Developing Raman Spectroscopy for Neuromuscular Disorders

Maria Plesia

Supervisors: Dr J.J.P. Alix, Dr R.J. Mead, Dr J.C. Day

Sheffield Institute for Translational Neuroscience
University of Sheffield

Submitted for the degree of Doctor of Philosophy (PhD)

October 2020

Acknowledgments

First and foremost, I wish to thank my supervisors, Dr James Alix, Dr Richard Mead and Dr John Day for giving me the opportunity to get involved in this project in the first place and learn so much. A huge thank you to Dr James Alix for his incredible guidance and constant support throughout the PhD and for his patience during the final months of writing up. I would also like to thank Dr John Day for always being so very generous with his time and knowledge and for his constant encouragement. Thank you to Dr Richard Mead for always willing to engage in interesting conversations about the project and for his ability to always see the positives in everything.

Thank you to Dr Oliver Stevens for helping me with the data analysis at the beginning of the project. A huge thank you to Dr Gavin Lloyd for always answering to my long emails, generously sharing his expertise and encouraging me throughout. I would also like to thank Dr Catherine Kendall for her support in the project and for being incredibly kind and helpful throughout. Many thanks to Ian Coldicott for all his help with the animal work throughout the project and for his patience in training me. A big thank you to Amy Keerie for her help with animal work whenever I needed it and her positivity.

A very big thank you to Matilde and Andreas for all their support and for bearing with me and my stress; going through PhDs together was so much more fun than doing it alone. I would also like to thank all my closest friends back in Greece for the beers and laughter whenever we managed to meet and for always finding a way to stay close and be there for me no matter the distance.

I would also like to thank my parents for their continuous support and encouragement, for always believing in my abilities and for teaching me to never give up. A big thank you to the Velliou and Nikolaou families for doing their best to keep me calm, supporting me and being there for me throughout. Finally, an enormous thank you to Harry for enduring all my panic and stress, always being incredibly supportive, always smiling and always believing we can make it.

Abstract

Background: Neuromuscular diseases are a broad group of conditions that affect nerves and muscles. Diagnosing these disorders can be very challenging and it often requires multiple investigations leading to delays in treatment. Additionally, monitoring disease progression also poses considerable challenges particularly for clinical trials. Spontaneous Raman spectroscopy can provide a rapid, label-free and highly specific molecular fingerprint of tissue. The purpose of this study was to explore whether Raman spectroscopy of muscle can be used to detect muscle pathology in preclinical murine models and human muscle tissue. In addition, I aimed to assess the effects of *in vivo* Raman recordings on muscle function.

Methods: In this study an *in vivo* intra-muscular fibre optic Raman technique was developed and tested in mouse models of two devastating human neuromuscular diseases: amyotrophic lateral sclerosis ($SOD1^{G93A}$ and $TDP\text{-}43^{Q331K}$) and Duchenne muscular dystrophy (*mdx*). Motor function of the animals was assessed after the recordings using the rotarod test and gait analysis in order to evaluate the effects of the *in vivo* experimental procedure. Post-mortem MRI images were acquired from the hind limbs of the animals to assess potential tissue damage. *Ex vivo* spectra from human muscle, acquired from patients with various neuromuscular disorders and healthy volunteers, were also recorded.

Results: The method was able to detect muscle pathology and discriminate between different preclinical neuromuscular disorders *in vivo* with high accuracy values. Additionally, the *in vivo* recordings appeared not to have affected the motor function of the mice and not to have caused any long-term tissue damage. The spectra acquired *ex vivo* from human muscle demonstrated similar featured to those observed in mice. Differences between healthy and diseased human tissue were apparent.

Conclusions: Spontaneous Raman spectroscopy shows promise as a translational research tool.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Expansion
ALS	amyotrophic lateral sclerosis
AUROC	area under the receiver operating characteristics
CCD	charge-coupled device
CK	creatine kinase
CMAPS	compound muscle action potentials
DMD	Duchenne muscular dystrophy
EMG	electromyography
fALS	familial amyotrophic lateral sclerosis
fdr	false discovery rate
LDA	linear discriminant analysis
LDF	linear discriminant function
LMN	lower motor neurones
LOMOCV	leave-one-mouse-out cross-validation
LOOCV	leave-one-spectrum-out cross-validation
LOSOCV	leave-one-sample-out cross validation
MND	motor neurone disease
MRI	magnetic resonance imaging
NA	numerical aperture
NCS	nerve conduction studies
NIR	near infrared
PCA	principal component analysis
PLS-DA	partial least squares discriminant analysis
QDA	quadratic discriminant analysis
RLSMOCV	repeated-leave-some-mice-out cross validation
ROS	reactive oxygen species
sALS	sporadic amyotrophic lateral sclerosis
SERS	surface enhanced Raman spectroscopy
SRS	stimulated Raman scattering
UMN	upper motor neurones
USS	ultrasound

1. Introduction

1.1 Neuromuscular Diseases

Neuromuscular diseases are a broad group of conditions that affect nerves and muscles and cause substantial morbidity and mortality. Their primary symptoms are muscle weakness and impairment of motor function. While some neuromuscular disorders, e.g. inflammatory myopathies, can be managed and treated effectively, many, e.g. motor neurone disease, are at present incurable. Moreover, diagnosis for many of these disorders can be quite challenging and often requires multiple clinical examinations and ancillary tests. Two of the mainstays of the current diagnostic pathway are electromyography (EMG) and muscle biopsy. In EMG a needle electrode is inserted into muscle and the electrical signals generated by the muscle are recorded and analysed. EMG can effectively differentiate between neurogenic disorders and disorders caused by primary muscle pathology (i.e. myogenic disorders). However, one major limitation of this technique is that findings are not specific for any given disease. For example, the EMG abnormalities seen in motor neurone disease (MND) are the same as those seen in peripheral neuropathies. A similar lack of specificity is also seen in myogenic disorders.

In muscle biopsy a small sample of tissue is removed for histological analysis. Thus, muscle biopsy is an invasive technique. Moreover, due to the small sample size pathology can be missed (Joyce, Oskarsson and Jin, 2015). Further commonly used ancillary tests include blood tests and imaging tests such as magnetic resonance imaging (MRI).

In addition to the diagnostic problems, monitoring disease progression has proven challenging for many of these disorders. Development of new biomarkers, which could be used for earlier diagnosis or monitoring disease progression, is therefore crucial and could aid timely clinical intervention and new treatment studies. A translational approach to biomarkers, linking laboratory and clinical studies may also improve our understanding of the pathophysiological mechanisms involved neuromuscular disorders.

In the following sections amyotrophic lateral sclerosis (ALS) and Duchenne muscular dystrophy (DMD) will be discussed as these are the neuromuscular disorders that are studied in detail in the preclinical aspect of this thesis. In addition, a brief introduction to the diagnostic challenge of neuromuscular disorders is also presented. These sections provide the context for the respective preclinical and human muscle sample experiments.

1.1.1 Diagnostic Challenges of Neuromuscular Disorders

Present diagnostic strategies in neuromuscular disease often include blood, neurophysiological and imaging tests. Blood tests are usually the first step in the investigations for primary muscle disease, known as myopathy. While easy to perform, blood tests can result in erroneous conclusions. For example, increased levels of serum creatinine kinase (the most widely used blood marker for muscle pathology, and the first test used to diagnose Duchenne muscular dystrophy) may also be seen in asymptomatic individuals (Prelle *et al.*, 2002) and can be poorly predictive of an underlying muscle disorder (Shaibani *et al.*, 2015).

Neurophysiological investigations are based around EMG. As already mentioned above EMG can effectively differentiate between diseases caused by nerve pathology and those caused by primary muscle disease. During the examination muscle contraction is performed and the

recruited motor units are recorded by the EMG needle; these are known as motor unit potentials. The configuration of these potentials is analysed, either qualitatively (by eye), or quantitatively, using specialist software. In myopathies, short duration, low amplitude, polyphasic potentials are present (Liguori *et al.*, 1997). However, similar potentials can be recorded during the phase of early nerve regeneration and in disorders of the neuromuscular junction. Thus, the context of such observations is crucial. Overall, the sensitivity of EMG to detect myopathic pathology has been reported between 54%-83% in various studies (Fuglsang-Frederiksen, 2006); one retrospective study in an adult setting reported sensitivity and specificity of EMG at 74% and 67%, respectively (Cardy and Potter, 2007). It is important to note that these values reflect the ability of the test to detect muscle pathology and not identify the exact diagnosis.

Imaging may also be of use in muscle disorders. For inflammatory myopathies magnetic resonance imaging (MRI) and ultrasound (USS), may be used and there are reports that suggesting that use of these tests can aid reaching a diagnosis by identifying muscle sites for biopsy (Jungbluth *et al.*, 2004; Tomasová Studynková *et al.*, 2007), as well as potentially identifying features more specific to particular diagnoses (Noto *et al.*, 2014). USS is inexpensive and has the advantage of being a bedside test but is highly operator dependent (Schiffenbauer, 2014).

As a result of the limitations of these methods, muscle biopsy remains the gold standard for diagnosis of myopathies as it can provide valuable information on muscle health and pathological features. However, there can be difficulties with muscles biopsies as well. It is an invasive and expensive procedure. Some muscles are difficult to biopsy and, as only small pieces of tissue are taken, pathology can be missed. It has been previously reported that a total of 86 biopsies were required to diagnose inclusion body myositis in 43 patients (Dahlbom, Lindberg and Oldfors, 2002), while in another report histological data from 258 biopsy samples led to a specific diagnosis in only 43% of the cases (Lai *et al.*, 2010). Some authors suggest simultaneous biopsy of multiple sites (Prayson, 2006); although decreases the possibility of missing pathology, this approach does expose patients to increased procedural risk (for example bleeding and infection) and discomfort.

Neurogenic disorders may also be difficult to diagnose. A good example of a difficult neurogenic diagnosis can be MND. The variable presentation of MND and the lack of a specific diagnostic test means that the diagnosis remains a clinical judgement, reached through exclusion of other disorders following a multi-modal approach, including radiological and neurophysiological testing (Vucic, Rothstein and Kiernan, 2014). Neurophysiological testing focuses on the pattern of non-specific EMG abnormalities across multiple cranio-spinal regions (Krarup, 2011). These findings are in both the El Escorial and Awaji-Shima diagnostic criteria, although only the latter suggests that these findings are equally important as bedside observation (Brooks *et al.*, 2000; de Carvalho *et al.*, 2008). The specificity of the EMG data varies with the highest sensitivity and specificity achieved in the limbs and the cranial and thoracic regions respectively (Jenkins *et al.*, 2016). Diagnosis of ALS and the challenges it entails are described in more detail in section 1.1.2.

In addition to diagnosis, the monitoring of neuromuscular disease progression and response to treatments poses significant challenges. This area is under intense investigation as a sensitive measure of new treatment efficacy could help in detecting subtle treatments effect and reduce the cost of clinical trials. In primary muscle disorders examples of difficulties faced include the contribution of steroid treatment to weakness in patients with inflammatory muscle disease,

when it can be difficult to understand if the weakness is due to the disease only, or if the steroid treatments also contributes to it.

1.1.2 Neuromuscular disease mouse models and translation: Differences in human and mouse skeletal muscle

Mice and humans share part of their genome and have many common physiological and pathological features (Rosenthal and Brown, 2007). Comparative analysis of human and mouse genomes has provided insight into gene homologues and allowed genomic manipulation that led to the creation of transgenic, knock-in and knockout mice that can be used as models of human pathology (Perlman, 2016). Furthermore, breeding and maintaining mice is easy and much more inexpensive than other larger mammals (Rosenthal and Brown, 2007; Justice and Dhillon, 2016; Perlman, 2016). Hence, the laboratory mouse is a widely employed model organism in human biology and disease research (Rosenthal and Brown, 2007; Justice and Dhillon, 2016; Perlman, 2016).

In regard to neuromuscular disease, a wide range of mouse models of various neuromuscular disorders have been developed and have offered valuable insights into pathophysiological mechanisms and the development of new drug candidates and therapies (Gurney *et al.*, 1994; Hsieh-Li *et al.*, 2000; Burgess, Cox and Seburn, 2016). However, preclinical studies are often hampered by poor translation that can partially be attributed to poor experimental design and biological noise in the mouse model backgrounds (Scott *et al.*, 2008; Mead *et al.*, 2011). Additionally, anatomical and physiological differences between mice and humans can lead to erroneous conclusions in preclinical studies and to mouse models that do not fully recapitulate the human disease phenotype and progression and, hence, to findings that are not translatable. Despite well documented similarities in various internal organ systems, like the musculoskeletal, endocrine and cardiovascular, notable differences in, for example, ontogeny, immunology and pathology of mice and human have also been reported (Rosenthal and Brown, 2007; Hu *et al.*, 2017). In the rest of this section differences in the musculoskeletal systems of the two species, and in particular, the skeletal muscle, which is of particular interest for this study, are presented.

Mice have a much smaller body mass than humans and are also quadrupedal animals. Hence, in order to meet their movement requirements, their muscles differ in size, architecture and geometry (Hu *et al.*, 2017). In a study comparing human and mouse neuromuscular junctions (NMJ) Jones *et al*. showed that human NMJs were only half the size of the NMJs that innervate mouse muscle and their axons were only a third of the calibre of their mouse counterparts. They also showed that the muscle fibres that human NMJs innervate are up to twice the diameter of those in mice (Jones *et al.*, 2017). Regarding muscle fibre length, it has been shown that the relative optimal fibre length, calculated as the ratio of the fibre length over the muscle belly length, allometrically decreases with body mass (Biewener, 1990, 2005). There is also an allometric increase of the muscle moment arm with the body mass (Eng *et al.*, 2008; Hu *et al.*, 2017). Hence, mouse skeletal muscle fibres have smaller optimal fibre length and exhibit larger moment arm. Body size plays an important role also in energy metabolism (Schiaffino and Reggiani, 2011). Energy metabolism for unit body mass is inversely related to body size. Thus, the metabolic activity of skeletal muscle, which is the most abundant tissue in the body, is higher in muscle of small than large animals and mitochondrial volume density in muscle fibres has been shown to be much smaller in humans than small mammals (Schiaffino and Reggiani, 2011).

Differences in gait kinematics as well as muscle geometry and architecture can lead to biomechanical differences between human and mice limbs as shown in a study conducted by Hu *et al*. (Hu *et al.*, 2017). In this study joint kinematics acquired from existing studies and musculoskeletal modelling were used to simulate muscle-tendon dynamics in order to compare fibre length changes of mouse hindlimb muscle and their counterparts in human lower limbs in walking (Hu *et al.*, 2017). It was shown that during walking, 19 out of 25 hindlimb mouse muscles experience significantly smaller fibre excursions (48 ± 19%) compared to the homolog muscles in humans (Hu *et al.*, 2017). It was also shown that reduced joint excursions and smaller muscle moment arms in mice musculoskeletal system primarily led to the smaller fibre excursions (Hu *et al.*, 2017). Such biomechanical differences during everyday activities, like gait, could lead to differences in disease phenotypes between mouse models and patients that can then lead to non-translatable pre-clinical findings (Hu *et al.*, 2017). An example of such model could be the *mdx* mouse model of DMD. As discussed in various sections in this thesis, the *mdx* mouse exhibits a mild phenotype compared to DMD patients. Differences in fibre excursions, i.e. in the magnitude of muscle lengthening and shortening in each gait cycle, suggest that mouse limbs could work under different repeated biomechanical loads in walking (Hu *et al.*, 2017). A large magnitude of muscle lengthening has been shown to cause more damage to muscles, particularly dystrophic ones (Petrof *et al.*, 1993; Brooks, Zerba and Faulkner, 1995; Consolino and Brooks, 2004). Thus, the smaller fibre excursions in the mouse hindlimb may cause less muscle damage and may contribute, along with other mechanisms, to the milder phenotype in *mdx* mice muscles compared to patients with DMD (Hu *et al.*, 2017).

Skeletal muscle is a heterogeneous tissue with muscle fibres that exhibit different morphological and functional characteristics (Scott, Stevens and Binder-Macleod, 2001). The diversity between muscle fibres stems from the existence of multiple myofibrillar proteins isomers (myosin isoforms being the most important), different metabolic enzymes present in each fibre (predominance of glycolytic or mitochondrial activities) and from differences in various subcellular systems, like intracellular calcium signalling (Schiaffino, 2010; Schiaffino and Reggiani, 2011). Hence, muscle fibres can be classified into groups based on properties such as contractile speed, myosin heavy chain (MyHC) expression, and metabolic capacity (Bloemberg and Quadrilatero, 2012). Due to the central role of myosin in muscle cell physiology and the different MyHC isoforms distributed in various fibres, MyHC is a very commonly used marker for fiber typing (Schiaffino and Reggiani, 2011). Muscle fibres of small mammals, like mouse and rat, contain four major MyHC isoforms, namely MyHCI (slow isoform) and MyHCIIa, MyHCIIx, MyHCIIb (fast isoforms) (Scott, Stevens and Binder-Macleod, 2001). Fibres expressing MyHCI are termed type I fibres, whereas fibres expressing MyHCIIa, MyHCIIx, and MyHCIIb are termed type IIA, type IIX, and type IIB fibres, respectively (Bloemberg and Quadrilatero, 2012). In addition, 'hybrid' fibres containing two MyHC isoforms can also be present in muscle in the following sequence: 1 ↔ 1/2A ↔ 2A ↔ 2A/2X ↔ 2X ↔ 2X/2B ↔ 2B (Schiaffino and Reggiani, 2011). However, this is not an obligatory pattern of MyHC gene expression since, for example, fibres co-expressing MyHCI and MyHCIIx but not MyHCIIa have been previously detected (Caiozzo *et al.*, 2003; Schiaffino and Reggiani, 2011). With the discovery of type 2X fibre types in rodents, it was shown that human fibres previously typed as 2B by ATPase staining actually contained the human orthologue of rat MyHCIIx (Smerdu *et al.*, 1994; Ennion *et al.*, 1995; Pette, Peuker and Staron, 1999). Similar to other large mammals, the

human MyHCIIb gene is only expressed in extraocular and laryngeal muscles (Andersen *et al.*, 2000; Smerdu and Cvetko, 2013). Hence, the main difference between human and mouse muscle fibres arises from the fact that, unlike rodent muscle, human limb muscles only express three myosin isoforms and do not contain type 2B fibres (Pette and Staron, 1997; Hilber *et al.*, 1999; Scott, Stevens and Binder-Macleod, 2001; Schiaffino, 2010; Schiaffino and Reggiani, 2011). Hybrid MyHC expression in different fibres allows, thus, for the following scheme in human skeletal muscle: 1 ↔ 1/2A ↔ 2A ↔ 2A/2X ↔ 2X (Talbot and Maves, 2016).

Apart from the fact that human skeletal muscle fibres contain less MyHC isoforms and have thus less hybrids, the relative proportions and metabolic properties of muscle fibres also vary between humans and rodents (Schiaffino, 2010). In general, there is greater abundance of slow fibres in human muscle, which is, thus, primarily composed by type 1 and type 2A fibres. Type 2X fibres are a relatively small component in most individuals (Schiaffino, 2010). On the contrary, mouse muscle primarily contains type 2B and 2X fibres, with type 2A fibres being a relatively limited component and type 1 fibres being extremely rare and mostly confined to specific muscles (for example the soleus) (Schiaffino, 2010). Additionally, the maximum velocity of muscle shortening within each fibre type decreases with increasing body size (Rome, Sosnicki and Goble, 1990; Seow and Ford, 1991; Pellegrino *et al.*, 2003; Marx, Olsson and Larsson, 2006). Hence, slow fibres in human skeletal muscle are slower than those in mouse muscle (Pellegrino *et al.*, 2003).

In regard to metabolic activity, staining for SDH activity is weaker in human compared to mouse muscle sections, showing that the oxidative enzyme complement is different in the two species and suggesting higher mitochondrial content and oxygen consumption in mice muscles (Schiaffino, 2010; Schiaffino and Reggiani, 2011). Additionally, whereas in human muscles the abundance of mitochondria and oxidative enzymes is greatest in type 1 fibres and lowest in 2X fibres, in mouse and rat muscles the oxidative potential is highest in 2A fibres and lowest in 2B fibres (Schiaffino, 2010; Schiaffino and Reggiani, 2011).

Differences between human and mouse skeletal muscle satellite cells have also been observed, despite the cells sharing characteristics of morphology and surface markers (Bareja *et al.*, 2014; Mierzejewski *et al.*, 2020). For example, apart from PAX7, which is the canonical marker of satellite cells in both species, mouse satellite cells also synthesise other markers, like CD34, c-MET, integrin a7 and nestin (Mierzejewski *et al.*, 2020). Such set of markers has not been detected in human satellite cells. Additionally, in differentiating mouse satellite cells cultured in vitro for one, two and three days both PAX7 and MYOD transcription factors were present in most of the nuclei (Zammit *et al.*, 2004). On the contrary, in human muscles MYOD was only found in a small number of satellite cells early on and only at half of the differentiating myoblasts after six days of in vitro culture suggesting, thus, differences in the myogenic differentiation between mice and humans (Bareja *et al.*, 2014; Mierzejewski *et al.*, 2020). Differences in satellite cells and myogenic differentiation can largely affect skeletal muscle growth and regeneration, and can, thus, potentially complicate the clinical translation of, for example, drugs validated in mouse models.

To summarise, there are large differences between human and mouse muscle, including differences in muscle fibre size, architecture, type, type profile, energy metabolism and biomechanical properties. It is, thus, important to keep in mind these species differences when trying to extrapolate conclusions derived from studies on mouse muscle from models of neuromuscular disease and cautious interpretation of the results is required.

1.1.3 Amyotrophic Lateral Sclerosis

Motor neurone disease (MND) is a progressive neurodegenerative disease that affects motor neurones in the brain and spinal cord. ALS is the most common form of MND and represents 75% of all MND cases (Hobson *et al.*, 2016).

Worldwide, ALS has an annual incidence rate of approximately 1-2 per 100,000 individuals, and a prevalence rate of approximately 5 per 100,000 people (Ferraiuolo *et al.*, 2011; Löscher and Feldman, 2014). The overall lifetime risk of ALS is higher for men (1:350) than women (1:400) (Kiernan *et al.*, 2011). The mean age of clinical onset of the disease is around 55-60 years, with an onset before the age of 40 and after the age of 80 being quite unusual (Ferraiuolo *et al.*, 2011; Kiernan *et al.*, 2011; Hobson *et al.*, 2016). Median survival of ALS patients is approximately 2-3 years from onset of symptoms (Couratier *et al.*, 2016). However, both age at disease onset and disease duration vary considerably among different patients (Régal *et al.*, 2006; Robberecht and Philips, 2013).

In ALS both upper motor neurones (UMN) in motor cortex and lower motor neurones (LMN) in the brainstem and spinal cord degenerate (Robberecht and Philips, 2013). Corticospinal neurones are directly or indirectly connected with spinal motor neurones which innervate skeletal muscles, controlling thus voluntary muscle movement. In ALS, corticospinal motor neurones degenerate and their descending axons in the lateral spinal cord harden and appear scarred; it is this appearance, together with muscle wasting, that led Charcot to propose the name ALS (Taylor, Brown and Cleveland, 2016). In addition, spinal and brainstem motor neurones are lost, which leads to secondary denervation and muscle wasting (Taylor, Brown and Cleveland, 2016). The initial axonal retraction and denervation of the lower motor neurones or the muscles is compensated by re-innervation from collateral sprouts of neighbouring surviving motor axons (Robberecht and Philips, 2013). During the course of the disease, as more resistant neurones degenerate, this mechanism fails. This leads to apparent loss of motor neurones, muscle atrophy and fasciculations (Robberecht and Philips, 2013). Moreover, accumulating evidence suggests that primary muscle degeneration may also be involved in ALS pathogenesis (Shi *et al.*, 2010; Moloney, de Winter and Verhaagen, 2014). ALS starts at a specific area and contiguously spreads to other parts of the body (Bäumer, Talbot and Turner, 2014). In many patients with ALS, prefrontal and temporal cortices are also affected, although the degree they are involved varies (Bäumer, Talbot and Turner, 2014). Deterioration of the frontal and temporal cortices results in executive dysfunction (up to 50% of ALS patients develop subtle dysexecutive syndrome), behavioural changes and for up to 15% of ALS cases to frontotemporal dementia (Bäumer, Talbot and Turner, 2014).

ALS is subdivided into two groups: familial ALS (fALS), that represents approximately 10% of cases and sporadic ALS (sALS) that accounts for 90% of all ALS cases (Wong and Martin, 2010; Robberecht and Philips, 2013; Al Sultan *et al.*, 2016). fALS is inherited with a Mendelian pattern and usually as an autosomal dominant trait. Currently, there are 26 known DNA loci associated with ALS and ALS-FTD (Al Sultan *et al.*, 2016). The most common genetic causes of FALS are mutations observed in one of the following genes: C9ORF72 (40-50%), SOD1 (20%), TARDBP (5%) and FUS (5%) (Al Sultan *et al.*, 2016). Mutations in these genes are also responsible for a small amount of SALS cases (approximately 10%). Identification of the genetic variants of ALS has aided the understanding of vital molecular mechanisms involved in ALS pathogenesis and progression.

Many molecular mechanisms have been identified and described including excitotoxicity, protein aggregation, oxidative stress, mitochondrial dysfunction, dysregulation of RNA processing, microglial activation and impaired axonal transport. There is also strong evidence that many of these mechanisms intersect at different points. However, for many of these mechanisms it is yet unclear if they are pathogenic or if they are a result of the disease. In addition to genetic and molecular mechanisms, several environmental factors have also been suggested to increase the risk of ALS and tested in several epidemiological studies. Examples of such factors are smoking, exercise and cyanotoxins (Al-Chalabi and Hardiman, 2013). However, there are not any definitive large-scale environmental risk factors identified yet.

There is no established diagnostic test for ALS, with the clinical examination findings of signs of UMN and LMN dysfunction being therefore the most important part of the diagnostic procedure (Löscher and Feldman, 2014). Ancillary investigations including neurophysiological studies and neuroimaging are also used to support the clinical diagnosis and to exclude rare mimics (Brooks *et al.*, 2000; de Carvalho *et al.*, 2008; Nzwalo *et al.*, 2014). LMN pathology can best be identified with electromyography (EMG) and nerve conduction studies (NCS) (Hobson *et al.*, 2016). EMG signs of acute and chronic denervation in various areas of the body support the clinical diagnosis of ALS (Cooper-Knock, Jenkins and Shaw, 2013). Moreover, rare mimics such as inclusion body myositis can usually be excluded. NCS are mainly employed in order to exclude ALS-mimicking disorders such as multifocal motor neuropathy with conduction block (Löscher and Feldman, 2014).

Involvement of UMN pathology can be investigated using neuroimaging studies (Hobson *et al.*, 2016). Brain or cervico-thoracic spine MRI is mainly used to exclude alternative pathological causes, such as structural intracranial or spinal pathology/compression (Hobson *et al.*, 2016; Cooper-Knock and Jenkins, 2013). Since ALS can have several varying clinical manifestations muscle biopsy, blood tests, genetic testing and several other types of investigations may also be used in certain cases in order to increase the diagnostic certainty (Cooper-Knock, Jenkins and Shaw, 2013).

When there are clear signs of LMN and UMN degeneration in conjunction with absence of sensory system involvement, the diagnosis ALS can be clear (Kraemer, Buerger and Berlit, 2010). However, due to limited, non-specific symptoms at disease onset and the necessity of multiple supportive investigations to decrease the likelihood of an incorrect diagnosis, there is a significant delay between symptom onset and confirmed diagnosis (Paganoni *et al.*, 2014). Moreover, atypical disease presentations as well as varying symptoms at early disease stages can render the diagnosis more challenging and time consuming (Kraemer, Buerger and Berlit, 2010; Gupta *et al.*, 2012; Cooper-Knock, Jenkins and Shaw, 2013). This delay between symptom onset and definitive diagnosis usually varies between 8 and 15 months constituting a very important proportion of disease duration (Paganoni *et al.*, 2014). Apart from the psychological effects that such a delay and uncertainty can have on patients, it also prevents them from entering clinical trials and/or starting treatment at an earlier disease stage, when fewer motor neurones may be been lost (Paganoni *et al.*, 2014). As a result, patients may not be able to fully benefit from developing or existing treatments (Paganoni *et al.*, 2014). Additionally, since it is more likely that any potential treatment will be more effective at an earlier stage, earlier disease diagnosis would also facilitate clinical trials.

To date there is only one widely accepted pharmacological treatment for ALS with a limited survival benefit, riluzole (Bensimon, Lacomblez and Meininger, 1994; Cooper-Knock, Jenkins and Shaw, 2013). The response of different patients to medication with riluzole varies, but the average life extension is approximately 3 months (Miller *et al.*, 2007). Additional to this drug, a new medication, edaverone, was developed and licenced recently (May 2017) in the USA. However, this drug is not yet licensed in Europe.

Symptomatic treatments are beneficial for patients, especially in terms of improving their quality of life. Such treatments often require a multidisciplinary approach with neurologists, specialist nurses, speech and language therapists, dieticians, physiotherapists, gastroenterologist and respiratory therapists being involved (Bäumer, Talbot and Turner, 2014). Example interventions that are often considered in such cases are gastrostomy (for patients that have difficulties in swallowing) and non-invasive ventilation. Apart from quality of life improvement, non-invasive ventilation also offers a modest survival advantage (Bäumer, Talbot and Turner, 2014).

Diagnosis only after the occurrence of significant neuronal damage and lack of efficient treatment, mark the need for efficient biomarkers in ALS. A biomarker is defined as "an objective measurement that acts as an indicator of normal biological processes, pathogenic processes or pharmacologic responses to therapeutic intervention" (Turner *et al.*, 2009). In that sense a biomarker could, depending on its characteristics, aid early disease diagnosis or even prognosis, monitor disease progress, categorise patients in terms of potential beneficial treatments and evaluate new therapeutic approaches (Turner *et al.*, 2013; Benatar *et al.*, 2016; Vucic, 2016). The field has gained a lot of attention over the past 10 years and promising candidates include proteotomic studies in biofluids such as CSF and blood, neurophysiological techniques and neuroimaging techniques (Radionuclide imaging, MRI) (Pradat and Dib, 2009; Turner *et al.*, 2013; Bame *et al.*, 2014; Benatar *et al.*, 2016). Despite the potential biomarkers indicated by these techniques there still is no validated clinically implemented biomarker in ALS (Bame *et al.*, 2014).

1.1.3.1 SOD1 Gene

The SOD1 gene comprises 5 exons interspersed with 4 introns and encodes an anti-oxidant ubiquitously expressed enzyme of 153 amino acids, known as Cu, Zn superoxide dismutase (Kaur, McKeown and Rashid, 2016). The SOD1 protein functions as a homodimer (Cooper-Knock, Jenkins and Shaw, 2013). Each subunit of the protein consists of an eight-stranded beta barrel and two metal ions, a copper and a zinc (Rakhit and Chakrabartty, 2006). The metal atoms play an essential role in the anti-oxidant catalytic activity of the enzyme. SOD1 protein acts as a free radical scavenger, catalysing the production of hydrogen peroxide and oxygen from free radical superoxide (Kaur, McKeown and Rashid, 2016). Superoxide anion arises as a by-product of aerobic metabolism, mainly because of electron leakage from the respiratory chain that leads to the incomplete reduction of molecular oxygen during oxidative phosphorylation in the mitochondria (Pasinelli and Brown, 2006; Cooper-Knock, Jenkins and Shaw, 2013). SOD1 catalyzes the inactivation of superoxide through cyclical reduction and oxidation of the copper ion in its active sites (Pasinelli and Brown, 2006; Barber and Shaw, 2010). SOD1 activity prevents cellular damage from superoxide (Kaur, McKeown and Rashid, 2016).

Mutations in the SOD1 gene were the first identified genetic cause of ALS and are thus the most widely studied (Rosen *et al.*, 1993; Hilton, White and Crouch, 2015). They account for 20% of

familial ALS cases and 2-7% of sporadic ones and are mainly associated with limb disease onset (Al Sultan *et al.*, 2016). Up to date more than 180 mutations throughout the whole gene have been found to be associated with the disease (Al Sultan *et al.*, 2016). Most of the currently identified mutations are point mutations with some frameshift mutations resulting in truncated proteins also present (Hilton, White and Crouch, 2015).The main pathological feature that distinguishes SOD1 ALS cases from other forms of ALS is the existence of neuronal cytoplasmic protein aggregates of mutated SOD1 proteins instead of TDP-43 protein aggregates as is the case in most of the other ALS cases (Cooper-Knock, Jenkins and Shaw, 2013).

Due to the large amount of different mutations as well as the wide range of physiological functions affected by the mutations, determining the cellular mechanisms involved in ALS pathogenesis due to SOD1 mutations has been challenging. Initially it was suspected that reduced dismutase activity mainly contributed to ALS pathogenesis (Hilton, White and Crouch, 2015). However, the fact that specific SOD1 mutants do retain their full enzymatic activity along with evidence that mutated SOD1 gene knockout in mice does not result in developing ALS led to rejection of that hypothesis (Reaume *et al.*, 1996; Barber and Shaw, 2010; Hilton, White and Crouch, 2015). Thus, it is now widely accepted that SOD1 mutations cause ALS through one or more gained toxic functions (gain of toxic function), the detailed nature of which still remains undetermined (Pasinelli and Brown, 2006). Various mutually compatible pathogenic mechanisms have been suggested to contribute to the toxicity of mutated SOD1 proteins and some of them will be discussed in the following paragraphs.

The important anti-oxidant enzymatic activity of superoxide dismutase suggests that oxidative stress could play a crucial role in the pathogenesis of SOD1-ALS (Bozzo, Mirra and Carrì, 2017). Oxidative stress results from imbalanced production and removal of reactive oxygen species (ROS) along with a decreased ability of the biological system to control damage induced by the elevated amount of ROS (Ferraiuolo *et al.*, 2011). ROS generation increases with age and oxidative stress could therefore be the mechanism that, in middle or later life, starts inhibiting the ability of a biological system to cope with a toxic insult such as a mutation, leading thus to neurodegeneration (Barber and Shaw, 2010; Ferraiuolo *et al.*, 2011). Moreover, there is strong evidence of oxidative damage (i.e. a role of oxidative stress in ALS) to several biomolecules (lipids, proteins, DNA, mRNA) in human post-mortem analysed tissue of familial and sporadic ALS patients as well as SOD1 mutated mouse models (Shibata *et al.*, 2001; Chang *et al.*, 2008; Ferraiuolo *et al.*, 2011; D'Amico *et al.*, 2013; Bozzo, Mirra and Carrì, 2017).

Various aberrant oxidative reactions catalysed by mutated SOD1 proteins, have been proposed to cause oxidative stress in SOD1 induced ALS (Pasinelli and Brown, 2006). The idea behind oxidative stress caused by any of these aberrant chemical reactions is that altered conformation of the mutated protein allows various different substrates to enter the active site and react with any of the metal ions and is thus still based on an altered catalytic activity of the enzyme (Barber and Shaw, 2010; Ferraiuolo *et al.*, 2011). However, it has been shown that mice with mutations that lead to a limited amount or complete lack of SOD1-bound copper and thus with a reduced or completely lost dismutase activity still develop motor neuron disease (Wong *et al.*, 2000; Wang *et al.*, 2003; Pasinelli and Brown, 2006). The contribution of aberrant chemistry in oxidative stress and ALS pathogenesis has thus come into question and new mechanisms that don't rely on the enzyme's catalytic activity have been proposed. For example, a proposed mechanism involves

dysregulation of transcription factor nuclear erythroid 2-related factor 2 (Nrf-2) mediated signaling anti-oxidant pathways due to a down-regulation of Nrf-2 (Cookson *et al.*, 2002; Barber and Shaw, 2010). Increased NADPH oxidase mediated superoxide production by mutated SOD1 proteins in microglia leading to prolonged ROS production has also been proposed (Harraz *et al.*, 2008). Although these mechanisms are likely to contribute to the progression of the disease, the degree to which they contribute to the gain of toxic function of the protein that leads to ALS pathogenesis still remains elusive.

