Accepted Manuscript

Investigating the use of visible and near infrared spectroscopy to predict sensory and texture attributes of beef M. longissimus thoracis et lumborum

Jamie Cafferky, Torres Sweeney, Paul Allen, Amna Sahar, Gerard Downey, Andrew Cromie, Ruth M. Hamill

PII: S0309-1740(18)30985-9
DOI: https://doi.org/10.1016/j.meatsci.2019.107915
Article Number: 107915
Reference: MESC 107915

To appear in: *Meat Science*

Received date: 18 October 2018
Revised date: 14 August 2019
Accepted date: 14 August 2019

Please cite this article as: J. Cafferky, T. Sweeney, P. Allen, et al., Investigating the use of visible and near infrared spectroscopy to predict sensory and texture attributes of beef M. longissimus thoracis et lumborum, Meat Science, https://doi.org/10.1016/j.meatsci.2019.107915

Investigating the use of visible and near infrared spectroscopy to predict sensory and texture attributes of beef *M. longissimus thoracis et lumborum*

Jamie Cafferky[a,b], Torres Sweeney[b], Paul Allen[a], Amna Sahar[a], Gerard Downey[a], Andrew Cromie[c], Ruth M. Hamill[1,*] ruth.hamill@teagasc.ie

[a]Department of Food Quality and Sensory Science, Teagasc Food Research Centre, Ashtown, Dublin 15, Ireland

[b]School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

[c]Irish Cattle Breeding Federation, Shinagh House, Bandon, Co. Cork, Ireland

[*]Corresponding author.

Abstract

The aim of this study was to calibrate chemometric models to predict beef *M. longissimus thoracis et lumborum* (*LTL*) sensory and textural values using visible-near infrared (VISNIR) spectroscopy. Spectra were collected on the cut surface of *LTL* steaks both on-line and off-line. Cooked *LTL* steaks were analysed by a trained beef sensory panel as well as undergoing WBSF analysis. The best coefficients of determination of cross validation (R^2CV) in the current study were for textural traits (WBSF = 0.22; stringiness = 0.22; crumbly texture = 0.41: all 3 models calibrated using 48 h post-mortem spectra), and some sensory flavour traits (fatty mouthfeel = 0.23; fatty after-effect = 0.28: both calibrated using 49 h post-mortem spectra). The results of this experiment indicate that VISNIR spectroscopy has potential to predict a range of sensory traits (particularly textural traits) with an acceptable level of accuracy at specific post-mortem times.

Keywords: Visible-near infrared spectroscopy; Chemometrics; Beef quality; Trained sensory panel; Shear force.

1. Introduction

Beef is one of the most valued products obtained from bovine animals (Food and Agriculture Organization of the United Nations, 2016). Eating satisfaction, sensory characteristics and nutritional content are the key factors that influence consumers in regard to purchasing fresh meat (De Marchi, 2013; Grunert, Bredahl, & Brunsø, 2004). Beef producers, manufacturers and retailers have a duty to offer meat that is safe, of a high quality and sustainable (Troy, Ojha, Kerry, & Tiwari, 2016). However, conventional methods to assess quality are expensive, destructive and time consuming (e.g. instrumental measurement of beef texture using Warner-Bratzler shear force (WBSF) and sensory analysis of meat samples using a trained sensory panel) (Prevolnik, Čandek-Potokar, & Škorjanc, 2004; Prieto, Ross, Navajas, Nute, Richardson, Hyslop, Simm, Roehe, 2009; Yancey, Apple, Meullenet, & Sawyer, 2010). Large scale, industry wide recording of beef quality phenotypes (such as instrumental tenderness and trained sensory panel scores) is not feasible, and is a limiting factor when developing breeding programmes for the improvement of beef eating quality (Cecchinato, De Marchi, Penasa, Albera, & Bittante, 2011). The drive to inform on product quality, maintain a consistently high level of quality and develop methods to identify animals that are genetically superior (in terms of meat quality) for breeding purposes is an important challenge that has led to the exploration of new technologies suitable as proxy measures to provide information on quality traits both online and inline on a larger scale (Troy et al., 2016).

Visible and near infrared (VISNIR) spectroscopy is a technology that has been proposed as an efficient, cost-effective tool for the simultaneous prediction of various important meat sensory quality traits (De Marchi, 2013; Dixit, Casado-Gavalda, R. Cama-Moncunill, X. Cama-Moncunill, Markiewicz-Keszycka, Cullen, & Sullivan, 2017; Prevolnik et al., 2004; Weeranantanaphan, Downey, Allen, & Sun, 2011). It involves the detection of the variation in the vibrations of –CH, –OH and –SH molecular bonds between samples of different quality

(Roggo, Chalus, Maurer, Lema-Martinez, Edmont, & Jent, 2007), which occur when a sample undergoes irradiation from a VISNIR light source and its reflectance is measured. Physical and sensory quality attributes in beef samples may potentially be indirectly predicted using VISNIR spectroscopy (Prieto, Roehe, Lavín, Batten, & Andrés, 2009) coupled with advanced chemometric regression analysis. Some of the main advantages of VISNIR spectroscopy are that it is a rapid, non-destructive and environmentally friendly objective technology, known for its ease of use and low-running cost (Dixit et al., 2017; Prieto, López-Campos, Aalhus, Dugan, Juárez, & Uttaro, 2014; Troy et al., 2016).

The aim of this study was to assess the potential of VISNIR spectroscopy as an early post-mortem, rapid predictor of instrumental texture, sensory texture, sensory juiciness and sensory flavour scores in 14 day aged beef samples, through chemometric modelling of VISNIR spectra collected at four different time-points (both on-line and off-line setting) and using three different mathematical spectral pre-treatments.

