Research in Beef Tenderness and Palatability in the Era of Big Data

Jean-François Hocquette¹, Marie-Pierre Ellies-Oury¹,², Isabelle Legrand³, David Pethick⁴, Graham Gardner⁴, Jerzy Wierzbicki⁵, and Rod Polkinghorne⁶

¹INRAE, Université Clermont Auvergne, VetAgroSup, UMR1213, Recherches sur les Herbivores, F-63122 Saint-Genès Champanelle, France
²Bordeaux Sciences Agro, Département Feed and Food, 1, cours du Général de Gaulle, CS 40201, 33175 Gradignan Cedex, France
³Institut de l’Elevage, Service Qualité des Carcasses et des Viandes, MRA-NA, 87060 Limoges Cedex 2, France
⁴School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia
⁵International Meat Research 3G Foundation, Polish Beef Association Ul. Smulikowskiego 4, 00-389 Warszawa, Poland
⁶International Meat Research 3G Foundation, 431 Timor Road, Murrurundi, NSW 2338, Australia

*Corresponding author. Email: jean-francois.hocquette@inrae.fr (Jean-François Hocquette)

Abstract: For decades, research has focused on predicting beef palatability using muscle biochemical traits and various biomarkers. In these approaches, a precise definition of the variable to predict (tenderness assessed by panelists, untrained consumers, or shear force) and repeatability of the measurements are crucial for creating significant data resources for the derivation of robust predictive models, as well as rigorous validation testing. This “big data” approach as well as the rapid evolution of new measurement technologies require careful definition of traits and transparent principles for data sharing and management. In this context, AgroPortal is a common platform to receive and host existing description systems, such as Animal Trait Ontology for Livestock. Furthermore, as in other fields, meat science researchers should improve the Findability, Accessibility, Interoperability, and Reusability of data (known as the FAIR Data Principles). For beef, strategic choices have been made in order to consider real consumers’ expectations, not well estimated correctly by lab approaches. Thus, Australia set up the Meat Standards Australia grading scheme based on beef assessment by untrained consumers. The ambitions of the International Meat Research 3G Foundation is to develop beef ontology, to set up an international database with a large number of consumers’ scores related to beef palatability and collected according to standard protocols of the Meat Standards Australia methodology. The Foundation also aims to support the beef industry by offering an international predictive model of beef palatability, flexible enough to take into account any local livestock characteristics or regional consumer specificity. This approach is supported by the United Nations Economic Commission for Europe, which promotes development of regulations and norms, technical cooperation, and exchange of best expertise and practices. This will substantially improve the transparency of data flow and price signaling between all participants of the value chain, from beef producers through to consumers at retail.

Key words: beef, tenderness, palatability, open science, standards, ontology

Introduction

Global interest in beef eating quality is not new; concern related to declining consumption—which is a hot topic in the United States, Australia, and Europe—has been present since at least the early 1990s. Growing recognition of the need to deliver a contemporary consumer product rather than a variable price-focused commodity led to extensive research in many countries aimed at better understanding the consumer issues and effective response. Significant early activity included
the introduction of the Beef Quality Audit and the concept of Palatability Assured Critical Control Point pathways in the US, as well as the introduction of the Meat and Livestock Commission Blueprint in the United Kingdom. Despite a great deal of research or initiatives to control beef eating quality, it is still difficult for the supply chain to guarantee consistent good-quality beef to the consumer. This is one major reason for dissatisfaction, which could influence the decrease in purchase intentions due to unsatisfactory eating experiences. Generally, consumers welcome the concept of a beef eating-quality guarantee system, but to create one, we need appropriate tools (Verbeke et al., 2010). Therefore, predicting eating quality (especially tenderness and flavor) at the consumer level has been the subject of active research with different strategies thanks to technical availability of new tools (such as genomics or robotics). Research has been mainly focused on instrumental tenderness, since flavor and other components of palatability are difficult to replicate through laboratory approaches. While it is understandable, the concentration on tenderness at the expense of other traits—in particular flavor—may have reduced the relevance to consumer satisfaction. Much consumer data indicate that tenderness accounts for 30% or less of the eating experience in, e.g., French (Legrand et al., 2013), Japanese (Polkinghorne, 2007), Irish (McCarthy et al., 2017), and US (O’Quinn et al., 2018) studies.