Protein instability and subsequent protein aggregation have also been proposed to play a crucial role in SOD1 toxicity since SOD1 inclusions have been found in motor neurons and astrocytes of transgenic SOD1 mice and fALS patients (Pasinelli and Brown, 2006; Ferraiuolo *et al.*, 2011). Protein instability caused either by mutations that lead to misfolding or demetallation, or by the dissociation of wild type and mutant SOD1 dimers into monomers due to oxidative stress related damage has been shown to render mutated SOD1 proteins more prone to aggregation (Barber and Shaw, 2010). Interestingly, it has also been reported that mutations associated with a more aggressive phenotype of the disease in transgenic SOD1 mice produce proteins that are more susceptible to formation of inclusions (Lindberg *et al.*, 2005; Sato *et al.*, 2005; Pasinelli and Brown, 2006). Moreover, Munch et al. suggested that SOD1 protein misfolding, once present, spreads through molecules and neighbouring cells in a prion-like way and SOD1 aggregates are built (Munch, O'Brien and Bertolotti, 2011). Suggested mechanisms by which these protein inclusions might contribute to SOD1 toxicity include decrease of dismutase activity due to aggregation and overwhelming of proteosomal activity which could lead to sequestering of specific heat shock proteins and subsequent impairment of their chaperone and/or anti-apoptotic function (Pasinelli and Brown, 2006; Kaur, McKeown and Rashid, 2016). Although protein inclusions are a hallmark of ALS and other neurodegenerative disorders it is still unknown if they are a cause of disease pathogenesis or a consequence of the disease. It is also still not clear if they are toxic or favourable in terms of sequestering harmful proteins (Barber and Shaw, 2010).

Mitochondrial dysfunction is a feature of SOD1 induced ALS and various studies have focused on its role in disease pathogenesis. Once again though it is unknown whether it is a primary or secondary pathological event. Mitochondria are important cellular organelles that are involved in various cellular functions such as energy metabolism, calcium homeostasis as well as apoptosis initiation and regulation (Cozzolino and Carrì, 2012). There are numerous examples of evident mitochondrial morphological abnormalities and dysfunction in cell lines, mutant mice and ALS patients including SOD1 protein aggregates in vacuolated mitochondrial intramembrane space, defects in respiratory chain function, impaired calcium buffering, and altered antioxidant defense mechanism. In addition to externally inserted mutated proteins, it is also believed that mutant SOD1 could directly damage mitochondria due to SOD1 protein expression within the mitochondrion through various toxic mechanisms, which could lead to initiation of apoptosis, change in the mitochondrial proteome and/or oxidative damage (Higgins *et al.*, 2002; Pasinelli and Brown, 2006; Barber and Shaw, 2010). Apart from oxidative stress, protein aggregation and apoptosis, there also is a possible crosstalk between mitochondrial dysfunction and other ALS pathogenesis related mechanisms such as excitotoxicity and ER stress.

Disrupted axonal transport and disorganization of axonal cytoskeleton are common pathological features of ALS and are thought to contribute to SOD1 mediated disease pathogenesis

(Robberecht and Philips, 2013). The cytoskeleton is responsible for determining the cell shape and it also aids intracellular movement of organelles. Disorganisation of cytoskeleton in terms of abnormal neurofilament aggregation is observed in ALS patients and transgenic mice (Barber and Shaw, 2010). This could be triggered by reduced expression of neurofilament–light (NF-L), or by ROS induced damage to NF-L subunits due to mutant SOD1 proteins and increased oxidative stress respectively (Barber and Shaw, 2010). Due to the crucial role of the cytoskeleton in intracellular transport and in the maintenance of axonal caliber, neurofilament aggregates have been suggested to physically affect axonal anterograde transport (Sasaki and Iwata, 1996; Rao and Nixon, 2003). Motor neurones are asymmetrical cells with long axons. Axonal transport of various components such as proteins and organelles from the cell body to the axons and synapses (anterograde transport) and vice versa (retrograde transport) is therefore essential for their survival. Slowing of anterograde and retrograde routes was evident in SOD1 transgenic mice before the initiation of neurodegeneration and in also a clinical feature in ALS patients (Ferraiuolo *et al.*, 2011). Moreover, it has been indicated that axonal transport is affected by mutant SOD1 in a cargo-specific way. Such an example is the disruption of the anterograde transport route of mitochondria but not the retrograde one (Kieran *et al.*, 2005; De Vos *et al.*, 2007; Bilsland *et al.*, 2010; Ferraiuolo *et al.*, 2011). There is also potential crosstalk between disrupted axonal transport and several other mechanisms that are related to ALS pathogenesis and progression. Defective transport of mitochondria due to impaired mitochondria function (mentioned above) could lead to impaired axonal transport of other cargoes as well due to lack of energy, required for that transport (Ferraiuolo *et al.*, 2011). Moreover, elevated levels of ROS are known to cause inhibition to axonal transport (Kaur, McKeown and Rashid, 2016).

Other mechanisms such as excitotoxicity, endoplasmatic reticulum stress, impairment of ubiquitin–proteasome system (UPS) and other protein degradation mechanisms have also been implicated in SOD1 mediated ALS pathogenesis, although the degree to which they contribute to the toxic gain of function is still unknown. Moreover, SOD1 mutants seem to play an important role in function of non-neuronal cells. Mutant SOD1 protein aggregates are also found in glial cells (Tobisawa *et al.*, 2003; Kaur, McKeown and Rashid, 2016). Furthermore, misfolded SOD1 proteins within microglia and astrocytes in conjunction with activation of these cells due to damage in motor neurons can lead to inflammation and a subsequent release of elevated levels of toxic factors (Boillée, Vande Velde and Cleveland, 2006; Kaur, McKeown and Rashid, 2016). Interaction of mutants with protein components of the neuroendocrine system could cause secretion of the mutated proteins which can then lead to microgliosis activation and neuronal cell death (Urushitani *et al.*, 2006; Kaur, McKeown and Rashid, 2016). Hence, it is obvious that SOD1 can affect many different factors and mechanisms involved in ALS pathogenesis and progression. It is therefore likely that it is a combination of several different pathways that are involved in the toxicity of SOD1.

1.1.3.2 Muscle in SOD1-mediated ALS

Skeletal muscle is affected in all ALS patients. As the disease progresses, progressive muscle denervation leads to muscle atrophy and weakness, resulting in loss of limb and bulbar function and, ultimately, respiratory failure. ALS has traditionally been described as disease that initially causes degeneration of motor neurons, which in turn causes muscle atrophy (Moloney, de Winter

and Verhaagen, 2014). However, studies on mutant SOD1 mice have shown that non-neuronal cells may also participate in disease pathogenesis (Moloney, de Winter and Verhaagen, 2014; Pansarasa *et al.*, 2014). With regard to skeletal muscle cells, it has been reported that neuromuscular junction abnormalities and muscle dysfunction occur before motor neuron degeneration (Dupuis and Loeffler, 2009; Musarò, 2010; Si *et al.*, 2014). In support of this hypothesis, two independent studies have shown that selective overexpression of mutant SOD1 in muscle only leads to muscle atrophy, possibly indicating muscle as a primary target of SOD1 toxicity (Dobrowolny *et al.*, 2008; Wong and Martin, 2010). In the study conducted by Wong et al. skeletal-muscle restricted expression of mutant SOD1 also led to neurodegeneration (Wong and Martin, 2010). Conversely, it was reported that although muscle-specific overexpression of the transcriptional coactivator, PGC-1a can delay muscle weakness, it does not affect survival (Da Cruz *et al.*, 2012; Wei *et al.*, 2013). Finally, it has been shown that disease progression is not affected by attenuation of mutant SOD1 solely in muscles (Miller *et al.*, 2006; Si *et al.*, 2014). Thus, further investigation is required in order to determine the degree to which skeletal muscle participates in triggering or maintaining, the disease. Despite this uncertainty, most of the pathophysiological mechanisms described above have also been implicated in muscle degeneration with mitochondrial dysfunction, protein aggregation and oxidative stress being the most investigated characteristics of muscle pathology.

Mutant protein aggregates (suggestive of biochemical abnormalities), elevated levels of ROS, as well as structurally altered and functionally impaired mitochondria have been found in SOD1 mouse muscle (Dobrowolny *et al.*, 2008; Halter *et al.*, 2010; Loeffler *et al.*, 2016). Decreased amounts of muscle heat shock proteins compared to the amount of such proteins in motor neurones have also been observed. This could potentially increase the intrinsic susceptibility of malformed mutant proteins and protein accumulation in the muscle (Wei *et al.*, 2013; Loeffler *et al.*, 2016). Luo et al. demonstrated that SOD1 protein aggregates and impaired mitochondrial dynamics are casually linked (Luo *et al.*, 2013). Mutant SOD1 protein aggregates formed in muscle mitochondria resulted in fragmentation of the mitochondrial network as well as in loss of mitochondrial membrane polarization (Luo *et al.*, 2013). However, the role of muscle SOD1 protein aggregates in protein misfolding and ROS production has come into question by Wei et al. (Wei *et al.*, 2012). In this study there were no SOD1 protein aggregates detected in skeletal muscle of SOD1 mice (Wei *et al.*, 2012). This was attributed to a better proteasomal activity in muscle (2 to 4-fold increased proteasomal activity was observed from pre-onset to symptomatic disease stage) (Wei *et al.*, 2012). Moreover, they also demonstrated that although there were higher levels of the soluble state of mutated SOD1 proteins, these didn't seem to directly affect mitochondrial release of ROS. Instead non-SOD1-containing protein fractions that undergo conformational changes due to mutant SOD1 toxicity seem to impair mitochondrial function and possibly affect mitochondrial ROS release (Wei *et al.*, 2012). Thus, although these mechanisms definitely play a crucial role in ALS muscle pathology, detailed understanding of their contribution and their potential crosstalk still remains elusive. Proteotomic and biochemical studies have additionally revealed changes in expression of regulatory proteins of cytoskeletal processes, iron and calcium homeostasis as well as apoptosis prompting thus the role of additional mechanisms in muscle degeneration (Chung and Suh, 2002; Capitanio *et al.*, 2012).

1.1.3.3 SOD1 Mouse Model

Various transgenic mouse models with over-expressed mutant SOD1 have been generated and used as a research model of ALS for more than 20 years. Being a well characterised mouse model, it has proven quite useful over the years in terms of identifying and testing potential pathogenic disease mechanisms, novel hypotheses and potential therapies. Moreover, structural and functional differences between mutated and wild-type SOD1 proteins have been studied (Hilton, White and Crouch, 2015). The developed disease phenotype exhibits many features that are also seen in human ALS cases including muscle denervation, weakness, atrophy and subsequent paralysis, axonal degeneration and demise of motor neurones in the spinal cord as well as activation of glial cells (Mead *et al.*, 2011; Robberecht and Philips, 2013). However, there are several limitations in the results obtained from animal-based research. Over-expression of the mutated gene, which is required in order to achieve the desired phenotype, is an important disadvantage of this mouse model as it can lead to experimental artifacts (Robberecht and Philips, 2013). Moreover, as SOD1 is only one of several genetic ALS variants, the importance and applicability of the findings obtained in the mutant SOD1 model to other types of familial and sporadic ALS cases needs to be demonstrated (Hilton, White and Crouch, 2015). The main source of criticism regarding the mutant SOD1 mouse model has been its utility in preclinical trials of new therapeutics. This arises from the fact that numerous therapeutic approaches that produced mild effects when tested in the mouse model were proven unsuccessful when tested in human (Benatar, 2007). Possible causes for that include biological noise and poor preclinical study designs (Ferraiuolo *et al.*, 2011; Mead *et al.*, 2011). Moreover, the fact that treatment studies in the mouse model sometimes start at a quite early, often pre-symptomatic disease stage in contrast with the human trials in which treatment is tested at a much later stage (usually after symptom onset) might also play a crucial role (Robberecht and Philips, 2013).

The G93A (glycine 93 changed to alanine) pseudo-wild type mutation, which does not affect the enzymatic activity of the protein, is a quite rare SOD1 mutation (Hilton, White and Crouch, 2015). However, it has been studied very thoroughly due to the fact that the SOD1-G93A transgenic mouse model was the first ALS animal model produced (generated in 1994) and it is up to date the most employed ALS model (Weydt *et al.*, 2003; Hatzipetros *et al.*, 2015; Browne and Abbott, 2016; Kim *et al.*, 2016). Multiple disease associated mechanisms and potential drug therapies have been tested in this mouse model despite the fact that due to the rareness of the mutation in human ALS cases, the applicability of the results to other types of ALS is uncertain.

1.1.3.3.1 Established *in vivo* Biomarkers

Due to the severe motor phenotype of the $SOD1^{G93A}$ mouse the main *in vivo* techniques employed to study disease onset and progression include tests that assess motor function like neurological scoring, rotarod running, grip strength and gait analysis (Weydt *et al.*, 2003). The phenotype of the $SOD1^{G93A}$ mice highly depends on the number of transgene copies with relatively high copy number models exhibiting a disease onset between 70-100 days and an endpoint between 120-160 days (Bame *et al.*, 2014; Pfohl, Halicek and Mitchell, 2015; Kim *et al.*, 2016). All of the above mentioned tests, or at least some of their outcome measures, have been shown to successfully detect changes in the motor function of transgenic (Tg) $SOD1^{G93A}$ mice at some point near or straight after disease onset (as defined in each study) and they usually remain significantly

different between the two groups as the disorder progresses until the endpoints (Miana-Mena *et al.*, 2005; Knippenberg *et al.*, 2010; Mancuso, Oliván, *et al.*, 2011; Mead *et al.*, 2011; Sun *et al.*, 2014; Oliván *et al.*, 2015). There is, however, no agreement in the literature as to which one of these tests is more sensitive, particularly in detecting changes at an earlier stage or after an intervention. Some studies claim that footprint and gait analyses can detect changes earlier, while others show that rotarod and/or grip endurance tests are more sensitive (Miana-Mena *et al.*, 2005; Knippenberg *et al.*, 2010; Mancuso, Oliván, *et al.*, 2011; Mead *et al.*, 2011; Sun *et al.*, 2014; Oliván *et al.*, 2015). Additionally, some of these tests have been shown to be very variable and require large animal numbers (rotarod, grip endurance), whereas others are time-consuming (gait and footprint analysis) or subjective (neurological scoring) (Knippenberg *et al.*, 2010; Mead *et al.*, 2011; Hatzipetros *et al.*, 2015). Imaging tools, like MRI, have also been shown to noninvasively detect changes in muscle due to disease progression (Marcuzzo *et al.*, 2011; Mead *et al.*, 2011; Caron *et al.*, 2015). For example, Marcuzzo et al. demonstrated that muscle volume in Tg mice was significantly reduced from control mice at 8 weeks of age and continued decreasing until the endpoint (18 weeks of age) (Marcuzzo *et al.*, 2011). It is worth mentioning that this difference was seen 4 weeks before a scoring system and the grip endurance test were able to detect significant differences between Tg and control mice (first observed at 12 weeks of age). Therefore, muscle atrophy and possibly biochemical changes take place in the muscle before disease onset, as this defined by functional tests (Marcuzzo *et al.*, 2011). The main disadvantages of MRI, as discussed above, are that it is a very expensive test and requires prolonged anaesthesia, rendering, thus, longitudinal measurements difficult.

Electrophysiological studies have also been shown to detect changes before functional assessment tests. For example, some studies document significant differences in certain electrophysiological parameters between Tg and control as early as 20 days of age, while impairment of classical parameters (scoring system, hanging wire, rotarod, gait analysis parameters) were not seen until after 90 days of age (Alves *et al.*, 2011). Other studies have also shown the ability of electrophysiological parameters to detect changes in muscle of Tg mice and successfully separate them from control at an early age (Mancuso, Santos-Nogueira, *et al.*, 2011; Li, Sung and Rutkove, 2013; Mancuso, Osta and Navarro, 2014). Electrophysiological recordings can, however, be technically challenging and require large animal numbers (Mancuso, Santos-Nogueira, *et al.*, 2011; Mancuso, Osta and Navarro, 2014).

In the rest of the thesis the terms MND and ALS were used interchangeably.

1.1.4 Duchenne Muscular Dystrophy

Muscular dystrophies are a group of inherited disorders, which cause progressive muscle weakness without an abnormality in the central or peripheral nervous system. Although the disorders share certain clinical features as well as dystrophic pathological findings in muscle biopsies they exhibit clinical, genetic and biochemical heterogeneity, which is the basis for their classification (Emery, 2002). DMD, which results from a mutation in the X-linked gene that encodes the protein dystrophin, is the most common and most severe muscular dystrophy in childhood (Biggar, 2006; Mah *et al.*, 2014). The incidence rate of DMD is approximately 1 in 3500 to 10000 boys (Mah *et al.*, 2014). Clinical symptoms of the disease usually appear between 2 and 5 years of age and early signs include delayed gait or alteration in gait, difficulties in jumping,

running and standing up as well as toe walking (Ge *et al.*, 2004; Ruiten, Bushby and Guglieri, 2017). Calf hypertrophy can also be observed at an early disease stage (Emery and Muntoni, 2015). Subsequently, all skeletal muscles (upper limb muscles and distal lower limb muscles) are affected and untreated children usually became wheelchair-dependent between the ages of 8–12 (Ge *et al.*, 2004; Ruiten, Bushby and Guglieri, 2017). Heart and respiratory muscles are also affected in a similar way at various stages of the disease and usually untreated patients die in their early twenties due to cardiac or respiratory failure (Ruiten, Bushby and Guglieri, 2017). Behavioural and cognitive problems as well as problems regarding language development are also present in some cases (Cyrulnik *et al.*, 2007; Flanigan, 2014).

Making a DMD diagnosis requires a clinical examination, blood tests, genetic testing and in some cases muscle biopsy. If DMD is suspected after clinical examination, a blood test is usually the first screening test undertaken in order to measure the concentration of serum creatine kinase (CK) (Ruiten, Bushby and Guglieri, 2017). CK levels are largely elevated in DMD with values in the first 5 years of life that are 10 – 20 (often 50-200) times higher than the upper limit of normal (Yiu and Kornberg, 2015). Elevated CK levels, though to a lower degree, can also be seen in newborns. Subsequently, a genetic test is performed in order to look for mutations in the dystrophin gene. Finally, in cases with clear clinical features of DMD but no apparent gene mutation, a muscle biopsy is undertaken in order to investigate the dystrophin expression in the muscles (Ruiten, Bushby and Guglieri, 2017).

Corticosteroids are currently the only pharmacological treatment effective in DMD. Corticosteroids do not treat DMD but they slow down disease progression in terms of prolonged ambulation (for up to 3 years) and decrease of cardiorepisratory function decline. (Balaban *et al.*, 2005; King *et al.*, 2007; Yiu and Kornberg, 2015). In most of the cases symptomatic treatment is also employed and a multidisciplinary approach is necessary when it comes to DMD management due to the multiple systems it affects. Novel treatment strategies include new drugs, cell and gene therapies (Cossu and Sampaolesi, 2007).

1.1.4.1 Dystrophin Gene

The dystrophin gene is located in the short arm of the X chromosome and is the largest gene in the human genome comprising 79 exons (Nowak and Davies, 2004; Biggar, 2006). Mutations that lead to DMD can be throughout the whole length of the gene, however, most of them are located between exons 2 and 10, or exons 45 and 55 (Ruiten, Bushby and Guglieri, 2017). More than half of the different mutations associated with DMD (60-70%) are large deletions resulting in frameshift errors, 7% to 11% are duplications and around 20% are small (point) mutations (White *et al.*, 2006; Oshima *et al.*, 2009; Aartsma-Rus, Ginjaar and Bushby, 2016). DMD results when the mutation in the dystrophin gene causes absolute lack of the protein. The dystrophin protein is a 427kDa protein and under normal circumstances it is expressed in the skeletal muscle, the brain and the heart. In the muscle, dystrophin is found in the inner cell membrane surface and is elevated at the costameres and cell-cell contact sites. Dystrophin forms dystrophin-associated glycoprotein complex (DAGC) at the sarcolemma by assembling several transmembrane (dystroglycan, sarcoglycan, sarcospan) and cytosolic (syntrophin, dystrobrevin and neuroneal nitric oxide synthase (nNOS)) proteins (McGreevy *et al.*, 2015). Dystrophin connects the intracellular cytoskeleton to the extracellular matrix via the DAGC. This complex plays thus a

crucial role in the stability of the muscle membrane and aids shock absorption during muscle contraction (Guiraud *et al.*, 2015). In addition to its structural function dystrophin is also believed to play a part in cellular signaling and gene expression (Cacchiarelli *et al.*, 2010; Constantin, 2014; Carr *et al.*, 2017).

1.1.4.2. Muscle and Disease Pathophysiology

It is known that DMD is caused due to a lack of dystrophin. However, the mechanisms that lead from the absence of the protein to muscle degeneration are yet to be fully clarified. Various histopathologic and biochemical changes occur in DMD muscle due to absence of dystrophin (Flanigan, 2014). Histological studies on muscle samples of DMD patients and animal models have revealed an ongoing process of repetitive degeneration and regeneration of myofibres (focal areas of regenerating and degenerating fibres and increased number of internal nuclei present) with progressive loss of regenerative capacity that leads to an increasing amount of necrotic muscle fibres infiltrated by monocytes and macrophages (Ge *et al.*, 2004). Necrotic fibres are progressively replaced by fat and connective tissue (increased amount of collagen and adipose present) which leads to the apparent fibrosis and chronic damage (Tanabe, Esaki and Nomura, 1986; Marshall, Williams and Goldspink, 1989; Bonilla, Tanji and Minetti, 1999; Ge *et al.*, 2004; Smith and Barton, 2014; Carr *et al.*, 2017; Ruiten, Bushby and Guglieri, 2017). Biochemical changes due to muscle wasting include reduction in muscle myosin, carnitine, and most glycolytic enzymes whereas alterations due to inflammation and invasion of fibroblasts involve an increase in enzymes present in fibroblasts and macrophages (such as NADP-linked dehydrogenases), proteases (cathepsins, lysosomal acid hydrolases, and calcium-activated proteases) and immune-related proteins (Emery, Muntoni and Quinlivan, 2015). Moreover, increased calcium levels and oxidative damage in terms of higher protein carbonylation have also been observed in the skeletal muscle of DMD patients and *mdx* mice (Fong *et al.*, 1990; Dunn and Radda, 1991; Kaczor *et al.*, 2007).

As already mentioned, the main function of dystrophin is to connect the intracellular cytoskeleton and contractile apparatus with the extracellular matrix ensuring thus the lateral transmission of mechanical stress during contraction and membrane stability (Emery, Muntoni and Quinlivan, 2015). An absence of this structural linkage can lead to membrane fragility and increased membrane permeability due to microlesions or other types of membrane damage induced by mechanical stress during muscle contractions (Consolino and Brooks, 2004; Deconinck and Dan, 2007; Guiraud *et al.*, 2015). Membrane damage and increased membrane permeability could lead to increased levels of intra-cellular calcium (and other micromolecules) and subsequently to cell dysfunction and death via activation of proteases (Marshall, Williams and Goldspink, 1989; Deconinck and Dan, 2007; Guiraud *et al.*, 2015). Continued cell death results in the apparent imbalance between degeneration and regeneration of muscle which then causes an inflammatory response and the observed muscle fibrosis with infiltration of adipose and connective tissue (Marshall, Williams and Goldspink, 1989; Deconinck and Dan, 2007; Guiraud *et al.*, 2015). This mechanism explains the main histological observations.

It has also been suggested that increased influx of calcium ions could be sequestered in mitochondria and could thus lead to mitochondrial overload and reduction in oxidative phosphorylation (Emery, Muntoni and Quinlivan, 2015). In support of these mechanisms,

proteotomic studies on DMD skeletal muscle of *mdx* mice have revealed changes in expression of proteins involved in metabolism and energy production e.g. a decrease in adenylate kinase 1 (AK 1), calcium homeostasis (for example upregulation of PP1) and cytoskeletal reorganization and maintenance (for example increased expression of RhoGDI-1, g-actin, and tropomyosin 1) (Ge *et al.*, 2004). Moreover, an increase of stress related chaperone proteins that play an important role in calcium storage and changes in the expression of mitochondrial proteins have also been observed in *mdx* mice (Carr *et al.*, 2017). Such changes could also be indicative of impaired calcium homeostasis and oxidative phosphorylation (Carr *et al.*, 2017). Comprehensive understanding of the above-mentioned mechanisms and other observed pathophysiological and biochemical changes is essential for the development of an effective treatment.

1.1.4.2 mdx Mouse Model

The *mdx* mouse model is the most used and best characterised DMD animal model (Granchelli *et al.*, 2013). The nonsense point mutation in exon 23 of the dystrophin gene that results in lack of dystrophin in this mouse model arose spontaneously in the early 1980s in a colony of C57BL/10ScSn mice (Bulfield *et al.*, 1984a). Despite the absence of dystrophin, *mdx* mice develop a milder phenotype of DMD than humans (Nowak and Davies, 2004). This is often attributed to the fact that mouse muscle exhibits a better and more robust regenerative capacity (Nowak and Davies, 2004). Histological studies of *mdx* skeletal muscle show ongoing degeneration and regeneration of muscle fibres, starting at about three weeks and persisting throughout the whole life of the mouse with a peak at around 12 months (Manning and O'Malley, 2015). Moreover, there is evidence of inflammation and 'pseudohypertrophy' (Grounds *et al.*, 2008). However, the lifespan of *mdx* is decreased by approximately 25% (whereas the decrease for DMD patients is approximately 75%) and fibrosis and loss of limb muscle function occur at a much more moderate degree (Coulton *et al.*, 1988; Manning and O'Malley, 2015; McGreevy *et al.*, 2015). Unlike skeletal and cardiac muscles which are much more mildly affected in mice, the degeneration of the diaphragm is similar to the pathology observed in humans (Stedman *et al.*, 1991; Guiraud *et al.*, 2015). Despite these disadvantages, this model has proven very useful in understanding the molecular changes and pathophysiological mechanisms that occur due to dystrophin deficiency, as well as in testing new therapeutic approaches (Manning and O'Malley, 2015).

1.1.4.2.1 Established *in vivo* Biomarkers

The main *in vivo* techniques used to study disease onset and progression and to evaluate potential treatments or the results of interventions for the mouse models of DMD are tests that probe muscle strength, function and coordination such as whole body or forelimb grip strength, two or four limb wire hanging and rotarod running (Grounds *et al.*, 2008). The main advantages of such tests are that they are non-invasive, inexpensive and usually not particularly time-consuming. Protocols and test regimes that combine these tests for the *mdx* mouse model have been established and have been shown not to affect the disease, making them suitable tests to study disease progression or the effect of potential therapeutic agents and rendering results from studies undertaken in different laboratories more comparable (Siegel *et al.*, 2009; Carlson *et al.*, 2010; Van Putten *et al.*, 2010; Sali *et al.*, 2012; Aartsma-Rus and Van Putten, 2014; Mcdonald *et al.*, 2015; Carlson, 2019; Putten, 2019). The main disadvantages of these techniques are that they

are confounded by animal weight, balance and behaviour. Also, due to the variability in the performance of the mice sometimes very large numbers of mice are required for reliable measurements (Grounds *et al.*, 2008; Carlson *et al.*, 2010; Carlson, 2019; Putten, 2019). Room conditions, such as room temperature, odours and noise, as well as time of the day can also affect the performance of the mice (Carlson, 2019; Putten, 2019). For the whole body tension and wire hanging tests it has also been shown that using more than one investigators for the tests introduces variability; thus, each longitudinal series of tests is best performed by a single investigator, or a high degree of standardization is needed to reduce inter-trial and inter-examiner variability (Carlson, 2010; Carlson *et al.*, 2010; Putten, 2019).

Numerous *mdx* studies have been published using these techniques. Various whole body tension parameters have been measured from one month to two years of age and shown to differ significantly between *mdx* and C57Bl/10 mice throughout the lifespan of *mdx* mice (Siegel *et al.*, 2009). Forelimb and hindlimb grip strength have also been shown to significantly vary between *mdx* and wild-type mice from 10 to 40 weeks but other outcome measures such as normalised duration of maintained grip has been reported to not vary significantly between the two groups (Spurney *et al.*, 2009; Mcdonald *et al.*, 2015). It has also been shown that between four and 13 weeks of age *mdx* mice perform worse in the two limb hanging test than age-matched control mice, with the mean longest time hanging ranging between 100 and 300 seconds for the *mdx* mice at the different time points and between 300 and 500 seconds for the wild-type mice (Aartsma-Rus and Van Putten, 2014). Other outcome measures were shown not to differ between the two groups at specific timepoints (Aartsma-Rus and Van Putten, 2014). The four limb wire hanging performance of *mdx* and C57Bl/10 mice differed less and rotarod running times did not differ for the two groups at any time point (Aartsma-Rus and Van Putten, 2014). Other studies have also reported that there was no difference in the rotarod performance between *mdx* and control mice at various timepoints, with some studies reporting significant differences in latency to fall between three and five months of age only (Spurney *et al.*, 2009; Mcdonald *et al.*, 2015). All the above show that tests that probe function and coordination, although useful, are not always able to identify the mild phenotype of the *mdx* mouse model and monitor progression consistently and are therefore typically used in combination with blood marker measurements and post-mortem tests like muscle force studies and histological analysis. The need for these additional tests renders longitudinal measurements very difficult (for example CK measurements) or impossible (post-mortem tests).

MRI has also been used to noninvasively detect disease activity at specific time points and evaluate the outcomes of therapeutic or other interventions in the *mdx* mouse model (Dunn and Zaim-Wadghiri, 1999; Amthor *et al.*, 2004; Walter *et al.*, 2005; Pratt *et al.*, 2013). Moreover, magnetic resonance imaging and spectroscopy have been employed to successfully monitor disease progression longitudinally (Pratt *et al.*, 2013; Heier *et al.*, 2014). In both studies foci of hyperintensity, present in the *mdx* tissue, were used as a measure of tissue heterogeneity, in order to quantitatively monitor changes in *mdx* muscle pathology. Pratt et al. demonstrated that heterogeneity was present in *mdx* muscle from 5 to 80 weeks of age, peaked between 9 and 13 and then dropped in a stable manner until 80 weeks of age (Pratt *et al.*, 2013). This study only used one *mdx* mouse and no control mice for comparison. Heier et al. also showed that the percentage of hyper-intense foci within the hind limb muscle tissue differed significantly between

mdx and wild-type mice at 6, 8, 10 and 12 weeks of age (Heier *et al.*, 2014). Spectroscopic measurements also revealed altered bioenergetics in *mdx* muscle compared to wild-type mice mainly at 6 weeks of age (Heier *et al.*, 2014). None of these studies compare the data acquired from *mdx* mice at different time-points in order to better evaluate if the technique is sensitive enough to detect changes compared to previous disease stages and not between healthy and diseased muscle. Nonetheless, they do show the potential of MRI to monitor onset and disease progression of DMD at the preclinical setting. The main limitations of MRI at the preclinical setting include that it is a very expensive method and the long scanning often required pose difficulties in maintaining body temperature and recovering the animal.

1.2 Raman Spectroscopy

Raman effect was discovered in 1928 by C.V. Raman, who used sunlight for excitation and a telescope to collect the Raman scattered light (Raman, 1928; Ferraro *et al.*, 2002). In 1934 Cohan et al. using a mercury lamp as the excitation source and an exposure time of 24 hours managed to study the Raman spectra of amino acids, exploring thus for the first time the biological applicability of Raman spectroscopy (Cohn *et al.*, 1934). In 1962, laser light sources for use with Raman spectroscopy were developed and facilitated Raman experimentation (Gilson *et al.*, 1970). 8 years later Lord et al. studied various proteins and their constituent amino acids in aqueous solutions using a laser-excited Raman system (Lord and Yu, 1970). Further technological advancements have led to Raman systems able to measure the chemical composition of various complex biological samples at a molecular level (Jermyn *et al.*, 2016). The Raman spectrum of the sample, known as its molecular fingerprint, offers essential qualitative and quantitative biochemical information (Noothalapati *et al.*, 2017). Molecular changes, that occur due the progression of a disease, or could cause a disease can be optically probed using Raman spectroscopy (Das *et al.*, 2006; Stone and Matousek, 2008; Rehman, et al., 2012). Thus, Raman spectroscopy can be used for disease prognosis and diagnosis.

In this field significant advances have been made particularly regarding cancerous pathology, with various applications demonstrating the ability of Raman spectroscopy to grade, diagnose and in some cases even prognosticate on various types of cancer using cells, tissue or biofluids (Butler *et al.*, 2016). The development of fibre-optic probes led to an increase in the *in vivo* diagnostic applications of the technique with examples in lung (McGregor *et al.*, 2016), cervical (Duraipandian *et al.*, 2012), and oesophageal (Wang *et al.*, 2015) cancer detection. Additionally to the wide range of cancer applications, the technique has been used to shed light on bone disease and, using biofluids like blood plasma and urine, on diabetes (Shafer-Peltier *et al.*, 2003), asthma and malaria (Hobro *et al.*, 2013). Moreover, recently application of Raman spectroscopy in monitoring and diagnosing of neurodegenerative diseases (Tian *et al.*, 2016) and myopathies (Gautam *et al.*, 2015) has started to gain attention.

Promising *in vivo* results (>80 % sensitivity and specificity, often with very short acquisition times) in various cancer studies across several different labs along with the fact that the technique is non-invasive and can be used for measurements of chemical composition (with high molecular specificity) and imaging of biological samples at a very low cost indicate the ability of Raman spectroscopy to be used in the clinical environment and clinical implementation is firmly on the horizon (Byrne *et al.*, 2015; Kong *et al.*, 2015; Jermyn *et al.*, 2016). One of the limitations regarding the translation of Raman spectroscopy into the clinic arises from the inherently weak Raman signal. This induces a limit in the speed of Raman systems, since short acquisition times can lead to the detection of weak signals with poor signal to noise ratios (Kong *et al.*, 2015). However, the development of new techniques that enhance the Raman signal along with advancements in Raman instrumentation constitute essential steps toward maximizing the speed of the technique. Additional challenges include standardising the data analysis methods and demonstrating the safety of Raman devices (Pence and Mahadevan-Jansen, 2016). Thus, studies that evaluate biological damage caused by Raman systems have to be conducted before the adoption of Raman technology in healthcare. Finally, the developed technology needs to address unmet needs and

provide new information regarding a clinical target. Thus, input from clinicians and constant multi-disciplinary collaborations are essential for delivery of the next steps.
In the subsequent sections the theory of Raman effect and Raman technology will be presented. Furthermore, the limited applications of Raman spectroscopy in neuromuscular disorders will be discussed.

1.2.1 Theory

1.2.1.1 Quantum Theory of Raman Effect

In quantum theory light consists of massless particles called photons. Upon interaction with matter photons can be transmitted through the material, absorbed or scattered by the molecules of the material.

