflect a compilation of traits that are indicators of tenderness or eating quality is not yet part among them, namely for the EUROP system[14]. Consequently, recent studies proposed to include further variables that are related to the fattening period factors for joint management of both carcass and meat qualities[15-17]. Thus, based on those studies and in an attempt of performing machine learning tools to sort beef carcasses into tenderness groups, we hypothesized that fattening period factors and carcass characteristics of young bulls, would be accurate for the prediction of tenderness potential. Moreover, the use of this approach would be easily implementable as an on-line evaluation tool to help the stakeholders for early management of carcasses according to their tenderness potential. In this trial, we hypothesized

that the rearing factors and carcass characteristics will play a more pivotal role than muscle properties to predict beef tenderness. Thus, the main objective of this study is to perform a holistic study by considering both rearing factors and carcass characteristics to predict beef tenderness classes using a decision tree-based learning method. Furthermore, in this work we propose for farmers, breeders' associations, and other industry stakeholders an original model-based approach, that would be continuously used to monitor animal and farm level information for joint management of beef tenderness or other meat sensory qualities.

Materials and Methods

Animal handling

A total of 308 young bulls that were handled in accordance with the French animal protection regulations (Code Rural, articles R214-64 to R214-71) were used. These animals constituted individual data of a dataset from the study by Gagaoua *et al.*[5] grouping experiments conducted in two INRA France experimental units: INRA-UE232 (Bourges) for animal rearing and handling and INRA-UE1414 Herbipôle (Theix) for lairage and slaughtering[4]. The dataset included Charolais (n = 116), Salers (n = 93), Limousin (n = 74), and Blond d'Aquitaine (n = 25) breeds[5]. The 4 animal's genetics reflect the distribution of the major suckled breeds slaughtered in France, namely in the Auvergne-Rhône-Alpes region. The animals were all born from mature cows, weaned at 32 weeks and then kept in an open shed. All the animals and irrespective of experiment and production year were raised and managed under similar conditions in the INRA research center. The animals were fed *ad libitum* feed consisting of forages (including straw, hay, grass silage, corn silage, and beet pulp silage) and concentrate containing soybean, urea, corn and barley grain. Diets were formulated according to the recommendations of INRA (2007) to meet NEg requirements needed for maintenance and theoretical body weight gain.

Rearing factors characterizing the fattening period

The rearing practices of the 308 young bulls were characterized by 10 rearing factors (**Table S1**) as described by Gagaoua *et al.*[5]. They included slaughter age (months), fattening duration (days), initial body weight at the beginning of the fattening period (initial body weight, kg), final body weight (kg), dry matter intake (DMI, kg DM/day), forage and concentrate in per cent (in the DM diet), energy intake (Mcal/day), average daily gain (ADG) for the fattening period, and feed efficiency (ADG/DMI ratio, kg/kg DM)[5].

Slaughtering and sampling

Before slaughter, all animals were fasted for 24 h and had free access to water. Slaughtering was performed in compliance with the French welfare regulations set at the level of European Directive (2001/88/EC) in the experimental slaughterhouse of INRA. The carcasses were not electrically stimulated and they were stored between 2 and 4°C until 24 h *post-mortem*. Muscle samples from the *Longissimus thoracis* (LT, mixed fast oxido-glycolytic muscle) were excised from the center of the 6th rib located at the right side of each carcass as illustrated in Gagaoua *et al.*[4]. The epimysium was carefully dissected and the samples were cut into steaks (2 cm thick) and placed in sealed plastic bags under vacuum and kept at 4°C for ageing (14 days). Each sample was then frozen and stored at −20°C until sensory evaluation.

Carcass grading and composition by the 6th rib dissection

The carcasses were graded under the EU beef carcass classification scheme. This was done at the slaughterhouse using experts familiar with the EUROP grid (Commission Regulation (EC) 1249/2008). Five conformation classes are defined, represented by the letters E, U, R, O, and P. The scoring consists of a visual assessment of carcass muscling where carcasses graded as 'E' have the most muscularity, and this decreases through to 'P' which have the least muscularity. European Union regulations allow for 3 subdivisions of each conformation class, high: "+", medium: "=" and low: "-". Hence, an incremental scale ranging from 1 to 15 was used, where 1 corresponds to P- (very low muscle development) and 15 to E+ (very high muscle development)[4]. At the same time, the fatness score of the carcasses, which describes the amount of fat on the outside of the carcass was numerically scored from 1 = leanest to 5 = fattest.

The carcass weight is typically measured within 2 h of slaughter and is expressed as cold carcass weight, which is 0.98 times the hot carcass weight according to the EUROP classification guidelines (Commission Regulation (EC) 1249/2008). The carcass weight was used to calculate the dressing per cent and carcass composition. For carcass composition (**Table S1**), muscle carcass per cent, fat carcass per cent, and bone carcass per cent were computed[4]. This was estimated from the composition of the 6th rib determined after physical dissection following standard commercial practice involving a close trimming and deboning of the muscle[4].

Meat tenderness evaluation

For tenderness evaluation, the sensory protocol recently described by our group was used[18-20]. Briefly, the steaks aged for 14 days were thawed at 5°C for 48 h before cooking and sensory

assessment at 55°C, the usually used end-point cooking temperature in France[18-20]. One hour before sensory assessment, the meat samples were cut into approximately 1.50 cm thick steaks and grilled on a double grooved plate griddle (SOFRACA, Morangis, France) heated to 300°C for 30 min before cooking. Steaks were heated for 2 min until the end-points temperature of 55°C in the geometric center of the steak was reached using a temperature probe (Type K, HANNA HI 98704, Newark, USA). After grilling, each steak was cut into 20 mm cubes that were immediately served to 12 panelists chosen according to the criteria described by Gagaoua *et al.* [18]. Thus, sensory panels for each experiment rated the steaks on a 10 cm unstructured line scale (from 0 to 10) for global tenderness, where 0 refers to extremely tough meat and 10 to extremely tender meat[18]. Within each sensory session, scores were averaged across panelists for each steak, and the means were used in the statistical analyses.

Statistical analyses

The data were analyzed using SAS 9.4 and XLSTAT 2017.19.6 software's. Before analysis, the data were scrutinized for outliers using Smirnov–Grubb's outlier test at 5% level.

As the main objective of this study is to implement decision trees based on clustering methods, we first categorized the carcasses according to their ribeye steaks tenderness scores[5]. Thus, we have used different unsupervised learning tools such as *k*-means, hierarchical cluster analysis (HCA) and partitioning around medoids (PAM) clustering algorithms[5, 21, 22]. The tenderness scores were mean centered prior to any clustering analysis by computing Z-scores. These latter represent the deviation of each trait observation relative to the mean of the corresponding animal irrespective of the breed and experiment. They were computed using the PROC STANDARD of SAS[23]. Using the residual scores of tenderness, the *k*-means and HCA algorithms gave the best results as detailed in Gagaoua *et al.*[5].