This review aims to describe the major strategies that have been developed to predict beef eating quality. In all of them, new issues appeared such as those related to standardization of methods, a prerequisite for big data approaches and modelling. Therefore, a worldwide strategy is required to develop interoperable data standards in order to share research efforts and research data. This is a prerequisite for modelling approaches with as much comparable data as possible to compensate for technical and biological variability of beef eating quality in predictive models. In addition, common standards across countries will facilitate beef trade and probably consumer satisfaction in any country. These are the objectives of the International Meat Research 3G Foundation and of the United Nations Economic Commission for Europe (UNECE).

Prediction of Beef Palatability Using Muscle Biochemical Traits, Their Genetic Markers, and Other Biomarkers

The scientific strategy

Currently, there is no simple technique to evaluate beef tenderness from carcasses (and/or live animals). This quality trait can only be assessed after slaughtering and processing of the muscle into meat, either by panelists or consumers for sensory analysis, and/or by mechanical measurements. These methods are constraining and costly, cannot be used routinely, and do not allow for optimal management and/or prediction of meat tenderness. In order to find alternative solutions, various research programs—at the national and international levels—have targeted genetic or biological markers for quality traits. From “expert” prediction (Ellies-Oury et al., 2013) and biochemical studies (Chriki et al., 2012) to functional genomics developments (Picard et al., 2015), biomarkers have been or are being explored in order to develop tools that can provide objective data to describe the tenderness potential of live animals and specific cuts (Berri et al., 2019) for beef.

Reference methods for eating quality of beef

Meat tenderness is often evaluated by sensory evaluation with panelists, which is a direct (and reference) method. It is also often estimated through its opposite, toughness. Indeed, mechanical methods using rheological instruments aim to simulate the resistance of meat cutting during chewing and thus to evaluate the hardness of the meat. The measured value represents the force required to shear a piece of meat perpendicular to the muscle fibers. These measurements can be made on raw or cooked meat, using a Warner-Bratzler cell (Bratzler, 1949; Warner, 1952) or a Salé cell (Salé, 1971). Nevertheless, the constraints involved in these evaluation methods (in particular their high cost, their moderate average repeatability, and the need for a significant quantity of meat to be scored) have led scientists to seek less invasive meat quality indicators (in particular of tenderness) that could be used as early predictors, especially before slaughtering.

It seems that the parallel implementation of consumer and expert panels could help to clarify the link between biomarkers and sensory scores. Because consumers are key actors in the generation of the value of bovine carcasses, it could be more appropriate to develop assessment of beef eating quality by untrained consumers, which has so far rarely been investigated, at least in Europe.

Biochemical muscle traits

To evaluate and characterize meat quality, with a view to predicting it, research was first conducted on understanding the biochemical properties of muscles
and their impacts on the sensory properties of meat (reviewed by Hocquette et al., 2012b, 2014). In the literature, a great number of studies have described relationships between tenderness and biochemical muscle traits, including intramuscular fat content and muscle fiber characteristics as well as contents of total and soluble collagen. In particular, it was initially indicated that it was possible to explain up to 30% of the variability in the tenderness of the longissimus thoracis muscle from Charolais young bulls by its biochemical traits (Renand et al., 2001). Other results, obtained with the rectus abdominis muscle (Oury et al., 2009, 2010), indicated that the biochemical properties of the muscle explained between 12% and 23% of the variability of the sensory descriptors of this muscle evaluated by panelists. However, thanks to a meta-analysis with much more data from different breeds, different cuts, and different sources, Chriki et al. (2013) concluded that the variability in tenderness that could be explained by some muscle biochemical properties is, in reality, limited (i.e., less than 6%). Furthermore, the contribution of the different biochemical traits to the sensory quality of meat varies according to the cut and to the animal type (Chriki et al., 2013). This underlines the necessity to analyze as much data as possible to get robust results and to avoid any bias due to limited data-sets restricted, e.g., to one cut or to one breed. However, when various experiments were not designed initially to ultimately link together, it appeared very difficult to integrate data from these experiments, which differ by units, scales, or laboratory methods. Ontology will help to address these issues (Hocquette et al., 2011).