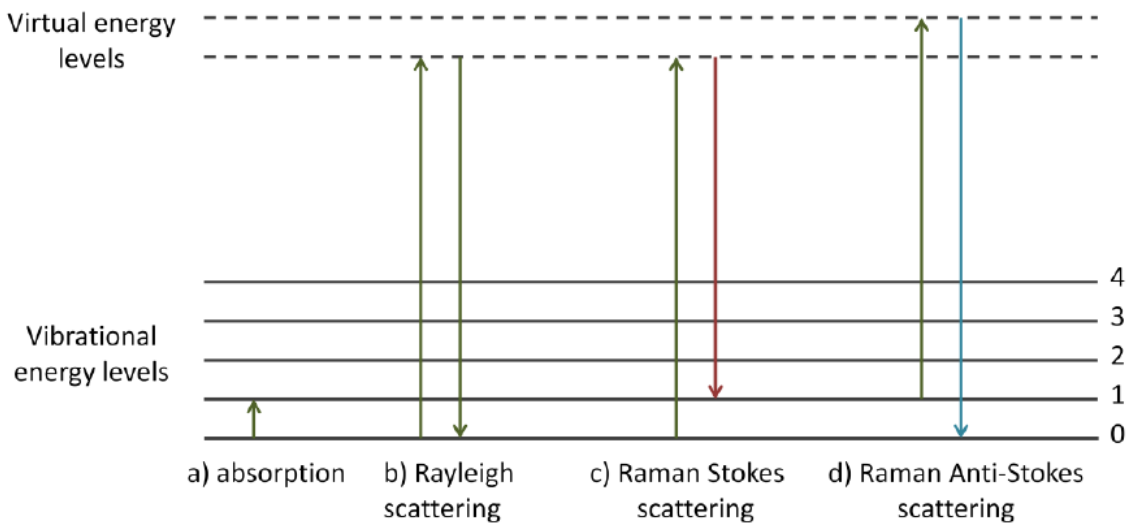


Figure 1.1: Absorption and different types of scattering.
The arrows depict the change in the molecular state due to photon absorption (upward arrows) and emission (downward arrows). The colour of the arrow describes the energy (i.e. colour) of the light.

Absorption occurs when the energy of the photons matches the energy gap between two molecular energy states. Once the molecule absorbs the photon, it gets promoted from the ground state to a higher (excited) energy level (Figure 1.1a). For scattering to occur, photon energy does not need to correspond to the energy gap between the energy levels of the molecule. Moreover, scattering is a two-photon process, during which photons get absorbed and re-emitted instantaneously.
In scattering there is energy transfer between the molecule and the incident photons (Straughan and Walker, 1976). This leads to the creation of an extremely short-lived complex, formed between the photons and the electrons in the electron cloud of the molecule, known as a molecular virtual state. However, this new molecular state is unstable, and light is emitted instantaneously as scattered radiation. Due to the fact that virtual states arise from the interaction of light with electrons their energy is dependent on the energy of the incident radiation.
In Rayleigh scattering the scattered photons have the same energy (i.e. the same frequency) with the incident photons, thus, the molecule, after being promoted to the virtual energy level, relaxes

back to its initial energy state (Fig 1.1b) (Smith and Dent, 2005). This occurs when only electron cloud distortion takes place during scattering and it is the dominant scattering process (Straughan and Walker, 1976).

On the other hand, in Raman scattering the interaction of photons with electrons induces nuclear motion (Straughan and Walker, 1976). This movement causes a noticeable change in the energy of the molecule and thus in the energy of the scattered photons. This leads to the molecule relaxing to a different vibrational state and to the scattered photon having different energy from the incident one (Rehman *et al.*, 2012). The change in the energy of the incident and scattered photon corresponds to the energy of the molecular vibration and is known as Raman shift. If a molecule, due to Raman scattering, undergoes a vibrational transition from the ground state to the first excited level, the scattered photon has less energy than the incident one (Fig.1.1c) (Smith and Dent, 2005). On the contrary, if the molecule undergoes a transition from a vibrationally excited state to the ground state, the scattered photon has more energy from the incident one (Fig. 1.1d). When the photon energy decreases (energy is offered to the molecular vibration), Raman scattering is known as Stokes scattering (Diem *et al.* 2008; Rehman *et al.*, 2012). When the photon gains energy it is known as Anti-Stokes scattering. Raman scattering is a much weaker process than Rayleigh scattering. Only one in every $10^6 – 10^{10}$ photons are scattered inelastically (Conroy *et al.*, 2005; Wachsmann-Hogiu *et al.*, 2009; Kong *et al.*, 2015).

1.2.1.2 Molecular Vibrations and Raman Spectroscopy

Molecules are made up of atoms, which are held together by chemical bonds. The energy of a molecule can be divided into translational energy, rotational energy and vibrational energy (Cothlup and Daly, 1975). If N is the number of atoms that make up a molecule there will be 3N degrees of freedom of motion for all the atoms in the molecule (each atom has 3 independent degrees of freedom of motion, in the x, y or z direction). Three of these degrees describe the translational movement of the molecule in space and three describe the rotational movement for non-linear molecules (two for linear molecules). Thus, there are 3N-6 vibrational degrees of freedom and therefore 3N-6 different vibrational modes for non-linear molecules (3N-5 for linear molecules) (Cothlup and Daly, 1975; Straughan and Walker, 1976, Smith and Dent, 2005).

During a molecular vibration the shape of the surrounding electron changes due to the change in the configuration of the nucleus. This can lead to a change of the molecular dipole moment or a change in the polarizability of the molecule. In order for the vibration of a molecule to be detected by Raman spectroscopy (Raman active), the polarizability of the molecule needs to be altered. Infrared spectroscopy detects the vibrational modes of the molecules that represent a change in the dipole moment.

Molecular vibrations take place when the atoms move periodically. This movement can induce a change in either the distance between the atoms (i.e. the bond length) or the angle between two bonds (Rehman *et al*. 2012). Depending on the change there are two different types of vibrations, stretching and bending (Kumar, 2013). Symmetrical and anti-symmetrical stretching as well as scissoring, rocking, wagging and twisting (bending modes) can be seen in figure 1.2. Vibrational transitions are the origin of Raman spectra. An example Raman spectrum can be seen in Figure 1.3.

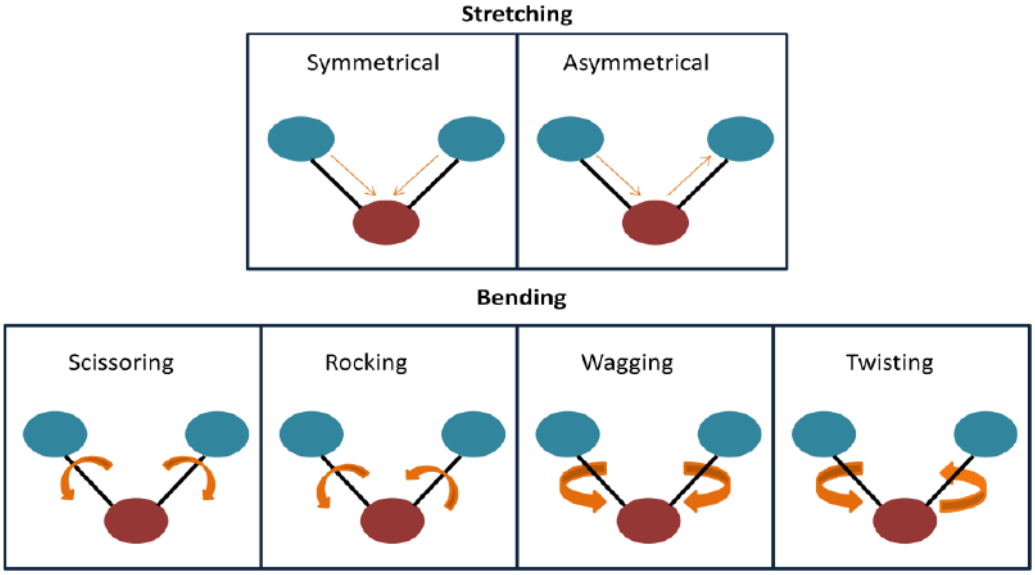


Figure 1.2: Vibrational modes.

The arrows depict the direction of blue atoms movement. Straight arrows are used when the movement induces change in the length of the bond (stretching) and curved arrows when the movement changes the angle of the bonds (bending).

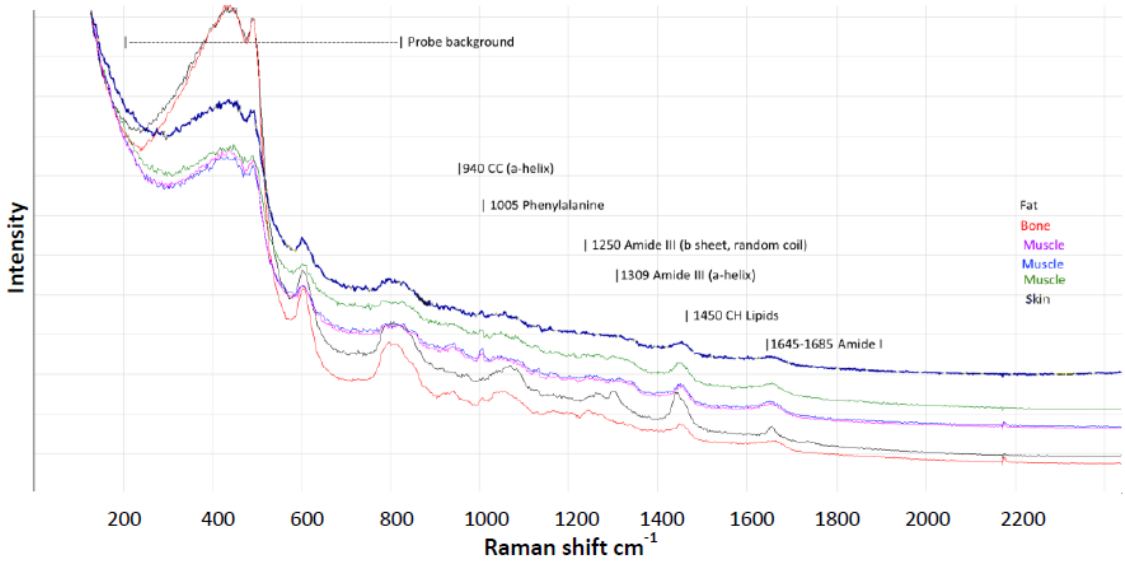


Figure 1.3: Raman spectra of different tissues.

The Raman peaks present in the spectra are assigned to specific biomolecules. The spectra were acquired using an 830nm laser (Laser power=60mW, Acquisition time=40s).

For a diatomic linear molecule (N=2), there is only one possible vibration. For an oxygen molecule this is a simple stretch of the O-O bond (Smith and Dent, 2005). This vibrational transition (from the ground to the first excited vibrational state) will appear as a band in the Raman spectrum of the molecule. In order to assign peaks of a Raman spectrum to molecular vibrations it is essential to understand that bonds between atoms in a molecule which are in close proximity and of similar energy interact with each other. The resulting Raman peak will correspond to the vibration of all

the atoms linked by these similar bonds. However, bonds between atoms that are well separated, or have a large difference in energy vibrations, can be treated separately (and will give different peaks in the Raman spectrum).

In Raman Spectroscopy the sample is illuminated with laser light, which is monochromatic radiation usually in the visible-infrared region (Ferraro *et al.*, 2002). The observed Rayleigh scattered light has the same frequency with the incident radiation, f_o. The Raman scattered light is weaker and has frequencies $= f_o \pm f_m$, where f_m is a vibrational energy of the molecule (Ferraro *et al.*, 2002). The $f_o + f_m$ are the anti- Stokes lines whereas the $f_o - f_m$ are the Stokes lines of the spectrum. Often in spectroscopy wavenumber is used instead of the frequency of light. The wavenumber is given by the following expression:

$$k = \frac{f}{c} = \frac{1}{\lambda}$$

where λ is the wavelength of light (in cm), c is the speed of light and f is the frequency of light (in 1/s) (Ferraro *et al.*, 2002).

1.2.2 Technology

1.2.2.1 Raman Micro-spectroscopy

Raman spectroscopy is most commonly performed using visible or near infrared (NIR) lasers. As a result, coupling of Raman spectrometers with bright-field or confocal microscopes is highly efficient (Delhaye and Dhamelincourt, 1975). Upon coupling with a microscope the spatial resolution of the recordings depends on the properties of the microscope system (Sato *et al.*, 2014). Thus, lateral spatial resolution is diffraction limited and is determined by the wavelength of the excitation laser light and the numerical aperture (NA) of the objective lens (Butler *et al.*, 2016). For wavelengths in the visible and NIR region and NA values ranging from 0.5 to 1.2, resolution values of about 0.5µm can be achieved. A conventional Raman micro-spectroscopy system can be seen in figure 1.4.

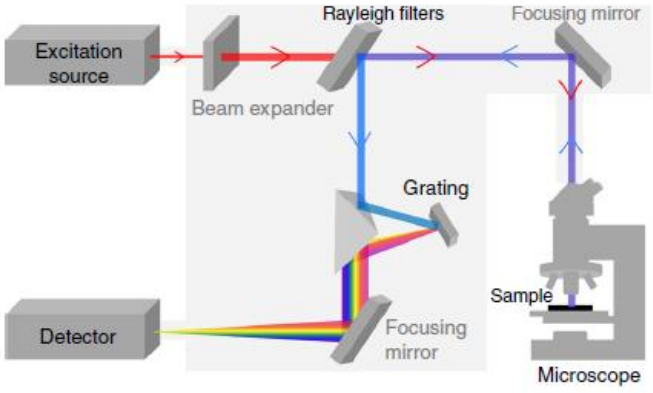


Figure 1.4: Generalised conventional Raman micro-spectroscopy system (adopted from Butler *et al.*, 2016).

The key components that comprise this system are a microscope, a laser source, a Rayleigh filter (usually notch filter), a monochromator and a detector. As already mentioned above Raman signal

is inherently weak. Therefore, it can easily be masked by autofluorescence generated in the biological samples. However, in the wavelength region from 700 to 900 nm (i.e. the window between melanin and water absorption) most tissues and body fluids exhibit minimum light absorption and thus minimum autofluorescence (Krafft *et al.*, 2009; Krafft, Dietzek and Popp, 2009). Therefore, systems designed for biomedical applications largely use excitation lasers in the NIR region; most frequently diode lasers at 785 or 830nm are being used. (Krafft *et al.*, 2009; Krafft, Dietzek and Popp, 2009). Notch filters are used to separate Rayleigh from Raman scattered light by blocking the elastically scattered photons. Subsequently, Raman scattered light is dispersed (by the monochromator) before reaching the detector, in order to separate and identify all the different Raman shifts in it. Usually cooled charge-coupled device (CCD) detectors are used to collect the dispersed Raman light.

The Raman spectrometer can also be coupled with a confocal microscope. In this case a pinhole is placed in front of the detector. Raman scattered light is passed through the pinhole before being detected (Diem, 2015). This arrangement prevents scattered (or fluorescent) light emerging from out-of-focus planes from passing efficiently through the pinhole, and thus being detected.

1.2.2.2 Fibre-optic Probes

Raman spectroscopy can be performed using optical fibres to deliver the excitation light and collect the Raman scattered light. The main advantage of these compact and flexible probes is that they render *in vivo* measurements much easier since they can be inserted into the human body. However, significant levels of photoluminescence and Raman scattered light are generated in both the excitation and collection silica optical fibres. This can easily mask the weak Raman light scattered by the sample (Kong *et al.*, 2015). Furthermore, Rayleigh scattered light must be prevented from entering the collection fibre (and thus again masking the Raman signal but also generating more photoluminescense in the collection fibre) (Motz *et al.*, 2005). Filters and lenses must therefore be implemented in a very small probe. In addition, in order to do real-time *in vivo* measurements enough Raman scattered light must be collected (i.e. efficient throughput collection) in a short time.

Various hand-held fibre optic probes have been previously reported. Hand-held probes of different configurations have used for various applications. These include:

- a probe consisting of a 1.27cm diameter illumination arm (with single excitation optical fibre) and a 2.53cm diameter collection arm (with a bundle of 58 collection fibres) for skin measurements (Huang *et al.*, 2001; Lui *et al.*, 2012)
- a pen-like probe made up of 19 illumination fibres interspersed with 38 collection fibres in a 3 mm ring, and 12 collection fibres placed in the center of the ring for arthroscopy of joint tissues (Esmonde-White *et al.*, 2011)
- a hand-held microprobe consisting of 7 collection fibres placed around a single excitation fibre for intraoperative detection of brain cancer (Desroches *et al.*, 2015)

Miniaturisation of fibre-optic probes allows coupling of the probes with endoscopes for real time *in vivo* measurements in hollow organs (Stevens *et al.*, 2016). Once again probes of various sizes and configurations have been reported. Huang et al. developed a 1.8mm diameter Raman probe comprising 32 collection fibres placed around a single excitation fibre that was coupled to a trimodal imaging endoscopic system for *in vivo* detection of gastric cancer (Huang, Teh, *et al.*, 2010). A 600µm

diameter fibre-optic probe consisting of 8 collection fibres surrounding a single excitation fibre that could be coupled to an intravascular endoscope was also reported (Komachi *et al.*, 2005).
Further miniaturization of the Raman fibre-optic allows insertion of the probes in needles (Kong *et al.*, 2015; Stevens *et al.*, 2016). Day et al. developed a needle probe made up of an excitation and a single collection fibre. The Raman probe was integrated into a hypodermic syringe with a 20ga needle. This configuration allows protection of the fibres but also access to solid organs such as breast, lymph nodes, or as tested in this project muscles (Day and Stone, 2013).
As Raman signal is very weak, recording spectra in short acquisition times (1 second or less) can be quite challenging with miniaturized probes. Delivery of excitation light and collection of Raman scattered light using the same optical fibre would maximize the collection efficiency of small probes, since the light cones of the illumination and collection beams would overlap (Kong Chong *et al.*, 1992). However, this arrangement suffers from a lot of background noise due to the Raman scattered light generated in the core of the optical fibre (Day and Stone, 2013). In order to avoid this problem different fibres are largely used for delivery and collection of light. Multiple collection fibres increase the collection efficiency but they also increase the size of the probe. Additionally, since the excitation and collection cones still don't overlap essential amount of scattered light is lost. By using beveled tip collection fibres, the collection light cones can be angled to increase the overlap with the excitation cone (Shim *et al.*, 1999). Another way to enhance the collection efficiency but also the excitation (by providing a collimated illumination spot), is to incorporate a ball lens at the tip of the probe (Motz *et al.*, 2005; Mo *et al.*, 2009).
As previously described optical fibre based probes suffer from background signal due to photoluminescence and Raman scattering that take place in the core of glass fibres. Filters must therefore be incorporated to block the unwanted light. Band pass filters are implemented in the excitation fibres. Band pass filters allow only wavelengths in a certain range to pass. Therefore, they can block all the Raman and fluorescent light generated in the excitation fibre and allow only laser light to pass. In addition, long pass or notch filters are implemented in the collection fibres for the same reason but also to prevent Rayleigh scattered light from entering the spectrometer. The filters can be implemented at the tip of the probe (Hattori *et al.*, 2007), at both ends (Huang, Teh, *et al.*, 2010), or at some point in the length of the probe (Day and Stone, 2013). An alternative way to avoid the generation of background signal in the optical fibres is to develop probes that work in the high wavenumber region (2400–3800 cm−1). In this spectral region there is still essential biological information and Raman scattering does not take place in the core of the optical fibres (Koljenovic *et al.*, 2007)). Several probes that work in the high wavenumber region have been reported (Koljenović *et al.*, 2005; Mo *et al.*, 2009; Lin *et al.*, 2016).

1.2.3 Raman Spectroscopy and Neuromuscular Disorders

The potential for Raman spectroscopy to aid the diagnosis and monitoring of neuromuscular disorders has recently started to gain attention. A microscopic Raman system was used to differentiate between different myopathies in *ex vivo* Drosophilla preparations (Drosophilla mounted on a slide) (Gautam *et al.*, 2015). Mutants related with nemaline-myopathies were distinguished from flies carrying mutations associated with cardiomyopathy. Disease progression was also monitored with spectra acquired from 2 days old and 12 days old flies. An overall increase in the levels of nucleic acids and β-sheet proteins, with a decrease in α-helix proteins

were observed as both types of myopathy progressed in all different types of mutants. Additionally, the same Raman system and classification algorithm were used to record spectra from a human sample with nemaline-myopathy and in the subsequent analysis the spectrum was grouped with the nemaline-myopathy fly spectra.

Recently, Tien et al. used SRS microscopy to image degeneration of peripheral nerves in SOD1-G93A mice (Tian *et al.*, 2016). SRS imaging exhibited comparable sensitivity to electromyography (EMG), which is the standard examination for MND, but could additionally offer information about the structure of peripheral nerve. This allowed the detection of lipid ovoids, probably resulting from degenerating myelinated cells, before EMG could detect evidence of neuronal degeneration, suggesting thus that Raman spectroscopy can be used for early disease detection. Moreover, in this study the potential of Raman spectroscopy for evaluation of new therapeutic drugs was explored. Reduction in the speed of peripheral nerve degeneration due to administration of minocycline could be detected in SRS images.

Spontaneous Raman micro-spectroscopy was used to study the spinal cord of $SOD1^{G93A}$ and control mice *ex vivo* (Picardi *et al.*, 2018). Spectra were acquired from grey and white matter sections from a small number of mice at 75 and 90 days of age (Picardi *et al.*, 2018). Several peaks and regions (450-650 cm^{-1}, 1050-1200 cm^{-1}, 1300 cm^{-1}, 1440 cm^{-1} and 1660 cm^{-1}) were shown to differ markedly between diseased and control mice from 75 days of age in the spectra acquired from the grey matter regions (Picardi *et al.*, 2018). The observed changes were tentatively associated with processes that signal the progression of ALS, like axonal demyelination and the structural degradation of associated lipids (Picardi *et al.*, 2018). The spectra acquired from white matter areas of diseased and control mice did not exhibit considerable differences (Picardi *et al.*, 2018).

A microscopic Raman system was used by Morasso et al to study plasma-derived small and large extracellular vesicles and blood plasma extracted from blood samples acquired from 20 sporadic ALS patients and 20 age- and sex-matched controls (Morasso *et al.*, 2020). The obtained results demonstrated that large extracellular vesicles derived from ALS patient plasma display an altered biochemical profile compared to those derived from healthy controls (Morasso *et al.*, 2020). Changes in the relative spectra were tentatively associated with alterations in lipids and phenylalanine metabolism and following PCA analysis an AUROC value of 0.84 was achieved (Morasso *et al.*, 2020). Finally, plasma and sEVs were shown to be rather homogeneous between patients and controls (Morasso *et al.*, 2020). Plasma surface-enhanced spectroscopy has also been recently used to study differences in the blood plasma of short- and long-duration ALS patients (Zhang *et al.*, 2019). SERS was shown to be able to distinguish with high sensitivity and specificity values between the two groups (AUROC=0.97) due to changes attributed to various metabolic pathways (Zhang *et al.*, 2019).

Recently, saliva acquired from ALS patients was also studied using Raman spectroscopy by Carolmagno et al. In this study saliva was acquired from 19 ALS patients, 10 patients with Parkinson's disease, 10 patients affected with Alzheimer's disease and 10 healthy controls (Carlomagno *et al.*, 2020). Multivariate analysis was used to study the obtained spectra and demonstrated a significant difference between the different groups, allowing precise discrimination of the spectra acquired from ALS patients' saliva (Carlomagno *et al.*, 2020). Spectral differences related to nucleic acids, glycogen and glucose and lipids were observed between the

'ALS' and 'healthy' groups (Carlomagno *et al.*, 2020). Finally, Raman data (more specifically scores of the principal components) were shown to correlate with clinical symptom scores (Carlomagno *et al.*, 2020).

While some of these studies employed technically challenging methods, rely on *ex vivo* analysis, or are based on a limited number of samples, they demonstrate the potential of the technique for the study of neuromuscular disorders. Further studies are required to investigate the potential of *in vivo* Raman systems and classification models to distinguish between different types of neuromuscular disorders and identify biomolecules that could be used for differentiation. Since, for these diseases, early detection and monitoring are important it will be useful to further explore the potential of Raman spectroscopy to facilitate earlier diagnoses (perhaps even at pre-symptomatic stages) and monitoring of the disease progression.

1.3 Aims and Hypotheses

1.3.1 Aims

The aim of this work is to explore the potential of Raman spectroscopy as a biomarker for neuromuscular disease. This will be done using *in vivo* measurements in preclinical murine models of MND (SOD1^{G93A} and TDP-43^{Q331K} mice) and DMD (*mdx* mice) and *ex vivo* human muscle samples.

1.3.2 Hypotheses

i)*In vivo* intra-muscular Raman spectroscopy can distinguish between neurogenic and myopathic disorders.

ii) *In vivo* intra-muscular Raman spectroscopy can detect changes over time in murine models of neuromuscular disease.

iii)*In vivo* intra-muscular Raman spectroscopy will not cause significant muscle injury and impair motor performance in mice.

iv)Raman spectroscopy of *ex vivo* human muscle samples can detect muscle pathology and distinguish between neuromuscular diseases.

2. Methods

2.1 Animal Experiments

The ethics statement for the animal experiments is presented in section 2.1.1. The animals used in each of the *in vivo* studies along with the study protocols are summarised in section 2.1.2. Finally, the experimental procedures performed as part of the animal studies are described in sections 2.1.4-2.17.

2.1.1 Ethics Statement

All mouse experiments were carried out in accord with the Animals (Scientific Procedures) Act 1986 under UK Home Office project license number 70/8587. The project methodology was reviewed and approved by the Ethical Review Committee Project Applications and Amendments Sub-Committee of the Sheffield University Ethical Review Committee and by the Animal Procedures Committee (London, UK). Animals were housed and cared for according to the Home Office Code of Practice for the Housing and Care of Animals Used in Scientific Procedures.

2.1.2 Animals and Study Protocols

All animals used in this study with the exception of C57BL/10ScSnOlaHsd mice were from in-house colonies. The $SOD1^{G93A}$ mice were transgenic mice (the human gene has been transferred to wild-type C57Bl/6 mice) that express a G93A mutant form of the human SOD1 gene. Male $SOD1^{G93A}$ were bred with female wild-type C57Bl/6 mice in an in-house colony. Thus, the SOD^{G93A} colony consisted of both $SOD1^{G93A}$ (transgenic) and C57Bl/6 (non-transgenic/wild-type) mice. The non-transgenic C57Bl/6 littermates were used as control mice for the $SOD1^{G93A}$ mice.

The *mdx* mutation of the DMD gene is a spontaneously arisen mutation. Hence, the *mdx* mice are not transgenic mice. In the *mdx* in-house colony female homozygous and male hemizygous mice for the Dmd^{mdx} allele were bred. Thus, the *mdx* colony only consisted of mice with the *mdx* mutation in the DMD gene. Wild-type C57BL/10ScSnOlaHsd (C57Bl/10) mice were used as control mice for the *mdx* colony and were bought from Envigo.

The TDP-43 mice are transgenic mice. The in-house colony TDP-43 colony consisted of TDP-43 (mice with the human *TARDBP* gene but without any mutation in the gene) and $TDP\text{-}43^{Q331K}$ (mice with the human *TARDBP* gene and a Q331K mutation). The TDP-43 mice were used as control mice for the $TDP\text{-}43^{Q331K}$. The C57Bl/6, C57Bl/10 and TDP-43 mice will be referred to as control mice in the next chapters.

Study 1: Neurogenic/Myopathic Neuromuscular Disorders

The aim of this study was to evaluate the potential of *in vivo* Raman spectroscopy of muscle to detect muscle pathology and to explore the potential of the technique to distinguish between muscle pathology induced by a neurogenic and a myopathic neuromuscular disorder. The $SOD1^{G93A}$ mouse model of MND and the *mdx* mouse model of DMD along with their control mice were used for this study. Additionally, potential tissue damage caused by the *in vivo* recordings was also investigated.

A total of 227 one and three months old mice were used in this study. Age and strain/genotype clustered the mice in the different groups. In each group some of the mice underwent the active

Raman procedure and some underwent the 'sham' procedure (see section 2.1.3.2). The groups with the number of mice used per group are shown in table 2.1.

Table 2.1: Number of mice used in each group in Study 1. *Age and strain/genotype clustered the mice in each group. Procedure clustered the mice in one of two subgroups.*

Group	Procedure	Number of animals
One month old	Raman	16
$SOD1^{G93A}$	Sham	10
One month old	Raman	16
C57Bl/6	Sham	6
One month old *mdx*	Raman	16
	Sham	16
One month old	Raman	16
C57Bl/10	Sham	16
Three months old	Raman	16
$SOD1^{G93A}$	Sham	16
Three months old	Raman	16
C57Bl/6	Sham	16
Three months old	Raman	16
mdx	Sham	16
Three months old	Raman	16
C57Bl/10	Sham	3

In order to assess the potential tissue damage caused by the technique the rotarod test and gait analysis (Catwalk) (see section 2.1.4) were performed on all the mice one week before and one day after the experimental procedure (for both active and 'sham' protocols) and on half of the cohort two weeks post-experiment. The protocol for this study along with the total animal number used in each step can be seen in figure 2.1.

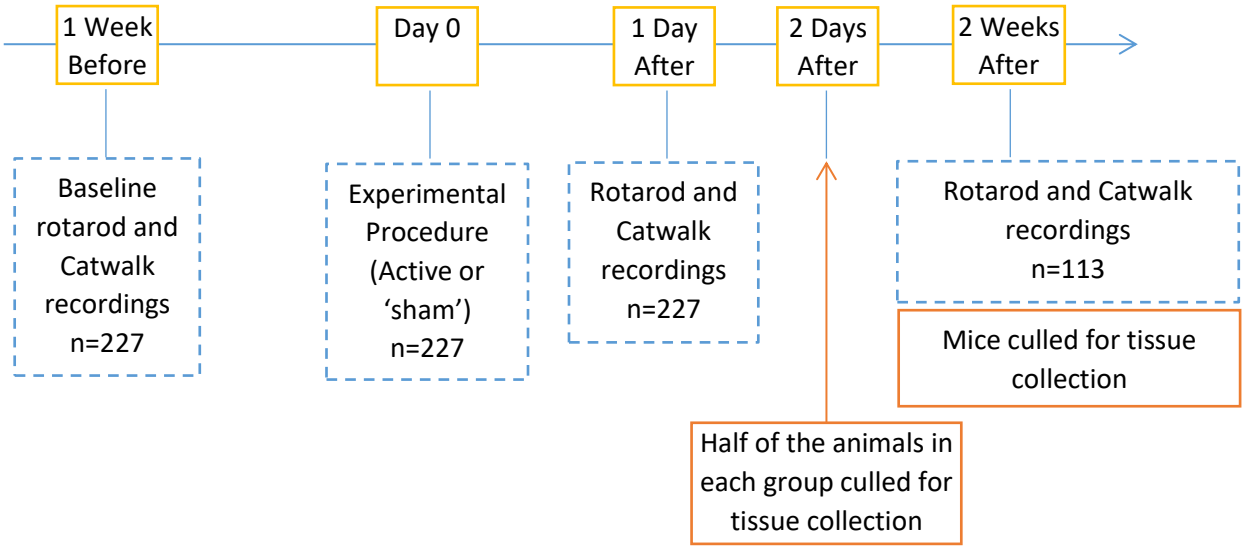


Figure 2.1: Study 1 Protocol.

Nine additional one month old C57Bl/6 mice that underwent the active experimental procedure were used for post-mortem MRI either six hours (n=3), two days (n=3) or two weeks (n=3) post-experiment (see section 2.1.5). Four, one month old C57Bl/6 mice that underwent the 'sham' experimental procedure were used for post-mortem MRI either six hours (n=2) or two days (n=2) post-injury.

Study 2: Intervention Study in *mdx*

The aim of this study was to explore the potential of the technique to detect more subtle changes in muscle pathology after an intervention. A total of 60 three months old female *mdx* and control mice were used for this study (table 2.2). Voluntary exercise using running wheels was used to exacerbate the muscle pathology of the *mdx* mice and the ability of *in vivo* Raman spectroscopy to detect the change in the muscle pathology was investigated. Both exercised and non-exercised *mdx* mice had the active Raman procedure (see section 2.1.3.2). Serum Creatine Kinase activity was measured as a broad indicator of muscle damage (see section 2.1.6.2). The protocol for this study along with the total animal number used in each step can be seen in figure 2.2.

Table 2.2: Number of mice used in each group in Study 2. *Voluntary exercise clustered the mice in each group. Procedure clustered the mice in subgroups.*

Group	Procedure	Number of animals
Exercised three months old *mdx* mice	Raman spectroscopy (Subgroup 1)	15
	Creatine Kinase Measurements (Subgroup 2)	10
Non-exercised three months old *mdx* mice	Raman spectroscopy (Subgroup 3)	15
	Creatine Kinase Measurements (Subgroup 4)	10
Non-exercised three months old C57Bl/10 mice	Creatine Kinase Measurements (Subgroup 5)	10

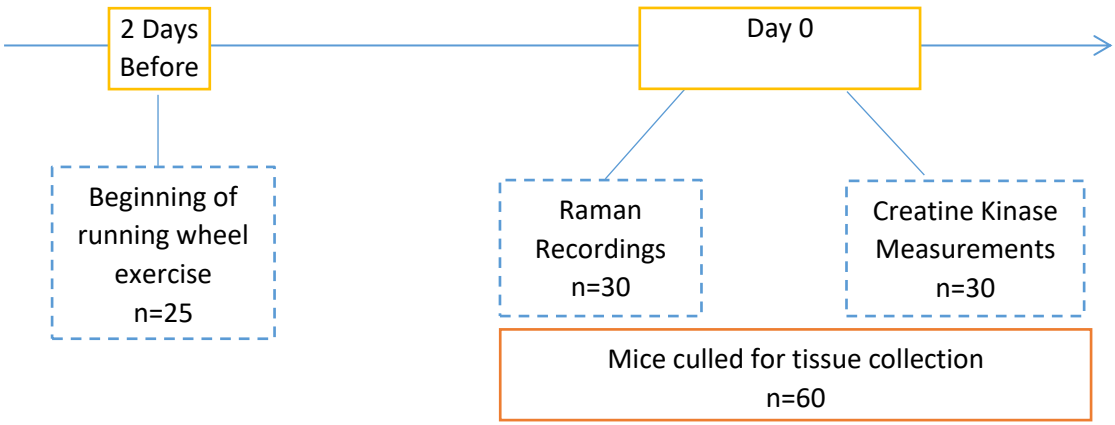


Figure 2.2: Study 2 Protocol.

Study 3: Preclinical MND Study

The aim of this study was to explore the potential of the technique to detect differences in muscle pathology between two mouse models of the same disorder. A total of 24 female three months old $SOD1^{G93A}$, TDP-43 and $TDP\text{-}43^{Q331K}$ mice was used for this study (table 2.3). All mice underwent the active Raman procedure (see section 2.1.3.2).

Table 2.3: Number of mice used in each group in Study 3.

Group	Number of animals
Three months old $SOD1^{G93A}$	8
One month old TDP-43	8
Three months old $TDP\text{-}43^{Q331K}$	8

Study 4: Male *mdx* Mice Study

This study aimed to investigate the performance of the models generated using Raman spectra acquired *in vivo* from male *mdx* mice and compare it with the performance of female *mdx* mice (used in Study 1). A total of 32 male *mdx* and control mice were used for this study (table 2.4). All mice underwent the active Raman procedure (see section 2.1.3.2).

Table 2.4: Number of mice used in each group in Study 4.