- *K*-means, is an iterative clustering method used to automatically partition a dataset into *k* groups where the number of clusters *k* is assumed to be fixed *a priori* by minimizing relative distances[21, 24]. This algorithm consists of two separate phases: the first phase is to define *k* centroids, one for each cluster; and the second phase is to take each point belonging to the given dataset and associate it to the nearest centroid.
- HCA is a cluster analysis, which seeks to build a hierarchy and binary tree of clusters[22, 25]. It classifies a dataset into groups that are internally homogeneous and externally isolated on the basis of measuring the similarity or dissimilarity between groups.

The average silhouette width (S_i) criterion was used for the validation of the clustering results to choose both the best clustering method and the number of clusters. S_i is based on the proximity among objects of a certain cluster, and the proximity of these objects to the nearest clusters[25] that refers to a succinct graphical representation method for validation of consistency within clusters (**Fig. 1**). The value of S_i ranges from −1 to +1, and can be interpreted as follow:

i) observations with a large S_i (almost 1) are very well clustered;

ii) a small S_i (around 0) means that the observation lies between two clusters and

iii) observations with a negative S_i are probably partitioned in the wrong cluster.

In this trial, the best results were obtained for *k*-means cluster algorithm using the Ward method. Thus, three clusters were identified and named tender (TEND+), medium (TEND=) and tough (TEND–).

Afterward and as for the classification analysis, different decision trees[26] using CHAID, C&RT and QUEST methods were performed using i) animal and fattening period characteristics; ii) carcass properties or iii) both, to predict the tenderness clusters identified by *k*-means[5]. Decision tree learning is a method commonly employed in data-mining as non-parametric supervised learning tools for classification and regression. For each decision tree, the accuracy was measured in addition to sensitivity and specificity.

The C&RT method gave in this study the best results leading to the greatest accuracy. It is the most popular decision tree method that provides critical variables threshold and allows identifying their directional influence on the outcomes[27]. C&RT method was implemented in order to show which rearing factors or carcass characteristics were more important and discriminated better the tenderness clusters. Using C&RT, the same predictor (splitter) variable or attribute may be used several times at different levels in the tree. Therefore, the measure of Gini impurity was used to categorize the target variables and identify the best splitters. The Gini index of a node with *n* objects and *c* possible clusters is defined as:

$$\text{Gini} = 1 - \sum_{j=1}^{n} \left(\frac{n_j}{n}\right)^2 \qquad (1)$$

where n_j is the number of objects from class *j* present in the node. If all of the observations in one node belong to one class, the Gini index will equal zero, which means the least impurity or greatest purity in that node. The greatest values are obtained when the same ratio of observations exists in the node.

Results

Beef tenderness, rearing factors and carcass characteristics

A general description analysis of the data was performed by computing mean, standard deviation and minimum and maximum ranges (**Table S1**). For tenderness scores evaluated by a trained panel, the coefficient of variation was 21.4%. After statistical normalization of tenderness scores for breed and experiment effects as well as existing interaction[5], the Z-scores allowed to build 3 tenderness clusters as detailed in Gagaoua *et al*.[5] (**Table S2**). The clusters were validated based on the silhouette widths (**Fig. 1**), and *k*-means gave a better S_i of 0.51 compared to HCA (S_i = 0.47) clustering method. As depicted in **Table S2**, the tough cluster named TEND–, grouped 108 animals that had an average tenderness score value of 4.14±0.57. The medium tenderness cluster named TEND=, grouped 141 animals and had an average score value of 5.46±0.43. The tender cluster named TEND+, grouped 59 animals and had an average tenderness score value of 6.95±0.62.

For the rearing factors, 7 of 10 variables differed among the 3 clusters (**Table S2**). The ribeye steaks of the TEND–were characterized compared to the TEND+ cluster by the highest age at slaughter (17.05 *vs.* 15.99 months), initial (412 *vs.* 367 kg) and final body weights (675 *vs.* 621), fattening duration (182 *vs.* 155 days), forage per cent (58.11 *vs.* 41.23%), and dry matter intake (8.99 *vs.* 8.04 kg DM/day). The concentrate per cent was the only variable higher in TEND+ than in TEND– cluster (**Table S2**). The three remaining factors: energy intake, average daily gain and feed efficiency were all similar among the clusters[5].

For the carcass characteristics, 6 of 12 variables differed between the tenderness clusters (**Table S2**). Three of them were the highest in TEND+ than TEND– and refer to fat estimates: fat carcass per cent (+14.45%), fat carcass weight of the 6th rib (+26.7%) and fatness carcass score (+21.7%). The 3 other variables that were the highest in TEND– comparatively to TEND+ cluster, refer to muscle carcass and are muscle carcass per cent (+3.66%), carcass weight (+8.94%) and muscle carcass weight (+11.83%).

Decision tree of beef tenderness prediction/classification using rearing factors

The final objective of performing decision trees is to create a model that predicts beef tenderness by learning simple decision rules inferred from the tenderness scores using i) rearing factors, ii) carcass characteristics or iii) both. Accordingly, the first partitioning pattern of tenderness clusters revealed by the C&RT is given in **Fig. 2**. The resultant model yielded a predictive accuracy of 70.78% compared to *k*-means (**Table 1**). From the whole dataset, the

first splitter was concentrate per cent and generated two groups. The first group contained 65 animals which were then differentiated by a threshold of 420 kg of initial body weight into TEND= (n=13) and TEND+ (n=52). The second group contained 243 animals, which at the end of the tree were partitioned as TEND– and TEND=. A fattening duration higher than 254 days (n=32) or a combination of fattening duration lower than this threshold associated with an initial body weight greater than 426 kg (n=60), allowed the categorization of all the TEND– samples. From the 10 rearing factors, concentrate per cent, initial body weight, fattening duration and dry matter intake (DMI) were the most important splitters. Overall, they led to tenderness clusters that were in line with those of *k*-means. Consequently, tough meat was scored below 5.0 points scores, tender meat was always scored over 6.0 and medium meat was scored over 5.0 and below 6.0.