Furthermore, results from research conducted with shear force or sensory scores by panelists as reference methods are not always confirmed when sensory scores from untrained consumers are used. Indeed, with scores from untrained consumers, only fat or moisture contents of muscles have the capacity to add any precision in a commercial eating-quality prediction model including muscle type (Bonny et al., 2015). These results raised the difficulties of evaluating the tenderness of beef with different methods.

**Other muscle biomarkers**

In the previous works, authors hypothesized that knowledge about biochemical properties of muscles of living animals could be mobilized to understand and explain the variability in beef tenderness. However, thanks to the previously mentioned unsuccessful results with these muscle biochemical traits alone, it seemed particularly useful to include more indicators of tenderness such as physicochemical changes related to post-mortem transformations of muscle fibers (Figure 1).

**Figure 1.** Contribution of muscle characteristics to the sensory quality of meat (color, juiciness, tenderness, and flavor liking).
Other authors (Whipple et al., 1990a, 1990b; El Masry et al., 2012) have proposed models that explained 76% and 83%, respectively, of the variability in meat toughness. These models were using (1) muscle characteristics related to maturation kinetics (protease inhibitor, pH) and fiber surfaces, or (2) near-infrared spectrometry data. However, they have been established on a small number of animals and/or muscles (n < 30) and are therefore difficult to generalize.

Research on muscle biochemical traits (which were mainly studied in the years 1990–2000 by classic biochemical methods) has been gradually supplemented by deeper developments in knowledge of muscle metabolism and cellular functioning thanks to the advent of more powerful analytical techniques (including the development of “omics”). This allowed new biomarkers to be taken into consideration. Thus, prediction models have been established using gene expression levels (Hocquette et al., 2012a) assessed by transcriptomics or protein biomarkers to predict toughness (Guillemin et al., 2011; Ouali et al., 2013) and/or sensory properties of meat. The equations established indicated that it is possible to explain, in the best cases, between 81% and 87% of the variability in tenderness or toughness of beef. However, as for biochemical traits, it is rarely possible to generalize these results because biomarkers of beef tenderness are often specific to an animal type (steer or young bull) or to environmental conditions (Hocquette et al., 2012a). Furthermore, the choice of biomarkers of interest appears to be muscle dependent (Guillemin et al., 2011), the equations being also dependent on the breed of the studied animals (Gagaoua et al., 2019). It is therefore important to include “cut/muscle” in the statistical analysis since biochemical traits have been shown to explain consumer acceptance across cuts or muscles but not within one muscle (Bonny et al., 2015).

Although the identification of biomarkers seems promising for some authors, there is still a long way to go. Indeed, recent studies showed that the associations between proteomic biomarkers and beef tenderness depend not only on animals, breeds, and livestock farming systems but also on how sensory scores by panelists are determined, which depend on the endpoint cooking temperature and on the country of origin of the panelists (Gagaoua et al., 2019). This raises, again, the issue of the reference methods to be used in such studies.

Recent proteomic analyses allowed the identification of a list of proteins considered as potential biomarkers of beef tenderness (Poulson et al., 2012; Gagaoua et al., 2015b; Picard et al., 2018). Nevertheless, such studies are scarce, relatively recent, and often performed with a moderate or relatively small number of samples (Ouali et al., 2013; Picard et al., 2014, 2015; Gagaoua et al., 2015a, 2017). Concerning tenderness, several candidate biomarkers have been identified (Picard and Gagaoua, 2017) and associated with apoptosis, oxidative stress, cytoskeletal proteins, proteolysis, oxidative metabolism, glycolytic metabolism, heat shock proteins, and transport and signaling (Picard and Gagaoua, 2017). They remain to be confirmed at a large scale with numerous data, which should be heterogeneous (in terms of muscles, breeds, sex, etc.) for proper and robust meta-analysis.