2. Materials and methods

2.1 Animals and sample preparation

Crossbred bull and steer progeny (18 ± 4 months old; n = 469 bull, n = 126 steer) from elite Irish beef breed artificial insemination bulls were obtained and reared under the same feeding and environmental conditions by the Irish Cattle Breeders Federation Tully Progeny Test Centre (Tully, Kildare, Ireland). Animals from eight beef breeds, with representative numbers for each breed in proportion to that of the Irish herd, were included as part of this study as follows; Limousin (LM; bull n = 200, steer n = 40), Charolais (CH; bull n = 107, steer n = 28), Simmental (SI; bull n = 50, steer n = 6), Belgian Blue (BB; bull n = 55, steer n = 6), Aberdeen Angus (AA; bull n = 30, steer n = 22), Salers (SA; bull n = 17, steer n = 12), Hereford (HE; bull n = 2, steer n = 8), Parthenais (PT; bull n = 8, steer = 4). Information on how animals were reared and subsequently slaughtered has been previously described in Cafferky, Hamill, Allen, O'Doherty, Cromie, and Sweeney (2019). At 24 h post-mortem, carcasses were quartered and at 48 h post-mortem loins were boned out. Twelve steaks with a thickness of 2.54 cm were sliced from the right-side *LTL* starting at the anterior end and vacuum packaged. These steaks were labelled 1-12 in order of cutting. The 4th steak on each loin was allocated for WBSF analysis and the 5th and 6th steaks for sensory analysis. Forty nine h PM steaks were vacuum packaged and aged for a further 12 days in a chill room set at 4 °C, then blast frozen and stored at -20 °C until analysis.

2.2 VISNIR spectral measurements

A portable ASD Labspec 5000 (ASD Inc., Boulder Colorado, USA) VISNIR spectrometer fitted with a high-intensity contact probe with a 10 mm spot size was used to collect spectra at four time-points, two on day 1 post-mortem and two on day 2 post-mortem. Prior to spectral acquisition the instrument was calibrated using a Spectralon tile as the white reference. VISNIR spectral measurements were recorded at quartering on the cut surface of the *LTL* muscle (5th rib) 1 day PM, immediately after cutting at the quartering stage (24 h PM) and after 1 h blooming in a chill room (25 h PM). This was repeated off-line on steak number 12 at 2 days PM, immediately after slicing (48 h PM) and again after 1 h blooming (49 h PM, respectively). The muscle was scanned in triplicate at 3 representative locations on the transverse surface of the *LTL*, that were free from excessive fat, observable connective tissue, bone and debris from cutting. For each of these three scans, 20 spectra were automatically collected by the instrument consecutively and averaged to reduce noise effect. Spectra were collected at 1 nm intervals between 350-2500 nm using the Indico Pro program (ASD Inc.) and saved in absorbance mode (log (1/R)). Spectral data was exported as JCAMP to The Unscrambler X version 10.3 (CAMO ASA, Oslo, Norway) for further chemometric analysis.

2.3 Warner-Bratzler shear force

Steaks (n = 469 bull, n = 126 steer) were thawed in unsealed plastic vacuum bags in a circulating water bath at room temperature (20 °C). Once thawed, steaks were trimmed of external fat, placed into a fresh unsealed plastic vacuum bag and immersed in a water bath for cooking (Grant Instruments Ltd., England) at 72 °C until an internal core temperature of 70 °C was reached (as measured using a temperature probe; Eirelec Ltd., Ireland). Samples were placed on the benchtop in opened unsealed vacuum bags (free of cooking juices) until they reached ambient temperature. Once samples had cooled to ambient temperature they were placed within new unsealed vacuum bags and left to temper overnight in a refrigerator at 4 °C. Shear force analysis was conducted following a modified version of American Meat Science Association (AMSA) Research Guidelines for Cookery, Sensory Evaluation and Instrumental Tenderness Measurements of Meat (AMSA, 2015). Seven cores per steak were taken with a 1.27 cm core and the fibres sheared perpendicular to the fibre direction using the Instron 4464 Universal testing machine (Instron Ltd., Buckinghamshire, UK), load cell of 500 N, cross head speed 250 mm/min and analysed using Bluehill 2 Software (Instron Ltd., Buckinghamshire, UK). Maximum peak force recorded during analysis was reported as Newton (N) shear force. Highest and lowest measurements were excluded with the average of the remaining 5 cores recorded as the result to reduce standard deviation (modified from AMSA 2015 guidelines).

2.4 Sensory analysis

Sensory analysis (n = 180 bulls, n = 29 steers) was conducted at the Teagasc Food Research Centre (Ashtown, Dublin 15, Ireland) following a modified version of AMSA (2015) guidelines. Samples were selected from the wider data set for trained sensory panel analysis by using the Kennard-Stone algorithm (Unscrambler X v.10.3 software package; CAMO, Trondheim, Norway), which selects a subset of spectra that have uniform distribution over the predictor space. This is designed to capture maximal spectral variability within the data set. Compusense 5.6 software was used to set up the project as a complete block design and to collect panellist scores. A 12 member trained sensory panel was used to assess the beef samples, using the average of 8 assessments per sample. Frozen 14 day aged samples were thawed at 4 °C for 24 h, then acclimatised to room temperature for 1 h prior to cooking. Steaks were cooked on a Velox CG-3 grill (Wantage, UK) at a temperature of 210 °C, until an internal temperature of 70 °C was reached, rested in foil for 4 min, cut into approximately 2.5 cm x 4 cm pieces, wrapped individually in foil, and served. Eight members of a 12 member trained sensory panel assessed each beef sample in the morning, using the average of 8 assessments per sample. Panelists had access to water and plain crackers as palate cleansers between each sample. Twelve sensory traits were scored (tenderness, juiciness, chewiness, stringiness, crumbly texture, fatty mouthfeel, difficulty to swallow, beef flavour, metallic flavour, beef after-effect, metallic after-effect and fatty after-effect) on a 0-100 line scale, with 0 being the lowest score and 100 the highest.