Group	Number of animals
One month old male *mdx*	8
One month old male C57Bl/10	8
Three months old male *mdx*	8
Three months old male C57Bl/10	7

Study 5: Combined Probe Study

The aim of this study was to test an in-house built combined probe, that can acquire both Raman and electrophysiological data. Moreover, this study aimed to compare the electrophysiological recordings acquired with the combined probe with standard EMG and NCS methods (ring electrodes, needle electrode). Raman spectra and compound muscle action potentials (CMAPs) and were recorded *in vivo* from 16 $SOD1^{G93A}$ and control mice (table 2.5).

Table 2.5: Number of mice used in each group in Study 5.

Group	Number of animals
Three months old $SOD1^{G93A}$ mice	8
Three months old C57Bl/6 mice	8

2.1.3 Raman Spectroscopy Experiments

2.1.3.1 Anaesthesia

Inhalational anesthesia was induced and maintained with Isoflurane. Mice were placed in the induction chamber of the anaesthetic machine until anaesthetised (5% isoflurane vapour, oxygen flow rate 4.0L/min). Consequently, animals were placed on a heating pad and anaesthesia maintained throughout the whole experimental procedure with inhalation of the anaesthetic agent through a nose cone (2% isoflurane vapour, oxygen flow 2.0L/min). Post-procedure they were placed in an incubator for approximately 15 minutes at 30˚C while recovering from anaesthesia. Once the mice were fully mobile and feeding, they were returned to their home cage.

2.1.3.2 Raman spectroscopy

A fibre-optic Raman needle probe system was used for the experimental procedure (Day and Stone, 2013). A schematic of the probe-spectrometer interface can be seen in figure 2.3. An 830 nm semiconductor laser (Innovative Photonics Solutions) is fiber coupled to an inline filter unit, which collimates the light through a laser line bandpass filter (Semrock, Inc.) to remove Raman and fluorescence generated in the fiber and refocuses the filtered light into a short length (15cm) of low-OH fiber with 105μm core and 0.22NA (Thor Labs, Inc.) to the distal end of the needle. An identical second fibre was used to collect scattered light at the sample and transmit it to the spectrometer via a similar inline filter unit containing a long pass filter that rejects the elastically scattered laser light. The excitation and collection fibres were sheathed in a stainless-steel tube with a 0.5 mm outer diameter which was then inserted into a 21G needle attached to a syringe. The spectrometer used was a Raman Explorer spectrograph (Headwall Photonics, Inc.), and iDus 420BR-DD CCD camera (Andor Technology, Ltd.) cooled to -70^{o}C. Acquisition time was 4 seconds for each spectrum and 10 spectra were acquired at each site (total acquisition time at each site 40 seconds).

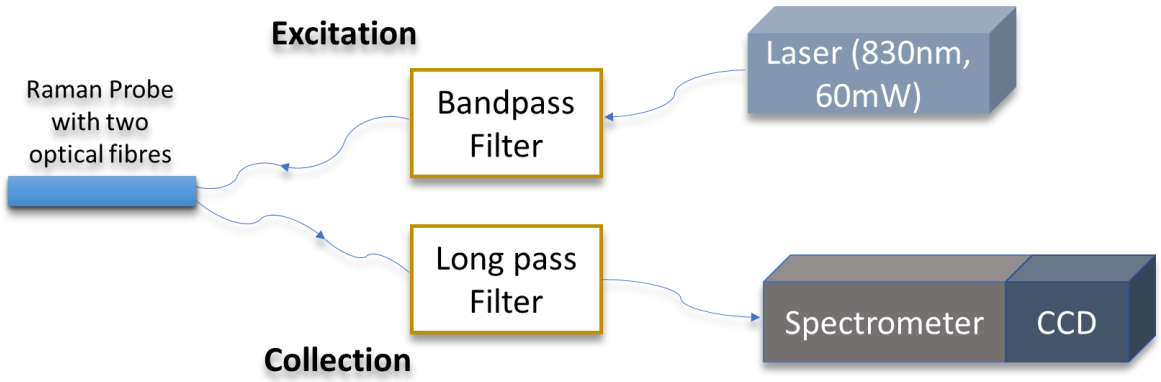


Figure 2.3: Raman system schematic.

Prior to the initiation of each set of experiments the system was calibrated in order to ensure the reproducibility of the CCD read-out so that recordings obtained on different days could be compared. Laser power output was measured and adjusted to 60 mW at the tip of the probe. Power output measurements were repeated often during the experiments to ensure the power remained consistent (at 60mW) throughout all the spectral recordings. Spectra from PTFE were acquired for wavelength calibration. Following the correction of PTFE offset an air background signal was also acquired in order to check visual consistency with previous measurements.

The *in vivo* experimental setup can be seen in figure 2.4. Hind-limbs of the anaesthetized animals were shaved prior to the experimental procedure. Active and sham procedures were performed with each animal having only one type of recording. For both procedures the needle was inserted in both the medial and lateral heads of both gastrocnemius muscles of each animal (4 sites/mouse) and the fibre-optic Raman probe was deployed through the tip of the needle by gently pushing the syringe. For the active Raman procedure, the laser was switched on, laser light shined into the muscle and the Raman spectra were acquired from the sites of interest. For the sham recordings the procedure was the same, but the laser was not switched on.

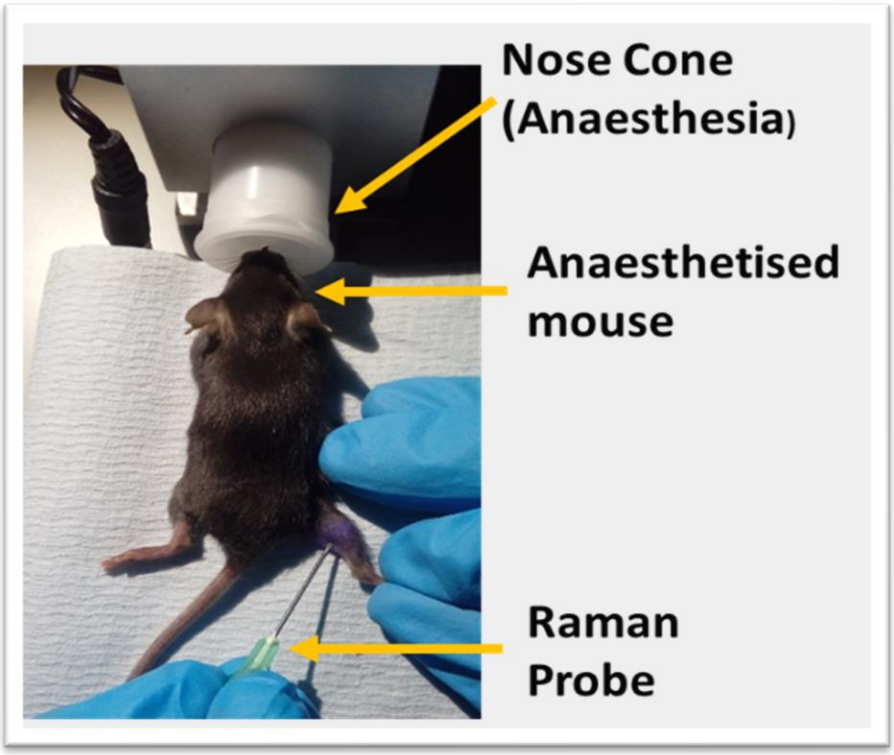


Figure 2.4: In vivo Raman experimental setup.
The probe is inserted in a 21G needle and the needle is inserted in the gastrocnemius muscle of the anaesthetised mouse

The fibre-optic Raman system described above was also used for the *ex vivo* recordings in mice muscle samples. The samples were stored at -80^{o}C and were left in room temperature to thaw prior the recordings. Power output of the laser was initially measured and adjusted to maximum power at the probe tip. For the recordings the samples were placed on white paper roll as shown in figure 2.5. Spectra were acquired at multiple sites of each sample (2-6 sites depending on the size of the sample). The probe was placed on the surface of the sample and ten spectra (acquisition time: 4s/spectrum) were acquired from each site. The ten spectra were then averaged prior data analysis.

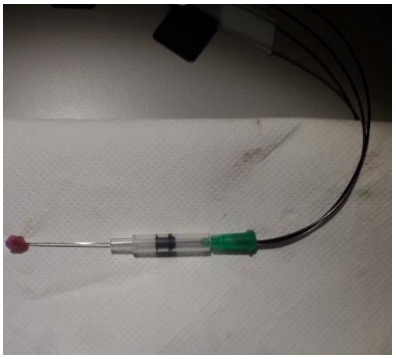


Figure 2.5: Ex vivo Raman experimental setup.
The sample was placed on white paper roll (animal tissue) and the spectra were recorded by placing the probe on the surface of the sample.

2.1.3.3 Combined electrophysiological and Raman recordings

Prior to the electrophysiological recordings the fur of both hindlimbs of the anaesthetised mice was shaved and the remaining fur was removed using depilatory cream to facilitate better contact.

All electrophysiological recordings were made using a Dantec Keypoint Focus EMG System (Optima, UK). To record the CMAP, twisted pair subdermal electrodes were used to apply supramaximal stimuli subcutaneously at the sciatic notch. Three different methods of recording the CMAPs were used whenever possible and compared: ring electrodes (standard method 1), needle electrode (standard method 2) and combined probe (new method). For the ring electrodes recordings, a standard 'belly-tendon' approach was used with one electrode placed over the muscles and another at the base of the ankle. A grounding electrode was placed in the base of the tail. For the recordings using the concentric needle electrode, the grounding electrode was again placed on the base of the tail and the needle electrode was placed in the gastrocnemius muscle. Similarly, when the combined probe was used to acquire the CMAPs, the grounding electrode was placed on the base of the tail and the probe was inserted in the gastrocnemius muscle. Two electrophysiological recordings were obtained when the combined probe was used with a recording of a Raman spectrum (acquisition time: 40s) in between (see section 2.1.3.2 for the Raman recording method). Whenever possible CMAPs were acquired from both hindlimbs of each mouse using all three methods in each hindlimb. When that was not feasible CMAP recordings were undertaken using one of the two standard methods and the combined probe.

A single, square wave electrical impulse of 0.1ms duration was applied to the sciatic notch in order to record the CMAPs. The stimulation intensity was increased until no further increase in the CMAP amplitude was observed (that is, a supramaximal response was obtained).

2.1.4 Post-experiment Motor Function Assessment

Motor function testing was carried out on all the animals in Study 1 in order to evaluate the impact of the *in vivo* Raman recordings on muscle function (see section 2.1.2 for time points and group sizes).

2.1.4.1 Accelerating Rotarod Test

All mice were trained on the rotarod test for three consecutive days prior the first set of performance recordings. For the performance recordings mice were placed on the rotarod for up to 300s. The rotarod (Ugo Basile 7650) was set to accelerate from 3 to 37 rpm in 270 seconds. Latency to fall was recorded in seconds for each mouse. Each mouse was allowed two runs and the best performance used for the data analysis. The rotarod apparatus can be seen in figure 2.6.

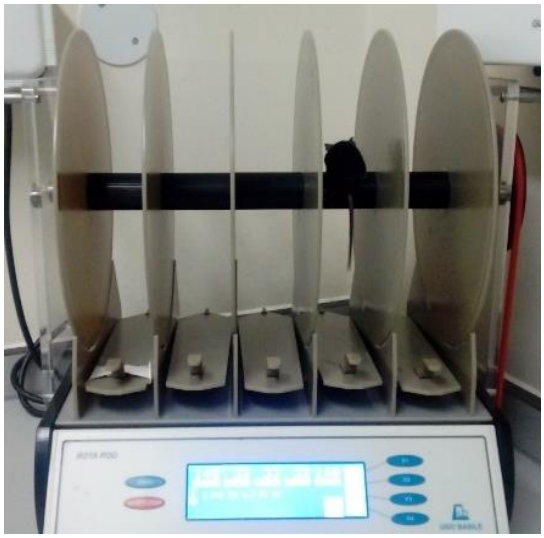


Figure 2.6: The rotarod apparatus.

2.1.4.2 Catwalk Gait Analysis

The CatWalk gait analysis system version 7.1 (figure 2.7) was used for gait analysis. The Catwalk system consists of an enclosed runway with a glass plate and two plastic panels perpendicular to the glass plate, an encased fluorescent tube that emits light inside the glass plate and a high speed colour camera positioned underneath the runway (Parvathy and Masocha, 2013a). Normally, the light emitted by the fluorescent lamp gets internally reflected and is therefore restricted to the glass surface plate. When the air in contact with the surface plate is replaced by a different medium, for example, when the animal's paw touches the glass floor, the light exits the glass and illuminates the contact area. As such, during an animal's run across the glass plate only the places of contact of the animal's paws with the glass floor light up (Deumens *et al.*, 2007). The run of the animal is recorded by the video camera underneath the runway (Deumens *et al.*, 2007). Data are acquired, stored, analysed and quantified by the CatWalk software 7.1 program.

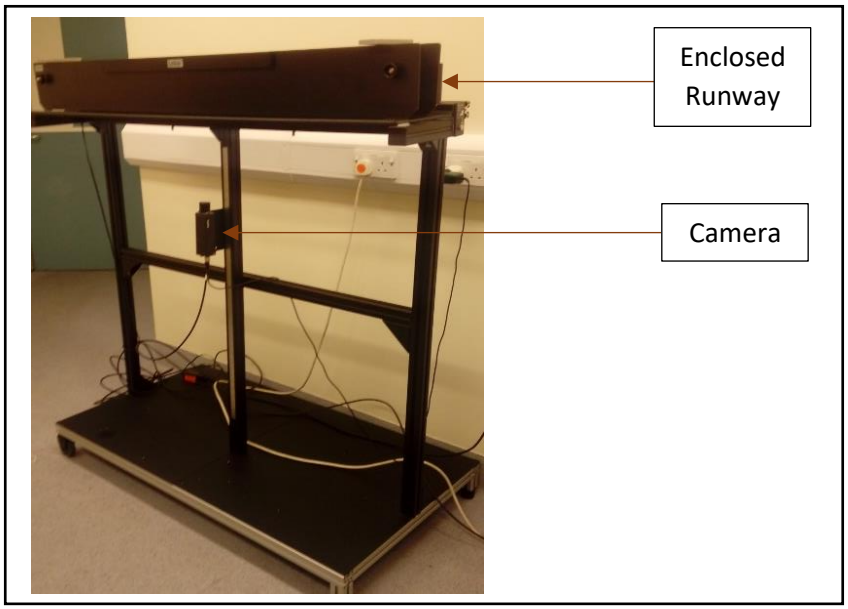


Figure 2.7: The Catwalk 7.1 system.

In this study mice were familiarised with the Catwalk runway and trained to cross it one day before the first recording. Motor performance of all the mice was then recorded before and after the Raman experimental procedure. Animals were placed on the catwalk apparatus in complete darkness and left to walk freely on the glass plate. Whenever possible six runs were acquired for each mouse and three were selected for analysis. Catwalk software was used to manually identify the footprints in each run (RF=Right Front, RH=Right Hind, LF=Left Front, LH= Left Hind) in the walkway panel, as can be seen in figure 2.8. After the footprint labelling, paw prints were displayed in the print view. The time-based gait diagram (i.e. duration of contact of the paws with the glass floor for all the step cycles in each run) was displayed in the timing view. Subsequently, the associated gait parameters were automatically calculated by the software. The data were then exported and collected using Excel.

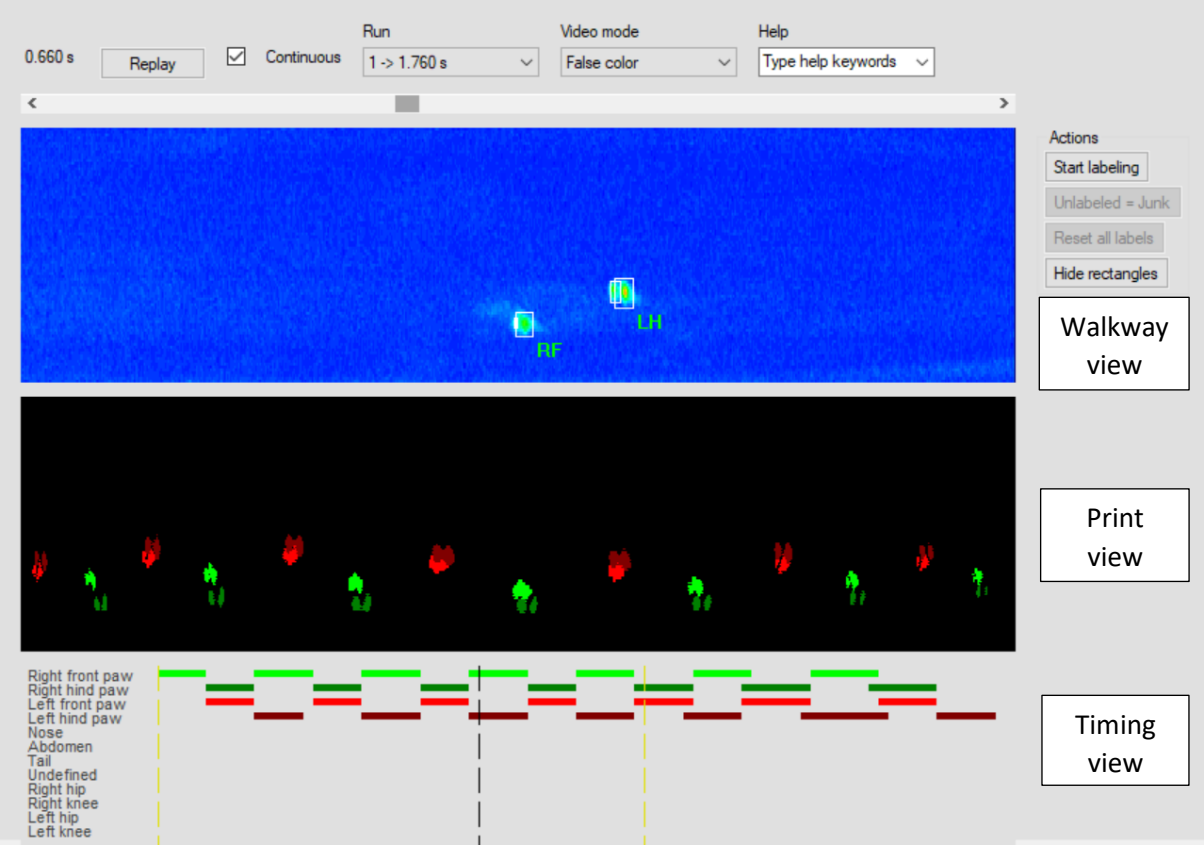


Figure 2.8: Catwalk 7.1 gait analysis software. *Each paw (digitised print) is manually labelled in every sequence in the walkway panel. Different colours are associated with the different paws in the print and timing panels. Paw prints are displayed in the print view. The bars in the timing view represent the stand for each paw and each step cycle.*

The gait parameters calculated by the Catwalk software were grouped in the following five categories a) Run characterization, b) Temporal parameters, c) Spatial parameters, d) Kinetic parameters, e) Interlimb coordination parameters (Caballero-Garrido *et al.*, 2017). The temporal, spatial and kinetic parameters as well as stride length are calculated for each paw separately. A full description of the parameters from each category that were analysed in this study is given in table 2.6. A graphical representation of some example parameters is given in figure 2.9.

Table 2.6: Full list and definitions of gait parameters. *Adjusted from* (Caballero-Garrido *et al.*, 2017).

Parameter	**Definition**
Run Characterisation	
Duration (s)	The duration of the entire run.
Temporal Parameters	
Stand or Stance Phase (s)	The duration of contact of a paw with the glass floor during a step cycle (fig. 2.9).

Swing (s)	The duration that a paw is not in contact with the glass floor during a step cycle (fig. 2.9).
Duty Cycle (%)	Stance phase duration as a percentage of the duration of the step cycle (fig. 2.9).
Initial Contact (s)	The time in seconds since the start of the run at which a paw makes contact with the glass floor (fig. 2.9).
Max Contact (s)	The time in seconds since the start of the run at which the largest part of a paw is in contact with the glass plate (fig. 2.9).

Spatial Parameters

Print Length (mm)	The length of the complete print (fig. 2.9). The complete print consists of the sum of all contacts of a paw with the glass plate.
Print Width (mm)	The width of the complete print (fig. 2.9).
Print Area (mm^2)	The surface area of the complete print (fig. 2.9).
Max Area (mm^2)	Maximum area of a paw that contacts the glass. (Print area at Max Contact).
Intensity (a.u)	The mean brightness of all pixels of the print at Max contact. The intensity is dependent on contact of the paw with glass plate and rises with increasing pressure. Thus, Intensity is a measure of weight support of each paw.

Kinematic Parameters

Swing Speed (m/s)	The speed of the paw during swing (fig. 2.9)
Stand Index	Measure for the speed at which the paw loses contact with the glass floor.

Interlimb Coordination

Base of Support (BOS) (mm)	The average distance between either the front paws or the hind paws (fig. 2.9).
Stride length (mm)	The distance between successive placements of the same paw (fig. 2.9).

Print positions (mm)	The distance between the position of placement of the hind paw and the position of placement of previously front paw on the same side of the body during one step cycle.
Step Sequence	Order in which the paws are placed on the glass floor.
Step Sequence: Number of Patterns	Number of patterns used through the entire run.
Step Sequence: Ca (%)	Cruciate step pattern (RF-LF-RH-LH).
Step Sequence: Cb (%)	Cruciate step pattern (LF-RF-LH-RH).
Step Sequence: Aa (%)	Alternate step pattern (RF-RH-LF-LH).
Step Sequence: Ab (%)	Alternate step pattern (LF-RH-RF-LH).
Step Sequence: Regularity Index (%)	The number of normal step sequence patterns (as described above) relative to the total number of all the paw placements in the run (fig. 2.9).
Support (%)	Percentage of a run when the animal is supported by zero, one, two three or four paws simultaneously on the glass plate in each Step Cycle.
Support: Zero (%)	Zero paws in contact with the glass plate
Support: Single (%)	One paw in contact with the glass plate.
Support: Diagonal (%)	Two paws in contact with the glass plate (LF-RH or RF-LH)
Support: Girdle (%)	Two paws in contact with the glass plate (LF-RF or RH-LH)
Support: Lateral (%)	Two paws in contact with the glass plate (LF-LH or RF-RH)
Support: Three (%)	Three paws in contact with the glass plate
Support: Four (%)	Four paws in contact with the glass plate

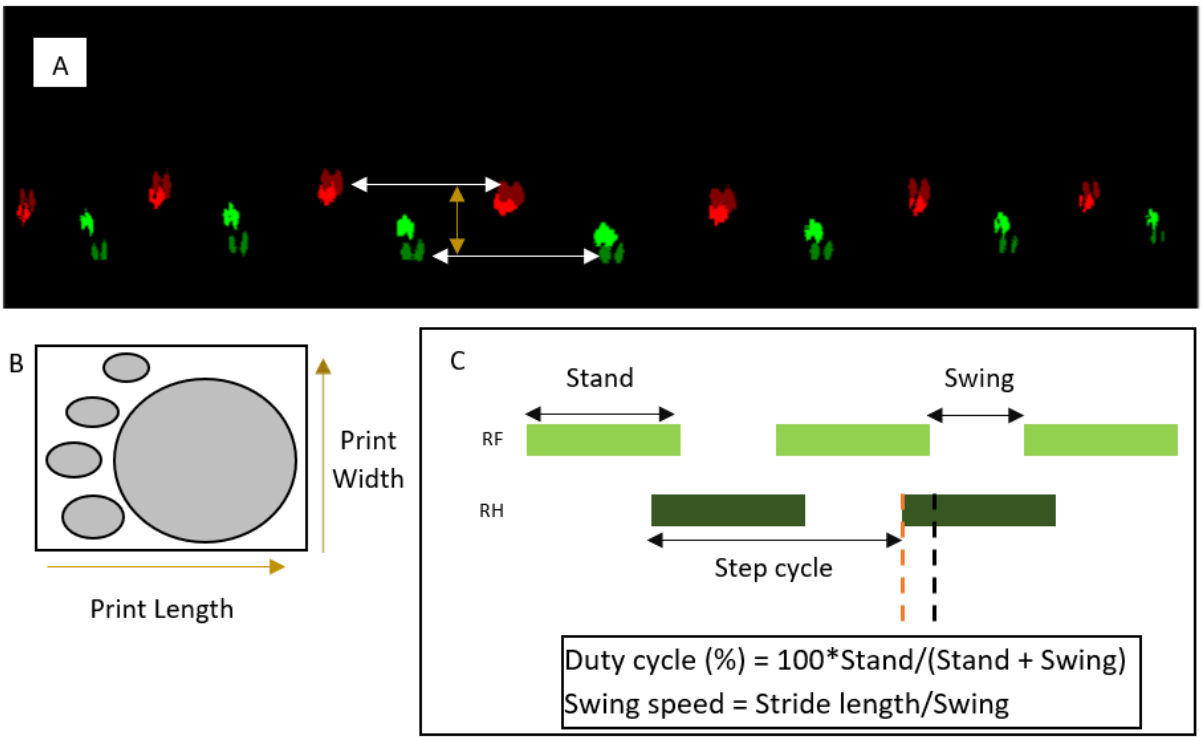


Figure 2.9: Graphical representation of gait parameters. *A) Stride lengths (in mm) for left and right hind paws (white arrows) and hindlimb base of support (in mm) (yellow arrow) shown in the print view panel. B) Print width (in mm), print length (in mm) and print area (mm^2) (shadowed area). C) Stand (in seconds), swing (in seconds), step cycle (in seconds), example of initial contact (in seconds) (orange dashed line) and max contact (in seconds) (black dashed line) of the second placement of the right hind paw shown in the time view panel. Formulas for the calculation of the duty cycle and swing speed displayed.*

2.1.5 Post-experiment Tissue Damage Assessment

Post-mortem MRI was used in a small cohort of animals from study 1 in order to assess the tissue damage induced by the *in vivo* Raman recordings (See section 2.1.2 for time points and group sizes).

2.1.5.1 Post-mortem Magnetic Resonance Imaging

The gastrocnemius muscles of both legs of the mice were imaged post-mortem, after the Raman recordings in order to assess potential muscle injury from the Raman recordings. Mice were culled either six hours (n=3), two days (n=3) or two weeks (n=3) post-experiment, the hind limbs were cut, placed into saline solution and transferred into the MRI unit, where they were scanned immediately. Each mouse leg was placed in a separate eppendorf tube filled with oil. The two eppendorfs were subsequently placed in a 50 ml falcon tube on a layer of oil soaked cotton, which was inserted in the scanner.

A 7 Tesla magnet (Bruker BioSpecAVANCE, 310 mm bore, MRI system B/C 70/30), with pre-installed 12 channel RT-shim system (B-S30) and fitted with an actively shielded, 116 mm inner diameter, water cooled, 3 coil gradient system (Bruker BioSpin MRI GmbH B-GA12. 400 mT/m maximum strength per axis with 80 ms ramps) was used for imaging. A ^{1}H birdcage volume resonator (Bruker, 300 MHz, 1 kWmax, outer diameter 114 mm/inner diameter 72 mm), placed

at the isocentre of the magnet was used for both radiofrequency (RF) transmission and reception. A workstation configured for use with ParaVisionTM 4.0 software operated the spectrometer. Following field shimming, off-resonance correction and RF gain setting a tri-plane FLASH sequence (Repetition time (TR) = 100 ms, Echo time (TE) = 6 ms, Flip angle = 30u, Number of Averages = 1, Field of view (FOV) = 80mm*80mm, slice thickness = 2mm, Matrix = 128*128, spatial resolution = 0.625mm*0.625mm) was used to localize the subject. Subsequently, fast rapid acquisition with refocused echo (RARE) sequences allowed low signal-to-noise ratio (SNR) visualisation of the area of interest and thus planning of axial high SNR RARE images (TR = 5500 ms - 6700ms, TE= 56 ms, Number of Averages=45-100 , FOV=20 mm*20 mm, Slice thickness = 0.3 mm, Matrix = 256*256) covering the entire region of interest. Fat suppression was used to prevent high intensity fat signal from obscuring the signal of interest (i.e. bright areas in the muscle due to injury).

2.1.6 Running wheel induced damage

Running wheel exercise was carried out on some of the animals in Study 2. Serum was collected from some of the animals in the same study (see section 2.1.2 for group sizes and study protocol).

2.1.6.1 Running Wheel Exercise

For the running wheel activity mice were caged individually. The running wheel consisted of a 37.8cm circumference plastic Fast Trac wheel attached on a 4 cm post fixed on the floor of the cage and it was placed in the corner of each cage. A magnet was glued to the underside of each wheel and a bicycle computer with reed switch was attached to the side of the cage. Mice were exercised voluntarily on the running wheel for 48 hours and the distance run by each individual mouse was measured and recorded daily. Food and water were administered ad libitum.

2.1.6.2 Serum CK Activity Measurements

The blood collection and CK measurements were done according to the treat-nmd neuromuscular network protocol (Carlson, 2014). Briefly, mice were anesthetised with an intraperitoneal injection of pentobarbital (100 µl) and blood was collected via cardiac puncture. Blood was allowed to clot for approximately four hours in room temperature. Clotted blood was subsequently centrifuged at 10,000 rpm for 10 minutes at 4°C and serum was collected. Serum was stored at -80°C.

Serum CK activity measurements were done using the Pointe Scientific Creatine Kinase (CK10) reagent. The reagent was prepared according to the manufacturer's instructions. The working reagent and serum were mixed in 96-well cell culture plates. Two aliquots of each serum sample were prepared. 100 µl of working reagent were added in each one of the used wells and the plate was pre-warmed at 37°C for 4 minutes. 2.5 µl of serum was then added to the wells and the absorbance was measured at 37°C every minute for three minutes at 340 nm using a microplate reader (PHERAstar). The absorbance of the two aliquots was averaged for each measurement. The differences in absorbance per minute ($^{\Delta A}/_{min}$) were calculated and averaged. To express the creatine kinase activity in units per litre ($^{U}/_{L}$) the following calculations were done:

$$\frac{U}{L} = \left(\frac{\Delta A/_{min}}{0.00622}\right) \times \frac{\text{Total volume}}{\text{Sample volume}}$$

$$\frac{U}{L} = \Delta A/_{min} \times 6592$$

2.1.7 Histology

2.1.7.1 Tissue Collection

Mice were sacrificed by cervical dislocation. After removing the skin, gastrocnemius muscles of both hind limbs were dissected from all the mice in Study 1. The quadriceps and tibialis anterior muscles of both hind limbs as well as the diaphragm muscle of all the mice in Study 2 were also dissected. Isopentane was used for tissue freezing. The isopentane was poured in a metal canister and the canister was placed in liquid nitrogen. After dissection, the muscles were embedded in optimal cutting temperature (OCT) compound, and immediately snap frozen in isopentane. The frozen samples were then stored at -80˚C.

2.1.7.2 Sectioning, H&E Staining and Imaging

Histological analysis was performed on gastrocnemius muscle acquired from one and three months old *mdx*, SOD1^{G93A}, C57Bl/6 and C57Bl/10 mice (n=3 samples/group). For sectioning, the muscle samples were acclimatised at -20˚C for an hour. 8 µm tissue sections were obtained using a cryostat (cryostat chamber and object temperature -20˚C) and mounted on uncoated, charged slides. Five sections were mounted per slide. Whenever possible twenty slides were acquired per sample with 80 µm intervals between every four slides. Slides were then stored at -20 ˚C.

Prior to staining, slides were left in room temperature to thaw for 30 minutes. The slides were then placed in 95% alcohol for 5 minutes and the cleared in tap water. Subsequently, the tissue sections were stained in Harris haematoxylin for 2 minutes and then washed in tap water. After being washed in Scott's tap water for a minute, the slides were stained in eosin for 5 minutes. Subsequently the tissue sections were washed in tap water and then quickly dehydrated in alcohols (75%, 90%, 100%, 100%). The sections were then cleared in xylene and mounted in DPX. The slides were imaged using a digital slide scanner (Nanozoomer series, Hamamatsu). The histological images were acquired using the NDP.view2 Viewing software (Hamamatsu).

2.2 Human Tissue

2.2.1 Ethics Statement and Tissue Collection

The collection and use of human tissue were approved by an NHS Research Ethics committee (reference 16/YH/0261).

Post-mortem human muscle from patients with a diagnosis of MND were obtained from Sheffield Brain Tissue Bank. Muscle collected with a core needle biopsy from patients with a diagnosis of mitochondrial disease, was obtained from Newcastle Mitochondrial Tissue Biobank. Healthy muscle tissue, collected during knee surgery from subjects with no known neuromuscular conditions, was also obtained from Newcastle Mitochondrial Tissue Biobank. Finally, open muscle biopsies were collected prospectively from patients undergoing investigation for possible neuromuscular disease at the Royal Hallamshire Hospital (Sheffield Teaching Hospitals NHS Foundation Trust).

2.2.2 Study protocol

The aims of this study were to:

- test if the fibre-optic Raman probe can record Raman spectra *ex vivo* from human muscle
- explore the potential of the technique to detect muscle pathology distinguish between healthy and diseased muscle
- to investigate at which stage of the diagnostic pathway Raman spectroscopy could be useful
- to compare the performance of the fibre-optic probe with that of a Raman microscope system

All the samples described in section 2.2.1 were used in this study. The number of samples in each group can be seen in table 2.7.

Table 2.7: Summary of the human muscle samples used in the study.

Group	Number of samples
MND tissue (Sheffield Brain Tissue Bank)	14
Mitochondrial disease (Newcastle Mitochondrial Tissue Biobank)	14
Tissue from healthy volunteers (Newcastle Mitochondrial Tissue Biobank)	10
Prospective muscle biopsies (Sheffield Teaching Hospitals)	39

The tissue samples were transferred to the Biophotonics Research Unit in Gloucester (Gloucestershire Hospitals NHS Foundation trust), where all the Raman recordings took place.

The samples were stored at -80°C and were left in room temperature to thaw prior the Raman recordings. *Ex vivo* Raman recordings were acquired with the fibre-optic Raman probe and with a Raman microscope system (see section 2.2.3) from all the samples except for four MND samples. Raman spectra from these samples were collected using the fibre-optic probe only.

2.2.3 Raman Spectroscopy

2.2.3.1 Optical Fibre Probe

The fibre-optic Raman system described in section 2.1.3.2 was also used for the *ex vivo* recordings in human samples. Power output of the laser was initially measured and adjusted to maximum power at the probe tip. The samples were placed on a Calcium Fluoride slide (figure 2.10). Spectra were acquired at multiple sites of each sample (2-6 sites depending on the size of the sample). The probe was placed on the surface of the sample and ten spectra (acquisition time: 4s/spectrum) were acquired from each site. The ten spectra were then averaged prior data analysis.

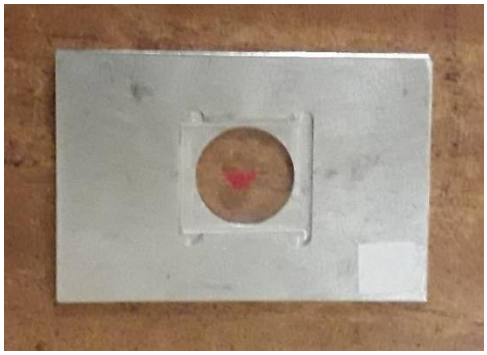


Figure 2.10: Ex vivo Raman experimental setup.
The sample was placed on a glass slide (human tissue) and the spectra were recorded by placing the probe on the surface of the sample.