Norms, Ontologies, Data Sharing, and Repeatability of Methods

The international new scientific context

Both the amount of data available and our ability to analyze it have greatly increased in the last decade. The increased availability of data is a fact in genetics and genomics thanks to high-throughput techniques, but despite important and recent progress, it is less obvious for phenotypic data, such as beef sensory traits. In parallel, new analytical techniques based on improved statistical methods as well as various forms of algorithms and of machine learning have revolutionized data analysis. Consequently, accurate prediction of biological phenomena has become feasible for the first time, at least in theory, thanks to the big data concept, which is perceived as the next frontier for innovation. For instance, it may increase the fit between consumers’ preferences and product features (Günther et al., 2017; Johnson et al., 2017). With the big data approach, in general, scientists are less and less limited by availability of data but are more constrained by their limited skills (such as modelling) to efficiently use them. In addition to data accessibility, their reproducibility, their interoperability, and their reusability are also important factors. This is the reason why a group of academic groups, industries, funding agencies and publishers (Wilkinson et al., 2016) have designed a set of principles (known as the FAIR data principles) to favor data sharing and reusability, across scientific fields.

How to share data

The first FAIR principle, Findability, means that data are assigned with a unique and relevant identifier,
are briefly described by metadata including the identifier, and are registered or indexed in a searchable resource. This principle is well-known in genetics and genomics. For instance, the Animal Quantitative Trait Loci Database is a repository for quantitative trait loci (https://www.animalgenome.org/cgi-bin/QTLdb/index). Gene Expression Omnibus is a public functional genomics data repository (https://www.ncbi.nlm.nih.gov/geo/). In contrast, in animal phenotyping (including meat science), this is less common. One main reason is historical: the characterization of animals and of their products has been the subject of research for centuries and therefore started in each country with poor communication between research organizations, resulting in low standardization of methods. On the contrary, when genomics exploded, scientists were already more or less communicating with a common wish to share resources. The lack of standardization also impairs the third FAIR principle, namely Interoperability, as described subsequently. To better share phenotypic data, some institutions, such as the French National Research Institute for Agriculture, Food and the Environment (INRAE), have created their own repository: the Data INRAE portal offers new services to facilitate the management, sharing, and retrieval of research data from INRAE (https://data.inra.fr/dataverse/root). The INRAE portal attributes a unique and relevant identifier to each set of data deposited in this repository. For instance, a dataset related to carcass yield has been recently deposited to compare carcass yield across breeds (Hocquette et al., 2019).

The second FAIR principle, Accessibility, means that data are retrievable, that the protocol is open and free, and that at least metadata—and sometimes even raw data—are accessible. Again, geneticists are in the habit of making their data accessible because they really need large amounts of data for genetic analyses. On the contrary, the meat industry is very conservative (Troy and Kerry, 2010), often dominated by a few retailers who are focused on the competition between each other. Consequently, they are not open to sharing data, which weakens product-science-based development and innovation in the meat sector. Again, any initiative such as the Data INRAE portal aims to share research data funded by public grants.

