final report

Investigating the viability of using MEXA for objective measurement of IMF and WBSF in Australian new and old seasons lambs

Abstract

The lamb Meat Standards Australia (MSA) grading system utilises no individual carcase measurements, apart from specifying minimum weight and fat thresholds, and therefore cannot differentiate eating quality between individual carcases. There is a strong call from industry to transition the lamb MSA grading system to an individual carcase grading system; however, this is impeded by the lack of individual carcase measurement for traits that influence lamb eating quality. Through extensive eating quality research, two traits have been identified as key drivers of eating quality in lamb: intramuscular fat percentage and Warner-Bratzler shear force. Technologies that can measure these two traits on-line at abattoir chain speed would enable the roll-out of an individual carcase grading system for lamb.

This study explored a number of analytical and calibration methods that are likely suitable for the development of Multi-energy X-ray Absorptiometry (MEXA) technology. It also highlighted the potential for MEXA to differentiate tissue components and predict complex multi-factorial traits such as shear force, as well as simpler traits such as intramuscular fat percentage, with good precision and accuracy. The promising outcomes from this work must be weighed with some conservatism, as the scale of the study provided only limited opportunity for validation testing, and the precision upon validation was significantly eroded. Nonetheless, these outcomes have highlighted the potential of this technology, and provided an opportunity to refine the analytical and calibration methods for future MEXA analysis. This will be crucial given MLA's investment in this technology.

Executive summary

The rationale for undertaking this study

This was an opportunistic study, undertaken while Spanish collaborators had access to the Multiscan Technologies Multi-energy X-ray Absorptiometry (MEXA) scanner. It considered the feasibility of a range of analytical and calibration methods that are likely suitable for the development of MEXA technology. This is crucial given MLA's investment in MEXA technology. We explored simple linear associations, partial least squares regression and, finally, extreme gradient boosting, with the latter suggesting some potential.

This study was also a chance to investigate the potential for MEXA to differentiate tissue components and predict complex multi-factorial traits such as Warner-Bratzler shear force, as well as simpler traits such as intramuscular fat percentage (IMF). Previous studies have already demonstrated that Dual-Energy X-ray Absorptiometry (DEXA) can measure the proportion of fat in a carcase (Gardner et al., 2018) based upon the ratio of attenuation of high and low energy x-rays passing through it. Fundamentally, this approach relies upon the physical principle of the ratio of the attenuation of high and low energy photons correlating with the atomic mass of the substance through which they pass (Pietrobelli et al., 1996). Differing atoms have varying k-edge attenuation which can influence these ratio values. On this basis, if multiple energies were detected as in MEXA (as opposed to only two energies as in DEXA), better differentiation of tissue types may be possible, particularly where multiple tissues may be involved, as is the case with shear force. As such, we explored MEXA's precision in differentiating tissue components and predicting both complex multi-factorial traits (i.e. shear force) and simpler traits (i.e. IMF).

The experimental design and methodology

Forty Merino-cross lambs were sourced from a commercial sheep flock located in Bordertown, South Australia. Half were born in September/October 2016, the other half in May/June 2017. All were slaughtered in October 2017 at a commercial abattoir as lambs (6 months old) and yearlings (12 months old) respectively. All lambs were maintained in the same paddock 3 weeks prior to slaughter on pasture with supplementary hay or pellet feeding when necessary. After slaughter, 3 samples 5cm long were taken from the *m.longissimus dorsi* near to the 12/13th rib, aged for 5 days and then frozen. The first sample was used to determine shear force, the second to determine intramuscular fat percentage, and the third was used for MEXA scanning.

Prior to MEXA scanning, samples were cut into two sections, creating sample 'repetition a' and sample 'repetition b'. All samples were scanned twice through the MEXA, effectively enabling a machine repetition comparison between 'run 1' and 'run 2'. The system acquired a spectroscopic image of the entire sample with each pixel containing an X-ray energy spectra of 128 channels. Within each spectroscopic image, a region of interest within the *longissimus dorsi* muscle was selected corresponding to a uniform section of the sample, while also maximising the sample area. Hence, the acquired information was a 3D matrix of variable pixel numbers x 128 channels.

The data was then transformed using three different methods. In the first method (referred to hereafter as calibration method 'none'), the mean of the raw pixel values for each sample was calculated for each of the 128 channels, resulting in 128 values for each sample. In the second method (referred to hereafter as calibration method 'minus'), the 128 values for each image were corrected against the clear-field region of the image. In the third method (referred to hereafter as calibration method 'log'), the 128 values for each image were corrected against the clear-field region of the image and log transformed.

Following the unsuccessful use of Partial Least Squares Regression to predict shear force, an alternative (machine learning) approach was taken. A predictive model was established using Extreme Gradient Boosting coupled with Bayesian optimisation, constructed using AutoStat® software. A 5-fold cross-validation procedure was used, with models trained in 4 groups and validated in the 5th, with this process repeated 5 times until models had been tested in each of the 5 groups. Shear force and intramuscular fat percentage were predicted within each of the 4 datasets, for run 1, repetition a and b, and for run 2, repetition a and b. Data from each calibration method ('none', 'minus', 'log') was analysed, and the average root mean square error of the prediction (RMSEP) across the 5 validation tests was recorded in each instance. A general linear model within AutoStat® was then used to generate the coefficient of determination (R^2), slope and bias of the relationship between actual versus predicted values for shear force and for intramuscular fat percentage.

The achievements of this study

This study demonstrated the potential of extreme gradient boosting as an analytical method suitable for exploring MEXA prediction of the commercially-relevant traits of shear force and intramuscular fat percentage, as evidenced by good precision within the trained models. The consistency of the trained predictions for the repeat scan of the same samples, and the within-sample repetitions, provide some level of confidence that the predictions were repeatable. This needs to be demonstrated upon validation to be confident that these results are not simply one-off random associations.

MEXA's capacity to predict both shear force and intramuscular fat percentage must be taken with cautious conservatism, as the models were imprecise and inaccurate upon validation. There were significant limitations to the rigour of the validation testing due to the experimental design of this opportunistic study. Nonetheless, these outcomes have highlighted the potential of this technology, and provided an opportunity to refine the analytical and calibration methods for future MEXA analysis.

The industry benefits of this study

MLA's strategy is to pursue objective carcase measurement. The findings from this study inform the numerous MLA projects that deploy x-ray technology for objective carcase measurement and the broader research being undertaken within the associated commercial companies, accelerating their commercial delivery. In particular, technologies (potentially MEXA) that can measure Warner-Bratzler shear force and intramuscular fat percentage—key drivers of eating quality in lamb—on-line at abattoir chain speed would meet industry's demand to transition from the existing pathways-based lamb MSA grading system to an individual carcase grading system.

In the short-term, the commercial deployment of a MEXA imaging system is to couple it with a conveyor, allowing product to smoothly pass between x-ray tube and detector. However, one of the limitations of the present methodology is the need to scan trimmed samples of muscle devoid of fat or bone tissue. This is due to the 2-dimensional nature of the image captured, and the confounding of multiple tissue types within each pixel. Commercial meat cuts consist of lean surrounded by layers of subcutaneous and intermuscular fat, and in some cases bone. In a 2-dimensional image, this results in mixtures of tissue types within each pixel, in particular confounding intramuscular fat depots with subcutaneous and intermuscular fat. We speculate that a machine learning approach may find relevant associations when analysing entire images that capture differentials between pixels; however, we suspect that this will be difficult to validate across datasets. The alternative would be to progress to a computed tomography MEXA image, enabling identification of tissue regions within 3-dimensions, and the isolation of key regions of interest (i.e. specific muscle groups). These devices are becoming more common-place, both in medical scanning and commercial baggage scanning industries, and should continue to be explored.

Table of contents

1   Background

Compared to other proteins, lamb meat is relatively expensive, hence consumer satisfaction is vital to underpin demand. Eating quality is strongly linked to consumer satisfaction (Grunert, Bredahl & Brunsø, 2004; Henchion et al., 2014), thus traits that predict eating quality are vital (Pethick et al., 2006). In an effort to predict the eating quality of the end product for consumers, the MSA grading system was developed in 1998 for beef (Polkinghorne et al., 2008) and in 2005 for sheepmeat. The beef system utilises a series of carcase measurements and other information about its on-farm and processing history to predict the eating quality of all commercially relevant cuts, enabling consumers to purchase beef based upon an expected eating quality (Polkinghorne et al., 2008). By contrast, the lamb system utilises no individual carcase measurements, apart from specifying minimum weight and fat thresholds, and therefore cannot differentiate eating quality between individual carcases. It is limited to simply identifying those that meet the minimum threshold for satisfactory consumer eating quality. There is a strong call from industry to transition the lamb MSA grading system to an individual carcase grading system; however, this is impeded by the lack of individual carcase measurements for traits that influence lamb eating quality.

