THE UNIVERSITY OF SYDNEY

Use of spectroscopic sensors in meat and livestock industries

Cassius E. O. Coombs

BSc (An. Vet. Bio. Sc.), MPhil

Sydney Institute of Agriculture

School of Life and Environmental Sciences

Faculty of Science

The University of Sydney

A thesis submitted to fulfil requirements of the degree of Doctor of Philosophy

February 2022

Statement of Originality

This is to certify that to the best of my knowledge; the content of this thesis is my own work. This thesis has not been submitted for any degree or any other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

I certify that this thesis, upon successful completion, will be made available for immediate use in the University of Sydney library.

The data that supports Chapter 2 cannot be publicly shared due to ethical or privacy reasons and may be shared upon reasonable request to myself if appropriate. My co-authors (Mario Fajardo, Robert Liddle, Luciano González) and I declare no conflicts of interest in Chapters 1, 2, 3, and 6 of the present thesis.

For Chapters 4 and 5, my co-authors Brendan Allman and Edward Morton work for Rapiscan Systems Pty. Ltd. who manufacture and commercialise similar equipment to the multi-sensory platform (AK198) used in these chapters. My other co-authors in these chapters (Marina Gimeno, Neil Horadagoda, Garth Tarr, Luciano González) declare no conflicts of interest.

“Life’s unfair, then you die” ~ Ted Bayas

“We are all just batteries for consumer culture” ~ Owen Cook

“No matter how far down the wrong road you’ve gone, TURN BACK!!!” ~ Eben Pagan

“Begin with the end in mind” ~ Stephen Covey

“You don’t get paid to be told what to do, you get paid to solve problems” ~ Jeff Downing

“If you’re feeling tired, don’t get tired” ~ Eric Jewson

Authorship attribution statement

Chapter 2 of this thesis is published and cited in the thesis as Coombs et al. (2021a). I co-designed the study with L. A. González, interpreted the analysis done by M. Fajardo and wrote the drafts of the manuscript.

Chapter 3 of this thesis is published and cited in the thesis as Coombs et al. (2021b). I co-designed this study with the co-authors, extracted the data with R. R. Liddle, analysed the data with L. A. González and wrote the drafts of the manuscript.

This thesis contains material published in Lillesand et al. (2021). This is in Chapter 1, Fig. 1 and is copied directly from the book.

This thesis contains material sourced from Rapiscan Systems Pty. Ltd. with permission granted by B. E. Allman and E. J. Morton of Rapiscan Systems Pty. Ltd., co-authors in Chapters 4 and 5. The image sourced from Rapiscan Systems is Fig. 8 in the present thesis (Chapter 4). The MATLAB algorithms in Chapters 4 and 5 were written by B. E. Allman, and the R scripts for statistical analysis were written by G. Tarr.

This thesis contains material published and in cited as Coombs et al. (2021c). This is in Chapter 6, Table 1. This study was published as a short conference paper; in this I designed the study and analysed the data together with L. A. González, extracted the spectral data from marked regions of interest using an algorithm developed by B. E. Allman, and wrote the drafts of the manuscript.

Abstract

Spectroscopic sensors such as near-infrared spectroscopy (NIRS), Raman spectroscopy, and hyperspectral (HS) imaging were trialled in the present thesis. The four experimental chapters were conducted with the purpose of investigating their potential for applications in the meat industries. The first study compared a handheld NIRS sensor connected to a smartphone (NIRvascan) against a benchtop NIRS sensor to measure chemical composition (pH, moisture, fat, protein) of beef and lamb retail cuts in three different sample presentations (fresh intact, freeze-dried, and freeze-dried followed by oven-dried). The smartphone NIRS sensor showed moderate precision in predicting fat concentration (r^2 = 0.78–0.81, residual prediction deviation - RPD = 2.1–2.3), and it was comparable to the benchtop NIRS sensor on processed meat. However, predictions from both NIRS sensors on fresh intact meat were insufficient for even rough screening in the industry (r^2 < 0.67, RPD < 2.0). The second study compared the NIRvascan sensor against a Raman spectrometer to differentiate grass-fed from grain-fed retail beef cuts. The NIRvascan was more accurate than the Raman when scanning lean tissue, whereas the Raman was more accurate on fat tissue (both >90% accuracy). The portable nature, low cost, smartphone connectivity, and similar accuracy of the NIRvascan made it a more viable option for industry and consumer use compared to larger and more expensive instruments. The third study used a multi-sensory platform containing two HS sensors (visible – VIS: 400–900 nm, and short-wave infrared – SWIR: 900–1700 nm) to classify beef and sheep organs by organ type (heart, kidney, liver, or lung). Hearts and livers were the most accurately identified organs (accuracy >90%), whereas kidneys and lungs were less accurate (50–70%). The fourth study used the same multi-sensory platform to identify sheep organs rejected or fit for human consumption. Hearts and livers were more accurately (78–96%) discriminated as diseased or healthy compared to kidneys and lungs (60–91%). In the latter two chapters, the SWIR and VIS HS sensors achieved similar accuracy although VIS was slightly more accurate in differentiating organ type and SWIR was slightly more accurate in differentiating organs rejected or not for human consumption. In conclusion, three of the four chapters showed the potential of these novel spectroscopic sensors for

applications in the meat industries and further studies with larger sample sizes are encouraged. Furthermore, the multi-sensory platform showed excellent potential for the meat industry because it is non-contact and can run at chain speed, therefore studies looking into chemical composition and discrimination of meat based on feeding regime using HS imaging in this platform should be conducted in the future.

Acknowledgements

This thesis could not be possible without the help of many individuals, and several teams that helped me along the way. Firstly, at home to my wonderful and supportive partner Steffi for everything including all the emotional support, keeping me well fed and encouraging a life away from the PhD. Lockdown gave me a lot of time to spend closer to you and our wonderful cat Tiger. To my mother, siblings, and friends (special mentions to Ben, Douglas, Luca, Betül, Mehdi and Irum) who inspired me along the way to follow in their footsteps and pursue a PhD, the road of the PhD is now over. For Dad and Grandpa, there is now a new Dr Coombs – I wish you could live to see this, even if it is a PhD.

Secondly to Luciano González, the best supervisor anyone could ask for, for all the late-night emails and run-arounds regarding scholarships, writing code for data analysis and for all the help with the admin, writing and keeping me on the straight and narrow path without going off on tangents. To the rest of the CCWF livestock team for providing great company and help at work and the experience of being there before me and always being able to teach me more - Augusto, Juan, Milad, Janine, Christie, Sam, Greg, Tony, Emilie, Claudia, and Holly.

To everyone from CCWF including Claudia K and Sonam for their assistance in freeze-drying and crude protein analysis, Kate and Lynne for keeping the place running (and Kate's invaluable JP services), and the IT assistance of Andrew and Rob.

To Rapiscan Systems, particularly Ed Morton and Brendan Allman, who were instrumental in ensuring this PhD happened and that we had a multi-sensory system and that I was given a stipend scholarship via Meat and Livestock Australia Donor Company (P.PSH.0886 – aviation security

equipment in the meat industry and V.DSP.2018 – offal sortation)! Additional thanks must go to the rest of the Rapiscan team who assisted me along the way - Russell, Greig, Loic, Sina, Xiaoshui, Kevin, Vince, Tim (at very late hours), ChiYuan, GimHui and Omega. Additional thanks to Mark Ovens and APR Intern who ensured I would intern at Rapiscan Systems in my first year, paving the way for success in this PhD.

To the meat suppliers - 1888 Certified (Chapter 3) and their collaborating farms for the provision of traced grass-fed beef, and to The Meat Man (Narellan), Dr Michelle Henry and the team at Gundagai Meat Processors for the provision of beef and sheep organs (Chapters 4 and 5).

To Meat and Livestock Australia for providing the top-up scholarship (B.STU.2001) and technical assistance grant for myself and Prof. González (B.STU.1805).

To the Australian Research Council (DP210100521) for your support to Dr Tarr for supervising me in undergoing this thesis.

Individual thanks to:

Garth Tarr – for your assistance in supervision and statistics, writing the script for the final two chapters, introducing me to R Markdown and more R tricks and paving the way for the papers in the future.

Marina Gimeno and Neil Horadagoda – for your help with the post-mortems and writing during the final two experimental chapters.

Mario Fajardo – for your assistance in writing the script for the first paper when I had no idea about spectroscopy or data analysis, and for helping with the writing of the first paper, even if it took close to two years to publish!

Joon Lee from Sydney Analytical for providing and training Robert and I on the Raman spectrometer

Budi, Mario (once again!), Ed Jones, Kim, Kasia, Svetlana and Ryan from ATP for all the help with the first experimental chapter and getting the first practical off the ground.

To the Honours students who helped shape me as a mentor – Luke, David, Schyla, Hans and particularly Robbie – your enthusiasm and contributions to this thesis and my motivation to keep writing cannot be understated. Many thanks for helping me grow as a researcher and a mentor.

Additional thanks to the undergraduate students I helped teach with Luciano out at Pye Farm, and Marina and Neil's veterinary pathology students.

List of Abbreviations

A	Absorbance spectra
Ad1	First derivative of absorbance spectra
Ad2	Second derivative of absorbance spectra
Afinal	Final absorbance spectral dataset and *ncomp* selected
AHDB	Agriculture and Horticulture Development Board
AOAC	Association of Official Analytical Chemists
cal	Calibration
CART	Classification and regression trees
CCD	Charge coupled device
CIE	Commission International l'Eclairage
CLA	Caseous lymphadenitis
COMB	Combination of VIS and SWIR hyperspectral sensors
CP	Crude protein
CSV	Comma separated values
CT	Computed tomography
CV	Cross-validation
DM	Dry matter
DOF	Days on feed
EPO	External parameter orthogonalization
FD	Freeze-dried
FDOD	Freeze-dried and oven-dried
fps	Frames per second
FI	Fresh intact
FIR	Far-infrared
GIMP	Gnu Image Manipulation Program

HS	Hyperspectral
IMF	Intramuscular fat
LED	Light-emitting diode
LDA	Linear discriminant analysis
LG-NIRS	Laboratory-grade near-infrared reflectance spectroscopy
LOOCV	Leave-one-out cross-validation
MEXA	Multi-energy X-ray
MIR	Mid-infrared
MLA	Meat and Livestock Australia
MSA	Meat Standards Australia
MSC	Multiplicative scatter correction
mtry	Number of variables available for splitting at each tree node
ncomp	Number of components available for partial least squares analyses
NDL	No detectable lesions
NFAS	National Feedlot Accreditation Scheme
NIR	Near-infrared reflectance
NIRS	Near-infrared reflectance spectroscopy
PC1	First principal component
PC2	Second principal component
PCA	Principal components analysis
PCAS	Pasturefed Cattle Assurance System
PLS-DA	Partial least squares discriminant analysis
PLSR	Partial least squares regression
QIR	Quartz infrared
R	Reflectance spectra
R^2 or r^2	Coefficient of determination

Rd1	First derivative of reflectance spectra
Rd2	Second derivative of reflectance spectra
RF	Random forest
Rfinal	Final reflectance spectral dataset and *ncomp* selected
RGB	Red-green-blue
RMSE	Root mean square error
$RMSE_{CV}$	Root mean square error of repeated cross-validation
ROI	Region of interest
RPD	Residual prediction deviation
SE	Standard error
SE_{CV}	Standard error of cross-validation
SEM	Standard error of the mean
SNV	Standard normal variate
SWIR	Short-wave infrared
USDA	United States Department of Agriculture
UV	Ultraviolet
val	Validation dataset
VIS	Visible

List of Publications Generated by the Thesis

Papers published in peer-reviewed journals

Coombs, C. E. O., Fajardo, M. & González, L. A. (2021). Comparison of smartphone and lab-grade NIR spectrometers to measure chemical composition of lamb and beef. *Animal Production Science* 61(16): 1723-1733. doi:10.1071/AN21069.

Coombs, C. E. O., Liddle, R. R. & González, L. A. (2021). Portable vibrational spectroscopic methods can discriminate between grass-fed and grain-fed beef. *Journal of Near Infrared Spectroscopy* 29(6): 321-329. doi:10.1177/09670335211049506.

Papers published in peer-reviewed conference proceedings

Coombs, C. E. O., Neely, L., Minasny, B., Fajardo, M. & González, L. A. (2019). Using a handheld near-infrared spectroscopy (NIRS) scanner to predict meat quality. In E. Puolanne (Ed.). *Proceedings of the 65th International Congress of Meat Science and Technology* (August 4-9, 2019). Meat for Diversifying Markets: Postersessions and Discussion, Potsdam (pp. 502-503). Event Lab, GmbH and DigICoMST.

Liddle, R. R., Coombs, C. E. O. & González, L. A. (2020). Two handheld spectral devices can differentiate grass- and grain-fed beef. In American Meat Science Association (Eds.). *Abstracts from the 2020 International Congress of Meat Science and Technology and the AMSA Reciprocal Meat Conference* (August 2-7, 2020). *Meat and Muscle Biology* 5(2): 100-101. Orlando, Florida. doi:10.22175/mmb.11683.

Coombs, C. E. O., Manan, H., Minasny, B. & González, L. A. (2021). Can X-ray fluorescence predict mineral concentrations of ground meat? In *Proceedings of the 33rd Australian Association of Animal Sciences Conference* (February 1-3, 2021). Fremantle (p. 142).

Liddle, R. R., Coombs, C. E. O. & González, L. A. (2021). Can a smartphone near-infrared spectroscopy sensor predict days on feed and marbling score? In *Proceedings of the 33rd Australian Association of Animal Sciences Conference* (February 1-3, 2021). Fremantle (p. 337).

Coombs, C. E. O., Allman, B. E., Morton, E. J. & González, L. A. (2021). Discrimination of grass-fed and grain-fed frozen beef using hyperspectral imaging. In *Proceedings of the 67th International Congress of Meat Science and Technology* (August 23-27, 2021), Kraków (p. 249). Symposium Cracoviense.

Table of Contents

List of Figures and Tables

1. Use of spectroscopic sensors in meat and livestock industries – literature review

1.1. Introduction

The demand for the delivery of a high-quality, safe meat product to the population has become crucial, particularly in recent years where quality has taken precedence over quantity for consumers (Polkinghorne & Thompson, 2010; Strong, 2004). Recent research has found that regardless of country of origin, consumers are willing to pay a higher price for higher quality meat (Pethick et al., 2018; Polkinghorne & Thompson, 2010; Thompson et al., 2010). Different measures and thresholds of meat quality exist. However, basic principles remain the same such as a bloomed cherry-red muscle colour at retail level (Mancini & Hunt, 2005; Suman et al., 2014), and tenderness of the cooked product (Miller et al., 2001; Savell & Cross, 1988). Marbling, or visible intramuscular fat (IMF), is a defining characteristic of meat quality as its content is linked to increased tenderness, juiciness, flavour, and overall liking (Frank et al., 2016; Miller, 2020; Pannier et al., 2014). These traits are prioritised by consumers even in developing countries including South Africa (Bonny et al., 2018; Thompson et al., 2010), China (O'Reilly et al., 2017) and Brazil (Magalhães et al., 2018).

At abattoir level, quality of beef and lamb is mostly measured subjectively, with the exception of pH, where meat is graded against scales for colour and marbling scores. Australian grading systems such as AUS-MEAT (AUS-MEAT, 2005; AUS-MEAT, 2018) and the consumer score-based Meat Standards Australia (MSA) (MLA, 2012; MLA, 2017a; Polkinghorne et al., 2008) are considered as the most comprehensive in the world, although they still rely heavily on subjective grading (Pannier et al., 2018; Polkinghorne & Thompson, 2010). Laboratory chemical measures for crude fat and protein exist, though their uses are expensive, time consuming and external to meat processing. To combat these issues, larger devices such as computed tomography (CT), X-ray attenuation (single-, dual-, or multi-energy) and ultrasound have been used at meat processing

plants to predict carcase yield to varying degrees of success (Scholz et al., 2015). However, the sheer expense and space requirements of such devices, as well as the training of staff to use them, has seen their uptake slow. Furthermore, meat quality has been the pre-eminent measure desired in processing plants around the world for several decades, with particular focus on IMF (Hwang et al., 2004; Pannier et al., 2018; Thompson, 2004). For this, smaller devices such as near-infrared reflectance (NIR) spectroscopy (NIRS) sensors have shown immense promise (Prieto et al., 2009; Prieto et al., 2017), with smaller handheld NIR devices a likely solution to this problem (Dixit et al., 2017). Near-infrared reflectance spectroscopy sensors and other devices that will be explored in this review include hyperspectral imaging (HS; Elmasry et al., 2012a) and Raman spectroscopy (Xu et al., 2020), which have been chosen due to being less expensive and easier to use.

Sensors, including various spectroscopy sensors, have also been used in animal production to remotely monitor animal performance, particularly weight, feed intake, growth rate, and disease status (Neethirajan et al., 2016). Animal health has been mentioned as a particularly crucial area of research for remote monitoring, particularly because this is labour-intensive for veterinarians and personnel both on-farm and in the abattoir, and potential zoonotic diseases raise issues for human health (Butler et al., 2003; Webber et al., 2012). However, only few studies have been successful in employing sensors in animal health monitoring, particularly in the abattoir (Thomas-Bachli et al., 2014). Thus, automatic detection of diseases in abattoirs has been an emerging area of interest via reports from AgResearch (2018) and Meat and Livestock Australia (Cook & Anderson, 2017), although very few peer-reviewed journal articles have been published beyond small pilot experiments.

Several studies have explored the potential of spectral sensors to the meat industries which will be detailed in the forthcoming sections, with the ensuing experimental chapters focusing on specific areas of the meat processing and supply chain. Paucities have been revealed in current literature, namely evaluation of meat quality, authentication of feeding regime, and identification of

offal type and defects including those arising from animal health issues. Consequently, spectroscopic sensor devices have been developed to allow for automation in these sectors of the meat supply chain and can be deployed based on successful experimental results and future feasibility studies. The use of optimal mathematical modelling for the data from each device and its predictive outcome is also discussed briefly throughout the sections of the review and discussed at the end of the review. The objective of the present review is to provide a solid background to the use of spectroscopic sensors in meat science, identify research gaps which can then be addressed in the experimental chapters of the present thesis, and propose further research to improve meat science evaluation with such sensors.

1.2. Spectroscopy in meat science

Spectroscopy has been widely investigated in meat science to measure eating quality (Prieto et al., 2009), chemical composition (Prieto et al., 2017), and many other attributes of meat (Dixit et al., 2017; Dixit et al., 2021). Spectroscopy is the measurement of the interaction between radiation intensity and matter as a function of wavelength (Williams et al., 2019). Different samples of matter have different spectral signatures generated during their interaction with electromagnetic radiation, with the signature across the electromagnetic spectrum allowing the physical and electronic structure to be determined at atomic, molecular, and macro scales (Wadoux et al., 2021). The electromagnetic spectrum defines spectroscopy as microwave, terahertz, infrared, near-infrared, visible (VIS), ultraviolet (UV), X-ray, and gamma-ray depending on wavelength and frequency (Fig. 1). Furthermore, the nature of the interaction between radiation energy and matter determines the type of spectroscopy which can be classified as absorption, emission, reflection, elastic scattering, inelastic scattering, and nuclear spectroscopy (Delpy & Cope, 1997). A vast variety of spectroscopic sensors and technologies exist for applications in meat science including spectrometers for VIS, vis-NIR, NIR, short-wave infrared (SWIR, 900–1800 nm) and Raman spectroscopies. In addition, hyperspectral imaging can be used to create a complete picture of the environment or objects of interest whereby each pixel contains reflectance intensity information throughout the VIS and NIR spectra (Elmasry et al., 2012a). The following section contains a review of these latter methodologies for application in meat science, specifically for purposes of chemical composition and classification for authentication.

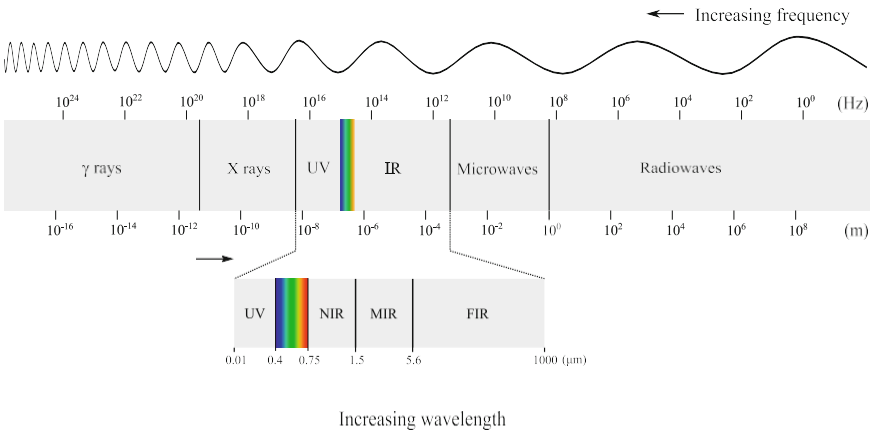

Fig. 1. The electromagnetic spectrum (Lillesand et al., 2015). UV – ultraviolet; NIR – near-infrared; MIR – mid-infrared; FIR – far-infrared.

1.2.1. Near-Infrared Reflectance Spectroscopy (NIRS)

Near-infrared reflectance spectroscopy (NIRS) is the section of the electromagnetic spectrum between 750 to 2500 nm. Traditionally, NIRS is used to predict chemical composition of soil (Wadoux et al., 2021), medicines (Zontov et al., 2016), ground grain, forages, and feed products (Williams, 2001), and dried and ground faecal matter from livestock (Dixon and Coates, 2009). Several studies have used NIRS to predict quality of meat, including tenderness, juiciness, flavour, colour, and pH (Prieto et al., 2009; Dixit et al., 2017; Dixit et al., 2021). In addition, NIRS sensors are used to predict chemical composition of meat because different chemical bonds produce vibrations reflected as peaks in the resulting spectrum where strong reflectance or absorption of energy is high (Ben-Gera & Norris, 1968; Osborne et al., 1993). Such chemical bonds correspond to protein (N-H), water (O-H) and fat (C=C, C-H) in the meat sample (Osborne et al., 1993; Prevolnik et al., 2004). Several sensors, particularly larger benchtop spectrometers, utilise the visible spectrum between 350 and 780 nm in addition to the NIR spectrum (vis-NIRS: 350–2500 nm) as it produces more accurate results with increased wavelength (Prieto et al., 2017). Smaller handheld sensors often utilise only the short-wave NIR spectrum (900–1600 nm) with a lower spectral resolution, reducing their performance (Zontov et al., 2016). However, selection of wavelength to only this range has

shown improved performance than the longer-wave NIR wavelength range (Byrne et al., 1998; Shackelford et al., 2004).

1.2.2. Raman spectroscopy

Another device that has received numerous publications in meat science is Raman spectroscopy. Raman spectroscopy was developed in 1928, and since the 1960s has been used due to its combination with lasers, whereupon interactions of photons and material result in vibrational energy resulting in backscattered frequency peaks existing at different relative wavenumbers, known as Raman shifts, due to different chemical bonds (Fowler et al., 2017; Xu et al., 2020). Raman spectroscopy has been traditionally used in the pharmaceutical industry as an authentication tool due to its ability to characterise molecular structures of chemicals (Huang et al., 2010). However, its use in meat science has resulted in variable accuracy and precision to predict meat quality and chemical composition (Cama-Moncunill et al., 2020; Fowler et al., 2015a; Fowler et al., 2017).

1.2.3. Hyperspectral imaging

Hyperspectral images are formed from measurements of the reflectance or absorption of light in multiple spatial points and multiple spectral wavelengths which allow the formation of an image with spectral information for each pixel. The HS images are downloaded as a hypercube with two spatial dimensions and one spectral dimension, resulting in hundreds of wavebands for each spatial dimension, which can be very large. However, these hypercube images can provide information about the shape, size, and colour of food products (Elmasry et al., 2012a). Each pixel within a HS image has a different spectral signature, and several pixels can form a region of interest (ROI) within an image where pixel spectra can be averaged, indices calculated, or prediction algorithms applied (Huang et al., 2014a). Hyperspectral imaging is a method of non-contact spectroscopy, which has

been used in studies with food and agricultural products integrating VIS and NIR spectra, with a pushbroom system or conveyor belt transporting an object of interest to be scanned by the HS sensors (Baeten et al., 2007; Elmasry et al., 2012a; Wold et al., 2006; Xu & Sun, 2017). Other types of sensors can also be a part of a line scanning HS system in order to provide a better spectral profile (Dixit et al., 2021). This provides a similar non-invasive line scanning method to real-time ultrasound, CT or X-ray attenuation as well as the detail for chemical parameters as demonstrated by NIRS, which provides significant advantages to the meat processing industry where sensors can become contaminated by contact with carcases and human operators (Elmasry et al., 2012a; Xu & Sun, 2017). Automation, particularly using multi-sensory platforms, is an emerging phenomenon in meat and animal science, with accurate traceability of livestock and products, plus information on the yield, composition and quality of meat and primal cuts desired by consumers (Scholz et al., 2015).

1.3. Applications of spectroscopy in meat science

Spectroscopy has found a plethora of applications in the meat industry but the most common and important include the measurement of meat eating quality, chemical composition, product authentication, and product safety for human consumption. This section describes these applications of spectroscopic sensors in the meat industry.

1.3.1. Meat quality and chemical composition

Near-infrared reflectance spectroscopy has been used predominantly in meat science for the prediction of chemical composition, where it has been more accurate compared to prediction of physical traits such as pH or eating quality traits including tenderness (Dixit et al., 2017; Prevolnik et al., 2004). Despite extensive research with NIRS to measure meat eating quality and chemical composition in the meat industry over the past two decades, NIRS sensors and devices are still not widely adopted in the industry for a variety of reasons that will be highlighted in the present section. Several studies have used the NIR spectrum to determine the concentration of water, protein, and fat in meat samples, whether fresh intact, minced, dried, or ground. Table 1 presents a detailed summary of previous research on the use of NIRS in measuring the chemical composition of meat with the main applications, accuracy and precision achieved, and considering some of the most important factors that can affect the precision and accuracy of predictions. A discussion of the factors affecting the accuracy of the prediction is discussed in section 4 below.

Studies over the past two decades have particularly focused on predicting IMF content within a given sample and making comparisons to conventional subjective marbling scores and objective chemical determination such as the Association of Official Analytical Chemists (AOAC) Soxhlet method (Helrich, 1990) using spectroscopy (Prieto et al., 2017). The NIR estimation of chemical IMF is believed to be more accurate than, and certainly have objectivity, compared to traditional marbling score for grading of meat (Bindon, 2004), particularly as different nations use

different marbling grading systems that are not based on chemical IMF concentration (Polkinghorne & Thompson, 2010; Strong, 2004). Consequently, few studies have sought to chemically measure IMF concentration against subjective marbling (Bindon, 2004; Ferguson, 2004; Harper et al., 2003; Johnston et al., 1999; Moore et al., 2010; Savell et al., 1986). In support of Bindon (2004) who suggested development of on-line NIRS to measure IMF, studies predicting chemical IMF both in laboratories using benchtop devices (Table 1) and on-site using portable devices (Table 2) have shown good precision and accuracy for research or industry screening. For example, Williams (2001) reported coefficient of determination of calibration (R^2_{cal}) of 0.50 to 0.83 and residual prediction deviation (RPD) >2 for various agricultural products to be sufficient for rough screening, and R^2_{cal} > 0.83 and RPD >3 for use of the calibration equations in research. These values were adapted for research in intact and homogenised meat (Barlocco et al., 2006). However, studies using NIRS on intact beef rib eye to predict marbling score as opposed to chemical fat have tended to yield very poor precision and accuracy (R^2_{cal} < 0.10; Magalhães et al., 2018). Therefore, the literature suggests that factors such as the characteristics of the trait being predicted, sample presentation and preparation, and the statistical methods for model development, and metrics used to present results could have a large influence on the interpretation and applicability of the results.

Table 1. Use of benchtop near-infrared spectroscopy sensors to determine chemical composition of meat. All precision and accuracy metrics are on the validation dataset unless indicated otherwise.

Brand	Wavelength range (nm)	Size (L x W x H, cm, kg)	Spectral interval and number of scans	Meat sample	Precision and accuracy metrics	Reference
Neotec 6350	1100-2500	NA	Spectral interval 2 nm 50 scans per sample	Ground beef and pork	Moisture (r_{val} = 0.93–0.98, bias = 0.01–0.42) Protein (r_{val} = 0.53–0.87, bias = 0.01–0.19) Fat (r_{val} = 0.99–1.00, bias = 0.06–0.18)	Lanza (1983)
FOSS NIRSystems 6500	400-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Intact and minced beef	Moisture (r^2_{CV} = 0.36–0.72, SE_{CV} = 10.4–15.5) Protein (r^2_{CV} = 0.50–0.79, SE_{CV} = 5.5–8.8) Fat (r^2_{CV} = 0.19–0.71, SE_{CV} = 4.7–8.1)	Cozzolino et al. (2000)
FOSS NIRSystems 6500	400-2500	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Intact and minced beef, lamb and chicken	Moisture (r^2_{CV} = 0.01–0.96, SE_{CV} = 6.9–33.1) Protein (r^2_{CV}= 0.26–0.97, SE_{CV} = 2.4–23.9) Fat (r^2_{CV} = 0.18–0.93, SE_{CV} = 4.7–46.9)	Cozzolino and Murray (2002)

					Beef pH (R^2_{CV} = 0.76–0.87, SE_{CV} = 0.13–0.18)	
FOSS NIRSystems 6500	400-2500	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Intact and minced beef	Moisture (R^2_{CV} = 0.09–0.41, SE_{CV} = 15.6–16.1) Protein (R^2_{CV} = 0.48–0.71, SE_{CV} = 20.5–23.9) Fat (R^2_{CV} = 0.89–0.92, SE_{CV} = 43.4–46.9)	Cozzolino et al. (2002)
FOSS NIRSystems 6500	400-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 30 scans per sample	Minced beef	Moisture (r^2_{CV} = 0.77, SE_{CV} = 0.58) Protein (r^2_{CV}= 0.82, SE_{CV} = 0.48) Fat (r^2_{CV} = 0.82, SE_{CV} = 0.44) Ash (r^2_{CV} = 0.66, SE_{CV} = 0.03)	Alomar et al. (2003)
FOSS NIRSystems 6500	400-2500	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Intact and minced pork	Moisture (r^2_{CV} = 0.66–0.90, SE_{CV} = 1.1–3.1, RPD = 1.4–3.9) Fat (r^2_{CV} = 0.30–0.87, SE_{CV} = 1.8–4.0, RPD = 1.1–2.3)	Barlocco et al. (2006)
FOSS NIRSystems 5000	1100-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm	Minced and freeze-dried beef	Moisture (r^2_{CV} = 0.91, SE_{CV} = 0.26–0.35) Protein (r^2_{CV} = 0.64, SE_{CV} = 0.20–0.33) Fat (r^2_{CV} = 0.99, SE_{CV} = 0.13–0.20) Ash (r^2_{CV} = 0.38–0.86, SE_{CV} = 0.03–0.09)	De Marchi et al. (2007)

Bran und Lübbe InfraAlyzer 500	1100-2500	58 x 73 x 50, 78.8 kg	Spectral interval 2 nm	Freeze-dried and ground mutton	Moisture (r^2_{val} = 0.92, SE_{val} = 0.38) Protein (r^2_{val} = 0.99, SE_{val} = 0.92) Fat (r^2_{val} = 0.99, SE_{val} = 0.43) Ash (r^2_{val} = 0.94, SE_{val} = 0.15)	Viljoen et al. (2007)
FOSS NIRSystems 6500	400-2500	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 1 scan per sample	Homogenised pork and other meats	Moisture (r^2_{val} = 0.67–0.96, RPD = 1.2–5.0) Protein (r^2_{val} = 0.11–0.96, RPD = 1.1–4.5) Fat (r^2_{val} = 0.94–0.96, RPD = 4.1–10.1)	Prevolnik et al. (2010)
Perten DA7200	950-1650	39.0 x 37.0 x 51.7 cm, 13 kg	Spectral interval 1 nm 4 scans per sample	Minced beef	Moisture (r^2_{CV} = 0.71, r^2_{val} = 0.93) Protein (r^2_{CV} = 0.81, r^2_{val} = 0.92) Fat (r^2_{CV} = 0.83, r^2_{val} = 0.91) pH (r^2_{CV} = 0.24, r^2_{val} = 0.72)	Yang et al. (2010)
SupNIR-1000	1000-1799	35.5 x 27.8 x 11.7, 6 kg	Spectral interval 1 nm 3 scans per sample	Intact and minced beef	Moisture (r^2_{val} = 0.81–0.86, SE_{val} = 0.49–0.70) Protein (r^2_{val} = 0.76–0.90, SE_{val} = 0.39–0.48) Fat (r^2_{val} = 0.66–0.94, SE_{val} = 0.55–0.73)	Sun et al. (2011)
FOSS NIRSystems 5000	1100-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 2 scans per sample	Minced beef	Fat (r^2_{CV} = 0.82, $RMSE_{CV}$ = 0.56, RPD = 2.13)	Cecchinato et al. (2012)

FOSS NIRSystems 6500	400-2500	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Ground and homogenised pork	Moisture (r^2_{val} = 0.97, SE_{val} = 0.38–0.63, RPD = 6.02–7.28) Protein (r^2_{val} = 0.88–0.90, SE_{val} = 0.55–0.76, RPD = 2.69–3.23) Fat (r^2_{val} = 0.99, SE_{val} = 0.38–0.50, RPD = 8.44–10.02)	Zamora-Rojas et al. (2012)
FOSS NIRSystems 6500	400-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 2 scans per sample	Dried and ground beef and lamb	Fat (r^2_{CV} = 0.95–0.98, SE_{CV} = 0.25, RPD = 4.60–6.59)	Mourot et al. (2014)
FOSS NIRSystems 6500	400-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Ground beef	Moisture (R^2_{cal} = 0.90, $RMSE_{CV}$ = 0.60, RPD = 2.13) Protein (R^2_{cal} = 0.85, $RMSE_{CV}$ = 0.48, RPD = 2.10) Fat (R^2_{cal} = 0.86, $RMSE_{CV}$ = 1.08, RPD = 2.01) pH (R^2_{cal} = 0.73, $RMSE_{CV}$ = 0.09, RPD = 1.14)	Prieto et al. (2014a)
SupNIR-1500	1000-1800 nm	35.5 x 27.8 x 11.7, 6 kg	Spectral interval 1 nm 3 scans per sample	Minced homogenised beef	Moisture (r^2_{val} = 0.94–1.00, SE_{val} = 1.22–4.95, RPD = 2.63–10.69) Protein (r^2_{val} = 0.95–0.98, SE_{val} = 0.69–1.40, RPD = 2.71–5.46) Fat (r^2_{val} = 1.00, SE_{val} = 0.98–1.21, RPD = 14.22–17.37)	Su et al. (2014)

StellarNet EPP2000-CXR-Srs plus EPP2000-InAs-512	400-1395	45.7 x 35.6 x 17.8	Spectral interval 5 nm 2 scans at 10 locations per sample	Intact pork	IMF (r^2_{CV} = 0.22, SE_{CV} = 1.09, RPD_{CV} = 1.2, r^2_{val} = 0.33, SE_{val} = 1.03, RPD_{val} = 1.3) pH (r^2_{CV} = 0.70, SE_{CV} = 0.11, RPD_{CV} = 2.1, r^2_{val} = 0.75, SEv_{al} = 0.11, RPD_{val} = 2.1)	Balage et al. (2015)
ASD FieldSpec Pro	350-2500	33 x 11.4 x 40. 6, 5.44 kg	Spectral resolution 3-10 nm 3 x 10 mm scans per sample at same location	Intact lamb	IMF (r^2_{CV} = 0.70, $RMSE_{CV}$ = 0.46, RPD_{CV} = 1.93, r^2_{val} = 0.66, $RMSE_{val}$ = 1.48, RPD_{val} = 1.79)	Pullanagari et al. (2015)
ASD LabSpec 5000 and LabSpec 4 Hi-Res	350-2500	36.8 x 29.2 x 12.7, 5.44 kg	Spectral resolution 1 nm 1 scan x 1 location per sample	Intact beef	IMF (r^2_{CV} = 0.70)	Pham et al. (2018)
FOSS NIRSystems 6500	400-2498	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Ground and freeze-dried/ground beef	Fat (r^2_{val} = 0.93, SE_{val} = 1.00–1.01, bias = 0.0)	Andueza et al. (2019)
ASD LabSpec 5000	350-2500	36.8 x 29.2 x 12.7, 5.44 kg	Spectral resolution 3-10 nm Spectral interval 1 nm 40 scans x 1 location per sample	Ground and freeze-dried and ground lamb	Fat (r^2_{CV} = 0.88–0.89, $RMSE_{CV}$ = 0.38–0.4, r^2_{val} = 0.75–0.83, $RMSE_{val}$ = 0.41–1.49)	Dixit et al. (2020)
ASD LabSpec 4 Hi-Res	350-2500	29.2 x 36.8 x 12.7, 5.44 kg	Spectral resolution 3-6 nm	Ground and freeze-dried and ground lamb	Fat (r^2_{CV} = 0.89, $RMSE_{CV}$ = 0.36, r^2_{val} = 0.79, $RMSE_{val}$ = 0.38)	Dixit et al. (2020)

			Spectral interval 1 nm 40 scans x 1 location per sample			
ASD LabSpec 5000	350-2500	36.8 x 29.2 x 12.7, 5.44 kg	Spectral resolution 3-10 nm Sampling interval 1.4-2 nm 40 scans x 6 locations per sample Scan time 7 s	Intact beef	IMF (r^2_{CV} = 0.88, SE_{CV} = 1.21, r^2_{val} = 0.88, SE_{val} = 1.16, bias = 0.08, RPD = 2.87) pH (r^2_{CV} = 0.90, SE_{CV} = 0.16, r^2_{val} = 0.84, SE_{val} = 0.19, bias = 0.02, RPD = 2.54)	Dixit et al. (2021)
ASD LabSpec 4 Hi-Res	350-2500	29.2 x 36.8 x 12.7, 5.44 kg	Spectral resolution 3-10 nm Sampling interval 1.4-2 nm Scan time 7 s Single line scan	Intact beef	IMF (r^2_{CV} = 0.91, SE_{CV} = 1.02, r^2_{val} = 0.89, SE_{val} = 1.12, bias = 0.02, RPD = 2.97) pH (r^2_{CV} = 0.87, SE_{CV} = 0.17, r^2_{val} = 0.86, SE_{val} = 0.18, bias = 0.01, RPD = 2.72)	Dixit et al. (2021)
ASD LabSpec 2500	350-1830	13 x 37 x 29, 6.2 kg	Spectral resolution 3-10 nm Sampling interval 1 nm 1 scan x 3 locations per sample	Intact beef	Moisture (r^2_{val} = 0.51, $RMSE_{val}$ = 1.25) Protein (r^2_{val} = 0.46, $RMSE_{val}$ = 0.51) Fat (r^2_{val} = 0.46, $RMSE_{val}$ = 1.06) Ash (r^2_{val} = 0.06, $RMSE_{val}$ = 0.06) pH (r^2_{val} = 0.19, $RMSE_{val}$ = 0.11)	Patel et al. (2021)

r^2_{CV} – coefficient of determination of cross-validation; r^2_{val} – coefficient of determination on the validation dataset; R^2_{Cal} – coefficient of determination on the calibration dataset; $RMSE_{CV}$ – root mean square error of cross-validation; $RMSE_{val}$ – root mean square error on the set-aside validation dataset; SE_{CV} - standard error of cross-validation; SE_{val} – standard error on the set-aside validation dataset; RPD – residual prediction deviation; IMF – intramuscular fat.

Note: Fat and IMF are used interchangeably as they are dependent on their mention in the original study.

The move towards smaller, handheld devices that can be used for on-site scanning in the industry was reviewed comprehensively by Dixit et al. (2017). Although benchtop devices encompass a greater wavelength range and often showed r^2 values exceeding 90% for predicting the concentration of fat and protein (Table 1), most of these findings took place in laboratories as the devices (even those labelled as portable) were too bulky to effectively be used in meat processing plants where thousands of carcases may need to be graded daily (Teixeira dos Santos et al., 2013). Additionally, ground meat was found to dominate in the most successful studies, which lacks applicability as a non-invasive tool that can be used on the surfaces of unprocessed meat such as carcases in abattoirs where meat grading is performed or even primal or retail cuts (Dixit et al., 2017). Consequently, Viljoen et al. (2007) suggested that the meat processing industry would benefit more from a device applicable to fresh meat grading than a very accurate (>95%) prediction that required preparation such as grinding, drying, or transportation to a laboratory. Despite this, few studies have explored truly handheld NIRS sensors, particularly those that connect via Bluetooth to a smartphone (Goi et al., 2022). In general, studies that have been published in the last decade have shown favourable, but not optimal, results for precision and accuracy of handheld sensors to measure chemical composition of beef and lamb (Table 2). However, the technology is advancing fast and future sensors may provide more accurate predictions. Observations from Table 2 indicate that handheld sensors can predict fat concentration of meat with r^2_{val} between 0.27 and 0.98, and protein with r^2_{val} between 0.23 to 0.85. It is important to note that different handheld sensors have a broad range of wavelengths, and this could also be affecting the results, in addition to sample processing and presentation (Teixeira dos Santos et al., 2013). In addition, studies on measurement of chemical composition using these sensors have been limited, often to only IMF (Table 2).

Table 2. Handheld NIR spectral sensors used to determine chemical composition of meat. All precision and accuracy metrics are on the validation dataset unless indicated otherwise.

Brand	Wavelength range (nm)	Size (L x W x H, cm, kg)	Spectral interval and number of scans	Meat sample	Precision and accuracy metrics	Reference
Polychromix Phazir 1624 MEMS-NIRS	1600-2400	25.4 x 15.2 x 29.4, 1.7 kg	Spectral resolution 12 nm Sampling interval 8 nm 10 scans x 4 locations per sample	Ground and homogenised pork	Moisture (r^2_{val} = 0.97, SE_{val} = 0.73, RPD = 5.81) Protein (r^2_{val} = 0.85, SE_{val} = 0.66, RPD = 2.51) Fat (r^2_{val} = 0.98, SE_{val} = 0.72, RPD = 8.08)	Zamora-Rojas et al. (2012)
Texas DLP NIRScan Nano EVM	900-1700	6.2 x 5.8 x 3.6, 84 g	Spectral resolution 10 nm Sampling interval 0.96-2.6 nm 1 scan x 3 locations per sample	Intact beef	IMF (r^2_{CV} = 0.67, $RMSE_{CV}$ = 2.46)	Pham et al. (2018)
ASD QualitySpec Trek	350-2500	30 x 10 x 31, 2.5 kg	Spectral resolution 3-9.8 nm Spectral interval 1 nm 1 scan x 1 location per sample	Intact beef	IMF (r^2_{CV} = 0.54, $RMSE_{CV}$ = 2.98)	Pham et al. (2018)

ASD QualitySpec Trek	350-2500	30 x 10 x 31, 2.5 kg	Spectral resolution 3-9.8 nm Spectral interval 1 nm 50 scans x 1 location per sample	Ground and freeze-dried and ground lamb	IMF (r^2_{CV} = 0.86, $RMSE_{CV}$ = 0.40, r^2val = 0.83, $RMSE_{val}$ = 0.43)	Dixit et al. (2020)
Texas DLP NIRScan Nano EVM	900-1700	6.2 x 5.8 x 3.6, 84 g	Spectral resolution 10 nm Sampling interval 0.96-2.6 nm 15 scans x 3 locations per sample	Ground and freeze-dried and ground lamb	IMF (r^2_{CV} = 0.85–0.88, $RMSE_{CV}$ = 0.38–0.47, r^2_{val} = 0.27–0.83, $RMSE_{val}$ = 0.38–1.28)	Dixit et al. (2020)
ASD TerraSpec Halo	350-2500	11.7 x 4.0 x 12.3, 2.5 kg	Spectral resolution 3-9.8 nm 1 scan x 5 locations per sample	Intact lamb	IMF (r^2_{CV} = 0.36–0.58, $RMSE_{CV}$ = 0.79-0.97)	Fowler et al. (2020)
Grain It Aurora NIR	950-1650	23 x 12 x 7, 2 kg	Spectral resolution 10 nm Sampling interval 2 nm 1 scan x 3 locations per sample	Intact beef	Moisture (r^2_{val} = 0.63, $RMSE_{val}$ = 1.07) Protein (r^2_{val} = 0.23, $RMSE_{val}$ = 0.62) Fat (r^2_{val} = 0.66, $RMSE_{val}$ = 0.80) Ash (r^2_{val} = 0.01, $RMSE_{val}$ = 0.06) pH (r^2_{val} = 0.04, $RMSE_{val}$ = 0.16)	Patel et al. (2021)
JDSU Micro NIR Pro	905-1649	4.5 x 4.5 x 4.2, 60 g	Spectral resolution 10 nm	Intact beef	Moisture (r^2_{val} = 0.70, $RMSE_{val}$ = 0.96)	Patel et al. (2021)

			Sampling interval 2 nm 1 scan x 3 locations per sample		Protein (r^2_{val} = 0.42, $RMSE_{val}$ = 0.52) Fat (r^2_{val} = 0.62, $RMSE_{val}$ = 0.86) Ash (r^2_{val} = 0.03, $RMSE_{val}$ = 0.06) pH (r^2_{val} = 0.19, $RMSE_{val}$ = 0.14)	
Consumer Physics SCiO	740-1070	4.0 x 1.9 x 6.8, 35 g	Spectral resolution 10 nm Sampling interval 1 nm 1 scan x 5 location per sample	Intact beef	Moisture (r^2_{CV} = 0.84, SE_{CV} = 0.60, RPD = 2.48) Protein (r^2_{CV} = 0.66, SE_{CV} = 0.59, RPD = 1.72) Fat (r^2_{CV} = 0.79, SE_{CV} = 0.80, RPD = 2.19) pH (r^2_{CV} = 0.52, SE_{CV} = 0.06, RPD = 1.46)	Goi et al. (2022)

r^2_{CV} – coefficient of determination of cross-validation; r^2_{val} – coefficient of determination on the validation dataset; $RMSE_{CV}$ – root mean square error of cross-validation; $RMSE_{val}$ – root mean square error on the set-aside validation dataset; SE_{CV} - standard error of cross-validation; SE_{val} – standard error on the set-aside validation dataset; RPD – residual prediction deviation; IMF – intramuscular fat.