Decision tree of beef tenderness prediction/classification using carcass characteristics

The second decision tree built using carcass characteristics is given in **Fig. 3**. The resultant model yielded a predictive accuracy of 67.21% that was slightly lower (-3.57%) than the first decision tree (**Table 1**). The first splitter was fatness carcass score. Therefore, if the mean fatness carcass score was higher than 2.42, the ribeye steaks belonged to the first group of 202 animals which, contained carcasses from the 3 clusters; and if they were less than 2.42, then, they belonged to the second group of 106 animals grouping carcasses of TEND– and TEND=. For this later cluster, carcass weight (≥511 kg) was the following predictor that identified the 12 carcasses of TEND= from those of 94 TEND– cluster. The 94 carcasses (mean tenderness scores of 4.75) were further partitioned by dressing per cent at a threshold of 60% into two groups of 41 (mean tenderness scores of 4.55; dressing is < 60%) and 53 carcasses (mean tenderness scores of 4.91; dressing is ≥60%). For the second group of 202 animals, it was carcass weight, which allowed the identification of two groups. If carcass weight was higher than 419 kg, 38 carcasses were partitioned and none of them contained tender meat; if their muscle carcass per cent (the third splitter) was higher than 70.6%, 13 were in TEND– cluster. However, if carcass weight was below the threshold of 419 kg, the 164 carcasses belonged to TEND= or TEND+. For the tender cluster, if dressing was higher than 60%, this allowed the identification of the 51 tender ribeye steaks. Therefore, the decision tree could simply and easily apply discrimination rule to identify based on fatness carcass score, carcass weight, dressing per cent and muscle carcass per cent the TEND–, TEND= or TEND+ clusters.

Decision tree of beef tenderness prediction/classification using both rearing factors and carcass characteristics

The last decision tree (**Fig. 4**) was built based on rearing factors and carcass characteristics. The resultant model yielded a predictive accuracy of 84.41% that was the highest compared to the first decision trees (**Table 1**). Among the independent variables, the first splitter of was fatness carcass score. Therefore, if the mean fatness carcass score was higher than 2.88, the ribeye steaks belonged to the first group of 166 animals which contain TEND+ and TEND= clusters only; and if they were less than 2.82, then, they belonged to the second group of 142 animals containing mainly carcasses of TEND– and 43 carcasses of the TEND=. After that and in line with the above results, concentrate per cent was the second splitter which partitioned the 166 carcasses into 60 TEND+ and 106 TEND=. Within the 60 carcasses of TEND+ cluster, 22 very tender ribeye steaks (mean tenderness of 6.84) that has muscle carcass higher than 71% could be distinguished from the 38 tender carcasses that has a mean tenderness of 6.31. The 106 TEND= carcasses were clustered using DMI at a threshold of 8.9 kg DM/day into two groups of 70 and 36 carcasses that has mean tenderness scores of 5.77 and 5.13, respectively. Within the 142 carcasses with fat carcass score below 2.88, fattening duration was the second predictor following fatness carcass score. Therefore, a fattening duration higher than 176 days grouped 49 carcasses in TEND–. However, the animals that has been fattened for a shorter period (<176 days) and at an initial body weight before fattening below 422 kg, led to medium (TEND=) ribeye steaks with a mean tenderness scores of 5.56. However, if they were fattened at an initial body weight more than and equal to 422 kg, all the ribeye steaks belonged to the TEND– cluster with a mean tenderness scores of 4.78. In summary, from the 22 attribute information 2 carcass characteristics (Fatness carcass score and muscle carcass, %) and 4 rearing factors (Fattening duration; concentrate per cent; DMI and initial body weight) were retained as predictors in this third decision tree. The rules suggest that the other attributes are insignificant and does not contribute much to the classification of ribeye steaks tenderness' according to their respective cluster.

Discussion

To the authors' knowledge, this study is the first to report the use of decision trees for the prediction/classification of beef tenderness using rearing factors and carcass characteristics. Furthermore, this study was undertaken by using variables that are at the same time easily measurable at the farm and slaughter levels to predict the tenderness potential of carcasses. Therefore, tenderness quality identification offered by the thresholding approach and decision tree allowed in this trial efficient classification of the ribeye steaks. This approach would challenge different actors of beef industry, including all levels of the beef chain from producer

to consumer, as it offers both feasibility and ability to objectively classify beef carcasses according to their tenderness potential.

Rearing factors as splitters (predictors) of meat tenderness clusters

In agreement with our hypothesis, we identified 4 important rearing factors (splitter attributes) for best management of the farming practices to achieve higher beef tenderness. It inferred from the regression trees that fattening duration is very important. Long fattening duration is not suitable as it leads to tough meat. This confirms our hypothesis dealing with the importance of the determination of the optimum thresholds of the factors governing the fattening period to achieve the best income and sustainability in cattle production[17]. Recently, we pointed out this importance, and proposed to take in account the fattening duration in relation with further traits such as fat tissue and muscle per cents of young bulls[5, 16, 17]. In cattle, several studies reported significant morphological changes that take place in the carcass and hence on muscle, during the fattening period and how these are to some extent confounded with age[2, 15]. However, the studies that have reported relationships between duration of fattening with beef sensory qualities are very scarce[28]. In line with our study, negative effect of long fattening duration on tenderness was reported[16]. Nevertheless, it is worthwhile to note that other factors such as age at slaughter and life weight of the animals, are confounded in this effect[28].

Initial body weight of the animals at the entrance of fattening period was retained as a splitter. This trial highlighted that an initial body weight threshold (~420 kg), *i.e.*, lower is better, is a good splitter of beef tenderness. As there were differences in initial body weight between clusters (**Table S2**), it allowed in this study to obtain differences at the end of the fattening period at both carcass and meat levels. Initial body weight is one of the most important factors affecting efficiency and profitability in beef cattle production[29]. Earlier studies had linked this factor with carcass characteristics[15, 16]. In addition, young cattle are known to have better feed conversion efficiency than older animals. In fact, as fat deposition is more energetic expensive than that of proteins, the initial body weight (*i.e.*, in wet tissue basis) of cattle was already considered before beginning feed.

Generally and from most described studies in the large literature, animals are being presented for slaughter when they achieved the desirable age and carcass characteristics. In many of the reported studies, feeding effects are further confounded with animal age and type, growth rate or carcass weight at slaughter. The large amount of literature, which considers the

effect of feeding on beef quality, has compared, almost exclusively, forage-based with grain-based diets[30]. In this study, the animals were managed under experimental conditions in order to avoid extreme effects or divergences between the animals. There is evidence and in agreement with our findings that concentrate-fed bulls produce more tender meat than forage-fed animals[31]. For example, earlier studies showed that increasing the concentrate intake of the animals increased the carcass fatness score[32] and thus intramuscular fat content. Accordingly, this positive effect on beef palatability was further linked to carcass fat deposition[33]. On another hand and in line with our findings, previous studies reported negative effects of forage finishing on beef tenderness[34]. Meanwhile, it is important to note that when forage-finished cattle have been compared with concentrate-finished cattle at a similar end-point such as final body weight, or degree of fattening, meat quality differences between diets were minimized[35].