How to reuse data

The third FAIR principle is Interoperability. This means first that data are in a formal, accessible, shared, and broadly applicable language for knowledge representation. Again, this is easier in genetics and genomics. Indeed, there are a relatively small number of techniques to study the DNA or RNA molecules. Scientists working in this area need large amounts of data and are motivated to define standards and to share their data. Therefore, standards were designed to describe the minimum amount of information that is essential to communicate about analyses in order to both enable interpretation of the experimental results with the least possible ambiguity and contribute to the reproducibility of the experiments. Such standards have been designed, such as Minimum Information About a Microarray Experiment by Brazma et al. (2001), Minimum Information About a Proteomics Experiment by Taylor et al. (2007), and Minimum Information About a Simulation Experiment by Waltemath et al. (2011). Unlike in DNA studies based on sequencing techniques, in biological science, there are many techniques and variants of these techniques to collect data for a single phenotype. This is, of course, more complicated for several phenotypes analyzed simultaneously. For instance, beef tenderness is a concept not precise enough: it is highly dependent on key elements of the measurement protocol, and it also depends on experts used to score beef (Gagaoua et al., 2016). More generally, confusion exists in meat science vocabulary, which results in misinformation and misunderstanding regarding meat descriptions (Seman et al., 2018). This is the reason why initiatives were taken to better define concepts and associated terms in meat science, such as the Meat Science Lexicon by the American Meat Science Association (Seman et al., 2018) or the Meat Dictionary by the French Meat Academy (Meat Academy, 2012). These resources should be organized in an ontology. An ontology is a formal and structured representation (often within a hierarchical structure) of a set of objects (for us, animal and meat traits or measurements) and of the relationships between these objects clearly defined with no ambiguity. The terms used should be machine readable to favor automated measurements or data use (Hocquette et al., 2012c). A unified phenotype ontology is crucial to facilitate comparisons of genes and phenotypes between individuals and across species and organisms (Hughes et al., 2008). To achieve the goals, the Animal Trait Ontology of Livestock (ATOL) was developed (http://www.atol-ontology.com/atol/) and should be strengthened in meat science.

The last FAIR principle is Reusability. This means that the data should be richly described, be released with a clear and accessible data-using license, be associated with detailed provenance, and meet domain-relevant community standards. As mentioned earlier,
to better share phenotypic data, INRAE has created its own repository: the Data INRAE portal (https://data.inra.fr/dataverse/root) is a tool to promote open science. It uses Dataverse, which is an open-source data repository software to support institutional research. The use of interoperable standards is also a key issue to make data reusable. To achieve this goal, a group of academics from different countries have developed AgroPortal, which is a vocabulary and ontology repository for agronomy, food, plant sciences, and biodiversity. This is a common platform to receive and host ontologies from varied areas related to agriculture, align them, and enable their use by informatics tools (Jonquet et al., 2018). For instance, by searching for a common term in meat science, such as “marbling,” different definitions of marbling can be found from 5 ontologies (including ATOL) at http://agroportal.lirmm.fr/.

**The Australian initiative**

Since no standards existed for describing consumer satisfaction (Polkinghorne and Thompson, 2010), scientists and professionals from Australia have designed the Meat Standards Australia (MSA) grading scheme (for a review, see Bonny et al., 2018b) taking into account lessons from past research as well as principles of data sharing and interoperability. Indeed, the MSA approach differs markedly from other systems currently employed for the prediction of beef eating quality. Firstly, it is based on responses of untrained consumers, those responses being considered as reference data unlike mechanical measurements or sensory scores from panelists mainly used in studies so far, as described earlier. Untrained consumers are asked to assess tenderness, juiciness, flavor liking, and overall liking for each beef cut following standard protocols (Watson et al., 2008a, 2008b). Then, these 4 scores are combined together to generate the 4-variable meat quality score (MQ4), which is a global eating-quality score (Figure 2). It was shown using the MSA approach that eating quality is poorly related to carcass fatness and conformation, which are the major traits considered in Europe for trading (Bonny et al., 2016). Secondly, the MSA grading scheme is based on large amounts of data obtained with the same standard protocols (Watson et al., 2008a, 2008b). Indeed, almost 100,000 untrained consumers were used over years to score around 700,000 beef samples from different sources in terms of breeds, animal types, cuts, etc., allowing robust modelling approaches, since the data are always comparable. By using this large database, the MSA grading model predicts MQ4, which is the palatability score of each individual cut based on 4 sensory descriptors from untrained consumer taste panels, namely tenderness, juiciness, flavor liking, and overall liking. MSA has identified and included in the model critical control points that impact palatability from the production, pre-slaughter, processing, and value-adding sectors of the beef supply chain using large-scale consumer testing. These Australian linked data may be complemented by other international studies conducted under identical protocols and currently amounting to a further 60,000 consumers and 420,000 individual beef samples.