Note: Fat and IMF are used interchangeably as they are dependent on their mention in the original study.

Hyperspectral imaging has been used successfully in beef to determine chemical composition (Kamruzzaman et al., 2012a) and quality traits such as pH, shear force and colour (Kamruzzaman et al., 2012b; Wang & Peng, 2018), and discrimination into different quality categories such as by tenderness grade (Konda Naganathan et al., 2008). In more recent studies, HS imaging showed very positive results in predicting IMF concentration in intact meat in real time (Dixit et al., 2021; Hitchman et al., 2021). Dixit et al. (2021) found that the VIS-NIR wavelength range (670–950 nm) showed greater accuracy and precision for pH and IMF concentration predictions than visible-only (470–630 nm), although the difference was much more pronounced for IMF (Table 4). The use of HS to classify samples into categories based on different spectral signatures has yielded high accuracy. Konda Naganathan et al. (2008) found that 100% of tough steaks were classified as tough, and more than 95% of tender steaks were classified as tender, although the sample size of tough steaks was very small.

Similar to the recent HS studies, charge coupled device (CCD) (Kuchida et al., 2000) and mirror-type cameras (Kuchida et al., 2001) have been developed to analyse rib eye marbling of cattle for commercial use. Due to the ability of the camera to predict the fat area ratio of a rib eye, as well as the overall and maximum coarseness of marbling particles and the number of small flecks (Nakahashi et al., 2008), the measure is denoted as objective marbling percentage and an alternative to chemical IMF measurement (Connolly et al., 2020; Kuchida et al., 2000; Maeda et al., 2014; Stewart et al., 2021). Very high precision and accuracy (r^2 = 0.96; bias -6.4 to 4.0 %) was found between crude fat and charge coupled objective marbling over a wide range of IMF concentration (2.9–39.8% IMF) (Kuchida et al., 2000). However, lower precision (r^2 = 0.52) was found using mirror-type cameras on Korean beef (6–25% IMF), though this was greater than the relationship between crude fat and subjective marbling (r^2 of 0.41) (Beak et al., 2021). It is worth noting that most of the above studies have been undertaken on Japanese and Korean cattle with 16-66% of the eye muscle area represented by marbling but most carcases in the global meat industry have significantly less marbling. Over the past 2-3 decades, video image analysis systems such as the VIAscan have been

developed for commercial objective evaluation of marbling against the United States Department of Agriculture (USDA) marbling score (Belk et al., 2000; Moore et al., 2010), AUS-MEAT marbling in Australia (Ferguson, 2004; Stewart et al., 2021) and against MSA traits in Poland (Konarska et al., 2017). Similarly, stationary CCD cameras using red-green-blue (RGB) representation have been commercially available and tested on European breed cattle against) USDA marbling (Gerrard et al., 1996) and both chemical IMF and MSA marbling in Australia (Stewart et al., 2021). In carcases where IMF is not overly high (1–20%) precision tends to not be particularly high ($r^2 < 0.80$; Konarska et al., 2017; Stewart et al., 2021). Better precision was found in an earlier study ($r^2 = 0.84$), although chemical IMF was not measured and a smaller range of subjective marbling score was obtained (Gerrard et al., 1996).

1.3.2. Meat authentication

The authentication of meat refers to the determination of meat products according to the length or type of feeding regime such as grass- or grain-fed, geographic origin such as region or country, species of pasture or concentrate fed, meat adulteration or substitution such as the mixing with cheaper animal products such as offal or horse meat with beef, and the addition of non-meat ingredients (Ballin, 2010). Spectroscopic sensors such as NIRS, Raman spectroscopy, and HS imaging have been widely researched for such applications in the livestock industries (Esteki et al., 2018; Lohumi et al., 2015; Monahan et al., 2018).

Studies that scanned materials other that meat with NIRS sensors proposed the use of NIRS as an authentication tool. For instance, Dixon and Coates (2009) scanned faecal matter of grazing animals to predict stable isotope concentration, which was then able to correctly determine the types of forages consumed. Studies on stable isotope content of beef have allowed for successful (>80% accuracy) determination of country of origin (Osorio et al., 2011), feeding regime of grass or grain (Osorio et al., 2011), or grazing of grass or grass/clover pasture (Moloney et al., 2018). Similarly, NIRS scanning of meat has been able to determine the type of pastures or forages

consumed by cattle regardless of sample preparation such as pasture/maize silage (Cozzolino et al., 2002) or barley/blend/maize (Barragán et al., 2021). However, differences in cattle breed and genetics within sheep flocks can affect adipose tissue deposition and therefore IMF concentration, which can in turn affect the performance of prediction models to authenticate meat products (Hitchman et al., 2021; Prieto et al., 2011). Similar differences occur due to feeding regime, whereupon grain-fed animals show increased fat deposition and more rapid weight gain than grass-fed animals (Priolo et al., 2002; Leheska et al., 2008), whereas fat deposition and IMF concentration are dependent on the length of time the animals consumed the high-grain diet (Duckett et al., 1993). This is not always the case and is often dependent on genotype and metabolites (Connolly et al., 2019). Furthermore, grass-fed beef is noted for its increased proportion of n-3 fatty acids (Duckett et al., 1993; Daley et al., 2010), whereas grain feeding is generally linked to increased palatability and more species-specific flavours in beef (Hwang & Joo, 2017) and lamb (Priolo et al., 2002). Two important points raised above are the need for handheld NIRS sensors and authentication tools to classify meat as grass-fed or grain-fed. Such sensors would provide an objective tool allowing for prevention of food fraud and product mislabelling that can be used in food supply chains (Lohumi et al., 2015). However, no studies have been published using portable, low-cost smartphone sensors to discriminate between grass-fed and grain-fed meat products. Furthermore, automated devices that can authenticate meat products at high speed have not been evaluated and multi-sensory platforms have the potential to achieve this.

Raman spectroscopy has traditionally been used as an authentication tool for detection of adulterants in olive oil, fruits, and meats (Lohumi et al., 2015). The ability of Raman spectroscopy as a tool to classify beef as grass-fed or grain-fed has been explored successfully in recent publications by Logan et al. (2020a; 2020b; 2021a; 2021b). However, the latter studies used only fat as a medium for scanning using Raman. Given its success in accurate classification of beef as grass-fed or grain-fed based on spectral signature (Table 3), Raman scanning of fat can be used as a baseline for comparison when testing the scanning of lean beef tissue, the ability for differentiation of other

species, and the ability for other technologies such as NIRS based on feeding regime. Hyperspectral imaging has also been used to detect concentrations of contaminants in mixtures of livestock products such as bacteria (Cheng & Sun, 2015), offal (Kamruzzaman et al., 2014), and other species (Kamruzzaman et al., 2013; Kamruzzaman et al., 2016) which are known to have different spectral signatures.

Table 3 shows studies that have used NIRS, Raman spectroscopy, and HS imaging to authenticate meat samples and provide an objective point of reference beyond simple product labelling. Such measures are necessary to combat food fraud and product mislabelling (Ballin, 2010; Esteki et al., 2018). Most of these studies have examined the presence of adulteration in mincemeat mixtures, whereupon products sold as minced beef may contain adulterants such as horsemeat, pork, trimmings, and offal from other species. One particularly noteworthy study was that from Morsy and Sun (2013) who obtained 100% accuracy in adulterant detection in fresh minced beef using both PLS-DA and LDA, and accuracy above 75% on frozen samples. Similarly, Dian et al. (2008) used vis-NIRS to differentiate grass-fed from grain-fed beef with 98% accuracy. However, these and other studies shown in Table 3 used benchtop NIR systems which affected the use of such technologies at processor, retailer, and consumer levels.

More recently, studies using portable NIRS and Raman sensors have shown promise, though not the same accuracy as benchtop systems. Raman spectroscopy of fat was used to discriminate grass- and grain-fed beef (several studies by Logan et al.; Table 3). Zając et al. (2014) identified five key Raman shifts for differentiating beef meat from horse meat (937, 879, 856, 829, and 480 cm^{-1}). Similarly, a further five shifts were found to differentiate grass-fed from grain-fed beef (1658, 1445, 1301, 1127, and 1069 cm^{-1}) (Logan et al., 2020a). In addition, a portable visible spectrometer was able to successfully differentiate grass-fed from grain-fed lamb with accuracy ranging from 80 to 89% (Dian et al., 2007). Although not truly handheld devices, the use of scanning probes on NIRS devices such as the LabSpec and AgriSpec models (from ASD, Inc.) provide the option to scan

carcases and cuts of meat in abattoirs, with feeding regime predicted in beef to varying degrees of accuracy (25–100%) depending on the feeding type and dataset used (Barragán et al., 2021). These findings therefore call for further research for miniaturised NIR devices seen in Table 2 to be evaluated as authentication tools. In addition, VIS and NIRS studies by Dian et al. (2007; 2008) and Raman spectroscopy studies by Logan et al. (2020a; 2020b) scanned only perirenal and point end brisket fat, respectively, and not the lean tissue of carcases or other muscles of commercial importance. Lean cuts without testable fat cover are commonly sold in supermarkets and a study that uses lean tissue scanning to authenticate the meat using NIR or Raman spectroscopy by feeding regime, species, or geographical origin could be of immense value to the meat industry (Ballin, 2010).

Table 3. Classification of meat and animal products for human consumption using NIR and Raman spectroscopies. All precision and accuracy metrics are on the validation dataset unless indicated otherwise.

Brand	Range	Size (L x W x H, cm, kg)	Spectral interval and number of scans	Meat sample	Precision and accuracy metrics	Reference
FOSS NIRSystems 6500	400-2500 nm	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Intact and minced beef	Maize silage fed (79% accurate) Pasture fed (82% accurate)	Cozzolino et al. (2002)
FOSS NIRSystems 6500	400-2500 nm	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 30 scans per sample	Ground beef	Friesian or Hereford (79% accurate) Muscle classification (90–98% accurate)	Alomar et al. (2003)
FOSS NIRSystems 6500	400-2500 nm	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 32 scans per sample	Intact and minced beef	Species classification by PCR (91% accurate) Species classification by PLS (96% accurate)	Cozzolino and Murray (2004)
Renishaw 2000 Raman probe	100-3000 cm^{-1}	NA	Laser 785 nm, 78 mW Spectral resolution 0.5-1 cm^{-1} 10 s integration x 1 accumulation x 15 scans per sample	Ground chicken and turkey breasts and legs	Classification of muscle and species (83% accurate)	Ellis et al. (2005)
FOSS NIRSystems 6500	400-2500 nm	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 2 nm 3 scans per sample	Minced lamb fat	Classification of grass-fed and grain-fed lamb (r^2_{CV} = 0.88, SE_{CV} = 0.17, 98% accuracy)	Dian et al. (2008)

Bruker MPA + VERTEX 70 NIRS	12500-3750 cm^{-1}	59 x 55 x 38.5, 30 kg	Spectral resolution 8 cm^{-1} 120 scans per sample	Minced beef adulteration with turkey meat	Classification into five classes of adulteration (71% accurate)	Alamprese et al. (2013)
FOSS NIRSystems 6500	400-2498 nm	29.5 x 38.8 x 54.6, 18.9 kg	Spectral interval 14 nm 32 scans x 2 locations per sample	Minced beef, pork, fat trim and offal	Detection of adulteration in beef (pork, fat trimming, offal) at 100% accuracy	Morsy and Sun (2013)
DeltaNu ExamineR Raman microscope	200-2000 cm^{-1}	NA	Laser 785 nm, 100 mW Spectral resolution 2 cm^{-1}	Ground beef and horse fat	No overlap in PCA clusters between 0, 25%, 50%, 75% and 100% mixtures Detection of horsemeat contamination if 25% or more	Boyacı et al. (2014)
ASD LabSpec 4	350-2500 nm	29.2 x 36.8 x 12.7, 5.44 kg	Spectral interval 1 nm 50 scans x 9 locations per sample	Intact beef	Detection of dark cutting (88–95% accurate)	Prieto et al. (2014b)
FOSS NIRSystems 6500	400-2498 nm	29.5 x 38.8 x 54.6, 18.9 kg	Spectral resolution 2 nm 32 scans x 2 locations per sample	Intact and minced beef	Detection of dark cutting (90–91% accurate)	Prieto et al. (2014b)
Renishaw inVia R4 Raman microscope	400-1500 cm^{-1}	111.6 x 93 x 61, 90 kg	Laser 785 nm, 6 mW Spectral resolution 0.5-1 cm^{-1} 10 s integration x 9 accumulations x 10 scans per sample	Intact chicken, beef, turkey, mutton, pork and horse	Detection of species (85–95% accurate)	De Biasio et al. (2015)

DXR SmartRaman	900-1800 cm^{-1}	69 x 97 x 46, 56.7 kg	Laser 780 nm, 100 mW Sampling interval 2 cm^{-1} 15 s integration x 20 accumulations x 3 locations per sample	Beefburgers made with rusk, water and offal	Detection of offal adulteration (81–100% accurate)	Zhao et al. (2015)
Metrohm MIRA 785	400-2300 cm^{-1}	4.53 x 8.82 x 12.65, 705 g	Laser 785 nm, 100 mW Spectral resolution 8-10 cm^{-1} 3 s integration x 5 accumulations x 3 locations per sample	Intact beef fat	Accuracy and precision of grass and grain discrimination model 96.5%	Logan et al. (2020b)
Metrohm MIRA 785	400-2300 cm^{-1}	4.53 x 8.82 x 12.65, 705 g	Laser 785 nm, 100 mW Spectral resolution 8-10 cm^{-1} 3 s integration x 5 accumulations x 3 locations per sample	Intact beef fat	Long-fed grain feeding (100% accuracy, 95% precision) Short-fed grain feeding (93% accuracy and precision) Supplement and grass feeding (95% accuracy and precision) Grass only feeding (90% accuracy, 93% precision)	Logan et al. (2021a)
Metrohm MIRA 785	400-2300 cm^{-1}	4.53 x 8.82 x 12.65, 705 g	Laser 785 nm, 100 mW Spectral resolution 8-10 cm^{-1}	Intact beef fat	Long-fed grain feeding (98% accuracy, 100% precision) Short-fed grain feeding (91.1% accuracy, 86% precision)	Logan et al. (2021b)

			3 s integration x 5 accumulations x 3 locations per sample		Supplement and grass feeding (92.1% accuracy, 82% precision) Grass only feeding (93.1% accuracy, 77% precision)	
ASD LabSpec 4	350-2500 nm	29.2 x 36.8 x 12.7, 5.44 kg	Spectral interval 1 nm 50 scans x 4 tissues per sample	Intact beef	Barley feeding (44–100% accurate) Corn feeding (56–88% accurate) Blend feeding (25–38% accurate)	Barragán et al. (2021)

r^2_{CV} – coefficient of determination of cross-validation; SE_{CV} - standard error of cross-validation; PCA – principal components analysis.

1.3.3. Livestock product safety

Following a comprehensive review of research papers, it was determined that almost no studies have used line scanning to determine the health status or defects of different organs, carcases or meat products. However, review articles summarising studies from the past two decades have heralded automation in the meat processing industry for health and safety hazard detection (Uzal et al., 2002; Webber et al., 2012). The majority of line scanning in this area has instead focused on live animals prior to slaughter (Ezanno et al., 2021; Neethirajan et al., 2017). One such area that has been proposed for automation has been the inspection of offal for evidence of animal diseases, where non-contact line scanning sensor technology could potentially prevent spread of potential zoonotic diseases (Samuel et al., 1980). The organs identified for inspection using automatic sensors were parenchymatous, known to be the most commonly condemned organs including hearts, lungs, livers, and kidneys (Cook & Anderson, 2017). The most commonly detected lesions causing rejection in the abattoir were presented by the Agriculture and Horticulture Development Board (AHDB, 2017).

AgResearch (2018) proposed the use of multi-sensory platforms to provide full examination of offal for both identification of the type and inspection of the health status. For instance, HS imaging of a selection of condemned organs was able to identify regions of cirrhosis in the liver, abnormal biomass such as cysts and abscesses in the liver, pneumonia, nephritis, and lesions compatible with *Cysticercus ovis* infection in the heart. However, cysts and abscesses were not identified 100% of the time, and diseases such as peritonitis and lung pleurisy were not easily identified. Dual-energy X-ray attenuation was able to identify cysts from liver cirrhosis and fluke worm (AgResearch, 2018). Cook and Anderson (2017) used CT to provide information on animal health status by scanning viscera, although this was a preliminary exploration. This latter study found that liver fluke, interstitial nephritis, and discolouration in the heart were not clearly detected by CT, although it was mentioned by the authors that augmentation with other sensing technology would have improved the results. It is hypothesised that sensors that can measure discolouration

and detect defects such as parasites in other agricultural products such as HS imaging may have the ability to detect these defects in offal (Elmasry et al., 2012a; Xu & Sun, 2017), but no scientific evidence exists to demonstrate this. However, detection using CT would be achievable for abnormalities presenting differences in fluid content such as cysts, abscesses and eggs from parasites such as tapeworm (Cook & Anderson, 2017). It is worth noting that these were pilot studies that were not published in peer-reviewed journals and their findings are to be considered as preliminary. Computed tomography (CT) has been used successfully to identify lesions on organs in a veterinary medicine context, although these studies scanned sedated live animals (Lee et al., 2009; Lee et al., 2011).

Hyperspectral imaging has not been used to differentiate organs in the abattoir, nor has it been used to identify transmissible health risks. It is believed that multi-sensory systems similar to those used previously in meat studies (Dixit et al., 2021; Hitchman et al., 2021) could provide insight into the spectral signatures of different organ types and identify spectral differences between healthy and unhealthy organs or tissues. This was seen in a prior study using HS to differentiate organs by spectral signature during exploratory surgery on a pig (Akbari et al., 2008). Once algorithms are developed for identification and disease detection purposes, multi-sensory systems can allow for automation in the offal sortation process in the meat industry and be distributed in abattoirs.

Hyperspectral sensors have also been used in line scanning systems to detect contamination and a range of defects related to product safety and health of slaughtered animals. For example, HS imaging has been used successfully in conjunction with PLSR and least squares support vector machines to estimate the total viable count of bacteria from spectra acquired within selected ROI of fish fillets where traditional plating measurement occurred (Cheng & Sun, 2015). Similarly, skin tumours in chicken carcases were identified using discriminant analysis with error rates <6% (Nakariyakul & Casasent, 2009).

1.4. Factors affecting the accuracy of predictions

Multiple factors can affect the accuracy of predictions from NIRS, Raman spectroscopy and HS imaging including the spectral range, spectral resolution or interval, number of scans per sample, sample presentation, and modelling method (Ben-Gera & Norris, 1968; Huang et al., 2014a; Prieto et al., 2017; Xu et al., 2020). Therefore, these factors were included in the tables to summarise and understand the differences between different spectroscopy systems used in previous studies. In addition, day-to-day variation has been mentioned to be a problem, particularly in smaller spectrometers (Dixit et al., 2020). These have been attributed to variations in temperature, humidity, dust, and other environmental factors (Williams et al., 2017).

Data pre-processing methods such as log(1/R), first and second derivatives, among others, can also have a large influence on the results (Huang et al., 2010; Zeaiter et al., 2005). Furthermore, it is important to note that the precision and accuracy of predictions could be presented in the literature for the calibration or model development dataset, for the resampling or cross-validation dataset, or for a set-aside or independent validation dataset. The best approach is to have a set-aside dataset for validation (often 15–40% of the total sample size), although this is not always possible due to smaller sample sizes (Dixit et al., 2017; Zeaiter et al., 2005). Resampling or cross-validation is frequently used when sample size is below 150 and various cross-validation methods including leave-one-out or k-fold are required (Dixit et al., 2017; Williams et al., 2017). Some studies have employed all three datasets, where k-fold cross-validation was used to prevent overfitting of calibration models, and a validation dataset was used to test the developed models (Balage et al., 2015; Dixit et al., 2020; Dixit et al., 2021; Kamruzzaman et al., 2012a).

These factors are limitations for comparing results from different studies and call for the development of a standard procedure for sample and data processing in meat science. In addition, more research is required to compare different sample and data processing methods. However, sample presentation could present challenges as it depends on the variable being predicted such as

pH or marbling score in intact meat or chemical crude protein or fat content in dried and ground meat. The use of NIRS, Raman and HS models on intact meat is desired due to this being their most common presentation at processor, retailer and consumer level. Similar to handheld devices, this is the most practical solution for the industry.

1.4.1. Spectral range, spectral resolution, and practicality

Comparison is essential to the trialling of new technology, with handheld NIRS sensors rarely used in literature (Table 2) compared to vis-NIRS benchtop sensors (Table 1). Traditionally, shorter wavelength ranges have limited the prediction accuracy of sensors. For instance, the use of vis-NIR (400–2500 nm) outperformed visible-only (400–700 nm) spectroscopy (r^2_{CV} = 0.88 vs. 0.68, correct classification 98% vs. 90%) in the discrimination of grass-fed and grain-fed lamb when homogenised perirenal fat was scanned (Dian et al., 2008).

Several benchtop vis-NIRS sensors have been deemed as “portable” with the addition of a scanning probe, although these still require a trolley, mains power, and computer connection to be used for on-line measures on carcases in an abattoir (De Marchi, 2013; Patel et al., 2021; Pullanagari et al., 2015). Accuracy and precision of chemical composition tends to be lower in handheld sensors compared to benchtop sensors (Dixit et al., 2017), although their low cost, lightweight nature, and ease of use upon entire cuts of meat make these more attractive to consumers (Teixeira dos Santos et al., 2013). Even less studies have used smartphone connectable sensors, which are billed as “palm-sized”, “mini”, or “micro” NIRS (Dixit et al., 2020; Goi et al., 2022; Patel et al., 2021; Pham et al., 2018). Such devices are marketed to consumers, and despite having shorter wavelength ranges than larger vis-NIRS sensors (900–1700 nm compared to 350–2500 nm), their accuracy has been deemed to be suitable for “rough screening” in the abattoir or by consumers and retailers (Goi et al., 2022).

The use of “rough screening” procedures in the abattoir (Williams, 2001) allows meat cuts to be sorted into groups, e.g., based on tenderness (Shackelford et al., 2005) or colour (Prieto et al., 2014b), which are very important traits for consumer selection. Similarly, the use of NIRS as an objective grading system of IMF, may allow for quality grading based on chemical IMF prediction to take place using a handheld sensor, as proposed by Bindon (2004), which is not the case for subjective marbling score grading (Savell et al., 1986). Despite these assertions, few studies have attempted comparisons of handheld sensors connecting to a smartphone with benchtop NIRS sensors for chemical parameters.

1.4.2. Sample preparation and product type

Prediction of chemical parameters has been found to be more precise in minced and freeze-dried-ground meat compared to intact meat (Table 1). This occurs due to the homogenisation of samples following mincing and further following freeze-drying and grinding, which is a requirement of sample preparation in the laboratory for the AOAC Soxhlet (IMF) and Kjeldahl (crude protein) analyses (Helrich, 1990). The use of inhomogeneous media such as FI meat affects the spectra and thus the accuracy of prediction models in comparison to more homogenous material such as minced or dried and ground meat, because vibrations can be obtained from different tissue types simultaneously (Delpy & Cope, 1997). This can be seen in previous studies where chemical content predictions on intact meat can be very promising with $r^2 > 0.60$ and often exceeding 0.8 (Kamruzzaman et al., 2012a; Sun et al., 2011; Yang et al., 2010). However, scanning of unprocessed meat cuts is still rare (Table 1; Table 2) and predictions generally fall below the required levels of precision and accuracy required for use in the industry (Williams, 2001). For instance, Cozzolino and Murray (2002) found CP and moisture predictions to have significantly lower precision in intact meat compared to minced meat, although limited differences were seen in IMF predictions. However, some of the smaller handheld sensors have been developed for use by consumers and retailers, and their use requires intact samples (Goi et al., 2022; Patel et al., 2021).

Originally, all NIR predictions of chemical composition took place on benchtop devices using homogenised samples in a small sample cup. For instance, Ben-Gera and Norris (1968) used only 2 g of sample in a 2 mm cup, while more modern studies have filled cups with diameter 50–70 mm and depth 10–25 mm with meat paste (Cozzolino et al., 2000; Su et al., 2014). Larger homogenised samples (150 g) have been used in 165 x 35 mm cuvettes (Alomar et al., 2003), whereas intact samples were scanned at dimensions up to 100 x 50 x 25 mm (Barlocco et al., 2006). This is a small sample of an experimental unit such as a steak, primal cut of muscle, or an entire animal. Similarly, a larger aperture size, or field of view area, of a sensor allows data from a larger portion of the sample to be captured, and this is rarely mentioned in NIRS studies although it is regularly measured in colorimetry studies (Holman et al., 2015). Measurement probes for benchtop NIRS ranging in aperture size from 2 to 55 mm (Shackelford et al., 2004) and 2.5 to 10 mm for handheld NIRS devices (Dixit et al., 2020) have allowed for the scanning of intact meat which allows the IMF concentration of a sample to be estimated on-line. Dixit et al. (2020) mentioned that the handheld sensor with a smaller viewed area produced a lower signal to noise ratio. The majority of benchtop NIRS systems scan the sample multiple times prior to obtaining the average spectra of a sample. The NIRS used by Lanza (1983) scanned a sample 50 times, with 5 scans per second, totalling 10 seconds of scanning, whereas ASD spectrometers have noted a 5 s time for a reading encompassing 50 scans using a probe (Barragán et al., 2021; Prieto et al., 2014b). Conversely, Dixit et al. (2021) noted a 7 s time per scan.

Devices that do not require contact with the samples are preferred for applications in the meat supply chain. Non-contact NIRS systems have also been configured in the form of line scanners with a conveyor belt, for example Dixit et al. (2021) collected spectra (100 mm area) from a 180 mm stand-off distance. Similarly, Goi et al. (2022) collected spectra from meat using a handheld NIRS held at a distance of 10 mm from the sample surface.

Similarly, collection of broader ranges of samples, such as more muscles and species, tend to improve precision and accuracy of NIRS prediction models, which also occurs in traits with greater

variation. This tends to result in greater precision and accuracy of IMF prediction models compared to chemical constituents with low variation between samples such as pH and CP (De Marchi et al., 2013). Prevolnik et al. (2010) found that using multiple muscles from pigs gave greater precision (r^2_{val} = 0.96) and accuracy (RPD = 4.5) to predict CP compared to only the *longissimus dorsi* (r^2_{val} = 0.11, RPD = 1.1). In the same study, use of different meat products significantly increased the precision and accuracy for moisture prediction, whereas IMF showed r^2_{val} > 0.90 and RPD > 4.0 regardless of muscle or species.

1.4.3. Methods for data processing and analysis

Pre-processing of NIR and HS data is necessary in almost all cases. Spectra are often cleaned to remove outliers and extreme values (outside confidence interval on a T^2 or Q residual plot) using principal components analysis (PCA) prior to analysis (Dixit et al., 2021; Hitchman et al., 2021) or predicted values more than 2.5 standard error from reference values (Andueza et al., 2019; Goi et al., 2022). The Lambert-Beer Law is commonly used to convert reflectance data recorded by the spectrometer to absorbance data by the equation absorbance = log(1/reflectance) due to the linear relationship between absorbance and concentration of the absorber (Norris et al., 1976), which is then predicted by multivariate calibration (Zeaiter et al., 2005). Further pre-processing of the spectra such as the calculation of the first or second derivative reduces multiplicative effects from the background and capture variations in the slope, respectively, and thereby may improve the accuracy and precision in most cases (Lanza, 1983). Spectral data can also be smoothed using combinations of Savitzky-Golay filters (Savitzky & Golay, 1964), Fourier transformation (Zeaiter et al., 2005), multiplicative scatter correction (Liao et al., 2010), standard normal variate (SNV) and detrend (Barnes et al., 1989). However, there are some studies where accuracy and precision are greater when no pre-processing methods are used (Kamruzzaman et al., 2015).

Multivariate calibration or model development methods include partial least squares regression (PLSR), PCA, and multiple linear regression. Of these, PLSR is dominant in meat science research on quality and chemical composition using NIRS (Dixit et al., 2017; Prieto et al., 2017), Raman spectroscopy (Fowler et al., 2017; Xu et al., 2020), and HS (Kamruzzaman et al., 2012b; Wang & Peng, 2018). The combination of PLSR with linear discriminant analysis (LDA), known as partial least squares discriminant analysis (PLS-DA) dominates classification studies in meat and agricultural product research (Ariana & Lu, 2010; Huang et al., 2014a). For HS imaging, more accurate results have been found when using it as a classification tool in discriminant analysis compared to prediction by linear regression (Huang et al., 2014a), for example PLS-DA modelling for beef in tenderness categories (Konda Naganathan et al., 2008). Conversely, Raman spectroscopy has mostly used PCA and hierarchical cluster analysis when it is used as a discrimination tool (Boyacı et al., 2014; Logan et al., 2020b; Xu et al., 2020).

Decision tree machine learning algorithm methods such as classification and regression trees (CART) and Cubist have been used in NIRS soil chemistry and quality analyses (Minasny & McBratney, 2008; Wadoux et al., 2021), although have not been used in meat studies. Tang et al. (2020) found that Cubist models outperformed PLSR in terms of precision in soil chemical properties. Similarly, the random forest (RF) method, an extension of decision trees, is known for its ability to achieve high accuracy on large datasets (Liaw & Weiner, 2002). Despite this, RF has been sparsely used in classification studies (Huang et al., 2014a; Minasny & McBratney, 2008), although in one study RF showed greater accuracy for rice cultivar classification compared to PLS-DA (Kong et al., 2013).

Williams et al. (2017) stated that the optimal NIRS prediction model is based on the highest correlation coefficient (r) or coefficient of determination (R^2), lowest standard error of cross-validation (SE_{CV}), lowest RMSE, and highest RPD. The use of the R^2 value on both calibration and validation (external or cross-validation) datasets allows for precision of the prediction models to be

determined, whereas Lin's concordance correlation coefficient, RPD, bias and SE values determine accuracy (Tedeschi, 2006). The RPD value can also be used to determine suitability of a given prediction model and NIR system to the meat or agricultural industry (Williams, 2014). Calibration equations developed from PLSR, or other statistical methods can then be used for predictions on future samples, provided these metrics are sufficiently high on both calibration and on independent set-aside or cross-validation datasets (Williams, 2014; Williams et al., 2017; Williams et al., 2019). However, sample sizes are often a concern when drawing conclusions because small sizes may lack the appropriate variance compared to the ranges expected on future samples. This can therefore affect the suitability of many devices and models for applications in the industry. Consequently, the RPD for most NIRS meat science projects falls between 2.0 and 2.4 which gives suitability for rough screening, such as classification in the abattoir or at retail level. An RPD value less than 2.0 is insufficient for prediction or classification, while above 3.0 is required for quality or process control in industry (Williams, 2001; Williams, 2014).

In classification studies employing PLS-DA, LDA, or RF (Kuhn, 2020); sensitivity refers to the proportion of true positives for a given positive class, specificity to the proportion of true negatives for a given positive class, precision (or, positive predictive value) to the proportion of true positives divided by marked positives, and accuracy (balanced accuracy) to the arithmetic mean of sensitivity and specificity. Overall accuracy of a dataset or table is the mean of the sensitivities of all classes.

1.5. Real-time line scanning in the meat industry

It is imperative that future development of spectral sensors in the meat industry focuses towards in-line scanning of meat and meat products. The non-contact nature of these sensors provides a safer alternative to handheld sensors in their ability to be installed in pilot plants prior to commercialisation in the industry and not allow for any cross-contamination. Furthermore, larger

systems have the ability to act as multi-sensory platforms whereupon HS imaging can take place along with other types of imaging including NIRS (Dixit et al., 2021) or X-ray (AgResearch, 2018). Early adoptions of conveyor belt technology for in-line scanning to predict chemical composition measured vis-NIRS light reflection from ground beef every 1/30 s (Anderson & Walker, 2003). However, the latter study only predicted fat concentration and was considered unable to scan non-uniform surfaces (Dixit et al., 2016; Dixit et al., 2017). Positive results have been reported for non-contact vis-NIRS moisture prediction in intact coalfish using NIR transflectance spectra (r^2_{CV} = 0.77–0.79), with moderate prediction using non-contact NIRS (r^2_{CV} = 0.55) (Wold et al., 2006). Similarly, prediction of pH in intact pork using an optoelectronic sensor showed high accuracy, with other predictions considered as reasonable (Liao et al., 2010). The use of beam splitting was found to improve the accuracy of predictions in minced beef when scanned at distances of 10.15 and 4.0 mm (Dixit et al., 2016). Multipoint NIR analysis employs the spectra from several pixels to be analysed to improve the accuracy of predictions compared to contact NIRS probes (Dixit et al., 2017).

As mentioned in the previous section (3.2), several HS imaging systems have been used successfully to determine contents of offal, or the presence of different species within minced meat mixtures (Table 4). These studies have shown that automation within the meat industry is possible and that such technologies could provide value and increase profits associated with reduced labour requirements in the long run (Dixit et al., 2017). Table 4 shows that, when presented, RPD values tend to be 3 or greater in line scanning systems, indicating the ability of the models to be used in industry for screening or even quality control purposes (Williams, 2014). These have been particularly high for detection of adulterants or contamination within minced meat. The very high accuracy for discrimination studies also shows the promise for HS systems, although in many cases more samples are required prior to deployment in the industry.

Table 4. Studies with on-line scanning of animal and meat products using different technologies such as NIRS and HS imaging.

Brand	Wavelength range	Spectral resolution, scan distance and number of scans	Meat sample	Precision and accuracy metrics	Reference
Perten DA-7000 NIR	900-1900 nm	Spatial resolution 3.3 cm Scan distance 12.2 cm	Homogenised beef	Fat (r^2_{val} = 0.83–0.93, SE_{val} = 2.15–2.28)	Anderson and Walker (2003)
TiTech Visionsort AS NIR	760-1040 nm	Spectral resolution 20 nm Spatial resolution 0.05 mm Scan distance 1-5 cm	Intact and homogenised fish	Moisture (r^2_{CV} = 0.55–0.94, $RMSE_{CV}$ = 0.6–1.9)	Wold et al. (2006)
Imperx IPX-2M30 HS	400-1000 nm	Spectral resolution 2.8 nm Spatial resolution 1600 x 1200 px Region of interest 200 x 600 px	Intact beef	Classification of steaks into tenderness categories (93.7% accurate)	Konda Naganathan et al. (2008)
Specim ImSpector V9 HS	447.3-951.2 nm	CCD camera 512 x 512 px Spectral resolution 10 nm Spatial resolution 1 nm Region of interest threshold 16 x 16 px	Entire chicken carcases	Detection of tumours (67.5-80% accurate)	Nakariyakul and Casasent (2009)

Ocean Optics USB4000-VIS-NIR	450-910 nm	Integration time 5 ms Scan distance 5 cm 3 scans per sample	Intact pork	Moisture (r^2_{val} = 0.79, $RMSE_{val}$ = 0.77) Fat (r^2_{val} = 0.77, $RMSE_{val}$ = 0.09) Protein (r^2_{val} = 0.76, $RMSE_{val}$ = 0.41) pH (r^2_{val} = 0.82, $RMSE_{val}$ = 0.10)	Liao et al. (2010)
Specim ImSpector N17E HS	890-1750 nm	Spectral resolution 6 nm Spatial resolution 0.58 mm/pixel, 320 x 256 px 2 scans per sample	Intact lamb	Moisture (r^2_{val} = 0.84–0.88, SE_{val} = 0.51–0.57, RPD_{val} = 2.35–2.63, r^2_{CV} = 0.86–0.91, SE_{CV} = 0.42–0.52, RPD_{CV} = 2.62–3.24) Protein (r^2_{val} = 0.63, SE_{val} = 0.34, RPD_{val} = 1.71, R^2_{CV} = 0.67, SE_{CV} = 0.33, RPD_{CV} = 1.73) Fat (r^2_{val} = 0.87–0.98, SE_{val} = 0.35–0.40, RPD_{val} = 3.20–3.66, r^2_{CV} = 0.90–0.91, SE_{CV} = 0.35–0.37, RPD_{CV} = 3.70–3.91) pH (r^2_{CV} = 0.65, $RMSE_{CV}$ = 0.09, RPD = 1.76)	Kamruzzaman et al. (2012a, b)
Specim ImSpector N17E HS	910-1700 nm	Spectral resolution 6 nm Spatial resolution 0.58 mm/pixel, 320 x 256 (0.06 mm/px) Sampling interval 3.34 nm	Minced lamb and pork	Detection of pork in lamb meat (r^2_{CV} = 0.98–0.99, $RMSE_{CV}$ = 1.42–1.45, RPD = 8.04–8.21)	Kamruzzaman et al. (2013)
Headwall Photonics HS spectrograph	940-1650 nm	Spectral resolution 4.8 nm	Intact pork	Pork marbling score (r_{val} = 0.77–0.90, $RMSE_{val}$ = 0.52–0.65)	Huang et al. (2014b)
Specim ImSpector N17E HS	910-1700 nm	Spectral resolution 6 nm	Minced lamb and offal	Detection of offal in minced lamb (r^2_{CV} = 0.97, $RMSE_{CV}$ = 1.80–1.84)	Kamruzzaman et al. (2014)

		Spatial resolution 0.58 mm/pixel, 320 x 256 (0.06 mm/px) Sampling interval 3.34 nm			
Specim ImSpector V10E HS	400-1000 nm	Spectral resolution 2.8 nm Spatial resolution 1004 x 1002 px Sampling interval 1.58 nm 1 scan per sample	Intact carp	Detection of total viable counts of bacteria (r^2_{CV} = 0.90–0.93, $RMSE_{CV}$ = 0.48–0.59, r^2_{val} = 0.90–0.93, $RMSE_{val}$ = 0.49–0.57, RPD = 3.13–3.89)	Cheng and Sun (2015)
Texas Instruments MC1002PF camera + V10C ImSpector HS	400-1000 nm	Sampling interval 5 nm Exposure time 9.4 ms	Minced beef and horsemeat	Detection of horsemeat in beef (r^2_{CV} = 0.99, SE_{CV} = 1.56, r^2_{val} = 0.98, SE_{val} = 2.23)	Kamruzzaman et al. (2015)
VideometerLab HS	405-970 nm	18 NIR wavelengths used to acquire multispectral images	Minced beef and pork	Detection of pork and beef (95% accurate) Detection of adulterants in pork and beef (89% accurate)	Ropodi et al. (2015)
Innopharma Labs NIR spectrometer	1515-2100 nm	Scan interval 5 nm Speed 100-210 rpm Scan distance 1-4 cm 1 scan (5 regions) per sample	Minced beef	Moisture (r^2_{val} = 0.94–0.98, $RMSE_{val}$ = 2.75–4.62) Fat (r^2_{val} = 0.95–0.99, $RMSE_{val}$ = 2.79–5.67) Protein (r^2_{val} = 0.90–0.95, $RMSE_{val}$ = 1.56–2.28) Ash (r^2_{val} = 0.95–0.99, $RMSE_{val}$ = 0.03–0.06)	Dixit et al. (2016)
Texas Instruments MC1002PF camera +	400-1000 nm	Sampling interval 5 nm Exposure time 9.4 ms	Minced beef and chicken	Detection of chicken in beef (r^2_{val} = 0.96, $RMSE_{val}$ = 1.79–3.18, RPD = 4.81–8.54)	Kamruzzaman et al. (2016)

Specim ImSpector V10C HS					
Xenics XEVA 1.7-320 + VIS InGaAs HS camera	550-1700 nm	Spectral resolution 5 nm Spatial resolution 320 x 150 px Write speed 25 frames/s Scan distance 37 cm	Intact beef	IMF (r^2_{val} = 0.48, RMSE = 0.54, bias = -0.57) pH (r^2_{val} = 0.20, RMSE = 0.08, bias = 0.64)	Craigie et al. (2017)
Allied-Vision Technology GX1660 CMOS HS	400-1000 nm	Spectral resolution 8 nm	Intact beef	Marbling score of beef (r_{val} = 0.95, SE_{val} = 0.3)	Aredo et al. (2017)
ASD LabSpec 4 Hi-Res NIR	350-2500 nm	Spectral resolution 3-10 nm Sampling interval 1.4-2 nm Scan time 7 s	Intact beef	IMF (r^2_{CV} = 0.91, SE_{CV} = 1.02, r^2_{val} = 0.89, SE_{val} = 1.12, bias = 0.02, RPD = 2.97) pH (r^2_{CV} = 0.87, SE_{CV} = 0.17, r^2_{val} = 0.86, SE_{val} = 0.18, bias = 0.01, RPD = 2.72)	Dixit et al. (2021)
Hyperspec EVNIR + VIS InGaAs HS camera	550-1700 nm	Spectral resolution 5 nm Spatial resolution 320 x 150 px Write speed 25 frames/s Scan distance 37 cm	Intact beef	IMF (r^2_{CV} = 0.92, SE_{CV} = 0.98, r^2_{val} = 0.90, SE_{val} = 1.06, bias = 0.85, RPD = 3.13) pH (r^2_{CV} = 0.91, SE_{CV} = 0.15, r^2_{val} = 0.89, SE_{val} = 0.16, bias = 0.91, RPD = 3.08)	Dixit et al. (2021)
Multi-camera system (VIS and VIS-NIRS HS)	470-630 nm and 670-950 nm	Spectral resolution 10 nm	Intact beef	IMF (r^2_{CV} = 0.29–0.82, SE_{CV} = 1.39–2.94, r^2_{val} = 0.44–0.72, SE_{val} = 1.76–2.47, bias = -0.03—0.26, RPD = 1.34–1.88)	Dixit et al. (2021)

	Spatial resolution 512 x 256 px (Vis) and 409 x 216 px (Vis-NIR) Write speed 7 frames/s Scan distance 20 cm	pH (r^2_{CV} = 0.75–0.84, SE_{CV} = 0.20–0.25, r^2_{val} = 0.75–0.77, SE_{val} = 0.23–0.24, bias = -0.03–0.04, RPD = 1.98–2.07)

r^2_{CV} – coefficient of determination of cross-validation; r^2_{val} – coefficient of determination on the validation dataset; $RMSE_{CV}$ – root mean square error of cross-validation; $RMSE_{val}$ – root mean square error on the set-aside validation dataset; SE_{CV} - standard error of cross-validation; SE_{val} – standard error on the set-aside validation dataset; RPD – residual prediction deviation; IMF – intramuscular fat.

Note: Fat and IMF are used interchangeably as they are dependent on their mention in the original study.

1.6. Conclusions and Hypothesis

1.6.1. Conclusions

Following a comprehensive review of the literature available for spectroscopic devices in the meat industry, several paucities were identified in the literature. These include, but are not limited to:

- Use of alternative modelling approaches other than partial least squares (i.e., PLSR, PLS-DA) for both prediction of chemical composition and discrimination.
- Use of handheld NIRS devices that connect to smartphones via Bluetooth for the prediction of meat quality and chemical composition.
- Use of the above devices and modelling approaches for the discrimination of meat based on feeding regime (grass-fed and grain-fed).
- Use of the above devices and modelling approaches for the prediction of DOF of a high-grain diet and visual marbling score using NIRS or Raman spectroscopy.
- Use of novel multi-sensory platforms and different modelling approaches for the identification of organs using line scanning tools in the abattoir.
- Use of novel multi-sensory platforms and modelling approaches for the discrimination of organs based on presence of defects or diseases in the abattoir, and whether diseased tissue differs in spectral signature.

To explore these paucities, the present thesis comprises the following experimental chapters:

Chapter 2: Use of a palm-sized NIR spectrometer (NIRvascan Nano) with Bluetooth connection to a smartphone application (NIRScan Nano) to evaluate chemical composition (moisture, protein, IMF and pH) of beef and lamb cuts from retail outlets. The resulting absorbance spectra were compared with those from a larger benchtop NIR spectrometer employing a scanning probe (ASD AgriSpec) for their ability to predict chemical composition and meat quality measured as

per previous methods (Coombs et al., 2017; Helrich, 1990; Mehnaz et al., 2019; MLA, 2017a). Chemical composition results at different sample presentations are presented in Chapter 2.

Chapter 3: Use of a palm-sized NIR spectrometer (NIRvascan) and Raman spectrometer (Bruker Bravo) to discriminate grass-fed and grain-fed status of beef cuts (including days on grain feed). The absorbance spectra from both devices were compared with each other for their predictions of feeding regime, MSA marbling score (as per MLA, 2017a), and DOF.

Chapter 4: Use of a prototype multi-sensory platform encompassing two HS sensors (visible and short-wave infrared) to identify parenchymatous organs (hearts, kidneys, livers, and lungs) from sheep and cattle from their spectral signatures.

Chapter 5: Use of a prototype multi-sensory platform encompassing two HS sensors (visible and short-wave infrared) to discriminate organs with defects and diseases based on acceptance or rejection from the abattoir sortation process, later confirmed by veterinary pathology. This technology was trialled for the purpose of automating the organ sortation process in the abattoir.

1.6.2. Hypothesis

The hypothesis of the present thesis was that novel spectroscopic devices can be used in the meat industry for prediction of chemical composition and discrimination of feeding regime at a consumer and retail levels, organ identification for the purposes of automatic organ sortation by type, and detection of defects and disease in livestock organs. The objective was to evaluate these novel devices and develop data analysis procedures to provide reliable, cost effective, and accurate alternatives within the meat industry as alternatives to the current methods of subjective carcase grading, external chemical analysis, trust in product labelling, and manual sortation and inspection of offal.

2. Comparison of smartphone and lab-grade NIR spectrometers to measure chemical composition of lamb and beef

Published in Animal Production Science 61: 1723-1733.