2.5 Chemometric analysis

All chemometric analysis in the current study was performed using The Unscrambler X v.10.3 software package (CAMO, Trondheim, Norway). All traits were tested for normal distribution and outliers within The Unscrambler software package. Post visual inspection, anomalous spectra were detected with the use of the H-statistic, which indicates the difference between a sample spectrum and the average spectrum of the set (Williams & Norris, 2001). All samples with a H-statistic greater than ≥ 3 standardised units from the mean spectrum was defined as a global H outlier and removed from the dataset. For the prediction of beef sensory and instrumental textural attributes using VISNIR spectroscopy (X data), 60 separate partial least squares regression models (PLSR) were developed for each of the 13 sensory and textural traits analysed (Y data), with 780 models constructed in total. PLSR models were developed using spectra collected at 4 post-mortem time-points which were: the quartering stage (on-line) on day 1 (24 and 25 h PM) and on a boned out piece of *LTL* sample (off-line) on day 2 (48 and 49 h PM). Models were developed for each of the 4 time-points with the use of 5 separate subset wavelength ranges (full, 350-2500 nm; clipped, 450-2300 nm; visible, 450-779 nm; near-NIR, 780-1099 nm; NIR, 1100-2300 nm) and 3 mathematical treatments (None, no treatment; Standard Normal Variate, SNV; Savitzky-Golay smoothing – 2nd derivative with 2nd polynomial and 21 smoothing points, SG). Spectra were subset into separate wavelength ranges to test the hypothesis that more robust models would be developed for specific traits by selecting specific ranges of the spectrum compared to those generated using the full spectrum. PLSR models were calibrated using the non-linear iterative partial least squares (NIPALS) algorithm, with the performance of models evaluated using a leave-one-out full internal cross validation (Prieto et al., 2014). The best fitting model for each trait was identified using a standard procedure in this field, by the model that had an optimal combination of the lowest Root Mean Square Error of Cross Validation (RMSECV) and the lowest number of PLS terms utilized; with this result validated using the explained variance test in the Unscrambler X 10.3 software package. Ratio Performance Deviation (RPD) was calculated as the Standard Deviation (SD) of a meat quality trait divided by the RMSECV (Williams, 2014).

2.6 Statistical analysis

Descriptive statistical analysis was performed using The Unscrambler X v.10.3 software package (CAMO, Trondheim, Norway). Pearson Correlations between beef *LTL* quality attributes were calculated using the CORR procedure in Statistical Analysis Software

(SAS) Version 9.4 (SAS Institute Inc., Cary, NC. USA) with a significance level of $P < 0.05$.

3. Results

3.1 Descriptive statistics and correlations between sensory traits

Sensory tenderness scores and WBSF values (Table 1) shared similar standard deviation values and coefficients of variation within the dataset (SD 10.47 and 10.8, respectively; CV 20.0% and 24.94%, respectively). Large coefficients of variation were found for the textural sensory traits chewiness, stringiness, crumbly texture, fatty mouthfeel, metallic after-effect and fatty after-effect (CV 41.49% to 84.95%).

Correlations among the twelve sensory attributes are displayed in Table 2. Sensory tenderness scores were inversely correlated with tougher textural attributes (chewiness, stringiness and difficulty swallowing, R = -0.72 to -0.86; $P < 0.001$), while tougher textural attributes were positively correlated to each other (R = 0.77 to 0.84; $P < 0.001$). Textural attributes associated with tender beef such as sensory tenderness and crumbliness were positively correlated with juiciness (R = 0.3 to 0.47; $P < 0.001$). The instrumental measurement of tenderness (WBSF) was negatively correlated with sensory tenderness, juiciness, crumbliness, beef flavour and beef after-effect, while positively correlated with chewiness, stringiness and difficulty to swallow ($P < 0.001$). Furthermore, juiciness was positively correlated with beef flavour and beef after-effect. No correlation was observed between WBSF values and fatty mouthfeel, metallic flavour, metallic after-effect and fatty after-effect ($P > 0.05$).

The flavour attributes: beef flavour, beef after-effect, fatty mouthfeel and fatty after-effect were all positively correlated with each other ($P < 0.001$), with a strong relationship between beef after-effect and beef flavour (R = 0.72; $P < 0.001$), as well as between fatty after-effect and fatty mouthfeel (R = 0.82; $P < 0.001$) (Table 2). The metallic sensory traits, metallic flavour and metallic after-effect were positively correlated (R = 0.67; P < 0.001); while metallic flavour and beef flavour were uncorrelated (P > 0.05). Positive correlations were present between metallic flavour and both fatty mouthfeel and fatty after-effect (P < 0.001).

Correlations between tenderness and fatty mouthfeel, fatty after-effect, metallic flavour and metallic after-effect (P < 0.001) were negative, while tenderness was positively correlated with beef after-effect (P < 0.05). No correlation was observed between tenderness and beef flavour (P > 0.05). Stringiness and crumbliness were both positively correlated with fatty mouthfeel, metallic flavour and metallic after-effect (P < 0.001). Crumbliness positively correlated with beef flavour and beef after-effect, while stringiness showed a positive correlation with fatty after-effect (P < 0.001).

3.2 Prediction of beef LTL textural and juiciness quality traits

Best fitting prediction models for beef sensory and textural quality traits using VISNIR spectroscopy are presented in Table 3. The highest coefficients of determination of cross validation (R^2CV) obtained for prediction of WBSF values in this study (R^2CV = 0.22, RMSECV = 10.1 N) were obtained using spectra collected at 48 h PM in the 450-2300 nm wavelength range. The best fitting model calibrated using spectra collected on-line 1 day PM was with the use of 24 h PM spectra (R^2CV = 0.14, RMSECV = 10 N), incorporating the entire VISNIR wavelength range recorded (350-2500 nm). In Figure 1, the spectra (48 h PM; 350-2500 nm) for the average of the five highest scoring WBSF samples (tough) and the average of the five lowest scoring WBSF samples (tender) are plotted together, with tougher samples having higher absorption from 650-1150 nm. However, absorption for the tougher samples was lower between 1400-2500 nm.