Through extensive eating quality research, two traits have been identified as key drivers of eating quality in lamb. The first is intramuscular fat percentage (IMF), which strongly influences juiciness, flavour, and—to a lesser extent—tenderness of the meat (Pannier et al., 2014a). A 4.5% change in IMF could increase lamb juiciness, flavour and overall liking scores by as much as 11 points across a 100 point scale (Pannier et al., 2018). The second trait is Warner-Bratzler shear force, a measurement closely related to the tenderness of meat and reflective of a number of intrinsic muscle characteristics such as sarcomere length, connective tissue content and the cross-linked structure of the connective tissue. In this case, a 37 Newton reduction in shear force was associated with an 11.6 unit increase in consumer tenderness scores for lamb meat (Pannier et al., 2014a). On this basis, technologies that can measure either IMF or shear force on-line at abattoir chain speed would enable the roll-out of an individual carcase grading system for lamb, using IMF and shear force as the key differentiators of eating quality.

One technology that may have the capacity to measure eating quality traits is Multi-Energy X-ray Absorptiometry (MEXA). Previous studies have already demonstrated that Dual-Energy X-ray Absorptiometry (DEXA) can measure the proportion of fat in a carcase (Gardner et al., 2018) based upon the ratio of attenuation of high and low energy x-rays passing through it. Fundamentally, this approach relies upon the physical principle of the ratio of the attenuation of high and low energy photons correlating with the atomic mass of the substance through which they pass (Pietrobelli et al., 1996). On this basis, x-ray attenuation ratio values should be able to determine the IMF content of dissected meat cuts where the subcutaneous fat has been removed. Another tissue that may vary in structure and total content within meat is collagen, a connective tissue protein that strongly influences shear force and therefore consumer tenderness scores (Young & Braggins, 1993). To date, no studies have explored the capacity to identify differing collagen levels using x-ray technologies.

Differing atoms have varying k-edge attenuation which can influence these ratio values. On this basis, if multiple energies were detected (as opposed to only two energies), better differentiation of tissue types may be possible, particularly where multiple tissues may be involved, as is the case with shear force. As such, MEXA offers potential to differentiate tissue components and predict complex multi-factorial traits such as shear force, or simpler traits such as IMF, with more precision.

2   Milestone description

This report delivers on all three milestones described in the contract V.TEC.1712. This report also appends another report compiled with our Spanish collaborators, which details corresponding estimation of these traits using near infra-red (NIR) imaging. Although not required as part of this contract, we have decided to report this centrally to enable direct comparison of technologies. The milestones specific to this project include:

2.1   Milestone 1

1.1 – Report on processing of samples for MEXA imagery and collation of data with phenotypic information.

2.2   Milestone 2

2.1 – Undertake analyses to develop and validate predictive models for Warner-Bratzler shear force.
2.2 – Undertake analyses to develop and validate predictive models for intramuscular fat percentage.

2.3   Milestone 3

3.1 – Assess the feasibility of MEXA for objectively measuring individual Australian lamb carcases.
3.2 – Assess the accuracy of objective measurements at chain speed in Australian lambs.
3.3 – Assess the suitability of MEXA for Meat Standards Australia objective measurements.

3   Project objectives

3.1   Objective 1

Collate MEXA data with phenotypic information of twenty new season (4-6 month old) and twenty old season (10-12 month old) lamb carcases for analysis.

3.2   Objective 2

Analyse MEXA data and validate predictive models for Warner-Bratzler shear force and intramuscular fat percentage.

3.3   Objective 3

Assess feasibility and accuracy of MEXA operating at chain speed.

4   Methodology

4.1   Samples

Forty Merino-cross lambs were sourced from a commercial sheep flock located in Bordertown, South Australia. Half were born in September/October 2016, the other half in May/June 2017. All were slaughtered in October 2017 at a commercial abattoir as lambs (6 months old) and yearlings (12 months old) respectively. All lambs were maintained in the same paddock 3 weeks prior to slaughter on pasture with supplemental hay or pellet feeding when necessary. After slaughter, 3 samples 5cm long were taken from the *m.longissimus dorsi* near to the 12/13th rib, aged for 5 days and then frozen. The first sample was used to determine shear force, the second to determine intramuscular fat percentage, and the third was used for MEXA scanning.

4.2   MEXA scanning

Prior to MEXA scanning, samples were cut into two sections, creating sample 'repetition a' and sample 'repetition b'. All samples were scanned twice through the MEXA, effectively enabling a machine repetition comparison between 'run 1' and 'run 2'.

An industrial X-ray prototype system with a multi energy detector (MXV-PACK 4010, Multiscan Technologies, Cocentaina, Spain) was used to scan the samples. The prototype had an X-ray spectrometric detector made of a semiconductor crystal (CDTe/CZT) with a pixel size of 0.8 mm. A belt conveyor transported the sample at a speed of 10 m/min. Simultaneously, X-rays were emitted from below the samples using a tungsten X-ray tube that was pulsed at 140kV. The resulting polychromatic array of photons was detected and transmitted radiation recorded across 128 channels, corresponding to energies from 20 to 160 keV. Thus, the system acquired a spectroscopic image of the entire sample with each pixel containing an X-ray energy spectra of 128 channels.

4.3   Spectral analysis

Firstly, within each spectroscopic image, a region of interest within the *longissimus dorsi* muscle was selected corresponding to a uniform section of the sample, while also maximising the sample area. Hence, the acquired information was a 3D matrix of variable pixel numbers x 128 channels. This data was then transformed using three different methods.

In the first method (referred to hereafter as calibration method 'none'), the mean of the raw pixel values for each sample was calculated for each of the 128 channels, resulting in 128 values for each sample.

In the second method (referred to hereafter as calibration method 'minus'), the 128 values for each image were corrected against the clear-field region of the image. This was done by isolating the clear-field region around each of the images and then calculating the average of these pixel values. This process was undertaken for each of the 128 channels. Thus, 128 separate clear-field values were calculated for each sample, with each clear-field value corresponding to the intensity of the incident radiation (I_o) at each pixel for that channel. The clear-field value was then subtracted from each of the image pixel values according to the following formula:

$$S_a = \frac{\sum_{i=}^{p} I_f^{a,i} - I_o^{a,i}}{p} \qquad \text{(Equation 1)}$$

where I_f is the intensity of the transmitted radiation through the sample and I_o is the intensity of the incident radiation at each pixel i of the region of interest, which contains p pixels. The calculation was done for each energy channel of the spectra a, which ranges from 1 to 128 and represents a given X-ray energy. According to Equation 1, an increase in the S_a value represents a decrease in the X-ray attenuation for channel a.

In the third method (referred to hereafter as calibration method 'log'), the 128 values for each image were corrected against the clear-field region of the image and log transformed. The clear-field region of each image was isolated, and the 128 separate clear-field values were calculated using the method described above, with each clear-field value corresponding to the intensity of the incident radiation (I_o) at each pixel for that channel. The pixel values of the sample were then expressed as a ratio against the incident radiation and log transformed according to the following formula:

$$S_b = \frac{-\sum_{i=1}^{p} \ln\left(\frac{I_f^{b,i}}{I_o^{b,i}}\right)}{p} \qquad \text{(Equation 2)}$$

where I_f is the intensity of the transmitted radiation and I_o is the intensity of the incident radiation at each pixel i of the region of interest, which contains p pixels. The calculation was done for each energy channel of spectra b, which ranges from 1 to 128 and represents a given X-ray energy. According to Equation 2, an increase in the S_b value represents an increase in the X-ray attenuation for channel b.

4.4   Tenderness and intramuscular fat percentage measurement

Approximately 65 g loin samples were used for shear force testing as described by Pannier et al. (2014a). Initially, samples were vacuum packed and aged for 5 days, and then frozen at −20 °C until subsequent testing. Frozen samples were cooked in plastic bags in a water bath for 35 minutes at 71 °C, and were cooled in running water for 30 minutes after cooking. Six cores (~3–4 cm long, 1 cm^2 cross sections) from each loin sample were cut and shear force was measured using a Lloyd texture analyser (Model LRX, Lloyd Instruments, Hampshire, UK) with a Warner–Bratzler shear blade fitted as described by Hopkins, Toohey, Warner, Kerr & van de Ven (2010).

To determine IMF%, approximately 40 g of diced loin muscle was collected. Samples were stored at −20 °C until they were subsequently freeze-dried using a Cuddon FD 1015 freeze dryer (Cuddon Freeze Dry, NZ). The IMF content was determined using a near infra-red procedure in a Technicon InfraAlyzer 450 (19 wavelengths) (Perry, Shorthose, Ferguson & Thompson, 2001). NIR readings were validated against chemical fat determinations using solvent extraction. IMF was expressed as percentage fat.