Abstract. Near-infrared reflectance spectroscopy (NIRS) has been extensively investigated for non-destructive and rapid determination of pH and chemical composition of meat including water, crude protein, intramuscular fat (IMF) and stable isotopes. Smaller, cheaper NIRS sensors that connect to a smartphone could enhance the accessibility and uptake of this technology by consumers. However, the limited wavelength range of these sensors could restrict the accuracy of predictions compared to benchtop laboratory NIRS models. This chapter aimed to compare the precision and accuracy metrics of predicting pH, water, crude protein and IMF of three sample presentations and two sensors. Fresh intact (FI) store-bought beef and lamb steak samples (n = 43) were ground and freeze-dried (FD), and then oven-dried to create freeze-dried oven-dried (FDOD) samples. All three forms of sample presentation (FI, FD, FDOD) were scanned using the smartphone and benchtop NIRS sensors. The IMF was the best predicted trait in FD and FDOD forms by the smartphone NIRS ($R^2 > 0.75$; RPD > 1.40), with limited differences between the two sensors. However, predictions on FI meat were poorer for all traits regardless of the NIRS scanner used ($R^2 \leq 0.67$; RPD ≤ 1.58) and not suitable for use in research or industry. The smartphone NIRS sensor showed accuracy and precision comparable to benchtop NIRS to predict meat composition. However, these preliminary results found that neither of the two sensors reliably predicted quality attributes for industry or consumer applications. Miniaturised NIRS sensors connected to smartphones could provide a practical solution to measure some meat quality attributes such as IMF, but the accuracy depends on sample presentation.

2.1. Introduction

Research has shown that consumers worldwide are willing to pay a premium for meat with high eating and nutritional quality, particularly in developed countries (Pethick et al., 2018; Polkinghorne & Thompson, 2010). Conventional measures of meat quality include, but are not limited to, colour of muscle and fat, pH, and marbling (AUS-MEAT, 2018) and chemical composition (Prieto et al., 2017). The issue with conventional meat grading procedures is that they are subjective, with the drive for on-line instrumental assessment providing the potential for a global system of objective carcase classification (Polkinghorne & Thompson, 2010). Chemical composition (fat, protein and water contents) is accurate and objective but measured in laboratories using wet chemistry methods which are invasive, destructive, expensive and time consuming due to the requirements of sample preparation and instrumentation (Elmasry et al., 2012a; Teixeira dos Santos et al., 2013). Therefore, both methods of measuring meat quality are not suitable for use by consumers or retailers and there is a need to develop consumer grade technologies to predict meat quality.

As a result, several scanning technologies have been developed to measure such quality traits in an accurate, objective and non-destructive manner. However, these are limited by their large size, lack of portability and high installation costs (Scholz et al., 2015). Near-infrared reflectance spectroscopy (NIRS) has been used extensively in research to predict eating and nutritional quality of meat and meat products (Dixit et al., 2017; Prieto et al., 2017; Qiao et al., 2015). However, prediction of meat quality and chemical composition by NIRS has been done to varied levels of precision and accuracy, as previously reviewed (Prevolnik et al., 2004; Prieto et al., 2009; Prieto et al., 2017).

The development of smaller, cheaper, and more portable NIRS scanners has provided the opportunity for multiple scans of the surface of heterogeneous fresh meat samples that could increase its accuracy of predictions (Teixeira dos Santos et al., 2013; Dixit et al., 2020). Despite these advances, research into the use of miniaturised, consumer grade NIR spectrometers for prediction of quality traits has not been extensive. Accuracy and precision have been generally inferior with miniaturised

NIR spectrometers due to the shorter wavelength range (900–1700 nm) compared to benchtop lab-grade NIRS (LG-NIRS) that include the visible spectrum and further into the NIR (350–2500 nm) (vis-NIRS; De Marchi, 2013). Recent on-line predictions using the latter device on FI meat have yielded promising results, particularly for chemical analyses (Cozzolino & Murray, 2002; Prieto et al., 2011; Pullanagari et al., 2015). Recent studies using the NIRS with shorter wavelength have found results comparable to benchtop NIRS to measure soil quality (Tang et al., 2020) and intramuscular fat (IMF) in beef (Pham et al., 2018; Sun et al., 2011). However, Dixit et al. (2020) found decreased accuracy and precision to predict IMF with miniaturised compared to larger LG-NIRS systems due to handheld NIRS being more subject to day-to-day variation.

A recent study on lamb reported that loin and topside IMF was predicted with moderate precision and accuracy (R^2 = 0.38–0.56; RMSE = 0.79–0.97) using a miniaturised vis-NIRS (350–2000 nm) in the abattoir 24 h post-mortem (Fowler et al., 2020). Similarly, Pham et al. (2018) found promising results comparing smartphone NIRS and LG-NIRS based on chemical IMF determination of beef (R^2 = 0.75, RMSE = 2.10) but did not include predictions of pH, water or CP. Conversely, the follow-up study on lamb IMF prediction yielded a much lower R^2 and higher RMSE for the handheld NIRS (0.27 and 1.28, respectively) compared to benchtop NIRS (0.76–0.83 and 0.34–0.41, respectively) (Dixit et al., 2020). Despite the studies of Pham et al. (2018) and Dixit et al. (2020) using smartphone NIRS, no traits other than IMF were predicted and only the ground and FD sample presentation was analysed.

The most precise predictions of chemical measures (r >0.96) have occurred on minced (Barlocco et al., 2006; Lanza, 1983; Sun et al., 2011) or dried and ground meat (Prieto et al., 2017; Su et al., 2014). However, analyses on ground or minced meat are not applicable for consumers purchasing intact meat, and not practical for the meat processing industry due to their destructive nature and increased sample processing requirement (Dixit et al., 2017; Prieto et al., 2009; Prieto et al., 2017; Viljoen et al., 2007). This has prompted suggestions that less accurate predictions on fresh

meat in the abattoir would have more benefit to the industry (Viljoen et al., 2007) and to consumers (Dixit et al., 2017).

Additionally, NIRS has been successfully used in animal production to predict the concentration of stable isotopes and then the type of forages consumed by grazing animals (Dixon & Coates, 2009). Scarce research on the use of stable isotopes (^{13}C and ^{15}N) demonstrated that these can be used to determine meat origin, nutritional background and thereby authentication of meat products (Bahar et al., 2009), which can be particularly useful at retail level where food fraud may occur, e.g., grass-fed and grain-fed meat (Prache et al., 2020). Prediction of meat stable isotope content by NIRS has not yet been investigated despite the many potential applications.

The aim of the present study is therefore to evaluate the potential of a miniaturised smartphone NIRS sensor and compare it to a conventional LG-NIRS to measure pH, water, IMF, CP, stable isotope ratios ($\delta^{13}C$ and $\delta^{15}N$), total C and total N on three different sample presentations (intact; FD and ground meat; freeze-dried, oven-dried and ground meat) of a small selection of store-bought meat cuts. It was hypothesised that prediction of laboratory measurements using the smartphone NIRS sensor will be comparable to those made by the LG-NIRS. In addition, it was hypothesised that chemical measurements are likely to be better predicted from scans on freeze-dried and ground meat (± oven-drying) because this homogenises the samples. Positive results from the present study would encourage further research with a larger sample size using the smartphone NIRS sensor.

2.2. Materials and methods

2.2.1. Sample collection and preparation

Forty-three retail cut samples of beef (n = 27) and lamb (n = 16) (Table 5) were purchased from butchers and supermarkets in Sydney, Australia. Sample collection was based to represent a broad range of marbling, muscle location, breed, colour, freshness and price. Subjective scoring for marbling and colour of muscle and fat was performed according to Australian grading standards (AUS-MEAT, 2018). Three different meat presentation types were used for scanning: (1) fresh intact (FI) upon opening of retail packaging; (2) freeze-dried (FD) following snap freezing (-80 °C); and (3) freeze-drying (72 h) and subsequently oven-dried at 65 °C overnight to obtain freeze-dried-oven-dried (FDOD).

Table 5. Retail meat samples used in this experiment.

Species	Common name	Muscle	Samples (*n*)
Beef	Cattleman's cutlet (Rib eye bone in)	LTL	6
Beef	Angus beef tenderloin	PM	4
Beef	Angus short ribs boneless	SV	7
Beef	Wagyu tenderloin fillet steak	PM	4
Beef	Angus scotch fillet (150 days grain-fed) — rib eye boneless	LTL	4
Beef	Australian Tomahawk steak (grain-fed) — rib eye bone in	LTL	2
Lamb	Cutlet	LTL	16

LTL - *m. longissimus thoracis et lumborum*; PM - *m. psoas major*; SV - *m. serratus ventralis*

2.2.2. Spectroscopy measurements

Six scans per cut of beef and three scans per cut of lamb were done using NIRS and instrumental colour scanners. Scans were at evenly spaced intervals along the lean (avoiding subcutaneous fat and connective tissue) surface of FI meat. The scanners used included the smartphone sensor (NIRvascan, Allied Scientific Pro., Gatineau, Quebec, Canada; 900–1700 nm; 6000:1 signal to noise ratio; 2 x 1 W tungsten halogen lamp; spectral bandwidth 3.5 nm per reading);

an LG-NIRS instrument (AgriSpec, ASD Inc., Boulder, CO, USA; 350–2500 nm; 14500:1 signal to noise ratio; 100 W halogen lamp; spectral bandwidth 1 nm per reading); and a colorimeter (CR-400 Chroma Meter, Konica Minolta Sensing Americas, Inc., Ramsey, NJ, USA) with aperture size of 8 mm, illuminant C and 2° standard observer calibrated on a white tile to evaluate instrumental colour (L*a*b*). The smartphone NIRS sensor (NIRvascan) was connected via Bluetooth to the NIRScan Nano smartphone application (KS Technologies, Colorado Springs, CO, USA) (Fig. 2). For both NIRS and colour determination, a glass microscopic slide (1 mm thickness) was used to separate the spectrometers and meat surfaces in order to prevent contamination of the apertures.

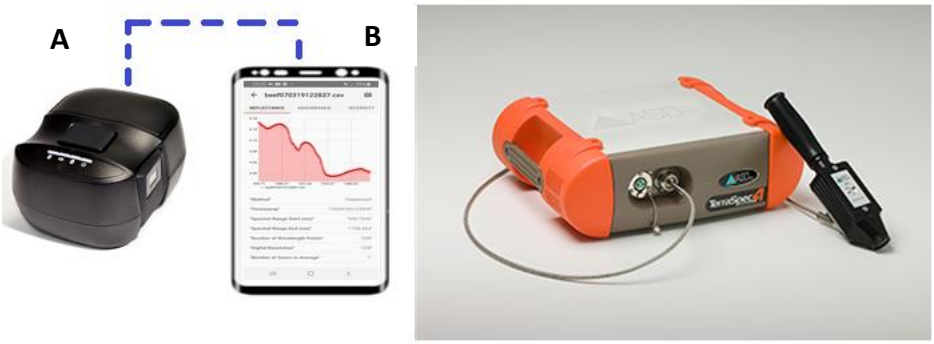

Fig. 2. Comparison of the two NIR spectrometers used to scan meat samples. **A)** Smartphone NIRS sensor: NIRvascan (Allied Scientific Pro., Gatineau, Canada) weighing 136 g and size 82.2 x 66 x 45 mm; **B)** Lab-grade NIRS (LG-NIRS) system: ASD AgriSpec (ASD Inc., Boulder, USA) weighing 5.44 kg and size 127 x 368 x 293 mm.

2.2.3. pH and temperature measurements

Muscle pH and temperature were measured by triplicate (beef) and duplicate (lamb) for each cut using a conventional pH-temperature probe (Hanna Instruments, Woonsocket, RI, USA) calibrated using buffer solutions at pH = 4.00 and pH = 7.00 as per MLA (2017a) requirements, and then averaged into one value for analysis.

2.2.4. Water determination

The samples were prepared for chemical analysis after scanning of intact cut and measurement of pH, temperature, colour and marbling. Sample preparation for water, CP, stable isotopes, C, N, and chemical IMF analyses was done on thirds of the steaks for each respective measurement following removal of subcutaneous fat, cut into multiple cubes of approximately 1 cm^3 and then snap-frozen at –80 °C and freeze-dried for 72 h in batches. The difference in weight before and after freeze-drying was used to calculate water content (%water = wet weight – dry weight × 100; Coombs et al., 2017) with the remainder considered as dry matter. Results were expressed as the average of technical duplicates.

2.2.5. Intramuscular fat determination

Intramuscular fat (IMF) was determined chemically on FD meat ground in an analytical cutting blade mill (IKA A11, IKA Works Inc., Rawang, Selangor, Malaysia) and then analysed using a modified Soxhlet method for chemical fat determination in lean muscle based on petroleum ether (Sigma-Aldrich, Castle Hill, NSW, Australia) extraction methods 960.39 and 945.16a (Helrich, 1990; Perry et al., 2001). Total ether extractable fat was determined using 1—2 g of lean FD muscle within an extraction thimble following 2 x 2 h flushes (Dow et al., 2011). Results were expressed as a single measure of the weight of the fat extracted by the Soxhlet apparatus following drying of the petroleum ether at 130 °C for 30 min divided by the total dried meat used (weight of dry extract / total dry weight) multiplied by the dry matter (DM) percentage (1 – %DM) to be expressed as the percentage of IMF of fresh wet meat.

2.2.6. Crude protein determination

Sub-samples (1.4-1.7 mg) of FDOD meat were taken for determination of total proportion of nitrogen (%N) using an isotope ratio mass spectrometer (Delta V Advantage, ThermoFisher Scientific, Bremen, Bremen, Germany) (Mehnaz et al., 2018). Total %N was then multiplied by 6.25 to determine CP content of meat per g of DM as per method 981.10 (Helrich, 1990). Results were expressed as a percentage of total fresh wet meat and expressed as the average of technical triplicates:

$$\%CPwet = \%CPdry \times \left(1 - \frac{\%DM}{100}\right) (1)$$

2.2.7. Stable isotope ratios and total C and N determination

Stable isotopes were determined using FDOD samples (0.75—0.95 mg; section 2.6) packed into tin capsules for dual carbon and nitrogen analysis as described by Bahar et al. (2009) and analysed using an isotope ratio mass spectrometer (section 2.6; Mehnaz et al., 2018). Delta notation (δ) was used to express the isotope ratios of ^{13}C and ^{15}N in the sample compared with reference values (Vienna Pee Dee Belemnite for ^{13}C, air-N_2 for ^{15}N) as per the equation:

$$\delta = \left(\left(\frac{Rsample}{Rreference}\right) - 1\right) \times 10^3$$

Where R is the ratio of heavy to light stable isotope and δ is the difference per thousand (‰). Total carbon and nitrogen contents in meat samples were also calculated and expressed as percentages. Results are expressed as the average of technical triplicates.

2.2.8. Dried and ground NIR spectroscopy

Following freeze-drying and grinding for determination of water and IMF contents (see sections 2.2.4 and 2.2.5), FD samples (n = 35) were re-scanned (three scans for lamb, six scans for beef) in Petri dishes covered by a glass microscope slide using both the smartphone NIRS and LG-NIRS

(see section 2.2.2). The non-scanned samples (n = 8) were not considered due to insufficient sample present following IMF analysis. Subsequently, following oven-drying for CP and stable isotopes determination (overnight at 65 °C; see sections 2.2.6 and 2.2.7), FDOD samples (n = 43) were re-scanned using the same procedure. These scans were performed to compare different sample presentations (Cozzolino et al., 2000; Cozzolino & Murray, 2002).

2.2.9. Prediction models and statistical analysis

Spectral data generated from the smartphone NIRS and LG-NIRS were averaged per sample (n = 43 for FI and FDOD, n = 35 for FD) and then inverse log transformed (log(1/R)) as described by Lanza (1983) prior to graphical presentation and development of Cubist prediction models for presentations FI (n = 210), FD (n = 183) and FDOD (n = 210). Cubist regression modelling was used as it has not been used in meat spectroscopy previously; instead, it has been used successfully in soil spectroscopy studies where it has often outperformed partial least squares regression (Minasny and McBratney, 2008; Peng et al., 2019; Tang et al., 2020). Smartphone NIRS data were trimmed between wavelength 950 and 1600 nm to eliminate spectral noise and data from both NIRS instruments were pre-processed using a Savitzky-Golay smoothing filter and standard normal variate (SNV). Prediction models were developed from the spectra of 75% of all samples (randomly generated training data) using a *Cubist* decision trees algorithm (Kuhn & Quinlan, 2020) in R (R Core Team, 2020). The remaining 25% of samples were used as the independent dataset for validation of the prediction models (test data). For each trait (pH, water, IMF, CP, $\delta^{13}C$, $\delta^{15}N$, total C, total N), 50 random bootstrapped samples were drawn from the training dataset in order to create 50 independent models. The performance of the models was calculated using the averaged 50 predictions with the validation dataset NIRS scans against the measured traits in the validation dataset (chemical data).

Laboratory measurements (pH, water, chemical IMF, CP, stable isotopes and total C and N) for each cut of meat were compared with the corresponding mean NIRS (smartphone and LG-NIRS)

values predicted by Cubist models. The goodness-of-fit, or precision, of the models was assessed on the test datasets using coefficients of determination (R^2) between the two NIRS prediction models and laboratory-measured values (ground truth) (Tedeschi, 2006). The RMSE and positive or negative bias of predictions were also obtained in R for model error assessment, or accuracy (Tedeschi, 2006). The RPD value was calculated as the ratio of standard deviation to standard error of the validation dataset (Williams, 2014). Goodness-of-fit was assessed by high R^2 and low RMSE and absolute value of bias; whereas applicability was assessed by RPD (Camacho-Tamayo et al., 2014; De Marchi, 2013; Williams, 2001; Williams et al., 2017). Williams (2001) stated that an R^2 above 0.83 was required for applications in research. However, for the purpose of meat quality predictions in this study, RPD > 2 was considered good for calibration and RPD > 3 good for analysis (Barlocco et al., 2006; Williams, 2014). Prediction models were considered as very high ($R^2 > 0.9$) high ($0.7 < R^2 \leq 0.9$), moderate ($0.5 < R^2 \leq 0.7$), low ($0.3 < R^2 \leq 0.5$), and very low ($R^2 \leq 0.3$) in precision (Williams et al., 2019).

2.3. Results

2.3.1. Physical and chemical measurements

Descriptive statistics for physical and chemical measurements are in Table 6. The largest variability and range between samples was observed for colour, marbling, and IMF, and lowest for water content and pH (Table 6). There were 10 samples with mean pH above 5.7 considered as dark cutters (data not shown). Regarding stable isotopes, lamb had greater ^{13}C depletion than beef (–24.32 vs. –21.49) and less increase in ^{15}N (5.98 vs. 8.01) compared with reference values (Table 6). A correlation matrix of reference measurements is provided in Table 7.

Table 6. Descriptive statistics for meat quality traits (n = 43) analysed and predicted using smartphone and laboratory-grade NIRS sensors.

Trait	Mean	SEM	SD	Min.	Max.	Median	CV (%)
pH	5.66	0.53	0.15	5.45	6.14	5.60	2.65
Temperature (°C)	17.50	0.18	1.89	9.80	19.80	17.70	10.80
Water (%)	71.10	0.31	2.87	65.77	78.77	71.16	4.04
IMF (% fresh)	6.24	0.52	3.38	0.83	16.19	5.94	54.17
CP (% fresh)	21.73	0.29	2.48	14.33	27.55	21.67	12.40
$\delta^{13}C$ (‰)	–22.54	0.23	2.18	–27.87	–17.80	–22.74	–9.65
$\delta^{15}N$ (‰)	7.25	0.14	1.28	5.43	10.07	6.90	17.62
Total C (%DM)	15.00	0.22	2.00	11.37	20.96	14.86	13.31
Total N (%DM)	3.43	0.05	0.42	2.27	4.88	3.47	12.41
CIE colour							
L*	36.28	0.21	3.06	26.98	43.05	36.47	8.43
a*	9.44	0.14	2.06	3.46	15.53	9.74	21.82
b*	3.75	0.08	1.17	0.47	7.85	3.73	31.20
AUS-MEAT							
Marbling	3.49	0.26	1.68	1	6	3	48.14
Meat colour	2.21	0.24	1.55	1B	7	2	70.14
Fat colour	2.14	0.24	1.55	0	5	2	72.43

Table 7. Pearson correlation matrix between meat quality traits and chemical composition of 43 beef and lamb samples.[1]

	pH	Water	IMF	CP	δ13C	δ15N	Total C
pH		–0.12 (0.43)	0.24 (0.11)	–0.18 (0.26)	0.19 (0.23)	0.56 (<0.001)	0.37 (0.01)
Water			–0.80 (<0.001)	–0.04 (0.79)	–0.01 (0.96)	–0.28 (0.07)	–0.76 (<0.001)
IMF				0.22 (0.16)	–0.05 (0.74)	0.22 (0.16)	0.74 (<0.001)
CP					0.11 (0.47)	–0.14 (0.36)	–0.07 (0.66)
δ13C						0.57 (<0.001)	–0.03 (0.83)
δ15N							0.40 (0.01)

[1] Pearson's correlation coefficients (r) between corresponding meat quality traits and *P*-values are in brackets.

2.3.2. Spectral data

The averaged spectra for the FI (Fig. 3), FD (Fig. 4) and FDOD (Fig. 5) samples were calculated for smartphone NIRS and LG-NIRS. Common peaks for both devices were identified at 950 nm, 1220 nm, and 1462 nm for FI samples (Fig. 3). These peaks corresponded to -OH (water), -CH_2 (fat), and -OH (water), respectively (Cozzolino et al., 2002; Osborne et al., 1993). Further peaks outside the 900–1700 nm spectrum occurred at 434 nm, 550 nm, 583 nm and 1898 nm which were not captured by NIRvascan (Fig. 2). The latter of these can be related to -NH_2 (CP) (Osborne et al., 1993). Freeze-dried meat had peaks at 1191 nm, 1516 nm and 1700 nm (Fig. 4), with similar peaks at 1188 nm, 1488 nm and 1700 nm occurring in FDOD meat (Fig. 5).

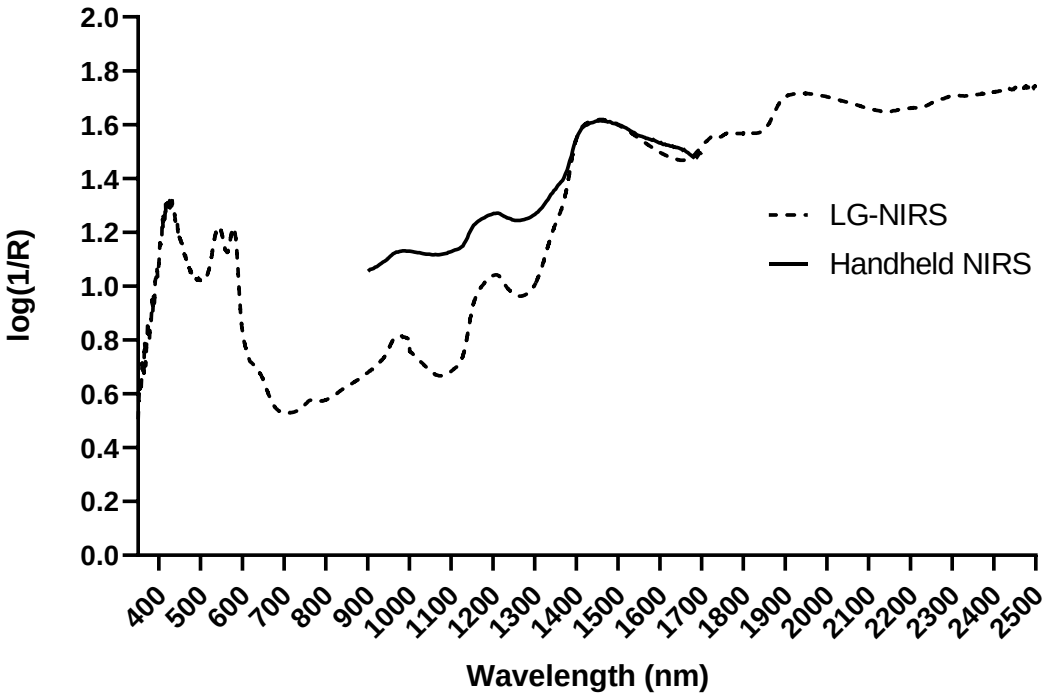

Fig. 3. Mean inverse log reflectance spectra (n = 210) generated by the handheld smartphone NIRS (NIRvascan, SEM = 0.12–0.16) and lab-grade NIRS (LG-NIRS, ASD AgriSpec, SEM for 350–2500 nm = 0.09–0.17; SEM for 900–1700 nm = 0.10–0.17) on 43 fresh intact (FI) retail meat samples.

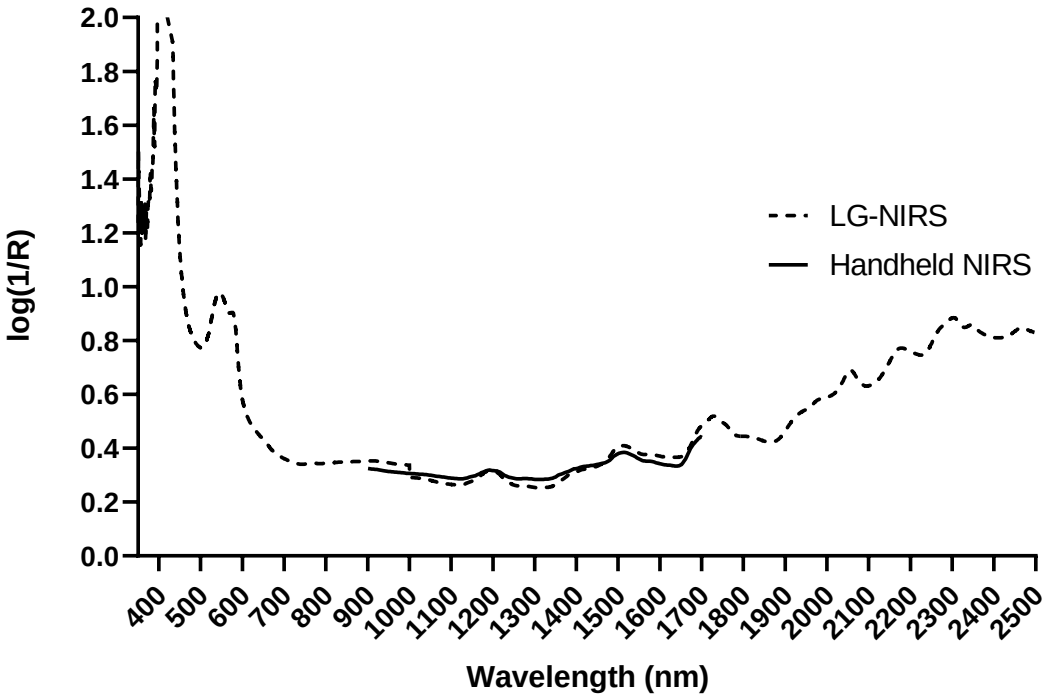

Fig. 4. Mean inverse log reflectance spectra (n = 183) generated by the handheld smartphone NIRS (NIRvascan, SEM = 0.08–0.09) and lab-grade NIRS (LG-NIRS, ASD AgriSpec, SEM for 350–2500 nm = 0.08–0.13; SEM for 900–1700 nm = 0.09–0.11) on 35 freeze-dried (FD) meat samples.

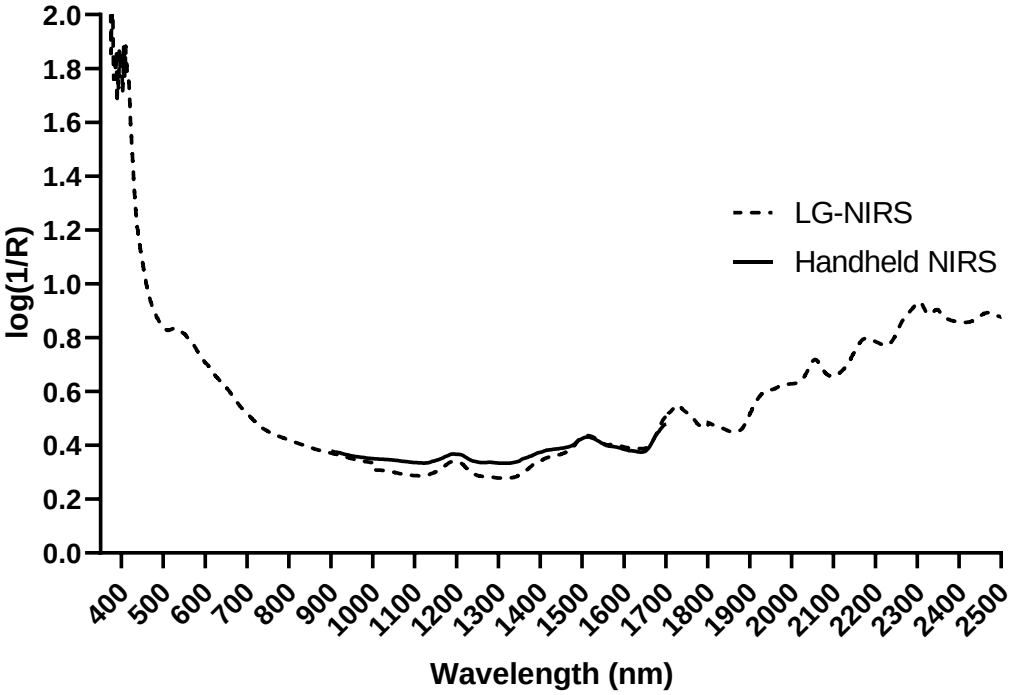

Fig. 5. Mean inverse log reflectance spectra (n = 210) generated by the handheld smartphone NIRS (NIRvascan, SEM = 0.08–0.09) and lab-grade NIRS (LG-NIRS, ASD AgriSpec, SEM for 350–2500 nm = 0.07–0.12; SEM for 900–1700 nm = 0.07–0.08) on 43 freeze-dried and oven-dried (FDOD) meat samples.

2.3.3. Prediction models

The average coefficient of determination for each device across all sample presentations and traits predicted was 0.53 and 0.56 for smartphone NIRS and LG-NIRS, respectively (Table 8). Similarly, RMSE (1.31 and 1.34) and RPD (1.55 and 1.54) were similar between both devices, although the smartphone NIRS showed lower mean bias than LG-NIRS (0.09 and 0.20, respectively). On average across sample presentations and devices, water and IMF concentrations were predicted with the highest accuracies (R^2 = 0.67 and 0.70, respectively). Total C and pH showed the lowest accuracy with average R^2 of 0.39 and 0.38, respectively (Table 8). However, differences amongst devices and sample presentations existed.

For example, the smartphone NIRS showed a low accuracy to predict pH, $\delta^{13}C$, $\delta^{15}N$, total C and total N from FI meat (R^2 = 0.37 to 0.44) and moderate to predict water content, whereas it predicted CP poorly (Table 4). Differences between the two NIR spectrometers were negligible for the majority of measurements except for IMF, CP and $\delta^{15}N$ for which LG-NIRS performed superiorly (Table 8).

Most predictions, regardless of the NIR spectrometer used, were more accurate and precise on FD and FDOD (average $R^2 > 0.58$) compared to FI meat ($R^2 < 0.49$) except for pH and total C, though both traits were better predicted in FDOD form by the LG-NIRS (Table 8). For FD meat, the smartphone NIRS was superior in predicting CP and $\delta^{13}C$. In contrast, LG-NIRS provided better predictions of $\delta^{15}N$. The FD treatment was particularly notable for its high precision to predict water, IMF and $\delta^{13}C$.

For FDOD meat, IMF and CP and showed higher accuracy and precision using smartphone NIRS; meanwhile LG-NIRS was more accurate predicting pH and total C.

Table 8. Goodness-of-fit statistics on the validation dataset (25%, n = 11) of various meat quality traits predicted using a smartphone NIRS (NIRvascan) and a laboratory-grade NIRS (ASD AgriSpec) against chemically measured meat quality traits.

Parameter	Smartphone NIRS				Laboratory-grade NIRS			
	r^2	RMSE	Bias	RPD	r^2	RMSE	Bias	RPD
Fresh intact (FI)								
pH	0.404	2.190	0.256	1.237	0.389	2.175	0.345	1.245
Water (%)	0.540	0.757	0.040	1.501	0.599	0.725	0.160	1.568
IMF (%)	0.534	2.725	0.392	1.420	0.664	2.454	2.454	1.577
CP (%)	0.253	3.096	0.564	1.157	0.509	2.583	0.385	1.387
δ13C (‰)	0.415	0.136	0.040	1.227	0.395	0.138	0.046	1.205
δ15N (‰)	0.394	0.911	–0.189	1.268	0.602	1.551	0.151	1.562
Total C (%)	0.442	0.344	0.042	1.259	0.397	0.345	0.053	1.256
Freeze-dried (FD)								
pH	0.256	2.441	–0.421	1.143	0.299	2.312	–0.312	1.207
Water (%)	0.702	0.645	–0.039	1.819	0.729	0.614	–0.058	1.910
IMF (%)	0.813	1.368	–0.278	2.312	0.812	1.347	–0.142	2.348
CP (%)	0.715	1.604	0.435	1.813	0.524	2.005	0.420	1.450
δ13C (‰)	0.826	0.068	0.007	2.351	0.699	0.090	0.027	1.765
δ15N (‰)	0.458	1.484	–0.112	1.338	0.721	1.055	–0.204	1.883
Total C (%)	0.255	0.391	–0.068	1.141	0.311	0.366	–0.047	1.218
Freeze-dried oven-dried (FDOD)								
pH	0.273	2.326	0.374	1.164	0.642	1.744	0.468	1.553
Water (%)	0.725	0.594	0.101	1.913	0.706	0.607	0.053	**1.873**
IMF (%)	0.779	1.844	0.328	2.099	0.572	2.586	–0.027	1.497
CP (%)	0.576	2.317	0.354	1.546	0.473	2.645	0.480	1.355
δ13C (‰)	0.544	0.115	0.031	1.451	0.608	0.105	0.022	1.587
δ15N (‰)	0.675	1.362	–0.011	1.780	0.705	1.374	–0.317	1.764
Total C (%)	0.278	0.371	0.061	1.168	0.638	0.279	0.077	1.552

2.4. Discussion

The present study tested the ability of a low-cost, handheld smartphone NIRS against the bulkier and more expensive, yet much wider wavelength range LG-NIRS. The ability of NIRS to detect chemical bonds (C-H, N-H, O-H corresponding to fat, protein and water, respectively) at 800–2500 nm makes it a viable solution for predicting chemical characteristics in meat and other food products (Osborne et al., 1993). Surprisingly, both NIRS devices performed similarly with moderate precision and accuracy to predict the traits of interest in the present study. The smartphone handheld NIRS sensor was developed for consumer use with the ability to instantly scan and upload data into the cloud via its Bluetooth connection to smartphones. Its shorter wavelength range (900–1700 nm) was not expected to significantly reduce its ability to predict chemical characteristics based on prior investigations (Coombs et al., 2019; Pham et al., 2018; Tang et al., 2020). This study was also novel and different from most previous work as it looked primarily at scanning FI meat with a view to deployment of such small handheld scanners for consumers, retailers, and in the meat processing industry measuring more parameters (pH, water, CP) and in multiple sample presentations compared to studies which only predicted IMF (Dixit et al., 2020; Fowler et al., 2020; Pham et al., 2018; Prieto et al., 2011). Model performance metrics generated were similar to other studies comparing portable NIRS vs. benchtop LG-NIRS (Tang et al., 2020; Williams et al., 2017). Prediction models for the smartphone NIRS in the present study showed promise compared to LG-NIRS in all three sample presentations (FI, FD, FDOD). However, neither NIRS sensor reached the accuracy needed for the prediction of analytical traits in FI format (all FI $R^2 < 0.70$; RPD < 2.0) desired for industry applications (Barlocco et al., 2006; Fowler et al., 2020; Pham et al., 2018; Williams, 2001).

The accuracy and precision (RMSE and R^2, respectively) of pH predictions of FI meat by the smartphone NIRS and LG-NIRS were low and similar to each other, and the precision and accuracy of smartphone NIRS predictions decreased in FD and FDOD formats. Previous studies have reported a range of accuracies to predict meat pH ranging from $R^2 = 0.31$ in FI beef using a near-infrared transmittance instrument at 850–1050 nm (De Marchi et al., 2013) to $R^2 = 0.62$ using benchtop LG-

NIRS at 350–1900 nm on fresh (De Marchi et al., 2013) and minced beef (Yang et al., 2010). The present study showed similar values to these (R^2 = 0.26–0.64), with the overall low precision in FI format (R^2 = 0.39–0.40) likely resulting from a low pH variation within samples, as mentioned in previous studies (De Marchi et al., 2013; Prieto et al., 2009). The smartphone NIRS prediction of pH was most accurate on FI meat, in agreement with De Marchi et al. (2013) who suggested that intact scanning yielded more precise predictions as this is the format pH is measured in the industry. The RPD in the present study for pH on FI meat (1.23–1.25) was poor but comparable to those found previously by De Marchi (2013) using vis-NIRS (1.08–1.52) on-line in the abattoir. The LG-NIRS in the present study predicted pH with the highest R^2 and RPD on FDOD meat compared to FI and FD. These results agree with earlier findings comparing portable and LG-NIRS using minced and intact beef, where ground meat showed higher precision and accuracy than fresh meat for prediction of pH (Cozzolino and Murray, 2002). However, the majority of studies using vis-NIRS (350–2500 nm) reported better predictions of pH compared to the present study (R^2 = 0.65–0.97) on FI beef and lamb (Andrés et al., 2008; Elmasry et al., 2012b; Kamruzzaman et al., 2012b; Qiao et al., 2015). As pH probes are non-destructive and objective, the only improvements from replacement of conventional measurement by non-invasive NIRS would be more rapid determination (scan + intervals ~ 10 seconds), real-time prediction and improvements in food safety by avoidance of cross-contamination between carcases (De Marchi, 2013). Further research could investigate non-contact of carcases by scanning at a distance (1.5–2.5 cm or longer), which may reduce the need for scan replication by increasing scanning area (Dixit et al., 2017; Osborne et al., 1993), while it could also allow for predictions of dark cutters (pH > 5.7; MLA, 2017a) and tenderness (Silva et al., 1999) to save time for retailers (Reis and Rosenvold, 2014).

Prediction of water content in the present study showed high accuracy and precision by both scanners on FD and FDOD meat (R^2 = 0.70–0.73; RMSE < 0.8), but it was lower on FI meat (R^2 = 0.54–0.60). Similarly, very high precision (R^2 > 0.84) was previously reported using NIRS on FI (Kamruzzaman et al., 2012b; Sun et al., 2011) and minced beef (Su et al., 2014; Sun et al., 2011; Yang et al., 2010), which suggested potential for use in quality control according to Williams et al. (2019). However, De

Marchi et al. (2013) found a significantly lower R^2 values of 0.01–0.12 on fresh ground meat. The ability to predict water content in dried and ground meat (FD and FDOD) was surprising, though this may have occurred due to high negative correlations with IMF and total C (–0.80 and –0.76, respectively).

The present study used samples with a wide range of chemical IMF content to represent consumer options in stores, denoted by visual marbling, as per large studies (Bindon, 2004; Savell et al., 1986). The mean IMF encountered was comparable to those found previously for beef (Lanza, 1983; Magalhães et al., 2018), lamb (Pullanagari et al., 2015) and mutton (Viljoen et al., 2007), although were higher than most prior studies on lamb (Dixit et al., 2020; Fowler et al., 2020; Kamruzzaman et al., 2012a; Karamichou et al., 2006) and beef (Cozzolino et al., 2000; Prieto et al., 2011). Marbling score and IMF are important measures of eating quality because of increased flavour, juiciness and liking (Frank et al., 2016; Su et al., 2014; Thompson, 2004). In agreement with our hypothesis, IMF was better predicted in FD compared to FI meat due to the homogenisation, lack of moisture and identical presentation for analysis. In a prior pilot study, Coombs et al. (2019) found predictions of subjective marbling score of LG-NIRS to be more precise than those from the NIRvascan used in the present study. The measurement of chemical IMF was made as it prompts the need for a more objective and global marbling grading system, such as NIRS-predicted based on accurate and precise predictions of chemical IMF content (Bindon, 2004; Polkinghorne & Thompson, 2010). The precision of the predictions of chemical IMF on FI meat in the present study were much higher than those reported for subjective marbling score by smartphone NIRS (Coombs et al., 2019) and LG-NIRS (Magalhães et al., 2018), which had $R^2 < 0.30$. Results on prediction of IMF of the present study are also similar to previous research with beef using vis-NIRS (350–1800 nm; $R^2 = 0.33–0.35$; Prieto et al., 2011) and lamb using handheld NIRS (900–1800 nm; $R^2 = 0.38–0.58$, Fowler et al., 2020). Other studies have yielded similar or better results compared to the present study (R^2 from 0.66–0.89) using on-line (portable NIRS) predictions of chemical IMF on FI meat (Kamruzzaman et al., 2012a; Pham et al., 2018; Pullanagari et al., 2015; Sun et al., 2011) or LG-NIRS on minced or FD meat (Cozzolino and Murray, 2002; Dixit et al., 2020; Yang et al., 2010). The predictions in the present study for FD meat showed

high precision and accuracy but were lower compared to other studies using minced beef without freeze-drying which reached values to allow applications for quality assurance as suggested by Williams et al. (2019) (R^2 > 0.92; Lanza, 1983; Su et al., 2014; Sun et al., 2011). Regarding the relatively high RPD for IMF predictions using FD meat in the present study, both NIRS prediction models seem applicable for "rough screening" (Williams et al., 2019), although increased sample size and optimisation of FI prediction models (poor RPD of 1.42–1.58) would need to occur for NIRS to be considered as an analytical or quality control tool in industry (Barlocco et al., 2006).

Crude protein concentrations in the present study were similar to those found previously for beef *m. longissimus lumborum* (Cozzolino et al., 2002) and lamb *m. semimembranosus* (Karamichou et al., 2006). Similar to IMF, CP was predicted with the highest precision on FD meat with both scanners, even though the chemical measurement occurred in FDOD form. Our FD results were similar in precision to those found on fresh intact lamb using NIRS (900–1700 nm) and hyperspectral (HS) camera by Kamruzzaman et al. (2012a) with a smaller CP range than the present study (21.06–24.05% CP; R^2 = 0.67). However, the precision of these results was much lower than those previous studies on minced beef where R^2 > 0.8 (Lanza, 1983; Sun et al., 2011; Yang et al., 2010). The shorter wavelength range of the NIRS-HS (Kamruzzaman et al., 2012a) was similar to the handheld NIRS in the present study, although its predictions on FI meat were more precise than the predictions by both scanners in the present study, showing the effect of the HS camera adding spatial resolution and detail. The LG-NIRS showed moderate precision (R^2 = 0.51) on FI meat, similar to an earlier study comparing intact and minced lamb (Cozzolino and Murray, 2002). The poorer prediction of CP compared with water and IMF contents in the present study agrees with previous results using fresh lamb (Kamruzzaman et al., 2012a), minced beef and pork (Lanza, 1983), ground mutton (Viljoen et al., 2007), ground lamb and emulsified beef (Kruggel et al., 1981). Kamruzzaman et al. (2012a) attributed this poorer prediction to a narrower range of CP data in comparison to water and IMF, however the present study had broader ranges of CP than those measured previously (Kamruzzaman et al., 2012a; Lanza, 1983; Sun et al., 2011; Yang et al., 2010). The results indicate that the predictions generated from the NIRS

scanners used in the present study showed low to moderate accuracy and precision and were low for FI meat, therefore these sensors would not be sufficient for abattoir screening.

Analysis of stable isotopes showed similar values to Australian (Nakashita et al., 2008) and Irish beef (Bahar et al., 2005). Increased depletion of ^{13}C in beef (δ13C, from –24 to –25‰) can indicate grass feeding compared to grain or legume feeding (Bahar et al., 2005; De Smet et al., 2004; Moloney et al., 2006). Meanwhile, less depletion ($\delta^{13}C$, from –22 to –18 ‰) can indicate C_4 forages being consumed by the animals such as maize (Bahar et al., 2005; Bahar et al., 2009) and longer (> 60 d) time on feed in the feedlot (Bahar et al., 2009). It is also possible that δ13C infers differences in production region (Baroni et al., 2011) although it is unknown if the differences between regions is related to the type of feed consumed, genetics or other environmental factors. Similarly, greater ^{15}N accumulation ($\delta^{15}N$ from 6 to 10 ‰) inferred grass feeding compared to grain or concentrate (Bahar et al., 2005; Moloney et al., 2006) or longer DOF from 60 to 220 d, indicating slow turnover within bovine muscle (Bahar et al., 2009). In the present study, cattleman's cutlet had the least ^{15}N and grain-fed Australian tomahawk steak the highest. Meanwhile, the least ^{13}C depletion occurred in grain-fed Australian tomahawk steak and Angus beef tenderloin, and the most occurring in 150 d grain-fed Angus Scotch fillet and lamb cutlets. It is important to highlight that the present study was not designed to assess differences between origin, cut or nutritional background because the objective was to obtain random samples containing multiple feeding systems, breeds and cuts to make the results more generalisable. Differences in the stable isotopes and total C and N contents between samples were small, likely because only one city was used for sampling, and therefore NIR predictions of these were poor in accuracy and precision. However, total C was predicted with higher R^2 by LG-NIRS compared to smartphone NIRS in the present study, which was similar to Tang et al. (2020) with air-dried soil. NIRS predictions of total C, N and stable isotope ratios have not yet been tested in meat and the results from the present study provide preliminary results and some avenues for exploration in future work, particularly for tracing meat origin and nutritional background.

Handheld NIR spectrometers such as the NIRvascan could find applications along the supply chain from the meat processing industry for grading to retailers for sourcing, marketing, restaurants and even consumers. Further testing in an industrial setting, such as a commercial abattoir, would improve sample size and add further support to its potential adoption, as it can take place on the exposed 12^{th}–13^{th} rib where grading occurs. Its potential use as an objective meat grading tool can decrease the costs associated with labour and improve transparency within the meat processing industry. Further testing including contactless distance scanning would also be beneficial towards adoption by processing and retail industries (Dixit et al., 2017). On-line detection of dark cut beef and sorting into groups using NIRS has shown promise previously at 95% for FI meat (Prieto et al., 2014b; Reis and Rosenvold, 2014) as has similar sorting for pale, soft, exudative pork at an industrial level (Prieto et al., 2017). Exploration of tenderness and shear force predictions using handheld NIRS could be another potential area of investigation, with LG-NIRS showing promise in identifying tender and tough meat post-processing (Shackelford et al., 2005; Yancey et al., 2010), although on-line predictions of these are still behind those of laboratory measures (Dixit et al., 2017). Additionally, current chemometric prediction models (water, CP, IMF) could potentially be improved through the use of HS imaging in combination with NIRS due to the capture of spatial distribution of fat and muscle on HS images (Kamruzzaman et al., 2012a; 2012b). Such prediction models have been shown to non-destructively determine nutritional quality in intact lamb samples with high accuracy.