The 49 h PM time-point yielded the highest R^2CV for prediction of sensory tenderness in the present study (R^2CV = 0.13, RMSECV = 8.6), while spectra collected 25 h PM gave the highest value for spectra collected 1 day PM (R^2CV = 0.06, RMSECV = 10.3). This 49 h PM model was calibrated using the full spectral wavelength (350-2500 nm). When 49 h PM spectral measurements (350-2500 nm) for the 5 most tender samples are averaged and plotted with the average spectra for the 5 toughest samples, the tougher samples had higher absorption between 550- 1450 nm (Figure 2).

Tougher sensory textural attributes were all predicted in a similar range with a low degree of accuracy (R^2CV = 0.1 to 0.22; RMSECV = 7.1 to 9.3). Best fitting models for all three sensory traits were built using spectra from the visible wavelength region, 450 – 779 nm, and collected 2 days PM (48 h PM for chewiness and stringiness; 49 h PM for difficulty to swallow).

Prediction of crumbly sensory texture gave the highest coefficient of determination of all traits analysed in the present study. The 48 h PM spectral model outperforms the 25 h PM model, with higher R^2CV (0.41, 48 h PM; 0.34, 25 h PM) and lower RMSECV values (9.4, 48 h PM; 10.5, 25 h PM). Best fitting models for crumbly texture (Table 3) were built using the full spectral wavelength range (350-2500 nm) and without the use of any mathematical pre-treatment as they were not found to improve accuracy of prediction or reduce RMSECV value.

Sensory panel juiciness scores predicted via VISNIR spectra had the lowest R^2 value of all traits in the current study, with the highest R^2CV achieved being 0.09 (Table 3). The best fitting PLSR models for juiciness shown in Table 3 were constructed using spectra collected off-line (49 h PM), when the muscle had been sliced and bloomed for 1 h. Both 24 h and 49 h PM models were shown to have the relatively low RMSECV value of 6.8.

3.3 Prediction of beef LTL sensory flavour

The best fitting model for beef flavour prediction (48 h PM) provided an R^2CV value of 0.13 and an RMSECV value of 6.4, and was constructed using the full spectral wavelength and no spectral pre-treatment. Beef after-effect was predicted with an R^2CV of 0.2, and a lower RMSECV (in comparison to beef flavour) of 5.4 (Table 3). The model for beef after-effect was calibrated using spectra from the near-NIR range, which is far narrower and in contrast with the full wavelength range which gave most accurate determinations for beef flavour. It is also worth noting that beef after-effect was the only sensory trait in this study where a spectral pre-treatment (SNV) model provided the best fitting model for both day 1 and 2 PM. In all other models within the present study, a spectral pre-treatment has only improved the accuracy of models for one PM day.

For prediction of fatty mouthfeel and fatty after-effect, the best fitting models for both traits (R^2CV = 0.23; 0.28 - Table 3) shared a number of similarities in their calibration, such as similar time-point (49 h PM), wavelength range utilized (780-1099 nm) and RMSECV value (3.2). In Figure 3, 25 h PM spectral measurements for the average of the 5 lowest scoring and the average of the 5 highest scoring fatty mouthfeel samples are plotted together (350-2500 nm). The 5 higher scoring samples have higher absorption between 550-2500 nm. Peaks are visible at 980, 1100 and 1400 nm for both high and low scoring fatty mouthfeel samples.

Best fitting predictive models for metallic flavour and metallic after-effect were found to be very similar (Table 3), with both models calibrated using spectra from the same time-point (24 h PM), the same optimal number of PLS terms utilized (10 terms), and the near-NIR range (780-1099 nm). Both models also had similar R^2CV predictions of 0.17 and 0.19, respectively.

4. Discussion

Methods currently employed for the measurement of meat quality do not allow for industry wide quality monitoring or the recording of these traits for breeding objectives. The ability to implement on-line systems to predict meat quality at a time-point compatible with many meat management systems (48 h PM) within the factory could be of utility to breeders and processors (Cecchinato et al., 2011; De Marchi, 2013). Prediction of instrumental and sensory textural information of carcasses on-line is desirable, as it allows for the formation of differentiated product lines of various grades of meat with higher quality and high added value (Balage, da Luz e Silva, Gomide, Bonin, & Figueira, 2015). Improved grading of meat based on quality may also lead to a fairer cost to the consumer and additionally a fairer payment to the producer.

The most accurate coefficients of determination for prediction of WBSF values in this study were obtained at the 48 h PM time-point (R^2CV = 0.22, RMSECV = 10.1. In Figure 1, differences in absorbance between the spectra of high scoring WBSF samples (tough) and low scoring WBSF samples (tender) can be seen along the VISNIR range (350-2500 nm), indicating that these spectral regions contain important information relevant to prediction of WBSF. This is in agreement with Hildrum, Nilsen, Mielnik, and Næs (1994) who reported tougher beef samples having higher absorbance values in the regions previously mentioned. Prieto et al. (2009) obtained a similar value (R^2CV = 0.23) for the prediction of 14 d aged beef *LTL* samples analysed using slice shear force, which is similar to WBSF, with both our study and theirs having low RPD values (1.07; 1.14, respectively). However, the present study used a much larger set of samples for WBSF prediction in comparison to Prieto et al. (2009), which possibly contributed to the improved RMSECV value (10.1, in comparison to 28.49). The RPD value of 1.07 reported in this study is considered too low for implementation of VISNIR spectroscopy into an abattoir for the prediction of shear force values, as RPD values < 2 are not recommended for quality control or grading purposes (Williams, 2014). Predictions of WBSF values in other meats using VISNIR spectroscopy have been reported by De Marchi, Penasa, Battagin, Zanetti, Pulici, and Cassandro (2011) in chicken breast (R^2CV = 0.41, RMSECV = 3.18) and Balage et al., (2015) in pork loin (R^2CV = 0.3, RMSECV = 4.98, RPD = 0.2). In the current study, R^2CV values for WBSF were greater than those obtained by De Marchi (2013), with similar standard error values reported. De Marchi (2013) showed the application of a mathematical pre-treatment upon spectra (multiplicative scatter correction and first derivative) had a negative effect upon predictive accuracy, in agreement with the present study. Ripoll, Albertí, Panea, Olleta, and Sañudo (2008) achieved higher coefficients of determination for beef *LTL* WBSF prediction (R^2 = 0.743). However, this value was achieved by collecting spectra on homogenized rather than intact sample, which previously has been shown to increase predictive capabilities in muscle tissues (Weeranantanaphan et al., 2011) by destroying and randomising the muscle fibre arrangements, therefore averaging the effect of scattering caused by fibres of different size (Prieto et al., 2009). Models necessitating a homogenization step are less practical when implementation in a commercial setting is being considered, as it is both destructive of the sample and time consuming.