4.5   Construction and cross-validation of predictive models

4.5.1   Partial Least Squares Regression Analysis

Calibration models were developed using Partial Least Squares Regression (PLSR) for shear force in lamb meat. Validations were made by a cross-validation (Venetian blinds) model. The best combination of pre-treatments was selected based on the lowest error of prediction and the lowest number of PLS factors. Partial Least Squares discriminant analysis (PLS-DA) was used as a linear classification method to predict the class (new or old). When needed, interval variable selection was applied to discard variables which may be adding complexity to a model and improve performance of a final model. This analysis method proved to be unsuccessful; however, as a point of reference, these results are still provided in the appended report submitted to the Spanish Institute of Food Research and Technology, the funding body supporting our Spanish collaborators.

4.5.2   Extreme Gradient Boosting Analysis

Following the unsuccessful use of PLSR to predict shear force, an alternative (machine learning) approach was taken. A predictive model was established using Extreme Gradient Boosting coupled with Bayesian optimisation, constructed using AutoStat® software. A 5-fold cross-validation procedure was used, with models trained in 4 groups and validated in the 5th, with this process repeated 5 times until models had been tested in each of the 5 groups. Shear force and intramuscular fat percentage were predicted within each of the 4 datasets, for run 1, repetition a and b, and for run 2, repetition a and b. Data from each calibration method ('none', 'minus', 'log') was analysed, and the average root mean square error of the prediction (RMSEP) across the 5 validation tests was recorded in each instance. The RMSEP represented the bounds within which 2/3 of the actual data lies from the predicted.

A general linear model within AutoStat® was then used to generate the coefficient of determination (R^2), slope and bias of the relationship between actual versus predicted values for shear force and for intramuscular fat percentage. Bias represented the difference between the predicted and actual values calculated at the mean of the observations, while slope reflected that this bias may differ at

observations above or below the mean. The RMSEP and R^2 are reported as indicators of precision and the slope and bias are reported as indicators of accuracy.

5 Results

The descriptive statistics for all forty samples are shown in Table 1 below.

Table 1. The descriptive statistics for intramuscular fat percentage (IMF), hot carcase weight (HCWT), fat depth measured at the C-site (Cfat), eye muscle area (EMA) and shear force as measured on day 5 (SF5).

|	IMF (%)	HCWT (kg)	Cfat (mm)	EMA (cm²)	SF5 (N)
Mean	4.7005	22.3350	3.2338	13.8720	34.8649
Standard Deviation	1.3232	2.1727	1.4765	2.7248	8.4468
Minimum	3.0769	19.2000	0.5800	9.3870	24.6823
Maximum	8.5317	31.9000	6.4200	21.2212	56.5589

5.1 Shear force prediction

The results for shear force prediction are shown in Table 2 plus Figures 1 and 2 below. Figure 1 represents the "trained" prediction, while Figure 2 represents the "validated" prediction.

Table 2. The root mean square error (RMSE), coefficient of determination (R^2), root mean square error of prediction (RMSEP), slope and bias for each combination of run (1, 2), repetition (a, b), and calibration method ('none', 'minus', 'log') for prediction of shear force (N).

Calibration	Statistic	Run 1 Rep a	Run 1 Rep b	Run 2 Rep a	Run 2 Rep b	Mean
'None'		*Training performance*				
	RMSE	0.2113	0.2317	0.2313	0.2700	0.2361
	R²	0.9413	0.9295	0.9315	0.9043	0.9267
		Validation performance				
	RMSEP	0.9375	1.1003	1.0303	0.9688	1.0092
	R²	0.0057	0.0618	0.0100	0.0043	0.0205
	slope	0.0928	-0.3356	-0.1509	-0.0898	-0.1209
	bias	1.0487	-0.2253	-0.9961	-0.1791	-0.0880
'Minus'		*Training performance*				
	RMSE	0.2682	0.3058	0.2050	0.2700	0.2623
	R²	0.9055	0.8772	0.9448	0.9043	0.9080
		Validation performance				
	RMSEP	1.0214	0.9236	1.0759	1.0469	1.0170
	R²	0.0447	0.0011	0.0472	0.0178	0.0277
	slope	-0.3060	0.0487	-0.2600	-0.1662	-0.1709
	bias	-0.0954	-0.0963	1.1748	0.9556	0.4847
'Log'		*Training performance*				
	RMSE	0.2441	0.2636	0.2502	0.2089	0.2417
	R²	0.9218	0.9088	0.9178	0.9427	0.9228
		Validation performance				
	RMSEP	1.0250	0.9006	1.0122	0.9378	0.9689
	R²	0.0825	0.0046	0.0087	0.0024	0.0246
	slope	-0.4228	0.1179	-0.1215	0.0704	-0.0890
	bias	-0.3820	-0.1766	0.9038	0.0542	0.0999

(A)   **(B)**

(C)   **(D)**
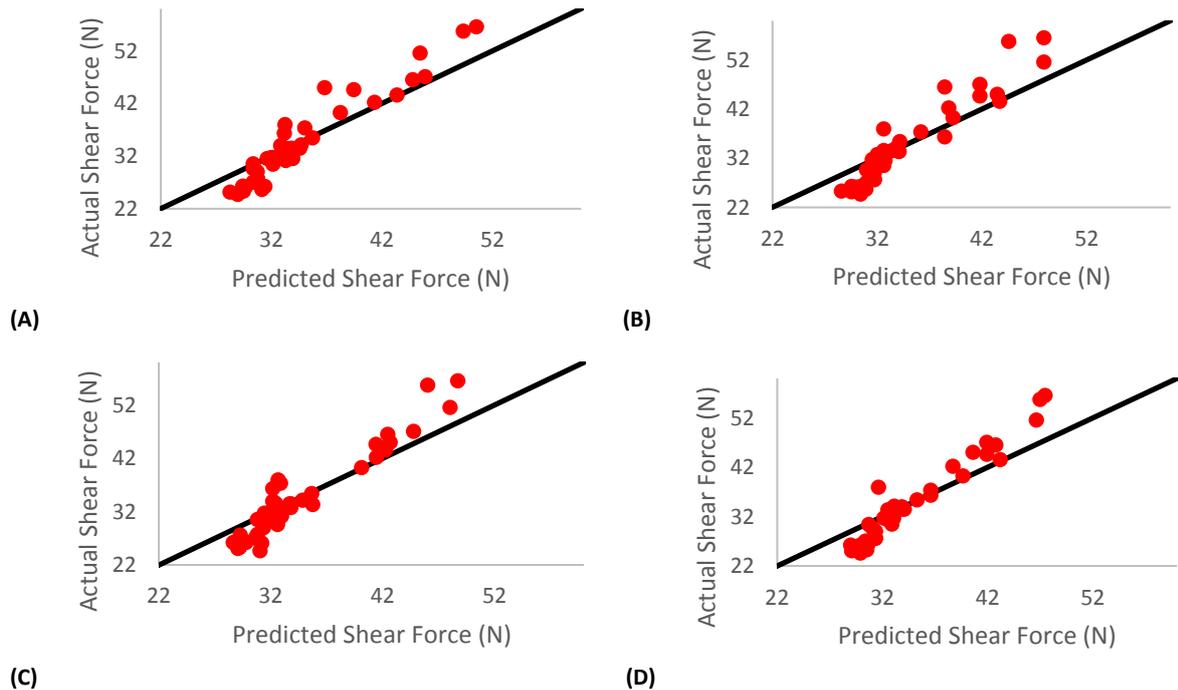
Figure 1. "Trained" predicted versus actual shear force for calibration method 'log': **(A)** run 1, repetition a; **(B)** run 1, repetition b; **(C)** run 2, repetition a; **(D)** run 1, repetition b. Line represents the outcome of a perfect fit between predicted and actual values.