The accuracy and precision of predictions reported in the present pilot study position the smartphone NIRS as a tool more suited to consumers, retailers or restaurants to monitor marbling or IMF, water, protein and pH of meat. The use of such a consumer grade scanner to detect fat and protein contents could be very useful for consumers at a retail level, particularly given the lack of difference between predictions from each NIR device. However, further studies with larger sample sizes are required before firm conclusions can be drawn. The use of an accurate and precise predictor of objective IMF or marbling in comparison to subjective marbling has been long hailed as an industry solution which can also improve consumer confidence in meat without destruction of the product to

obtain chemical measures (Bindon, 2004; Dixit et al., 2017). For instance, a previous study found AUS-MEAT marbling to explain just 30–40% of the variation in chemical IMF, consistent with most Australian studies conducted (Harper et al., 2003). This is less than the R^2 reported in the present study using the smartphone NIRS on FI meat. Store-labelled "lean meat" is expected to contain less than 2% fat (Pedersen et al., 2003), significantly less than the IMF determined chemically in the majority of studies cited.

Limitations of the present study include the smaller sample size, the small size of samples used for water content measurement due to size restrictions of the freeze-dryer used, and the experimental design of chemical measurements taken from random slices rather than corresponding to scanning location, all of which may have reduced the predictive capacity of the NIRS scanners. Due to these limitations, these conclusions are to be interpreted with caution.

2.5. Conclusion

The use of an affordable handheld, smartphone NIRS scanning system for meat was trialled in the present pilot study with a small number of samples to explore the potential of this technology to measure meat quality attributes. The low-cost consumer grade NIRS sensor showed similar accuracy and precision of predictions to a laboratory-grade NIRS despite its limited wavelength range. Despite promising predictions compared to LG-NIRS, the miniaturised smartphone NIRS instrument did not yield enough precision nor accuracy on intact meat samples required for adoption as a consumer tool based on results from the present study. However, with respect to the small sample size used, the smartphone NIRS sensor deserves further research with larger datasets to increase generalisation of the prediction equations, for example testing in a commercial abattoir setting where all meat is fresh (not aged) and official grading is performed.

3. Portable vibrational spectroscopic methods can discriminate between grass-fed and grain-fed beef

Published in Journal of Near Infrared Spectroscopy 29(6): 321-329.

Abstract. This chapter analysed the ability for portable near-infrared reflectance (NIR) and Raman spectroscopy sensors to differentiate between grass-fed and grain-fed beef. Scans were made on lean and fat surfaces of 108 beef steak samples labelled as grass-fed (n=54) and grain-fed (n=54), with partial least squares discriminant analysis (PLS-DA) and linear discriminant analysis (LDA) used to develop discrimination models which were tested on independent datasets. Furthermore, PLS-DA was used to predict visual marbling score and days on feed (DOF). The NIR spectra accurately discriminated between grass- and grain-fed beef on both fat (91.7%, n = 92) and lean (88.5%, n = 96), as did Raman (fat 95.2%, n = 82; lean 69.6%, n = 68). Fat scanning using NIRS moderately predicted DOF (r^2_{val} = 0.53), though Raman and NIRS lean prediction models for DOF and marbling were less precise (r^2_{val} < 0.50). It can be concluded that portable NIR and Raman spectrometers can be used successfully to differentiate grass-fed from grain-fed beef and therefore aid retail and consumer confidence.

3.1. Introduction

Globally, beef products are labelled by feed type (grass or grain), breed, country of origin and cut (muscle type). Higher-end retailers sometimes supply meat grading qualities such as AUS-MEAT (2018) marbling score or Meat Standards Australia (MSA) certification (MLA, 2017a; Polkinghorne & Thompson, 2010). In general, these qualities are not included when labelling beef at retail level, but an unspoken confidence exists amongst consumers that the labelling is accurate. Consumers worldwide have been known to pay premium prices for certain characteristics that denote good eating quality such as certain cuts, tenderness and marbling (Bonny et al., 2018; Pethick et al., 2018; Polkinghorne et al., 2008). Grass-fed beef is lower in fat, and is considered to be of superior nutritional quality, safety and environmental friendliness compared to grain-fed beef, though often at the expense of eating quality (Prache et al., 2005; Van Elswyk & McNeill, 2014). However, the industry has no robust tools to confirm that the labelling of products is accurate, which may result in product manipulation and mislabelling by producers, processors and retailers being prevalent (Logan et al., 2020a; Prache et al., 2020; Xu et al., 2020). In Australia, certification of grass-fed beef is assessed by the Pasturefed Cattle Assurance System (PCAS), based on solely grazing on pasture and consumption of any amount of grain being banned (PCAS, 2016). Grain-fed beef is governed by the National Feedlot Accreditation Scheme with a list of certified grain types for feeding and a minimum number of days finished on feed (AUS-MEAT, 2018).

The use of analytical methods such as Raman spectroscopy (Logan et al., 2020a; Logan et al., 2020b; Xu et al., 2020) and near-infrared reflectance spectroscopy (NIRS) have been used previously to differentiate beef based on feeding regime (Cozzolino et al., 2002; Prieto et al., 2014a). Cozzolino et al. (2002) found accuracies of 79–83% in classification of beef either from pasture or maize silage feeding. However, the latter study lacked industry applications because it scanned homogenised ground beef using a large and non-portable NIRS. In contrast, Prieto et al. (2014a) were not able to reliably differentiate between ground beef products from cattle fed either sunflower or flaxseed, although chemical composition was predicted accurately.

Near-infrared spectroscopy of intact beef has predicted various meat quality traits, such as pH, fatty acid profile, shear force and colour to various accuracies (Andrés et al., 2008; Byrne et al., 1998; Magalhães et al., 2018; Prieto et al., 2011; Prieto et al., 2017). However, no evidence of studies using NIRS scanning of intact beef for purposes of discriminating grass-fed and grain-fed feeding regime have been found. Recent studies have shown successful discrimination analysis of feeding regime (grass-fed and grain-fed) on intact subcutaneous fat from beef brisket (accuracy 85–98.5%; Logan et al., 2020b; Logan et al., 2021a; Logan et al., 2021b) using a handheld Raman spectrometer, although no studies have discriminated beef based on feeding regime using Raman spectra collected from lean tissue (Chapter 1). This evidence presents an opportunity for future use of both NIRS and Raman in meat processing or retail sectors for authentication. Raman has also been used on intact beef to analyse structure, molecular interactions, fatty acid concentration and species (Xu et al., 2020; Aalhus et al., 2014; Ellis et al., 2005; Fowler et al., 2015b; Fowler et al., 2017; Fowler et al., 2018). In a recent study, a lab-grade vis-NIR spectrometer was able to discriminate barley and corn finished beef loins at an accuracy of 90–94% (Barragán et al., 2021).

One issue with conventional analytical methodologies is their large size, high cost, and non-portable nature. For this reason, palm-sized near-infrared spectrometers have been developed for industry and consumer applications (Dixit et al., 2017; Teixeira dos Santos et al., 2013). Despite shorter wavelength ranges (900–1700 nm) often restricting predictive capabilities (Dixit et al., 2017; Dixit et al., 2020; Teixeira dos Santos et al., 2013), these handheld spectrometers are much cheaper, smaller and can connect to a smartphone. Indeed, some studies have shown comparable predictions from low-cost smartphone NIR spectrometers to larger benchtop spectrometers to measure meat quality and chemical attributes (Coombs et al., 2019; Pham et al., 2018), and soil quality (Tang et al., 2020). However, visual marbling or chemical fat content have been predicted with low to moderate precision (r^2_{CV} or r^2_{val} = 0.27–0.75) using handheld NIR (Coombs et al., 2018; Dixit et al., 2020; Fowler et al., 2020) and Raman spectrometers (Fowler et al., 2015b). Interestingly, none of these studies compared Raman with low-cost consumer grade NIR spectrometers to predict DOF or differentiate grass-fed and

grain-fed beef, which could expand the adoption and use of NIR technology for commercial applications.

The aims of the present study were therefore to use two portable vibrational spectrometers (Raman and NIR) to differentiate between grass-fed and grain-fed beef without prior information on the exact diet received by the animal, the cut name, or the muscle type on a selection of whole, unhomogenised store-bought cuts labelled as grass-fed or grain-fed for 100, 150 or 300 minimum days on feed (DOF). Furthermore, DOF and visual scoring of MSA marbling were predicted using these spectroscopic methods scanning both lean and fat tissues. In previous studies, Raman spectroscopy has been successful in differentiating grass-fed and grain-fed beef and in predicting marbling. However, handheld NIR spectrometers designed for consumer use have not been trialled for discrimination of feeding regime, which could greatly benefit the consumer and retail industry with an objective, portable, practical, and low-cost classification system. Furthermore, the ability to predict DOF has not been assessed previously, and its accurate prediction shows great potential for the meat supply chain.

3.2. Materials and Methods

3.2.1. Sampling and experimental design

A total of 108 beef steak samples, comprising 54 grass-fed and 54 grain-fed samples, were used in this study. All samples were purchased locally from butchers and supermarkets, with 42 of the 54 grass-fed samples sourced from a certified butcher which partners with five specific trustworthy farms, providing full production, animal, and background information (https://www.1888certified.com.au). The remaining 12 grass-fed samples were PCAS certified (PCAS, 2016). Grain-fed samples were all certified by the National Feedlot Accreditation Scheme (NFAS) and labelled with minimum DOF information in the categories of 100 (n = 12), 150 (n = 23) and 300 (n = 19) (https://www.nh-foods.com.au/facilities/whyalla-beef/). Nine different cuts (bolar blade, T-bone, porterhouse/New York, rib eye, rump, scotch, shin, tri-tip, and short ribs) were used across the entire study, though cut type was a random effect and not included as a factor for analysis in this study. Visual scoring of marbling was performed quantitatively according to MSA standards (MLA, 2017a).

3.2.2. Near-infrared reflectance and Raman spectroscopy

A palm-sized (8.2 x 6.6 x 4.5 cm, 136 g) handheld NIR spectrometer based on the Texas Instruments' Digital Mirror Device (NIRvascan, Allied Scientific Pro., Gatineau, Canada) was connected via Bluetooth to a smartphone using the NIRScan Nano application (KS Technologies, Colorado Springs, USA) and used to acquire reflectance spectra between 900 and 1700 nm (3.5 nm intervals, 1.69 x 0.025 mm slit size, 25–50 ms scan time) on meat. The Raman spectrometer (Bravo, Bruker Optics, Billerica, USA) acquired spectra between 300 and 3200 cm^{-1} (2 cm^{-1} intervals, 1 x 1 mm slit size, 5 second scan) using Duo Laser^{TM} excitation (power < 100 mW; 700–1100 nm). Both spectrometers were calibrated using inbuilt wavenumber calibration. Spectra were acquired on both lean and fat surfaces, but eight grass-fed samples (rib eye steaks) had no testable fat. Ten scans were performed on the surface of each lean meat sample (n = 54 grass, n = 54 grain) and four scans on the fat surface (n = 46 grass, n = 54 grain).

It is worth noting that the size (27 x 15.6 x 6.2 cm, 1.5 kg) of the Raman affected its portability. Both NIR and Raman spectra were acquired on meat covered by plastic wrap to limit cross-contamination and eliminate probe cleaning time. Due to the availability of beef steaks, sampling was performed in two separate sessions.

3.2.3. Analysis of spectral data

Lean and fat spectral datasets for each of the Raman and NIR spectrometers were kept separate for all cases. Raman data were automatically downloaded to the Bruker OPUS software (version 8.5, Bruker Optics, Ettlingen, Germany) and then converted to comma separated values (CSV) format in Microsoft Excel. The NIR spectra were automatically stored in the NIRScan Nano smartphone application in CSV format and downloaded to Microsoft Excel. All Raman and NIR spectra for the same sample (repetition) were averaged using OPUS and Microsoft Excel, respectively, prior to statistical analysis and model development. Averaged spectral data per sample were imported into R (R Core Team, 2020), where NIRS data were inverse log transformed (log(1/R)) as per Lanza (1983) and no missing values were obtained. To detect outliers, both NIR and Raman spectral data were scaled and centred, then analysed via principal components analysis (PCA) model to identify Q residuals and Hotelling T^2 values plotted with two components using the *mdatools* package in R (Kucheryavskiy, 2020; Kucheryavskiy, 2021). Outliers were defined as having orthogonal and score distances > 20 on the residual plot (Kucheryavskiy, 2021). For Raman, lean data had eleven outliers which required removal (n = 97) and fat data had fourteen (n = 86). No outliers were detected for either fat or lean NIRS data. Data were further pre-processed using standard normal variate (SNV) by scaling of rows as described by Huang et al. (2010).

3.2.4. Discrimination models for grass-fed and grain-fed beef

Calibration models were developed using the *Caret* package (Kuhn, 2020) in R with 75% of the spectral data used for the training dataset and 25% for the independent validation dataset using an inbuilt function of the *tidymodels* package (Kuhn & Wickham, 2020). Partial least squares discriminant analysis (PLS-DA) and linear discriminant analysis (LDA) were used to determine the accuracy, sensitivity, and specificity of the trained model to discriminate between grass-fed or grain-fed beef on the independent dataset. Four training models for each spectrometer and tissue type were developed (NIRS lean, NIRS fat, Raman lean and Raman fat) and repeated 10-fold cross-validation was performed in *Caret* with five repetitions. Model development and selection (15 components based on maximised receiver operator curve = 1.00) was done in the *Caret* package of R (Kuhn, 2020). Model evaluation for discriminant analysis was done using a two-class (grass or grain) confusion matrix of the classified predictions, applying the training or calibration models (PLS-DA and LDA) to both the calibration and external validation datasets with the metrics generated from the confusion matrix being sensitivity, specificity, precision, and accuracy (Kuhn, 2020).

For each of the four models, variable importance plots were obtained in R using the *Caret* package (scaled from 0 to 100 as per Kuhn, 2020) to determine the important regions of each spectrum to discriminate between grass-fed and grain-fed beef. Thresholds were defined as the most important wavelengths in the NIR spectrum or wavenumbers in the Raman spectrum (importance > 40).

3.2.5. Predictions of MSA marbling and DOF

A linear regression model in R was used to assess the difference in marbling score between DOF groups (0, 100, 150 and 300), with DOF being the fixed effect. Partial least squares regression (PLSR) models (75% training, 25% validation) with resampling using cross-validation (10 buckets repeated 5 times) were used to develop prediction models of MSA marbling and DOF on the training dataset and then applied to the validation dataset. The optimal PLSR model (3 components) with the lowest root mean

square error of repeated cross-validation ($RMSE_{CV}$) was selected. The resulting prediction models were then used to determine the goodness-of-fit of the predictions against the observed values on the validation dataset using the *Spectracus* package in R (Fajardo et al., 2019). Goodness-of-fit of each PLSR model (NIRS and Raman of lean and fat) was assessed through the coefficient of determination (r^2_{val}), root mean square error ($RMSE_{val}$), residual prediction deviation (RPD) and bias of the external validation dataset (Williams et al., 2017).

3.3. Results

Descriptive statistics of cut type, marbling and DOF are presented in Table 9. The mean marbling score was 692.6 and mean DOF was 95.8 days. Results of the linear regression model determined that marbling score was not significantly affected by feeding regime, where grain-fed beef had an MSA score of 726 ± 29.5 and grass-fed had an MSA score of 645 ± 32.7 ($P = 0.07$; data not shown). However, DOF affected MSA marbling with 150 DOF showing the lowest MSA marbling score and 1000 DOF showing the highest ($P < 0.05$; Table 9).

Table 9. Descriptive statistics for days on feed (DOF) and Meat Standards Australia (MSA) marbling score.

DOF (days)	n	Mean (± SEM) Marbling	Minimum Marbling	Maximum Marbling
0	54	659.3 (16.04)^{c}	400	800
100	12	791.7 (66.10)^{b}	400	1000
150	23	465.2 (16.96)^{d}	300	600
300	19	1000 (0.00)^{a}	1000	1000

$^{a, b\ c, d}$ Superscripts with a different letter indicate a significant difference ($P < 0.05$).

The NIR spectral data were obtained and were found to be highly consistent, with no outliers found for lean or fat. Mean raw absorbance spectra for lean and fat of each feeding system is presented in Fig. 6. Grain-fed beef had greater absorbance than grass-fed beef across the entire spectrum for lean although the largest difference was beyond 1380 nm (Fig. 6a). For fat, grass-fed beef had greater reflectance than grain-fed beef at wavelengths shorter than 1350 nm, and grain-fed beef had greater reflectance at 1445 to 1470 nm (Fig. 6b).

Raman intensity spectra were obtained, and data cleaning was performed to remove outliers above distances of 20 on Hotelling-T^2 plots formed following PCA modelling. High numbers of outliers were found for both lean (10.2%, n = 11) and fat (14%, n = 14). Mean cleaned Raman spectral data was presented in Fig. 2 for lean (Fig. 7a) and fat (Fig. 7b).

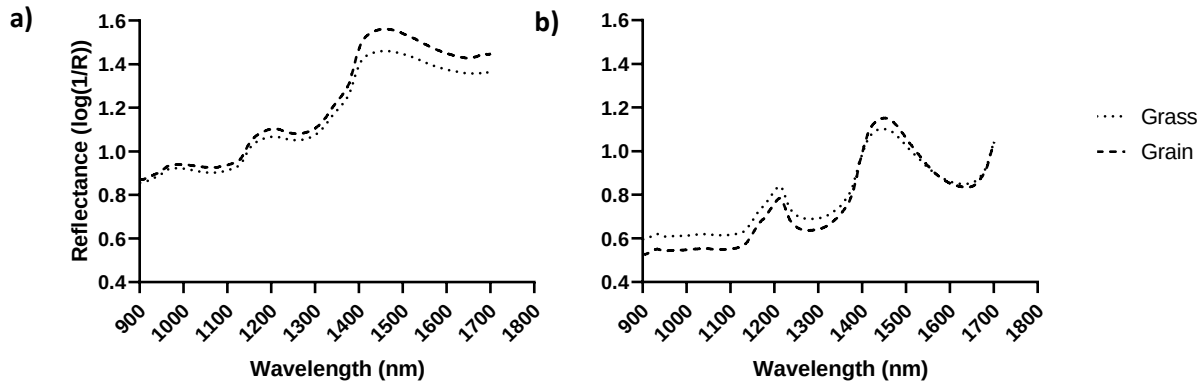

Fig. 6. Mean near-infrared absorbance (log(1/R)) spectra of grass-fed and grain-fed beef for **a)** lean muscle tissue (n = 108) and **b)** fat tissue (n = 100).

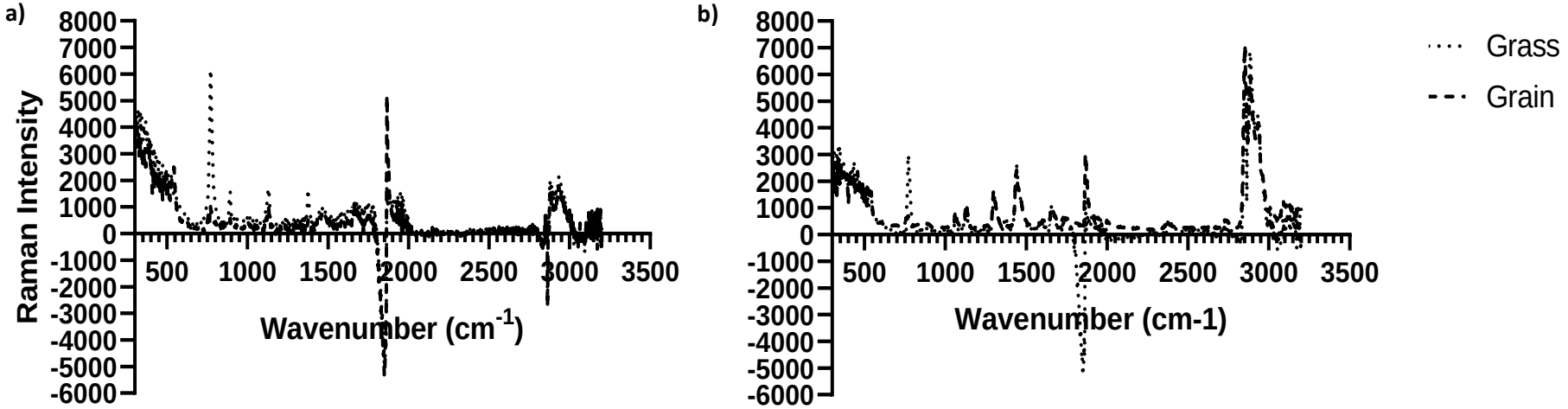

Fig. 7. Mean cleaned (post-outlier removal) Raman spectra of grass-fed and grain-fed beef for **a)** lean tissue (n = 97) and **b)** fat tissue (n = 86).

Model performance metrics indicated that the models performed nearly perfect on the training datasets with accuracies of 100% for both lean and fat using either NIR or Raman spectroscopy, regardless of PLS-DA or LDA, except for Raman on lean where it was 91–96% accurate (Table 10). However, model performance on the validation dataset was not as good with accuracy ranging between 65 and 95%; similar to the calibration dataset, Raman on lean showed significantly lower accuracy (65–69%) while other models showed accuracies above 80% (Table 10).

Model performance on the validation dataset indicated that discrimination between grass-fed and grain-fed on lean was more accurate with NIR spectroscopy compared to Raman spectroscopy, whereas discrimination accuracy of fat was greater using Raman (Table 10). Generally, sensitivity, specificity and precision were similar between NIR and Raman spectroscopy for fat, although NIRS was superior to Raman spectroscopy for lean (Table 10). Limited differences occurred for goodness-of-fit metrics between PLS-DA and LDA models, with LDA slightly better for NIR lean and fat spectra, and PLS-DA slightly better for Raman fat spectra. Raman scanning of lean showed PLS-DA to be superior to LDA in specificity and precision, though inferior in sensitivity and identical in accuracy (Table 10). Fat scanning using the NIRS was found to have excellent specificity and precision (100%) for differentiating grass-fed and grain-fed beef using PLS-DA and LDA, while lean scanning showed higher accuracy and sensitivity (>84%, Table 10). For Raman, all metrics were superior scanning fat compared to lean.

Modelling with PLS-DA indicated that the most important wavelengths for discriminant analyses occurred at the upper limits of the NIR spectrum for both lean and fat (1686–1700 nm), while other areas of high importance occurred at 954–994 nm, 1040–1070 nm, 1128–1162 nm, 1199–1225 nm and 1391–1401 nm for lean, and 1388–1409 nm for fat. Raman spectroscopy showed more importance for differentiation between wavelengths of 762–782 cm^{-1} and 1864–1874 cm^{-1} for lean and 770–772 cm^{-1}, 1820–1864 cm^{-1} and 2862 cm^{-1} for fat.

The NIR spectrometer showed poor precision in predicting the MSA marbling score and DOF (r^2_{val} < 0.35). Precision and accuracy were greatest for predicting DOF scanning fat with the Raman spectrometer, although these were still moderate (Table 11).

Table 10. Discrimination between grass-fed and grain-fed beef following scanning of lean (n = 108) and fat (n = 100) surfaces using a smartphone NIR spectrometer and a portable Raman spectrometer using partial least square discriminant analysis (PLS-DA) and linear discriminant analysis (LDA) on the training (calibration) and validation datasets.

	Calibration* dataset (n = 81)			
Model	Sensitivity	Specificity	Precision	Accuracy
NIR Lean PLS-DA	1.000	1.000	1.000	1.000
NIR Lean LDA	1.000	1.000	1.000	1.000
NIR Fat PLS-DA	1.000	1.000	1.000	1.000
NIR Fat LDA	1.000	1.000	1.000	1.000
Raman Lean PLS-DA	0.934	0.956	0.969	0.959
Raman Lean LDA	0.909	0.902	0.882	0.906
Raman Fat PLS-DA	1.000	1.000	1.000	1.000
Raman Fat LDA	1.000	1.000	1.000	1.000
	Validation dataset (n = 27)			
	Sensitivity	Specificity	Precision	Accuracy
NIR Lean PLS-DA	0.846	0.923	0.917	0.885
NIR Lean LDA	0.923	0.923	0.923	0.923
NIR Fat PLS-DA	0.692	1.000	1.000	0.846
NIR Fat LDA	0.769	1.000	1.000	0.885
Raman Lean PLS-DA	0.500	0.846	0.714	0.695
Raman Lean LDA	0.600	0.692	0.600	0.652
Raman Fat PLS-DA	0.900	1.000	1.000	0.952
Raman Fat LDA	0.900	0.818	0.818	0.859

* Models directly applied to the training dataset.

Table 11. Goodness-of-fit of partial least squares regression (PLSR) models to predict marbling score (Marbling) and days on feed (DOF, days) using NIR and Raman spectroscopic analysis of lean (n = 108) and fat tissues (n = 100) of beef. *

Model	r^2_{val}	$RMSE_{val}$	RPD	Bias
NIR Lean Marbling	0.162	623.6	0.374	-589.0
NIR Lean DOF	0.320	97.61	1.145	0.737
NIR Fat Marbling	-0.045	257.9	0.941	13.05
NIR Fat DOF	-0.039	596.0	0.190	574.3
Raman Lean Marbling	<0.001	259.8	0.856	62.07
Raman Lean DOF	0.134	632.5	0.187	616.9
Raman Fat Marbling	0.010	619.6	0.385	-576.5
Raman Fat DOF	0.468	83.00	1.330	31.74

* Results shown for the validation dataset only.

3.4. Discussion

The two spectrometers investigated in this pilot study showed high accuracy (> 85%) for the purposes of discrimination between grass- and grain-fed beef cuts using both PLS-DA and LDA, with the exception of Raman prediction models on lean with moderate accuracy (60–70%). The most promising finding of the study was that the smartphone NIR spectrometer, designed for consumer and retailer use, was able to differentiate grass-fed from grain-fed beef with high precision, sensitivity, specificity, and accuracy using PLS-DA and LDA. Its advantages compared to the Raman used include smaller size, lower cost, and easier connectivity to a smartphone instead of a laptop (Coombs et al., 2019; Dixit et al., 2020; Pham et al., 2018; Tang et al., 2020; Teixeira dos Santos, 2013), along with its more consistent spectra with no outliers detected in the present study. The presence of outliers for the Raman spectra may be due to the more difficult use of this bulkier instrument and the sensitivity to movements of the probe while scanning the samples compared to the NIRS. The Raman spectrometer required more extensive training for its use, in addition to more data cleaning post-scanning in the present study, which along with its higher cost rendered it less efficient. Possible explanations for greater accuracy of the NIR spectrometer include smoother spectra and the ability to detect larger differences between the spectra of both feeding systems. Findings from this study infer that grass-fed and grain-fed beef can be reliably discriminated from each other using a consumer grade smartphone spectrometer, irrespective of the cut used.

The majority of the differences between grass-fed and grain-fed beef occurred at the higher end of the NIR spectrum for lean (1400–1700 nm) but at shorter wavelengths for fat (900–1400 nm), the lean being similar to findings of Cozzolino et al. (2002) who scanned homogenised beef to discriminate between grass-fed and silage-fed beef with a benchtop NIRS of wider wavelength range (400–3500 nm) compared to the present study. Interestingly, the areas considered most important in the present study corresponded to areas of the greatest change in spectral reflectance rather than the largest separation between grass-fed and grain-fed spectra, except for the highest importance at 1700 nm (data not shown). The 1700 nm region showed the greatest separation between spectra and

this region captures C-H bonds in fatty acids (Manley, 2014), with grain-fed lean tending to have greater fat content than grass-fed beef. The peaks identified in the mean NIR spectrum of lean beef in the present study were 960–980 and 1400–1450 nm, corresponding to –OH bonds and indicating water content (Cozzolino et al., 2002; Delpy & Cope, 1997; Manley, 2014; Prieto et al., 2011). Another peak occurred between 1130 and 1270 nm corresponding to C-H bond vibrations, and differences in intramuscular fat (IMF) content or fatty acid profile (Prieto et al., 2011). These results suggest that fat content and composition may help differentiating grass-fed from grain-fed beef (Logan et al., 2020a). It is likely that cut selection of samples in the present study affected the marbling and thereby the fat content, and the use of one cut only in future studies may alleviate this (Logan et al., 2020a; Prieto et al., 2014a). However, the present study was designed to explore the potential value of these sensors across a range of cuts and marbling scores so that the experimental design justified a random selection of cuts.

The use of fat scanning in addition to lean to differentiate grain-fed and grass-fed beef showed similar and slightly improved results to those presented by Cozzolino et al. (2002) who used homogenised beef. The Raman spectrometer produced more accurate classification scanning fat compared to lean, likely due to the increased differences in spectral signatures of grass-fed and grain-fed fat in the shorter half of the spectrum (< 1750 cm^{-1}), with a more pronounced peak occurring between 1130 and 1270 cm^{-1} (C-H fat bonds) for Raman compared to lean surface scans. The NIR spectrometer produced negligible differences in discrimination accuracy between lean and fat. However, fat scans showed greater precision and specificity, likely due to the spectral differences between grass- and grain-fed beef occurring over longer wavelength ranges for fat (900–1350, 1400–1520 nm) compared to lean (1400–1700 nm).

Previous research demonstrated differences in fatty acid profile that were picked up by Raman spectroscopy, with key shifts occurring at 1069, 1127, 1301 and 1445 cm^{-1} (Logan et al., 2020a). Similar differences were found in the present study using Raman, where the most notable differences between grass-fed and grain-fed were important between 760 and 780 cm^{-1} and between 1820 and

1860 cm^{-1} when fat was scanned. The latter shift corresponds to carbon-carbon double bonds (HORIBA, 2020), whereupon grass-fed beef showed higher C-C double bonding than grain-fed beef, which infers higher proportions of unsaturated fatty acids (Logan et al., 2020a).

Dian et al. (2008) obtained 89.1–98.6% accuracy in differentiating grass-fed and grain-fed lamb carcases, which was greater compared to the present study, through use of the visible spectrum (trimmed data from a vis-NIRS down to 400–700 nm) as opposed to the NIR spectrum (900–1700 nm). Carotenoid pigments are known to absorb light between 450 and 510 nm, within the visible spectrum, with this region producing higher absorbance in grass-fed beef (Knight et al., 2001; Prache & Thierez, 1999; Realini et al., 2004). Measuring absorbance and the Commission Internationale de l'Eclairage (CIE) b* value (CIE, 1978) in fat using a colorimeter would likely produce strong differences between grass-fed and grain-fed beef and aid in their authentication (Aoriera et al., 2017; Dunne et al., 2009). A recent study using a spectrometer with a similar range showed predictions of CIE colour to be comparable to a benchtop laboratory-grade vis-NIRS with r^2_{val} around 0.7 (Patel et al., 2021). Additionally, measurements of vitamin E, fatty acid profile, β-carotene and fat colour could further enhance grass-fed and grain-fed discrimination (Duckett et al., 2009; Logan et al., 2020a; Logan et al., 2021a; Luciano et al., 2011).

Predictions of marbling score and days on feed were moderate to poor using both spectrometers, however it was promising that fat was a better predictor of DOF compared to lean, and lean was a better predictor of visual marbling score compared to fat. This likely occurred because fat content, increase of saturated fatty acids and decrease in carotenoid content, which affect the spectral signature of fat, are later maturing traits associated with longer DOF (Dunne et al., 2009). Scanning fat tissue could not infer greater amounts of IMF, which is scored visually on lean tissue (MLA, 2017a). Marbling score has not been traditionally predicted as well as chemical IMF, and the results of the present study reflect this and are comparable to earlier work using portable NIRS (900–1700 nm) to predict marbling score or IMF on intact meat where r^2_{CV} or $r^2_{val} < 0.30$ (Coombs et al., 2019; Dixit et al., 2020; Magalhães et al., 2018). Indeed, previous research demonstrated that the best

results for handheld NIRS to predict IMF resulted in r^2_{CV} values between 0.5 and 0.67 (Fowler et al., 2020; Pham et al., 2018). Care was taken in the present study to use a larger number of scans per sample than previously mentioned studies that encompassed a greater portion of the surface area of each steak, which was designed to improve predictions of marbling. The high bias found for the prediction of MSA marbling infers a lack of model precision which could have arisen due to the wide range of cuts of meat.

The fact that MSA and other marbling grading systems are subjective (Polkinghorne & Thompson, 2010) may have affected the accuracy of these predictions. The prediction of IMF would be a more accurate and objective measure, but marbling score is still the industry standard for beef quality in Australia (AUS-MEAT, 2018; MLA, 2017a). Nevertheless, promising results ($r^2_{val} \geq 0.65$) have been found previously for IMF prediction using portable vis-NIR spectrometers (350–2500 nm) in beef (Sun et al., 2011) and lamb (Dixit et al., 2020; Pullanagari et al., 2015).

The present study attempted to predict the length of time animals were fed a high-grain diet in a feedlot through DOF which is an important aspect for many consumers when making purchase decisions as it often infers Wagyu genetics and increased fat content, marbling and desirable flavours associated with these (Chapter 1; Frank et al., 2016; Logan et al., 2021b; Miller, 2020; Van Elswyk & McNeill, 2014). The predictions of DOF and marbling were made with the objective to infer whether fat content would lead to the high accuracy of discriminating grass- and grain-fed beef. However, the predictions of DOF were poor in the present study except for scanning of fat using the Raman spectrometer, which showed moderate precision. It is therefore recommended that future work should use larger datasets with staggered DOF samples across a wider range, such as 0, 70, 100, 150, 300, 450 and 600 DOF. For instance, a recent study using Raman scanning subcutaneous fat was able to classify beef (n = 520) into different DOF groups (grass, grass + supplement, 70 and 100 DOF) at 87.1% accuracy and 86.8% precision by multi-class PLS-DA instead of PLSR (Logan et al., 2021b), as was used in the present study. Fat proved to be a better predictor for DOF than lean for Raman, which may have occurred due to changes in the fatty acid profile with DOF (Logan et al., 2021b; Van Elswyk

& McNeill, 2014). Fat colour would be expected to be lighter from longer grain feeding durations due to decreased β-carotene derived from pasture feeding (Duckett et al., 2009; Logan et al., 2020a). However, the spectrometers used in the present study do not measure reflectance in the visible range and that may explain the low accuracy for predicting DOF.

Despite similar prediction results for differentiating grass-fed and grain-fed beef by both spectrometers, the results of the present study indicate that NIRS showed greater ability to differentiate grass-fed and grain-fed beef without the need for data cleaning. The scientific discrimination methods presented in this paper could save time and costs, eliminating the need for chemical analysis to authenticate or certify meat products (Xu et al., 2020). This could combat mislabelling of beef as grass-fed when it may in fact be finished on grain, or the other way around. However, the predictions of DOF and marbling score were not sufficiently accurate for either industry or retail applications (Barlocco et al., 2006; Williams et al., 2017).

Raman spectroscopy was not able to produce greater spectral differences between grass-fed and grain-fed lean compared to NIRS. However, its ability to discriminate accurately following fat scanning was comparable to NIRS after data cleaning, which was required due to the high variability of Raman spectra within and between samples. The PLS-DA accuracy of Raman in the present study was comparable to recent findings scanning fat (85–96%; Logan et al., 2021b) although lean was less accurate for discrimination purposes, being 69.5% compared to 96.5–98.5% (Logan et al., 2020b; Logan et al., 2021a). Similarly, precision for fat using PLS-DA in both the present study and Logan et al. (2021b) was 100%. For future work with Raman, care would need to be taken when scanning to reduce analytical variability because Raman spectral scans recorded in the OPUS software are averages of the total number of scans and therefore, one being incorrect may compromise the entire sample. This was also reflected in the high number of outliers, which were not detected in recent studies (Logan et al., 2020b; Logan et al., 2021a) This finding shows that operator experience and usage may affect the results and training operators on using Raman spectroscopy correctly would be

needed to avoid fluorescence reflection from raw samples, which in turn affects the amount of sub-samples (Qiao & van Kempen, 2004).

3.5. Conclusions

The present preliminary study highlighted the potential for vibrational spectroscopic devices to classify beef as grass-fed or grain-fed, with particularly promising results for the low-cost, smartphone NIR spectrometer. The high variability of Raman spectroscopy indicated the need for further research in the field to refine both the technology and scanning technique as these likely affected the prediction models. Marbling score could not be predicted accurately with either spectrometer used in the present trial. The NIR spectrometer proved to be a better predictor than the Raman spectrometer for feeding regime when scanning lean, though the reverse was true when scanning fat. The NIR sensor was better at predicting both DOF and marbling score using fat and lean scanning, compared to Raman. The lower number of outliers, lower cost, smaller size, and smartphone connection of the consumer grade NIR spectrometer makes it an attractive proposition for authentication and differentiation of grass-fed and grain-fed beef at a retail level which can improve consumer confidence.

4. Differentiation of sheep and cattle internal organs using visible and short-wave infrared hyperspectral imaging

Abstract. Automatic identification and sorting of livestock organs in the meat processing industry could reduce costs and improve efficiency. Two hyperspectral sensors encompassing the visible (400–900 nm) and short-wave infrared (900–1700 nm) spectra were used to identify the organs by type. A total of 104 parenchymatous organs of cattle and sheep (heart, kidney, liver and lung) were scanned in a multi-sensory system which encompassed both sensors along a conveyor belt. Spectral data was obtained and averaged following manual mark-up of 3 to 8 regions of interest of each organ. Two methods were evaluated to classify organs: partial least squares discriminant analysis (PLS-DA) and random forest (RF). In addition, classification models were obtained with the smoothed reflectance and absorbance, and the first and second derivatives of the spectra to assess if one was superior to the rest. The in-sample accuracy for the visible, short-wave infrared, and combination of both sensors was higher for PLS-DA compared to RF. The accuracy of the classification models was not significantly different between data pre-processing methods or between visible and short-wave infrared sensors. Hyperspectral sensors seem promising to identify organs from slaughtered animals which could be useful for the automation of quality and process control in the food supply chain such as abattoirs.

4.1. Introduction

The routine differentiation of livestock organs at abattoirs by meat inspectors occurs based on organ type combined with the gross exclusion diagnosis of potential diseases rendering offal safe for human consumption by trained meat inspectors and a supervising veterinarian (Wilson et al., 2019). This differentiation is a manual labour-intensive process and includes the potential for human error. Automation of the post-mortem process in abattoirs has been trialled in previous studies using non-contact and non-invasive imaging methods including X-ray attenuation, computed tomography (CT), and hyperspectral (HS) imaging for livestock body composition analysis (Elmasry et al., 2012a; Scholz et al., 2015). Such systems have allowed for meat and organs to not be destroyed or contaminated during analysis, with rapid on-line technologies providing instant feedback at processor chain speed and no need for sample preparation or external transportation (Elmasry et al., 2012a; Scholz et al., 2015; Teixeira dos Santos et al., 2013). However, these non-invasive imaging systems generate extensive data which require pre-processing methods and algorithm development to be sufficiently accurate and efficient for use in industry (Elmasry et al., 2012a; Scholz et al., 2015).

The characteristics, analysis, and applications of spectral imagery in meat quality evaluation were comprehensively reviewed by Elmasry et al. (2012a), who concluded that HS imaging systems can be used successfully as quality control tools in meat processing industries. Hyperspectral imaging measures the reflectance of light in multiple narrow bands along the light spectrum and has shown great potential in animal industries (Kumar et al., 2016; Xu & Sun, 2017). These HS technologies can be split into fractions of visible (VIS; 400–900 nm) and short-wave infrared (SWIR; 900–1700 nm). In the agriculture sector, HS has been used for prediction of quality, safety, contamination detection, microbial spoilage, and chemical composition of fruits, cereal grains, animal feed and meat (Baeten et al., 2007; Cheng & Sun, 2015; Huang et al., 2014a; Kamruzzaman et al., 2012a). The spectral data downloaded from HS devices are compared between two objects of interest, which can be differentiated based on differing spectral signatures by peaks or differing intensities at certain

wavelengths, with classification occurring through development of prediction models (Huang et al., 2014a).

In the sphere of organs, differentiation and segmentation has occurred based on different spectral intensities of porcine arteries, veins and organs including liver and colon (Akbari et al., 2009), and five tissues (peritoneum, urinary bladder, spleen, small intestine and colon) during open exploratory surgery on a pig (Akbari et al., 2008). However, these studies were limited by a sample size of one animal and no further studies have attempted to differentiate animal organs by type. Despite this, studies examining HS sensors have successfully used seven SWIR wavelengths to differentiate offal from lamb muscle (Kamruzzaman et al., 2014), and similarly three VIS and two SWIR wavelengths to differentiate beef from chicken in mince mixtures (Kamruzzaman et al., 2015). In addition, a benchtop spectrometer containing VIS and SWIR sensors were successful to differentiate beef, lamb, pork, and chicken meats from one another (Cozzolino & Murray, 2004).

Different data pre-processing and machine learning methods to analyse spectral data are common (Zeaiter et al., 2005). However, the comparison of methods is rare in scientific literature, and it is unclear which methods may be superior. For instance, several studies use absorbance instead of reflectance data (Lanza, 1983), others use first- and second-order derivatives to capture changes in the spectra (Ritthiruangdej et al., 2011), and others have combined these with smoothing of the spectra such as centred moving average, multiplicative scatter correction (MSC), detrending, standard normal variate (SNV) or Savitzky-Golay filtering (Savitzky & Golay, 1964) to reduce non-chemical background and baseline signals from spectra (Huang et al., 2014a; Prieto et al., 2017; Zeaiter et al., 2005). Kamruzzaman et al. (2012b) used centred moving average of reflectance spectra and found no improvements with derivatives, MSC and SNV, while Kamruzzaman et al. (2015) concluded that raw absorbance spectra were optimal.

The aims of the present study were to: 1) investigate the differences between livestock organs in spectral signatures generated from VIS and SWIR imagery; 2) explore the potential of these to differentiate bovine and ovine parenchymatous organs (heart, kidney, liver, and lung); and 3) evaluate

the effect of different data pre-processing techniques and machine learning methods on the accuracy of organ classification. Both reflectance and absorbance data, and their first and second derivatives as pre-processing methods, were used as predictors with partial least squares discriminant analysis (PLS-DA) and random forest (RF) algorithms. It was hypothesised that a multi-sensory platform could provide a spectral profile of individual organs that can be used for development of discrimination algorithms for the automation of this process into food safety and quality control in the red meat industry.

4.2. Materials and Methods

No animals were slaughtered for the purpose of this study, with offal being obtained from an abattoir and a butcher. Therefore, animal ethics approval was not required.

4.2.1. Sample collection and scanning procedure

A total of 104 parenchymatous bovine and ovine organs were collected from a collaborating abattoir and local butcher and maintained at refrigerator temperatures (1–4 °C) prior to scanning (Table 12). The organs included in this study were heart (n = 33), kidney (n = 20), liver (n = 29) and lung (n = 20).

A prototype multi-sensory platform consisting of dual-view multi-energy X-ray and a VIS and SWIR HS imaging system (Rapiscan Inspection System AK198, Rapiscan Systems Pte Ltd., Singapore) connected to a Cube computer running Ubuntu (Linux OS) was used for the imaging of the organs (Fig. 8). Organs were placed in a sealed tray with a transparent acrylic lid to ensure HS penetration and double containment, which was placed within protective lead curtains for scanning. A conveyor transported the samples from end to end (6.64 s for 1260 mm, 189.8 m/s) with both sides protected by lead shielding while X-rays were on. The X-rays (2 x 160 keV tungsten tubes), light-emitting diode (LED) strip lamp (VIS) and a quartz infrared (QIR) lamp (SWIR) provided light sources for the HS sensors.

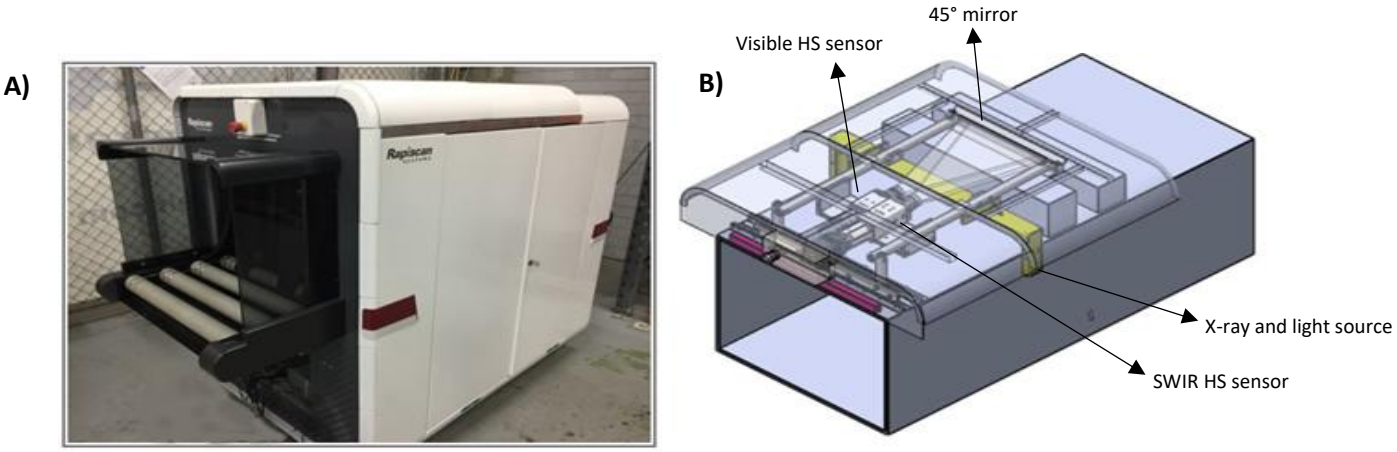

Fig. 8. Rapiscan multi-sensory imaging system used to scan livestock parenchymatous organs. **A)** The external view of the complete prototype imaging system (AK198). **B)** A schematic showing placement of the hyperspectral sensors within the imaging system. *Source*: Rapiscan Systems Pte Ltd.