2.2.3.2 Microscope

Following the recordings with the fibre-optic probe Raman spectra were also acquired using 830nm excitation wavelength with an acquisition time of 40s (in each site) using a Renishaw Raman spectrometer system 1000 (Renishaw Plc. Wotton-under-Edge, UK). Spectra were acquired using a x50 objective from multiple sites in each sample (2-6 sites depending on the size of the sample. The power at the objective was 30mW.

2.3 Combined Probe Construction

A fibre-optic Raman probe like the one used for the rest of the experiments (see section) was constructed. To build the combined probe the metal bore containing the two optical fibres was inserted in a heat shrink (Wall size: 0.00035'', ID: 0.020) and was heated using a hot air gun. Approximately 1 mm was left without insulation at both ends of the metal bore to allow for electrical contact (figure 2.11). The standard connectors from the EMG machine were then looped around the needle and the uninsulated end of the probe (contact site-figure 2.11) to make the contacts.

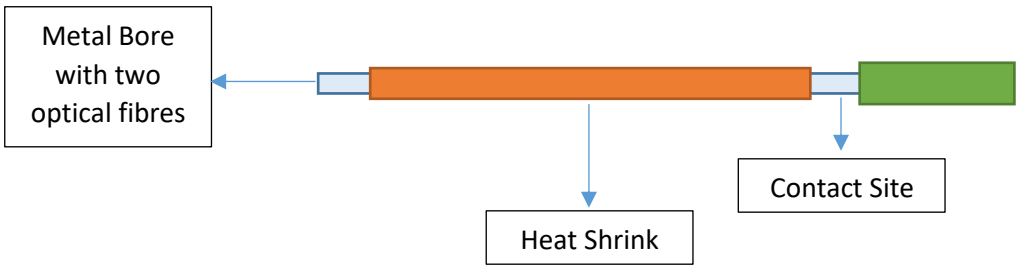


Figure 2.11: Combined Raman/EMG probe schematic.

2.4 Data Analysis

2.4.1 Spectral Analysis

Spectral analysis was done in Matlab environment (Matlab R2019b The MathWorks, Inc.,Natick, MA). Raw spectra were interpolated to integer wavenumber spacings between 900 and 1800 cm^{-1}, normalised using standard normal variate normalisation (SNV) and mean-centred (Barnes, Dhanoa and Lister, 1989). Spectra were windowed between 900 cm^{-1} and 1800 cm^{-1} as outside this region the spectra were dominated by background related to silica Raman signal generated in the optical fibres (before 900 cm^{-1}) or consisted of uninformative noise (after 1800 cm^{-1}). More specifically, there is a peak located around 800cm^{-1} that has been previously shown to be related to fused silica (Saavedra *et al.*, 2014). Since, this peak was very prominent in the Raman spectra acquired with the fibre-optic probe the spectra region up to 900cm^{-1} was excluded from the analysis in order to avoid the artefact of the probe fibres affecting the data analysis and being used for classification.

Principal component fed linear discriminant analysis (PCA-LDA), principal component fed quadratic discriminant analysis (PCA-QDA) and partial least squares discriminant analysis (PLS-DA) classification models were built for all the datasets. PCA was used to process the spectra and the most appropriate principal components (largest difference between the groups of interest) were selected using student's t-test (comparison between two groups) or analysis of variance (ANOVA) (comparison between three or more groups) applied to the PCs followed by false discovery rate (fdr) correction (Q=0.05). PCs up to PC12 were examined as components above that were found not to contain discriminatory information. In PLS-DA selection of the optimal number was done by increasing the number of components included to build the classification model until the ability of the algorithm to accurately predict the spectra no longer increased.

The classification performance of the different models was validated using leave-one-spectrum-out (LOO), leave-one-mouse/sample-out (LOMO/LOSO) and repeated leave-some-mice (RLSMOCV) out cross-validation (CV). As can be seen in figure 2.12 in the leave-out-cross validation the dataset is split into a training and test set. The classification model is built using the training set data. The test set is then projected on the model to evaluate its performance. The process is repeated until each spectrum has been left out once and its label (group) has been predicted by the model. In order to ensure that there is no prior information of the test dataset in the training set model, components to include were selected for each training dataset (without including the spectra of the given test set). In the case of the repeated leave-some-out cross-validation the entire cross-validation process was repeated 100 times using different, randomly selected combination of left out mice.

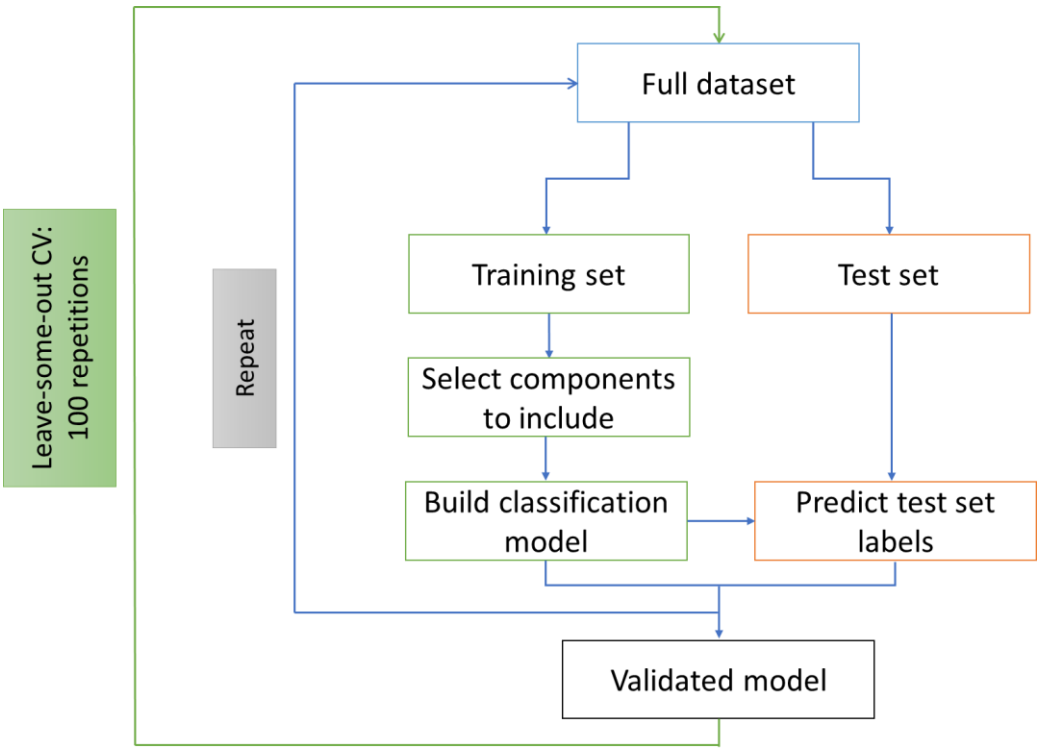


Figure 2.12: Leave-out cross-validation.
The data set is split into training and test sets. The model is generated using the spectra in the training set and validated using the label predictions from the data in the test set. The cycle is repeated until each spectrum is left out once. For repeated leave-some-mice/samples out the whole process was repeated 100 times.

For the *in vivo* animal work each animal group consisted of the same number of mice (see methods section 2.1.2 for animal numbers). For models generated from mice groups consisting of 16 mice each, the repeated leave-two-mice/group cross-validation was performed. For example, in the case that two groups were used to build the model, four mice (two from each group) were left out in each step of the cross-validation. This was done to avoid biasing the model by leaving out 4 mice of the same group in any step. For models generated from mice groups consisting of 8 mice leave-one-mouse/group cross-validation was performed. The human samples (*ex vivo* work) were not split in equally sized groups and this cross-validation was not possible.
Background subtraction was performed for clarity but not as a pre-processing step (Lieber and Mahadevan-Jansen, 2003).
PCA, LDA and PLS-DA techniques are presented in the following sections.

2.4.1.1 Principal Components Analysis

Principal Component Analysis is a method for exploratory data analysis (unsupervised learning) and a dimensionality reduction technique. The main aim of dimensionality reduction techniques is to simplify the data (describe a given data set with a reduced number of variables) while preserving as much of the variation in the original data set as possible. PCA computes a new set of uncorrelated latent variables that are linear combinations of the original variables. These new

variables or coordinates (Principal Components) are defined by the directions that maximise the variance in the dataset and the most important information present in the data can usually be described using the first few PCs. A new orthogonal coordinate system can, thus, be formed using only the most informative dimensions (only the PCs of interest) (figure 2.13). Each Raman spectrum can be accurately fitted to a linear combination of these PCs.

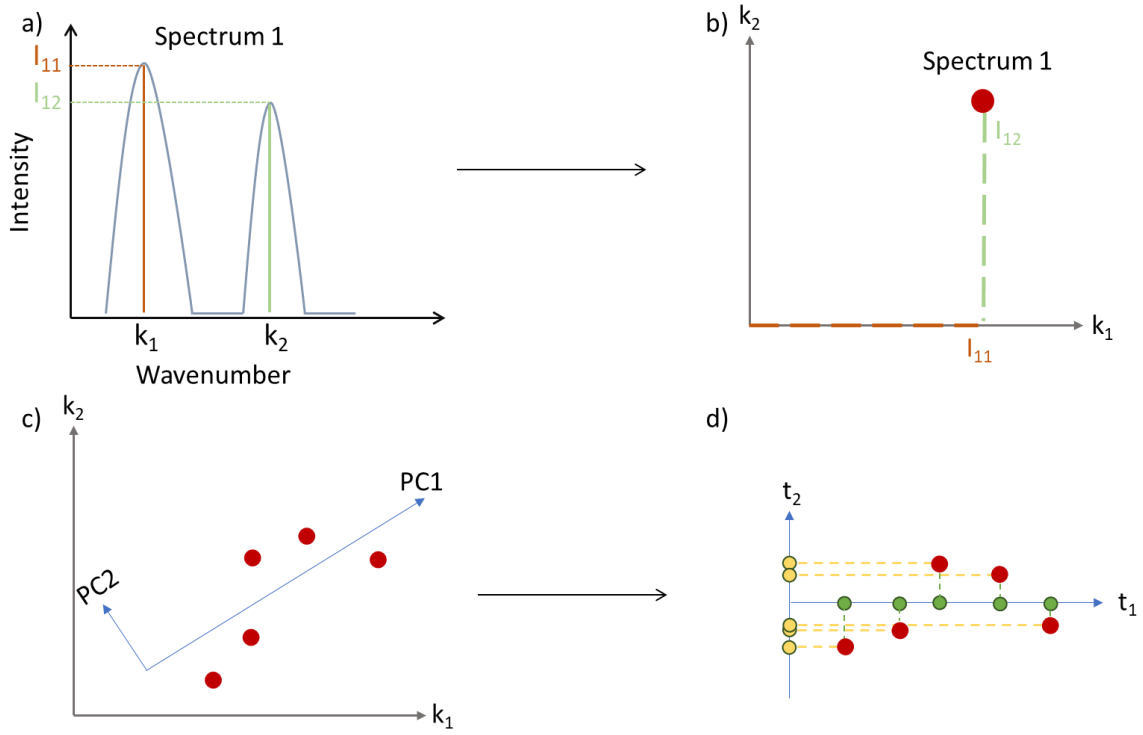


Figure 2.13: Illustration of principal component analysis for example data set with two original variables. *a) Example spectrum with two wavenumbers. b) The spectrum represented as a data point in the spectral space (space with as many dimensions as the original variables). c) More spectra (red dots) in the spectral space. The two principal components (directions of maximum variance) are also illustrated (blue arrows). d) The spectra are projected on the direction of the principal components (green and yellow dots) and the score of each spectrum for each PC (t_1 for PC1 and t_2 for PC2) can be computed by projecting the data in the direction of the PCs (dashed lines).*

Mathematically, the eigenvectors of the covariance matrix of the original variables represent the directions of the axes where there is most variance. The eigenvalues associated with each eigenvector give the amount of variance present in this direction (in this PC). Thus, if eigenvectors are ranked in order of decreasing eigenvalues, the principal components are ranked in order of significance (Varmuza and Filzmoser, 2009).

For a data set X consisting of n observations (spectra) and m original variables the direction with the most variance is the first principal component and is defined by a loading vector:

$$\boldsymbol{p_1} = (p_1 \dots . . p_m)$$

The scores of each component, which are orthogonal projection coordinates on the direction of the respective loading vector, are linear combinations of the original variables and the respective

loadings (Varmuza and Filzmoser, 2009). For observation i, defined by the vector $\boldsymbol{x_i}$ with elements $x_{i1} \dots x_{im}$ the score t_{i1} for PC1 is:

$$t_{i1} = x_{i1}\, p_1 + \cdots + x_{im}\, p_m \qquad i = 1 \dots . n$$

And so, for all n observations arranged as rows in a matrix $\boldsymbol{X}$ the score vector, $\boldsymbol{t_1}$, of PC1 is obtained by:

$$\boldsymbol{t_1} = \boldsymbol{X} \cdot \boldsymbol{p_1}$$

The second principal component (PC2) is defined as an orthogonal direction to PC1 and possessing the largest residual variance of the scores (Varmuza and Filzmoser, 2009). PCs up to the number of the original variables can be calculated and each PC is orthogonal to all the previously computed ones. Due to the fact that the amount of explained variance decreases in every next PC, the variances of the later PCs are often very small or zero (Varmuza and Filzmoser, 2009). These PCs represent noise in the dataset and can be removed from subsequent analysis (dimensionality reduction).

If all the loading vectors are collected as columns in the loading matrix, $\boldsymbol{P}$, and all score vectors are collected in the score matrix, $\boldsymbol{T}$, (figure 2.14) the PCA scores are computed by the following equation:

$$\boldsymbol{T} = \boldsymbol{X} \cdot \boldsymbol{P}$$

The loading vectors are orthogonal to each other and their lengths are normalised to 1:

$$\boldsymbol{p_j^T} \cdot \boldsymbol{p_j} = 1 \qquad j = 1 \dots . m$$
$$\boldsymbol{p_j^T} \cdot \boldsymbol{p_k} = 0 \qquad j, k = 1 \dots . m \; and \; j \neq k$$

The PCA scores are also orthogonal to each other and uncorrelated (Varmuza and Filzmoser, 2009).

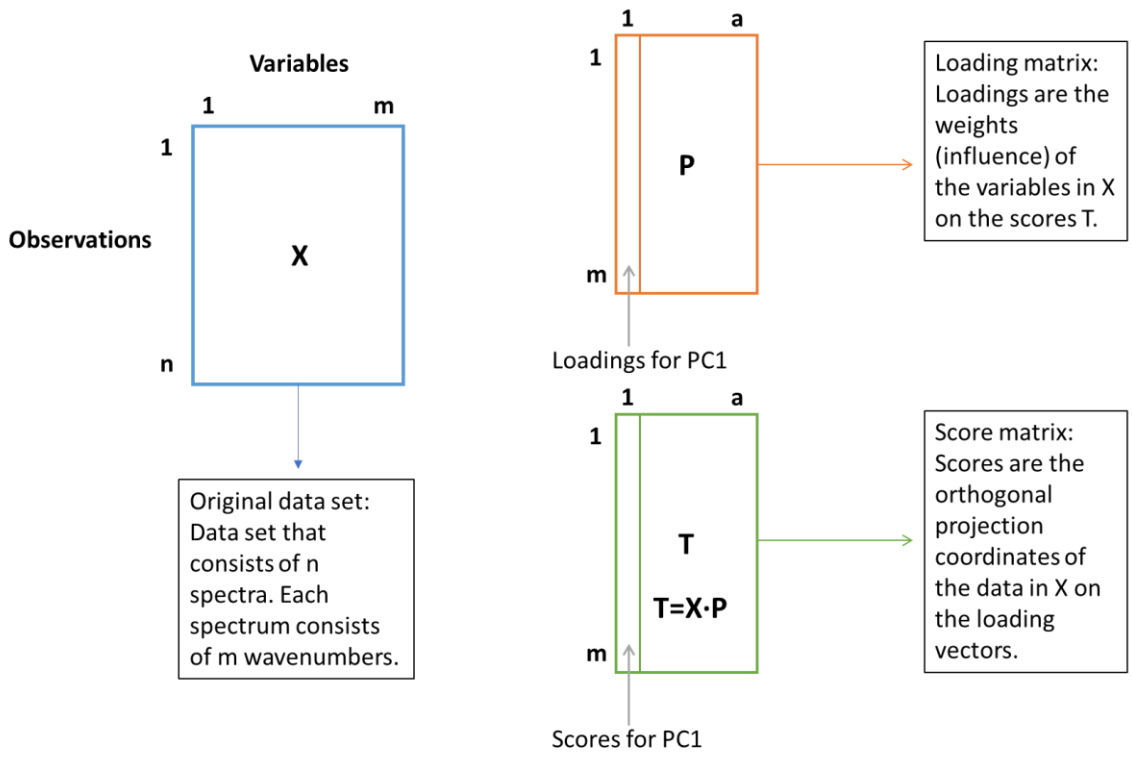


Figure 2.14: Graphical representation of PCA matrices.
Since PCA is used for dimensionality reduction the number of components used (a) is usually smaller than the number of original variables (m). Adjusted from (Varmuza and Filzmoser, 2009).

The PCA scores can be used to reconstruct the original dataset (X-matrix) (Varmuza and Filzmoser, 2009). To do that, usually, only the PCs that preserve the most important information are employed. Thus, an approximated X-matrix (X_a) with reduced noise can be constructed. If all possible PCs were used, the error (residual) matrix E_{res} would be zero but that would lead to overparameterisation of the spectra and would defeat the purpose of using PCA for dimensionality reduction (Varmuza and Filzmoser, 2009).

$$\boldsymbol{X_a = T \cdot P^T} \qquad \boldsymbol{X = T \cdot P^T + E_{res}} \qquad \boldsymbol{E_{res} = X - X_a}$$

As already mentioned above PCA is an unsupervised method, which means that there is no knowledge of the different groups present in the dataset (e.g. diseased/healthy spectra). Thus, PCA identifies the directions in the spectral space that maximise variation in the data independently of the sample groups and is in many cases not ideal for spectral discrimination and classification (Notingher *et al.*, 2004). PC's that are diagnostically significant (variance in the data is due to differences between healthy and diseased tissues) are therefore often used for further analysis with 'supervised' techniques like multiple regression or discriminant analysis.

2.4.1.2 Linear Discriminant Analysis

Linear discriminant analysis is a supervised method that is often used as a classification technique. It can be described using two different approaches, the Bayesian and the Fisher approach. The

non-Bayesian approach will be used in this chapter. LDA finds the directions in the spectral space (or PC's space if used after PCA) that maximise the separability of the different classes. These directions, also known as Linear Discriminant Functions (LDs), are linear combinations of the original variables. Utilising the Fisher's criterion, LDs are defined by the directions that maximise the inter-group variance, while minimising the intra-group variability of specified groups (Conrad and Bonello, 2016)(Conrad and Bonello, 2016). For a data set X with n observations belonging in two different groups and m original variables the linear discriminant function is:

$$y = b_1 x_1 + \cdots + b_m x_m$$

The coefficients $b_1 \dots b_m$ form a decision vector b and projecting the observations on an axis defined by the decision vector give the discriminant scores of the observations y_i (Varmuza and Filzmoser, 2009). A natural measure of class separation between the projected points is the difference between their means. In order to have good class separation and hence good classification performance in a given direction, the separation of the means needs to be as large as possible in that direction relative to some measure of dispersion of the observations within each class (that needs to be minimised).
If y_1 is the arithmetic mean of the discriminant scores of the observations belonging in the first group and $y_2$is the arithmetic mean of the discriminant scores for the second group, the LD is defined by the vector b_{fisher} that maximises the function:

$$J(b) = \frac{|y_1 - y_2|}{S_y}$$

Where S_y is the square root of the pooled variance, which is a weighted sum of the variances of S_1^2 and S_2^2 of y for the two groups (i.e. measure of the dispersion of the observations within each class) (Varmuza and Filzmoser, 2009).

Using b_{fisher} and if $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ are the arithmetic mean vectors of the data from groups 1 and 2 respectively the classification threshold is given by the following equation:

$$y_0 = \frac{\boldsymbol{b_{fisher}^T x_1} + \boldsymbol{b_{fisher}^T x_2}}{2}$$

Which is the mean of the scores obtained by projecting the group means on the discrimination direction (Fig. 2.15).

The LDA classification model built using the original data set can then be used as a predictive model of group membership for new observations. The discriminant score y_i of a new observation $\boldsymbol{x_i}$ is:

$$y_i = \boldsymbol{b_{fisher}^T x_i}$$

The new observation is classified to one of the two groups by comparing y_i to y_0 (Varmuza and Filzmoser, 2009).

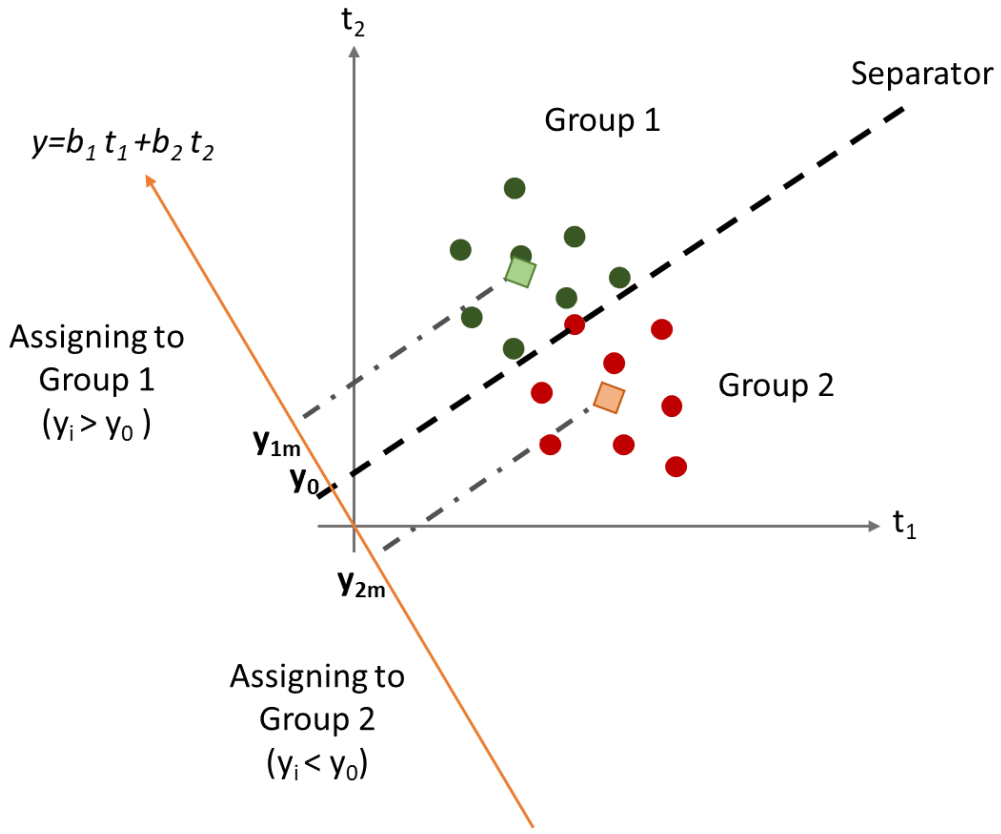


Figure 2.15: Fisher discriminant rule in a 2-D PC's space.
Score scatter plot of PC1 and PC2 scores with observations shown as dots. A discriminant variable y is computed, and the group assignment is done by comparing the discriminant score of an observation y_i to a threshold value y_0 (average of the projected group means y_{1m} and y_{2m} - dashed line). The diamonds represent the group means and they are projected (using the dashed-dotted line) on the discriminant variable y. The dashed line is the separator of the two classes in the 2-D space. For n-dimensional data (i.e. in an n-dimensional space) the separation line becomes a hyperplane. Adjusted from (Varmuza and Filzmoser, 2009).

For two classes there is only one linear discriminant function (shown in figure 6). LDA can be used to build a classification model when there are more than two groups. For n classes there are $(n-1)$ LDs. Once the classification model is developed using a given dataset, the membership of unknown samples to one of the defined classes can be predicted (predictive model) (Ballabio and Consonni, 2013). A limitation of LDA is that the number of original variables is required to be smaller than the number of observations (Notingher *et al.*, 2004). If there are more variables than samples PCA can be performed as the first step of the multivariate analysis and LDA is then performed on PC scores of relevant PCs. LDA requires an assumption of equal covariance matrices between the input variables of the different classes. This means it does not take into account any

differences in the variance structures of each group (Brereton and Lloyd, 2014). Unlike LDA, QDA allows for different variance structures for each group (Brereton and Lloyd, 2014). In QDA the separator (boundary that separates the classes) is no longer linear; it becomes quadratic.
Classifying spectra that have been used to generate the classification model in order to assess the predictive performance of the PCA-LDA model can lead to over-optimistic results. 'Overfitting' can be avoided by using some of the spectra of the original dataset to build the model and some other to assess its performance. However, this solution leads to less spectra being used to build the model. Instead leave-out cross-validation can be used.

2.4.1.3 Partial Least Squares-Discriminant Analysis (PLS-DA)

PLS was initially introduced as a regression method. And as such it was proposed to handle continuous response variables (Lee, Liong and Jemain, 2018). In order to use it as a discriminant analysis tool the response variable must contain the group information and is therefore categorical (replaced by dummy variables describing the different categories) (Lee, Liong and Jemain, 2018). PLS-DA is a dimensionality reduction and classification method that has been used extensively in chemometrics for prediction model construction. PLS-DA computes new latent variables (LVs-PLS components) that are linear combinations of the original variables to model the relevant sources of data variability (Ballabio and Consonni, 2013). The LVs are defined by the directions that have a maximum covariance between the data and the class variable. This criterion ensures maximum variance of the data in the spectral space and high correlation with the interesting property (the group label) (Varmuza and Filzmoser, 2009). Like in PCA, a new coordinate system can thus be formed using only the most informative dimensions (dimensionality reduction). Similarly, the scores are the coordinates of the observations in the direction of the LVs and the loadings are the coefficients of the original variables in the linear combinations which determine the LVs and as such they can be interpreted as the influence of each original variable on each LV (Ballabio and Consonni, 2013). The description of PLS-DA as a supervised PCA needs to happen with caution though since the actual score and loading matrices computed by the two different methods differ in some of their properties. For example, in PLS-DA, unlike in PCA, the loadings are not orthogonal (Brereton and Lloyd, 2014). The fundamental PLS-DA equations for a matrix X (original dataset) consisting of n observations belonging in two different groups are:

$$\boldsymbol{X} = \boldsymbol{T} \cdot \boldsymbol{P^T} + \boldsymbol{E}$$

$$\boldsymbol{c} = \boldsymbol{T} \cdot \boldsymbol{q^T} + \boldsymbol{f}$$

Where T and P and q are score and loading matrices respectively. E and f are residuals (Brereton and Lloyd, 2014).
After the classification model is built prediction of group membership for new or training dataset observations can be achieved by calculating the c_i value (class value) and comparing it with a threshold value (usually the value between the two class labels).

Thus, PLS-DA is a supervised method that combines dimensionality reduction and discriminant analysis into one algorithm. Unlike LDA the number of observations does not need to be larger than the number of variables and it does not assume equal covariance matrices between the two groups (Lee, Liong and Jemain, 2018). PLS-DA can also be extended to the case where the dataset is split into more than two groups. Similar to LDA, PLS-DA is also susceptible to overfitting. Thus, cross-validation is an important step when building a PLS-DA classifier (Ruiz-Perez *et al.*, 2018).

2.4.1.4 Measures of Predictive Models

Three very common measures of the ability of a statistical model to correctly predict the class of the observations of a given data set are the sensitivity, specificity and accuracy. The confusion matrix is often used to describe the performance of a classifier. Figure 2.8 shows the confusion matrix of a data set with two classes.

Table 2.8: Confusion matrix of a two-group model.

		Predicted Class	
		1	**2**
True Class	**1**	True Positive (TP)	False Negative (FN)
	2	False Positive (FP)	True Negative (TN)

Using class 1 as the positive class the sensitivity, specificity and accuracy are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN} \times 100$$

$$Specificity = \frac{TN}{TN + FP} \times 100$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

Thus, the sensitivity is the proportion of the positive class observations that were correctly classified by the model. Specificity is the proportion of the negative class observations that were classified as negatives by the model. Finally, accuracy is the percentage of all the observations that were correctly classified. A perfectly accurate model would have zero false positives and false negatives and would result in sensitivity, specificity and accuracy values of 100%.

Receiver Operating Characteristic (ROC) curves are plots of the true positive rate (sensitivity) against the false positive rate (1-specificity) and illustrate the ability of a classifier to separate between classes as the discrimination threshold is varied. Thus, the area under an ROC curve is also a measure of how well the model can distinguish between the two classes. A model with AUC near to one is well able to separate the two classes. A model with AUC close to 0.5 approaches a random coin toss. Sensitivity, specificity, accuracy and AUC are often used in medicine to describe the ability of a diagnostic test to correctly predict the presence or absence of a medical condition. In that case the diseased group is considered to be the positive class. Hence, the sensitivity of a

diagnostic test is defined as the ability of the test to correctly identify the individuals with the disease and the specificity of the test is its ability to correctly classify the individuals without the disease.

In the case of assigning a sensitivity, specificity and accuracy to the two-group models used in this study the medical approach was used when one of the two groups consisted of spectra acquired from diseased tissue and the other one from healthy. In these cases, the diseased group was considered as the positive class and the sensitivity demonstrated the ability of the model to correctly identify the diseased spectra. When this was not the case one of the two groups was chosen as the positive class and this will be stated in the results section. For the multi-group models the performance indices for each group were calculated by using each one group as the positive class and the rest of the groups as the negative class (one-versus-all approach).

2.4.2 Rotarod and Catwalk Data Analysis

GraphPad Prism version 8 was used for statistical analysis of rotarod data (GraphPad, San Diego, CA, USA). The rotarod performance data were not normally distributed for all the different sub-groups and time points. All rotarod recordings at two time points were analysed using paired sample t-test for normally distributed data and Wilcoxon matched-pairs signed rank test for non-normally distributed data. For the analysis of rotarod performance at three time points repeated measures one-way ANOVA with Turkey's correction for multiple comparisons (when ANOVA was significant) was used for normally distributed data and the Friedman test for non-parametric data. Mean and standard deviation are shown in the graphs that display normally distributed data. Median and interquartile range are shown in the graphs that display non-normally distributed rotarod data.

Matlab (Matlab R2019b The MathWorks, Inc.,Natick, MA) was used for statistical analysis of the Catwalk gait parameters. Multiple paired sample t-tests (one for each parameter) were performed between two time points (one week before vs one day after-data collected from the full cohort in Study 1 and one week before vs two weeks after – data collected from half the cohort, which had the third recording) followed by Benjamini & Hochberg/Yekutieli false discovery rate control. The adjusted p-values (q-values) were also calculated and shown when significant.

2.4.3 Running Wheel Distance, CK Levels and EMG performance Data

GraphPad Prism version 8 was used for statistical analysis of creatine kinase levels and EMG performance in studies 2 and 5 respectively (GraphPad, San Diego, CA, USA). In study 2 student's t-test and one-way ANOVA were used in order to compare the distance run by two different mice groups and the CK levels of the different mice groups (exercised *mdx,* non-exercised *mdx* and C57Bl/10 mice) respectively. In study 5, mixed-effect model repeated measures ANOVA was with Turkey post-test was used to compare the CMAP amplitudes acquired from the same mice with the different techniques. Student's t-test was used to compare the CMAP amplitude of $SOD1^{G93A}$ with that of control mice. Paired t-test was used to compare CMAP amplitude from electrophysiological recordings acquired with the combined probe from the same mice before and after the Raman recordings.

3. Results

3.1 Animal Experiments

3.1.1 Summary of experiments

3.1.1.2 In vivo

A total of 8640 spectra were recorded from gastrocnemius muscles of mice *in vivo*. A summary of the spectra acquired for each of the studies can be seen in Table 3.1. As already described in the methods (see section 2.1.3.2) for each mouse the probe was inserted in both medial and lateral heads of both gastrocnemius muscles (4 sites/mouse). Acquisition time was 4 seconds for each spectrum and 10 spectra were acquired at each site (total acquisition time at each site 40 seconds). The ten spectra acquired in each site were averaged prior the multivariate analysis resulting in four spectra per mouse.

Table 3.1: Summary table of in vivo experiments.

Study	Mice Groups	Number of mice/group	Number of averaged spectra
Study 1	One month old *mdx*	16	512
Neurogenic/Myopathic Neuromuscular Disorders	Three months old *mdx*		
	One month old $SOD1^{G93A}$		
	Three months old $SOD1^{G93A}$		
	One month old C57Bl/6		
	Three months old C57Bl/6		
	One month old C57Bl/10		
	Three months old C57Bl/10		
Study 2	Three months old exercised *mdx*	16	128
Intervention Study	Three months old non-exercised *mdx*		
Study 3	Three months old $SOD1^{G93A}$	8	96
MND Study	Three months old TDP-43		
	Three months old $TDP\text{-}43^{q331k}$		
Study 4	One month old male *mdx*	8	124
Male *mdx* mice	Three months old male *mdx*	8	
	One month old male C57Bl/10	8	
	One months old male C57Bl/10	7	

Study 5	Three months old $SOD1^{G93A}$	8
Combined Probe Study	Three months old C57Bl/6	

The first study aimed to explore the ability of the technique to distinguish between healthy and diseased muscle, to detect muscle pathology at different disease stages and to discriminate between neuromuscular disorders of different origin. The second and third studies aimed to investigate the potential of detecting more subtle pathological changes or differences in the muscle, as could be the case after an intervention (Intervention study), or between different animal models of the same disorder (MND study). The fourth study aimed to assess the performance of the models generated with male *mdx* mice and compare it to that of the models generated using female *mdx* mice. Finally, the fifth study aimed to test the combined probe, that can acquire both Raman and electrophysiological data.

3.1.1.3 Ex vivo

A total of 3000 spectra were recorded from gastrocnemius, tibialis anterior, quadriceps and diaphragm muscles of exercised and non-exercised *mdx* mice *ex vivo*. A summary of the *ex vivo* spectra acquired for study 2 can be seen in Table 3.2. As already described in the methods (see section 2.1.3.2) for each sample the probe was inserted in multiple sites. Acquisition time was 4 seconds for each spectrum and 10 spectra were acquired at each site (total acquisition time at each site 40 seconds). The ten spectra acquired in each site were averaged prior the multivariate analysis.

Table 3.2: Summary table of ex vivo recordings.

Muscle	Group	Number of samples	Total number of averaged spectra
Gastrocnemius	Exercised *mdx* mice	3	18
	Non-exercised *mdx* mice	4	24
Tibialis Anterior	Exercised *mdx* mice	11	44
	Non-exercised *mdx* mice	12	48
Quadriceps	Exercised *mdx* mice	8	48
	Non-exercised *mdx* mice	10	60
Diaphragm	Exercised *mdx* mice	8	32
	Non-exercised *mdx* mice	9	36

Ex vivo measurements were undertaken in order to explore if the various muscles were affected differently by the running wheel exercise. Moreover, the ability of the technique to distinguish between different muscles was also assessed.

3.1.2 Study 1: Neurogenic/Myopathic Neuromuscular Disorders

The work presented in the following sections as Study 1 has been published in ACS Chemical Neuroscience (Plesia *et al.*, 2021).