Key grading factors used in the prediction model include Bos indicus content (which induces a decrease in beef palatability), animal sex, ossification

---

**Figure 2.** Contribution of animal and carcass characteristics and of post-mortem actions to the sensory quality of meat (tenderness, juiciness, flavor liking, and overall liking) according to the Meat Standards Australia principles. HGP = hormonal growth promotant.
(an indicator of animal physiological maturity), carcass weight, the use of hormonal growth promotants, marbling and rib fat thickness, carcass hanging method, ultimate pH, aging time, muscle name, and cooking method (Figure 2). At the end, palatability for individual muscles is predicted for a specific cooking method such as grilling, roasting, stir fry, and variations for thin slicing (reviewed by Bonny et al., 2018b). The new model version will also include a very large negative effect on eating quality for high-oxygen modified atmosphere packaging.

In Australia, cattle within the MSA system must come from registered producers and undergo best practice management and stress minimization such as no mixing of different groups of cattle. At slaughter, all carcasses are graded once ultimate pH is obtained, and then the grading data are used by the MSA model at the abattoir to predict the cut × cook outcome (i.e., the MQ4). MSA currently grades around 3.5 million carcasses per year in Australia (MLA, 2018, 2019). This strategy has been possible thanks to sharing and reusability of standardized sensory scores from untrained consumers.

Prediction of Beef Quality Using New Measurement Technologies

The Australian context

Within the Australian processing sector, there have been increasing calls to replace the existing human-based measurements populating the MSA model with technologies to objectively measure these same traits. Examples of traits currently “scored” by human graders include marbling, rib fat depth, ossification, and eye muscle area, although in practice eye muscle area isn’t used in the MSA model. There is a comprehensive system for accrediting and calibrating these human graders; however, significant bias in interpretation of these traits is still known to exist (Thompson, 2016). On this basis, the fundamental principal for introducing technologies to replace human grading is that the technology must perform as well as, or better than, the existing human grader in their precision, accuracy, and repeatability for the trait in question. These technologies need to be rapid—keeping up with the fastest chain speeds. They also need to have the potential to be automated, or be adaptable to existing automation within processing facilities. Crucially, standards for error tolerances must be established to enable the initial accreditation and subsequent industry auditing of each technology. Aligning these standards will further enhance the internationalization of these technologies and the prediction of eating quality. While benefits may accrue from more repeatable objective measurement technologies, further potential improvement in estimation may arise from measurement of additional factors that cannot be observed by the human eye. Known improvement could be achieved by incorporating genomic markers if, or perhaps when, commercially feasible technologies become available. Examples include near-infrared spectroscopy, Ramon, and nuclear magnetic resonance, all of which have the potential to directly measure muscle components that influence eating-quality relationships.

Other benefits of new measurement technologies

In some cases, these measurement technologies provide the opportunity to introduce new traits into existing eating-quality prediction systems. Examples include intramuscular fat percentage rather than visual marbling score, shear force, and wholecarcass measures of lean meat yield rather than simple indicators such as rib fat depth. A range of technologies are being tested for determining intramuscular fat percentage. Hyperspectral and Red Green Blue cameras capture an image of the eye muscle at the same location where the visual marbling grade is determined and therefore could be trained to objectively measure this trait. However, other devices are being tested that do not require this cut surface, capturing imagery through insertable probes or using scanning technologies such as nuclear magnetic resonance or X-rays (Anderson et al., 2018). These technologies determine the total chemical fat percentage of the muscle region of interest. Although this, in turn, could be transformed into the equivalent visual marbling score, this also opens the opportunity to directly input chemical fat percentage into the eating-quality prediction model. Stewart et al. (2018) have demonstrated the potential for this value to predict eating quality, a prediction that performs as well as or better than visual marbling. This has the added advantage of training objective measurement technologies on a repeatable and objective laboratory measurement, rather than subjective human grading data, which can be prone to imprecision and bias (Thompson, 2016).