Sensory tenderness predictions resulting from spectra collected offline in the laboratory at 49 h post-mortem in the present study were similar to that reported by Prieto et al. (2009), but lower than those reported by Ripoll et al. (2008) (R^2CV = 0.98). The result achieved by Ripoll et al. (2008) may be attributed to the use of animals with differing maturity rates and varying fatness and conformation grades, leading to greater variation within the dataset (tenderness CV = 32.92%) and in turn higher coefficients of determination. Moreover, in their study spectral measurements and sensory panel analysis were carried out on the same day (7 d PM). Again, when implementation in the factory setting is the ultimate goal, collecting spectra offline at 7 d PM may be less desired. In Figure 2, tougher sensory tenderness samples had higher absorbance between 550-1450 nm. This is most likely due to tougher samples having reduced sarcomere length and therefore deeper penetration path length (Prieto, Ross, et al., 2009).

The textural sensory traits associated with toughness – chewiness, stringiness and difficulty to swallow, were poorly predicted in the current study (Table 3). Interestingly, all 3 traits were highly positively correlated to each other (Table 2). Coefficients of determination of calibration (R^2Cal) for chewiness obtained in the current study (R^2Cal = 0.38) are lower than in other studies (Liu, Lyon, Windham, Realini, Pringle, & Duckett, 2003; Rødbotten, Nilsen, & Hildrum, 2000), who reported R^2Cal values of 0.58 and 0.38, respectively. The R^2CV values are not presented in the studies mentioned so cannot be compared. It is notable that the models presented by Liu et al. (2003) were constructed using just 24 carcasses from the similar Angus and Hereford breeds; with the possibility that this animal group used for calibration is too homogenous for prediction of chewiness in a larger scale.

The R^2CV values reported in the current study for the prediction of juiciness by way of VISNIR spectroscopy are in agreement with that of Prieto et al. (2009), who also achieved a low R^2CV value (0.13), and in disagreement with Liu et al. (2003) and Ripoll et al. (2008) who reported more accurate predictions of juiciness (R^2 = 0.5-0.54). Rødbotten et al. (2000) were unable to obtain any predictive model for beef sensory juiciness scores in their study even after the application of mathematical pre-treatment to the spectral data (multiplicative scatter correction), underlining that prediction of juiciness scores is a complex affair and a difficult attribute to consistently predict with accuracy. Clearly, juiciness is a more complex trait than previously considered and aspects of the physical structure may impact the release of juice in the mouth which was not well-predicted here.

Predictive models obtained in the current study for beef flavour and beef after-effect were unsatisfactory (Table 3). Low predictions for beef flavour were also reported by Prieto et al. (2009) (R^2CV = 0.26; Standard Error of Cross Validation, SECV = 0.56) and Byrne, Downey, Troy, and Buckley (1998) (R^2CV = 0.24; SECV= 0.39), while Andrés, Murray, Navajas, Fisher, Lambe, and Bünger (2007) reported low predictions of flavour in lamb (R^2CV = 0.27). Beef flavour variation is most commonly due to changes in fatty acid profile. Fatty acid content was not examined in the current study, however the C – H molecular bonds of fatty acids (1100-1400 nm) (Prieto, Andrés, Giráldez, Mantecón, & Lavín, 2008; Prieto, Roehe, et al., 2009; Pullanagari, Yule, & Agnew, 2015) are located within the spectral regions that provided the most accurate prediction of the beef flavour trait (350-2500 nm). It is also worth noting that beef after-effect was the only sensory trait in this study where a spectral pre-treatment (SNV) model provided the best fitting model for both day 1 and 2. In all other models within the present study, spectral pre-treatment has only improved the accuracy of models for one PM day. One possible explanation is that every trait (except beef after-effect) where SNV was shown to improve the performance of a model was calibrated using spectra collected at day 1 PM. It is possible that these spectra collected post quartering, on-line, and under sub-optimal conditions had the most to gain from a mathematical pre-processing step.

Fatty mouthfeel and fatty after-effect are another pair of sensory traits in the current study that are both highly positively correlated and comparable to each other. R^2CV predictions for fatty after-effect were marginally more accurate of the two (0.28 compared to 0.23 for mouthfeel), with this model calibrated with the use of Savitzky-Golay smoothing. This was the only model within the present study where SG smoothing increased the predictive capabilities of a model, in comparison to untreated spectra. Interestingly, it was observed that R^2 calibration models for both traits related to fattiness had higher predictive values in models constructed with day 1 PM spectra (R^2Cal = 0.46-0.53), indicating that some information for these traits is captured in spectra recorded on the day of quartering. The most accurate models for the prediction of fat related sensory traits in the current study were calibrated using spectra within the near-NIR range (780-1099 nm), which contains the third stretching overtone of IMF C – H bonds (950 nm) as well as the O – H bond for water (890 nm) (Shenk, Westerhaus, & Workman, 2007), two regions known for their role in prediction of IMF/fat sensory traits in meat (Andrés et al., 2007). In Figure 3, peaks are visible at 980, 1100 and 1400 nm for high and low scoring fatty mouthfeel samples when they are plotted against each other. The peak visible at 980 nm is presumably due to water (Liu et al., 2003), while the peaks visible at 1100 nm and 1400 nm are most likely related to C – H molecular bonds of fatty acids (Williams & Norris, 2001).