DMI is linearly related with the composition of the diet[36]. This was among the rearing factors that were retained in the equations models of the percentages of fat and muscle carcass tissues[16]. This result highlights the impact of feeding, supported in this study by the strong splitting power of concentrate per cent needed to achieve high tenderness quality. This further agrees with the reduction of DMI when the proportion of concentrate per cent in the diet was increased. In the same trend than this work, an earlier study reported similar findings in cattle reared to evaluate the impact of fattening feeding in such a way to improve carcass and meat eating quality[28]. Thus, we suggest addressing in details the role that DMI would play in relation with meat tenderness (within forage and/or concentrate diet-type), especially since it is a meaningful factor affecting animal performance.

Carcass characteristics as predictors (splitters) of meat tenderness clusters

We have recently identified a set of indicators to better characterize beef carcasses at the slaughterhouse level in addition to the EUROP system[14]. This study, confirms well that a combination of factors based or not on the EUROP system would be helpful to discriminate soon after slaughter the tenderness potential of carcasses. Fatness carcass score and muscle carcass % were retained in the overall decision tree as good splitters. Thus, the hypothesis that among carcass characteristics some of them would have an impact on the palatability of beef cuts was very strongly supported by the results of this trial.

Fatness carcass score was the first best attribute splitting the young bull's carcasses. If the fatness carcass score was higher than 2.88, the meat was tenderer. This finding agrees well with the important role of fat in the palatability of beef[37]. This confirms the proposal to use the

subjective grading of carcasses by their EUROP fat cover scores as a threshold in carcasses segregation[38] and hence, in this study for tenderness categorization. Our findings evidence the power of fatness scoring to categorize beef sensory quality. Fatness score known also as marbling was developed as an indicator of quality and palatability as it describes the amount and distribution of visible intramuscular fat[39]. Although the chemically extracted fat content is not exactly the same as fatness score due to the fact that some invisible fat deposits that cannot be seen by visual appraisal can also be determined by chemical analysis, the relationship between them has been reported as linear[39]. The involvement of fat in meat tenderness would be explained by the role that it would play during chewing though lubrication[37], stating that fat present in and around the muscle fibers and perimysium during mastication results in a tenderer beef[37, 40].

The studies that have linked muscle carcass per cent with meat tenderness are very scarce. In this study, an increase of muscle carcass was at the expense of the meat tenderness. In all cases, these would to some extent related to muscle protein turnover, known to affect the final meat quality[41] with a strong dependency on energy intake and protein supply during fattening. Carcass weight, closely related with lean carcass per cent[42] was reported to affect meat quality. For example, it is known to affect in a curvilinear fashion and negatively the final muscle pH[43]. In addition, carcass weight and pre-slaughter growth rate have been shown to alter meat quality, specifically tenderness[44]. In the present experiment, increasing concentrate proportion decreased the carcass muscle and improved fat content, consistent with previous studies[45].

Conclusion

This study showed the usefulness of the used set of methodologies when dealing with rearing factors (from farm) or carcass characteristics (to the slaughterhouse) of animals that may have experienced divergent rearing management practices. At the same time, considering that both statistical approaches (clustering and decision tree) led independently to similar results, the set of attributes (splitters) selected by the decision tree process may be considered objective and therefore reliable for classifying carcasses according to their tenderness potential. Overall, in the classifications performed with the whole dataset, the last decision tree selected more rearing factors unambiguously and the classification/regression accuracy was improved in comparison to the others. This provides evidence and in agreement to our proposal, the usefulness of taking in account information available during the fattening period of the animal to predict its final meat quality, herein tenderness. Furthermore, the present results are valid from a practical point of view as they can be easily used by the various actors involved in the

meat sector, whatever their position. To sum up concerning the take home message of this study, a ribeye steak of the studied animals was considered tender (TEND+) if it matched the following rule:

IF (fatness carcass score ≥2.88) **AND** (concentrate ≥ 82%) [**AND** (muscle carcass ≥71%)] **THEN** meat was [very] tender.

Acknowledgments

The authors acknowledge the financial support given to M.G from project S3-23000846 funded by Région Auvergne-Rhône-Alpes and FEDER. This study was designed to support the FBEA project (Filière Bovins Engraissement Auvergne-Rhône-Alpes) in collaboration with ARIA-ARA (Association Régionale des Industries Agro-Alimentaires d'Auvergne-Rhône-Alpes) and Chambre Régionale d'Agriculture d'Auvergne-Rhône-Alpes. The authors convey special thanks to the different INRA research units, namely INRA-UE232 (Bourges) for their help in animal handling and data collection and to all the involved persons especially Dr. Gilles Renand. The authors thank INRA Herbipôle for animal management and slaughtering. The authors would like to thank the two anonymous reviewers for their helpful and constructive comments that have greatly helped improve the presentation of this article.

References

1. Destefanis G, Brugiapaglia A, Barge MT and Dal Molin E, Relationship between beef consumer tenderness perception and Warner-Bratzler shear force. *Meat science* **78**:153-156 (2008).
2. Ferguson DM, Bruce HL, Thompson JM, Egan AF, Perry D and Shorthose WR, Factors affecting beef palatability - farmgate to chilled carcass. *Australian Journal of Experimental Agriculture* **41**:879-891 (2001).
3. Koohmaraie M, Biochemical factors regulating the toughening and tenderization processes of meat. *Meat science* **43**:193-201 (1996).
4. Gagaoua M, Picard B and Monteils V, Associations among animal, carcass, muscle characteristics, and fresh meat color traits in Charolais cattle. *Meat science* **140**:145-156 (2018).
5. Gagaoua M, Picard B, Soulat J and Monteils V, Clustering of sensory eating qualities of beef: Consistencies and differences within carcass, muscle, animal characteristics and rearing factors. *Livestock Science* **214**:245-258 (2018).
6. Picard B and Gagaoua M, Chapter 11 - Proteomic Investigations of Beef Tenderness, in *Proteomics in Food Science: from farm to fork*. Academic Press, pp 177-197 (2017).
7. Ouali A, Gagaoua M, Boudida Y, Becila S, Boudjellal A, Herrera-Mendez CH and Sentandreu MA, Biomarkers of meat tenderness: present knowledge and perspectives in regards to our current understanding of the mechanisms involved. *Meat science* **95**:854-870 (2013).
8. England EM, Scheffler TL, Kasten SC, Matarneh SK and Gerrard DE, Exploring the unknowns involved in the transformation of muscle to meat. *Meat science* **95**:837-843 (2013).
9. Picard B, Gagaoua M, Micol D, Cassar-Malek I, Hocquette JF and Terlouw CE, Inverse relationships between biomarkers and beef tenderness according to contractile and metabolic properties of the muscle. *J Agric Food Chem* **62**:9808-9818 (2014).
10. Miller MF, Carr MA, Ramsey CB, Crockett KL and Hoover LC, Consumer thresholds for establishing the value of beef tenderness. *J Anim Sci* **79**:3062-3068 (2001).
11. Berri C, Picard B, Lebret B, Andueza Urra JD, Vautier A, Chartrin P, Beauclercq S, Lefèvre F, Legrand I and Hocquette J-F, Predicting quality of meats: myth or reality? , in *16 Journées Sciences du Muscle et Technologies des Viandes*, Ed. ADIV, Paris, p n.p. (2016).