(A)   **(B)**

(C)   **(D)**
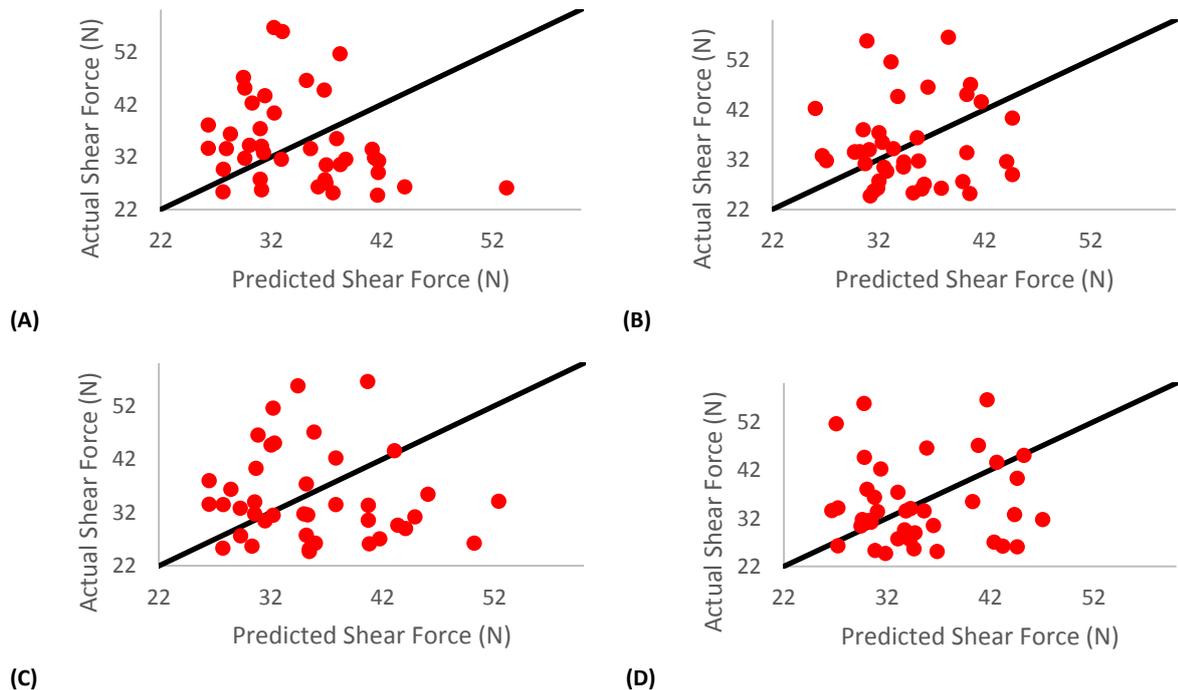
Figure 2. "Validated" predicted versus actual shear force for calibration method 'log': **(A)** run 1, repetition a; **(B)** run 1, repetition b; **(C)** run 2, repetition a; **(D)** run 1, repetition b. Line represents the outcome of a perfect fit between predicted and actual values.

Within the training analysis, the precision of predicting shear force was high across each of the 'none', 'minus', and 'log' calibration methods, and consistent across runs and sample repetitions. However, upon validation, this performance was markedly diminished, with the magnitude of the RMSEP, indicatively about 9.8070N across all tests, representing about 30% of the full 31.8766N range in shear force raw data.

Given the poor precision upon validation, the bias and slope estimates were highly variable across each of the 'none', 'minus', and 'log' calibration methods.

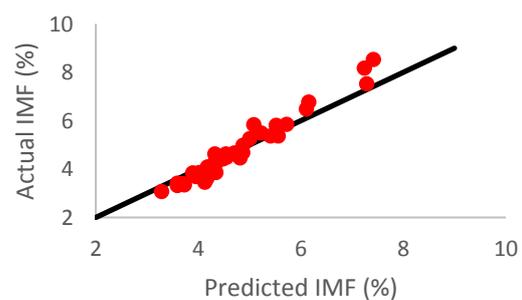## 5.2 Intramuscular fat percentage prediction

The results for intramuscular far percentage prediction are shown in Table 3 plus Figures 3 and 4 below. Figure 3 represents the "trained" prediction, while Figure 2 represents the "validated" prediction.

Table 3. The root mean square error (RMSE), coefficient of determination (R^2), root mean square error of prediction (RMSEP), slope and bias for each combination of run (1, 2), repetition (a, b), and calibration method ('none', 'minus', 'log') for prediction of intramuscular fat (%).

Calibration	Statistic	Run 1 Rep a	Run 1 Rep b	Run 2 Rep a	Run 2 Rep b	Mean
		Training performance				
'None'	**RMSE**	0.1452	0.3171	0.1324	0.1801	0.1937
	R^2	0.9883	0.9440	0.9904	0.9819	0.9762
		Validation performance				
	RMSEP	1.5006	1.4004	1.6140	1.5989	1.5285
	R^2	0.0049	0.0205	0.0060	0.0302	0.0154
	slope	0.0994	0.2409	-0.1143	-0.2955	-0.0174
	bias	0.0725	0.1755	0.1884	0.1140	0.1376
'Minus'		*Training performance*				
	RMSE	0.4516	0.2018	0.1917	0.3546	0.2817
	R^2	0.8865	0.9773	0.9795	0.9300	0.9433
		Validation performance				
	RMSEP	1.5676	1.6046	1.5443	1.4851	1.5504
	R^2	0.0475	0.0154	0.0083	0.0003	0.0179
	slope	-0.4111	-0.1966	-0.1562	-0.0337	-0.1994
	bias	0.0539	0.1752	0.0904	-0.0532	0.0666
'Log'		*Training performance*				
	RMSE	0.2496	0.2457	0.2354	0.2743	0.2513
	R^2	0.9653	0.9664	0.9692	0.9581	0.9648
		Validation performance				
	RMSEP	1.6167	1.7386	1.5724	1.5489	1.6192
	R^2	0.0908	0.0180	0.0015	0.0052	0.0289
	slope	-0.5816	-0.1779	-0.0562	-0.1176	-0.2333
	bias	0.1653	0.3271	0.0737	0.2033	0.1924
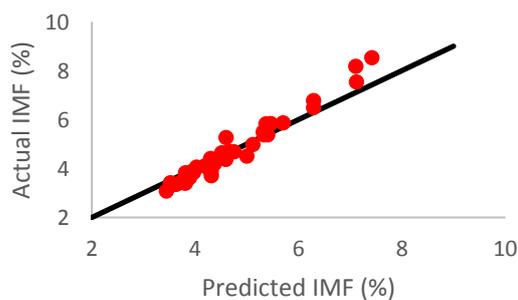
(A)   **(B)**
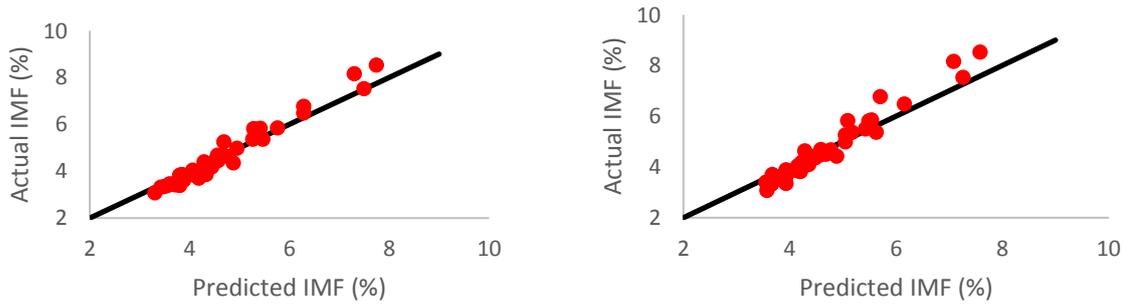
(C)   **(D)**

Figure 3. "Trained" predicted versus actual intramuscular fat percentage for calibration method 'log': **(A)** run 1, repetition a; **(B)** run 1, repetition b; **(C)** run 2, repetition a; **(D)** run 1, repetition b. Line represents the outcome of a perfect fit between predicted and actual values.

(A)   **(B)**
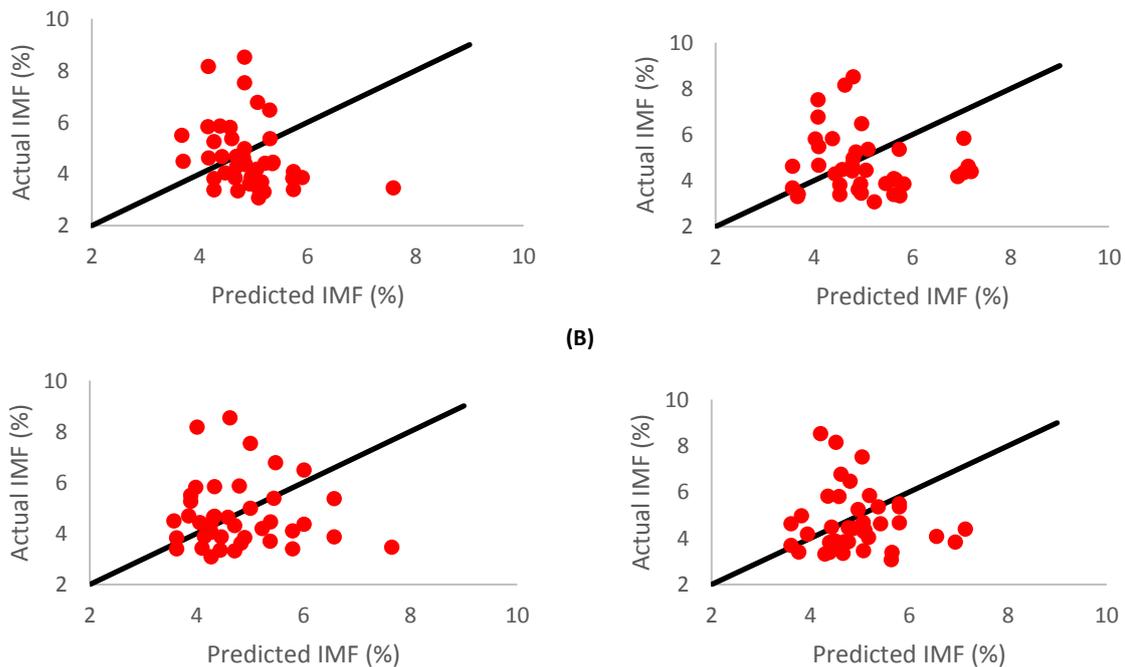
(C)   **(D)**

Figure 4. "Validated" predicted versus actual intramuscular fat percentage for calibration method 'log': **(A)** run 1, repetition a; **(B)** run 1, repetition b; **(C)** run 2, repetition a; **(D)** run 1, repetition b. Line represents the outcome of a perfect fit between predicted and actual values.