To the best of our knowledge, no other reports exist in the literature relating to the use of visible and near infrared spectroscopy for the prediction of the sensory traits beef after-effect, fatty mouthfeel, fatty after-effect, metallic flavour and metallic after-effect in beef *LTL*. While the CV values for fatty mouthfeel, fatty after-effect, metallic flavour and metallic after-effect were high (Table 1) and thus might be expected to generate acceptable models, despite a variety of pre-processing approaches and spectral subsets being explored, the models generated were lower quality in comparison to other traits. However, it is considered that texture, which was well predicted with good models for WBSF, stringiness and crumbliness are among the most important sensory traits for beef, suggesting that VISNIR spectroscopy is of value in sensory quality prediction for this foodstuff.

5. Conclusion

The results of this research demonstrate that VISNIR spectroscopy applied at standard cutting time-points showed varying levels of performance depending on trait, and prediction was best for texture traits. Spectra collected at 2 days PM provided best fitting models for 11 of the 13 sensory traits assessed. Traits with higher coefficients of variation were predicted with greater accuracy than traits with lower levels of variation. Chemometric models constructed using different regions of the VISNIR range as well as a number of mathematical treatments were investigated in an effort to maximise overall model performance, in particular predictive efficacy and robustness. For the majority of traits analysed, mathematical pre-treatments did not improve model performance. Models calibrated using the full VISNIR spectral wavelength range (350 – 2500 nm) and the visible wavelength range (450 – 779 nm) provided the best fitting models in the majority of traits analysed. However, for some traits which did not produce high quality models, subsetting of spectral ranges resulted in some improvement to the model fit. The value of VISNIR spectroscopy to prediction of beef quality for breeders is that spectra can be recorded non-destructively, at time-points compatible with meat management systems. With this data, information can be captured regarding a very large range of traits, both technological and sensory, once calibrations, such as those in the present study are available. This study provides the first calibrations for many traits derived from trained sensory panels and may have relevance to future breeding programmes for beef sensory quality.

Acknowledgements

This work is funded by the BreedQuality project (11/SF/311) which is supported by The Irish Department of Food, Agriculture and the Marine (DAFM) under the National Development Plan 2007-2013. We acknowledge the Irish Cattle Breeding Federation for supply of samples, Carol Griffin and Carmel Farrell for conducting the trained sensory panel and Eugene Vesey for his assistance.

Figure 1. Line plot showing the differences in 48 h PM averaged spectra between the five samples with the lowest WBSF values (tender), seen in red, compared to the averaged five samples with the highest WBSF values (tough), seen in blue.

Figure 2. Line plot showing the differences in 49 h PM spectra between the five samples with the lowest sensory tenderness values (tough), seen in blue, compared to the five samples with the highest tenderness values (tender), seen in red.

Figure 3. Line plot showing the differences in 25 h PM spectra between the five samples with the lowest fatty mouthfeel values, seen in red, compared to the five samples with the highest fatty mouthfeel values, seen in blue.

Competing interest:

The authors declare they have no competing interests

References

AMSA. (2015). Research guidelines for cookery, sensory evaluation and instrumental tenderness measurements of meat. *American Meat Science Association*. Retrieved from https://www.meatscience.org/docs/default-source/publications-resources/amsa-sensory-and-tenderness-evaluation-guidelines/research-guide/2015-amsa-sensory-guidelines-1-0.pdf?sfvrsn=6

Andrés, S., Murray, I., Navajas, E. A., Fisher, A. V., Lambe, N. R., & Bünger, L. (2007). Prediction of sensory characteristics of lamb meat samples by near infrared reflectance spectroscopy. *Meat Science*, *76*(3), 509–516. doi:10.1016/j.meatsci.2007.01.011

Balage, J. M., da Luz e Silva, S., Gomide, C. A., Bonin, M. de N., & Figueira, A. C. (2015). Predicting pork quality using Vis/NIR spectroscopy. *Meat Science*, *108*, 37–43. doi:10.1016/j.meatsci.2015.04.018

Byrne, C. E., Downey, G., Troy, D. J., & Buckley, D. J. (1998). Non-destructive prediction of selected quality attributes of beef by near-infrared reflectance spectroscopy between 750 and 1098 nm. *Meat Science*, *49*(4), 399–409. doi:10.1016/S0309-1740(98)00005-9

Cafferky, J., Hamill, R. M., Allen, P., O'Doherty, J. V., Cromie, A., & Sweeney, T. (2019).

Effect of Breed and Gender on Meat Quality of M. longissimus thoracis et lumborum Muscle from Crossbred Beef Bulls and Steers. *Foods*, *8*(5). doi:doi:10.3390/foods8050173

Cecchinato, A., de Marchi, M., Penasa, M., Albera, A., & Bittante, G. (2011). Near-infrared reflectance spectroscopy predictions as indicator traits in breeding programs for enhanced beef quality. *Journal of Animal Science*, *89*(9), 2687–2695. doi:10.2527/jas.2010-3740

De Marchi, M. (2013). On-line prediction of beef quality traits using near infrared spectroscopy. *Meat Science*, *94*(4), 455–460. doi:10.1016/j.meatsci.2013.03.003

De Marchi, M., Penasa, M., Battagin, M., Zanetti, E., Pulici, C., & Cassandro, M. (2011). Feasibility of the direct application of near-infrared reflectance spectroscopy on intact chicken breasts to predict meat color and physical traits. *Poultry Science*, *90*(7), 1594–1599. doi:10.3382/ps.2010-01239

Dixit, Y., Casado-Gavalda, M. P., Cama-Moncunill, R., Cama-Moncunill, X., Markiewicz-Keszycka, M., Cullen, P. J., & Sullivan, C. (2017). Developments and Challenges in Online NIR Spectroscopy for Meat Processing. *Comprehensive Reviews in Food Science and Food Safety*, *16*(6), 1172–1187. doi:10.1111/1541-4337.12295

Food and Agriculture Organization of the United Nations. (2016). Meat & meat products. Retrieved November 30, 2017, from http://www.fao.org/ag/againfo/themes/en/meat/home.html

Grunert, K. G., Bredahl, L., & Brunsø, K. (2004). Consumer perception of meat quality and implications for product development in the meat sector — a review §, *66*, 259–272. doi:10.1016/S0309-1740(03)00130-X

Hildrum, K. I., Nilsen, B. N., Mielnik, M., & Næs, T. (1994). Prediction of sensory characteristics of beef by near-infrared spectroscopy. *Meat Science*, *38*(1), 67–80.