3.1.2.1 Classification Models Performance

PCA-LDA, PCA-QDA and PLS-DA were used to generate classification models. The performance of the models was validated using leave-one-spectrum out (LOOCV), leave-one-mouse-out (LOMOCV) and repeated leave-some-mice-out (RLSMOCV) cross-validation (CV). The results for the different two group models cross-validated using LOOCV and LOMOCV are presented in section 3.1.2.1.1. Repeated cross-validation is presented in section 3.1.2.1.2.

3.1.2.1.1 Two Group Models

a) One month old *mdx* vs. C57Bl/10 mice

This is a very early disease stage for the *mdx* mice. Thus, the model generated using the spectra from one month old *mdx* and control mice aimed to explore the ability of the technique to detect muscle pathology at an early stage. The predictive capability of the models using different analysis and CV methods for the one month old *mdx* mice can be seen in table 3.3. The ROC curves, comparing the three different approaches for LOOCV and LOMOCV can be seen in figure 3.1.

Table 3.3: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods for the one month old mdx and C57Bl/10 mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	73.4%	65.5%	70.0%
Leave-one-mouse-out CV	70.3%	62.1%	66.4%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	81.3%	63.8%	72.9%
Leave-one-mouse-out CV	79.7%	62.1%	71.3%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	84.4%	82.8%	83.6%
Leave-one-mouse-out CV	84.4%	75.9%	80.3%

Using LOMOCV the ability to correctly classify spectra decreased for PCA-LDA and PLS-DA models. The sensitivity of the PCA-QDA model also decreased when LOMOCV was used. However, the specificity of the model slightly increased leading thus to a smaller drop in the accuracy. The area under the ROC curves also decreased for LOMOCV. However, the AUC values did not fall below 0.7. In general, the models attained better sensitivity than specificity values with most sensitivity values (for different methods and CVs) being above 70%.
For both CV approaches the PLS-DA model performed better than the PCA related models. The LOMO cross-validated PLS-DA model demonstrated a good performance for the early disease stage achieving an accuracy of 80.3% and an AUC of 0.89.

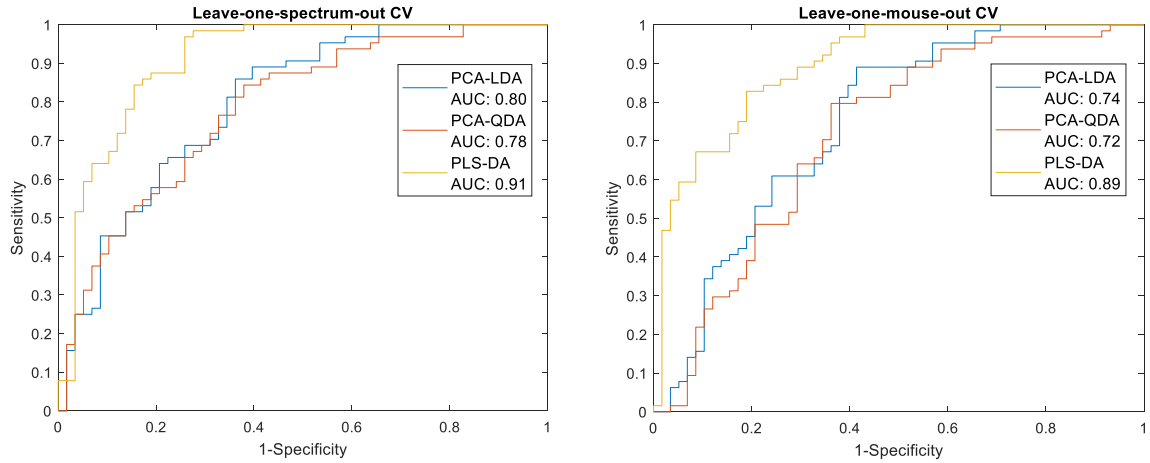


Figure 3.1: ROC curves for the cross validated classification models for the one month old mdx and C57Bl/10 mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

b) Three months old *mdx* vs. C57Bl/10 mice

This is a more established disease stage. Thus, the model generated using the spectra from three months old *mdx* and control mice aimed to explore the ability of the technique to detect muscle pathology at a later stage. The predictive capability of the models generated using different techniques and CV methods for the three months old *mdx* mice can be seen in table 3.4.

The classification performance of the models was better for the three months old *mdx* mice compared to one month old mice. The different cross-validation approaches did not have a big impact on the models, with PLS-DA achieving same classification performance with both CV methods and PCA-QDA achieving very similar ones. Only the PCA-LDA model demonstrated a slightly decreased accuracy, mainly due to a drop in specificity, when using LOMOCV. The PLS-DA model demonstrated the best classification performance, attaining a sensitivity of 93.8% and specificity of 92.2%.

Table 3.4: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods for the three months old mdx and C57Bl/10 mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	90.6%	75%	82.8%
Leave-one-mouse-out CV	89.1%	72%	80.5%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	87.5%	75%	81.3%
Leave-one-mouse-out CV	85.9%	76.6%	81.3%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	93.8%	92.2%	92.9%
Leave-one-mouse-out CV	93.8%	92.2%	92.9%

The ROC curves for the different techniques and CV methods can be seen in figure 3.2. The AUROC was equal or above 0.85 for all the models generated using PCA and it was above 0.96 for both PLS-DA models.

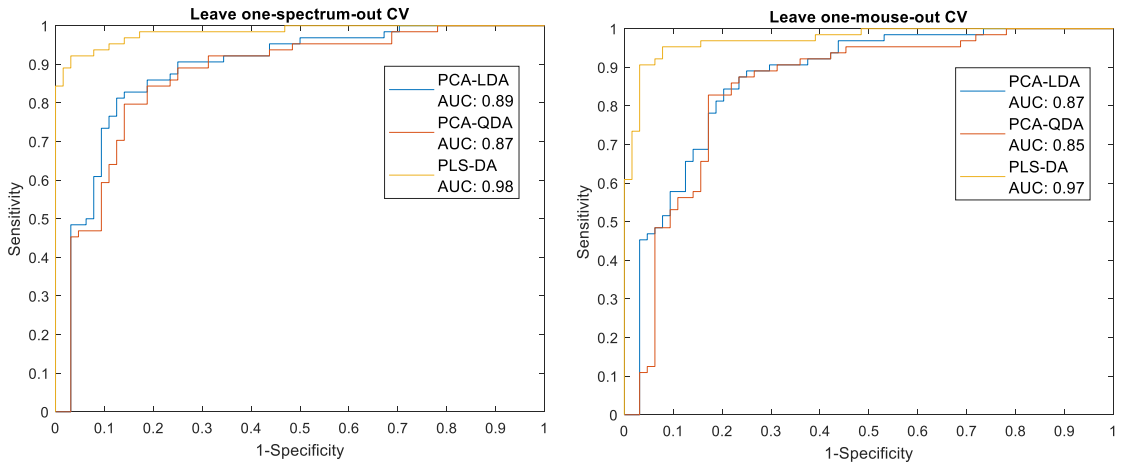


Figure 3.2: ROC curves for the cross validated classification models for the three months old mdx and C57Bl/10 mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

c) One vs. Three months old *mdx* mice

The model built using the spectra of diseased mice at different ages aimed to investigate the ability of the technique to detect the changes in the muscle tissue as the disorder progresses from an early stage to a more established one. The predictive capability of the models using different analysis techniques and CV methods can be seen in table 3.5. All the models yielded better sensitivity than specificity values when distinguishing between one and three months old *mdx* mice. This indicates that the models were able to identify more accurately the three months old *mdx* mice compared to the one month old. Using LOMOCV only the specificity of the PCA-LDA model decreased considerably. The PCA-QDA model demonstrated a very similar performance with the different CV methods, yielding sensitivity, specificity and accuracy values above 80%. LOMOCV slightly increased the performance of the PLS-DA model compared to LOOCV. The PCA-LDA models achieved the best sensitivity values. However, the PLS-DA LOMO cross-validated model had the highest accuracy and more balanced sensitivity and specificity values and was therefore considered to have demonstrated the best performance.

Table 3.5: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods* for the o*ne and three months old mdx mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	98.4%	76.6%	88.7%
Leave-one-mouse-out CV	96.8%	69.0%	82.8%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	90.6%	80%	85.2%
Leave-one-mouse-out CV	89%	80%	84.4%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	95.3%	84.4%	89.8%
Leave-one-mouse-out CV	95.3%	87.5%	91.4%

The ROC curves for the different techniques and CV methods can be seen in figure 3.3. The AUROC 0.90 or above for all the models generated using PCA. Both PLS-DA models yielded an AUC of 0.96.

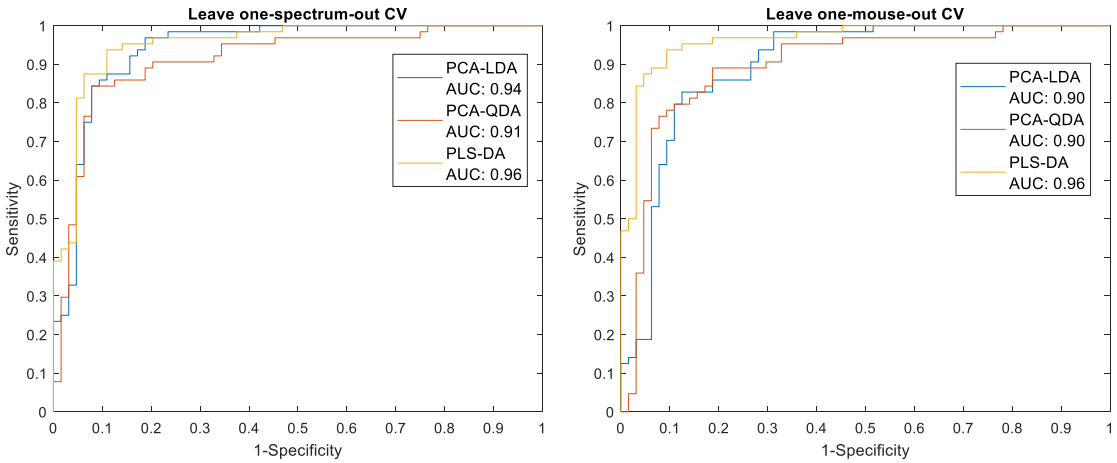


Figure 3.3: ROC curves for the cross validated classification models for the one and three months old mdx mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

d) One month old SOD1^{G93A} vs. C57Bl/6 mice

This is a pre-symptomatic disease stage for the SOD1^{G93A} mice. Therefore, this model investigated the ability of the technique to detect very subtle changes that might take place in the muscle before the appearance of any symptoms. There were no PC scores significantly different between one month old SOD1^{G93A} and C57Bl/6 mice. Hence, LDA and QDA could not be performed. The predictive capability of the PLS-DA model using different CV methods can be seen in table 3.6 and the respective ROC curves can be seen in figure 3.4.

As can be seen from the accuracy values achieved with both CV methods and from the ROC curves the PLS-DA models were not able to discriminate between one month old SOD1^{G93A} and C57Bl/6 mice. Using LOMOCV the sensitivity, specificity and accuracy values dropped below 50% indicating that the technique was not able to detect differences in the muscle that would allow it to correctly classify spectra in the different groups.

Table 3.6: Two group PLS-DA classification model performance evaluated using different CV methods* for the *one month old SOD1^{G93A} and C57Bl/6 mice.

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	53.3%	56.1%	54.7%
Leave-one-mouse-out CV	41.7%	47.1%	44.4%

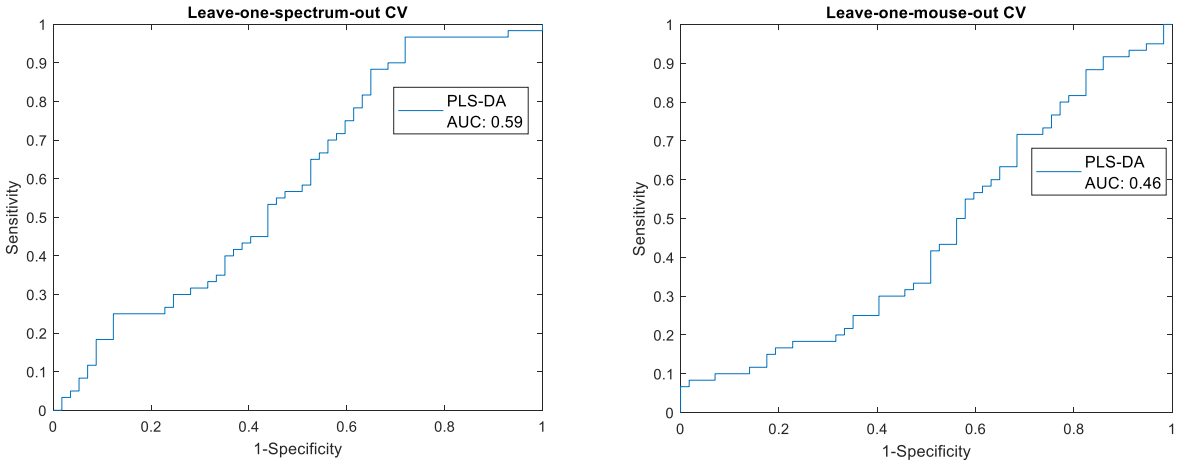


Figure 3.4: ROC curves for the cross validated classification PLS-DA model for the one month old SOD1^{G93A} and C57Bl/6 mice*. ROC curves for leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) is also displayed.*

e) Three months old SOD1^{G93A} vs. C57Bl/6 mice

This is an established disease stage for the SOD1^{G93A} mice. Thus, the model generated using the spectra from three months old SOD1^{G93A} and C57Bl/6 mice aimed to explore the ability of the technique to detect muscle pathology resulting from a neurogenic disorder at an established disease stage. The predictive capability of the models using different analysis and CV methods for the three months old SOD1^{G93A} mice can be seen in table 3.7 The ROC curves, comparing the three different approaches for LOOCV and LOMOCV can be seen in figure 3.5.

Table 3.7: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods for the three months old SOD1^{G93A} and C57Bl/6 mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	83.9%	85%	84.4%
Leave-one-mouse-out CV	82.3%	75%	78.7%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	79%	83.3%	81.1%
Leave-one-mouse-out CV	79%	75%	77%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	77.4%	70%	73.8%
Leave-one-mouse-out CV	79.3%	70%	74.6

Using LOMOCV the ability to correctly classify spectra decreased for PCA-LDA and QDA models. For both models this was primarily due to a decrease in specificity. For the PLS-DA model the specificity remained unchanged and the sensitivity increased when LOMOCV was implemented, leading to an increased accuracy. The PCA-LDA model achieved the best performance, with a sensitivity above 80% for both cross-validation methods. The area under the ROC curves also decreased for LOMOCV for the PCA-LDA and QDA models and increased for the PLS-DA model. AUC values remained above 0.8 for all models and CV approaches. PCA-LDA demonstrated the demonstrated an AUC of 0.86 for the leave-one-mouse-out CV and had the highest AUC for both CV methods.

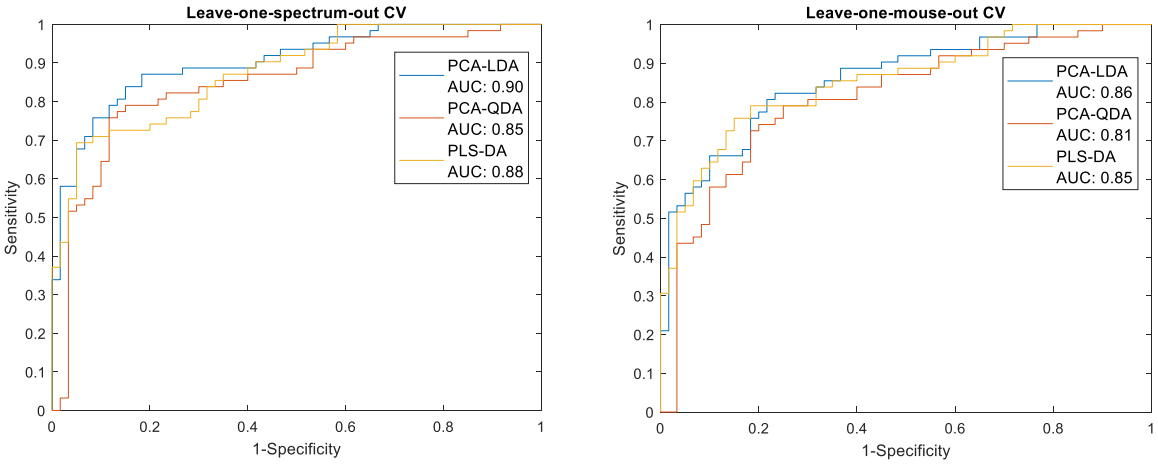


Figure 3.5: ROC curves for the cross validated classification models for the three months old SOD1^{G93A} and C57Bl/6 mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

f) One vs. Three months old SOD1^{G93A} mice

The model built using the spectra of SOD1^{G93A} mice at different ages aimed to investigate the ability of the technique to monitor disease progression in terms of detecting the changes in the muscle tissue as the disorder progresses from a pre-symptomatic stage to a more established one. The predictive capability of the models using different analysis techniques and CV methods can be seen in table 3.8. The sensitivity and specificity values were very balanced for all the models and CV techniques. Using LOMOCV there was a very small decrease in the ability of all the models to correctly classify spectra. Specificity, sensitivity and accuracy above 85% were achieved using PCA-LDA with both CV techniques. The PCA-QDA model demonstrated a very similar performance when using LOSOCV and slightly decreased one when using LOMOCV, still yielding sensitivity, specificity and accuracy values above 80% though. The PLS-DA model achieved the best performance with all accuracy values being above 95%.

Table 3.8: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods for the one and three months old SOD1^{G93A} mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	87.1%	88.3%	87.17%
Leave-one-mouse-out CV	85.5%	86.7%	86.1%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	85.5%	86.7%	86.1%
Leave-one-mouse-out CV	82.3%	85%	83.6%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	95.2%	98.3%	96.7%
Leave-one-mouse-out CV	93.5%	98.3%	95.9%

The ROC curves for the different techniques and CV methods can be seen in figure 3.6. The AUROC was 0.90 or above for all the models. Both PLS-DA models yielded an AUC of 0.99.

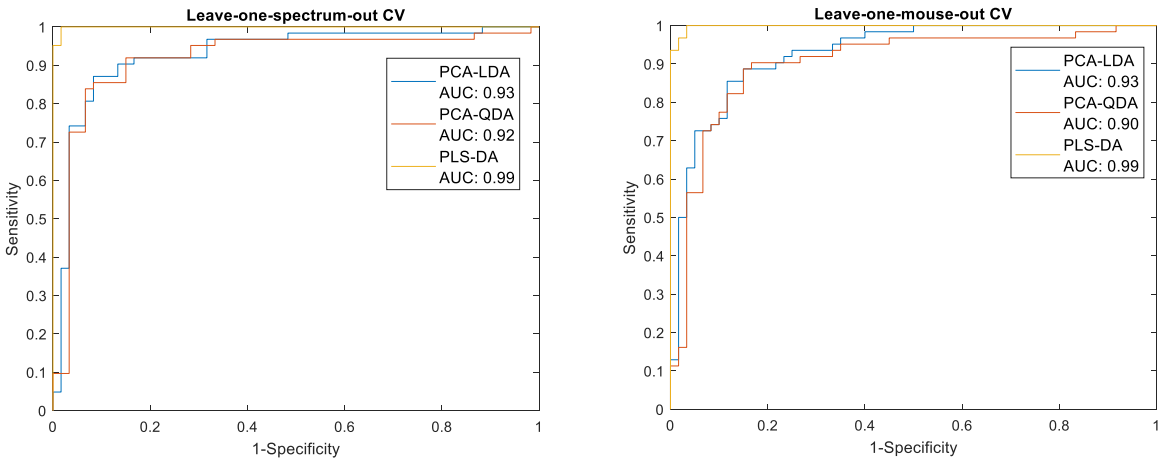


Figure 3.6: ROC curves for the cross validated classification models for the one and three months old SOD1^{G93A} mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

g) One month old *mdx* vs. SOD1^{G93A} mice

This is a very early disease stage for both *mdx and* SOD1^{G93A} mice. Thus, the model generated using the spectra from one month old diseased mice aimed to explore the ability of the technique to detect muscle pathology for disorders of different origin at an early (*mdx*) and pre-symptomatic (SOD1^{G93A}) disease stage. The predictive capability of the models using different analysis and CV methods for the one month old *mdx and* SOD1^{G93A} mice can be seen in table 3.9. The ROC curves, comparing the three different approaches for LOSOCV and LOMOCV can be seen in figure 3.7.

Table 3.9: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods for the one month old mdx and SOD1^{G93A} mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	92.2%	96.7%	94.5%
Leave-one-mouse-out CV	90.6%	98.3%	94.4%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	92.2%	96.7%	94.4%
Leave-one-mouse-out CV	89.1%	95%	91.9%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	93.8%	98.3%	95.7%
Leave-one-mouse-out CV	93.8%	98.3%	95.7%

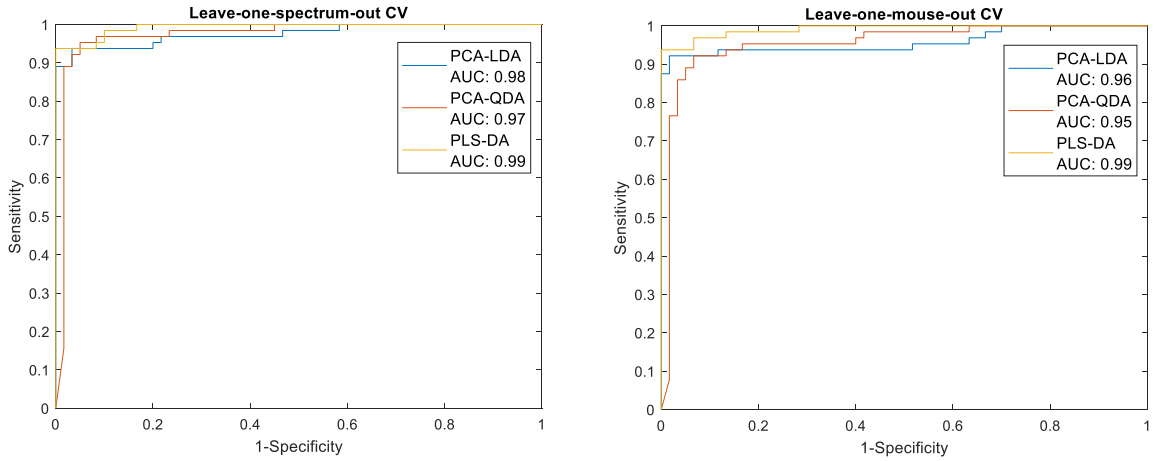


Figure 3.7: ROC curves for the cross validated classification models for the one month old SOD1^{G93A} and mdx mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

Using LOMOCV the ability to correctly classify spectra slightly decreased for PCA-LDA and QDA models whereas the performance of the PLS-DA model remained unchanged. Except for the LOMO cross-validated PCA-QDA sensitivity all other performance indices were above 90% for all models and CV methods. The area under the ROC curves also decreased slightly for PCA-LDA and QDA models for LOMOCV.

For both CV approaches the PLS-DA model performed better than the PCA related models. The LOMO cross-validated PLS-DA model demonstrated a very good performance for distinguishing between SOD1^{G93A} and *mdx* mice at an early disease stage achieving an accuracy of 95.7% and an AUC of 0.99.

h) Three months old *mdx* vs. SOD1^{G93A} mice

This is a more established disease stage for both mouse models. Thus, the model generated using the spectra from three months diseased mice aimed to explore the ability of the technique to distinguish between more established muscle pathology of the two different neuromuscular disorders. The predictive capability of the models generated using different techniques and CV methods for the three months old *mdx* mice can be seen in table 3.10.

The different cross-validation approaches did not have a big impact on the PCA related models, with only the sensitivity values of both models slightly decreasing when using LOMOCV. Both models achieved accuracy above 80% for both CV methods. Both models (and for both CV methods) demonstrated a better ability to correctly classify *mdx* mice (higher sensitivity than specificity values). Using LOMOCV both sensitivity and specificity of the PLS-DA model decreased. However, the PLS-DA demonstrated the most balanced sensitivity and specificity values for both CV methods with higher (for LOOCV) and very similar (LOMOCV) accuracy values and so it was considered to have achieved the best performance.

Table 3.10: Two group PCA-LDA, PCA-QDA and PLS-DA classification models performance evaluated using different CV methods for the three months old mdx and SOD1^{G93A} mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	95.4%	74.2%	84.9%
Leave-one-mouse-out CV	93.8%	74.2%	84.1%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	90.6%	77.4%	84.1%
Leave-one-mouse-out CV	87.5%	77.4%	82.5%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	93.8%	85.5%	89.7%
Leave-one-mouse-out CV	87.5%	80.0%	83.4%

The ROC curves for the different techniques and CV methods can be seen in figure 3.8. The AUROC was equal or above 0.89 for all the models. The PLS-DA model generated the ROC curve with the highest AUC for both CV methods, yielding an AUC of 0.96 and 0.93 for LOSOCV and LOMOCV respectively.

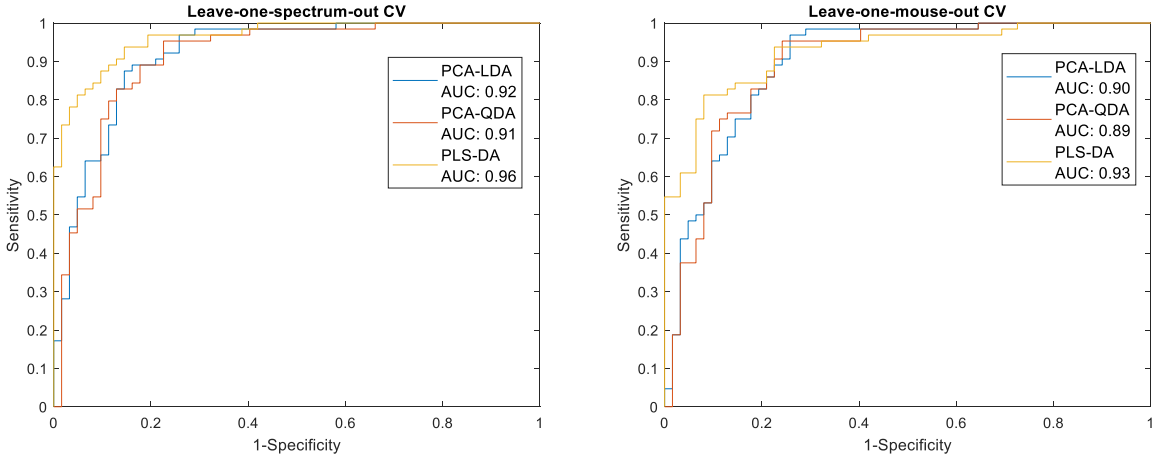


Figure 3.8: ROC curves for the cross validated classification models for the three months old SOD1^{G93A} and mdx mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

3.1.2.1.2 Repeated Leave-Some-Mice-out Cross-Validation (RLSMOCV)

Repeated cross validation, has been employed in order to validate the classification performance of the models more robustly and to evaluate how much it can vary for the same data set (Lloyd *et al.*, 2012, 2013). In this case cross validation has been performed by randomly selecting two mice from each mouse group (four in total) to 'leave out' and use in the test set.

a) *mdx* mice

Table 3.11 displays the classification performance of the cross-validated PCA-LDA, PCA-QDA and PLS-DA models for the one month old *mdx* mice. The mean sensitivity, specificity and accuracy values (+/-) one standard deviation and the coefficients of variability (CoV) are shown. The PLS-DA model achieved the best performance with an accuracy of 76.3%. The coefficients of variability were similar for the different models with the sensitivity being the most variable parameter in all three. For all three approaches RLSMOCV yielded a similar performance with LOMOCV.

Table 3.11: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for one month old mdx and C57Bl/10 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

One month old mdx/ C57B/10 mice	***Sensitivity*** ***(+/- Standard Deviation, CoV)***	***Specificity*** ***(+/- Standard Deviation, CoV)***	***Accuracy*** ***(+/- Standard Deviation, CoV)***
PCA-LDA	71.3%	65.7%	68.6%
	(+/-3.1%, CoV: 4.3%)	(+/-2.6%, CoV:4.0%)	(+/-2.0%, CoV:3.0%)
PCA-QDA	76.7%	63.1%	70.4%
	(+/-3.1%, CoV:4.0%)	(+/-1.8%, CoV:2.9%)	(+/-1.9%, CoV:2.7%)
PLS-DA	80.5%	71.6%	76.3%
	(+/-4.0%, CoV:5%)	(+/-3.4%, CoV:4.7%)	(+/-3.1%, CoV:4.1%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model can be seen in figure 3.9. The mean AUC value for each model (+/-) one standard deviation is also shown. The PCA-LDA and QDA ROC curves had mean AUCs of 0.76 and 0.74 respectively whereas the PLS-DA ROC curve had a mean AUC of 0.85.

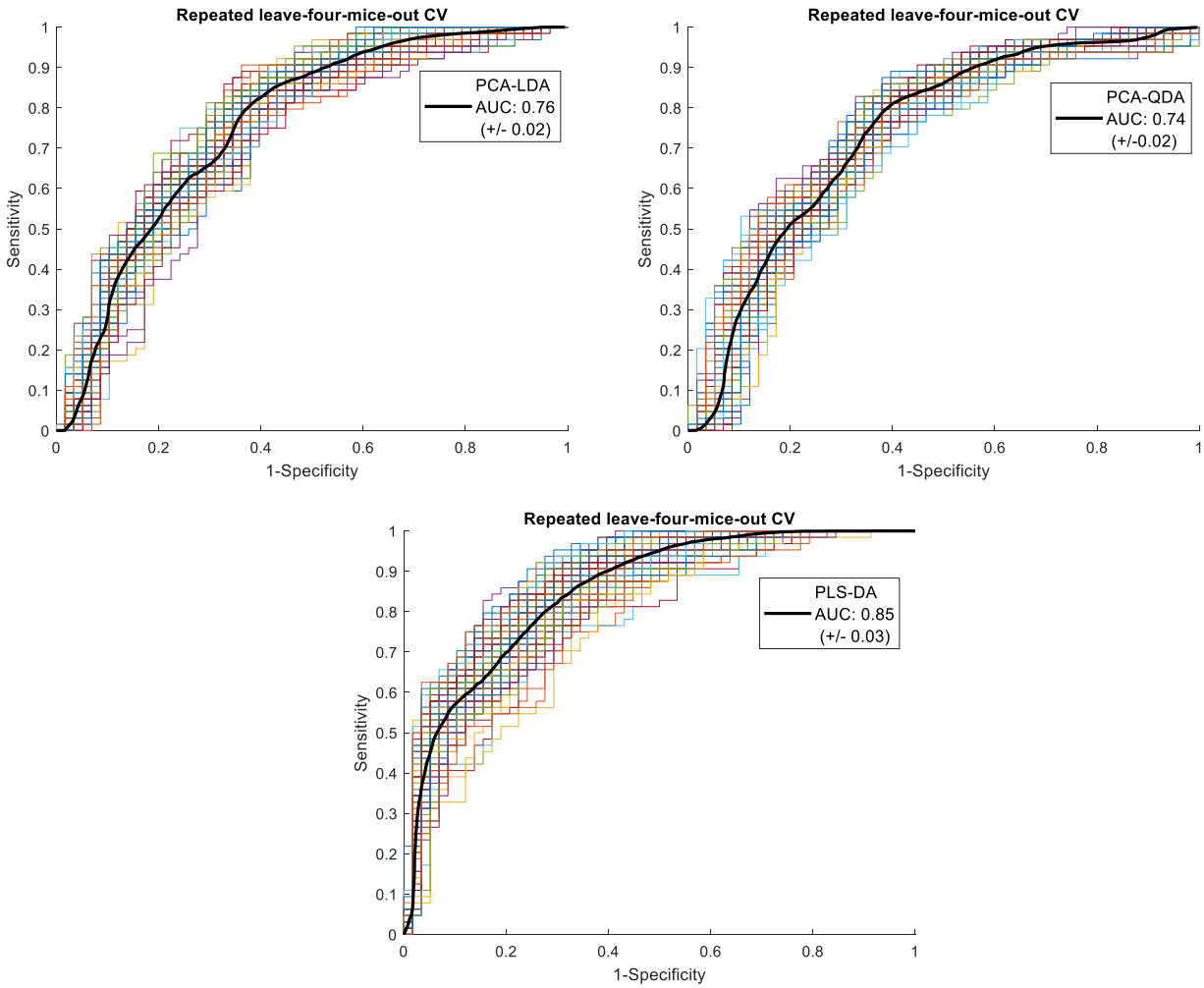


Figure 3.9: ROC curves generated from repeated cross-validation of one month old mdx and C57Bl/10 mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

Table 3.12 displays the classification performance of the cross-validated PCA-LDA, PCA-QDA and PLS-DA models for the three months old *mdx* mice. The performance of all models is improved compared to the performance of the models for the one month old mice and the coefficients of variability decreased for all the classification indices and models. The PCA-LDA and QDA models attained similar accuracy values, with the LDA model having a better sensitivity. The PLS-DA model achieved the best performance with a sensitivity of 94.2%, a specificity of 88.7% and an accuracy of 91.3%. The coefficients of variability were similar for the different models. The specificity was the most variable parameter in all three models.

Table 3.12: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three months old mdx and C57Bl/10 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

Three months old mdx / C57B/10 mice	***Sensitivity*** ***(+/- Standard Deviation, CoV)***	***Specificity*** ***(+/- Standard Deviation, CoV)***	***Accuracy*** ***(+/- Standard Deviation, CoV)***
PCA-LDA	91.6%	76.4%	84.1%
	(+/-2.2, CoV: 2.4%)	(+/-3.0, CoV: 3.8%)	(+/-1.7, CoV: 2.0%)
PCA-QDA	88.6%	78.4%	83.5%
	(+/-1.8, CoV: 2.0%)	(+/-2.0, CoV: 2.6%)	(+/-1.2, CoV: 1.4%)
PLS-DA	94.2%	88.7%	91.3%
	(+/-2.3, CoV: 2.4%)	(+/-2.9, CoV: 3.3%)	(+/-1.8, CoV: 2.0%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model for the three months old *mdx* mice can be seen in figure 3.10. The PCA-LDA and QDA ROC curves had a mean AUC of 0.90 and 0.91 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.96.