Similar to the shear force results, the precision of predicting the percentage of intramuscular fat within the training analysis was high across each of the 'none', 'minus', and 'log' calibration methods, and consistent across runs and sample repetitions. However, as was the case with shear force, the precision diminished upon validation, with the magnitude of the RMSEP, indicatively about 1.55% across all tests, representing about 28% of the full 5.45% unit range in IMF% data. Furthermore, the R^2 indicated very little predictive power of the validated models.

This loss of precision resulted in relatively meaningless slope and bias estimates, which varied somewhat randomly between tests.

6   Discussion

6.1   MEXA effectiveness in predicting SF5 and IMF%

6.1.1   Adequate prediction, though interpret results conservatively

This was an opportunistic study, undertaken while Spanish collaborators had access to the Multiscan Technologies MEXA scanner. One of the key goals of this study was to explore a range of different approaches that can be utilised for analysing and interpreting MEXA images. In this case, we have explored simple linear associations, partial least squares regression and, finally, extreme gradient boosting, with the latter suggesting some potential.

This study has also shown that MEXA may have capacity to predict both shear force and intramuscular fat percentage, as indicated by the performance of trained models. This likely indicates simple associations that are informing this prediction. However, this statement must be taken with cautious conservatism, as the models were imprecise and inaccurate upon validation. It is too early to conclude that this technology will not work, as there were significant limitations to the rigour of the validation testing. Most importantly, the dataset was small and thus validation testing was undertaken on only 8 animals per test—far from ideal for this type of modelling. Another factor to consider is that there was minimal opportunity to validate across either diverse animal genotypes or environmental conditions that may affect the machine reading. Although not genetically diverse, the lambs at least varied in age, which is likely responsible for some of the phenotypic divergence in SF5 and IMF% that the MEXA has been able to detect in the trained models. Nonetheless, the actual phenotypic range for both SF5 and IMF% was relatively small compared to that seen in resource flock experimental datasets where IMF% ranged between 1.5% and 9.8% (Pannier et al., 2014b) and SF5 ranged between 16.6N and 104.8N (Stewart, McGilchrist, Gardner & Pethick, 2018). This limited range would also limit our ability to construct a robust model predicting these traits. Lastly, our Spanish collaborators anecdotally noted that the amount of fluid exudate from each sample upon defrosting varied substantially. While their technicians were careful to handle the samples in a uniform fashion, it is likely that this variable purge has impacted the MEXA readings adding considerable random variation. Given the quantitative nature of this technology, we expect that this would certainly have restricted the performance of our validation analysis.

The physical principal behind the effectiveness of MEXA is similar to that of DEXA, in which the log of the ratio of attenuation for two different energy levels correlates with atomic mass. Our log transformation method was an attempt to represent that correlation. However, MEXA may be more effective than DEXA because each of the energy level contrasts will potentially enable classification of both the density and the Z-effective value of different substances across an atomic mass continuum. The Z-effective value, or the effective nuclear charge, is related to the shielding effect of an atom. This property is linked to the balance between the pull of the protons on valence electrons and the repulsion forces from inner electrons. A point to note is that it is likely that many of the 128 energy levels captured by the MEXA detector are superfluous. While each additional energy level detected may enhance our ability to differentiate individual atomic components that contribute to the eating quality within meat. However, we speculate that with the addition of several of these energy levels, most of the atomic components determining meat quality traits would be characterised.

The most surprising trait successfully predicted was shear force. This trait is an outcome of both the inherent muscle structure (with the entire muscle architecture consisting of connective tissues, fibre types and sarcomere lengths) and the way that muscle deteriorates post-mortem (i.e. aging through proteolysis). We expect that MEXA will only detect variation in atomic mass of substances in the meat

and, as such, it must be reflecting inherent muscle structures at the point of slaughter—not the subsequent proteolytic change through aging.

Lastly, the consistency of the trained predictions for the repeat scan of the same samples, and the within-sample repetitions, provide some level of confidence that the predictions were repeatable. Nonetheless, this needs to be demonstrated upon validation to be confident that these results are not simply one-off random associations.

6.1.2 Extreme Gradient Boosting versus Partial Least Squares Regression

Partial least squares regression is effective at discriminating signals within a spectra more optimised for trait identification; however, in this study, this analytical approach did not work. By contrast with the extreme gradient boosting methodology, PLS was not effective at either the training or the validation phase. As shown in the appended report, PLS models demonstrated poor precision even at the training phase of this analysis, as opposed to extreme gradient boosting, which at least demonstrated good precision within the trained models. This demonstrates the potential of extreme gradient boosting as an analytical method suitable for exploring MEXA prediction of commercially-relevant traits. Extreme gradient boosting is an ensemble learning algorithm, somewhat similar to "random forest", using a collection (or ensemble) of weaker individual models to arrive at a final predictive model. The extensive information available across the 128 energy levels, and the highly correlated nature of this data, appear well-suited to this machine learning methodology.

6.1.3 No real pressure on calibration

The calibration method had no obvious impact upon the results. In addition to the log transformation method, we also tested a crude background correction method (i.e. 'minus') and simply using the raw pixel values (i.e. 'none') to differentiate these calibration methods. On the basis of this limited dataset, there was actually little difference between these methods, although this may be due to the machine learning approach, which is likely to have captured the inherent associations "blindly", avoiding the need for structured calibration.

Alternatively, it is also possible that calibration has little bearing on this dataset. The samples scanned within this study were all imaged within the same 10 minute window. An electrical device like the MEXA is not likely to have been exposed to very much "machine-drift" across such a short period. Had these samples been scanned across several hours, days, or weeks, or across different MEXA devices, then it is likely that calibration would have had a more significant impact upon our results.

6.2 Industrial capacity for deploying MEXA

6.2.1 Clean and trimmed sample required

In the short-term, the commercial deployment of a MEXA imaging system is to couple it with a conveyor, allowing product to smoothly pass between x-ray tube and detector. However, one of the limitations of the present methodology is the need to scan trimmed samples of muscle devoid of fat or bone tissue. This is due to the 2-dimensional nature of the image captured, and the confounding of multiple tissue types within each pixel. Commercial meat cuts consist of lean surrounded by layers of subcutaneous and intermuscular fat, and in some cases bone. In a 2-dimensional image, this results in mixtures of tissue types within each pixel, in particular confounding intramuscular fat depots with subcutaneous and intermuscular fat. We speculate that a machine learning approach may find relevant associations when analysing entire images that capture differentials between pixels; however, we suspect that this will be difficult to validate across datasets. The alternative would be to progress to a computed tomography MEXA image, enabling identification of tissue regions within 3-dimensions, and the isolation of key regions of interest (i.e. specific muscle groups). These devices

are becoming more common-place, both in medical scanning and commercial baggage scanning industries, and should continue to be explored.

7 Conclusions and recommendations

This report highlights the potential for MEXA to predict key eating quality traits, however the limited capacity for validation testing indicated substantial loss of precision and accuracy for predicting 5-day aged shear force and intramuscular fat percentage. Given the small scale of the study, which provided only limited opportunity for validation testing, we cannot yet conclude that this technology won't work. Future experimental data should be gathered across a more balanced dataset that includes structures to validate across time and environmental effects. Importantly, this study has enabled us to explore a number of analytical and calibration methods that are likely suitable for future MEXA analysis. This will be crucial given MLA's investment in this technology.

In future work, it is recommended to calibrate MEXA across a larger dataset consisting of more genetically and phenotypically diverse animals, across varying environmental conditions, across time (hours/days/weeks), and across different MEXA devices.

8 Key messages

- Partial least squares regression models demonstrated poor precision even at the training phase of this analysis, as opposed to extreme gradient boosting, which demonstrated good precision within the trained models. This demonstrates the potential of extreme gradient boosting as an analytical method suitable for exploring MEXA prediction of commercially-relevant traits, such as Warner-Bratzler shear force and intramuscular fat percentage (IMF).
- Validation testing of MEXA's prediction of shear force and IMF, demonstrated a substantial loss in precision and accuracy. However, this needs to be treated with some conservatism as there were significant limitations to the rigour of the validation testing.
- The consistency of the trained predictions for the repeat scan of the same samples, and the within-sample repetitions, provide some level of confidence that the predictions were repeatable. Nonetheless, this needs to be demonstrated upon validation to be confident that these results are not simply one-off random associations.
- The most surprising trait successfully predicted was shear force. This trait is an outcome of both the inherent muscle structure (with the entire muscle architecture consisting of connective tissues, fibre types and sarcomere lengths) and the way that muscle deteriorates post-mortem (i.e. aging through proteolysis). We expect that MEXA will only detect variation in atomic mass of substances in the meat and, as such, it must be reflecting inherent muscle structures at the point of slaughter—not the subsequent proteolytic change through aging.