The HS imaging system consisted of two sensors covering the spectral ranges from 400–900 nm (VIS) and 900–1700 nm (SWIR). The VIS (Basler Ace GigE, Photonic Science, East Sussex, UK) and SWIR (Snake A/C GigE v3 AK081, Photonic Science, East Sussex, UK) sensors were powered by 12 V power supply units and fitted with Specim spectrographs (VNIR V10E and NIR V17E, respectively) and a Grade 1 InGaAs detector with air cooled housing. Spectral resolutions were 3 and 5 nm for VIS and SWIR, respectively, with both sensors capturing 200 spectral slices per second. Exposure time, image size (width, length, offset) and acquisition rate were controlled by Ubuntu (Linux OS) computer programs (eBUSPlayer SDK, Pleora Technologies, Kanata, Canada) and stream2camstodisk command line (B. E. Allman, personal communication May 19, 2020) in the Aravis environment of Linux. Spectral increment was approximately 1.5 nm between contiguous bands, with 300 bands for VIS and 512 for SWIR. The VIS sensor had a 1920 x 1200 (spectral x spatial) pixel sensor, spectral binned four times and offset 70 pixels, spatial dimension was not binned, and offset was 550 pixels, equalling 300 bands. The SWIR sensor had a 640 x 512 (spatial x spectral) pixel sensor, offset by 64 pixels and the area captured was 256 pixels. These dimensions were chosen in order to get 150 frames per second (fps)

for both HS sensors. Exposure times were calculated from 150 Hz, resulting in a 6.666 ms refresh rate, so exposure times were 6.4 ms for VIS and 4 ms for SWIR to download data at this refresh rate.

4.2.2. Extraction and analysis of spectral data

Scanned images (PNG format) trimmed to comprise the tray containing organ samples were constructed from 200–400 frames generated by the HS sensors using MATLAB programming language (MATLAB R2021a, Mathworks Inc., USA). ImageJ (version 1.53a; National Institutes of Health, Bethesda, MD) was used to mark-up regions of interest (ROI) with 7 x 7 pixels in size upon each complete organ image avoiding visible fat. Three to eight ROI were marked-up upon each organ depending on the organ's size, with larger organs having more ROI than smaller organs. Gnu Image Manipulation Program software (version 2.10.18; GIMP Development Team, 2020) was used for image viewing and manipulation to obtain pixel values of the ROI, which were then written into a MATLAB algorithm developed by B. E. Allman (personal communication, February 11, 2021) to obtain reflectance spectra for each image. Output spectral data (VIS and SWIR) were averaged per organ.

4.2.3. Data processing and outlier removal

Mean reflectance HS data per organ were imported into R software (R Core Team, 2020). Both VIS and SWIR spectra were subjected to a principal components analysis (PCA) model as per Logan et al. (2020). Each dataset was independently visualised using PCA (Q residuals and Hotelling T^2 values) with 2 components using the *mdatools* package (Kucheryavskiy, 2020) to detect outliers defined as observations with orthogonal and score distances > 20 on the residual plot (Kucheryavskiy, 2021). Three outliers were detected and removed from VIS and one from SWIR (Table 12). Subsequently, all datasets were trimmed manually to remove machine artifact effects at the start and end of each spectrum which presented as flat regions. The final spectra for analysis contained wavelengths from 470.5 to 800.5 nm for VIS and 1000.5 to 1600.5 nm for SWIR. A combination dataset (COMB) was created by merging the trimmed VIS and SWIR spectra. To smooth the spectra and avoid spectral

noise, trimmed centred moving average equations were used with a window length of 5 and 20% trim for VIS, whereas SWIR and COMB used a window length of 11 and 10% trim. Cubic polynomial Savitzky-Golay filters (Savitzky & Golay, 1964; Stevens & Ramirez-Lopez, 2020) with identical window lengths were also fitted but did not smooth the spectra as effectively as centred moving average and were therefore not considered. Both reflectance (R) and absorbance (A = 1/log(R) as per Lanza, 1983) spectra were subsequently pre-processed using first (d1) and second (d2) derivatives, with all these datasets used to develop subsequent classification models of the organs. All spectral datasets (R, Rd1, Rd2, A, Ad1, Ad2 for VIS, SWIR and COMB) were centred and scaled before model development. Data processing was implemented using the *tidyverse* suite of packages (Wickham et al., 2019).

Table 12. Description and number of bovine and ovine parenchymatous organs used to develop automatic identification algorithms from visible (VIS) and short-wave infrared (SWIR) hyperspectral sensors following removal of outliers.

Organ type	VIS	SWIR	COMB
Heart	32	32	31
Kidney	20	20	20
Liver	28	29	28
Lung	21	22	21
Total	101	103	100

4.2.4. Statistical modelling

4.2.4.1. Classification model development

Classification models using spectral data from three datasets (VIS, SWIR, COMB) and six pre-processing treatments (R, Rd1, Rd2, A, Ad1, Ad2) were tuned using leave-one-out cross-validation (LOOCV). The choice of LOOCV was primarily due to the relatively low sample sizes. The PLS-DA and RF methods used the *pls* and *randomForest* functions within the *Caret* package (Kuhn, 2020) to differentiate organ type. Model metrics for goodness-of-fit were evaluated using the multi-class summary in the *Caret* package (Kuhn, 2020). Model tuning was achieved using a number of components (*ncomp*) ranging

from 1 to 25 for PLS-DA and number of variables available for splitting at each tree node (*mtry*) between 300 to 500 for RF (Kuhn, 2008; Liaw & Wiener, 2002), based on the highest accuracy and the lowest log loss, respectively, on the LOOCV data. Plots for *ncomp* were visually assessed for the minimal *ncomp* to reach the peak in order to prevent overfitting of PLS-DA models. After the optimal tuning parameters were obtained, the final model was run using the *pls* package (Mevik et al., 2020).

Accuracy, precision, sensitivity, specificity and coefficient of agreement (Kappa) were the model metrics obtained by resampling the PLS-DA and RF discrimination models using LOOCV (Williams, 2001; Williams et al., 2019). The best model among all datasets with six pre-treatments was selected based on LOOCV accuracy and Kappa for determination of the in-sample accuracy (Kuhn, 2008). Following this, the sensitivity, specificity, precision and balanced accuracy were obtained per organ and HS sensor following PLS-DA and RF modelling. Sensitivity corresponds to the inverse of the out of bag error for each organ. Wavelength variable importance (scaled from 0 to 100) of the COMB dataset was determined using the *varImp* function in the *Caret* package (Kuhn, 2020).

4.2.4.2. Principal components analysis model development

Principal components analysis (PCA) modelling of the three datasets was completed and visualised using the R package *ggfortify* (Horikoshi & Tang, 2016; Tang et al., 2016).

4.3. Results

4.3.1. Reflectance spectroscopy

Visible reflectance spectra for the four organs are shown in Fig. 9a, with heart and lung showing greater reflectance compared to liver and kidney across the range between 500 and 850 nm. Liver and kidney had similar spectral signatures throughout the VIS spectrum except between 500 and 600 nm where kidneys had slightly greater intensity. Similarly, hearts showed greater intensity compared to lungs between 500 and 600 nm but both organs showed similar intensity between 600 and 850 nm. Much stronger separation of the spectra occurred in the SWIR region, particularly between 1050 to 1300 nm where lung showed greater reflectance than heart, followed by liver and finally kidney with the lowest intensity (Fig. 9b).

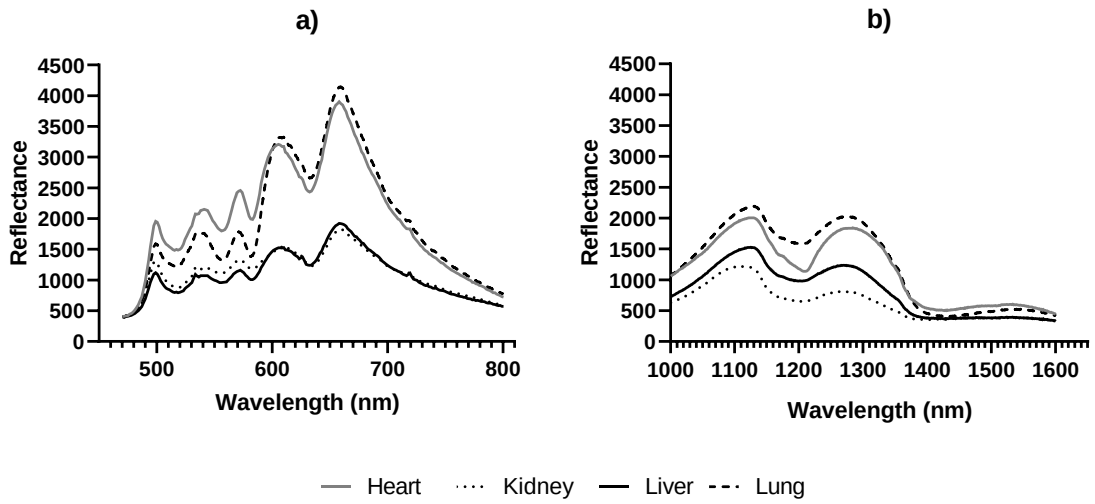

Fig. 9. Trimmed centred moving average **a)** visible (470.5–800.5 nm); and **b)** short-wave infrared (1000.5–1600.5 nm) spectra for livestock organs by organ type (heart, kidney, liver, and lung).

4.3.2. Partial least squares discriminant analysis

4.3.2.1. Visible spectra

Classification results of the different mathematical pre-processing methods using PLS-DA are shown in Table 13. For VIS spectra, absorbance (A) was selected as the best model because it showed 88% accuracy and 84% Kappa on the LOOCV dataset (*ncomp* = 9) for PLS-DA. However, the accuracy of reflectance (R) was similar to A. The in-sample accuracy for A was 96% with only four samples misclassified (Table 14). First and second derivatives of R and A yielded slightly lower accuracy compared to the raw data. All hearts were correctly classified using the VIS spectra along with 18 of 20 kidneys, 27 of 28 livers, and 20 of 21 lungs (Table 14).

4.3.2.2. SWIR spectra

For SWIR spectra, first derivative of absorbance (Ad1) was selected for PLS-DA because of the highest accuracy on the LOOCV dataset (accuracy 92%, Kappa 90%, *ncomp* = 21), although this did not differ from raw A or second derivative of A of the SWIR spectra (Table 13). In-sample accuracy of 98-99% indicated overfitting therefore *ncomp* was reduced to 6 where accuracy was slightly reduced to 88%. Three hearts were misclassified as lungs, four lungs as livers and two livers as kidneys (Table 14).

4.3.2.3. Combination VIS and SWIR spectra

Combination of VIS and SWIR spectra resulted in very high overfitting where all datasets resulted in 100% in-sample accuracy and all *ncomp* ≥ 16 (Table 13). The *ncomp* was reduced to 6 giving in-sample accuracies ranging from 79 to 93% (data not shown). The Ad1 treatment was selected as it had the highest LOOCV accuracy (88%) and PLS-DA modelling of COMB data correctly classified all 28 livers, 28 of 31 hearts, 18 of 20 kidneys, and 19 of 21 lungs (Table 14).

Table 13. Partial least squares discriminant analysis (PLS-DA) classification accuracy and coefficient of agreement (Kappa, κ) from visible (VIS), short-wave infrared (SWIR) and combination VIS and SWIR (COMB) hyperspectral sensors to differentiate bovine and ovine hearts, kidneys, livers, and lungs using various pre-processing methods on the leave-one-out cross-validation (LOOCV) and in-sample datasets.

Spectra		LOOCV dataset			In-sample dataset	
	ncomp[1]	Precision	Accuracy	κ	Accuracy	κ
VIS						
R	11	0.87	0.87	0.83	0.96	0.95
Rd1	16	0.85	0.85	0.80	1.00	1.00
Rd2	6	0.77	0.78	0.70	0.90	0.87
A	**9**	**0.88**	**0.88**	**0.84**	**0.96**	**0.95**
Ad1	11	0.86	0.86	0.81	0.99	0.99
Ad2	9	0.81	0.82	0.76	0.96	0.95
Afinal	**9**	**0.88**	**0.88**	**0.84**	**0.96**	**0.95**
SWIR						
R	24	0.91	0.91	0.88	0.99	0.99
Rd1	20	0.91	0.90	0.87	0.99	0.99
Rd2	18	0.91	0.91	0.88	0.99	0.99
A	24	0.92	0.92	0.90	0.98	0.97
Ad1	**21**	**0.92**	**0.92**	**0.90**	**0.98**	**0.97**
Ad2	18	0.92	0.92	0.90	0.98	0.97
Afinal	**6**	**0.92**	**0.92**	**0.90**	**0.88**	**0.84**
COMB						
R	21	0.94	0.94	0.92	1.00	1.00
Rd1	24	0.97	0.97	0.96	1.00	1.00
Rd2	20	0.98	0.97	0.96	1.00	1.00
A	22	0.94	0.94	0.92	1.00	1.00
Ad1	**19**	**0.98**	**0.98**	**0.97**	**1.00**	**1.00**
Ad2	16	0.96	0.96	0.95	1.00	1.00
Afinal	**6**	**0.88**	**0.88**	**0.84**	**0.93**	**0.91**

[1] number of components selected for PLS-DA; R – reflectance; Rd1 – first derivative of reflectance; Rd2 – second derivative of reflectance; A – absorbance; Ad1 – first derivative of absorbance; Ad2 – second derivative of absorbance; **bold** indicates the dataset used for reduced *ncomp* and subsequent determination (Afinal or Rfinal).

Table 14. Performance of visible (VIS), short-wave infrared (SWIR) and combination VIS and SWIR (COMB) hyperspectral sensors in identifying the type of organs using partial least squares discriminant analysis on the in-sample dataset.

Spectra	Predicted number of each organ			
	Heart	Kidney	Liver	Lung
VIS				
Heart	32	1	0	0
Kidney	0	18	0	0
Liver	0	1	27	1
Lung	0	0	1	20
Accuracy (%)	100	90	96	95
SWIR				
Heart	28	0	0	1
Kidney	1	19	2	0
Liver	0	1	27	4
Lung	3	0	0	17
Accuracy (%)	88	95	93	77
COMB				
Heart	28	1	0	0
Kidney	1	18	0	0
Liver	0	1	28	2
Lung	2	0	0	19
Accuracy (%)	90	90	100	90

Table 15 shows the goodness-of-fit metrics for VIS, SWIR and COMB for all organs. All metrics across organs ranged between 77 and 100% with VIS producing better results than SWIR, being 90% or more across metrics. Hearts and livers were the best classified organs by VIS and COMB, showing sensitivity, specificity, precision and accuracy above 90%. On the other hand, SWIR showed the greatest accuracy and sensitivity in classifying kidneys.

Table 15. Livestock organ classification from hyperspectral sensors using partial least squares discriminant analysis (PLS-DA) for visible (VIS), short-wave infrared (SWIR) and combination VIS and SWIR (COMB) hyperspectral sensors on the in-sample dataset.

VIS (A)	Heart	Kidney	Liver	Lung
Sensitivity	1.00	0.90	0.96	0.95
Specificity	0.99	1.00	0.97	0.99
Precision	0.97	1.00	0.93	0.95
Accuracy	0.99	0.95	0.97	0.97
SWIR (Ad1)				
Sensitivity	0.88	0.95	0.93	0.77
Specificity	0.99	0.96	0.93	0.96
Precision	0.97	0.86	0.84	0.85
Accuracy	0.93	0.96	0.93	0.87
COMB (Ad1)				
Sensitivity	0.90	0.90	1.00	0.90
Specificity	0.99	0.99	0.96	0.97
Precision	0.97	0.95	0.90	0.90
Accuracy	0.94	0.94	0.98	0.94

(A) - raw absorbance data; (Ad1) – first derivative of absorbance data. These data were selected due to the highest accuracy on the LOOCV dataset (Table 13).

4.3.2.4. Variable importance for PLS-DA

Variable importance for PLS-DA predictions of each organ type by combined spectral model (VIS and SWIR using Ad1 pre-processing) is shown in Fig. 10. Hearts showed the greatest importance at 540 nm; kidneys at 780 and 1100 nm; livers at 580 and 780 nm; and lungs at 580–600 nm. Livers showed greater variable importance peaks than lungs and kidneys, which in turn were generally higher than hearts, although hearts tended to exhibit more stability across the spectra.

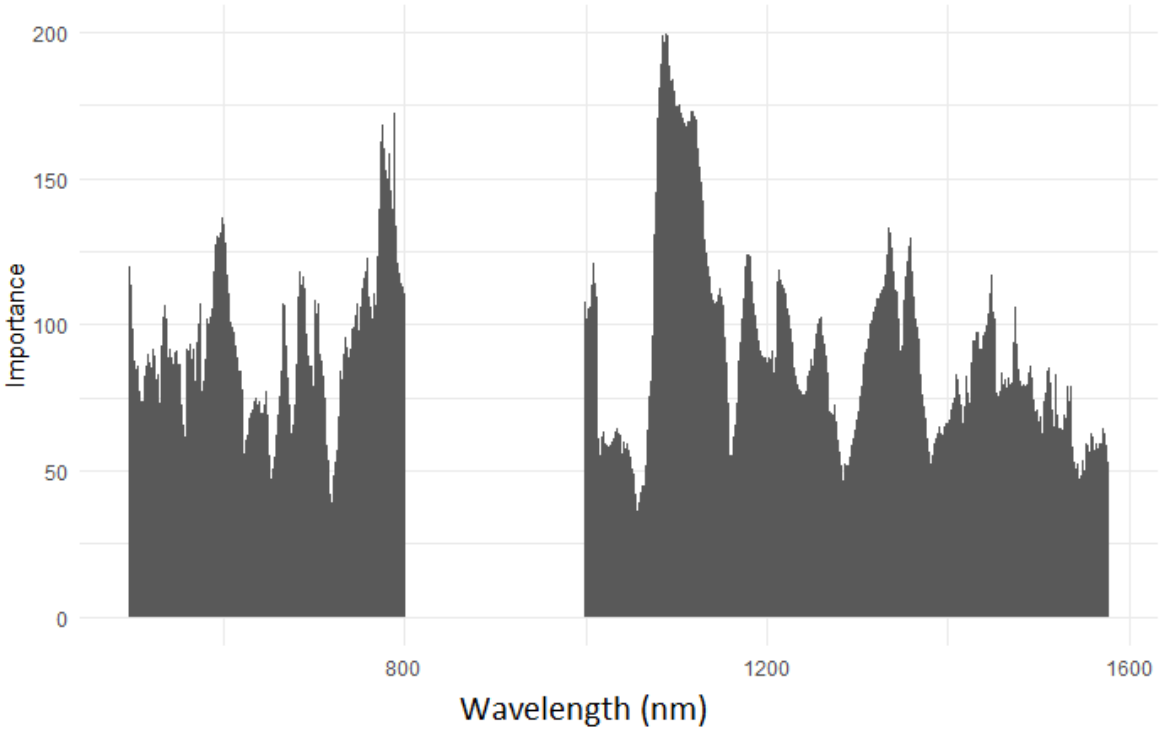

Fig. 10. Variable importance for combination of visible (470.5–800.5 nm) and short-wave infrared (1000.5–1600.5 nm) spectra using partial least squares discriminant analysis to identify bovine and ovine organs by type.

4.3.3. Random forest

4.3.3.1. Random forest modelling on VIS, SWIR and COMB

For RF, SWIR and COMB had Ad1 selected with the highest accuracy on the LOOCV dataset (all ≥ 85%) however the accuracy across different datasets showed an accuracy between 68 and 89% (Table 16). Visible HS data had Rd2 selected with the highest accuracy on the LOOCV dataset. The *mtry* ranged from 310 to 490 with no overfitting. In-sample accuracies were slightly less than accuracies on the LOOCV dataset for all spectral data. Misclassification occurred by all spectra for each organ type (Table 17). Table 17 shows overall accuracies to be highest for livers using SWIR and COMB, and for hearts using VIS data. Table 18 shows the RF in-sample classification metrics of organ type, where livers were the best classified overall by RF (sensitivity > 89%, accuracy > 90%). However, hearts showed greater

sensitivity and equal accuracy and precision than livers by RF using VIS data. Lungs were the least correctly classified organ regardless of spectral data. However, all sensitivities and balanced accuracies were greater than or equal to 80% regardless of organ.

Table 16. Random forest algorithm classification accuracy and coefficient of agreement (Kappa, κ) from visible (VIS), short-wave infrared (SWIR) and combination VIS and SWIR (COMB) hyperspectral sensors to differentiate bovine and ovine hearts, kidneys, livers, and lungs on the leave-one-out cross-validation (LOOCV) and in-sample datasets.

Spectra	LOOCV dataset				In-sample dataset	
	mtry	Precision	Accuracy	κ	Accuracy	κ
VIS						
R	410	0.72	0.72	0.62	0.69	0.58
Rd1	310	0.80	0.81	0.75	0.80	0.73
Rd2	**480**	**0.87**	**0.86**	**0.81**	**0.84**	**0.78**
A	480	0.71	0.72	0.62	0.68	0.57
Ad1	440	0.86	0.85	0.80	0.82	0.76
Ad2	450	0.82	0.81	0.75	0.80	0.73
SWIR						
R	320	0.82	0.82	0.75	0.81	0.74
Rd1	500	0.84	0.83	0.78	0.82	0.75
Rd2	340	0.83	0.83	0.78	0.84	0.79
A	330	0.82	0.82	0.75	0.82	0.75
Ad1	**490**	**0.86**	**0.85**	**0.80**	**0.84**	**0.79**
Ad2	370	0.83	0.83	0.78	0.83	0.76
COMB						
R	310	0.83	0.83	0.77	0.83	0.77
Rd1	460	0.87	0.87	0.82	0.85	0.80
Rd2	440	0.86	0.86	0.81	0.82	0.76
A	390	0.84	0.84	0.78	0.82	0.76
Ad1	**450**	**0.90**	**0.89**	**0.85**	**0.87**	**0.82**
Ad2	310	0.86	0.85	0.80	0.85	0.80

mtry – number of nodes available for random sampling at each split when developing tree models; R – reflectance; Rd1 – first derivative of reflectance; Rd2 – second derivative of reflectance; A – absorbance; Ad1 – first derivative of absorbance; Ad2 – second derivative of absorbance; **bold** indicates the dataset used for final determination.

Table 17. Livestock organ classification from hyperspectral sensors using random forest (RF) classification for visible (VIS), short-wave infrared (SWIR) and combination VIS and SWIR (COMB) hyperspectral sensors on the in-sample dataset.

Spectra	Predicted number of each organ			
	Heart	Kidney	Liver	Lung
VIS				
Heart	29	1	1	4
Kidney	0	16	2	1
Liver	1	3	25	1
Lung	2	0	0	15
Accuracy (%)	91	80	89	71
SWIR				
Heart	27	2	0	0
Kidney	2	17	0	0
Liver	0	1	27	6
Lung	3	0	2	16
Accuracy	84	85	93	73
COMB				
Heart	27	1	0	3
Kidney	0	17	0	0
Liver	0	2	27	2
Lung	4	0	1	16
Accuracy (%)	87	85	96	76

Table 18. Random forest model metrics per organ for visible (VIS), short-wave infrared (SWIR) and combination VIS and SWIR (COMB) hyperspectral sensors on the in-sample dataset.

VIS (Rd2)	Heart	Kidney	Liver	Lung
Sensitivity	0.91	0.80	0.89	0.71
Specificity	0.91	0.96	0.93	0.98
Precision	0.83	0.84	0.83	0.88
Accuracy	0.91	0.88	0.91	0.84
SWIR (Ad1)				
Sensitivity	0.84	0.85	0.93	0.73
Specificity	0.97	0.98	0.91	0.94
Precision	0.93	0.89	0.79	0.76
Accuracy	0.91	0.91	0.92	0.83
COMB (Ad1)				
Sensitivity	0.87	0.85	0.96	0.76
Specificity	0.94	1.00	0.94	0.94
Precision	0.87	1.00	0.87	0.76
Accuracy	0.91	0.93	0.95	0.85

(Rd2) – second derivative of reflectance data was selected; (Ad1) – first derivative of absorbance data was selected. These data were selected due to the highest accuracy on the LOOCV dataset (Table 16).

4.3.3.2. *Random forest variable importance*

Combination VIS and SWIR (Ad1 pre-processing) variable importance of RF modelling is shown in Fig. 11 on an overall basis rather than by organ types. In comparison to the PLS-DA clearer peaks were seen across the VIS and SWIR spectra at 590–600, 740–780, 1080–1140, and 1180–1200 nm.

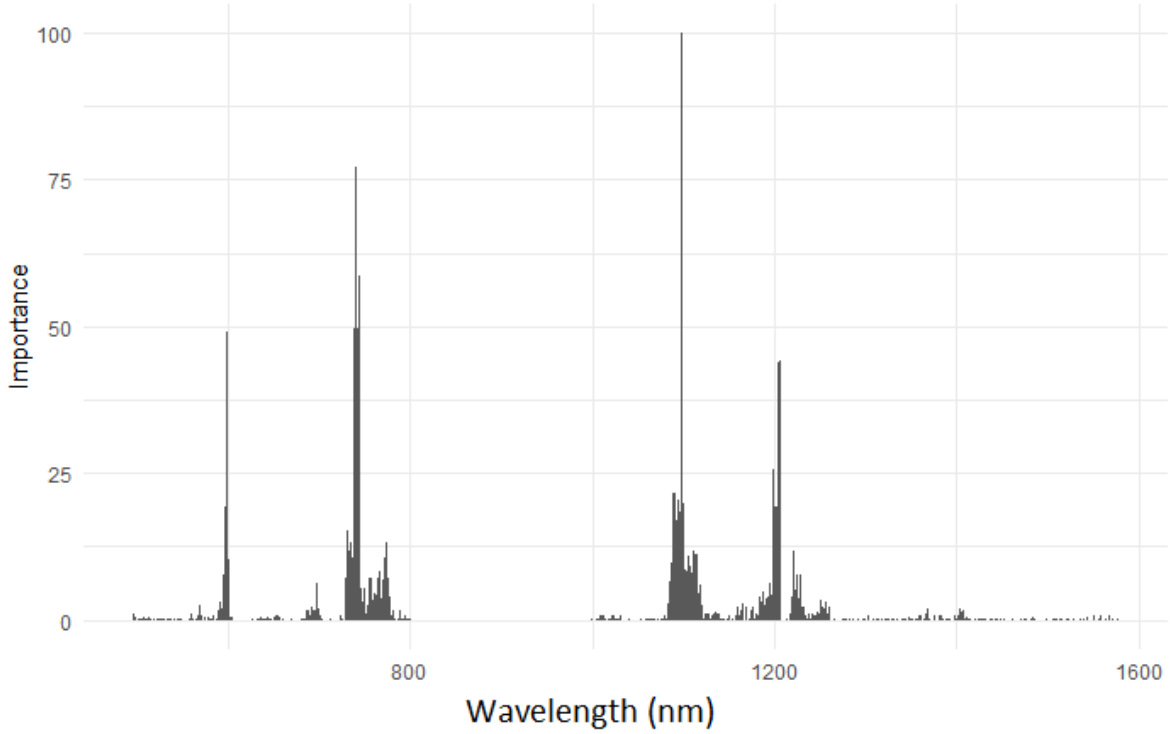

Fig. 11. Variable importance for combination of visible (470.5–800.5 nm) and short-wave infrared (1000.5–1600.5 nm) spectra using random forest modelling to identify bovine and ovine organs by type.

4.3.4. Principal components analysis

The PCA score plots for PC1 against PC2 for Ad1 data were selected for both VIS and SWIR spectra because this series resulted in the highest accuracy. Fig. 12 visually demonstrates that the datapoints cluster together for each organ type. The PCA showed similar results to PLS-DA and RF, where different organs showed different spectral features for identification and hearts tended to be the most clustered organ with the least overlapping.

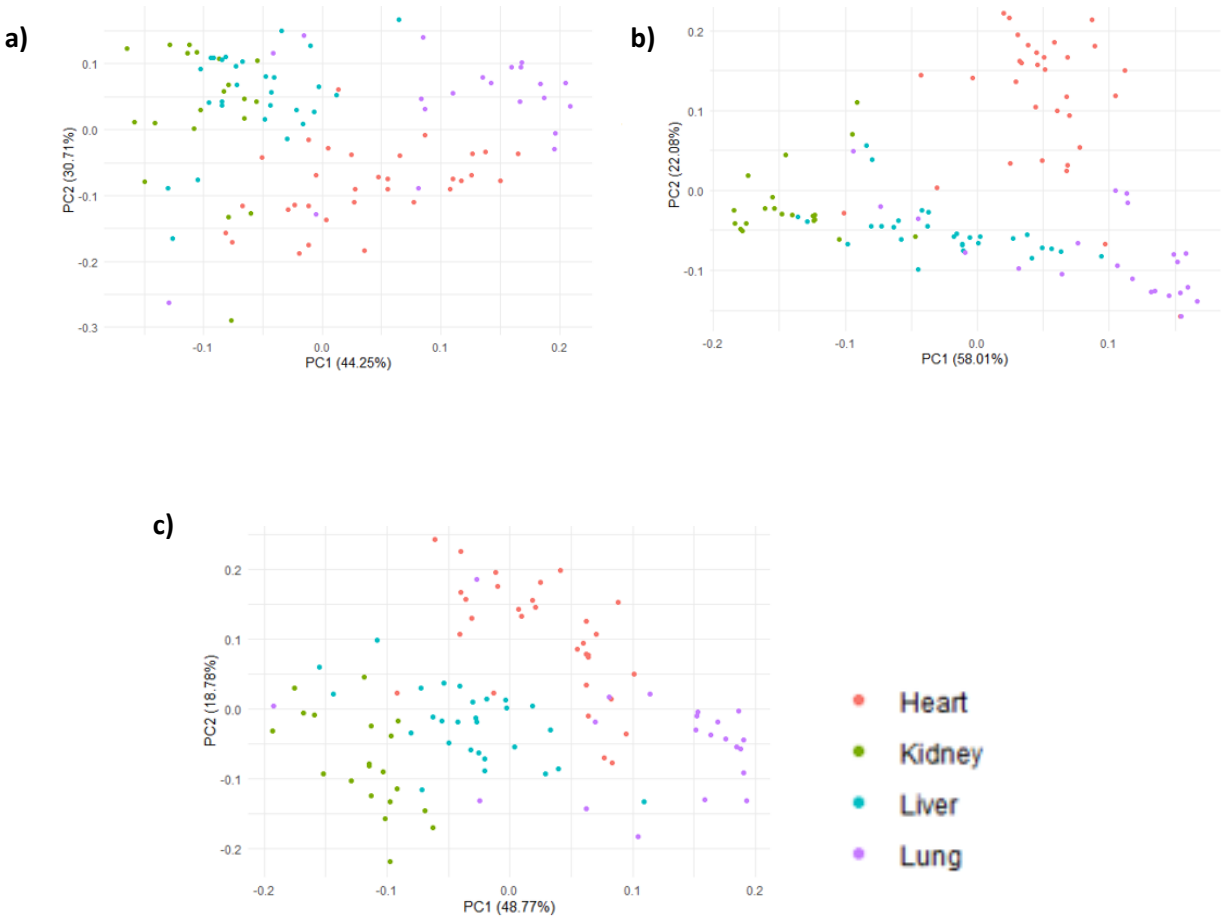

Fig. 12. Principal components analysis (PC1 vs. PC2) plots for the first derivative of absorbance hyperspectral data of **a)** visible (VIS); **b)** short-wave infrared (SWIR); and **c)** combination VIS and SWIR spectra for classification of bovine and ovine parenchymatous organs by type.

4.4. Discussion

The objective of the present study was to explore the potential of VIS and SWIR hyperspectral data to classify parenchymatous organs by type (heart, kidney, liver and lung), to assess the suitability of different data pre-processing techniques for HS data, and to compare RF modelling with the more conventionally used PLS-DA. This study was an exploration of the use of HS imaging in organ identification and inspection within the meat processing industry, potentially leading to automation and quality control. Results demonstrated that automated classification for organ type could be done correctly 95% of the time for VIS and 87% for SWIR using PLS-DA without overfitting. These, in addition to the RF accuracy of 85% for automated classification using the combination of VIS and SWIR sensors, highlights the promise for potential uses of a multi-sensory platform in the beef and sheep meat industries. Potential applications include automated animal organ identification and sorting, processing using robotics, and quality assurance replacing tedious manual procedures normally done manually by meat inspectors and veterinarians (Webber et al., 2012). However, it is important to note that the present study dealt with classification of organs from both sheep and cattle together. The objective of the present study was to differentiate organs independent of origin. However, it is important to note that large-scale processing plants or abattoirs either slaughter one or the other species, whereas smaller ones often slaughter both species. The differentiation of species using HS sensors is a potential avenue of exploration, though was not undertaken in the present study because it was not the objective.

Both VIS and SWIR sensors produced similar accuracy to classify organs by type although VIS was slightly better and more consistent across pre-processing and classification methods. It was expected that SWIR would be superior given it comprises a good portion of the near-infrared spectrum which is known to be able to detect C-H and N-H bonds (Osborne et al., 1993; Prieto et al., 2017). Results from Baeten et al. (2007) and Kamruzzaman et al. (2014) showed SWIR to perform superior to VIS regions to differentiate fruits and agri-food, and the amount of offal addition to a meat mixture, respectively. However, these studies used only one HS sensor encompassing both regions and one

classification method (PLS-DA) was used in these studies. The combination of the two spectral regions (VIS and SWIR) in portable and benchtop devices were also promising in studies of meat quality (Prieto et al., 2014a; Prieto et al., 2017) and microbial spoilage in fish fillets (Cheng & Sun, 2015). The use of such non-invasive devices in the meat processing industry has been hailed for a long period (Scholz et al., 2015). However, a lack of consistent accuracy, particularly for quality traits, in conjunction with the high costs of installation, has held such technologies back from industry adoption (Dixit et al., 2017; Prieto et al., 2017). The present study presents a different application for such technologies, whereupon organ differentiation can take place objectively within the abattoir and the cost of qualified staff could offset the installation and maintenance cost of the multi-sensory platform. Furthermore, such a platform could also add further value to the data collected by predicting chemical composition, quality control, and detection of health issues as demonstrated in other studies (Ariana & Lu, 2008; Kamruzzaman et al., 2012a; 2012b; Yang et al., 2009).

A novel aspect of the present study was the use of RF as an alternative classification method and its comparison to the conventional PLS-DA to discriminate organs based on spectral signature. Random forest is a classification algorithm that has found multiple applications because of its efficiency to handle large datasets and achieve high accuracy (Liaw & Wiener, 2002). However, decision tree RF modelling has been sparsely used in HS classification studies of food (Huang et al., 2014a; Kong et al., 2013; Xu & Sun, 2017), in comparison to the commonly used PLS-DA and linear discriminant analysis (LDA) (Baeten et al., 2007; Dixit et al., 2017; Elmasry et al., 2012a; Huang et al., 2014a). Positive results for RF classification were found in the present study, with accuracy compared to PLS-DA being very similar on the LOOCV dataset and slightly lower on the in-sample dataset. For COMB, RF produced greater LOOCV and in-sample accuracies than PLS-DA. Kong et al. (2013) found that RF modelling was superior to PLS-DA when classifying rice seed cultivars in the SWIR spectrum. In most previous qualitative studies with HS, PLS-DA has been used instead of RF and other classification methods such as PCA, LDA, support vector machines, band ratio and artificial neural networks (Huang et al., 2014a) because PLS-DA provides a combination of partial least squares

regression and LDA (Ariana & Lu, 2010). Nevertheless, LDA has shown high accuracy (>94%) in discriminating pork quality classes (Liu et al., 2010) and insect-damaged from normal wheat kernels (Singh et al., 2009) using HS imaging. One study compared PLS-DA and LDA for SWIR spectral data (950–1650 nm) to classify geographic origin of Chinese lamb and obtained classification accuracies of 89% and 75%, respectively, on the validation dataset (Sun et al., 2012). These results were similar to those from the present study comparing PLS-DA and RF.

Findings from the present study indicate VIS and SWIR imaging methods, used alone and in combination, produced high accuracy for the spectral differentiation of individual organs. However, these results are to be interpreted with caution due to the small sample size. Larger trials with larger sample size are required to build on this pilot study for differentiation of organs within a multi-species abattoir or supply chain, where one or two incorrectly classified organs will not severely affect the model metrics. For instance, Cozzolino and Murray (2004) showed very similar results to the present study when differentiating meat by species, although SWIR and COMB (94–96%) showed more correct classification than VIS (85%).

The present study provided good overall accuracy (>80%) when using VIS and SWIR HS sensors individually and in combination to differentiate organs by type. However, one limitation was the time required to download, mark-up, extract and analyse the spectral data, all processes that can be automated based on the results of the present study. Despite a rapid scanning time (5–6 s), image sizes of 48 KB for VIS frames and 13 MB for SWIR frames were large, with some scans having up to 1000 frames downloaded per scan which can be time and space consuming. This excess consumption may require a high-performance computer and the time required for image download processing may slow the uptake of these automated HS processes in commercial conditions, as any technology would need to be run at chain speed (Gardner et al., 2018). Similar issues with image file size were reported by Elmasry et al. (2012a). However, these limitations could be easily overcome and the whole process could be fully automated. The small sample size of the present pilot study resulted in no testing of the calibration model against an independent dataset, although LOOCV is an accepted and widely used

method for model evaluation in VIS and SWIR HS studies (Kamruzzaman et al., 2012a; Konda Naganathan et al., 2008; Xu & Sun, 2017). The automatic segmentation of an image into "organ" and "background" would allow for organ identification based on shape analysis and for discriminant analysis of each ROI for classification. Such programming would be needed prior to deployment in the industry and a larger training library is also essential in improving accuracy of the model, particularly regarding beef offal which was small in the present study.

Other characteristics that could be examined and quantified by HS imaging to assist with the uptake of these technologies include protein, fat, and mineral concentration of organs. Prior studies have investigated the prediction of these parameters with SWIR HS imaging using lamb meat (Kamruzzaman et al., 2012b), and near-infrared reflectance spectroscopy has also been prevalent to varying levels of accuracy (Lanza, 1983; Prieto et al., 2017). However, the use of ground meat as opposed to intact meat for the most successful of these studies has similarly slowed the progress of uptake in processing plants (Dixit et al., 2017).

The differences in reflectance and absorbance intensity between organ types arise from differences in the chemical composition, colour, and tissue morphology, which provide a spectral signature to each organ (Akbari et al., 2009; Xu & Sun, 2017). Biel et al. (2019) found that of livers, hearts and kidneys, livers had the most protein, P and K; hearts had the most fat; and kidneys had the most Ca and Na. In a study on lamb offal nutritional composition, lungs had significantly more moisture and Fe than the other organs, whereas heart had more fat, liver more Zn and kidney more Na (Bester et al., 2018). This may correspond to the findings of the present study where hearts and lungs had stronger reflectance than livers and kidneys. However, a study on pork found that Raman spectral reflectance was higher in heart, followed by kidney and lowest in liver (Hu et al., 2017) which agrees with the present study.

4.5. Conclusion

The present pilot study showed that visible, short-wave infrared, and combined HS sensors with wavelengths between 400 and 1700 nm could be implemented in a multi-sensory imaging system with the potential for future commercial applications at chain speed. Both sensors were similarly accurate, and the system was very effective at differentiating livestock organs by type with good accuracy and sensitivity. The PLS-DA algorithms were slightly more accurate for differentiation compared to RF and data pre-processing methods did not provide significant advantages, except for first derivative. Improvements in sample size and in streamlining the analytical process to provide information in real-time could allow such systems to be deployed into the meat processing industry as control tools for authentication of livestock organs by organ type. The value of the system could potentially increase by including other characteristics such as identification of species and disease, contamination, and other quality control outcomes.

5. A preliminary investigation into the automatic detection of diseased lamb organs using hyperspectral imaging

Abstract. The post-mortem inspection process of livestock viscera at abattoirs is expensive and gruelling, but it is essential for the detection and condemnation of edible organs and carcase due to food safety issues. Lesions in hearts, kidneys, livers, and lungs are amongst the most common offal defects found in abattoirs. Visible (VIS) and short-wave infrared (SWIR) hyperspectral imaging implemented in a multi-sensory platform were used to differentiate between sheep parenchymatous organs passed as fit (Healthy, n = 42) or not fit (Diseased, n = 47) for human consumption. Partial least squares discriminant analysis (PLS-DA) and random forest (RF) were used to classify organs as healthy or diseased in heart (n = 28), kidney (n = 15), liver (n = 24), and lung (n = 22). PLS-DA produced equal or greater classification accuracy and sensitivity than RF for all organs except for lung when VIS sensors were used (means 84.4% and 78.3%, respectively). Livers and hearts (86.9%) showed higher accuracy than lungs and kidneys (75.9%). Limited differences occurred between VIS and SWIR sensors, although only one sensor tended to be more accurate compared to a combination of both. SWIR outperformed VIS in accuracy across all organs (84.8% vs. 76.3%), and the combination of VIS and SWIR was also accurate (83.0%). The use of hyperspectral imaging is an attractive proposition for the meat processing industry as a non-invasive imaging technology to detect defects in offal, and it can also provide automatic detection, saving time and labour costs.

5.1. Introduction

The automation of the post-mortem meat and animal product inspection and processing in the abattoir has been sought for a long time. One such area of automation has been the sortation of offal based on fitness for human consumption via the detection of defects, contamination, or infectious disease (Thomas-Bachli et al., 2014; Webber et al., 2012). In the abattoir, such examinations are carried out by meat inspectors under the supervision of a veterinarian (Webber et al., 2012; Wilson et al., 2019). This inspection is of vital public health importance due to the removal of potential zoonotic diseases from the processing chain and thereby limiting human exposure (Butler et al., 2003). In addition, this inspection has economic benefits in surveillance of diseases and providing feedback to producers, which could be made quicker with automation (Thomas-Bachli et al., 2014). Furthermore, previous research in sheep found correlation between lung lesions (pleuritis, abscessation and pneumonia) with reduced daily gain, reduced carcase fat, and longer time to reach optimal carcase weight, which along with potential condemnation, affect returns to producers (Goodwin-Ray et al., 2008; Jones et al., 1982; Lacasta et al., 2008). These findings highlight the economic benefits of automatic detection and rapid reporting of animal health to producers.

In recent decades, automation in the meat processing industry has become prevalent (Nade et al., 2005; Scholz et al., 2015; Toohey et al., 2018). Similarly, veterinary medicine studies have used CT to detect abnormalities in anaesthetised cattle prior to euthanasia and organ removal (Lee et al., 2009; Lee et al., 2011). The use of modern non-contact sensor technology to automate the process of meat and offal inspection to detect health and safety hazards has been suggested, with expectations to provide a more rapid, accurate, and sensitive measurement, with greater maintenance of food safety (Neethirajan et al., 2017; Uzal et al., 2002; Webber et al., 2012). Sensors also provide automatic decision making and sorting tools for characterisation and deviation detection in processing plants (Neethirajan et al., 2017). Non-invasive inspection can also prevent potential spread of zoonotic diseases or cross-contamination of infectious diseases between carcases or organs (Samuel et al., 1980; Uzal et al., 2002). Despite this, the post-mortem scanning of internal organs has

been scarce, with high costs and practicality affected by the physical size of imaging systems presenting constraints to their commercial uptake (Neethirajan et al., 2017; Scholz et al., 2015; Webber et al., 2012). Similarly, image size and the slow speed of computers present a constraint when a large number of organs need to be scanned on a commercial production line (Chapter 4; Elmasry et al., 2012a).

The use of hyperspectral (HS) imaging, which can be split into visible (VIS: 400–900 nm) and short-wave infrared (SWIR: 900–1700 nm), has been successfully trialled previously to detect microbial spoilage in fish (Cheng and Sun, 2015), skin tumours in poultry (Nakariyakul and Casasent, 2009), and offal contamination in meat mixtures (Kamruzzaman et al., 2014). The latter studies all showed precision and accuracy greater than 90%. The suitability of HS imaging for the meat and food industries has been the subject of several reviews (Baeten et al., 2007; Elmasry et al., 2012a; Huang et al., 2014a; Xu and Sun, 2017). Prediction algorithms can be developed to differentiate animal tissues and abnormalities from measurements of size, texture, colour, shape, and spectral signatures of regions of interest (ROI) (Elmasry et al., 2012a; Xu and Sun, 2017). However, none of these reviews examined HS sensor technology as a tool to detect diseases in sheep organs.

Diseases commonly encountered post-mortem in parenchymatous organs in include liver fluke (*Fasciola hepatica*) and cysticercosis (*Cysticercosis tenuicollis*) in the liver, caseous lymphadenitis (CLA; *Corynebacterium pseudotuberculosis*) and pneumonia (*Pasteurella spp.* and *Mycoplasma spp.*) in lungs, cysticercosis (*Cysticercus ovis*) in the heart, and interstitial nephritis in the kidney (AHDB, 2017). In addition, hydatid cysts (*Echniococcus granulosis*), abscesses, and haemorrhage can occur in several organs. These animal health issues lead to condemnation of the organ for human consumption and may even result in condemnation of the entire carcase. For instance, the carcase is usually condemned if *C. ovis* cysts are found in three or more locations, or if emaciation occurs concurrently with liver cysts or CLA, or secondary infections such as septicaemia occur from pneumonia (AHDB,

2017). Poor body scores and excessive abscessation can also lead to entire carcase condemnation, which result in significant economic losses (Arsenault et al., 2003; Uzal et al., 2002).

The present study selected parenchymatous organs (hearts, kidneys, livers, and lungs) from sheep that were and were not deemed fit for human consumption by the inspectors at a commercial abattoir. Organs were scanned using a multi-sensory platform encompassing VIS and SWIR sensors, and then examined grossly by veterinary pathologists to confirm abnormalities. These spectral data were then analysed using both partial least squares discriminant analysis (PLS-DA) and random forest (RF) as per Chapter 4 to classify each organ as healthy or diseased. It was hypothesised that both PLS-DA and RF would have good and comparable accuracy in classifying the organs as healthy or diseased.

5.2. Materials and Methods

All offal used in the present study were sourced from a collaborating abattoir and animal ethics approval was not necessary.

5.2.1. Sample collection and storage

A total of 89 sheep parenchymatous organs were collected from a collaborating abattoir in New South Wales, Australia, following commercial slaughter. Organs were processed as per Australian guidelines including palpation and incisions if needed (Wilson et al., 2019) and were stored chilled (1–4 °C) when not examined or scanned. Organs were described as hearts, kidneys, livers, or lungs, and either healthy or diseased, with those diseased considered as not fit for human consumption (Table 19). All organs were transported from the abattoir to the laboratory, scanned entire with the multi-sensory platform, and then examined for abnormalities by experienced veterinary pathologists.