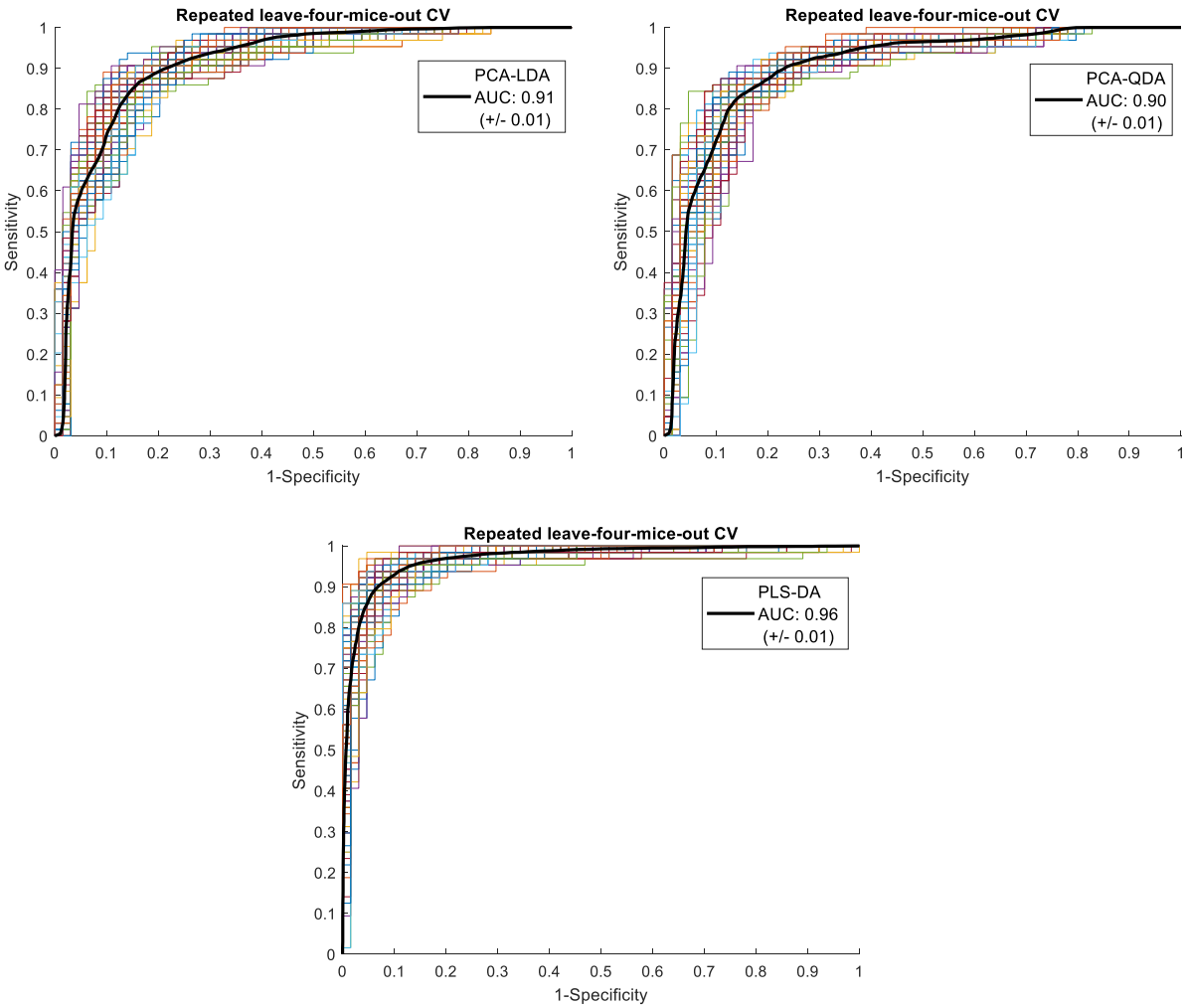


Figure 3.10: ROC curves generated from repeated cross-validation of three months old mdx/C57Bl/10 mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

The classification performance of the RLSMOCV PCA-LDA, PCA-QDA and PLS-DA models for distinguishing between one and three months old *mdx* mice can be seen in table 3.13. The ROC curves generated in each of the repetitions and the mean ROC curve for each model can be seen in figure 3.11. The PCA-LDA and PLS-DA models attained similar sensitivities, both around 95.5%, but the PLS-DA model achieved much better specificity, having thus, a better overall performance (Accuracy 88% and mean AUC 0.96). The coefficients of variability were similar for the different models, remaining below 5% for all the indices. The PCA-QDA model had the most balanced sensitivity and specificity values but the lowest accuracy and mean AUC values.

Table 3.13: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for one and three months old mdx mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

One/Three months old mdx mice	***Sensitivity*** ***(+/- Standard Deviation, CoV)***	***Specificity*** ***(+/- Standard Deviation, CoV)***	***Accuracy*** ***(+/- Standard Deviation, CoV)***
PCA-LDA	95.6% (+/-1.8, CoV: 1.9%)	71.6% (+/-3.4, CoV: 4.7%)	83.6% (+/-2.0, CoV: 2.4%)
PCA-QDA	89.1% (+/-1.5, CoV: 1.7%)	77.6% (+/-2.5, CoV: 3.2%)	83.4% (+/-1.5, CoV: 1.8%)
PLS-DA	95.4 % (+/-1.6, CoV: 1.9%)	80.6% (+/-3.2, CoV: 4.0%)	88% (+/-2.0, CoV: 2.3%)

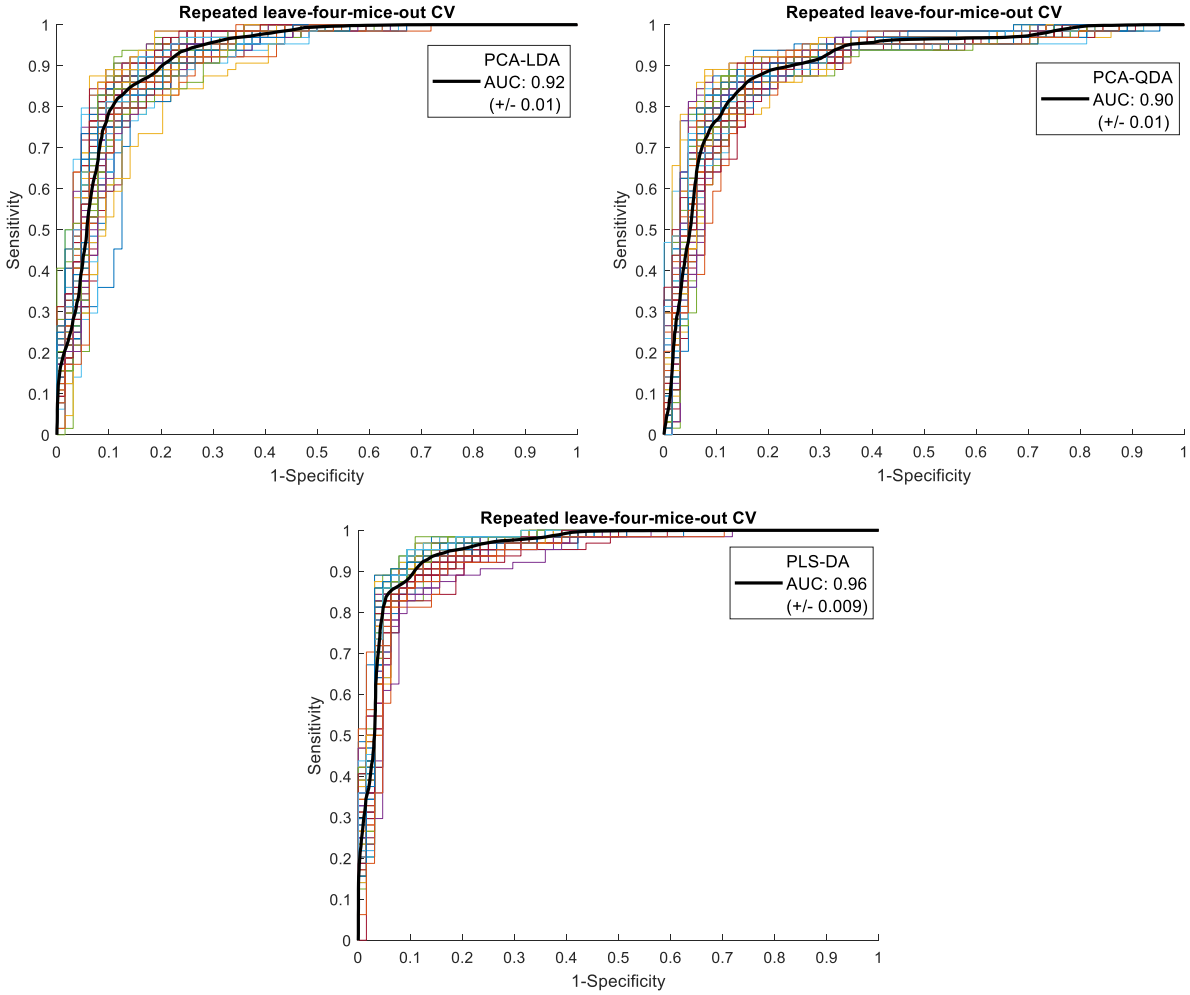


Figure 3.11: ROC curves generated from repeated cross-validation of one and three months old mdx mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

An additional three group model was built using the spectra of the one and three months old *mdx* mice as well as the spectra of the three months old C57Bl/10 mice to investigate the ability of the technique to simultaneously separate diseased from healthy tissue and the different disease stages. This model could be more useful in a clinical setting, where the diseased group will be more heterogeneous since it will most probably include patients at different disease stages. The model was validated using RLSMOCV and the results for the different models are displayed below. Two mice from each group were left out in each CV step. Thus, each test set consisted of six mice in total. All models showed high sensitivity and specificity for the three months old *mdx* mice, with all the values being above 85%. The PCA-LDA model achieved the best performance for this group. However, it demonstrated the lowest sensitivity for the healthy mice. In general, all the models achieved good accuracy for one month old *mdx* mice and three months old C57Bl/10 mice though with much higher specificities than sensitivities.

Table 3.14: Repeated cross-validated PCA-LDA, PCA-QDA and PLS-DA three group model classification performance for mdx and C57Bl/10 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
	(+/- Standard Deviation, CoV)	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
3 months old mdx	95.6%	87.4%	90.2%
	(+/-1.6, CoV: 2.4%)	(+/-1.3, CoV: 3.8%)	(+/-1.1, CoV: 2.0%)
1 month old mdx	70.2%	84.0%	79.4%
	(+/-3.2, CoV: 2.0%)	(+/-1.5, CoV: 2.6%)	(+/-1.5, CoV: 1.4%)
3 months old C57Bl/10	58.4%	90.6%	79.9%
	(+/-3, CoV: 2.4%)	(+/-1.9, CoV: 3.3%)	(+/-1.6, CoV: 2.0%)

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
	(+/- Standard Deviation, CoV)	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
3 months old mdx	86.3%	89.4%	88.4%
	(+/-3.0, CoV: 3.5%)	(+/-1.4, CoV: 1.5%)	(+/-1.2, CoV: 1.4%)
1 month old mdx	67.3%	87.2%	80.6%
	(+/-3.3, CoV: 4.9%)	(+/-1.5, CoV: 1.7%)	(+/-1.6, CoV: 2.0%)
3 months old C57Bl/10	69.9%	85.2%	80.1%
	(+/-2.9, CoV: 2.9%)	(+/-2.0, CoV: 2.3%)	(+/-1.8, CoV: 2.2%)

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
	(+/- Standard Deviation, CoV)	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
3 months old mdx	92.7%	88.3%	89.8%
	(+/-1.9, CoV: 2.0%)	(+/-2.0, CoV: 2.3%)	(+/-1.4, CoV: 1.6%)
1 month old mdx	66.7%	86.6%	80.0%
	(+/-3.4, CoV: 5.1%)	(+/-2.0, CoV: 2.2%)	(+/-1.5, CoV: 1.9%)

3 months old C57Bl/10	68% (+/-5.6, CoV: 8.2%)	88.9% (+/-1.9, CoV: 2.1%)	81.9% (+/-2.1, CoV: 2.6%)

b) SOD1^{G93A} mice

The classification performance of the RLSMOCV PCA-LDA, PCA-QDA and PLS-DA models for distinguishing between three months old SOD1^{G93A} and C57Bl/6 mice can be seen in table 3.15. The PCA-LDA model demonstrated the best performance with sensitivity, specificity and accuracy values of 82.5%, 76% and 79.3% respectively. The coefficients of variability were similar for the different models, with the specificity being the most variable index.

Table 3.15: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three months old SOD1^{G93A} and C57Bl/6 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

Three months old SOD1^{G93A}/C57Bl/6 mice	***Sensitivity (+/- Standard Deviation, CoV)***	***Specificity (+/- Standard Deviation, CoV)***	***Accuracy (+/- Standard Deviation, CoV)***
PCA-LDA	82.5% (+/-2.0, CoV: 2.4%)	76% (+/-3.1, CoV: 4.0%)	79.3% (+/-2.3, CoV: 2.9%)
PCA-QDA	78.8% (+/-2.8, CoV: 3.6%)	72.6% (+/-3.4, CoV: 4.6%)	75.8% (+/-2.2, CoV: 2.9%)
PLS-DA	80.0 % (+/-2.7, CoV: 3.4%)	74.8% (+/-4.2, CoV: 5.6%)	77.4% (+/-2.6, CoV: 3.4%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model for the three months old SOD1^{G93A} mice can be seen in figure 3.12. The PCA-LDA and QDA ROC curves had a mean AUC of 0.86 and 0.81 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.85.

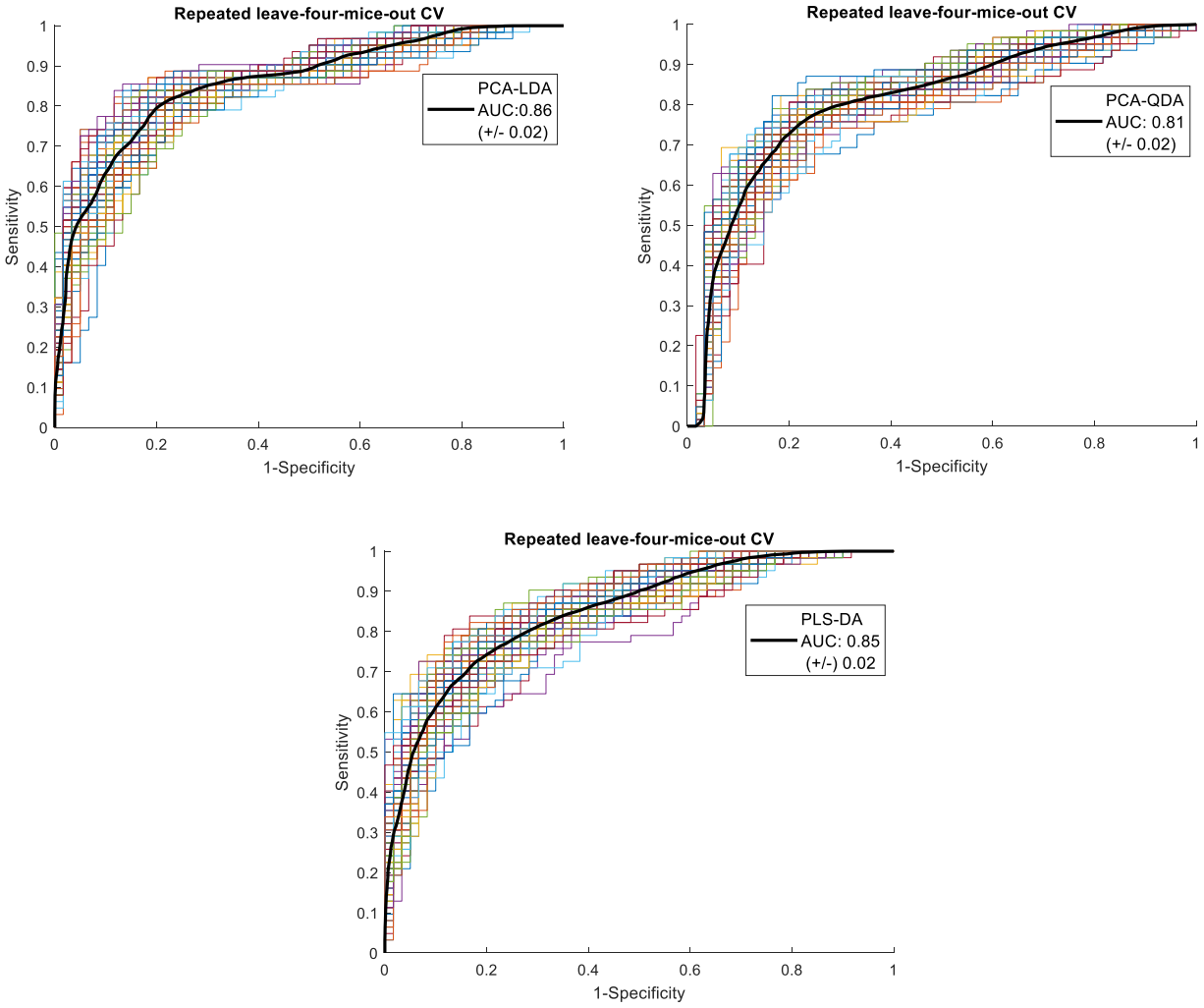


Figure 3.12: ROC curves generated from repeated cross-validation of three months old SOD1^{G93A} and C57Bl/6 mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

The classification performance of the RLSMOCV PCA-LDA, PCA-QDA and PLS-DA models for distinguishing between one and three months old SOD1^{G93A} can be seen in Table 3.16. The ROC curves generated in each of the repetitions and the mean ROC curve for each model can be seen in figure 3.13. The PCA-LDA and PCA-QDA models had a similar performance with both models achieving an accuracy value around 85% and mean AUC around 0.90. The PLS-DA model achieved much better sensitivity and specificity, both around 95%, and a mean AUC of 0.99. The coefficients of variability remained below 5% for all the indices and models.

Table 3.16: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for one and three months old SOD1^{G93A} mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

	Sensitivity	***Specificity***	***Accuracy***
One/Three months old SOD1^{G93A} mice	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
PCA-LDA	85.5%	86.6%	85.9%
	(+/-3.8, CoV: 4.4%)	(+/-2.5, CoV: 2.9%)	(+/-2.3, CoV: 2.7%)
PCA-QDA	83.9%	86.6%	85.2%
	(+/-3.5, CoV: 4.1%)	(+/-2.7, CoV: 3.1%)	(+/-2.2, CoV: 2.6%)
PLS-DA	94.3%	96.9%	95.6%
	(+/-1.8, CoV: 1.9%)	(+/-2.0, CoV: 2.1%)	(+/-1.5, CoV: 2.0%)

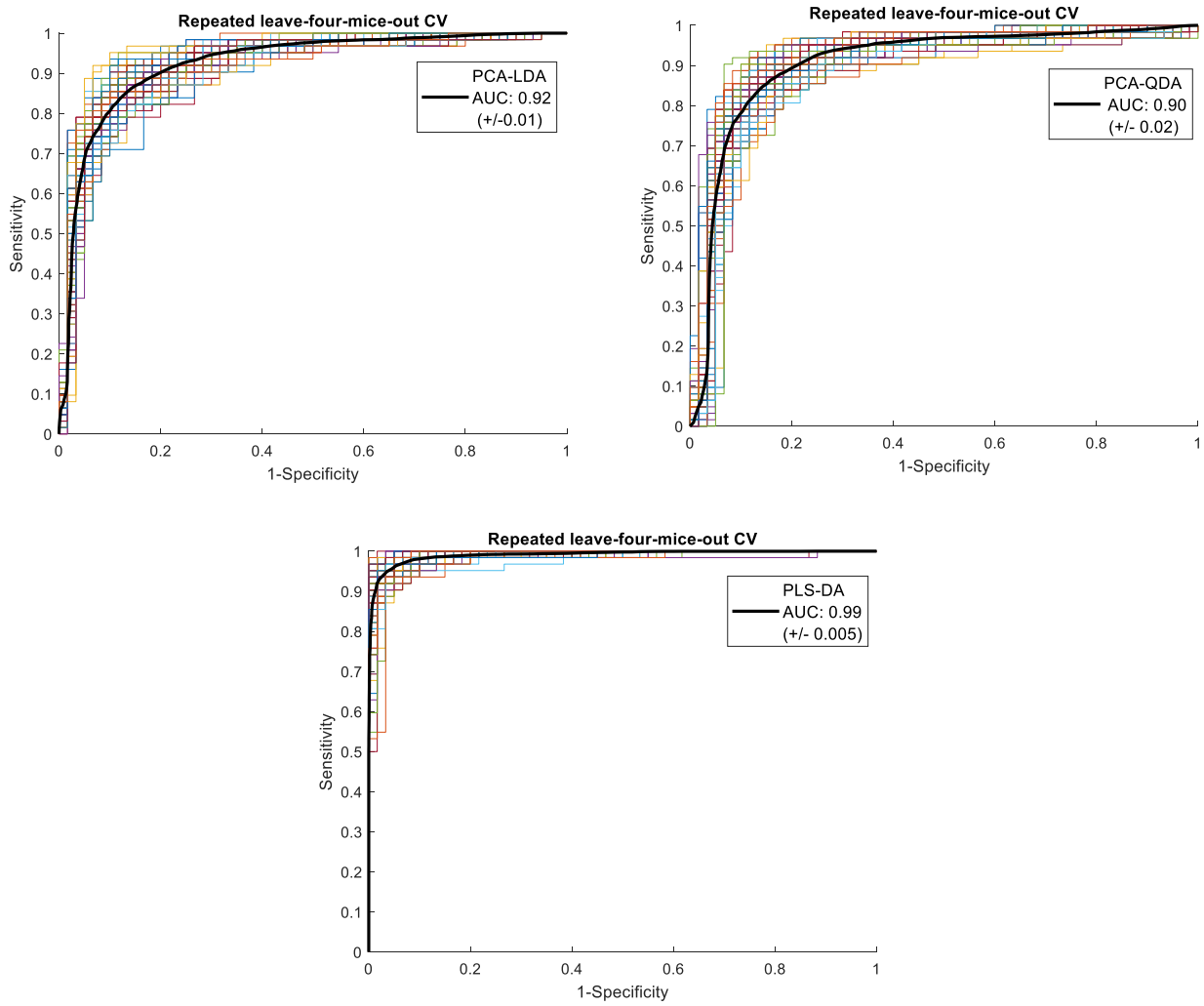


Figure 3.13: ROC curves generated from repeated cross-validation of one and three months old SOD1^{G93A} mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

An additional three group model was built using the spectra of the one and three months old SOD1^{G93A} mice as well as the spectra of the three months old C57Bl/6 mice to investigate the ability of the technique to simultaneously separate diseased from healthy tissue and the different disease stages. The model was validated using RLSMOCV and the results for the different models are displayed below. Two mice from each group were left out in each CV step. Thus, each test set consisted of six mice in total.

Table 3.17: Repeated CV PCA-LDA three group model classification performance for mdx and C57Bl/10 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
	(+/- Standard Deviation, CoV)	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
3 months old SOD^{G93A}	78.7%	88.6%	85.2%
	(+/-1.8, CoV: 2.3%)	(+/-1.8, CoV: 2.0%)	(+/-1.4, CoV: 1.6%)
1 month old SOD^{G93A}	94.1%	95.7%	95.2%
	(+/-1.4, CoV: 1.5%)	(+/-1, CoV: 1.0%)	(+/-0.8, CoV: 0.8%)
3 months old C57Bl/6	74%	88.9%	84.0%
	(+/-3.7, CoV: 5.0%)	(+/-1.9, CoV: 2.4%)	(+/-1.5, CoV: 1.2%)

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
	(+/- Standard Deviation, CoV)	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
3 months old SOD^{G93A}	76.5%	87.4%	83.8%
	(+/-2.8, CoV: 3.6%)	(+/-2.3, CoV: 2.6%)	(+/-2.0, CoV: 2.4%)
1 month old SOD^{G93A}	89.3%	94.2%	92.6%
	(+/-1.5, CoV: 1.7%)	(+/-1.1, CoV: 1.2%)	(+/-0.8, CoV: 0.9%)
3 months old C57Bl/6	70.5%	86.5%	81.2%
	(+/-4.5, CoV: 6.4%)	(+/-1.6, CoV: 1.9%)	(+/-1.9, CoV: 2.3%)

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
	(+/- Standard Deviation, CoV)	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
3 months old SOD^{G93A}	80.7%	89.2%	86.3%
	(+/-2.6, CoV: 3.2%)	(+/-1.3, CoV: 1.4%)	(+/-1.3, CoV: 1.6%)
1 month old SOD^{G93A}	95.0%	98.1%	97.1%
	(+/-3.4, CoV: 3.5%)	(+/-2.0, CoV: 2.1%)	(+/-0.9, CoV: 0.9%)
3 months old C57Bl/6	80.5%	90.7%	87.3%

	(+/-2.8, CoV: 3.5%)	(+/-1.5, CoV: 1.7%)	(+/-1.4, CoV: 1.6%)

c) *mdx* vs. SOD1^{G93A} mice

Table 3.18 displays the classification performance of the cross-validated PCA-LDA, PCA-QDA and PLS-DA models for the one month old *mdx and* SOD1^{G93A} mice. The ROC curves generated in each of the repetitions and the mean ROC curve for each model can be seen in figure 3.14. The PCA-LDA and QDA models attained similar accuracy values, with the LDA model having slightly better sensitivity and specificity. The PLS-DA model demonstrated the best performance mainly due to higher sensitivity. The coefficients of variability were similar for the different models and remained below 2.5% for all the indices. The PCA-LDA and QDA ROC curves had a mean AUC of 0.97 and 0.96 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.98.

Table 3.18: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for one month old mdx and SOD1^{G93A} mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

	Sensitivity	***Specificity***	***Accuracy***
One month old mdx/SOD1^{G93A} mice	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
PCA-LDA	89.9%	97.1%	93.4%
	(+/-1.8, CoV: 2.0%)	(+/-1.9, CoV: 2.0%)	(+/-1.4, CoV: 1.5%)
PCA-QDA	89.3%	95.3%	92.2%
	(+/-2.1, CoV: 2.4%)	(+/-1.8, CoV: 1.9%)	(+/-1.5, CoV: 1.4%)
PLS-DA	93.4%	97.3%	95.3%
	(+/-0.8, CoV: 0.9%)	(+/-1.3, CoV: 1.3%)	(+/-1, CoV: 1.0%)

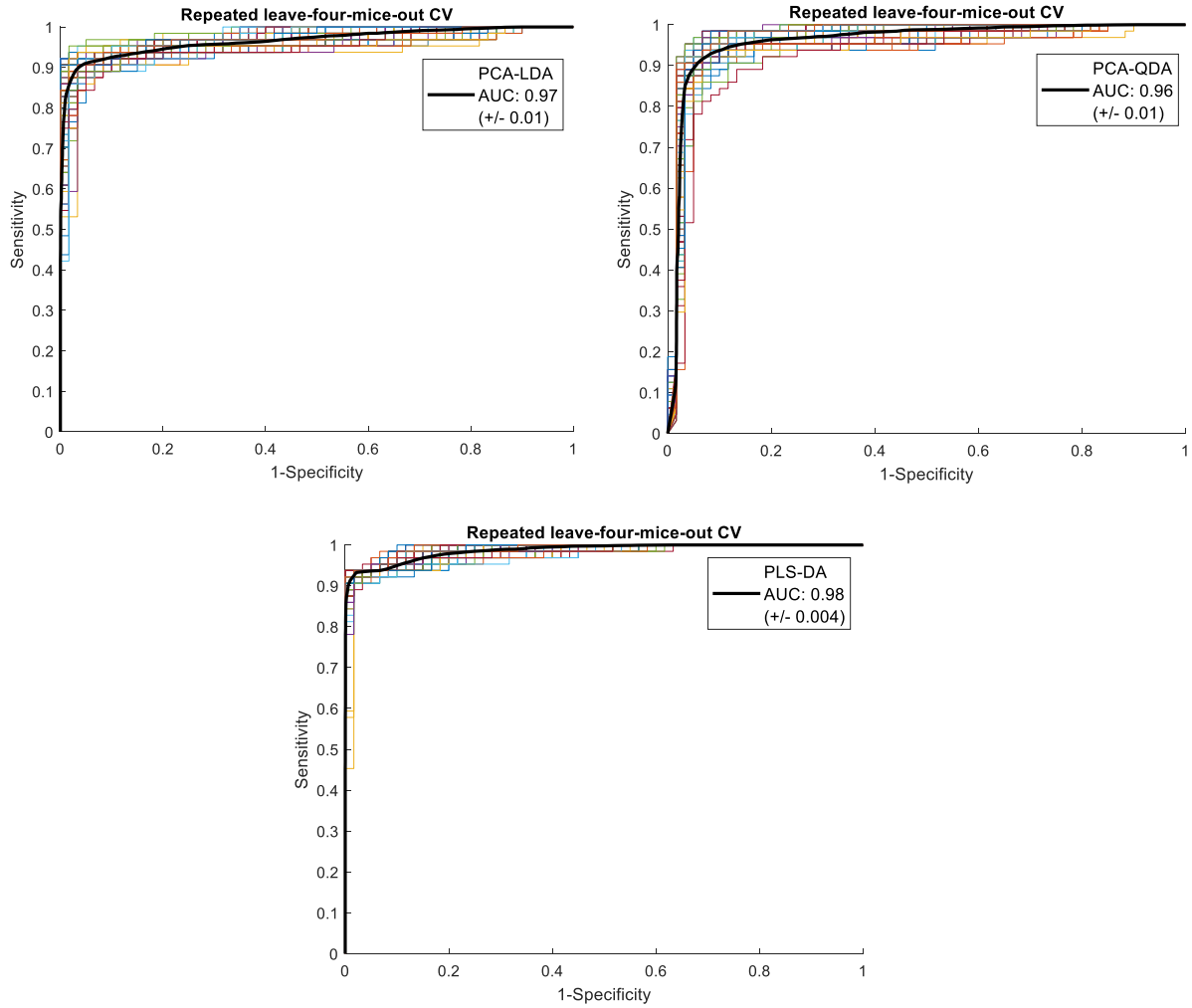


Figure 3.14: ROC curves generated from repeated cross-validation of one months old SOD1^{G93A} and mdx mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

The classification performance of the RLSMOCV PCA-LDA, PCA-QDA and PLS-DA models for distinguishing between three months old *mdx* and SOD1^{G93A} mice can be seen in Table 3.19. The ROC curves generated in each of the repetitions and the mean ROC curve for each model can be seen in figure 3.15. The different models attained similar accuracies, all around 83%. The PCA-LDA model yielded the highest sensitivity but the lowest specificity. Hence, the PLS-DA model was considered to have achieved the best performance due to the highest accuracy and mean AUC and more balanced sensitivity and specificity values. The PCA-LDA and QDA ROC curves had a mean AUC of 0.89 whereas the PLS-DA curves demonstrated a mean AUC of 0.92.

Table 3.19: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three month old mdx and SOD1^{G93A} mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

	Sensitivity	***Specificity***	***Accuracy***
Three months old mdx/SOD1^{G93A} mice	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***	***(+/- Standard Deviation, CoV)***
PCA-LDA	93.5%	73.3%	83.5%
	(+/-2.6, CoV: 2.8%)	(+/-1.7, CoV: 2.3%)	(+/-1.6, CoV: 1.9%)
PCA-QDA	87.4%	78.5%	83.0%
	(+/-2.8, CoV: 3.2%)	(+/-2.1, CoV: 2.7%)	(+/-1.6, CoV: 1.4%))
PLS-DA	88.2%	78.2%	83.2%
	(+/-3.8, CoV: 4.3%)	(+/-4.8, CoV: 6.1%)	(+/-3.0, CoV: 3.6%))

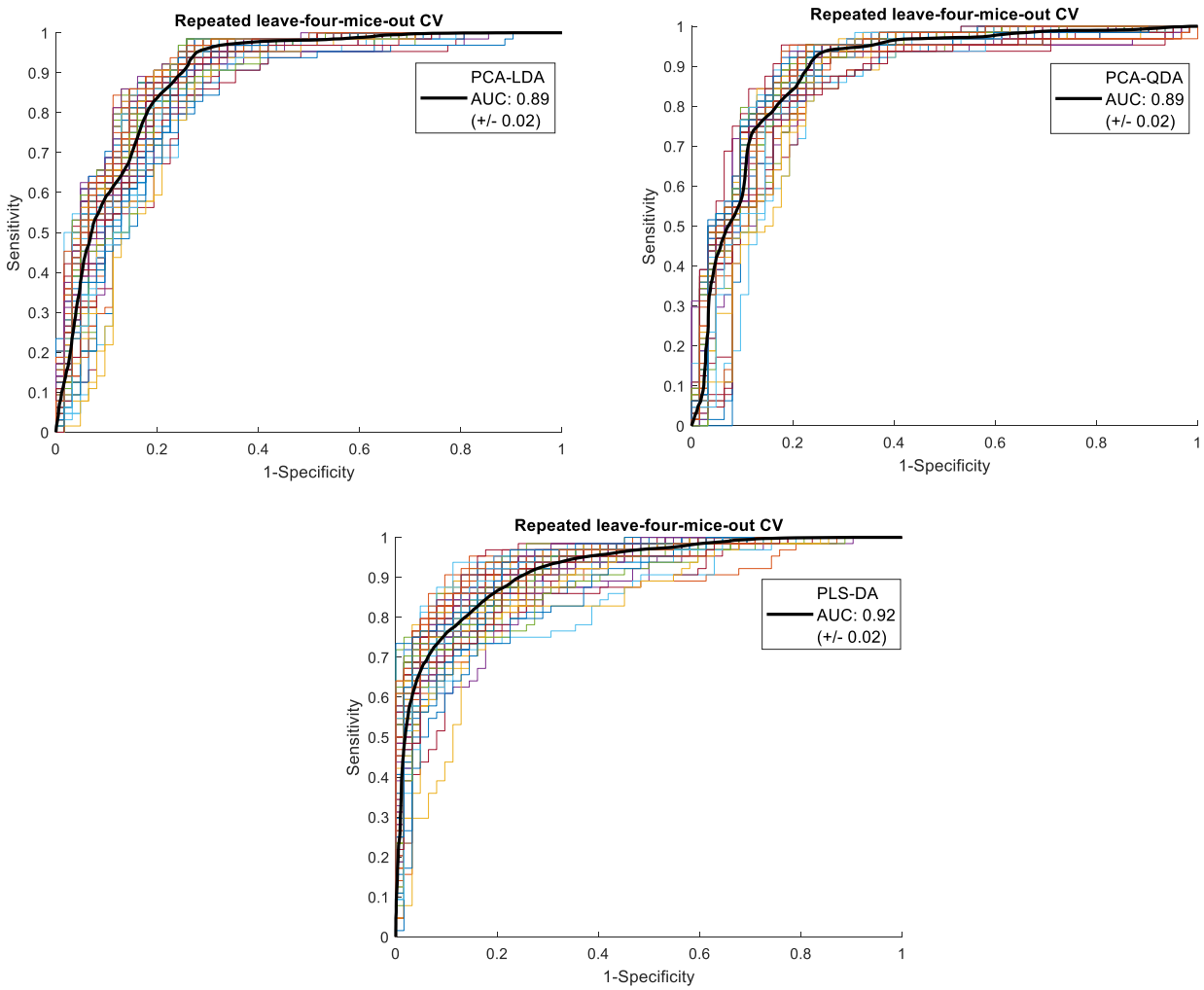


Figure 3.15: ROC curves generated from repeated cross-validation of three months old SOD1^{G93A} and mdx mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

3.1.2.2 Understanding the Basis of Classification

3.1.2.2.1 Mean and Difference Spectra

The mean and difference spectra of the different mice groups were plotted to visually examine the most prominent peaks and to identify the major differences between the groups of each model. In figures 3.16 and 3.17 the mean spectra of each group (+/-) standard deviation are shown.