12. Hocquette JF, Botreau R, Legrand I, Polkinghorne R, Pethick DW, Lherm M, Picard B, Doreau M and Terlouw EMC, Win–win strategies for high beef quality, consumer satisfaction, and farm efficiency, low environmental impacts and improved animal welfare. *Anim Prod Sci* **54**:1537-1548 (2014).
13. Polkinghorne RJ and Thompson JM, Meat standards and grading: A world view. *Meat science* **86**:227-235 (2010).
14. Monteils V, Sibra C, Ellies-Oury MP, Botreau R, De la Torre A and Laurent C, A set of indicators to better characterize beef carcasses at the slaughterhouse level in addition to the EUROP system. *Livestock Science* **202**:44-51 (2017).
15. Oury MP, Agabriel J, Agabriel C, Micol D, Picard B, Blanquet J, Labouré H, Roux M and Dumont R, Relationship between rearing practices and eating quality traits of the muscle rectus abdominis of Charolais heifers. *Livestock Science* **111**:242-254 (2007).
16. Soulat J, Picard B, Leger S and Monteils V, Prediction of beef carcass and meat traits from rearing factors in young bulls and cull cows. *J Anim Sci* **94**:1712-1726 (2016).
17. Gagaoua M, Monteils V, Couvreur S and Picard B, Identification of Biomarkers Associated with the Rearing Practices, Carcass Characteristics, and Beef Quality: An Integrative Approach. *J Agric Food Chem* **65**:8264-8278 (2017).
18. Gagaoua M, Micol D, Picard B, Terlouw CE, Moloney AP, Juin H, Meteau K, Scollan N, Richardson I and Hocquette JF, Inter-laboratory assessment by trained panelists from France and the United Kingdom of beef cooked at two different end-point temperatures. *Meat science* **122**:90-96 (2016).
19. Gagaoua M, Bonnet M, Ellies-Oury MP, De Koning L and Picard B, Reverse phase protein arrays for the identification/validation of biomarkers of beef texture and their use for early classification of carcasses. *Food Chemistry* **250**:245-252 (2018).
20. Gagaoua M, Micol D, Richardson RI, Hocquette JF, Terlouw EMC, Meteau K, Juin H, Moloney AP, Nuernberg K, Scollan ND, Boudjellal A and Picard B, Relationships between overall liking score and sensory meat attributes in different types of beef cattle, in *Proceedings of the 59th International Congress of Meat Science and Technology*, Ed, Izmir, Turkey, p 4 (2013).
21. Likas A, Vlassis N and J. Verbeek J, The global k-means clustering algorithm. *Pattern Recognition* **36**:451-461 (2003).
22. Valletta JJ, Torney C, Kings M, Thornton A and Madden J, Applications of machine learning in animal behaviour studies. *Animal Behaviour* **124**:203-220 (2017).
23. Gagaoua M, Terlouw EM, Boudjellal A and Picard B, Coherent correlation networks among protein biomarkers of beef tenderness: What they reveal. *J Proteomics* **128**:365-374 (2015).
24. Gagaoua M, Bonnet M, De Koning L and Picard B, Reverse Phase Protein array for the quantification and validation of protein biomarkers of beef qualities: The case of meat color from Charolais breed. *Meat science* **145**:308-319 (2018).
25. Rousseeuw PJ, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**:53-65 (1987).
26. Ture M, Tokatli F and Kurt I, Using Kaplan–Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4.5 and ID3) in determining recurrence-free survival of breast cancer patients. *Expert Systems with Applications* **36**:2017-2026 (2009).
27. Breiman L, Friedman J, Stone CJ and Olshen RA, *Classification and regression trees*. CRC press (1984).
28. Vestergaard M, Madsen NT, Bligaard HB, Bredahl L, Rasmussen PT and Andersen HR, Consequences of two or four months of finishing feeding of culled dry dairy cows on carcass characteristics and technological and sensory meat quality. *Meat science* **76**:635-643 (2007).