5.2.2. Scanning procedure

All organs were scanned using a non-contact multi-sensory imaging platform encompassing HS technology and multi-energy X-ray attenuation (Rapiscan Inspection System AK198, Rapiscan Systems Pte Ltd, Singapore) as previously described (see section 4.2.1 and Fig. 8). The multi-sensory system included a conveyor belt (6.64 s, 1260 mm, 0.19 m/s) and it was run by an Ubuntu (Linux) Cube computer program, which controlled exposure time, image size, and acquisition rate. Two HS sensors covered the spectral range from 400–900 nm (VIS) and 900–1700 nm (SWIR). The VIS (Basler Ace GigE, Photonic Science, East Sussex, UK) and SWIR (Snake A/C GigE v3 AK081, Photonic Science, East Sussex, UK) sensors were powered by 12 V power supply units and fitted with Specim spectrographs (VNIRV10E and NIR V17E, respectively) and a Grade 1 InGaAs detector with air cooled housing. Spectral resolutions were 3 and 5 nm for VIS and SWIR, respectively, with both sensors capturing 200

spectral slices per second. Spectral increment (1.5 nm, VIS 300 bands, and SWIR 512 bands), sensor resolutions (VIS 1920 x 1200 px; SWIR 640 x 512 px), frame rate (150 fps), refresh rate (150 Hz, 6.66 ms) and exposure times (VIS 6.4 ms; SWIR 4.0 ms) were used as per Chapter 4 (see section 4.2.1).

5.2.3. Organs' gross examination

All organs were systematically examined by veterinary pathologists for gross lesions. Should a lesion be present the following data were recorded: location, distribution, demarcation, colour, shape, appearance of the cut surface, and consistency. The most likely cause of the lesion was also recorded to the organ in question. Organs were also examined for off-colours and inconsistency in texture by palpation, with sectioning and sampling for histopathology in some instances, to confirm the identity of the lesions.

5.2.4. Image analysis and extraction of spectral data

As per Chapter 4 (see section 4.2.2), HS images (both VIS and SWIR) were downloaded as 200–300 slices in PNG format, with trimming done to remove slices beyond 10 frames either side of the tray the samples were scanned in. MATLAB programming language (MATLAB R2021a, Mathworks Inc., USA) was used to construct complete organ images from individual frames, with these images viewed using Gnu Image Manipulation Program (GIMP) software (version 2.10.18; GIMP Development Team, 2020) and marked-up using ImageJ (version 1.53a; National Institutes of Health, Bethesda, MD). Three to eight evenly spaced ROI (7 x 7 px, 10.2 mm width x 9.0 mm length) were marked-up for each organ depending on the size of the organ. Pixel dimensions of the ROI were written into the MATLAB algorithm referenced in Chapter 4 (section 4.2.2) to extract spectral data. All HS data (both VIS and SWIR) were averaged for each organ, with organs classified by their organ type (heart, kidney, liver, or lung) and their status (Diseased or Healthy, based on the classification at the abattoir and corroborated by the pathologists).

5.2.5. Outlier removal and spectral trimming

Subsequently, mean reflectance spectral data from VIS, SWIR, and combination of both sensors (COMB) were imported into R (version 4.0.2; R Core Team, 2020). A principal components analysis (PCA) model using the *mdatools* package (Kucheryavskiy, 2020) was used to visualise each dataset and identify outliers with orthogonal and score distances > 20 on the Q residual plot (Kucheryavskiy, 2021). Two organs (one diseased heart and one diseased lung) were removed from the VIS and COMB datasets because these were detected as outliers. As per the data pre-processing methods of Chapter 4 (see section 4.2.3), the spectra between 470.5–800.5 nm (VIS), and 1000.5–1600.5 nm (SWIR) were retained with those outside these ranges presenting as flat endpoints.

In the present study comparisons were made between the disease statuses of different organs, therefore all datasets (VIS, SWIR, and COMB) were analysed separately according to their organ type (heart, kidney, liver, and lung). As per the results obtained by Chapter 4 (see section 4.2.3), all reflectance spectra were smoothed using the trimmed centred moving average method with a sliding window (length = 11) and then transformed to obtain the first derivative of absorbance where absorbance = log(1/reflectance). All data processing was done in R (R Core Team, 2020) with the assistance of the suite of *tidyverse* packages (Wickham et al., 2019).

5.2.6. Classification model development

Classification models were developed using two machine learning methods within the *Caret* package (Kuhn, 2020) i.e., partial least squares discriminant analysis (PLS-DA) and random forest (RF). Resampling for model tuning and evaluation of all datasets was done using the LOOCV method and all datasets were centred and scaled. The number of components (*ncomp*) for PLS-DA models of each organ type were selected according to the highest accuracy. The number of variables available for

splitting at each tree node (*mtry*) used for the RF models was selected using the lowest log loss (Kuhn, 2008; Liaw and Weiner, 2002).

Model evaluation for diseased or healthy status was done by fitting tuned models using the LOOCV of each of the four organ types, three spectral ranges, and both classification methods. For the estimated models, goodness-of-fit model metrics including sensitivity, specificity, precision, accuracy, and coefficient of agreement (Kappa) were generated with Diseased being the positive class.

5.3. Results

5.3.1. Abattoir sortation and post-mortem

All organs labelled by the abattoir as healthy or fit for human consumption (n = 42), were similarly labelled by the veterinary pathologists with no detectable lesions. However, 16 of the 47 organs labelled as diseased at the abattoir presented no detectable gross lesions other than discolouration and were therefore rejected for human consumption (Table 19). Discolouration was the most common and predominant reason for rejection of hearts and kidneys (73.7%), whereas mineralisation in the liver (36.4%), CLA (35.3%) and pneumonia (29.4%) in lungs were also significant in the present study (Table 19).

Table 19. Gross interpretation of lesions found in sheep organs after slaughter.

Lesion	Heart	Kidney	Liver	Lung	Total
Haemorrhage	1	1	1	1	4
Mineralisation	0	0	4	0	4
Pneumonia	n/a	n/a	n/a	8	8
Caseous lymphadenitis	0	0	1*	6**	7
Pyelonephritis	0	1	0	0	1
Atelectasis	0	0	0	2	2
Myocarditis	1	0	0	0	1
Hepatitis	0	0	3	0	3
Abscessation	0	0	0	2	2
Myocarditis	1	0	0	0	1
Bronchiectasis	0	0	0	1	1
Discolouration but NDL***	10	4	2	0	16
Total Diseased	**13**	**6**	**11**	**17**	**47**
Total Healthy	**15**	**9**	**13**	**5**	**42**
Total	**28**	**15**	**24**	**22**	**89**

* Found in the hepatic lymph node.

** Found in the mediastinal lymph node.

*** NDL – no detectable lesions. Some lungs showed more than one lesion.

5.3.2. Spectral data and model tuning

Figs. 1 and 2 show the transformed VIS and SWIR absorbance spectra, respectively, split by organ type. The VIS spectra showed no complete separation of diseased organs from healthy organs (Fig. 13). However, the region between 600 and 660 nm showed the greatest separation between healthy and diseased status, particularly for hearts and livers, with diseased organs showing greater absorbance than healthy organs at 600 to 620 nm, and the reverse occurring from 620 to 660 nm. The SWIR spectrum of organs showed a more evident and longer region of healthy and diseased organ differentiation than VIS (1020–1320 nm), regardless of organ type (Fig. 14). Across this range, the clearest differentiation between healthy and diseased organs occurred in kidneys, where absorbance was greater for healthy than diseased kidneys.

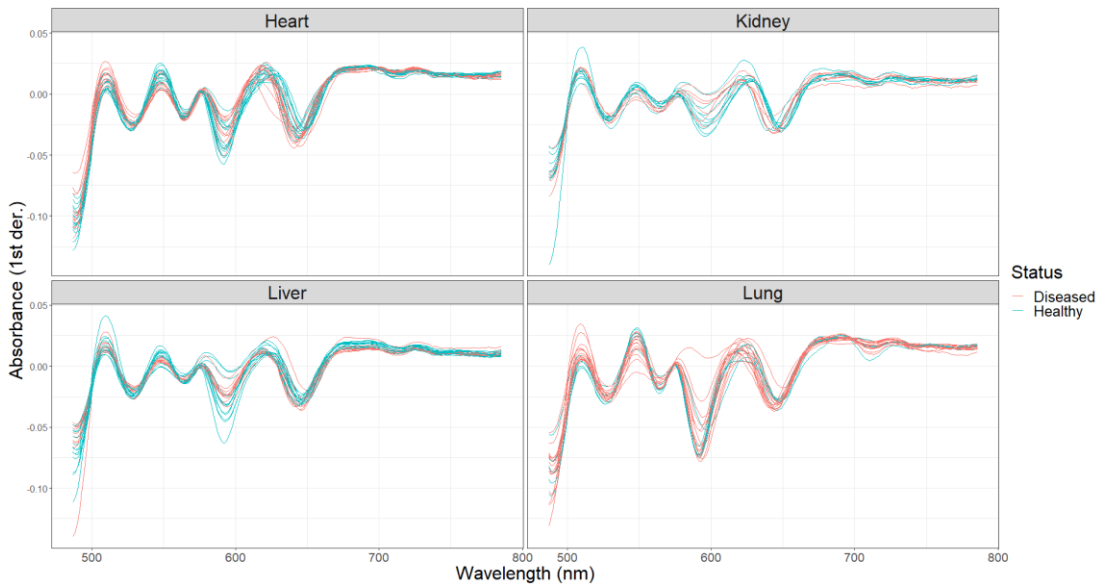

Fig. 13. Transformed (first derivative) and smoothed (centred moving average) visible absorbance hyperspectral data of healthy and diseased sheep organs.

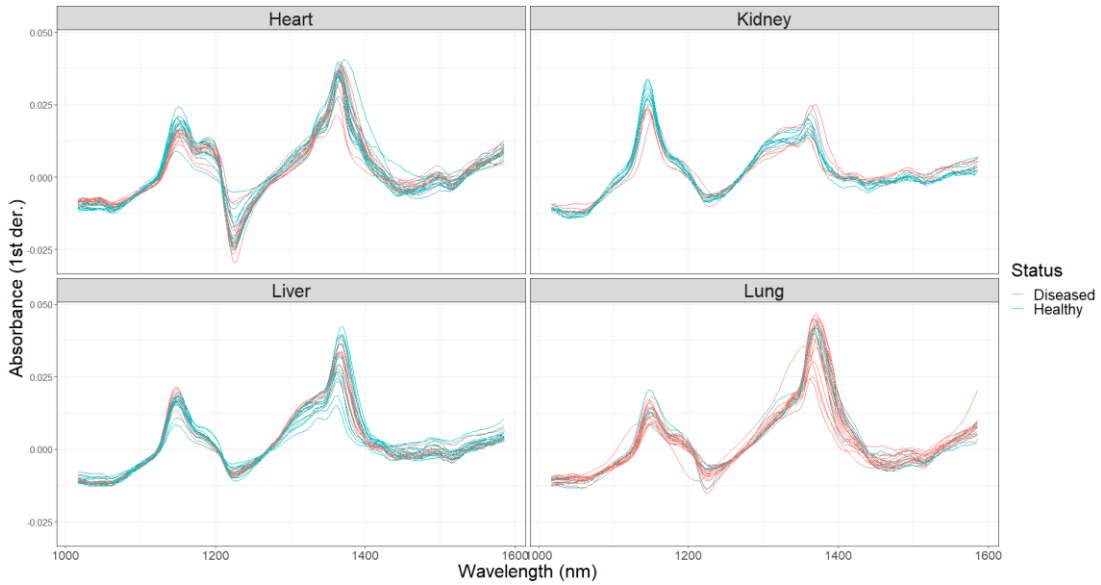

Fig. 14. Transformed (first derivative) and smoothed (centred moving average) short-wave infrared absorbance hyperspectral data of healthy and diseased sheep organs.

Due to the small sample size, *ncomp* of the PLS-DA models tended to be small, particularly for liver and lung, with lung being the most unbalanced organ dataset in terms of numbers of healthy and diseased (Table 20). Conversely, heart was the most abundant organ, and all PLS-DA models used *ncomp* of 5 or more (Table 20). It is worth noting that *ncomp* of 10 was originally selected for COMB differentiation of liver, although this model overfit with 100% for all goodness-of-fit metrics on the training dataset but much lower on the validation dataset (data not shown). The smallest organ in total number (kidney) had the lowest *mtry* values for SWIR, though the highest for VIS and COMB. Meanwhile, the most abundant (heart) had the greatest *mtry* values for SWIR, and the lowest for COMB (Table 20).

5.3.3. Classification of organs as healthy or diseased

Table 20 shows the results for differentiating sheep organs based on healthy or diseased status using three different spectra (VIS, SWIR, and COMB) and two different classification methods (PLS-DA and

RF). Hearts and livers were correctly classified as diseased or healthy using PLS-DA above 85% regardless of the sensor used. All diseased lungs were correctly identified as diseased using SWIR and PLS-DA, whereas all healthy kidneys were correctly identified as healthy using the same methods. The combination of the two sensors predicted disease status better than both individual sensors with PLS-DA modelling for livers, and with RF modelling for hearts. However, COMB did not outperform both individual sensors for any other organ or classification method.

Classification and detection of diseased kidneys tended to be the weakest among the organs used in the present study, showing sensitivities below 67% and accuracies below 87%. In contrast, lungs showed the lowest specificity across sensors and models with RF modelling below 20% and Kappa below 18%.

Regarding machine learning methods, PLS-DA outperformed RF in classification accuracy except for differentiating healthy and diseased lungs using the VIS sensor (Table 20). However, the majority of these differences were small, or RF was equal in accuracy to PLS-DA, which occurred for differentiating healthy from diseased hearts using COMB and livers using SWIR.

Table 20. Goodness-of-fit statistics from hyperspectral (HS) data used to differentiate between diseased or healthy organs of sheep from commercial slaughter.

Spectral range	Method	Statistic					
		Tuning	Accuracy	Kappa	Sensitivity	Specificity	Precision
Heart							
VIS	PLS-DA	5	0.85	0.70	0.75	0.93	0.90
	RF	430	0.78	0.54	0.67	0.87	0.80
SWIR	PLS-DA	7	0.89	0.79	0.92	0.87	0.86
	RF	480	0.82	0.64	0.77	0.87	0.83
COMB	PLS-DA	5	0.85	0.70	0.83	0.87	0.83
	RF	300	0.85	0.70	0.83	0.87	0.83
Kidney							
VIS	PLS-DA	7	0.73	0.44	0.67	0.78	0.67
	RF	470	0.60	0.17	0.50	0.67	0.50
SWIR	PLS-DA	6	0.87	0.71	0.67	1.00	1.00
	RF	320	0.73	0.41	0.50	0.89	0.75
COMB	PLS-DA	1	0.80	0.57	0.67	0.89	0.80
	RF	470	0.73	0.41	0.50	0.89	0.75
Liver							
VIS	PLS-DA	4	0.88	0.75	0.82	0.92	0.90
	RF	450	0.83	0.66	0.73	0.92	0.89
SWIR	PLS-DA	2	0.92	0.83	0.91	0.92	0.91
	RF	410	0.92	0.83	0.91	0.92	0.91
COMB	PLS-DA	4	0.96	0.92	0.91	1.00	1.00
	RF	390	0.88	0.75	0.82	0.92	0.90
Lung							
VIS	PLS-DA	1	0.67	-0.16	0.88	0.00	0.74
	RF	330	0.76	0.17	0.93	0.20	0.79
SWIR	PLS-DA	6	0.91	0.70	1.00	0.60	0.90
	RF	430	0.73	-0.08	0.94	0.00	0.76
COMB	PLS-DA	2	0.81	0.39	0.94	0.40	0.83
	RF	380	0.76	0.17	0.94	0.20	0.79

VIS visible; SWIR – short-wave infrared; COMB – combination VIS and SWIR; PLS-DA – partial least squares discriminant analysis; RF – random forest; Tuning – number of components (*ncomp*) for PLS-DA and number of nodes available for each tree splitting (*mtry*) for RF.

5.4. Discussion

In examination of the organs, it was determined that CLA was the most common disease of sheep resulting in offal rejection for human consumption in the present study (35.3%), which was consistent with Australian sheep findings reported by Webber et al. (2012). Pneumonia and CLA made up the majority of rejected lungs (64.7%), whereas the most common disease for rejected livers was focal parenchymal mineralisation (36.4%). On the other hand, most of the condemned hearts and kidneys in the present study were rejected due to discolouration. From a food safety perspective, CLA dominance provides limited risk to the entire carcase (Murray, 1986) or human health if there are no open wounds (Arsenault et al., 2003). However, labour associated with slicing to detect CLA lesions and abscesses, carcase trimming, and disinfection, has been mentioned as associated losses for processors, along with reduced performance and animal deaths for producers (Arsenault et al., 2003). For this reason, the automatic detection and subsequent excision of CLA lesions could enhance the post-mortem process.

The measurement of instrumental colour (CIE L*, a*, b*) and detection of its changes have been previously explored using HS imaging and computer vision (Chen and Kim, 2004; Elmasry et al., 2012b; Xu and Sun, 2017). For instance, SWIR HS (964–1631 nm) was used successfully with instrumental colorimetry to measure colour changes in salmon fillets (Wu et al., 2012). Such technologies could be similarly applied using different organs, although the colorimeter would provide contact with the organs. Hyperspectral analysis can detect different surface textures within the same organ (Xu and Sun, 2017), which could assist with identifying specific regions of disease such as CLA within an organ, and also detect specific diseases that may be present. However, more samples would be required to develop a reliable training model that can be evaluated with an independent dataset rather than using resampling as used in the present study. In addition, more samples possessing one particular pathology or lesion type, such as CLA in the mediastinal lymph nodes of the lung (Arsenault et al., 2003; Webber et al., 2012), could also improve the prediction models for identification of specific diseases and locations within organs.

The collection of several condemned organs with similar pathologies may also improve HS prediction models and allow them to detect additional lesions and identify specific diseases and their locations within organs. For instance, CLA in the mediastinal lymph nodes of the lung could be used for this based on the results of the present study and others (Arsenault et al., 2003; Webber et al., 2012). In turn, this could reduce the costs of veterinary inspection in the abattoir, particularly given the lack of zoonotic diseases (e.g., liver fluke) encountered in the present study.

Hyperspectral prediction models were better at detecting diseased than healthy lungs as reflected by the high sensitivity and low specificity. However, this would likely mean healthy lungs would be marked as diseased by the model. The other organs tended to have higher specificity than sensitivity, although the values were much closer than in lungs. These sensitivities showed that, while promising, this system requires substantial improvements (Xu and Sun, 2017). The combination of the two HS sensors spanning VIS and SWIR ranges (400–1700 nm) did not necessarily provide greater accuracy than either of the two sensors separately, which was a finding similar to Chapter 4 of the present thesis where the same sensors were used to differentiate organs by type. Despite the small sample size, the positive goodness-of-fit statistics sorting hearts and livers as diseased or healthy encountered in the present study constitutes a good feasibility study rather than for industry purposes (Williams et al., 2017). However, the use of samples with more gross lesions present may provide greater accuracy and sensitivity, which was true in the present study for livers and lungs, respectively. Larger datasets and more gross lesions could allow for combination with the study presented in Chapter 4 of the present thesis to also discriminate between different organ types as well as between healthy and diseased organs, which would be ideal prior to trialling such technologies in abattoirs.

The use of non-invasive imaging devices could assist the biosecurity procedures in preventing diseases bypassing into markets as a result of human error. In this context, studies on breast cancer screening in humans have reported artificial intelligence models as being more accurate in mammogram performance than human screeners (McKinney et al., 2020). However, similar studies

using artificial intelligence in animal health surveillance are still in their infancy, though improved accuracy and decision making have been noted (Ezanno et al., 2021). The present study did not attain high accuracy for all organs but provides a solid pilot study encouraging a large-scale study to be conducted. Few studies have taken to imaging animal organs, with most that employed non-invasive imaging devices providing information on whole carcase, meat, or skin composition and health (Elmasry et al., 2012a; 2012b; Nade et al., 2005; Scholz et al., 2015). However, some veterinary medicine studies scanned entire live sedated animals scanned with CT for detection of lesions such as abscesses in cattle (Lee et al., 2009; Lee et al., 2011) and screening for cystic echinococcosis in sheep (Mao et al., 2017). Subsequently, animals were euthanised and lesions were examined using histopathology (Lee et al., 2009; Lee et al., 2011; Mao et al., 2017). However, these studies tended to be exploratory in nature with a small number of samples and use of simple regression and Pearson correlation analyses, rather than focusing on surveillance of large numbers of animals entering processing plants with more advanced machine learning modelling.

Regarding machine learning classification methods, PLS-DA tended to be more accurate than RF in the majority of the models, which was also seen in the prior study differentiating beef and sheep organs by type (Chapter 4, see section 4.3). However, the accuracy of RF was not overly lower than PLS-DA and was equal in several models. Random forest modelling is a more modern and alternative classification method to the conventionally used PLS-DA (Huang et al., 2014a; Liaw and Wiener, 2002). Previous findings using RF for classification of rice cultivars (Kong et al., 2013) and in plant ecological studies (Cutler et al., 2007; Lawrence et al., 2006) were positive. However, few studies have employed RF in studies with HS data classifying animal products. In the present study only discrimination of diseased and healthy lungs using VIS showed improved accuracy, sensitivity and precision for RF compared to PLS-DA, with lower specificity and Kappa. Similarly, only kidneys using VIS data showed improved specificity and Kappa for RF compared to PLS-DA, with equal precision and accuracy. Overall, it can be concluded that RF showed promise but was generally lower in accuracy as a machine learning method compared to PLS-DA.

It was an interesting finding that hearts were one of the most accurately classified organs for diseased or healthy status given that only 10 of the 13 hearts condemned by the abattoir had evidence of pathological processes at gross examination as opposed to appearance defects. On the other hand, all condemned lungs and the majority of livers had evidence of clinical diseases associated with them. This finding of more gross lesions present in the liver and lungs, resulting in production losses, was demonstrated by previous studies (Arsenault et al., 2003; Mao et al., 2017). Furthermore, Mao et al. (2017) found that the presence of cysts in both livers and lungs of sheep was highly correlated, and that the use of CT on live sheep successfully identified 86% of the total number of cysts found during a necropsy. The CT showed better results for determining the diameter, location, and type of cyst. Similar results were found in the present study for sheep livers and lungs using SWIR HS and PLS-DA as a classification method (91–92% accuracy). However, the present study did not identify the size, type, or location within the organ of the lesions, which could assist in the extraction of spectral data from the lesions themselves. Therefore, future studies should include an accurate demarcation of the lesion and differentiation by type with larger sample size, which could be further improved with combination of X-ray and HS imaging. In doing so, rapid and objective post-mortem health feedback from processors to producers could be made (Lee et al., 2011; Webber et al., 2012). Hyperspectral imaging could also assist with animal and herd health management post-mortem at a veterinary medicine level (Neethirajan et al., 2017), providing diagnoses and identifying potential disease outbreaks as has been done previously in exploratory studies using CT (Lee et al., 2009; Mao et al., 2017).

5.5. Conclusion

Visible and short-wave infrared HS imaging can be used to determine the disease status of sheep organs, with this pilot study proving that classification accuracy was adequate overall and was particularly successful for livers and hearts. However, it is worth noting that the diseases were often identified with the naked eye, palpation, or were presented as discoloured to abattoir inspectors. Future studies could use HS imaging, alone or in combination with other technologies such as multi-energy X-ray or CT, to identify specific pathologies within individual organs and their locations. However, they would need to use more samples, particularly those presenting similar lesions. This could be very beneficial for the meat processing and veterinary medicine industries. Hyperspectral imaging technologies are non-invasive and non-contact, with the potential to enable automatic sorting of livestock organs and disease detection in the abattoir, allowing animal health reports to be provided for producers.

6. General Discussion and Conclusions

The global meat and livestock industry needs objectivity and automation to measure product quality and detect products that are unsuitable for the intended market (Dixit et al., 2017). The present thesis provided pilot investigations into four key areas, 1) meat quality and chemical composition, 2) authentication of meat by feeding regime, 3) classification of organs by type, and detection of defects, and 4) diseases in organs. Beef and sheep were used in the present thesis due to their high importance in the Australian agricultural industries, consumption, and exports (MLA, 2017a; MLA, 2017b). Retailers and consumers are becoming more aware of food health and safety and demand more information and certainty about them (Pethick et al., 2018; Toohey et al., 2018; Van Elswyk & Mcneill, 2014). The first two experimental chapters of the present thesis evaluated low-cost, handheld devices to objectively measure the chemical composition and feeding regime of meat which could also see applications for retailers and consumers (Chapter 2; Chapter 3). Chapters 4 and 5 explored the use of hyperspectral sensors to automatically identify organs by type and effects for applications in the automation of processes. The sample sizes were small as the pilot investigations were conducted to explore the potential of the technologies and determine if further investment from the meat industries was encouraged.

The general hypothesis and objective of the present thesis was that spectroscopic sensors (NIRS, Raman spectroscopy, HS imaging) could provide objective measures of chemical composition, feeding regime, organ type, and presence of defects or diseases in organs. Except for Chapter 3, which employed a large dataset with equal numbers of grass-fed and grain-fed beef, these hypotheses were proven on small datasets, with accuracy exceeding 80% in most cases. Chapter 5 showed that detection of organs by defect or disease was strongly dependent on organ type, with hearts and livers showing greater accuracy than kidneys and lungs.

6.1. Applications and value of novel spectroscopic sensors

The lack of objective measurements for meat grading has led to several nations developing their own independent grading systems, although the results of some papers have concluded that numerous similarities between these exist (Bonny et al., 2018; Pethick et al., 2018; Polkinghorne & Thompson, 2010; Strong, 2004). The most important objective measure of eating quality found in the review of several studies (Chapter 1, section 1.3.1) was the quantification of intramuscular fat (IMF) due to its effect on tenderness, juiciness, flavour, and overall liking (Frank et al., 2016). The IMF is correlated to subjective marbling scores as demonstrated by Savell et al. (1986), although subsequent studies have been sparse (Bindon, 2004; Cameron et al., 1994; Dow et al., 2011; Kruk et al., 2002; Ueda et al., 2007). The present thesis showed moderate to good precision (r^2_{val} = 0.53–0.81) between chemical and objective measures of IMF using on-line handheld NIR spectroscopy (Chapter 2, section 2.3.3). However, the precision was much stronger on FD and ground meat than FI meat which suggests that sample homogenisation and drying is required for more accurate predictions, as was demonstrated in soil samples (Wadoux et al., 2021). Previous studies measured objective marbling content as a percentage via charge coupled or mirror-type cameras in Wagyu rib eyes (Connolly et al., 2020; Kuchida et al., 2000; Nakahashi et al., 2008). Similarly, red-green-blue camera technology connected to smartphones via Bluetooth have been used commercially to measure marbling scores (MB, MSA and AUS-MEAT) and IMF content in Wagyu rib eyes (Condon, 2020). Meat eating quality probes based on laser reflectance have also been used commercially to measure IMF content (Sim, 2021a). The fact that these devices are non-contact and can provide the information instantly can allow a more streamlined approach into industry uptake. Smartphone NIRS sensors, as used in Chapters 2 and 3 of the present thesis, instead present a valuable screening tool for consumers. However, the precision and accuracy of NIRS to predict IMF, crude protein, moisture content, and pH of commercial FI beef and lamb cuts were low as demonstrated in Chapter 2.

One study alternative that could improve the accuracy of IMF prediction by NIRS sensors is the use of discrimination methods to sort cuts into groups of high, moderate, and low marbling (as

well as CP, moisture, and pH), similar to the statistical analyses performed in Chapters 3–5 (PLS-DA, LDA, RF). Prior on-line NIRS studies have found 78% accuracy for tenderness (shear force) above median (Shackelford et al., 2005) and 90% accuracy for dark cutting (Prieto et al., 2014b) using official methods from the USA and Canada, respectively. Based on these findings and the discriminant analyses in Chapter 3, it is possible to create a reliable non-destructive prediction model for meat quality classes differentiated by IMF concentration and other measurements of chemical composition. However, this approach may need a larger number of samples to allow sortation into classes (i.e., 0–1%, 1–3%, 3–5%, 5–10%, 10–15%, 15+% IMF, or high-medium-low CP) because the results of the present thesis and other studies have shown inconsistent results and low accuracy of NIRS when predicting chemical composition in intact meat treated as linear variables.

Alternatively, hyperspectral (HS) imaging could present a non-contact and accurate alternative to measure eating quality, although the cost and sheer size of installation of such systems in abattoirs is high. However, the advantage of these systems is that they can run automatically and uninterrupted at chain speed constantly measuring quality attributes on a conveyor belt (Hitchman et al., 2021). Therefore, the cost savings in labour and speed to obtain objective measures 'on the fly' and to use these data for automation processes such as sorting of products can offset the installation cost. Chapters 4 and 5 both encountered large image sizes and long processing times, which were mentioned in the review by Elmasry et al. (2012) as barriers to industry adoption. Nevertheless, this is unlikely to be a limitation with recent and constant advances in computing and information technologies. Recent studies have shown that using a multi-sensory system containing both HS and NIRS sensors can predict IMF concentration and pH of intact beef reliably and with high precision and accuracy (r^2_{val} > 0.90; RPD > 3; Dixit et al., 2021), scores consistent with industry determination (Williams, 2001). Similarly, the HS imaging system used in the present thesis has demonstrated at chain speed the successful discrimination of grass-fed and grain-fed frozen beef without contact (Coombs et al., 2021c; Table 21).

Table 21. Preliminary results to classify frozen steaks as grass-fed or grain-fed using two hyperspectral (HS) sensors in a multi-sensory platform using three different machine learning methods.

	Visible HS sensor				
Method	Sensitivity	Specificity	Precision	Accuracy	Kappa
PLS-DA	0.929	0.929	0.929	0.929	0.857
LDA	1.000	0.929	0.933	0.964	0.929
RF	0.857	0.857	0.857	0.857	0.714
	Short-wave infrared HS sensor				
PLS-DA	0.786	0.929	0.917	0.857	0.714
LDA	0.786	0.929	0.917	0.857	0.714
RF	0.786	0.857	0.846	0.821	0.643

Adapted from Coombs et al. (2021c). PLS-DA: partial least squares discriminant analysis; LDA: linear discriminant analysis; RF: random forest.

In addition, the accuracy and applications of multi-sensory platforms could be augmented with extra sensors such as X-ray as in the platform used in the present thesis. Unfortunately, time and length of the thesis did not allow for inclusion of X-ray data, but research is ongoing to determine the potential of additional information from X-ray data to measure product quality and safety. Furthermore, robust HS models for lamb IMF prediction incorporating flock and year of slaughter data are in progress (Hitchman et al., 2021). It is worth noting that Meat Standards Australia (MSA) does not provide marbling grading for sheepmeat (Fowler et al., 2020; MLA, 2012; Toohey et al., 2018) despite its importance to eating quality (Allen, 2021). However, objective automatic marbling score is a trait being investigated for its inclusion in official sheepmeat grading (Allen, 2021; Jacob & Calnan, 2018; Pannier et al., 2018; Sim, 2021a) and a price grid has been launched for additional payments for high IMF (>5%) lamb carcases to producers (Sim, 2021b).

The discrimination of beef cuts into grass- and grain- feeding regime was successful using both NIRS and Raman spectroscopy, with NIRS outperforming Raman on lean surfaces and Raman outperforming NIRS on fat surfaces (Chapter 3). The handheld, smartphone connectable NIRS sensor

was used because its small size, low cost, speed of analysis and lack of outliers would increase its practicality to the meat industry as an authentication tool (Chapter 3). In contrast, the Raman sensor used in Chapter 3 did not pose these advantages to the same extent as the NIRS sensor. Conducting larger studies with more samples would provide greater information on the accuracy and robustness of prediction models which could enhance processor uptake of these technologies. Using the same cut type for comparison of MSA marbling and DOF data may improve upon the very low accuracy of NIRS and Raman in Chapter 3 and can determine if a larger trial is warranted.

Chapters 4 and 5 explored a multi-sensory platform that is designed to mimic airport baggage scanners (Paulus et al., 2017) with the trials explored to determine the efficacy of installation within commercial abattoirs which has been done in pilot plants using dual-energy X-ray attenuation (Gardner et al., 2010; Gardner et al., 2018) and HS imaging (Dixit et al., 2021). These systems primarily focused on carcase composition and meat quality, respectively. In contrast, the present thesis trialled HS imaging to automatically detect and identify organs by type and defects (Chapters 4–5). Whereas Chapter 4 mixed beef and sheep organs to classify by type only, Chapter 5 used only one species and focused on sheep organs. The collaborating abattoir marked several hearts and some kidneys and livers as defective, although a subsequent examination by veterinary pathologists found no detectable lesions and only discolouration. This is a reason to support a naming convention of ‘organs with defects’ instead of ‘diseased organs’ or ‘condemned organs’, which were used in Chapter 5. In contrast, most livers and all lungs were found to possess diseases, though none seemed to be zoonotic. It can be concluded that the potential of HS upon the smaller sample size used in Chapter 5 means that a larger study with equal numbers of diseased and healthy organs is warranted prior to development of methods for use in processing plants.

AgResearch (2018) and Cook and Anderson (2017) conducted similar pilot studies to the results presented in Chapter 5 focused on animal health using X-ray and computed tomography, respectively, although these were not published in peer-reviewed journals nor were they scientifically

analysed. The differentiation of organs by type in the present thesis was successful, although the mixing of beef and sheep organs was unconventional and may have confounded the results. However, the reasons to do this analysis was the small sample size to evaluate the classification models properly and the similarity of the spectral data for organs of both species as shown in Fig. 15 below. It was noted in Chapter 4 that smaller abattoirs in Australia often process multiple species while larger ones tend to focus on one species only (Toohey et al., 2018). For these reasons, the results from an analysis with the same data as Chapter 3 but classifying organs by species and type is presented in the following section.

6.2. Reflectance spectroscopy to identify organs by species and type

The multi-sensory platform used in the present thesis could be used in organ processing scenarios under commercial conditions such as abattoirs or processing plants. One example is where organs are mixed and need to be identified by both species and type. The spectra of each organ and results of classification algorithms to differentiate organs by species and type are presented in this section. Visible (VIS) and short-wave infrared (SWIR) reflectance spectra for each of the four organs and each species are shown in Fig. 15, with hearts and lungs showing higher intensity compared to livers and kidneys in both VIS and SWIR regions. For VIS spectra the difference in intensity was shown to occur between 500 and 850 nm (Fig. 15A and B). Hearts and lungs had similar spectral signatures throughout the VIS spectrum except between 500 and 600 nm where hearts had greater intensity. Similarly, kidneys showed slightly greater intensity than livers between 500 and 600 nm. Much stronger differentiation occurred in the SWIR region, particularly between 1050 to 1350 nm where lung was greatest in intensity, followed by heart, then liver, and then kidney with the lowest intensity (Figs. 1C and 1D). Differences between species were negligible for both spectra (Fig. 15).

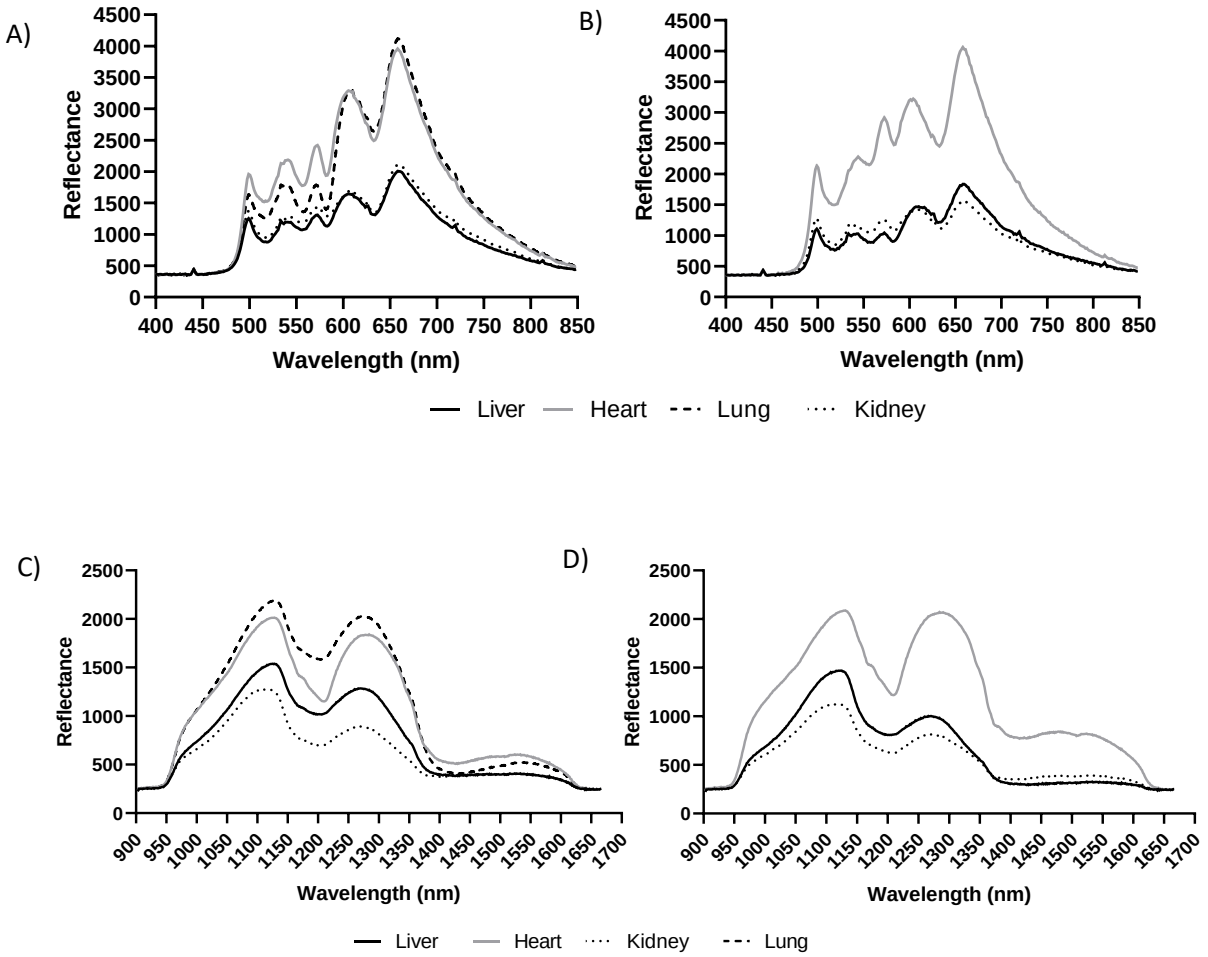

Fig. 15. Mean visible (panel A and B) and short-wave infrared (panel C and D) reflectance spectra for each of the organs from sheep (panel A and C) and beef (panel B and D).

The datasets were pooled across species to develop algorithms for organ and species differentiation as if a mix of organs from both species were scanned through the platform. These algorithms were developed and validated using 5-fold cross-validation. When using both species, the three different spectral regions used in the present study (VIS, SWIR, and combination VIS and SWIR – COMB) each performed best using a different discriminant analysis model for predictions. Predictions from VIS spectral data were most accurate with linear discriminant analysis (LDA) modelling (Fig. 16). In contrast, predictions from SWIR data were most accurate using random forest

(RF) modelling (Fig. 17), and COMB predictions with partial least squares discriminant analysis (PLS-DA) modelling (Fig. 18). However, the three classification methods evaluated with either VIS or SWIR yielded similar accuracies to each other, but COMB was much more accurate with PLS-DA compared to RF and LDA. The PLS-DA model for COMB also had the highest sensitivity, specificity, area under the curve, and Kappa coefficient of agreement (Fig. 18). Using COMB yielded the highest overall accuracy (71–92%) and Kappa (63–89%) compared to the use of each sensor separately, regardless of the discriminant analysis method used. Therefore, these results show that HS imaging is accurate for industry applications to differentiate between species and organs when these are mixed and passed through the platform.

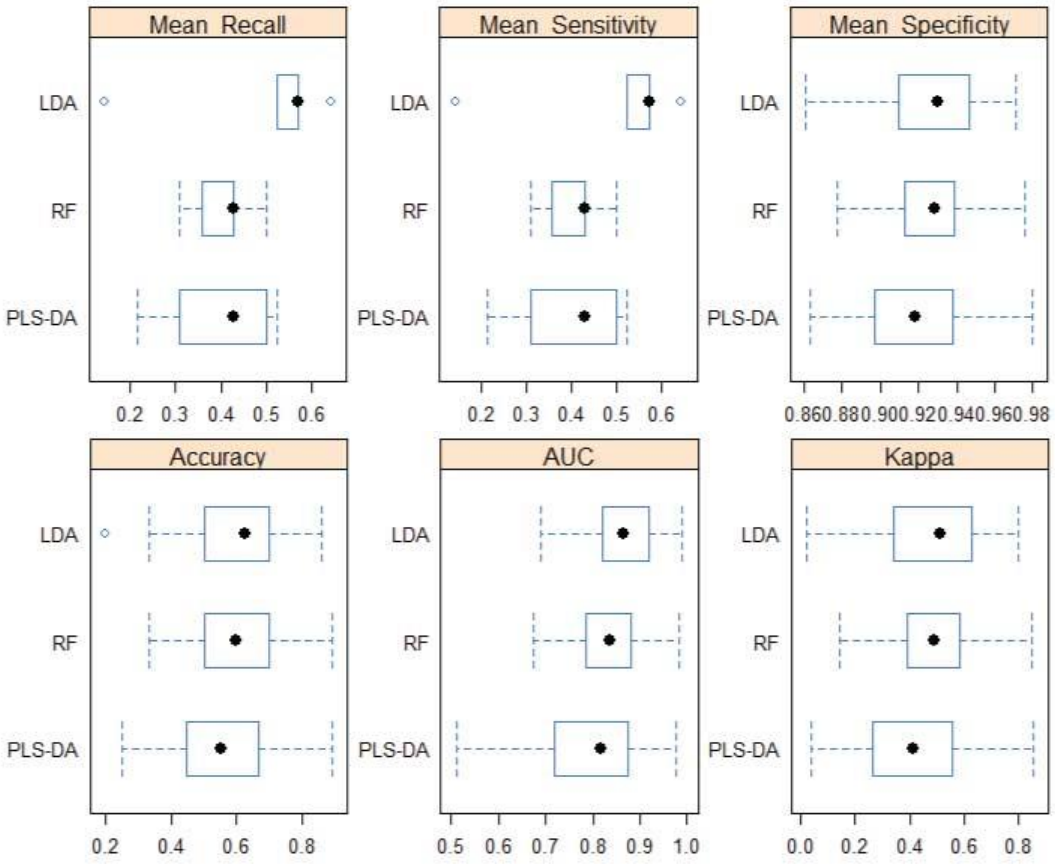

Fig. 16. Model metrics for the classification of organs from sheep and cattle by both species and type using a visible (VIS) hyperspectral sensor with partial least squares discriminant analysis (PLS-DA), linear discriminant analysis (LDA) and random forest (RF).

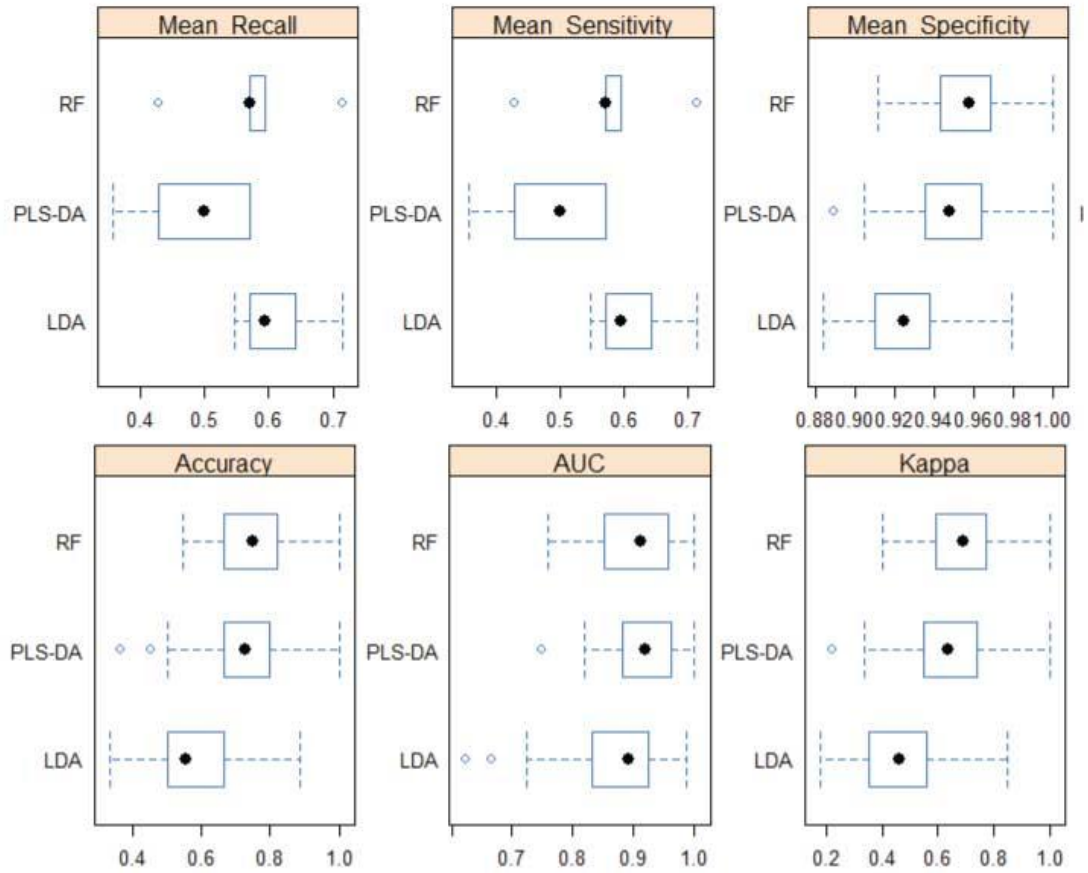

Fig. 17. Model metrics for the classification of organs from sheep and cattle by both species and type using a short-wave infrared (SWIR) hyperspectral sensor with partial least squares discriminant analysis (PLS-DA), linear discriminant analysis (LDA) and random forest (RF).