Significant advances are also being made in the measurement of whole-carcass lean meat yield. Dual-energy X-ray absorptiometry, and carcass 3D imaging
systems, are showing great promise for determining the percentage of fat, lean, and bone in a carcass (Gardner et al., 2018). Although the MSA model already has proxy indicators of lean meat yield, with rib fat and carcass weight inputs, it is possible that a more detailed phenotype would enhance this prediction.

Crucially, as new consumer data are generated, it is becoming increasingly important to capture these additional measures, future-proofing the MSA model and making it more agile to change.

**Ambitions of the International Meat Research 3G Foundation**

**The international context**

Much of the work in beef science was internationally collaborative as meat scientists sought to develop appropriate science-based approaches to support industry in improving the consumer experience no matter the country. This is the reason why the International Meat Research 3G Foundation was launched recently. However, some decades ago, the advent of industry crisis related to *Escherichia coli* 0157:H7 and bovine spongiform encephalopathy demanded a shift of research funding and priority in both Europe and the US, whereas Australia—spared from direct exposure to either and with a significant quality challenge—maintained a primary research focus on eating quality. Early research in Australia established that existing systems were not fully effective and forced a fundamental reappraisal with measurement directly by consumer sensory testing as a fundamental base.

Critical steps were the development of rigorous test protocols and expression of consumer sensory standards such as the MQ4, leading to a realization that accurate prediction of an individual beef meal result needed to assign an outcome to cooked individual-meal–sized portions rather than a generic carcass description. As described earlier, a further fundamental decision was to accumulate data from all related experiments and commercial product evaluation no matter the country. Data are stored in a common database utilizing consistent description to enable issues to be evaluated over multiple disparate base studies.

Formal MSA international collaboration began with consumer studies in South Korea (Thompson et al., 2008) and Northern Ireland (Chong et al., 2019) and were followed by further projects in the US (Polkinghorne, 2007), Japan (Polkinghorne et al., 2011), Ireland (McCarthy et al., 2017), France (Legrand et al., 2013), and South Africa (Strydom et al., 2019). Independent extensive projects utilizing identical protocols were conducted in New Zealand (Garmyn et al., 2019), the US (O’Quinn et al., 2018), and Poland (Guzek et al., 2015; Pogorzelski et al., 2020), with other countries also extending earlier programs or conducting new studies (such as those in the UK, Wales, and France in the past year), the more recent projects being under the auspices of the Foundation.

Many of the research teams and individuals have collaborated closely over this period and concluded that, for maximum global beef industry benefit and research efficiency, it would be highly beneficial to pool data for analysis and to develop further industry applications, which is the aim of the International Meat Research 3G Foundation.

The various consumer studies had also established that global consumer groups were more similar than different with very similar sensory response despite cultural differences (Bonny et al., 2017, 2018a). Cattle raising systems and environments were complementary across countries. An American feedlot system and a Japanese or Irish housing system both finished cattle but involved different ration ingredients and on-farm management; Welsh and New Zealand grasslands were often very different from Australian and South African rangelands; and breeds used in France or Poland differed from those in Australia. Furthermore, use or not of hormonal implants as well as different proportions of young bulls, older cows, or dairy-cross cattle overlapped as well. Within this diversity was substantial crossover, providing considerable benefit from combining all data to enable the intricate interaction of genotype, phenotype, and environment to be more accurately established.