Liu, Y., Lyon, B. G., Windham, W. R., Realini, C. E., Pringle, T. D. D., & Duckett, S. (2003). Prediction of color, texture, and sensory characteristics of beef steaks by visible and near infrared reflectance spectroscopy. A feasibility study. *Meat Science*, *65*(3), 1107–1115. doi:10.1016/S0309-1740(02)00328-5

Prevolnik, M., Čandek-Potokar, M., & Škorjanc, D. (2004). Ability of NIR spectroscopy to predict meat chemical composition and quality - A review. *Czech Journal of Animal Science*, *49*(11), 500–510.

Prieto, N., Andrés, S., Giráldez, F. J., Mantecón, A. R., & Lavín, P. (2008). Ability of near infrared reflectance spectroscopy (NIRS) to estimate physical parameters of adult steers (oxen) and young cattle meat samples. *Meat Science*, *79*(4), 692–699. doi:10.1016/j.meatsci.2007.10.035

Prieto, N., Lopez-Campos, O., Aalhus, J. L., Dugan, M. E. R., Juarez, M., & Uttaro, B. (2014). Use of near infrared spectroscopy for estimating meat chemical composition, quality traits and fatty acid content from cattle fed sunflower or flaxseed. *Meat Science*, *98*(2), 279–288. doi:10.1016/j.meatsci.2014.06.005

Prieto, N., Roehe, R., Lavín, P., Batten, G., & Andrés, S. (2009). Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review. *Meat Science*, *83*(2), 175–186. doi:10.1016/j.meatsci.2009.04.016

Prieto, N., Ross, D. W., Navajas, E. A., Nute, G. R., Richardson, R. I., Hyslop, J. J., Simm, G., Roehe, R. (2009). On-line application of visible and near infrared reflectance spectroscopy to predict chemical-physical and sensory characteristics of beef quality. *Meat Science*, *83*(1), 96–103. doi:10.1016/j.meatsci.2009.04.005

Pullanagari, R. R., Yule, I. J., & Agnew, M. (2015). On-line prediction of lamb fatty acid composition by visible near infrared spectroscopy. *Meat Science*, *100*, 156–163. doi:10.1016/j.meatsci.2014.10.008

Ripoll, G., Albertí, P., Panea, B., Olleta, J. L., & Sañudo, C. (2008). Near-infrared reflectance spectroscopy for predicting chemical, instrumental and sensory quality of beef. *Meat Science*, *80*(3), 697–702. doi:10.1016/j.meatsci.2008.03.009

Rødbotten, R., Nilsen, B. N., & Hildrum, K. I. (2000). Prediction of beef quality attributes from early post mortem near infrared reflectance spectra. *Food Chemistry*, *69*(4), 427–436. doi:10.1016/S0308-8146(00)00059-5

Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., & Jent, N. (2007). A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *Journal of*

Pharmaceutical and Biomedical Analysis, *44*(3 SPEC. ISS.), 683–700. doi:10.1016/j.jpba.2007.03.023

Shenk, J.S., Workman, J.J., & Westerhaus, M.O. (2007). Application of NIR spectroscopy to agricultural products. In D. Burns & E. Ciurczak (Eds.), *Handbook of near infrared analysis* (3rd ed., pp. 347–389). New York: CRC Press.

Troy, D. J., Ojha, K. S., Kerry, J. P., & Tiwari, B. K. (2016). Sustainable and consumer-friendly emerging technologies for application within the meat industry: An overview. *Meat Science*, *120*, 2–9. doi:10.1016/j.meatsci.2016.04.002

Weeranantanaphan, J., Downey, G., Allen, P., & Sun, D.-W. (2011). A review of near infrared spectroscopy in muscle food analysis: 2005-2010. *J Near Infrared Spectrosc*, *19*(March), vii–ix. doi:10.1255/jnirs.xxx

Williams, P. (2014). The RPD statistic: A tutorial note. *NIR News*, *25*, 22–26.

Williams, P., & Norris, K. H. (2001). *Near-infrared technology : in the agricultural and food industries*. American Association of Cereal Chemists.

Yancey, J. W. S., Apple, J. K., Meullenet, J. F., & Sawyer, J. T. (2010). Consumer responses for tenderness and overall impression can be predicted by visible and near-infrared spectroscopy, Meullenet-Owens razor shear, and Warner-Bratzler shear force. *Meat Science*, *85*(3), 487–492. doi:10.1016/j.meatsci.2010.02.020

Table 1. Range, mean, standard deviation (SD) and coefficient of variation (CV) of thirteen beef *LTL* eating quality traits (sensory scores measured on a scale of 0-100).

Attribute	*n*	Range	Mean	SD	CV (%)
WBSF (N)	595	22.05 – 96.77	43.3	10.8	24.94
Tenderness	209	23.25 – 70.75	52.34	10.47	20.0
Chewiness	209	6.33 – 53.33	26.92	11.17	41.49
Stringiness	209	0.36 – 52.55	11.56	9.45	81.75
Crumbliness	209	1.4 – 56.94	19.22	12.65	65.82

Juiciness	209	17.5 – 54.35	33.26	6.88	20.69
Difficulty Swallowing	209	0.83 – 46	12.33	7.78	63.1
Fatty mouthfeel	209	0.21 – 20.37	5.7	4.45	78.07
Beef flavour	209	7.11 – 60.87	40.25	7.11	17.66
Metallic flavour	209	1.42 – 29.5	13.9	5.52	39.71
Beef AE	209	6.22 – 45	29.88	6.22	20.81
Metallic AE	209	2.71 – 37.62	17.3	7.41	42.83
Fatty AE	209	0.5 – 24.25	5.58	4.74	84.95

WBSF, Warner-Bratzler shear force; N, Newtons; SD, Standard Deviation; CV, Coefficient of Variation; AE, after-effect; Sensory scores are measured on a scale of 0-100

Table 2. Pearson correlation coefficients between instrumental and trained panel sensory quality attributes of beef *LTL*.