9   Bibliography

Gardner GE, Starling S, Charnley J, Hocking-Edwards J, Peterse J, Williams A (2018) Calibration of an on-line dual energy X-ray absorptiometer for estimating carcase composition in lamb at abattoir chain-speed. *Meat Science* **144**, 91-99.

Grunert KG, Bredahl L, Brunsø K (2004) Consumer perception of meat quality and implications for product development in the meat sector—a review. *Meat Science* **66**(2), 259-272.

Henchion M, McCarthy M, Resconi VC, Troy D (2014) Meat consumption: Trends and quality matters. *Meat Science* **98**(3), 561-568.

Hopkins DL, Toohey ES, Warner RD, Kerr MJ, van de Ven R (2010) Measuring the shear force of lamb meat cooked from frozen samples: comparison of two laboratories. *Animal Production Science* **50**, 382-385.

Pannier L, Gardner GE, Pethick DW (2018) Effect of Merino sheep age on consumer sensory scores, carcass and instrumental meat quality measurements. *Animal Production Science* **59**(7), 1349-1359.

Pannier L, Gardner GE, Pearce KL, McDonagh M, Ball AJ, Jacob RH, Pethick DW (2014a) Associations of sire estimated breeding values and objective meat quality measurements with sensory scores in Australian lamb. *Meat Science* **96**(2), 1076-1087.

Pannier L, Pethick DW, Geesink GH, Ball AJ, Jacob RH, Gardner GE (2014b) Intramuscular fat in the *longissimus* muscle is reduced in lambs from sires selected for leanness. *Meat Science* **96**(2), 1068-1075.

Perry D, Shorthose WR, Ferguson DM, Thompson JM (2001) Methods used in the CRC program for the determination of carcass yield and beef quality. *Australian Journal of Experimental Agriculture* **41**(7), 953-957.

Pethick DW, Warner RD, Banks RG (2006) Genetic improvement of lamb — industry issues and the need for integrated research. *Australian Journal of Agricultural Research* **57**, 591–592.

Pietrobelli A, Formica C, Wang Z, Heymsfield SB (1996) Dual-energy X-ray absorptiometry body composition model: review of physical concepts. *American Journal of Physiology* **271**(6), E941-E951.

Polkinghorne R, Thompson JM, Watson R, Gee A, Porter M (2008) Evolution of the Meat Standards Australia (MSA) beef grading system. *Australian Journal of Experimental Agriculture* **48**, 1351-1359.

Stewart SM, McGilchrist P, Gardner GE, Pethick DW (2018) Lamb loin tenderness is not associated with plasma indicators of pre-slaughter stress. *Meat Science* **137**, 147-152.

Young OA, Braggins TJ (1993) Tenderness of ovine *semimembranosus*: Is collagen concentration or solubility the critical factor? *Meat Science* **35**(2), 213-222.

10 Appendix

Non-destructive spectrometric technologies to determine sheep age at sacrifice according to meat tenderness

Fulladosa, E.[1], Gardner, G.E.[2], Williams, A.[2], Muñoz, I.[1], Font-i-Furnols, M.[1], Gou, P[1]

[1]*IRTA, XaRTA, Food Technology, Finca Camps i Armet, E-17121 Monells, Girona, Catalonia.*
[2]*Murdoch University, South Street, Murdoch, 6150, Australia.*

1. Introduction

The feasibility of vis-NIR technology to predict sensory and instrumental tenderness or aging times has been evaluated previously by several authors using high performance NIR instruments (Barbin, Valous & Sun, 2013; Hildrum, Nilsen, Mielnik & Næs, 1994; Kamruzzaman, ElMasry, Sun & Allen, 2013; Ripoll, Albertí, Panea, Olleta & Sañudo, 2008). Recently, low-cost NIR devices connected to smartphones specially designed for consumers have become available in the market. These devices show promising results for the qualitative classification of horticultural products according to quality attributes (Li, Qian, Shi, Medlicott & East, 2018), for the determination of egg storage time (Coronel-Reyes, Ramirez-Morales, Fernandez-Blanco, Rivero & Pazos, 2018) and for the evaluation of total antioxidant capacity in gluten free grains (Wiedemair & Huck, 2018). However, there is currently little work related to meat tenderness in the literature. On the other hand, Kroger et al (2006) suggested that dual energy X-ray was suitable for tenderness estimation, showing a coefficient of determination of 0.69 for a multiple non-linear regression. Soladoye et al (2018) reported that dual energy X-ray and NIR spectroscopy can be effectively used as a means of classifying pork bellies according to softness. New multi energy X-ray sensors have also shown the capacity to monitor meat aging during chilling-vacuum packaging of beef samples (Austrich et al 2017). However, compositional variations of meat (such as fat content) are likely to influence the spectra, hence the ability to evaluate tenderness or aging time in samples with variable composition needs to be investigated.

The aim of this study was to evaluate the feasibility of NIR and multi energy X-ray spectroscopy to determine animal age at slaughter from the characterization of the instrumental tenderness of meat. We assessed the performance of a multi energy X-ray sensor, a high performance NIR system and a portable low-cost NIR system to predict shear-force in sheep samples and to discriminate between age/tenderness groups and between tenderness groups.

2. Material and methods

2.1. Samples

Merino-cross lambs were sourced from a commercial sheep flock located in Bordertown (South Australia). Half were born in September/October 2016 (n=20), the other half in May/June 2017 (n=20), all were slaughtered in October 2017 at a commercial abattoir as lambs (6 months old) and yearlings (12 months old). All lambs were maintained in the same paddock 3 weeks prior to slaughter on pasture with supplemental hay or pellet feeding when necessary. After slaughter 2 samples 5cm long were taken from the *longissimus dorsi* near to the 12/13th rib, aged for 5 days and then frozen. One sample was used to determine shear force; the other was used for scanning using 3 different spectrometric devices.

2.2. Spectrometric acquisition systems

Spectra were acquired using two NIR-based systems (a high performance and a low-cost device) and a multi energy X-ray spectroscopic sensor as described below.

a. High performance NIR device

Spectra were recorded on a Fourier Transform NIR spectrometer model Matrix-F duplex (Bruker Optik GmbH, Germany). NIR spectra were collected using the OPUS™ software (Bruker Optik GmbH, Germany). NIR reflectance spectra were collected over the 12,000–4000 cm−1 spectral region (corresponding to a wavelength interval of 830–2500 nm) using an on-contact probe IN 268–2 (Solvias

AG, Switzerland). This probe uses a bifurcated (Y-shaped) fibre optic bundle to illuminate the sample (four fibres input) and to collect the diffuse reflection (four fibres output). The measured spot size is 3 mm ø. Using this probe, spectra were acquired on-contact with the sample. Each spectrum was obtained from 16 scans performed at 8 cm^{-1} resolution. All spectra were recorded in reflectance mode with respect to a reference standard made of bakelite.

Data acquisition was performed on the region of interest (*Longissimus dorsi* muscle) (Figure 1), in which instrumental tenderness determination was later performed. For each device, five spectra were acquired in this muscle at 15 ℃. Mean spectra was calculated and used for development and validation of the predictive models.

b.   Low-cost NIR device
The SCiO handheld smartphone-based NIR spectrometer (Consumer Physics, Israel) was used to collect spectra. A reflection spectra in a range between 740 nm and 1070 nm with a 1 nm resolution were collected. Samples were scanned using SCiO$^{(TM)}$ shade accessory that helps to avoid the influence of external light and to keep the same 10 mm distance in all the collected spectra. The measured spot size was 20 mm^2. Spectra were stored in a cloud-based data set and using a research license of Scio Lab, spectral signals were downloaded and imported into Matlab in order to develop and optimize the chemometric models. Data acquisition was performed following the procedure described in section a.

c.   Multi energy X-ray device
An industrial X-ray prototype system with a multi energy detector (MXV-PACK 4010, Multiscan Technologies, Cocentaina, Spain) was used to scan the samples. The prototype had an X-ray spectrometric detector made of a semiconductor crystal (CDTe/CZT) with a pixel size of 0.8 mm. A belt conveyor transported the sample at a speed of 10 m/min.