Central to these efforts was the use of common protocols and measurements (Watson et al., 2008a, 2008b). With the adoption of common sensory test protocols, consumer data were complementary with many animal and carcass traits also common or readily translated. Others, such as marbling and ossification, were not used in some regions and led to work with the UNECE Specialized Section on Standardization of Meat to establish and document extensions to the UNECE Bovine Language. A working group with Poland as lead rapporteur developed recommendations for beef grading input standards based on MSA protocols, which were subsequently accepted and now provide a formal base for data aggregation and conversion where appropriate.
Organization and missions of the International Meat Research 3G Foundation

The not-for-profit Foundation, called the International Meat Research 3G Foundation (https://imr3gfoundation.org/), was incorporated under Polish law to provide a practical structure to facilitate scientific collaboration, data storage, and utilization and to provide a platform for commercial application under the auspices of UNECE. Formal structures include a Management Council charged with legal responsibility for governance, business functions, and delivery of commercial activity and a Scientific Reference Group with responsibility for scientific standards, collaboration, and peer review. The Scientific Reference Group decisions are strictly evidence based and relate to data, whereas commercial and industry-related activity is the responsibility of the Council (Figure 3). It is anticipated that the structure will provide a useful forum for ongoing scientific collaboration across countries and organizations.

A major project is DATABank, which is establishing a cloud-based data storage and management system that can provide secure and confidential data storage for members and facility to pool data for agreed purposes as desired. Supporting largely open-source software is being expanded to provide easy access for researchers or students to assist with trial design and application aligned to standard protocols to ensure compatibility. Considerable effort is being made through the ontology working party to ensure standardized description and linkage to other international standards such as ATOL or the International Committee for Animal Recording. A data analysis technical group is foreshadowed as data are accumulated and made available for scientific investigation and potential development of eating-quality prediction models, with the capacity to relate consumer populations to alternative production systems and regions.

The Foundation has responsibility for training human graders in carcass chiller assessment in Europe, and potentially further regions. The Foundation is collaborating with AUS-MEAT, the Australian Standards organization responsible for holding and monitoring the use of official standards across industry sectors and for mandatory accreditation of red meat exporters, to ensure uniform application of the UNECE standards. This includes human grader correlation through a computerized quality assurance program (the On-Site Correlation and Practice System, or OSCAP). Training courses were run in Wales and France in 2019 and will expand further in conjunction with supporting applied meat science courses for industry participants. It is anticipated that the Foundation may provide access to eating-quality prediction models for beef grading on a commercial basis to encourage uniform consumer-based standards within a cost-effective framework. Above all, long-term benefit will accrue from rigorous scientific collaboration related to providing a sophisticated consumer focus and understanding to support long-term beef industry sustainability and relevance.

Conclusions

The prediction of beef eating quality has been the subject of active research for decades. Large amounts of research have been focused on the various production and post-mortem factors, which regulate muscle biochemical characteristics and thus beef tenderness and eating quality. As described in this review, some conflicting results appeared depending on the breed, the cut, or the country. Some conflicting results are due to the various methods to assess beef eating quality (protocols of shear force, or of sensory tests). Nevertheless, the proportion of variance explained by any variability in muscle biochemical trait was low based on meta-analysis, despite some encouraging results. This was attributed to our relatively low understanding of muscle biochemistry and has justified important developments in omics approaches, which are much more powerful thanks to the advent of high-throughput techniques. The goal was to capture new genes related to beef eating quality and thus new biological mechanisms inaccessible by classic biochemistry. This approach was impaired by low numbers of animals that can be simultaneously studied. In parallel, automatic methods were developed thanks to the advent of robotics. In all cases, conflicting results were still there. This is now attributed to the limited number of studied samples in some genomics or robotics experiments due
to their still high cost despite significant decreases. Furthermore, it can also be attributed to the lack of standards to compare beef eating quality across countries or research labs taking into account diversity in beef production. Consequently, as in other biological fields, the big data approach is expecting to solve these limitations, first by developing some standards and ontologies and second by compensating for the high variability of data by acquiring a large number of data supposed to strengthen the robustness of modeling approaches. This is one of the main goals of the International Meat Research 3G Foundation. The Foundation is indeed promoting data sharing, based on the most advanced beef grading system, MSA, which has been undergoing development since the 1990s, always with the same protocols to record the most powerful determinants of beef eating quality on a large scale across countries.

**Literature Cited**