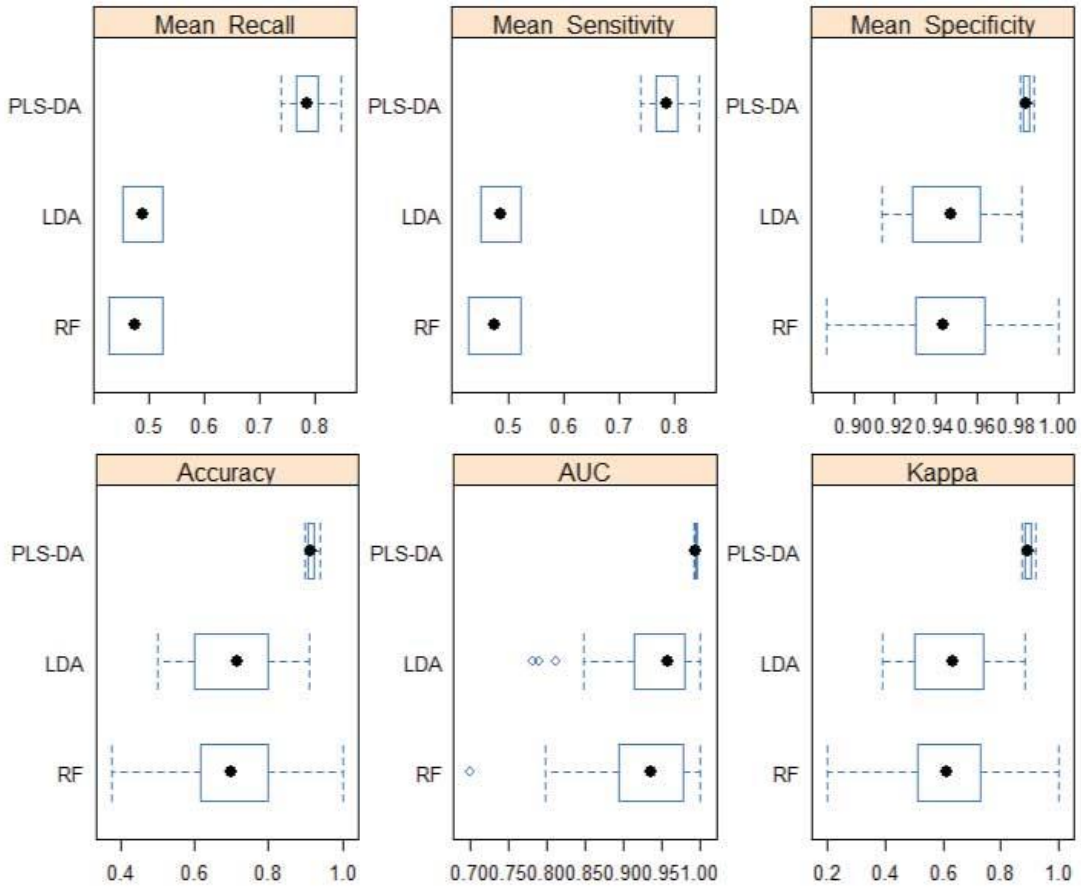

Fig. 18. Model metrics for the classification of organs from sheep and cattle by both species and type using a combination of visible and short-wave infrared hyperspectral sensors (COMB) with partial least squares discriminant analysis (PLS-DA), linear discriminant analysis (LDA) and random forest (RF).

6.3. Data processing

Spectral data pre-processing before prediction model development using different machine learning approaches may be necessary to improve the accuracy of the predictions (Zeaiter et al., 2005). For example, mathematical treatments that may assist in IMF quantification include outlier detection and removal, absorbance transformation, standard normal variate (SNV), detrending and centred moving average (Patel et al., 2021), and multiplicative scatter correction (Goi et al., 2022). The first experimental chapter (Chapter 2) used the Beer-Lambert Law ($absorbance = log(\frac{1}{reflectance})$)

(Zeaiter et al., 2005) followed by Cubist regression to develop prediction models which is commonly used for NIRS data in soil (Minasny & McBratney, 2008). However, there was no attempt to evaluate different statistical approaches for pre-processing the spectra or to use alternate machine learning methods to develop the prediction models. Similarly, different mathematical treatments to pre-process spectra were not explored in Chapter 3, which simply used absorbance spectra for both NIRS and Raman with SNV. Derivatives and smoothing were not used in either Chapter 2 or Chapter 3.

Chapter 1 stated that not every mathematical treatment improves the accuracy of the calibration model followed with leave-one-out cross-validation. However, multiple pre-processing methods should be explored to develop models with optimal performance onto larger data sets. Results for devices encompassing shorter NIR wavelength ranges (900–1700 nm) such as the NIRvascan (Chapters 2–3) and the SWIR HS sensor (Chapters 4–5) have been shown to have comparable accuracy to devices with longer wavelengths (Chapter 2). Furthermore, Shackelford et al. (2004) recommended trimming the spectra outside the vis-NIRS range of 450–1400 nm because it showed more accurate results than 350–2500 nm. Similar results were found after selecting regions of Raman spectroscopy for differentiation instead of the entire Raman spectrum (Logan et al., 2020a; Zając et al., 2014). Short-wave NIR spectra (750–1100; Byrne et al., 1998, and 750–1300 nm; Shackelford et al., 2004; 2005) have been shown to outperform the longer-wave spectra in meat quality and discrimination analyses. However, these shorter wavelength ranges could provide further increases in prediction accuracy if appropriate pre-processing methods are used to remove spectral noise, additive, and multiplicative effects (Dixit et al., 2017). Chapter 4 explored mathematical treatments for spectra pre-processing in the form of first and second derivatives and compared reflectance and absorbance spectra. The Savitzky-Golay filter was trialled although it did not smooth the spectra to the same extent as centred moving average (Fig. 19). In most cases the most accurate model was absorbance with first derivative and as a result this pre-processing method was used in Chapter 5. However, the differences in the accuracy of the predictions with different

data pre-processing methods were not large in most instances and therefore, data pre-processing may not be critical in all cases.

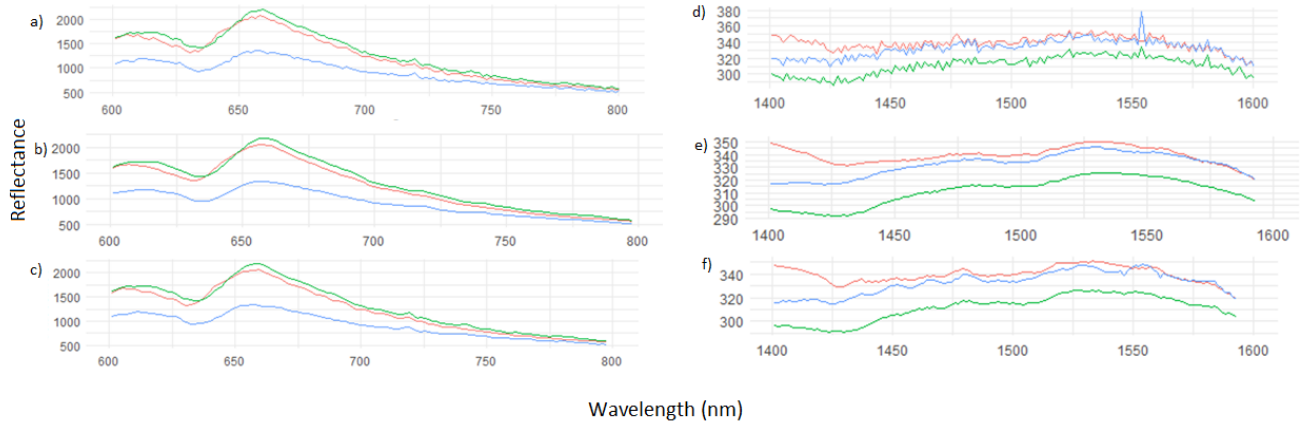

Fig. 19. Spectra from Chapter 4 (n = 3) after undergoing different smoothing treatments: **a)** raw VIS; **b)** centred moving average of VIS; **c)** Savitzky-Golay filtered VIS; **d)** raw SWIR; **e)** centred moving average SWIR; **f)** Savitzky-Golay filtered SWIR.

Wadoux et al. (2021) suggested that the presence of moisture may confound the measurements of chemical constituents of soil using NIRS, with soil products frequently dried prior to chemical determination. Drying of meat is impractical, although it is the method used prior to determination of IMF and CP by AOAC official methods (Helrich, 1990; Chapter 2). Equations developed for removal of moisture effects such as external parameter orthogonalization (EPO) may be applicable for use in fresh meat studies, as these have been previously used in studies on fruit regarding temperature effects (Roger et al., 2003) and soil regarding moisture effects (Minasny et al., 2011). The latter study used laboratory conditioned (air-dried) samples as their calibration dataset and field (fresh) samples as their validation dataset (Minasny et al., 2011). The effect of ambient temperature variations on meat in NIRS studies have not been analysed (Dixit et al., 2017), although maintenance of a constant temperature could improve the precision of predictions, e.g., an

abattoir boning room. Similarly, the selection of Cubist regression modelling in Chapter 2 followed improved predictions compared to PLSR in soil (Minasny & McBratney, 2008; Tang et al., 2020), whereas the use of RF discrimination modelling followed improved classifications of rice cultivars compared to PLS-DA (Kong et al., 2013). Bootstrapping was used as the resampling method for the spectral data in Chapter 2, and this can be additionally used prior to PLSR analyses (known as bagging PLSR) through use of randomly generating multiple datasets and prediction models with the predictions averaged over several bootstrapped samples to improve accuracy (McBratney et al., 2006).

It is clear from the results of the present thesis that prediction of chemical composition of FI meat was the weakest of the pilot experiments conducted because most precision and accuracy metrics (r^2, RPD) did not meet standards for even rough screening ($r^2_{val} < 0.60$; RPD < 2.0; Williams, 2001). Therefore, further research with larger datasets with the same sensor may not be warranted although sorting samples into groups by IMF content and trimming and pre-processing of spectra may improve the accuracy of the prediction models developed. Similarly, MSA marbling score and DOF did not meet these minimum standards for rough industry screening (Chapter 3). This is disappointing due to the potential of the smartphone NIRS sensor as an authentication tool (Chapter 3), and the performance of similar sized sensors for evaluating IMF in prior studies (Chapter 1; Table 2). These sensors (Texas Instruments DLP NIRScan Nano, JDSU Micro NIR Pro, SCiO), each weighing less than 100 g, showed r^2 on the validation dataset between 0.62 and 0.79 for IMF concentration in FI meat (Goi et al., 2022; Patel et al., 2021; Pham et al., 2018). Moisture was predicted between r^2_{val} of 0.70 and 0.84, although CP and pH showed RPD below 2 (Goi et al., 2022; Patel et al., 2021). It remains to be seen whether analysis over several sampling days would improve the accuracy and precision of the prediction models, as suggested in previous studies (Dixit et al., 2017; Williams et al., 2017).

In contrast, the use of a smartphone NIRS sensor to classify beef as grass-fed or grain-fed showed predictions with high accuracy, particularly when scanning the lean surfaces of steaks (Chapter 3). This could allow for a protocol to be developed for processors, retailers, and consumers, supported by the fact that the NIRS sensor in the present thesis scanned steaks through plastic wrap without differences. Goi et al. (2022) used a smaller NIRS sensor and scanned steaks at 1 cm from the meat sample, whereas other studies have determined IMF using non-contact multi-sensory systems (Dixit et al., 2021; Hitchman et al., 2021). The present thesis determined grass- or grain-feeding regime using the non-contact multi-sensory system from Chapters 4 and 5 (Table 21; Coombs et al., 2021c). More studies are required to determine the effects on model statistics between different muscles, and effects on acquiring spectra from regions of interest (ROI) on entire commercial primal cuts (AUS-MEAT, 2005) instead of steaks cut from these, as they are small enough to be scanned through the system yet have increased relevance to processors and retailers. Furthermore, the ROI from the two HS imaging chapters (Chapters 4–5) were relatively small (7 x 7 px), questioning whether these constitute an adequate representative sample. Numbers of ROI per sample were also based on the organ's size. This method could be of use in the future as it allows for specific diseased regions of an organ to be marked-up for the algorithm to identify. Alternatively, algorithms could scan entire organs and then search for abnormal regions, which could be assisted by X-ray spectroscopy. Similarly, identifying different components of an organ sampled from the abattoir such as lymph nodes, fat, and bile ducts could assist in identification of the organ, as well as detection of defects, diseases, or abnormalities. However, this would require larger ROI such as marking-up entire organs from an HS image which has been done in prior studies (Elmasry et al., 2012; Manley, 2014; Nakariyakul & Casasent, 2009). This may be the most likely method to be trialled in subsequent studies and compared against the pilot studies of the present thesis. Another area of future research and development is computer vision to analyse not only spectral signature but also shape analysis and perhaps even 3D reconstruction of scanned organs (Xu & Sun, 2017; Zia et al., 2015). Prior to deployment in the meat industry, further work is also required in obtaining

optimal HS images where the object of interest is entirely visible and clearly seen, which can be time consuming when considering the multitude of sizes, shapes, colours, and a moving conveyor belt providing variation (Elmasry et al., 2012; Prieto et al., 2017). As previously mentioned, the use of multi-energy X-ray data to develop upon existing classification models is also encouraged and a few examples from work in the present thesis are presented in the section below.

6.4. Potential value of X-ray imagery

Hyperspectral sensors can only measure the electromagnetic radiation from the surface of products and thus, cannot measure the characteristics inside organs to detect defects or abnormalities below the surface. However, the multi-sensory platform of the present thesis also collected data from a multi-energy X-ray sensor which can penetrate tissues much further than the devices employed in the prior chapters, where penetration was in the mm range (Zou et al., 2016). The platform contains six X-ray sensors that penetrate at different depths and could identify abnormalities that the HS sensors cannot. Unfortunately, there was limited time in the present thesis to analyse and present the large datasets that were collected. Therefore, this section presents some examples of the data collected to demonstrate potential applications in the meat industries which could be above and beyond the HS sensors.

Based on visual inspection, several multi-energy X-ray attenuation (MEXA) images generated by the multi-sensory platform showed the shape and features of organs, and in some cases, allowed for the marking-up of defects such as caseous lymphadenitis (CLA) (Fig. 20). In Fig. 20B, the MEXA image shows an uncut CLA lesion that was felt during palpation by inspectors at the abattoir and was then confirmed by veterinary pathologists after examination (Fig. 20A). On the other hand, Fig. 21 shows that in other organs such as kidney with pyelonephritis, the defect is not noticeable to the naked eye, although it may become apparent following image analysis or if dissected organs showing the lesion are scanned (i.e., in Fig. 21D). However, the MEXA image of the defective kidney (Fig. 21E)

appears lighter in colour than the healthy kidney passed for human consumption (Fig. 21B). This anecdotal information suggests that diseases may cause a change in tissue density that may be reflected in X-ray absorption. This data could allow the development of algorithms with threshold values for heart, liver, lung, kidney, and different defects that could replace palpation in the abattoir. Interestingly, these differences in absorption have been found to detect foreign bodies in food products (Zou et al., 2016) and therefore, this device could also be used to detect foreign bodies such as needles from vaccinations or other metal bodies. With a more comprehensive dataset, differences in pixel value across an organ or ROI within an organ can assist in detecting organs with defects that may be rejected by the processor. Further preliminary exploration into the use of MEXA for organ differentiation by type and health status could lead to commercialisation of multi-sensory platforms and their installation in processing plants.

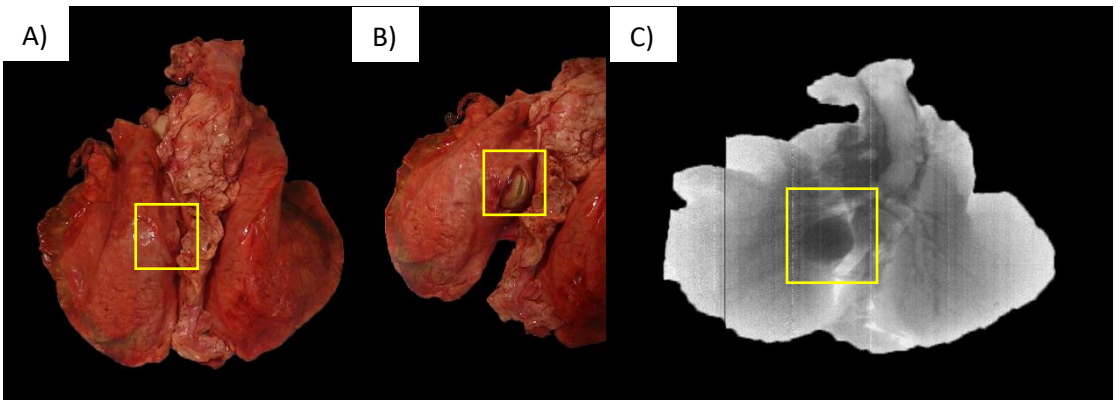

Fig. 20. A) Photograph of unprocessed sheep lung; **B)** Photograph of same sheep lung post-incision; and **C)** Multi-energy X-ray (MEXA) image of the unprocessed sheep lung showing a caseous lymphadenitis (CLA) lesion.

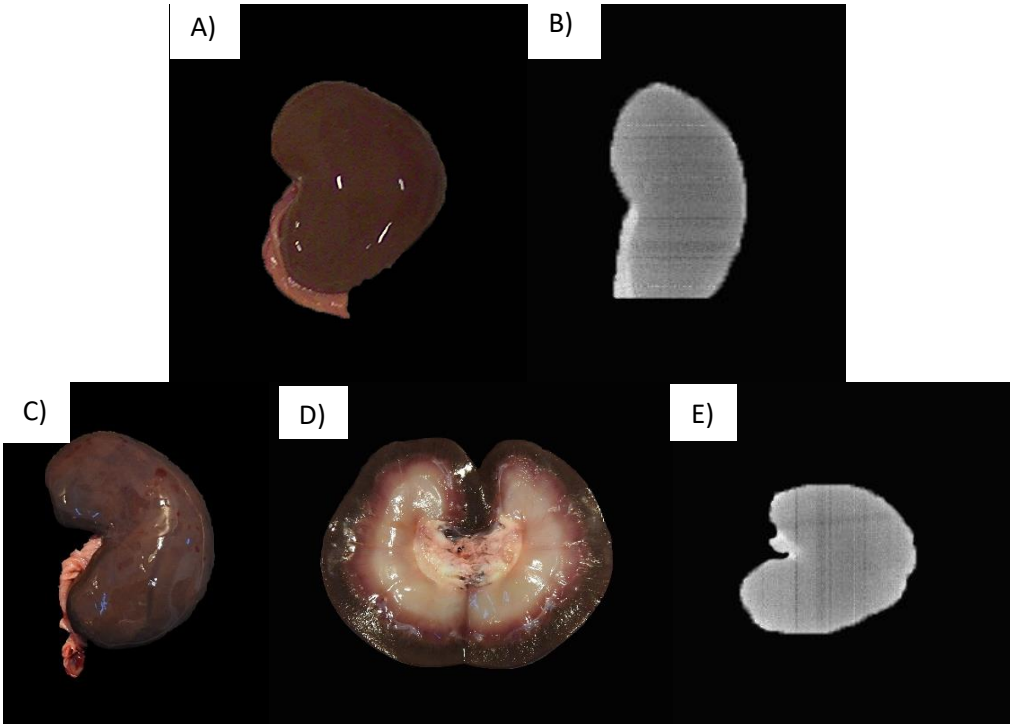

Fig. 21. A) Photograph of sheep kidney passed as fit for human consumption; **B)** MEXA image of sheep kidney passed as fit for human consumption; **C)** Photograph of sheep kidney rejected at the abattoir due to defects; **D)** Photograph of bisected kidney showing focal lesion suggestive of pyelonephritis; and **E)** MEXA image of rejected kidney with pyelonephritis.

Finally, another area of research and development with the multi-sensory platform is the identification of the type of disease rather than just normal vs. abnormal. For example, this could include the differentiation between lung abscesses, CLA, lung consolidation, and pleurisy as described by Blakebrough-Hall et al. (2020). Similarly, abnormalities in livers could be differentiated between fluke worms, abscesses, and different types of hepatitis. This may require more advanced data analytics than those used in the present thesis including shape measurements integrated with spectral and intensity information. The development of a large database or library will be required to achieve this which should contain the number of samples required for each type of defect or disease.

6.5. Conclusion

The present thesis concludes that the novel spectroscopic sensors evaluated are promising to classify meat by feeding regime, and organs by type and disease status. However, larger datasets and refinement of the methodology are required to validate these exploratory results, such as number of scans and ROI required for repeatability, and selection of specific regions with defects within an organ. This may provide greater accuracy, precision, sensitivity, and specificity metrics to support subsequent deployment of smartphone NIRS sensors and multi-sensory imaging platforms containing VIS and SWIR HS sensors in the meat industries. Further research should focus on developing an automated workflow for analysis, detection, and sortation of products. Based on the results of the present thesis, smartphone NIRS sensors could be used in the industry as tools to validate feeding regime of beef after the results of the present thesis are validated with larger datasets. However, it is worth noting that installation of larger multi-sensory platforms would provide greater long-term benefits despite the upfront cost due to their ability to provide valuable information on the verification of feeding regime, health assessment, and organ identification without making contact with samples. Further testing of both X-ray and HS imaging within the multi-sensory platform on larger datasets can extend to predicting chemical composition such as fat and protein contents of meat samples. It is therefore concluded that the novel sensors explored in this thesis have enormous potential to automate the meat grading, authentication, and offal sortation and inspection processes that take place in the meat processing industry and afterwards such as retail.

7. List of References

Aalhus, J. L., López-Campos, Ó., Prieto, N., Rodas-González, A., Dugan, M. E. R., Uttaro, B. & Juárez, M. (2014). Review: Canadian beef grading – Opportunities to identify carcass and meat quality traits valued by consumers. *Canadian Journal of Animal Science* 94(4): 545–556. doi:10.4141/cjas-2014-038.

AgResearch. (2018). *Sensing for offal grading and enablement of automation*, Final report for project 2016-1003. Sydney: Australian Meat Processor Corporation.

AHDB. (2017). *Abattoir post-mortem conditions guide*. In O. Doxon (Ed.). Kenilworth, UK: Agriculture and Horticulture Development Board. Retrieved September 20, 2021, from https://projectblue.blob.core.windows.net/media/Default/Imported%20Publication%20Docs/AHDB%20Beef%20and%20Lamb/AbattoirGuide_1555_180917_WEB.pdf.

Akbari, H., Kosugi, Y., Kojima, K. & Tanaka, N. (2008). Wavelet-based compression and segmentation of hyperspectral images in surgery. In T. Dohi et al. (Eds.). *Proceedings of the 4th International Worksop of Medical Imaging and Augmented Reality* (August 1-2, 2008), Tokyo (pp. 142-149). Berlin: Springer. doi:10.1007/978-3-540-79982-5_16.

Akbari, H., Kosugi, Y., Kojima, K. & Tanaka, N. (2009). Blood vessel detection and artery-vein differentiation using hyperspectral imaging. In *Proceedings of the 31st International Conference of IEEE Engineering in Medicine and Biology Society* (September 2-6, 2009). Engineering the Future of Biomedicine, Minneapolis (pp. 1461-1464). Institute of Electrical and Electronics Engineers. doi:10.1109/IEMBS.2009.5332920.

Alamprese, C., Casale, M., Sinelli, N., Lanteri, S. & Casiraghi, E. (2013). Detection of minced beef adulteration with turkey meat by UV-vis, NIR and MIR spectroscopy. *LWT – Food Science and Technology* 53(1): 225-232. doi:10.1016/j.lwt.2013.01.027.

Allen, P. (2021). Recent developments in the objective measurement of carcass and meat quality for industrial application. *Meat Science* 181: 108601. doi:10.1016/j.meatsci.2021.108601.

Alomar, D., Gallo, C., Castañeda, M. & Fuchslocher, R. (2003). Chemical and discriminant analysis of bovine meat by near infrared reflectance spectroscopy (NIRS). *Meat Science* 63(4): 441-450. doi:10.1016/S0309-1740(02)00101-8.

Anderson, N. M. & Walker, P. N. (2003). Measuring fat content of ground beef stream using on-line visible/NIR spectroscopy. *Transactions of the ASAE* 46(1): 117-124. doi:10.13031/2013.12537.

Andrés, S., Silva, A., Soares-Pereira, A. L., Martins, C., Bruno-Soares, A. M. & Murray, I. (2008). The use of visible and near infrared reflectance spectroscopy to predict beef *M. longissimus thoracis et lumborum* quality attributes. *Meat Science* 76(2): 217-224. doi:10.1016/j.meatsci.2007.06.019.

Andueza, D., Listrat, A., Durand, D., Normand. J., Mourot, B. P. & Gruffat, D. (2019). Prediction of beef fatty acid composition by visible near-infrared spectroscopy was improved by preliminary freeze-drying. *Meat Science* 158: 107910. doi:10.1016/j.meatsci.2019.107910.

Aredo, V., Velásquez, L. & Siche, R. (2017). Prediction of beef marbling using hyperspectral imaging (HIS) and partial least squares regression (PLSR). *Scientia Agropecuaria* 8(2): 169-174. doi:10.17268/sci.agropecu.2017.02.09.

Ariana, D. P. & Lu, R. (2010). Hyperspectral waveband selection for internal defect detection of pickling cucumbers and whole pickles. *Computers And Electronics in Agriculture* 74(1): 137-144. doi:10.1016/j.compag.2010.07.008.

Aroeira, C. N., Torres Filho, R. A., Fontes, P. R., Ramos, A. L. S., Gomide, L. A. M., Ladeira, M. M. & Ramos, E. M. (2017). Effect of freezing prior to aging on myoglobin redox forms and CIE color of beef from Nellore and Aberdeen Angus cattle. *Meat Science* 125: 16-21. doi:10.1016/j.meatsci.2016.11.010.

Arsenault, J., Girard, C., Dubreuil, P., Daignault, D., Galarneau, J-R., Boisclair, J., Simard, C. & Bélanger, D. (2003). Prevalence of and carcass condemnation from maedi-visna, paratuberculosis and caseous

lymphadenitis in culled sheep from Quebec, Canada. *Preventative Veterinary Medicine* 59(1-2): 67-81. doi:10.1016/S1067-5877(03)00060-6.

AUS-MEAT. (2005). *Handbook of Australian meat processing* (7th Ed.). Brisbane: AUS-MEAT Limited.

AUS-MEAT. (2018). *Handbook of Australian beef processing: The AUS-MEAT language* (4th Ed.). Brisbane: AUS-MEAT Limited.

Baeten, V., Fernández Pierna, J. A. & Dardenne, P. (2007). Hyperspectral imaging techniques: an attractive solution for the analysis of biological and agricultural materials. In H. F. Grahn & P. Geladi (Eds.), *Techniques and Applications of Hyperspectral Image Analysis* (pp. 289-311). John Wiley & Sons, Ltd. doi:10.1002/9780470010884.ch12.

Bahar, B., Moloney, A. P., Monahan, F. J., Harrison, S. M., Zazzo, A., Scrimgeour, C. M., Begley, I. S. & Schmidt, O. (2009) Turnover of carbon, nitrogen, and sulphur in bovine *longissimus dorsi* and *psoas major* muscles: Implications for isotopic authentication of meat. *Journal of Animal Science* 87(3): 905-913. doi:10.2527/jas.2008-1360.

Bahar, B., Monahan, F. J., Moloney, A. P., O'Kiely, P., Scrimgeour, C. M. & Schmidt, O. (2005) Alteration of the carbon and nitrogen stable isotope composition of beef by substitution of grass silage with maize silage. *Rapid Communications in Mass Spectrometry* 19(14): 1937-1942. doi:10.1002/rcm.2007.

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

Ballin, N. Z. (2010). Authentication of meat and meat products. *Meat Science* 86(3): 577-587. doi:10.1016/j.meatsci.2010.06.001.

Barlocco, N., Vadell, E., Ballesteros, F., Galietta, G. & Cozzolino, D. (2006). Predicting intramuscular fat, moisture and Warner-Bratzler shear force in pork muscle using near infrared reflectance spectroscopy. *Animal Science* 82(1): 111-116. doi:10.1079/ASC20055.

Barnes, R. J., Dhanoa, M. S. & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy* 43(5): 772-777. doi:10.1366/0003702894202201.

Baroni, M. V., Podio, N. S., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., Gallegos, E., Gautier, E., Peral-García, P., Hoogewerff, J. & Wunderlin, D. A. (2011) How much do soil and water contribute to the composition of meat? A case study: Meat from three areas of Argentina. *Journal of Agricultural and Food Chemistry* 59(20): 11117-11128. doi:10.1021/jf2023929.

Barragán, W., Aalhus, J. L., Penner, G., Dugan, M. E. R., Juárez, M., López-Campos, Ó., Vahmani, P., Segura, J., Angulo, J. & Prieto, N. (2021). Authentication of barley-finished beef using visible and near-infrared spectroscopy (Vis-NIRS) and different discrimination approaches. *Meat Science* 172: 108342. doi:10.1016/j.meatsci.2020.108342.

Beak, S-Y., Park, S. J., Fassah, D. M., Kim, H. J., Kim, M., Jo, C. & Baik, M. (2021). Relationships among carcass traits, auction price, and image analysis traits of marbling characteristics in Korean cattle beef. *Meat Science* 171: 108268. doi:10.1016/j.meatsci.2020.108268.

Belk, K. E., Scanga, J. A., Wyle, M., Wulf, D. W., Tatum, J. D., Regan, J. O. & Smith, G. C. (2000). The use of video image analysis and instrumentation to predict beef palatability. In *Proceedings of the 53rd Reciprocal Meat Conference*, Columbus, Ohio (pp. 10-15).

Ben-Gera, I. & Norris, K. H. (1968). Direct spectrophotometric determination of fat and moisture in meat products. *Journal of Food Science* 33(1): 64-67. doi:10.1111/j.1365-2621.1968.tb00885.x.

Bester, M., Schönfeldt, H. C., Pretorius, B. & Hall, N. G. (2018). The nutrient content of selected South African lamb and mutton organ meats (offal). *Journal of Food Chemistry* 238: 3-8. doi:10.1016/j.foodchem.2017.05.075.

Biel, W., Czerniawska-Piątkowska, E. & Kowalczyk, A. (2019). Offal chemical composition from veal, beef, and lamb maintained in organic production systems. *Animals* 9(8): 489. doi:10.3390/ani9080489.

Bindon, B. M. (2004). A review of genetic and non-genetic opportunities for manipulation of marbling. *Australian Journal of Experimental Agriculture* 44(7): 687-696. doi:10.1071/EA02173.

Blakebrough-Hall, C., Hick, P. & González, L. A. (2020). Predicting bovine respiratory disease outcome using latent class analysis. *Journal of Animal Science* 98(12): skaa381. doi:10.1093/jas/skaa381.

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

Boyacı, İ. H., Temiz, H. T., Uysal, R. S., Velioğlu, H. M., Yadegari, R. J. & Rishkan, M. M. (2014). A novel method for discrimination of beef and horsemeat using Raman spectroscopy. *Food Chemistry* 148: 37-41. doi:10.1016/j.foodchem.2013.10.006.

Butler, R. J., Murray, J. G. & Tidswell, S. (2003). Quality assurance and meat inspection in Australia. *Revue Scientifique et Technique de OIE* 22(2): 697-712. doi:10.20506/rst.22.2.1430.

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.

Cama-Moncunill, R., Cafferky, J., Augier, C., Sweeney, T., Allen, P., Ferragina, A., Sullivan, C., Cromie, A. & Hamill, R. M. (2020). Prediction of Warner-Bratzler shear force, intramuscular fat, drip-loss and

cook loss in beef via Raman spectroscopy and chemometrics. *Meat Science* 167: 108157. doi:10.1016/j.meatsci.2020.108157.

Camacho-Tamayo, J., Rubiano, Y. & Hurtado M. P. (2014). Near-infrared (NIR) diffuse reflectance spectroscopy for the prediction of carbon and nitrogen in an Oxisol. *Agronómía Colombiana* 32(1): 86-94. doi:10.15446/agron.colomb.v32n1.38967.

Cameron, P. J., Zembayashi, M., Lunt, D. K., Mitsuhashi, T., Mitsumoto, M., Ozawa, S. & Smith, S. B. (1994). Relationship between Japanese Beef Marbling Standard and intramuscular lipid in the M. longissimus thoracis of Japanese Black and American Wagyu cattle. *Meat Science* 38(2): 361-364. doi:10.1016/0309-1740(94)90125-2.

Cecchinato, A., De Marchi, M., Penasa, M., Casellas, J., Schiavon, S. & Bittante, G. (2012). Genetic analysis of beef fatty acid composition predicted by near-infrared spectroscopy. *Journal of Animal Science* 90(2): 429-438. doi:10.2527/jas.2011-4150.

Chen, Y-R. & Kim. M. S. (2004). Visible/NIR imaging spectroscopy for assessing quality and safety of agro-foods. In A. M. C. Davies & A. Garrido-Varo (Eds.). *Proceedings of the 11th International Conference on Near Infrared Spectroscopy* (April 6-11, 2003). Stretching NIR Light to the Limit, Córdoba (pp. 67-68). Agriculture Research Service, U.S. Department of Agriculture.

Cheng, J-H. & Sun, D-W. (2015). Rapid and non-invasive detection of fish microbial spoilage by visible and near infrared hyperspectral imaging and multivariate analysis. *LWT – Food Science and Technology* 62(2): 1060-1068. doi:10.1016/j.lwt.2015.01.021.

CIE. (1978). *Recommendations on uniform color spaces – color equations, psychometric color terms*. Paris: Commission Internationale de l'Eclairage (CIE).

Condon, J. (2020). *Smart-phone based carcase grading camera shows big potential*. Beef Central. Retrieved November 26, 2021, from https://www.beefcentral.com/news/smart-phone-based-carcase-grading-camera-shows-big-potential/.

Connolly, A., Dona, A., Wilkinson-White, L., Hamblin, D., D'Occhio, M. & González, L. A. (2019). Relationship of the blood metabolome to subsequent carcass traits at slaughter in feedlot Wagyu crossbred steers. *Scientific Reports* 9: 15139. doi:10.1038/s41598-019-51655-2.

Connolly, A., Dona, A., Hamblin, D., D'Occhio, M. J. & González, L. A. (2020). Changes in the blood metabolome of Wagyu crossbred steers with time in the feedlot and relationships with marbling. *Scientific Reports* 10: 18987. doi:10.1038/s41598-020-76101-6.

Cook, J. & Anderson, F. (2017). *Beef and lamb OCM with CT in situ further development*, Final report for project A.TEC.0123. Sydney: Meat and Livestock Australia Limited.

Coombs, C. E. O., Allman, B. E., Morton, E. J. & González, L. A. (2021c). Discrimination of grass-fed and grain-fed frozen beef using hyperspectral imaging. In *Proceedings of the 67th International Congress of Meat Science and Technology* (August 23-27, 2021), Kraków (p. 249). Symposium Cracoviense.

Coombs, C. E. O., Fajardo, M. & González, L. A. (2021a). Comparison of smartphone and lab-grade NIR spectrometers to measure chemical composition of lamb and beef. *Animal Production Science* 61(16): 1723-1733. doi:10.1071/AN21069.

Coombs, C. E. O., Holman, B. W. B., Collins, D., Friend, M. A. & Hopkins, D. L. (2017). Effects of chilled-then-frozen storage (up to 52 weeks) on lamb *M. longissimus lumborum* quality and safety parameters. *Meat Science* 134: 86-97. doi:10.1016/j.meatsci.2017.07.017.

Coombs, C. E. O., Liddle, R. R. & González, L. A. (2021b). Portable vibrational spectroscopic methods can discriminate between grass-fed and grain-fed beef. *Journal of Near Infrared Spectroscopy* 29(6): 321-329. doi:10.1177/09670335211049506.

Coombs, C. E. O., Neely, L., Minasny, B., Fajardo, M. & González, L. A. (2019). Using a handheld near-infrared spectroscopy (NIRS) scanner to predict meat quality. In E. Puolanne (Ed.). *Proceedings of the*

65th International Congress of Meat Science and Technology (August 4-9, 2019). Meat for Diversifying Markets: Postersessions and Discussion, Potsdam (pp. 502-503). Event Lab, GmbH and DigICoMST.

Cozzolino, D., De Mattos, D. & Vaz Martins, D. (2002). Visible/near infrared reflectance spectroscopy for predicting composition and tracing system of production of beef muscle. *Animal Science* 74(3): 477-484. doi:10.1017/S1357729800052632.

Cozzolino, D. & Murray, I. (2002). Effect of sample presentation and animal muscle species on the analysis of meat by near infrared reflectance spectroscopy. *Journal of Near Infrared Spectroscopy* 10(1): 37-44. doi:10.1255/jnirs.319.

Cozzolino, D. & Murray, I. (2004). Identification of animal meat muscles by visible and near infrared reflectance spectroscopy. LWT - Food Science and Technology 37(4): 447-452. doi:10.1016/j.lwt.2003.10.01.3.

Cozzolino, D., Murray, I., Scaife, J. R. & Paterson, R. (2000). Study of dissected lamb muscles by visible and near infrared spectroscopy for composition assessment. *Animal Science* 70(3): 417-423. doi:10.1017/S1357729800051766.

Craigie, C. R., Johnson, P. L., Shorten, P. R., Charteris, A., Maclennan, G., Tate, M. L., Agnew, M. P., Taukiri, K. R., Stuart, A. D. & Reis, M. M. (2017). Application of Hyperspectral imaging to predict the pH, intramuscular fatty acid content and composition of lamb M. longissimus lumborum at 24h post mortem. *Meat Science* 132: 19-28. doi:10.1016/j.meatsci.2017.04.010.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology* 88(11): 2783-2792. doi:10.1890/07-0539.1.

Daley, C. A., Abbott, A., Doyle, P. S., Nader, G. A. & Larson, S. (2010). A review of fatty acid profiles and antioxidant content in grass-fed and grain-fed beef. *Nutrition Journal* 9: 10. doi:10.1186/1475-2891-9-10.

De Biasio, M., Stampfer, P., Leitner, R., Huck, C. W., Wiedemair, V., Balthasar, D. 2015. Micro-Raman spectroscopy for meat type detection. In M. A. Druy et al. (Eds.). *Proceedings of the 8th Sensing Technology and Applications Vol. 9482* (June 3, 2015). Next-generation spectroscopic technologies VIII, Baltimore (94821J). doi:10.1117/12.2176321.

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., Berzaghi, P., Boukha, A., Mirisola, M. & Galol, L. (2007). Use of near infrared spectroscopy for assessment of beef quality traits. *Italian Journal of Animal Science* 6(S1): 421-423. doi:10.4081/ijas.2007.1s.421.

De Marchi, M., Penasa, M., Cecchinato, A. & Bittante, G. (2013). The relevance of different near infrared technologies and sample treatments for predicting meat quality traits in commercial beef cuts. *Meat Science* 93(2): 329-335. doi:10.1016/j.meatsci.2012.09.013.

De Smet S, Balcaen, A., Claeys, E., Boeckx, P. & Van Cleemput, O. (2004). Stable carbon isotope analysis of different tissues of beef animals in relation to their diet. *Rapid Communications in Mass Spectroscopy* 18(11): 1227-1232. doi:10.1002.rcm.1471.

Delpy, D. T. & Cope, M. (1997). Quantification in tissue near-infrared spectroscopy. *Philosophical Transactions of the Royal Society B* 352(1354): 649-659. doi:10.1098/rstb.1997.0046.

Dian, P. H. M., Andueza, D., Barbosa, C. M. P., Amoreux, S., Jestin, M., Carvalho, P. C. F., Prado, I. N. & Prache, S. (2007). Methodological developments in the use of visible reflectance spectroscopy for discriminating pasture-fed from concentrate-fed lamb carcasses. *Animal* 1(8): 1198-1208. doi:10.1017/S175173110700047X.

Dian, P. H. M., Andueza, D., Jestin, M., Prado, I. M. & Prache, S. (2008). Comparison of visible and near-infrared reflectance spectroscopy to discriminate between pasture-fed and concentrate-fed lamb carcasses. *Meat Science* 80(4): 1157-1164. doi:10.1016/j.meatsci.2008.05.009.

Dixit, Y., Casado-Gavalda, M. P., Cama-Moncunill, R., Cama-Moncunill, X., Jacoby, F., Cullen, P. J. & Sullivan, C. (2016). Multipoint NIR spectrometry and collimated light for predicting the composition of meat samples with high standoff distances. *Journal of Food Engineering* 175: 58-64. doi:10.1016/j.foodeng.2015.12.004.

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.

Dixit, Y., Hitchman, S., Hicks, T. M., Lim, P., Wong, C. K., Holibar, L., Gordon, K. C., Loeffen, M., Farouk, M. M., Craigie, C. R. & Reis, M. M. (2021). Non-invasive spectroscopic and imaging systems for prediction of beef quality in a meat processing pilot plant. *Meat Science* 181: 108410. doi:10.1016/j.meatsci.2020.108410.

Dixit, Y., Pham, H. Q., Realini, C. E., Agnew, M. P., Craigie, C. R. & Reis, M. M. (2020). Evaluating the performance of a miniaturized NIR spectrophotometer for predicting intramuscular fat in lamb: A comparison with benchtop and hand-held Vis-NIR spectrophotometers. *Meat Science* 162: 108026. doi:10.1016/j.meatsci.2019.108026.

Dixon, R. & Coates, D. (2009). Review: Near infrared spectroscopy of faeces to evaluate the nutrition and physiology of herbivores. *Journal of Near Infrared Spectroscopy* 17(1): 1-31. doi:10.1255/jnirs.822.

Dow, D. L., Wiegand, B. R., Ellersieck, M. R. & Lorenzen, C. L. (2011). Prediction of fat percentage within marbling score on beef longissimus muscle using 3 different fat determination methods. *Journal of Animal Science* 89(4): 1173-1179. doi:10.2527/jas.2010-3382.

Duckett, S. K., Neel, J. P. S., Fontenot, J. P. & Clapham, W. M. (2009). Effect of winter stocker growth rate and finishing system on: III. Tissue proximate, fatty acid, vitamin, and cholesterol content. *Journal of Animal Science* 87(9): 2961-2970. doi:10.2527/jas.2009-1850.

Duckett, S. K., Wagner, D. G., Yates, L. D., Dolezal, H. G. & May, S. G. (1993). Effects of time on feed on beef nutrient consumption. *Journal of Animal Science* 71(8): 2079-2088. doi:10.2527/1993.7182079x.

Dunne, P. G., Monahan, F. J., O'Mara, F. P. & Moloney, A. P. (2009). Colour of bovine subcutaneous tissue: A review of contributory factors, associations with carcass and meat quality and its potential utility in authentication of dietary history. *Meat Science* 81(1): 28-45. doi:10.1016/j.meatsci.2008.06.013.

Ellis, D. I., Broadhurst, D., Clarke, S. J. & Goodacre, R. (2005). Rapid identification of closely related muscle foods by vibrational spectroscopy and machine learning. *Analyst* 130(12): 1648-1654. doi:10.1039/b511484e.

Elmasry, G., Barbin, D. F., Sun, D-W. & Allen, P. (2012a). Meat quality evaluation by hyperspectral imaging technique: An overview. *Critical Reviews in Food Science and Nutrition* 52(8): 689-711. doi:10.1080/10408398.2010.507908.

Elmasry, G., Sun, D-W. & Allen, P. (2012b). Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of beef. *Journal of Food Engineering* 110(1): 127-140. doi:10.1016/j.foodeng.2011.11.028.

Esteki, M., Shahsavari, Z. & Simal-Gandara, J. (2018). Use of spectroscopic methods in combination with linear discriminant analysis for authentication of food products. *Food Control* 91: 100-112. doi:10.1016/j.foodcont.2018.03.031.

Ezanno, P., Picault, S., Beaunée, G., Bailly, X., Muñoz, F., Duboz, R., Monod, H. & Guégan, J-F. (2021). Research perspectives on animal health in the era of artificial intelligence. *Veterinary Research* 52: 40. doi:10.1186/s13567-021-00902-4.

Fajardo, M., Campbell, S., Malone, B., Minasny, B. & Nelson, M. (2019). *Spectracus: Functions for environmental spectral data manipulation*. R package version 0.6.

Ferguson, D. M. (2004). Objective on-line assessment of marbling: a brief review. *Australian Journal of Experimental Agriculture* 44(7): 681-685. doi:10.10171/EA02161.

Fowler, S. M., Morris, S. & Hopkins, D. L. (2020). Preliminary investigation for the prediction of intramuscular fat content of lamb in-situ using a hand-held NIR spectroscopic device. *Meat Science* 166: 108153. doi:10.1016/j.meatsci.2020.108153.

Fowler, S. M., Schmidt, H., van de Ven, R., Wynn, P. & Hopkins, D. L. (2015a). Predicting meat quality traits of ovine *m. semimembranosus*, both fresh and following freezing and thawing, using a hand held Raman spectrometer. *Meat Science* 108: 138-144. doi:10.1016/j.meatsci.2015.06.016.

Fowler, S. M., Ponnampalam, E. N., Schmidt, H., Wynn, P. & Hopkins, D. L. (2015b). Prediction of intramuscular fat content and major fatty acid groups of lamb *M. longissimus lumborum* using Raman spectroscopy. *Meat Science* 110: 70-75. doi:10.1016/j.meatsci.2015.06.016.

Fowler, S. M., Schmidt, H., Scheier, R. & Hopkins, D. L. (2017). Raman spectroscopy for predicting meat quality traits. In F. Toldrá & L. M. L. Nollet (Eds.), *Advanced Technologies for Meat Processing* (2nd edn, pp. 83-112). CRC Press. doi:10.1201/9781315152752-3.

Fowler, S. M., Schmidt, H., van de Ven, R. & Hopkins, D. L. (2018). Preliminary investigation into the use of Raman spectroscopy to predict meat and eating quality of beef loins. *Meat Science* 138: 53-58. doi:10.1016/j.meatsci.2018.01.002.

Frank, D., Joo, S-T. & Warner, R. (2016). Consumer acceptability of intramuscular fat. *Korean Journal for Food Science of Animal Resources* 36(6): 699-708. doi:10.5851/kosfa.2016.36.6.699.

Gardner, G. E., 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. doi:10.1016/j.meatsci.2018.06.020.