	Tenderness	Juiciness	Chewiness	Stringiness	Crumbliness	Fatty mouthfeel	Difficulty Swallowing	Beef flavour	Metallic flavour	Beefy AE	Metallic AE	Fatty AE
WBSF	- 0.39 ***	- 0.26 ***	0.43 ***	0.33 ***	- 0.39 ***	0.01	0.44 ***	- 0.22 ***	0.12	- 0.27 ***	0.03	- 0.07
Tenderness		0.3 ***	- 0.86 ***	- 0.72 ***	0.1	- 0.39 ***	- 0.74 ***	0.09	- 0.27 ***	0.2 *	- 0.35 ***	- 0.27 ***
Juiciness			- 0.26 ***	- 0.92	0.47 ***	0.24 **	- 0.26 ***	0.33 ***	0.17	0.38 ***	0.23 **	0.29 ***
Chewiness				0.77 ***	- 0.13	0.43 ***	0.84 ***	- 0.09	0.36 ***	- 0.14	0.41 ***	0.31 ***
Stringiness					0.12	0.61 ***	0.79 ***	0.03	0.39 ***	- 0.02	0.42 ***	0.45 ***
Crumbliness						0.47 ***	- 0.08	0.35 ***	0.34 ***	0.26 ***	0.43 ***	0.16
Fatty mouthfeel							0.46 ***	0.29 ***	0.54 ***	0.2 *	0.61 ***	0.82 ***
Difficulty swallowin								- 0.14 *	0.38 ***	- 0.2 *	0.39 ***	0.34 ***

g				
Beef flavour	0.04	0.72 ***	0.22 *	0.39 ***
Metallic flavour		0.03	0.67 ***	0.53 ***
Beef AE			0.33 ***	0.36 ***
Metallic AE				0.58 ***

Table 3. Best fitting predictions of thirteen sensory traits on beef *LTL* using VISNIR spectral measurements.

Variable	Day PM	PM Time	Math treatment	Range (nm)	*n*	*p*	R²Cal	RMSEC	R²CV	RMSECV	RPD
WBSF (N)	1	24	None	350 - 2500	595	8	0.23	9.53	0.14	10	1.08
	2	48	None	450 - 2300	450	9	0.3	9.52	0.22	10.1	1.07
Tenderness	1	25	None	450 - 779	209	5	0.12	9.9	0.06	10.3	1.02
	2	49	None	350 - 2500	160	6	0.25	8	0.13	8.6	1.22
Juiciness	1	24	SNV	350 - 2500	205	4	0.15	6.4	0.04	6.8	1.01
	2	49	None	450 - 779	160	2	0.11	6.6	0.09	6.8	1.01
Chewiness	1	25	None	450 - 779	209	6	0.18	10.2	0.11	10.6	1.05
	2	48	None	450 - 779	166	10	0.38	7.9	0.17	9.3	1.2
Stringiness	1	24	SNV	780 - 1099	205	10	0.36	7.6	0.15	8.8	1.07
	2	48	None	450 - 779	166	9	0.37	6.9	0.22	7.7	1.23
Crumbliness	1	25	None	350 - 2500	209	10	0.54	8.5	0.34	10.5	1.2
	2	48	None	350 - 2500	166	7	0.52	8.4	0.41	9.4	1.35
Fatty mouthfeel	1	25	SNV	350 - 2500	209	10	0.53	3	0.17	4	1.11
	2	49	None	780 - 1099	160	10	0.36	2.9	0.23	3.2	1.37
Difficulty swallowing	1	25	None	450 - 779	209	6	0.17	7.1	0.09	7.5	1.04
	2	49	None	450 -	16	1	0.34	6.1	0.1	7.1	1.1

				779	0	0					
Beef flavour	1	24	SNV	450 - 779	205	9	0.25	6.2	0.11	6.8	1.04
	2	48	None	350 - 2500	166	5	0.23	6	0.13	6.4	1.11
Metallic flavour	1	24	None	780 - 1099	205	10	0.36	4.38	0.17	5	1.1
	2	48	None	450 - 779	166	9	0.32	4.26	0.11	4.8	1.15
Beef AE	1	24	SNV	1100 - 2300	205	3	0.09	5.9	0.04	6.1	1.08
	2	49	SNV	780 - 1099	160	10	0.38	4.8	0.2	5.4	1.22
Metallic AE	1	24	None	780 - 1099	205	10	0.36	5.9	0.19	6.6	1.12
	2	48	None	350 - 2500	166	6	0.23	6.3	0.08	6.9	1.07
Fatty AE	1	24	None	350 - 2500	205	9	0.46	3.5	0.26	4.1	1.16
	2	49	SG	780 - 1099	160	7	0.41	2.9	0.28	3.2	1.48

PM, post-mortem; *n*, number of samples; *p*, number of PLS terms utilized in the calibration equation; R^2Cal, coefficient of determination of calibration; RMSEC, root mean square error of calibration; R^2CV, coefficient of determination of cross validation; RMSECV, root mean square error of cross validation; RPD, ratio performance deviation; WBSF, Warner-Bratzler shear force; (N), Newton; None, no mathematical treatment; SG, Savitzky-Golay smoothing – 2nd derivative with 2nd polynomial and 21 smoothing points; SNV, standard normal variate; Diff. swallow, AE, after-effect