29. Koknaroglu H, Loy DD and Hoffman MP, Effect of housing, initial weight and season on feedlot performance of steers in Iowa. *South African Journal of Animal Science* **35**:282-290 (2005).
30. Patterson DC, Moore CA and Steen RWJ, The effects of plane of nutrition and slaughter weight on the performance and carcass composition of continental beef bulls given high forage diets. *Animal Science* **58**:41-47 (2010).
31. Moloney AP, Mooney MT, Kerry JP and Troy DJ, Producing tender and flavoursome beef with enhanced nutritional characteristics. *Proceedings of the Nutrition Society* **60**:221-229 (2007).
32. Keane MG, Drennan MJ and Moloney AP, Comparison of supplementary concentrate levels with grass silage, separate or total mixed ration feeding, and duration of finishing in beef steers. *Livestock Science* **103**:169-180 (2006).
33. Alfaia CPM, Alves SP, Martins SIV, Costa ASH, Fontes CMGA, Lemos JPC, Bessa RJB and Prates JAM, Effect of the feeding system on intramuscular fatty acids and conjugated linoleic acid isomers of beef cattle, with emphasis on their nutritional value and discriminatory ability. *Food Chemistry* **114**:939-946 (2009).
34. Mitchell GE, Reed AW and Rogers SA, Influence of Feeding Regimen on the Sensory Qualities and Fatty Acid Contents of Beef Steaks. *Journal of Food Science* **56**:1102-1103 (1991).
35. Bidner TD, Schupp AR, Montgomery RE and Carpenter JC, Acceptability of Beef Finished on All-Forage, Forage-Plus-Grain or High-Energy Diets. *Journal of Animal Science* **53**:1180-1187 (1981).
36. Juniper DT, Browne EM, Fisher AV, Bryant MJ, Nute GR and Beever DE, Intake, growth and meat quality of steers given diets based on varying proportions of maize silage and grass silage. *Animal Science* **81**:159-170 (2007).
37. Savell J and Cross H, The role of fat in the palatability of beef, pork, and lamb. *Designing foods: Animal product options in the marketplace*:345-355 (1988).
38. Varona L, Moreno C and Altarriba J, A model with heterogeneous thresholds for subjective traits: Fat cover and conformation score in the Pirenaica beef cattle1. *Journal of Animal Science* **87**:1210-1217 (2009).
39. Dow DL, Wiegand BR, Ellersieck MR and Lorenzen CL, Prediction of fat percentage within marbling score on beef longissimus muscle using 3 different fat determination methods. *J Anim Sci* **89**:1173-1179 (2011).
40. Gagaoua M, Terlouw EMC, Micol D, Hocquette JF, Moloney AP, Nuernberg K, Bauchart D, Boudjellal A, Scollan ND, Richardson RI and Picard B, Sensory quality of meat from eight different types of cattle in relation with their biochemical characteristics. *Journal of Integrative Agriculture* **15**:1550-1563 (2016).
41. Oksbjerg N and Therkildsen M, Chapter 3 - Myogenesis and Muscle Growth and Meat Quality A2 - Purslow, Peter P, in *New Aspects of Meat Quality*. Woodhead Publishing, pp 33-62 (2017).
42. Wagner JJ, Lusby KS, Oltjen JW, Rakestraw J, Wettemann RP and Walters LE, Carcass Composition in Mature Hereford Cows: Estimation and Effect on Daily Metabolizable Energy Requirement During Winter. *Journal of Animal Science* **66**:603-612 (1988).
43. McGilchrist P, Alston CL, Gardner GE, Thomson KL and Pethick DW, Beef carcasses with larger eye muscle areas, lower ossification scores and improved nutrition have a lower incidence of dark cutting. *Meat science* **92**:474-480 (2012).
44. Spanier AM, McMillin KW and Miller JA, Enzyme Activity Levels in Beef: Effect of Postmortem Aging and End-point Cooking Temperature. *Journal of Food Science* **55**:318-322 (1990).

45. Manninen M, Jauhiainen L, Ruusunen M, Soveri T, Koho N and Pösö R, Effects of concentrate type and level on the performance and health of finishing Hereford bulls given a grass silage-based diet and reared in cold conditions. *Livestock Science* **127**:227-237 (2010).

Table 1. Accuracy of classifications and prediction of tender (TEND+), medium (TEND=) and tough (TEND–) beef clusters by decision trees using the rearing factors and/or carcass characteristics compared to those of *k*-means clustering based on tenderness scores with 10-fold cross validation.

Clustering method / Tenderness classes	*k*-means clusters	Decision tree 1[†]	Decision tree 2[‡]	Decision tree 3[§]
TEND+	**59**	52	51	60
TEND=	**141**	164	150	149
TEND–	**108**	92	107	99
Classification accuracy (%) [¶]		70.78	67.21	84.41
Sensitivity (weighted average)		0.710	0.681	0.834
Specificity (weighted average)		0.289	0.319	0.165

[†] Decision tree was built using rearing factors.

[‡] Decision tree was built using carcass characteristics.

[§] Decision tree was built using both the rearing factors and carcass characteristics.

[¶] "%" of predictions that are correct [*Accuracy* = (TN +TP)/(TN+TP+FN+FP)]. TP, FN, TN, and FP denote the numbers of true positives, false negatives, true negatives, and false positives, respectively.

Figure captions

Fig. 1. Silhouette widths (S_i) plots for *k*-means on the left and HCA on the right. For the two clustering methods, three clusters we retained as best results and validated using *Si* and Davies-Bouldin indices in all cases.

Fig. 2. C&RT decision tree based on rearing factors for the classification and prediction of tenderness beef cuts firstly identified using *k*-means clustering on the 308 young bulls. The distribution of the animals in each tenderness cluster were used for accuracy measurement. At the beginning of the C&RT method, all of the data are concentrated at a root node located at the top of the tree. Then this latter is divided into two child nodes on the basis of an independent variable (splitter), herein concentrate (%), that creates the best homogeneity. The cut‐off value of each dividing splitter was calculated from the data of all the subjects, using the Gini index. In fact, the data in each child node are more homogenous than those in the upper parent node. This process is continued repeatedly for each child node until all of the data in each node have the greatest possible homogeneity. This node is called a terminal node or "leaf" and has no branches. Then, each tenderness cluster is identified by the gradient color meaning TEND– (Tough, with blue color); TEND= (Medium, with green color and TEND+ (Tender, with orange to red color).

Fig. 3. C&RT decision tree based on carcass characteristics for the classification and prediction of tenderness meat cuts firstly identified using *k*-means clustering on the 308 young bulls. The distribution of the animals in each tenderness cluster were used for accuracy measurement. At the beginning of the C&RT method, all of the data are concentrated at a root node located at the top of the tree. Then this latter is divided into two child nodes on the basis of an independent variable (splitter), herein fat carcass score, that creates the best homogeneity. The cut‐off value of each dividing splitter was calculated from the data of all the subjects, using the Gini index. In fact, the data in each child node are more homogenous than those in the upper parent node. This process is continued repeatedly for each child node until all of the data in each node have the greatest possible homogeneity. This node is called a terminal node or "leaf" and has no branches. Then, each tenderness cluster is identified by the gradient color meaning TEND– (Tough, with blue color); TEND= (Medium, with green to yellow color and TEND+ (Tender, with red color).

Fig. 4. C&RT decision tree based on both rearing practices factors and carcass characteristics for the classification and prediction of tenderness meat cuts firstly identified using *k*-means

clustering on the 308 young bulls. The distribution of the animals in each tenderness cluster were used for accuracy measurement. At the beginning of the C&RT method, all of the data are concentrated at a root node located at the top of the tree. Then this latter is divided into two child nodes on the basis of an independent variable (splitter), herein fat carcass score, that creates the best homogeneity. The cut‐off value of each dividing splitter was calculated from the data of all the subjects, using the Gini index. In fact, the data in each child node are more homogenous than those in the upper parent node. This process is continued repeatedly for each child node until all of the data in each node have the greatest possible homogeneity. This node is called a terminal node or "leaf" and has no branches. Then, each tenderness cluster is identified by the gradient color meaning TEND– (Tough, with blue color); TEND= (Medium, with green to yellow color and TEND+ (Tender, with orange to red color).