Simultaneously, X-rays were emitted from below the samples using a tungsten X-ray tube that was pulsed at 140kV. The resulting polychromatic array of photons was detected and transmitted radiation recorded across 128 channels, corresponding to energies from 20 to 160 keV. Thus the system acquired a spectroscopic image of the sample with each pixel containing an X-ray energy spectra of 128 channels. Within each spectroscopic image a region of interest (*Longissimus dorsi* muscle) was identified that corresponded to a uniform section of the sample, while also maximising the sample area. So the size of the acquired information was a 3D matrix of variable pixel numbers x 128 channels which was analysed using a Matlab script written in house (MATLAB, Ver. 7.7.0, The Mathworks Inc., Natick, MA, USA). The mean X-ray attenuation (S_a) for the energy channel *a* of the selected ROI was calculated after background correction and log-transformation as described in equation 1.

$$S_a = \frac{- \sum_{i=1}^{p} \ln\left(\frac{I_f^{a,i}}{I_o^{a,i}}\right)}{p} \qquad \text{(Eq. 1)}$$

where I_f is the intensity of the transmitted radiation and I_o is the intensity of the incident radiation at each pixel *i* of the ROI which contains *p* pixels. The calculation was done for each energy channel of the spectra *a,* that ranges from 1 to 128 and represents a given X-ray energy. According to Eq. 1, an increase of S_a value represents an increase of the X-ray attenuation for channel *a*.

2.3 Tenderness determination
Approximately 65 g loin samples were used for shear force testing as described by Pannier et al. 2014. Initially samples were vacuum packed and aged for 5 days, and then frozen at −20 °C until subsequent testing. Frozen samples were cooked in plastic bags in a water bath for 35 min at 71 °C, and were cooled in running water for 30 min after cooking. Six cores (~3–4 cm long, 1 cm2 cross sectional) from each loin sample were cut and shear force was measured using a Lloyd texture analyser (Model LRX, Lloyd Instruments, Hampshire, UK) with a Warner–Bratzler shear blade fitted as described by Hopkins, Toohey, Warner, Kerr, and van de Ven (2010).

2.4 Construction and cross-validation of predictive models
Spectra were subjected to various pre-treatments including mean center, vector normalisation, constant offset elimination, straight line subtraction, min-max normalisation, multiplicative scatter correction and first derivative. All spectra were processed and multivariate calibration models constructed with

the aid of the software in Matlab® environment, version 7.1 (The MathWorks, Inc) and applying algorithms from the PLS Toolbox, version 8.1.1 (Eigenvector Research, Inc., Wenatchee, WA, USA).

Calibration models were developed by using Partial Least Squares regression (PLSR) for shear force in lamb meat. Validations were made by a cross validation (Venetian blinds) model. The best combination of pre-treatments was selected based on the lowest error of prediction and the lowest number of PLS factors. Partial Least Squares discriminant analysis (PLS-DA) was used as a linear classification method to predict the class (new or old). When needed, interval variable selection was applied to discard variables which may be adding complexity to a model and improve performance of a final model.

The goodness of fit of the models was assessed using the coefficient of determination (R^2) obtained in the cross-validation set of samples, and the root mean square error of cross validation (RMSECV). Furthermore, in order to evaluate the predictive ability of the calibration models, the residual predictive deviation (RPD) statistic was used. This qualitative measure for the assessment of the validation results is the relationship between the standard deviation (SD) of the reference values and RMSECV (Williams & Norris, 1987). Conzen (2006) considered a model good for screening with a $RPD > 3$, good for quality control with a $RPD > 5$, and excellent for all analytical tasks with a $RPD > 8$.

3. Results and discussion

3.1 Samples characterization

Tenderness of samples ranged from 24.68 to 56.56 N. Table 1 shows significant differences in shear force between young and old animals ($p=0.01$). Mean shear force was 38.45 ± 8.99 N for young animals and 31.65 ± 7.25 N for old animals. This is an unusual trend, as older animals were expected to have greater shear force, as has been shown by previous authors (Hopkins et al 2007). None-the-less these samples still provided substantial variation for instrument prediction testing. Water content of samples was not significantly different.

Table 1. Lamb meat samples characterization in terms of tenderness for young and old animals.

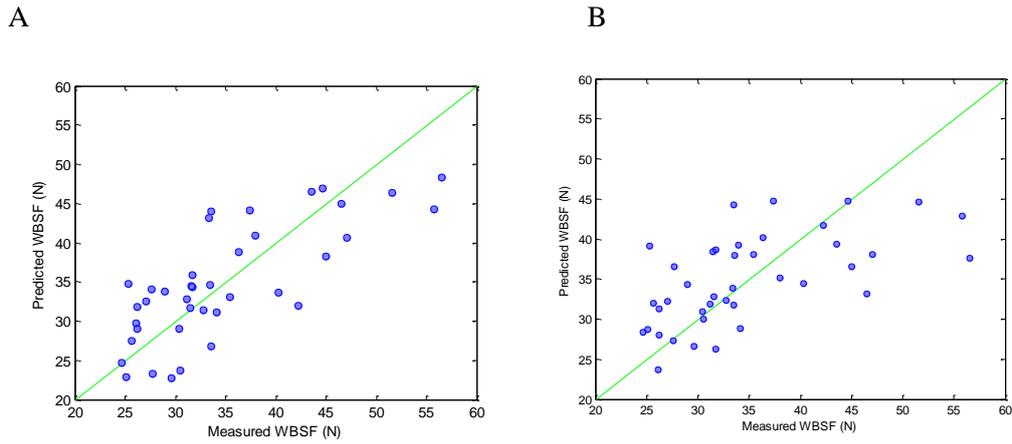|		Shear force (N)			
Animal age	**n**	**Mean**	**Desvest**	**min**	**max**
Young	21	38.45[a]	8.99	25.32	56.56
Old	21	31.65[b]	7.25	24.68	55.79

[ab] Different letters mean significant differences ($p<0.05$)

3.2 Predictive ability of lamb meat tenderness

PLSR models and internal cross-validation were developed using the spectra acquired with the different spectrometric systems. R^2 values were low for both calibration and cross-validation data sets for the three tested spectrometric systems. Mean centering and first derivative were found to improve predictive results. However, for NIR this was only after selection of the most representative spectral intervals for estimating shear force (Table 2). In the case of high performance NIR (that acquires from 830 to 2500 nm) the most informative wavelength were from 956 to 960 nm, from 1053 to 1058 nm and from 1640 to 1660 nm. The models used these specific regions of the spectra and 8 PLS factors to achieve a RMSECV of 5.54. In the case of low-cost NIR device (that acquires from 740 to 1070), the most representative wavelengths were found to be from 780 to 790 nm and from 1040 to 1050 nm. The models used these specific regions of the spectra and 3 PLS factors to achieve a RMSECV of 0.69. However, R^2_{cv} were still low in both cases showing values of 0.578 and 0.359 (see Figure 1). Multi energy X-ray showed no capacity to predict lamb meat tenderness.

Table 2. Performance metrics in iPLSR models for WBSF using the three investigated systems.

	Multi energy X-ray	**NIR-high performance**	**NIR-low cost**
Used spectral interval	20-40 keV	956-960 nm, 1053-1058 nm, 1640-1660 nm	780-790 nm, 1040-050nm
PLS factors	3	8	3
$RMSE_C$	7.66	1.93	6.19
R^2_{cal}	0.15	0.947	0.448
$RMSE_{CV}$	9.02	5.54	6.69
R^2_{CV}	0.005	0.578	0.359

A

B

Figure 1. Relationship between the measured and predicted shear force using PLSR models obtained from high performance NIR spectra (a) and low-cost NIR spectra (b).