Gardner, G. E., Williams, A., Siddell, J., Ball, A. J., Mortimer, S., Jacob, R. H., Pearce, K. L., Hocking Edwards, J. E., Rowe, J. B. & Pethick, D. W. (2010). Using Australian Sheep Breeding Values to increase lean meat yield percentage. *Animal Production Science* 50(12): 1098-1106. doi:10.1071/AN10144.

Gerrard, D. E., Cao, X. & Tan, J. (1996). Beef marbling and color score determination by image processing. *Journal of Food Science* 61(1): 145-148. doi:10.1111/j.1365-2621.1996.tb14745.x.

GIMP Development Team. (2020). *GNU Image Manipulation Program (Version 2.10.18)*. GIMP, University of California, Berkeley, CA. http://gimp.org/.

Goi, A., Hocquette, J-F., Pellattiero, E. & De Marchi, M. (2022). Handheld near-infrared spectrometer allows *on-line* prediction of beef quality traits. *Meat Science* 184: 108694. doi:10.1016/j.meatsci.2021.108694.

Goodwin-Ray, K. A., Stevenson, M. A., Heuer, C. & Cogger, N. (2008). Economic effect of pneumonia and pleurisy in lambs in New Zealand. *New Zealand Veterinary Journal* 56(3): 107-114. doi:10.1080/00480169.2008.36818.

Harper, G. S., Oddy, V. H., Pethick, D., Tume, R. K., Barendse, W. J. & Hygate, L. (2003). *Biological determinants of intramuscular fat deposition in beef cattle: current mechanistic knowledge and sources of variation*. Final Report, Meat Research Corporation Project FLOT208. Sydney: Meat and Livestock Australia Limited.

Helrich, K. (Ed.) (1990). *Official methods of analysis of the Association of Official Analytical Chemists* (15th edn). Arlington: Association of Official Analytical Chemists.

Hitchman, S., Loeffen, M. P. F., Reis, M. M. & Craigie, C. R. (2021). Robustness of hyperspectral imaging and PLSR model predictions of intramuscular fat in lamb M. longissimus lumborum across several flocks and years. *Meat Science* 179: 108492. doi:10.1016/j.meatsci.2021.108492.

Holman, B. W. B., Ponnampalam, E. N., van de Ven, R. J., Kerr, M. G. & Hopkins, D. L. (2015). Lamb meat colour values (HunterLab CIE and reflectance) are influenced by aperture size (5 mm v. 25 mm). *Meat Science* 100: 202-208. doi:10.1016/j.meatsci.2014.10.006.

HORIBA. (2020). *Raman data and analysis – Raman bands*. HORIBA Jobin Yvon, Inc. Retrieved June 18, 2020, from https://static.horiba.com/fileadmin/Horiba/Technology/Measurement_Techniques/Molecular_Spectroscopy/Raman_Spectroscopy/Raman_Academy/Raman_Tutorial/Raman_bands.pdf.

Horikoshi, M. & Tang, Y. (2016). *ggfortify: Data visualization tools for statistical analysis results*. R package version 0.4.12. https://CRAN.R-project.org/package=ggfortify.

Hu, X., Zou, L., Huang, X. & Lu, X. (2017). Detection and quantification of offal content in ground beef meat using vibrational spectroscopic-based analysis. *Scientific Reports 7*: 15162. doi:10.1038/s41598-017-15389-3.

Huang, H., Liu, L. & Ngadi, M. O. (2014a). Recent developments in hyperspectral imaging for assessment of food quality and safety. *Sensors* 14(4): 7248-7276. doi: 10.3390/s140407248.

Huang, H., Liu, L., Ngadi, M. O., Gariépy, C. & Prasher, S. O. (2014b). Near-infrared spectral image analysis of pork marbling based on Gabor filter and wide line detector techniques. *Applied Spectroscopy* 68(3): 332-339. doi:10.1366/13-07242.

Huang, J., Romero-Torres, S. & Moshgbar, M. (2010). Practical considerations in data pre-treatment for NIR and Raman spectroscopy. *American Pharmaceutical Review* 13(6): 116-127.

Hwang, I. H., Park, B. Y., Cho, S. H., Kim, J. H. & Lee, J. M. (2004). Meat quality of highly marbled imported beef with reference to Hanwoo beef. *Journal of Animal Science and Technology* 46(4): 659-666. doi:10.5187/JAST.2004.46.4.659.

Hwang, Y-H. & Joo, S-T. (2017). Fatty acid profiles, meat quality, and sensory palatability of grain-fed and grass-fed beef from Hanwoo, American, and Australian crossbred cattle. *Korean Journal of Food Science of Animal Resources* 37(2): 153-161. doi:10.5851/kosfa.2017.37.2.153.

Jacob, R. & Calnan, H. (2018). *Improving lamb lean meat yield – a technical guide for the Australian lamb and sheep meat industry*. Sydney: Meat and Livestock Australia Limited.

Johnston, D. J., Reverter, A., Thompson, J. M. & Perry, D. (1999). Genetic and phenotypic relationships between four methods of assessing intramuscular fat in beef carcases. *Proceedings of the 13th Conference of Association for the Advancement of Animal Breeding and Genetics* (July 4-7, 1999). Rising to the Challenge: Breeding for the 21st Century Consumer, Mandurah (pp. 345-348).

Jones, G. E., Field, A. C., Gilmour, J. S., Rae, A. G., Nettleton, P. F. & McLauchlan, M. (1982). Effects of experimental chronic pneumonia on bodyweight, feed intake and carcase composition of lambs. *The Veterinary Record* 110(8): 168-173. doi:10.1136/vr.110.8.168.

Kamruzzaman, M., ElMasry, G., Sun, D-W. & Allen, P. (2012a). Non-destructive prediction and visualization of chemical composition in lamb meat using NIR hyperspectral imaging and multivariate regression. *Innovative Food Science & Emerging Technologies* 16: 218-226. doi:10.1016/j.ifset.2012.06.003.

Kamruzzaman, M., ElMasry, G., Sun, D-W. & Allen, P. (2012b). Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis. *Analytica Chimica Acta* 714: 57-67. doi:10.1016/j.aca.2011.11.037.

Kamruzzaman, M., Haque, M. E. & Ali, M. R. (2014). Hyperspectral imaging technique for offal quantification in minced meat. *Journal of the Bangladesh Agricultural University* 12(1): 189-194. doi:10.3329/jbau.v12i1.21411.

Kamruzzaman, M., Makino, Y., Oshita, S. & Liu, S. (2015). Assessment of visible near-infrared hyperspectral imaging as a tool for detection of horsemeat adulteration in minced beef. *Food and Bioprocess Technology* 8(5): 1054-1062. doi:10.1007/s11947-015-1470-7.

Kamruzzaman, M., Makino, Y. & Oshita, S. (2016). Rapid and non-destructive detection of chicken adulteration in minced beef using visible near-infrared hyperspectral imaging and machine learning. *Journal of Food Engineering* 170: 8-15. doi:10.1016/j.foodeng.2015.08.023.

Kamruzzaman, M., Sun, D-W., ElMasry, G. & Allen, P. (2013). Fast detection and visualization of minced lamb meat adulteration using NIR hyperspectral imaging and multivariate image analysis. *Talanta* 103: 130-136. doi:10.1016/j.talanta.2012.10.020.

Karamichou, E., Richardson, R. I., Nute, G. R., McLean, K. A. & Bishop, S. C. (2006). Genetic analyses of carcass composition, as assessed by X-ray computed tomography, and meat quality traits in Scottish blackface sheep. *Animal Science* 82(2): 151-162. doi:10.1079/ASC200518.

Knight, T. W., Death, A. F., Lambert, M. G. & McDougall, D. B. (2001). The rate of reduction in carotenoid concentration in fat of steers fed a low carotenoid ration, and the role of increasing carcass fatness. *Australian Journal of Agricultural Research* 52(10): 1023-1032. doi:10.1071/AR01017.

Konarska, M., Kuchida, K., Tarr, G. & Polkinghorne, R. J. (2017). Relationships between marbling measures across principal muscles. *Meat Science* 123: 67-78. doi:10.1016/j.meatsci.2016.09.005.

Konda Naganathan, G., Grimes, L. M., Subbiah, J., Calkins, C. R., Samal, A. & Meyer, G. E. (2008). Visible/near-infrared hyperspectral imaging for beef tenderness prediction. *Computers and Electronics in Agriculture* 64(2): 225-233. doi:10.1016/j.compag.2008.05.020.

Kong, W., Zhang, C., Liu, F., Nie, P. & He, Y. (2013). Rice seed cultivar identification using near-infrared hyperspectral imaging and multivariate data analysis. *Sensors* 13(7): 8916-8927. doi:10.3390/s130708916.

Kruggel, W. G., Field, R. A., Riley, M. L., Radloff, H. D. & Horton, K. M. (1981). Near infrared reflectance determination of fat, protein and moisture in fresh meat. *Journal of the Association of Official Analytical Chemists* 64(3): 692-696. doi:10.1093/jaoac/64.3.692.

Kruk, Z. A., Pitchford, W. S., Siebert, B. D., Deland, M. P. B. & Bottema, C. D. K. (2002). Factors affecting estimation of marbling in cattle and the relationship between marbling scores and intramuscular fat. *Proceedings of the 24th Biennial Animal Production in Australia Conference* (July 4-7, 2002), Adelaide (pp. 129-132). Australian Society of Animal Production.

Kucheryavskiy, S. (2020). mdatools – R package for chemometrics. *Chemometrics and Intelligent Laboratory Systems* 198: 103937. doi:10.1016/j.chemolab.2020.103937.

Kucheryavskiy, S. (2021). *Package 'mdatools' – Multivariate data analysis for chemometrics*. Vignette for R package version 0.11.5. Retrieved April 30, 2021, from https://cran.r-project.org/web/packages/mdatools/mdatools.pdf.

Kuchida, K., Kono, S., Konishi, K., Van Vleck, L. D., Suzuki, M. & Miyoshi, S. (2000). Prediction of crude fat content of longissimus muscle of beef using the ratio of fat area calculated from computer image analysis: comparison of regression equations for prediction using different input devices at different stations. *Journal of Animal Science* 78(4): 799-803. doi:10.2527/2000.784799x.

Kuchida, K., Suzuki, M. & Miyoshi, S. (2001). Development of photographing equipment for the cross-section of carcass and prediction of BMS number by using obtained image from that equipment. *Animal Science Journal* 72(8): J224-J231. doi:10.2508/chikusan.72.8_224.

Kuhn, M. (2008). Building predictive models in *R* using the *caret* package. *Journal of Statistical Software* 28(5): 1-26. doi:10.18637/jss.v028.i05.

Kuhn, M. (2020). *Caret: Classification and regression training*. R package version 6.0-86. Vignette retrieved February 25, 2022, from https://cran.r-project.org/web/packages/caret/caret.pdf.

Kuhn, M. & Quinlan, R. (2020). *Cubist: Rule- and instance-based regression modelling*. R package version 0.2.3. Retrieved February 10, 2021, from https://CRAN.R-project.org/package=Cubist.

Kuhn, M. & Wickham, H. (2020). *Tidymodels: A collection of packages for modelling and machine learning using tidyverse principles.* R package version 0.1.4. Retrieved June 15, 2021, from https://www.tidymodels.org/.

Kumar, A., Saxena, S., Shrivastava, S., Bharti, V., Kumar, U. & Dhama, K. (2016). Hyperspectral imaging (HSI): Applications in animal and dairy sector. *Journal of Experimental Biology and Agricultural Sciences* 4(4): 448-461. doi:10.18006/2016.4(4).448.461.

Lacasta, D. Ferrer, L. M., Ramos, J. J., González, J. M. & De las Heras, M. (2008). Influence of climatic factors on the development of pneumonia in lambs. *Small Ruminant Research* 80(1-3): 28-32. doi:10.1016/j.smallrumres.2008.08.004.

Lanza, E. (1983). Determination of moisture, protein, fat and calories in raw pork and beef by near infrared spectroscopy. *Journal of Food Science* 48(2): 471-474. doi:10.1111/j.1365-2621.1983.tb10769.x.

Lawrence, R. L., Wood, S. D. & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment* 100(3): 356-362. doi:10.1016/j.rse.2005.10.014.

Lee, K. J., Kishimoto, M., Shimizu, J., Kobayashi, Y., Matsumoto, K., Sasaki, N., Ishii, M., Inokuma, H., Iwasaki, T., Miyake, Y. & Yamada, K. (2011). Use of contrast-enhanced CT in the diagnosis of abscesses in cattle. *Journal of Veterinary Medical Science* 73(1): 113-115. doi:10.1292/jvms.10-0233.

Lee, K., Yamada, K., Tsuneda, R., Kishimoto, M., Shimizu, J., Kobayashi, Y., Furuoka, H., Matsui, T., Sasaki, N., Ishii, M., Inokuma, H., Miyake, Y. & Iwasaki, T. (2009). Clinical experience of using multidetector-row CT for the diagnosis of disorders in cattle. *The Veterinary Record* 165(19): 559-562. doi:10.1136/vr.165.19.559.

Leheska, J. M., Thompson, L. D., Howe, J. C., Hentges, E., Boyce, J., Brooks, J. C., Shriver, B., Hoover, L. & Miller, M. F. (2008). Effects of conventional and grass-feeding systems on the nutrient composition of beef. *Journal of Animal Science* 86(12): 3575-3585. doi:10.2527/jas.2007-0565.

Liao, Y-T., Fan, Y-X. & Cheng, F. (2010). On-line prediction of fresh pork quality using visible/near-infrared reflectance spectroscopy. *Meat Science* 86(4): 901-907. doi:10.1016/j.meatsci.2010.07.011.

Liaw, A. & Wiener, M. R. (2002). Classification and regression by randomForest. *R News* 2(3): 18-22.

Lillesand, T., Kiefer, R. W. & Chipman, J. (2015). *Remote sensing and image interpretation* (7th edn). New York: Wiley.

Liu, L., Ngadi, M. O., Prasher, S. O. & Gariépy, C. (2010). Categorization of pork quality using Gabor filter-based hyperspectral imaging technology. *Journal of Food Engineering* 99(3): 284-293. doi:10.1016/j.jfoodeng.2010.03.001.

Logan, B. G., Hopkins, D. L., Schmidtke, L., Morris, S. & Fowler, S. M. (2020a). Preliminary investigation into the use of Raman spectroscopy for the verification of Australian grass and grain fed beef. *Meat Science* 160: 107970. doi:10.1016/j.meatsci.2019.107970.

Logan, B. G., Hopkins, D. L., Schmidtke, L. M. & Fowler, S. M. (2020b). Analysis of Raman spectra for the verification of Australian grass- and grain-fed beef using principal component analysis and partial least square models. *Journal of Raman Spectroscopy* 51(11): 2338-2346. doi:10.1002/jnirs.5983

Logan, B. G., Hopkins, D. L., Schmidtke, L. M., Morris, S. & Fowler, S. M. (2021a). Classification of Southern Australian grass- and grain-fed beef. *Food Analytical Methods* 14(8): 1730-1743. doi:10.1007/s12161-021-02010-7.

Logan, B. G., Hopkins, D. L., Schmidtke, L. M. & Fowler, S. M. (2021b). Authenticating common Australian beef production systems using Raman spectroscopy. *Food Control* 121: 107652. doi:10.1016/j.foodcont.2020.107652.

Lohumi, S., Lee, S., Lee, H. & Cho, B-K. (2015). A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science and Technology* 46: 85-98. doi:10.1016/j.tifs.2015.08.003.

Luciano, G., Moloney, A. P., Priolo, A., Röhrle, F. T., Vasta, V., Biondi, L., López-Andrés, P., Grasso, S. & Monahan, F. J. (2011). Vitamin E and polyunsaturated fatty acids in bovine muscle and the oxidative stability of beef from cattle receiving grass or concentrate-based rations. *Journal of Animal Science* 89(11): 3759-3768. doi:10.2527/jas.2010-3795.

Maeda, S., Grose, J., Kato, K. & Kuchida, K. (2014). Comparing AUS-MEAT marbling scores using image analysis traits to estimate genetic parameters for marbling of Japanese Black cattle in Australia. *Animal Production Science* 54(5): 557-563. doi:10.1071/AN12368.

Magalhães, A. F. B., Teixeira, G. H. A., Ríos, A. C. H., Silva, D. B. S., Mota, L. F. M., Muniz, M. M. M., de Morais, C. D. M., de Lima, K. M. G., Cunha Júnior, L. C., Baldi, F., Cavalheiro, R., de Oliveira, H. N., Chardulo, L. A. L. & de Albuquerque, L. G. (2018). Prediction of meat quality traits in Nelore cattle by near-infrared reflectance spectroscopy. *Journal of Animal Science* 96(10): 4229-4237. doi:10.1093/jas/sky284.

Mancini, R. M. & Hunt, M. C. (2005). Current research in meat color. *Meat Science* 71(1): 100-121. doi:10.1016/j.meatsci.2005.03.003.

Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews* 43(24): 8200-8214. doi:10.1039/c4cs00062e.

Mao, R., Qi, H., Pei, L., Hao, J., Dong, J., Jiang, T., Ainiwaer, A., Shang, G., Xu, L., Shou, X., Zhang, S., Wu, G., Lu, P., Bao, Y. & Li, H. (2017). CT scanning in identification of sheep cystic echinococcosis. *BioMed Research International* 2017: 4639202. doi:10.1155/2017/4639202.

McBratney, A. B., Minasny, B. & Viscarra Rossel, R. (2006). Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. *Geoderma* 136(1-2): 272-278. doi:10.1016/j.geoderma.2006.03.051.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., Ledsam, J. R., Melnick, D., Mostofi, H., Peng, L., Reicher, J. J., Romera-Paredes, B., Sidebottom, R., Suleyman, M., Tse, D., Young, K. C., De Fauw, J. & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature* 577(7788): 89-94. doi:10.1038/s41586-019-1799-6.

Mehnaz, K. R., Keitel, C. & Dijkstra, F. A. (2018). Effects of carbon and phosphorus addition on microbial respiration, N_2O emission, and gross nitrogen mineralization in a phosphorus-limited grassland soil. *Biology and Fertility of Soils* 54(4): 481-493. doi:10.1007/s00374-018-1274-9.

Mevik, B-H., Wehrens, R. & Liland, K. H. (2020). *pls: Partial least squares and principal component regression*. R package version 2.7-3. https://CRAN.R-project.org/package=pls.

Miller, R. (2020). Drivers of consumer liking for beef, pork, and lamb: a review. *Foods* 9(4): 428. doi:10.3390/foods9040428.

Miller, M. F., Carr, M. A., Ramsey, C. B., Crockett, K. L. & Hoover, L. C. (2001). Consumer thresholds for establishing the value of beef tenderness. *Journal of Animal Science* 79(12): 3062-3068. doi:10.2527/2001.79123062x.

Minasny, B. & McBratney, A. B. (2008). Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemometrics and Intelligent Laboratory Systems* 94(1): 72-79. doi:10.1016/j.chemolab.2008.06.003.

Minasny, B., McBratney, A. B., Bellon-Maurel, V., Roger, J-M., Gobrecht, A., Ferrand, L. & Joalland, S. (2011). Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma* 167-168: 118-124. doi:10.1016/j.geoderma.2011.09.008.

MLA. (2012). *Meat Standards Australia: Sheepmeat*. Sydney: Meat and Livestock Australia Limited.

MLA. (2017a). *Meat Standards Australia: Beef information kit*. Sydney: Meat and Livestock Australia Limited.

MLA. (2017b). *Meat Standards Australia: Annual Outcomes Report 2016-17: Producing excellence in eating quality.* Sydney: Meat and Livestock Australia Limited.

Moloney, A. P., Bahar, B., Schmidt, O., Monahan, F. J., O'Kiely, P. & Scrimgeour, C. (2006). Differentiation of beef according to the pre-slaughter diet of cattle using the stable isotope ratios of carbon and nitrogen. In D. Troy et al. (Eds.). *Proceedings of the 52nd International Congress of Meat Science and Technology* (August 13-18, 2006). Harnessing and exploiting global opportunities, Dublin (pp. 665-666). Wageningen Academic. doi:10.3920/978-90-8686-579-6.

Moloney, A. P., O'Riordan, E. G., Schmidt, O. & Monahan, F. J. (2018). The fatty acid profile and stable isotope ratios of C and N of muscle from cattle that grazed grass or grass/clover pasture before slaughter and their discriminatory potential. *Irish Journal of Agricultural and Food Research* 57: 84-94. doi:10.1515/ijafr-2018-0009.

Monahan, F. J., Schmidt, O. & Moloney, A. P. (2018). Meat provenance: Authentication of geographical origin and dietary background of meat. *Meat Science* 144: 2-14. doi:10.1016/j.meatsci.2018.05.008.

Moore, C. B., Bass, P. D., Green, M. D., Chapman, P. L., O'Connor, M. E., Yates, L. D., Scanga, J. A., Tatum, J. D., Smith, G. C. & Belk, K. E. (2010). Establishing an appropriate mode of comparison for measuring the performance of marbling score output from video image analysis beef cattle grading systems. *Journal of Animal Science* 88(7): 2464-2475. doi:10.2527/jas.2009-2593.

Morsy, N. & Sun, D-W. (2013). Robust linear and non-linear models of NIR spectroscopy for detection and quantification of adulterants in fresh and frozen-thawed minced beef. *Meat Science* 93(2): 292-302. doi:10.1016/j.meatsci.2012.09.005.

Mourot, B. P., Gruffat, D., Durand, D., Chesneau, G., Prache, S., Mairesse, G. & Andueza, D. (2014). New approach to improve the calibration of main fatty acids by near-infrared reflectance spectroscopy in ruminant meat. *Animal Production Science* 54(10): 1848-1852. doi:10.1071/AN14328.

Murray, G. (1986). Ante-mortem and post-mortem meat inspection: an Australian Inspection Service perspective. *Australian Veterinary Journal* 63(7): 211-215. doi:10.1111/j.1751-0813.1986.tb02997.x.

Nade, T., Fujita, K., Fujii, M., Yoshida, M., Haryu, T., Misumi, S. & Okumura, T. (2005). Development of x-ray computed tomography for live standing cattle. *Animal Science Journal* 76(7): 513-517. doi:10.1111/j.1740-0929.2005.00298.x.

Nakahashi, Y., Maruyama, S., Seki, S., Hidaka, S. & Kuchida, K. (2008). Relationships between monounsaturated fatty acids of marbling flecks and image analysis traits in longissimus muscle for Japanese Black steers. *Journal of Animal Science* 86(12): 3551-3556. doi:10.2527/jas.2008-0947.

Nakariyakul, S. & Casasent, D. P. (2009). Fast feature selection algorithm for poultry skin tumor detection in hyperspectral data. *Journal of Food Engineering* 94(3-4): 358-365. doi:10.1016/j.foodeng.2009.04.001.

Nakashita, R., Suzuki, Y., Akamatsu, F., Iizumi, Y., Korenaga, T. & Chikaraishi, Y. (2008). Stable carbon, nitrogen, and oxygen isotope analysis as a potential tool for verifying geographical origin of beef. *Analytica Chimica Acta* 617(1-2): 148-152. doi:10.1016/j.aca.2008.03.048.

Neethirajan, S., Tuteja, S. K., Huang, S-T. & Kelton, D. (2017). Recent advancement in biosensors technology for animal and livestock health management. *Biosensors and Bioelectronics* 98: 398-407. doi:10.1016/j.bios.2017.07.015.

Norris, K. H., Barnes, R. F., Moore, J. E. & Shenk, J. S. (1976). Predicting forage quality by infrared reflectance spectroscopy. *Journal of Animal Science* 43(4): 889-897. doi:20.2527/jas1976.434889x.

O'Reilly, R. A., Pannier, L., Gardner, G. E., Garmyn, A. J., Jacob, R. H., Luo, H., Meng, Q., Miller, M. F. & Pethick, D. W. (2017). IMF and eating quality in sheepmeat: a comparison of American, Chinese and Australian consumers. In D. Troy et al. (Eds.). *Proceedings of the 63rd International Congress of Meat Science and Technology* (August 13-18, 2017). Nurturing locally, growing globally, Cork (pp. 799-800). Wageningen Academic. doi:10.3920/978-90-8686-860-5.

Osborne, B. G., Fearn, T. & Hindle, P. H. (1993). Applications of near infrared spectroscopy in food and beverage analysis. In Browning D. (Ed.). *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis* (2nd edn, pp. 145-199). Harlow: Longman Scientific and Technical.

Osorio, M. T., Moloney, A. P., Schmidt, O. & Monahan, F. J. (2011). Multielement isotope analysis of bovine muscle for determination of international geographical origin of meat. *Journal of Agricultural and Food Chemistry* 59(7): 3285-3294. doi:10.1021/jf1040433.

Pannier, L., Gardner, G. E., O'Reilly, R. A. & Pethick, D. W. (2018). Factors affecting lamb eating quality and the potential for their integration into an MSA sheepmeat grading model. *Meat Science* 144: 43-52. doi:10.1016/j.meatsci.2018.06.035.

Pannier, L., Gardner, G. E., Pearce, K. L., McDonough, M., Ball, A. J., Jacob, R. H. & Pethick, D. W. (2014). Associations of sire estimated breeding values and objective meat quality measurements with sensory scores in Australian lamb. *Meat Science* 96(2B): 1076-1087. doi:10.1016/j.meatsci.2013.07.037.

Patel, N., Toledo-Alvarado, H. & Bittante, G. (2021). Performance of different portable and hand-held near-infrared spectrometers for predicting beef quality composition and quality characteristics in the abattoir without meat sampling. *Meat Science* 178: 108518. doi:10.1016/j.meatsci.2021.108518.

Paulus, C., Moulin, V., Perion, D., Radisson, P. & Verger, L. (2017). Multi-energy x-ray detectors to improve air-cargo security. In A. Ashok et al. (Eds.). *Proceedings of SPIE Defense + Security Vol 10187 (*April 9-13, 2017). Anomaly Detection and Imaging with X-rays II, Anaheim, California. Society of Photo-Optical Instrumentation Engineers (SPIE) Digital Library. doi:10.1117/12.2262585.

PCAS. (2016). *Pasturefed cattle assurance system standards (version 5.7)*. Pasturefed Cattle Assurance System. Retrieved June 3, 2020, from http://www.pcaspasturefed.com.au/assets/documents/Standards/PCAS%20Standards.v5.7%20270916(Final).pdf.

Pedersen, D. K., Morel, S., Andersen, H. J. & Engelsen, S. B. (2003). Early prediction of water-holding capacity in meat by multivariate vibrational spectroscopy. *Meat Science* 65(1): 581-592. doi:10.1016/S0309-1740(02)00251-6.

Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B. & Shi, Z. (2019). Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* 337: 1309-1319. doi:10.1016/j.geoderma.2018.08.006.

Perry, D., Shorthose, W. R., Ferguson, D. M. & Thompson, J. M. (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. doi:10.1071/EA00092.

Pethick, D. W., McGilchrist, P., Polkinghorne, R., Warner, R., Tarr, G., Garmyn, A., Thompson, J. & Hocquette, J-F. (2018). International research on beef and lamb eating quality (Travaux de recherche internationaux sur la qualité sensorielle de la viande ovine et bovine). *Viandes & Produits Carnés* 34: 1-11.

Pham, H. Q., Tu, C., Dixit, Y., Agnew, M., Craigie, C. & Reis, M. M. (2018). Mini NIR spectrophotometer for predicting intramuscular fat in beef: A comparison with high resolution NIR spectrophotometers. In E. Puolanne (Ed.). *Proceedings of the 64th International Congress of Meat Science and Technology* (August 12-17, 2018), Melbourne. YRD Pty. Ltd. and DigICoMST.

Polkinghorne, R. J. & Thompson, J. M. (2010). Meat standards and grading: A world view. *Meat Science* 86(1): 227-235. doi:10.1016/j.meatsci.2010.05.010.

Polkinghorne, R., Thompson, J. M., Watson, R., Gee, A. & Porter, M. (2008). Evolution of the Meat Standards Australia (MSA) beef grading system. *Australian Journal of Experimental Agriculture* 48(11): 1351-1359. doi:10.1071/EA07177.

Prache, S., Cornu, A., Berdagué, J. L. & Priolo, A. (2005). Traceability of animal feeding diet in the meat and milk of small ruminants. *Small Ruminant Research* 59(2-3): 157–168. doi:10.1016/j.smallrumres.2005.05.004.

Prache, S., Martin, B. & Coppa, M. (2020). Review: Authentication of grass-fed meat and dairy products from cattle and sheep. *Animal* 14(4): 854-863. doi:10.1017/S1751731119002568.

Prache, S. & Theriez, M. (1999). Traceability of lamb production systems: carotenoids in plasma and adipose tissue. *Animal Science* 69(1): 29-36. doi:10.1017/S1357729800051067.

Prevolnik, M., Čandek-Potokar, M. & Škorjanc, D. (2004). Ability of NIR spectroscopy to predict meat chemical composition and quality: a review. *Czechoslovak Journal of Animal Science* 49(11): 500-510. doi:10.17221/4337-CJAS.

Prevolnik, M., Škrelp, M., Škorjanc, D. & Čandek-Potokar, M. (2010). Application of near infrared spectroscopy to predict chemical composition of meat and meat products. *Technologija Mesa* 51(2): 133-142.

Prieto, N., López-Campos, Ó., Aalhus, J. L., Dugan, M. E. R., Juárez, M. & Uttaro, B. (2014a). 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., López-Campos, Ó., Zijlstra, R. T., Uttaro, B. & Aalhus, J. L. (2014b). Discrimination of beef dark cutters using visible and near infrared reflectance spectroscopy. *Canadian Journal of Animal Science* 94(3): 445-454. doi:10.4141/cjas-2014-024.

Prieto, N., Pawluczyk, O., Dugan, M. E. R. & Aalhus, J. L. (2017). A review of the principles and application of near-infrared spectroscopy to characterise meat, fat, and meat products. *Applied Spectroscopy* 71(7): 1403-1426. doi:10.1177/0003702817709299.

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., Richardson, R. I., Hyslop, J. J., Simm, G. & Roehe, R. (2011). Online prediction of fatty acid profiles in crossbred Limousin and Aberdeen Angus beef cattle using near infrared reflectance spectroscopy. *Animal* 5(1): 155-165. doi:10.1017/S1751731110001618.

Priolo, A., Micol, D., Agabriel, J., Prache, S. & Dransfield, E. (2002). Effect of grass or concentrate feeding systems on lamb carcass and meat quality. *Meat Science* 62(2): 179-185. doi:10.1016/S0309-1740(01)00244-3.

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.

Qiao, T., Ren, J., Craigie, C., Zabalza, J., Maltin, C. & Marshall, S. (2015). Quantitative prediction of beef quality using visible and NIR spectroscopy with large data samples under industry conditions. *Journal of Applied Spectroscopy* 82(1): 137-144. doi:10.1007/s10812-015-0076-1.

Qiao, Y. & van Kempen, T. A. T. G. (2004). Technical note: Comparison of Raman, mid, and near infrared spectroscopy for predicting the amino acid content in animal meals. *Journal of Animal Science* 82(9): 2596-2600. doi:10.2527/2004.8292596x.

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing: Vienna. https://R-project.org/.

Realini, C. E., Duckett, S. K. & Windham, W. R. (2004). Effect of vitamin C addition to ground beef from grass-fed or grain-fed sources on color and lipid stability, and prediction of fatty acid composition by near-infrared reflectance analysis. *Meat Science* 68(1): 35-43. doi:10.1016/j.meatsci.2004.02.002.

Reis, M. M. & Rosenvold, K. (2014). Early on-line classification of beef carcasses based on ultimate pH by near infrared spectroscopy. *Meat Science* 96(2A): 862-869. doi:10.1016/j.meatsci.2013.10.016.

Ritthiruangdej, P., Ritthiron, R., Shinzawa, H. & Ozaki, H. (2011). Non-destructive and rapid analysis of chemical compositions in Thai steamed pork sausages by near-infrared spectroscopy. *Food Chemistry* 129(2): 684-692. doi:10.1016/j.foodchem.2011.04.110.

Roger, J-M., Chauchard, F. & Bellon-Maurel, V. (2003). EPO-PLS external parameter orthogonalization of PLS application to temperature-independent measurement of sugar content of intact fruits. *Chemometrics and Intelligent Laboratory Systems* 66(2): 191-204. doi:10.1016/S0169-7439(03)00051-0.

Ropodi, A. I., Pavlidis, D. E., Mohareb, F., Panagou, E. Z. & Nychas, G-J. E. (2015). Multispectral image analysis approach to detect adulteration of beef and pork in raw meats. *Food Research International* 67: 12-18. doi:10.1016/j.foodres.2014.10.032.

Samuel, J. L., O'Boyle, D. A., Mathers, W. J. & Frost, A. J. (1980). The contamination with Salmonella of bovine livers in an abattoir. *Australian Veterinary Journal* 56(11): 526-528. doi:10.1111/j.1751-0813.1980.tb02579.x.

Savell, J. & Cross, H. (1988). The role of fat in the palatability of beef, pork and lamb. In D.L. Call et al. (Eds.), *Designing foods: Animal production in the marketplace* (pp. 345-355). Washington, D.C.: National Academy Press. 345-355.

Savell, J. W., Cross, H. R. & Smith, G. C. (1986). Percentage ether extractable fat and moisture content of beef *Longissimus* muscle as related to USDA marbling score. *Journal of Food Science* 51(3): 838-839. doi:10.1111/j.1365-2621.1986.tb13946.x.

Savitzky, A. & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36(8): 1627-1639. doi:10.1021/ac60214a047.

Scholz, A. M., Bunger, L., Kongsro, J., Baulain, U. & Mitchell, A. D. (2015). Non-invasive methods for the determination of body and carcass composition in livestock: dual energy X-ray absorptiometry, computed tomography, magnetic resonance imaging and ultrasound: a review. *Animal* 9(7): 1250-1264.

Shackelford, S.D., Wheeler, T. L. & Koohmaraie, M. (2004). Development of optimal protocol for visible and near-infrared reflectance spectroscopic evaluation of meat quality. Meat Science 68(3): 371-381. doi:10.1016/j.meatsci.2004.01.013.

Shackelford, S. D., Wheeler, T. L. & Koohmaraie, M. (2005). On-line classification of US Select beef carcasses for *longissimus* tenderness using visible and near-infrared reflectance spectroscopy. *Meat Science* 69(3): 409-415. doi:10.1016/j.meatsci.2004.08.011.

Silva, J. A., Patarata, L. & Martins, C. (1999). Influence of ultimate pH on bovine meat tenderness during ageing. *Meat Science* 52(4): 453-459.

Sim, T. (2021a). *Gundagai launches world-first meat yield and IMF lamb grid*. Sheep Central. Retrieved November 26, 2021, from https://www.sheepcentral.com/gundagai-launches-world-first-meat-yield-and-imf-lamb-grid/.

Sim, T. (2021b). *Early results in Gundagai's world-first lamb IMF grid*. Sheep Central. Retrieved November 26, 2021, from https://www.sheepcentral.com/early-results-in-on-gundagais-world-first-lamb-imf-grid/.

Singh, C. B., Jayas, D. S., Paliwal, J. & White, N. D. G. (2009). Detection of insect-damaged wheat kernels using near-infrared hyperspectral imaging. *Journal of Stored Products Research* 45(3): 151-158. doi:10.1016/j.jspr.2008.12.002.

Stevens, A. & Ramirez-Lopez, L. (2020). An introduction to the *prospectr* package. Vignette for R package version 0.2.1. Retrieved October 22, 2021, from https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html.

Stewart, S. M., Gardner, G. E., Williams, A., Pethick, D. W., McGilchrist, P. & Kuchida, K. (2021). Association between visual marbling score and chemical intramuscular fat with camera marbling percentage in Australian beef carcasses. *Meat Science* 181: 108369. doi:10.1016/j.meatsci.2020.108369.

Strong, J. (2004). Differences in carcass grading schemes used in the USA, Japan and Australia. *Australian Journal of Experimental Agriculture* 44(7): 675-680. doi:10.1071/EA02172.

Su, H., Sha, K., Zhang, L., Zhang, Q., Xu, Y., Zhang, R., Li, H. & Sun, B. (2014). Development of near infrared reflectance spectroscopy to predict chemical composition with a wide range of variability in beef. *Meat Science* 98(2): 110-114. doi:10.1016/j.meatsci.2013.12.019.

Suman, S. P., Hunt, M. C., Nair, M. N. & Rentfrow, G. (2014). Improving beef color stability: Practical strategies and underlying mechanisms. *Meat Science* 98(3), 490-504. doi:10.1016/j.meatsci.2014.06.032.

Sun, S., Guo, B., Wei, Y. & Fan, M. (2012). Classification of geographical origins and prediction of $\delta^{13}C$ and $\delta 1^{5}N$ values of meat by near infrared reflectance spectroscopy. *Food Chemistry* 135(2): 508-514. doi:10.1016/j.foodchem.2012.05.004.

Sun, X., Lu, L., Zhang, J., Zhang, S. & Sun, B. (2011). Research on prediction chemical composition of beef by near infrared reflectance spectroscopy. *Spectroscopy and Spectral Analysis* 31(2): 379-383. doi:10.3964/j.issn.1000-0593(2011)02-0379-05.

Tang, Y., Horikoshi, M. & Li, W. (2016). *ggfortify*: Unified interface to visualize statistical result of popular R packages. *The R Journal* 8(2): 474-485. doi:10.32614/RJ-2016-060.

Tang, Y., Jones, E. & Minasny, B. (2020). Evaluating low-cost portable near infrared sensors for rapid analysis of soils from South Eastern Australia. *Geoderma Regional* 20: e00240. doi:10.1016/j.geodrs.2019.e00240.

Tedeschi, L. O. (2006). Assessment of the adequacy of mathematical models. *Agricultural Systems* 89(2-3): 225-247. doi:10.1016/j.agsy.2005.11.004.

Teixeira dos Santos, C. A., Lopo, M., Páscoa, R. N. M. J. & Lopes, J. A. (2013). A review on the applications of portable near-infrared spectrometers in the agro-food industry. *Applied Spectroscopy* 67(11): 1215-1233. doi:10.1366/13-07228.

Thomas-Bachli, A. L., Pearl, D. L., Friendship, R. N. & Berke, O. 2014. Exploring relationships between whole carcass condemnation abattoir data, non-disease factors and disease outbreaks in swine herds in Ontario (2001-2007). *BMC Research Notes* 7: 185. doi:10.1186/1756-0500-7-185.

Thompson, J. M. (2004). The effects of marbling on flavour and juiciness scores of cooked beef, after adjusting to a constant tenderness. *Australian Journal of Experimental Agriculture* 44(7): 645-652. doi:10.1071/EA02171.

Thompson, J., Polkinghorne, R., Gee, A., Motiang, D., Strydom, P., Mashau, M., Ng'ambi, J., de Kock, R. & Burrow, H. (2010). Beef palatability in the Republic of South Africa: implications for niche-marketing strategies. *ACIAR Technical Reports* No. 72. Canberra: Australian Centre for International Agricultural Research.

Toohey, E. S., van de Ven, R. & Hopkins, D. L. (2018). The value of objective online measurement technology: Australian red meat processor perspective. *Animal Production Science* 58(8): 1559-1565. doi:10.1071/AN17775.

Ueda, Y., Watanabe, A., Higuchi, M., Shingu, H., Kushibiki, S. & Shinoda, M. (2007). Effects of intramuscular fat deposition on the beef traits of Japanese Black steers (Wagyu). *Animal Science Journal* 78(2): 189-194. doi:10.1111/j.1740-0929.2007.00424.x.

Uzal, F. A., More, S. J., Dobrenov, B. & Kelly, W. R. (2002). Assessment of organoleptic post-mortem inspection techniques for bovine offal. *Australian Veterinary Journal* 80(1): 70-74. doi:10.1111/j.1751-0813.2002.tb12841.x.

Van Elswyk, M. E. & McNeill, S. H. (2014). Impact of grass/forage feeding versus grain finishing on beef nutrients and sensory quality: The U.S. experience. *Meat Science* 96(1): 535-540. doi:10.1016/j.meatsci.2013.08.010.

Viljoen, M., Hoffman, L. C. & Brand, T. S. (2007). Prediction of the chemical composition of mutton with near infrared reflectance spectroscopy. *Small Ruminant Research* 69(1-3): 88-94. doi:10.1016/j.smallrumres.2005.12.019.

Wadoux, A. M. J-C., Malone, B., Minasny, B., Fajardo, M. & McBratney, A. B. (2021). Soil spectral interference with R: Analysing digital soil spectra using the R programming environment. In A. E. Hartemink & A. B. McBratney (Eds.), *Progress in Soil Science*. Springer. doi:10.1007/978-3-030-64896-1.

Wang, W. & Peng, Y. (2018). Hyperspectral imaging for assessing quality and safety of meat. In A. I. L. Maldonado et al. (Eds.), *Hyperspectral Imaging in Agriculture, Food and Environment* (pp. 65-84). InTech Open. doi:10.5772/intechopen.74371.

Webber, J. J., Dobrenov, B., Lloyd, J. & Jordan, D. (2012). Meat inspection in the Australian red-meat industries: past, present and future. *Australian Veterinary Journal* 90(9): 363-369. doi:10.1111/j.1751-0813.2012.00972.x.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Milton Bache, S., Müller, K.,

Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. & Yutani, H. (2019). Welcome to the Tidyverse. *The Journal of Open Source Software* 4(43): 1686. doi:10.21105/joss.01686.

Williams, P. C. (2001). Implementation of near-infrared technology. In P. C. Williams & K. H. Norris (Eds.), *Near Infrared Technology in the Agricultural and Food Industries* (2nd edn, pp. 145-171). St Paul, Minnesota: American Association of Cereal Chemists.

Williams, P. C. (2014). The RPD statistic: a tutorial note. *NIR News* 25(1): 22-26. doi:10.1255/nim.1419.

Williams, P., Antoniszyn, J. & Manley, M. (2019). *Near-infrared technology – getting the best out of light*. African Sun Media. doi:10.18820/9781928480310.

Williams, P., Dardenne, P. & Flinn, P. (2017). Tutorial: Items to be included in a report on a near infrared spectroscopy project. *Journal of Near Infrared Spectroscopy* 25(2): 85-90. doi:10.1177/0967033517702395.

Wilson, C. S., Jenkins, D. J., Barnes, T. S. & Brookes, V. J. (2019). Evaluation of the diagnostic sensitivity and specificity of meat inspection for hepatic hydatid disease in beef cattle in an Australian abattoir. *Preventative Veterinary Medicine* 167: 9-15. doi:10.1016/j.prevetmed.2019.03.014.

Wold, J. P., Johansen, I-R., Haugholt, K. H., Tschudi, J., Thielemann, J., Segtnan, V. H., Narum, B. & Wold, E. (2006). Non-contact transflectance near infrared imaging for representative on-line sampling of dried salted coalfish (Bacalao). *Journal of Near Infrared Spectroscopy* 14(1): 59-66. doi:10.1255/jnirs.587.

Wu, D., Sun, D-W. & He, Y. (2012). Application of long-wave near infrared hyperspectral imaging for measurement of color distribution in salmon fillet. *Innovative Food Science & Emerging Technologies* 16: 361-372. doi:10.1016/j.ifset.2012.08.003.

Xu, J-L. & Sun, D-W. (2017). Hyperspectral imaging technique for on-line monitoring of meat quality and safety. In F. Toldrá & L. M. L. Nollet (Eds.), *Advanced Technologies for Meat Processing* (2nd edn, pp. 17-82). CRC Press. doi:10.1201/9781315152752-2.

Xu, Y., Zhong, P., Jiang, A., Shen, X., Li, X., Xu, Z., Shen, Y., Sun, Y. & Lei, H. (2020). Raman spectroscopy coupled with chemometrics for food authentication: a review. *TrAC Trends in Analytical Chemistry* 131: 116017. doi:10.1016/j.trac.2020.116017.

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.

Yang, C-C., Chao, K. & Kim, M. S. (2009). Machine vision system for online inspection of freshly slaughtered chickens. *Sensing and Instrumentation for Food Quality and Safety* 3(1): 70-80. doi:10.1007/s11694-008-9067-8.

Yang, J. S., Meng, Q. X., Ren, L. P., Zhou, Z. M. & Xie, X. X. (2010). Rapid evaluation of beef quality by NIRS technology. *Spectroscopy and Spectral Analysis* 30(3): 685-687. doi:10.3964/j.issn.1000-0593(2010)03-0685-03.

Zając, A., Hanuza, J. & Dymińska, L. (2014). Raman spectroscopy in determination of horse meat content in the mixture with other meats. *Food Chemistry* 156: 333-338. doi:10.1016/j.foodchem.2014.02.002.

Zamora-Rojas, E., Pérez-Marín, D., De Pedro-Sanz, E., Guerrero-Ginel, J. E. & Garrido-Varo, A. (2012). Handheld NIRS analysis for routine meat quality control: Database transfer from at-line instruments. *Chemometrics and Intelligent Laboratory Systems* 114: 30-35. doi:10.1016/j.chemolab.2012.02.001.

Zeaiter, M., Roger, J-M. & Bellon-Maurel, V. (2005). Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. *TrAC Trends in Analytical Chemistry* 24(5): 437-445. doi:10.1016/j.trac.2004.11.023.

Zhao, M., Downey, G. & O'Donnell, C. P. (2015). Dispersive Raman spectroscopy and multivariate data analysis to detect offal adulteration of thawed beefburgers. *Journal of Agricultural and Food Chemistry* 63(5): 1433-1441. doi:10.1021/jf5041959.

Zia, A., Liang, J., Zhou, J. & Gao, Y. (2015). 3D reconstruction from hyperspectral images. In *Proceedings of the IEEE Workshop on Applications of Computer Vision* (January 5-9, 2015), Waikoloa, Hawaii (pp. 318-335). doi:10.1109/WACV.2015.49.

Zontov, Y. V., Balyklova, K. S., Titova, A. V., Rodionova, O. Y. & Pomerantsev, A. L. (2016). Chemometric aided NIR portable instrument for rapid assessment of medicine quality. *Journal of Pharmaceutical and Biomedical Analysis* 131: 87-93. doi:10.1016/j.jpba.2016.08.008.

Zou, X., Huang, X. & Povey, M. (2016). Non-invasive sensing for food reassurance. *Analyst* 141(5): 1587-1610. doi:10.1039/C5AN02152A.