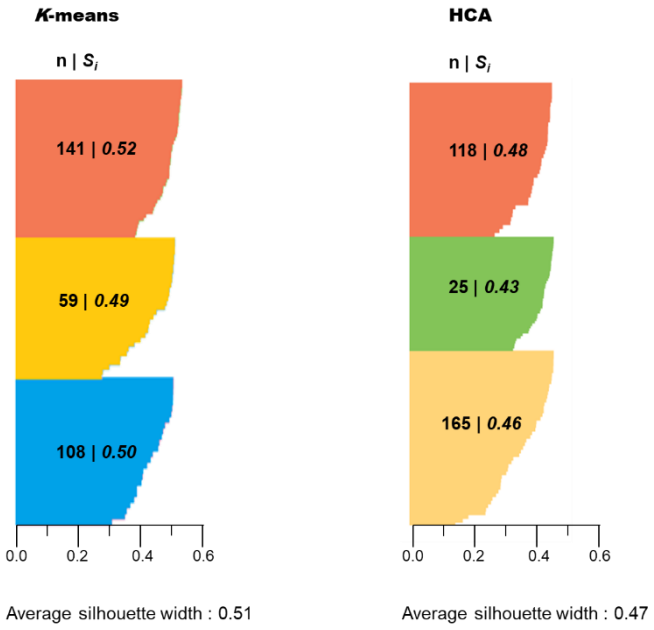

Fig. 1.

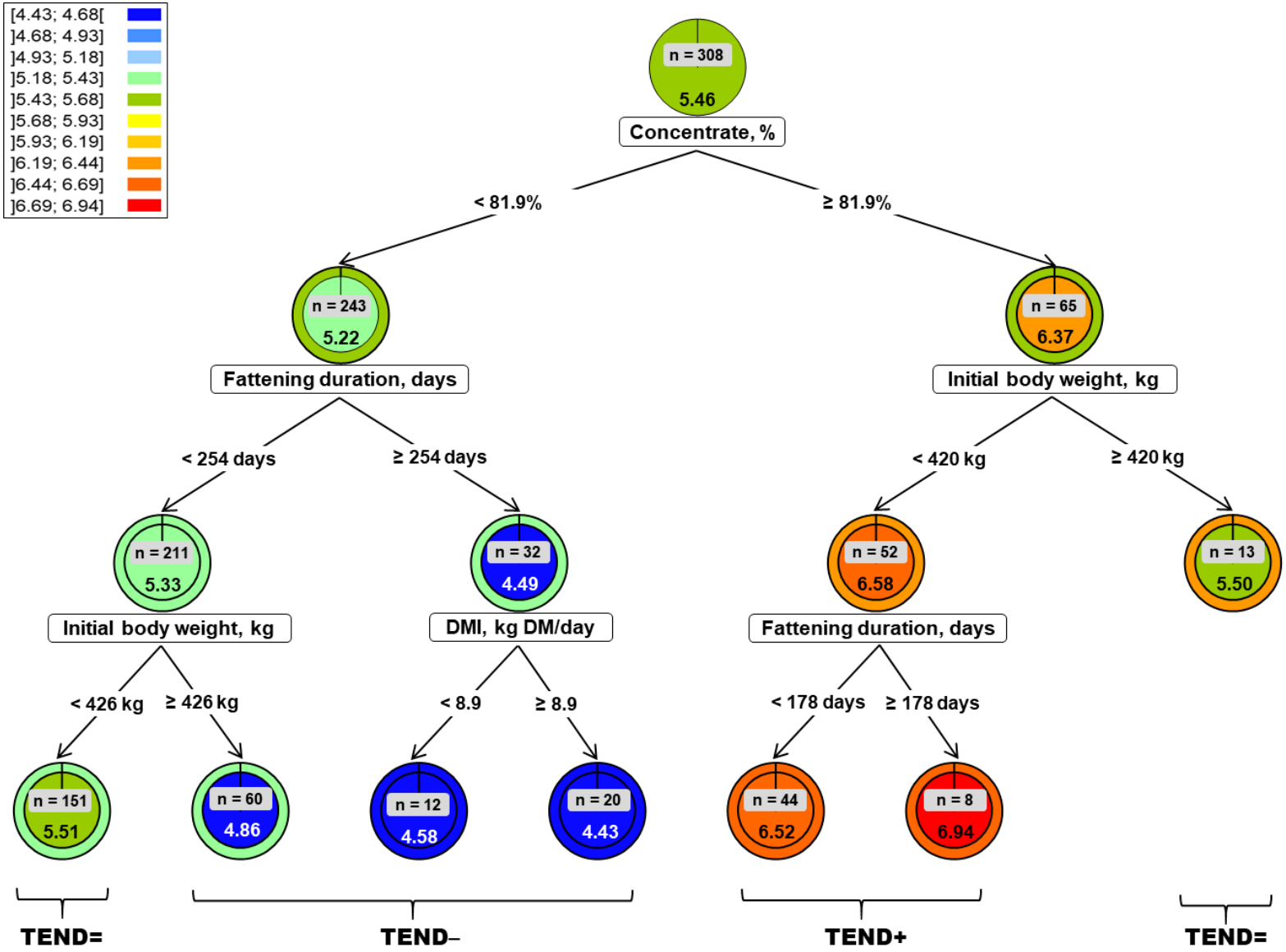

Fig. 2.