3.3 Classification ability according to animal's age

Figure 2 shows the average spectra for the different age groups for the multi energy X-ray device (a), high performance NIR device (b) and low-cost NIR device (c), before and after pre-processing was applied.
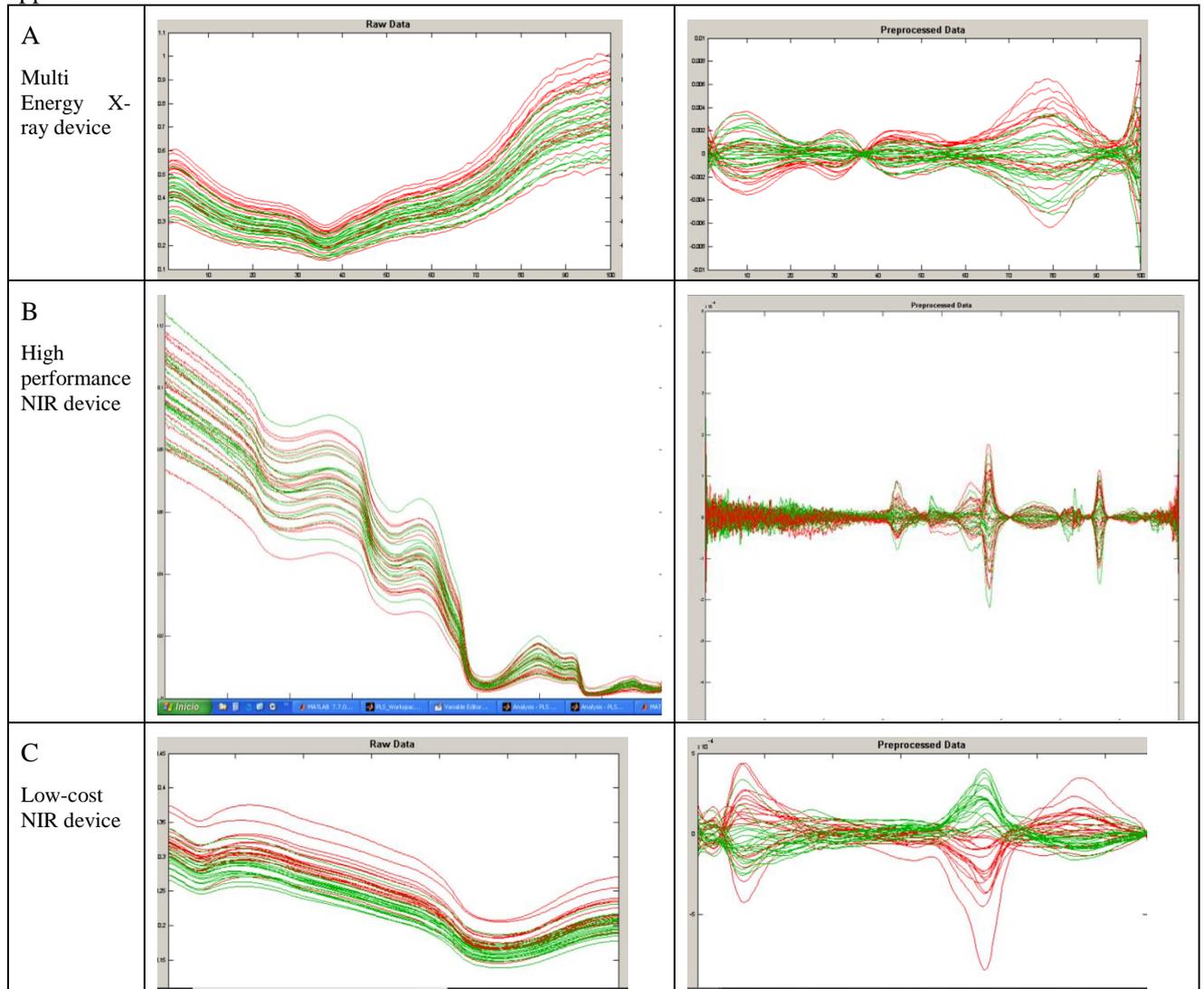
Figure 2. Average spectra for the two age groups and the same spectra after pre-processing (mean center and first derivative). Interval selection for iPLSDA is shown.

The results for the PLSDA and iPLSDA models for classification of samples according to animal age are shown in Table 3. The lowest error rate (%) was found for the high performance NIR where accuracy of calibration and prediction reached 100% for all the samples when using the most representative interval selection. The highest error was found when using multi energy X-ray (27.5%), suggesting that currently this technology is non-acceptable for implementation within the food industry. This might partly be due to the variations in thickness found between samples which may cause variation in the attenuation values, producing higher errors. However, low cost NIR devices are shown to be able to discriminate accurately (with 2.5% of error rate in cross validation when using specific spectral intervals).

Table 3. Confusion matrix for discrimination of young and old animals by PLSDA models obtained from multi energy X-ray, high performance and low cost NIR systems.

	multi energy X-ray		**NIR-high performance**		**NIR-low cost**	
PLSDA						
PLS factors	2		4		5	
Used spectra region	20-130 keV		830-2500 nm		740-1070 nm	
Calibration	*Actual Young*	*Actual Old*	*Actual Young*	*Actual Old*	*Actual Young*	*Actual Old*
Predicted Young	12	4	19	1	18	2
Predicted Old	7	17	0	19	1	20
Error rate (%)	27.5%		2.5%		7.6%	
Cross-validation						
Predicted Young	12	4	13	2	18	1
Predicted Old	7	17	6	18	2	18
Error rate (%)	27.5%		20.5%		7.6%	
iPLSDA						
PLS factors	2		3		5	
Used spectra region	20-30 keV		1358-1364, 1465-1480, 1935-1948 nm		900-940 nm	
Calibration	*Actual Young*	*Actual Old*	*Actual Young*	*Actual Old*	*Actual Young*	*Actual Old*
Predicted Young	13	2	19	0	19	0
Predicted Old	6	19	0	20	0	20
Error rate (%)	20%		0%		0%	
Cross-validation						
Predicted Young	11	3	19	0	19	1
Predicted Old	8	18	0	20	0	19
Error rate (%)	27.5%		0%		2.5%	

Conclusions

This study has demonstrated the ability of spectroscopic methods, especially high performance and low-cost NIR, to provide a fast and effective method for classifying meat from old or young sheep animals at slaughter. The potential of studied spectroscopic devices in the prediction of sheep age needs to be further investigated on a larger number of samples.

Acknowledgements

References

Barbin DF, Valous NA, Sun DW (2013) Tenderness prediction in porcine longissimus dorsi muscles using instrumental measurements along with NIR hyperspectral and computer vision imagery. *Innovative Food Science & Emerging Technologies* **20**, 335-342.

Coronel-Reyes J, Ramirez-Morales I, Fernandez-Blanco E, Rivero D, Pazos A (2018) Determination of egg storage time at room temperature using a low-cost NIR spectrometer and machine learning techniques. *Computers and Electronics in Agriculture* **145**, 1-10.

Fulladosa E, Austrich A, Muñoz I, Guerrero L, Benedito J, Lorenzo JM, Gou P (2018) Texture characterization of dry-cured ham using multi energy X-ray analysis. *Food Control* **89**, 46-53.

Hildrum KI, Nilsen BN, Mielnik M, Næs T (1994) Prediction of sensory characteristics of beef by near-infrared spectroscopy. *Meat Science* **38**(1), 67-80.

Hopkins DL, Toohey ES, Warner RD, Kerr MJ, van de Ven R (2010) Measuring the shear force of lamb meat cooked from frozen samples: comparison of two laboratories. *Animal Production Science* **50**, 382-385.

Hopkins DL, Stanley DF, Martin LC, Toohey ES, Gilmour AR (2007) *Australian Journal of Experimental Agriculture* **47**, 1155-1164.

Kamruzzaman M, ElMasry G, Sun DW, Allen P (2013) Non-destructive assessment of instrumental and sensory tenderness of lamb meat using NIR hyperspectral imaging. *Food Chemistry* **141**(1), 389-396.

Kröger C, Bartle CM, West JG, Purchas RW, Devine CE (2006) Meat tenderness evaluation using dual energy X-ray absorptiometry (DEXA). *Computers and Electronics in Agriculture* **54**(2), 93-100.

Li M, Qian Z, Shi B, Medlicott J, East A (2018) Evaluating the performance of a consumer scale SCiO™ molecular sensor to predict quality of horticultural products. *Postharvest Biology and Technology* **145**, 183-192.

Liu Y, Lyon BG, Windham WR, Realini CE, Pringle TDD, 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**, 1107-1115.

Moran L, Andres S, Allen P, Moloney AP (2018) Visible and near infrared spectroscopy as an authentication tool: Preliminary investigation of the prediction of the ageing time of beef steaks. *Meat Science* **142**, 52-58.

Pannier L, Gardner GE, Pearce KL, McDonagh M, Ball AJ, Jacob RH, Pethick DW (2014) Associations of sire estimated breeding values and objective meat quality measurements with sensory scores in Australian lamb. *Meat Science* **96**(2), 1076-1087.

Prieto N, Navajas EA, Richardson RI, Ross DW, Hyslop JJ, Simm G, Roehe R (2010) Predicting beef cuts composition, fatty acids and meat quality characteristics by spiral computed tomography. *Meat Science* **86**(3), 770-779.

Prieto N, Ross DW, Navajas EA, Nute GR, Richardson R., Hyslop JJ, 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**, 96-103.

Ripoll G, Albertí P, Panea B, Olleta JL, Sañudo C (2008) Near-infrared reflectance spectroscopy for predicting chemical, instrumental and sensory quality of beef. *Meat Science* **80**(3), 697-702.

Soladoye OP, Prieto N, Lopez-Campos O, Aalhus JL, Uttaro B, Roberts JC, Larsen I, Shand P, Gariépy C, Juárez M (2018) Potential of near infrared (NIR) spectroscopy and dual energy X-ray absorptiometry (DXA) in predicting pork belly softness. *Meat Science* **142**, 1-4.

van Oeckel MJ, Warnants N, Boucqué CV (1999) Comparison of different methods for measuring water holding capacity and juiciness of pork versus on-line screening methods. *Meat Science* **51**, 313-320.

Wiedemair V, Huck CW (2018) Evaluation of the performance of three hand-held near-infrared spectrometer through investigation of total antioxidant capacity in gluten-free grains. *Talanta* **189**, 233-240.