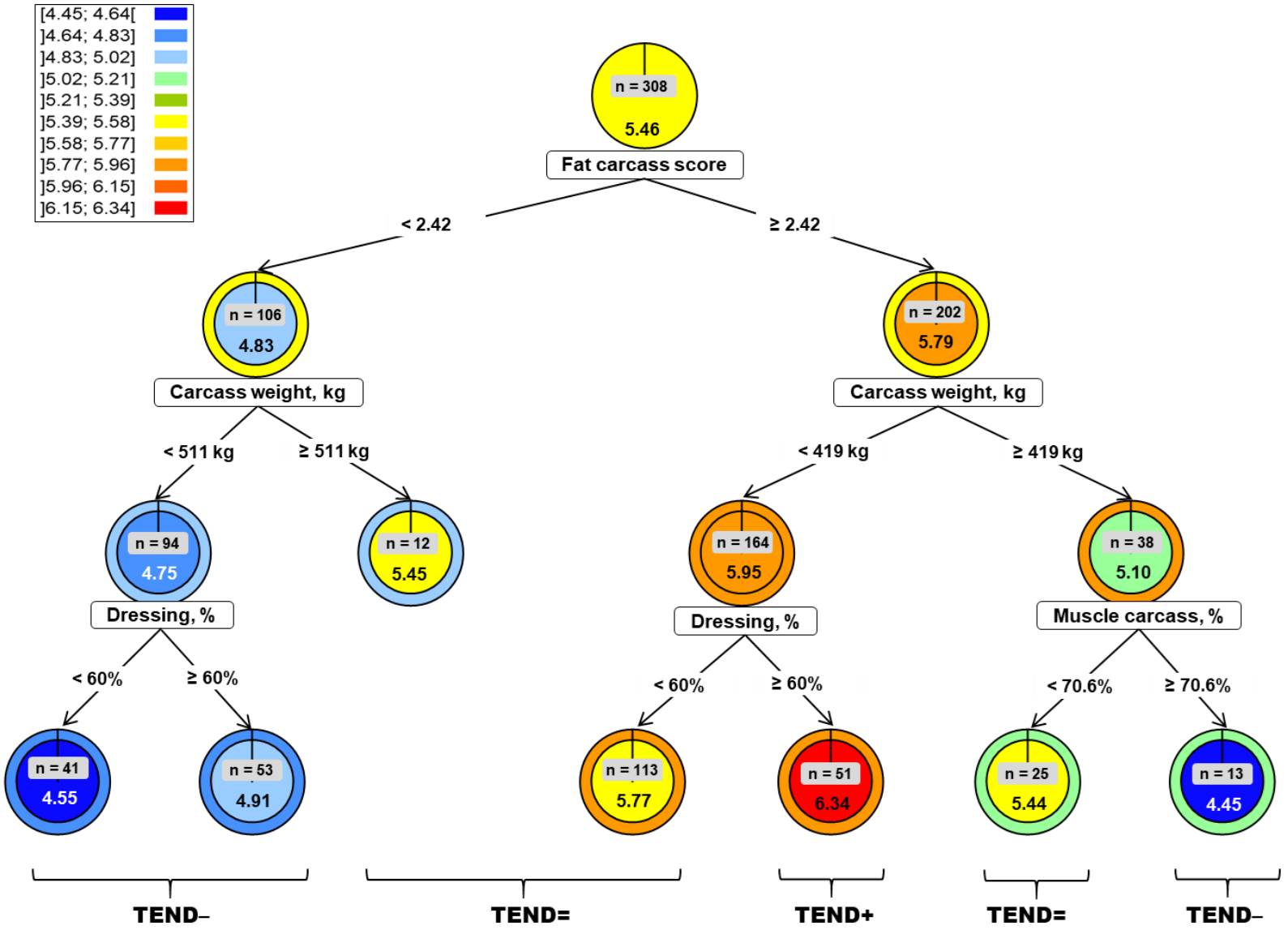

Fig. 3.

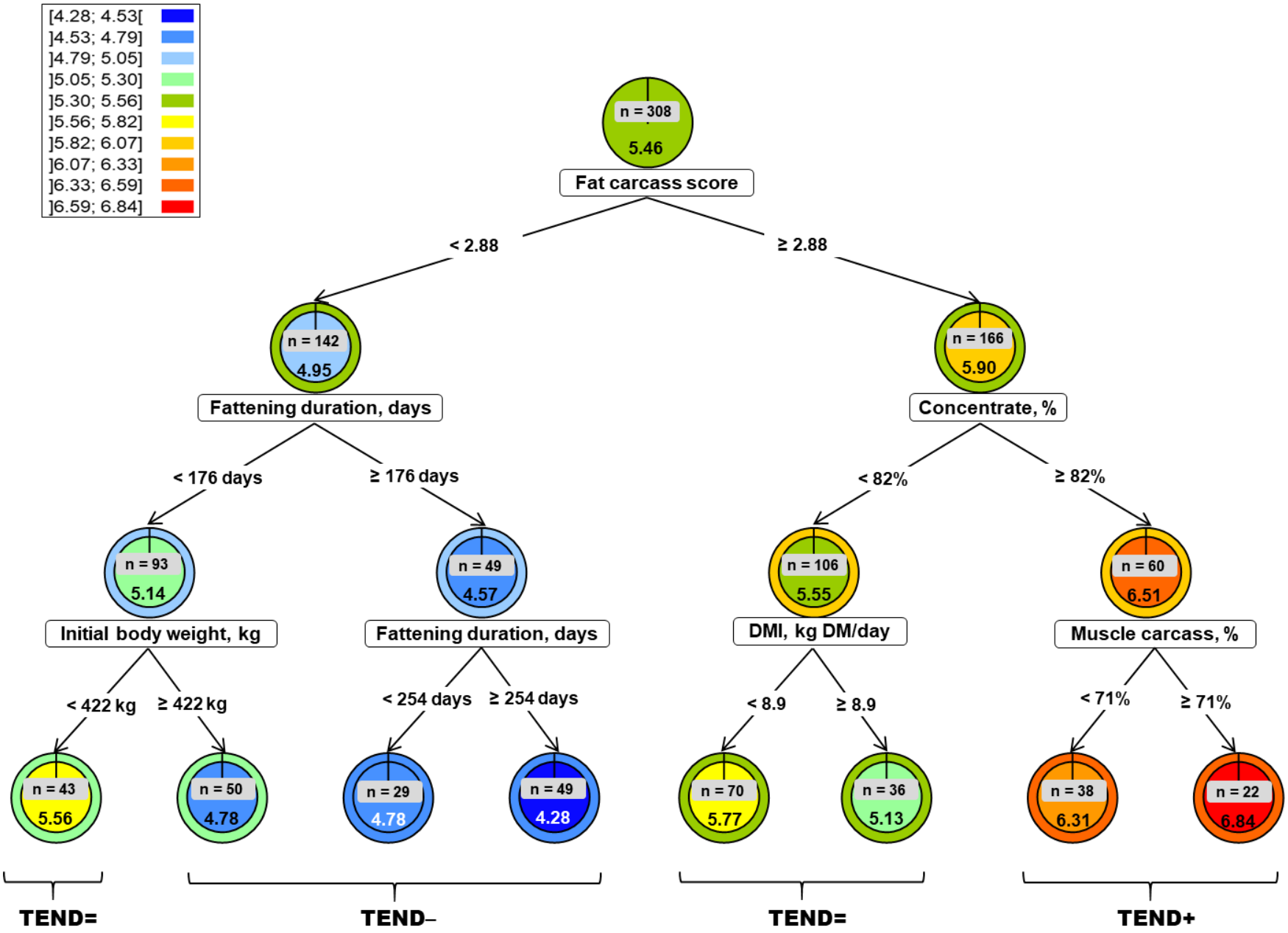

Fig. 4.