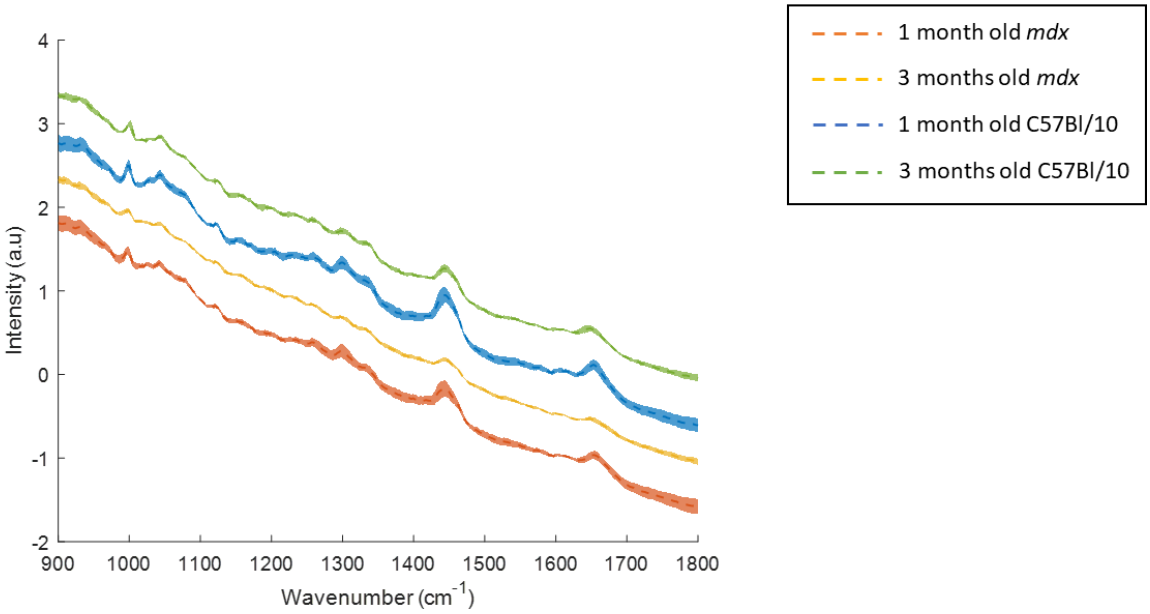


Figure 3.16: Mean Raman spectra for mdx and C57Bl/10 mice groups.

The mean spectra for mdx and control mice of one and three months of age are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

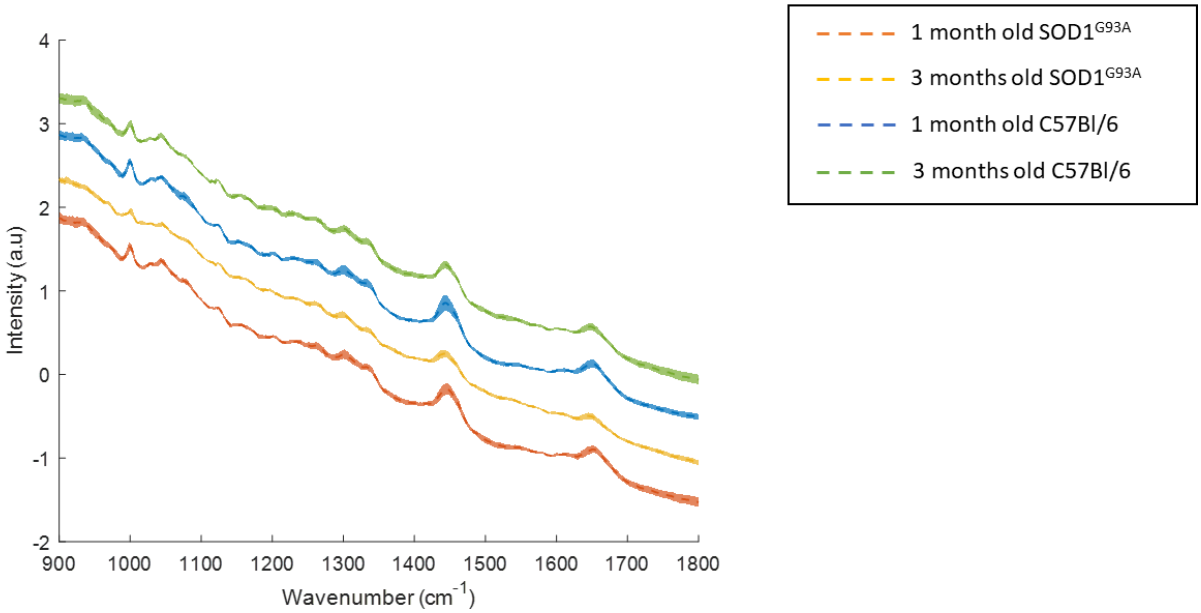


Figure 3.17: Mean Raman spectra for SOD1^{G93A} and C57Bl/6 mice groups.
The mean spectra for SOD1^{G93A} and C57Bl/6 mice of one and three months of age are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

Figures 3.18-3.25 show the mean and difference spectra of the combinations of groups used to build the two group models. The mean spectra of the different groups are very similar, with the same prominent peaks being present in all of them. Despite the signal to noise ratio of the spectra acquired with the probe being limited by the fluorescent background, biochemically reliable peaks could be seen in the mean and difference spectra.

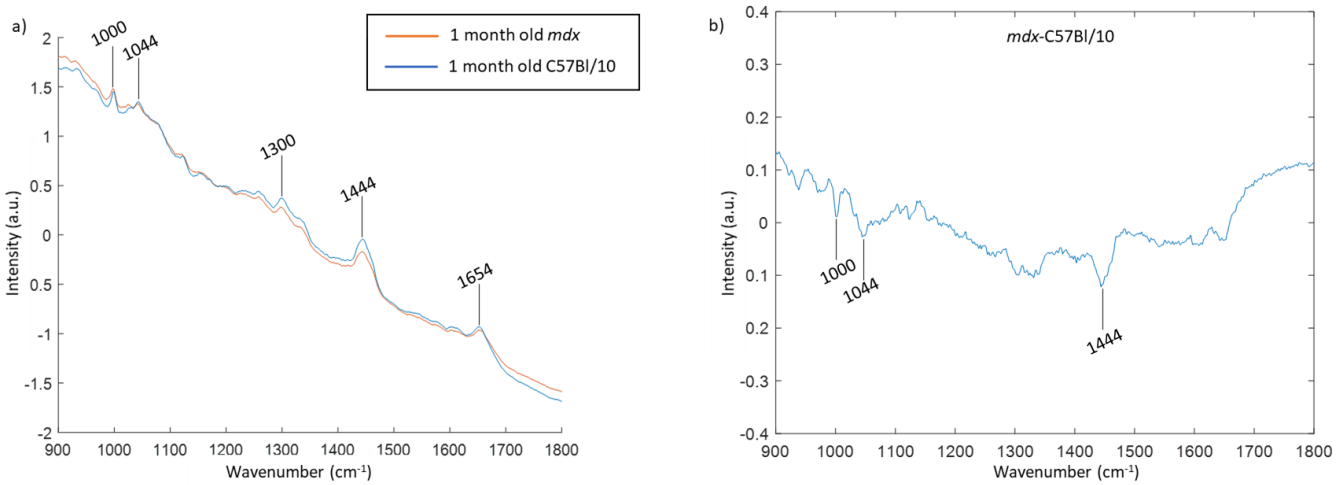


Figure 3.18: Mean and difference spectra of one month old mdx and C57Bl/10 mice.
a) Mean spectra for one and month old mdx and C57Bl/10 mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

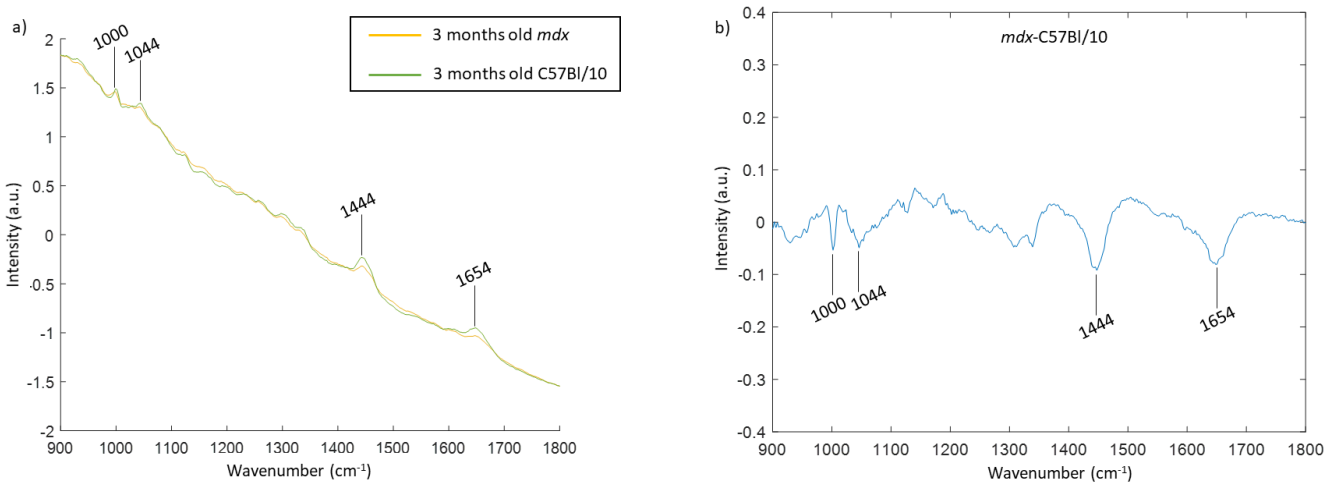


Figure 3.19: Mean and difference spectra of three months old mdx and C57Bl/10 mice.
a) Mean spectra for three months old mdx and C57Bl/10 mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

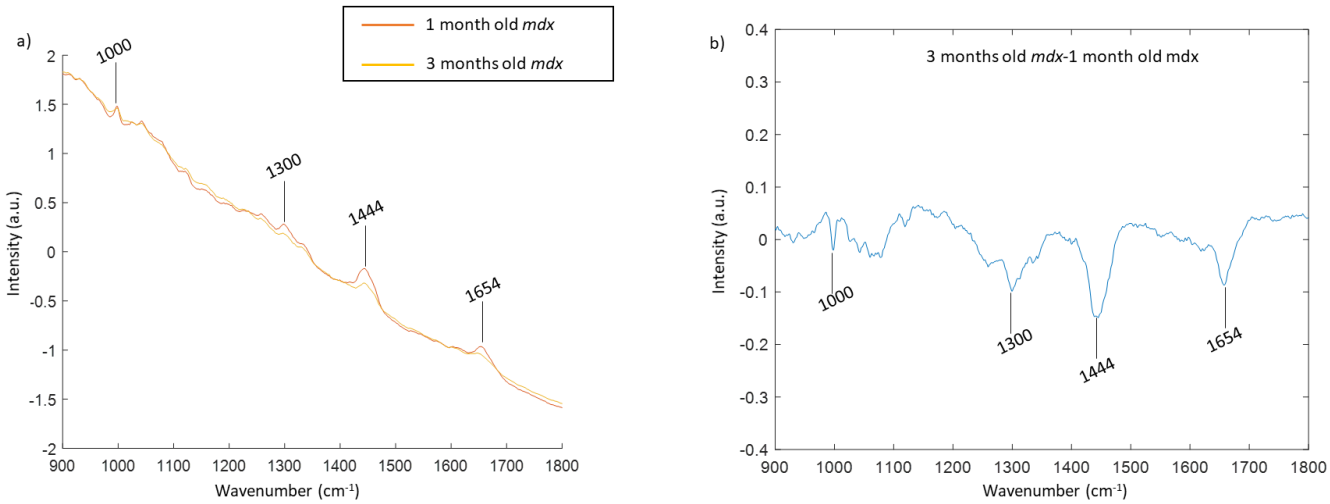


Figure 3.20: Mean and difference spectra of one and three month old mdx and C57Bl/10 mice.
a) Mean spectra for one and three months old mdx mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

A decrease in the prominent peaks located around 935cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1300 cm^{-1}, 1330 cm^{-1}, 1445 cm^{-1}, 1655 cm^{-1} as the disorder progresses can be seen in the difference spectra of the *mdx* and C57Bl/10 mice and one and three months old *mdx* mice. Tentative peak assignments for these peaks can be seen in table 3.20. A more detailed table with the references for each assignment found in this and all the following tables can be found in Appendix A.

Table 3.20: Prominent Raman peaks in mean and difference spectra of the mdx and C57Bl/10 mice and tentative peak assignments. *A decrease in these peaks was observed as the disorder progressed for the mdx mice.*

Wavenumber (cm^{-1})	Vibrational Modes	Major Assignments
935	C-C stretching	Proteins
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1300	CH_2 twisting	Lipids, Amide III (proteins)
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1444	CH_2 bending	Proteins, Lipids
1654	C=O stretching, C=C stretch	Amide I (proteins), Lipids

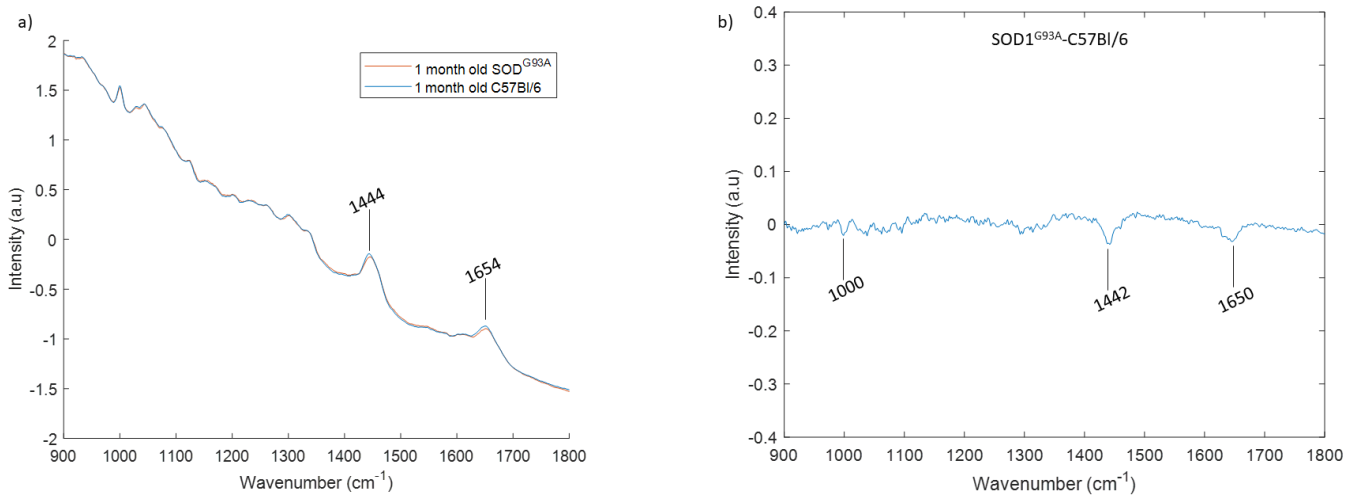


Figure 3.21: Mean and difference spectra of one month old SOD1^{G93A} and C57Bl/6 mice.

a) Mean spectra for one month old SOD1^{G93A} and C57Bl/6 mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

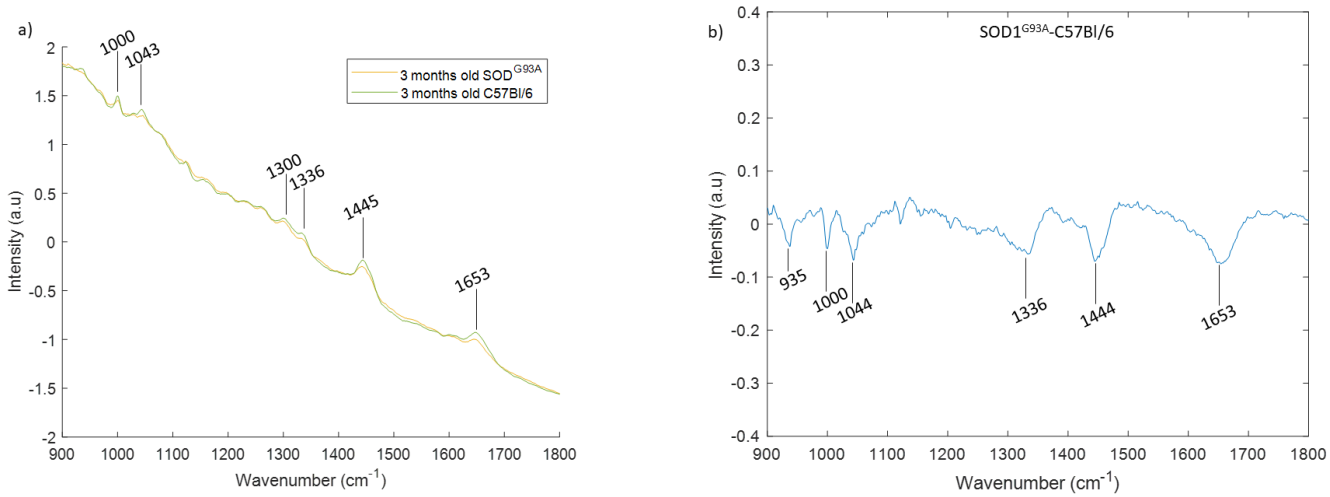


Figure 3.22: Mean and difference spectra of three months old SOD1^{G93A} and C57Bl/6 mice.
a) Mean spectra for three months old SOD1^{G93A} and C57Bl/6 mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

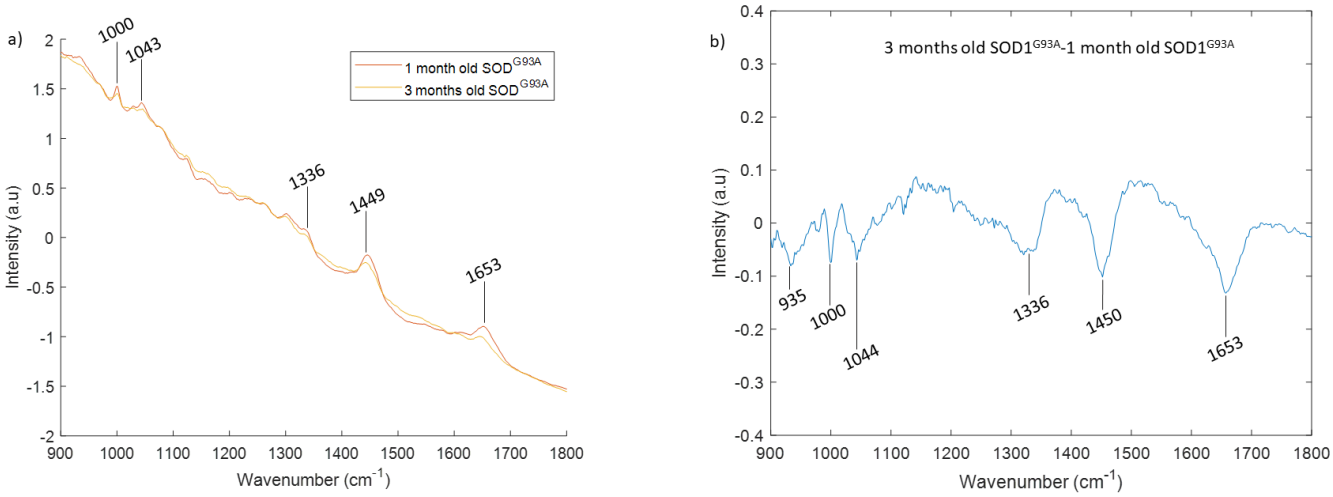


Figure 3.23: Mean and difference spectra of one and three months old SOD1^{G93A} mice.
a) Mean spectra for one and three months old SOD1^{G93A} mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

The differences observed in the intensity of the most prominent peaks in the mean and difference spectra of the one month old SOD1^{G93A} when compared to the C57Bl/6 mice were very subtle. Larger differences were observed in the intensity of the peaks when comparing spectra acquired from three months old SOD1^{G93A} to C57Bl/6 and one month old SOD1^{G93A} mice. Similar to the *mdx* mice, the most prominent peaks present in the spectra also decreased as the disorder progressed for the SOD1^{G93A} mice. Table 3.21 summarises these peaks and their tentative assignments.

Table 3.21: Prominent Raman peaks in mean and difference spectra of the SOD1^{G93A} and C57Bl/6 mice and tentative peak assignments. *A decrease in these peaks was observed as the disorder progressed for the SOD1^{G93A} mice.*

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
935	C-C stretching	Proteins
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1445-1450	CH_2 bending, CH_2CH_3 bending	Proteins, Lipids
1654	C=O stretching, C=C stretch	Amide I (proteins), Lipids

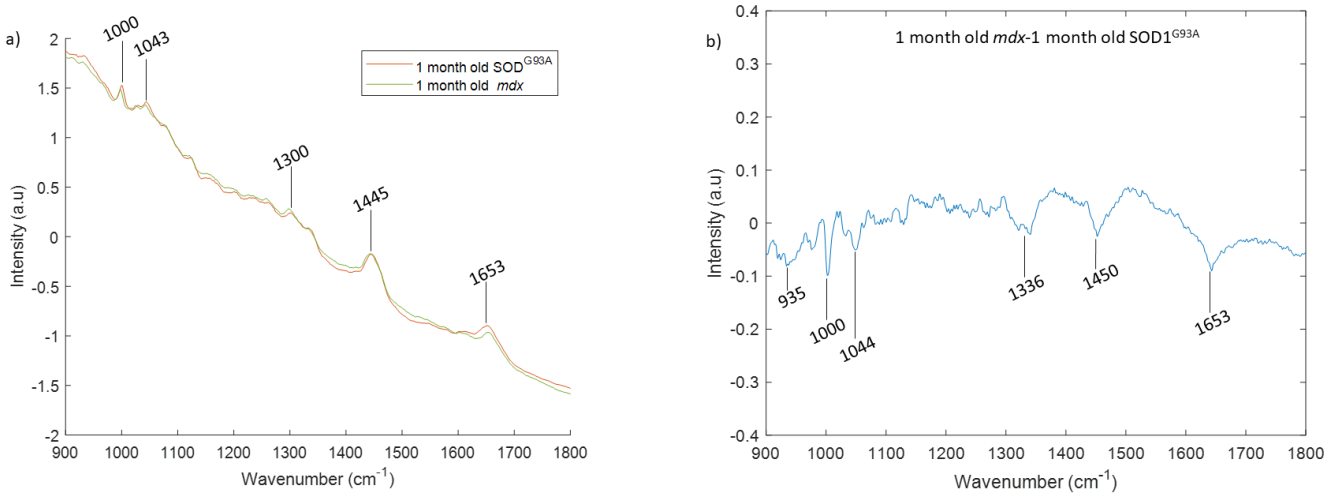


Figure 3.24: Mean and difference spectra of one month old mdx and SOD1^{G93A} mice.
a) Mean spectra for one month old mdx and SOD1^{G93A} mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

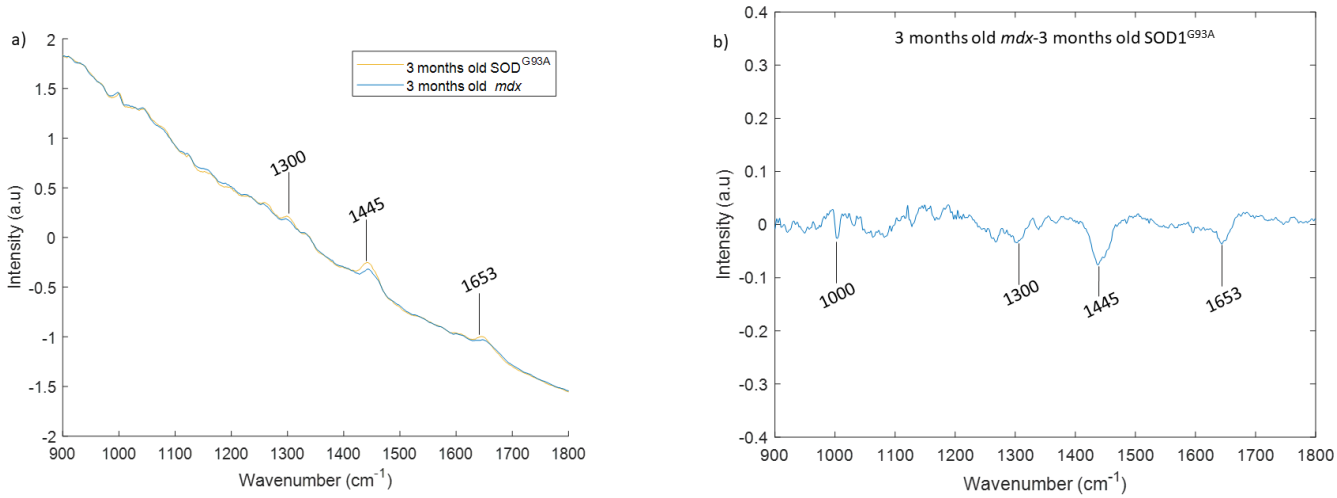


Figure 3.25: Mean and difference spectra of three months old mdx and SOD1^{G93A} mice.
a) Mean spectra for three months old mdx and SOD1^{G93A} mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

Similar differences were observed in the mean and difference spectra of *mdx* and SOD1^{G93A} mice at different ages. The difference in the intensity in the bands located around 1000 cm^{-1}, 1044 cm^{-1} and 1653 cm^{-1} appeared to have decreased in the spectra acquired from the older mice. Table 3.22 summarises these peaks and their tentative assignments.

Table 3.22: Prominent Raman peaks in mean and difference spectra of the SOD1^{G93A} and mdx mice and tentative peak assignments.

Wavenumber (cm^{-1})	Vibrational Mode s	Assignment
935	C-C stretching	Proteins
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1445-1450	CH_2 bending, CH_2CH_3 bending	Proteins, Lipids
1654	C=O stretching, C=C stretch	Amide I (proteins), Lipids

Background subtraction was used to aid identification of more spectral features present in the spectra. The background subtracted mean spectra of the different groups can be seen in figures 3.26 and 3.27. A larger number of peaks were identifiable in these mean spectra.

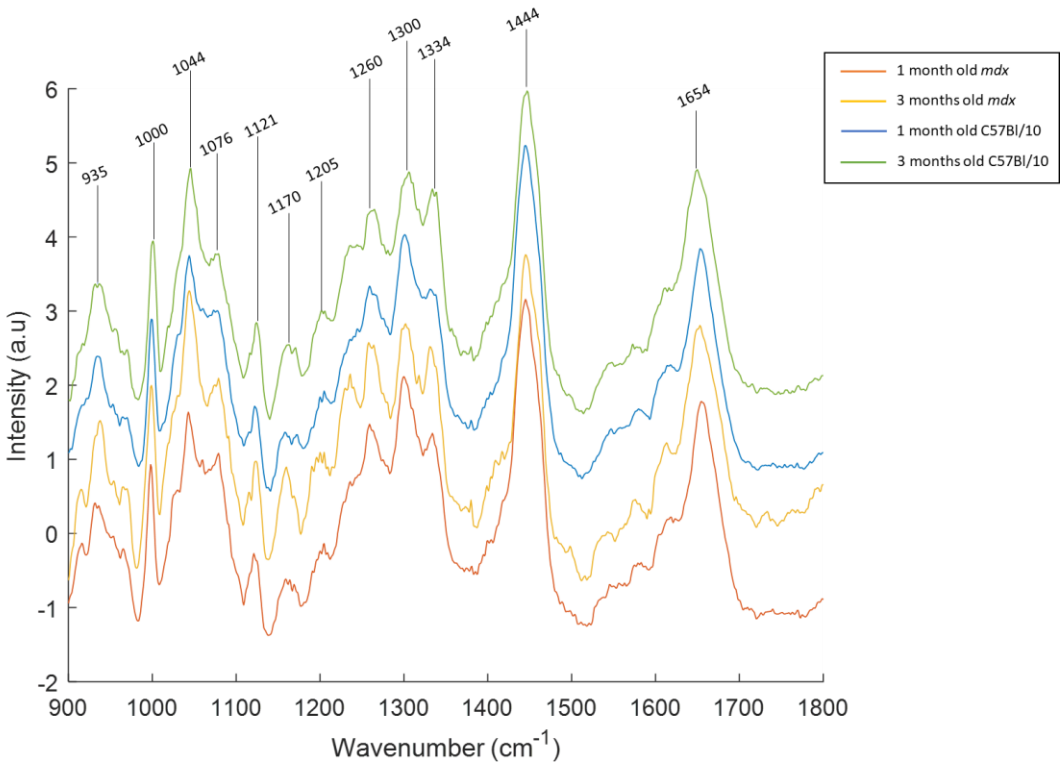


Figure 3.26: Background subtracted mean spectra of mdx and C57Bl/10 mice groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

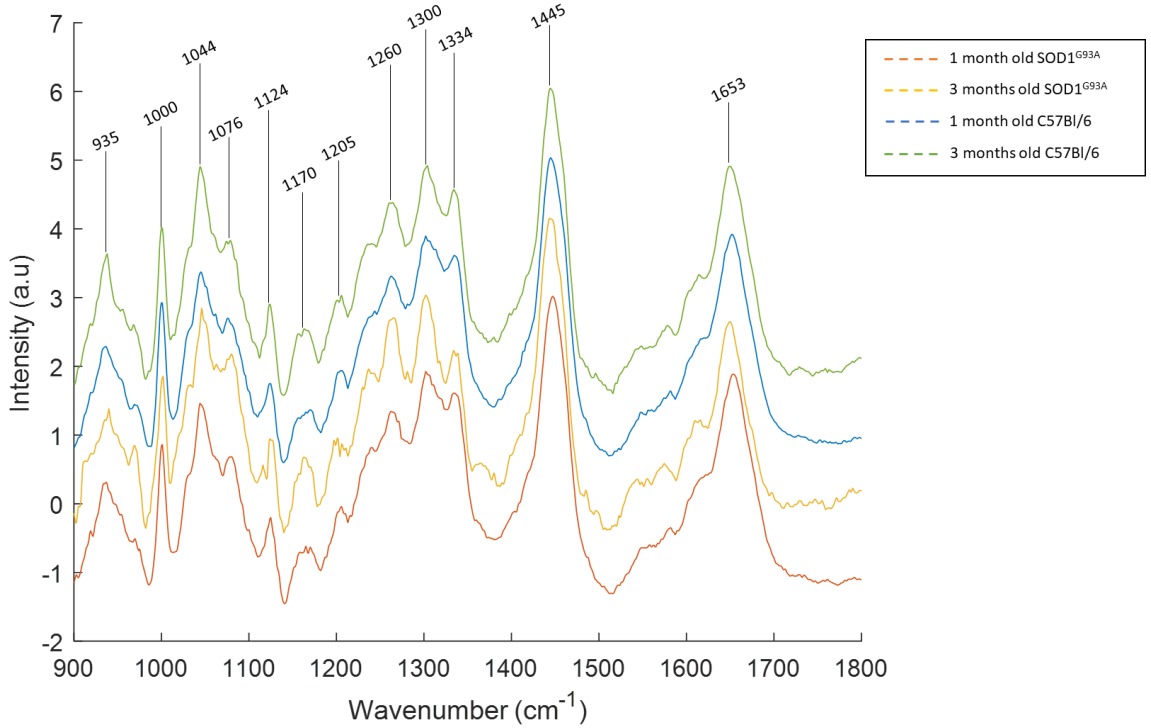


Figure 3.27: Background subtracted mean spectra of SOD1^{G93A} and C57Bl/6 mice groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

Additional shoulder peaks can be seen around 975 cm^{-1}, 1030 cm^{-1}, 1550 cm^{-1}, 1575 cm^{-1} and 1615 cm^{-1}. Tentative peak assignments for the peaks present in the background subtracted mean spectra are presented in table 3.23.

Table 3.23: Prominent Raman peaks in background subtracted spectra and tentative peak assignments. *Amino acids are specified when the peaks are largely associated with them in the literature.*

Wavenumber (cm^{-1})	Example Vibrational Mode	Assignment
935	C-C stretching	Proteins
950		Proteins (Valine, Proline, Phenylalanine)
975	CH_2 deformation	Proteins, Nucleic Acids
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1030	C-H bending	Proteins (Phenylalanine, Proline)
1044		Proteins
1076	C-C stretching, C-O stretching	Proteins (Tryptophan), Lipids, Nucleic Acids
1121	C-C stretching, C-N stretching	Proteins, Lipid
1170	C-H bending	Proteins (Tyrosine, Phenylalanine)
1205	$\nu(C\text{-}C_6H_5)$	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)
1260	C-N stretching	Amide III (proteins), Lipids
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1444	CH_2 bending	Proteins, Lipids
1550	$\nu(C\text{-}C)$	Proteins (Tryptohan)
1570	Ring breathing modes of the DNA/RNA bases	Nucleic Acids
1615	C=C stretching	Proteins (Tyrosine)
1654	C=O stretching, C=C stretching	Amide I (proteins), Lipids

Table 3.24 compares the peaks present in the spectra of the main muscle constituents and of skeletal muscle found in the literature with the peaks present in the subtracted mean spectra. A

more detailed list of peaks of muscle related components with the references can be found in Appendix B.

Table 3.24: Prominent Raman peaks in background subtracted spectra compared with peaks of major muscle components and skeletal muscle found in the literature.

Wavenumber (cm^{-1})	Myosin	Tropomyosin	Actin	Type I Collagen	Muscle Fibre	Skeletal Muscle
935	✓	✓		✓	✓	✓
950						
975	✓				✓	
1000	✓	✓	✓	✓	✓	✓
1030	✓			✓		
1044	✓				✓	
1076	✓		✓			✓
1121	✓	✓			✓	✓
1170	✓	✓				✓
1205	✓	✓			✓	
1260	✓			✓		✓
1300	✓	✓				✓
1335	✓	✓		✓	✓	✓
1444	✓	✓	✓	✓	✓	✓
1550	✓				✓	✓
1580					✓	✓
1610	✓					
1654	✓	✓	✓	✓	✓	✓

As can be seen in table 3.24 the peaks identified in the Raman muscle spectra were consistent with previously reported peaks from skeletal muscle. As expected, most of these peaks are associated with proteins probably due to the high protein content in the muscle and many of them can be found in the spectra of myosin, actin, tropomyosin and collagen (Table 3.24). The Raman spectrum of skeletal muscle is complex due to several overlapping bands (different protein bands as well as protein and lipid overlap). Band overlapping often leads to broadening of the Raman peaks, as is often the case with the peaks in the Amide I region (and can be seen in the mean spectra). Thus, identification of individual constituents from Raman spectra acquired from muscle tissue is quite complex.

3.1.2.2.2 Multivariate Analysis

As can be seen in the mean and difference figures in the previous section the most prominent peaks were present in the mean spectra of all the different groups. Despite differences in the intensity of these peaks clear visual differences were not easily apparent. Hence, multivariate techniques were employed in order to utilise and further elucidate the biochemical features present in the spectra. The PCA loadings and the linear discriminant function (LDF) as well as the PLS weights were plotted in order to illustrate the important peaks for spectral classification. PLS weights were used instead of loadings because the weight matrix is orthogonal whereas the loadings matrix is not. Thus, the weights for each component can be interpreted independently of the weights for the other components.

Assigning specific peaks to one of the two groups used in each model was not trivial since the two groups were not always clearly separated around zero in the corresponding score histograms or score plots. Additionally, the peaks in the loadings or weight plots were most often not centred around zero probably due to the presence of background in the spectra. Hence, a combination of the information found in the difference spectra and the loading/weight plots was used in an attempt to better understand the differences present in the spectra from the different groups. The score histograms and loading/weight plots in which the separation of the spectra from the different groups was best for each model are shown in the following sections. The loading/weight plots for the rest of the PCs and components used to build each model are shown in Appendix C. Tables 3.25-3.31 summarise and compare the most prominent peaks found in all the loading/weight plots (from all the components used for each model) and the ones found in the difference spectra of each two-group model.

a) One month old *mdx* vs. C57Bl/10 mice

Using Student's t-tests followed by fdr correction PC2 and PC3 scores were found to be significantly different between the one month old *mdx* and C57Bl/10 mice. PC3 scores were found to have the largest difference between the healthy and diseased mice (q_{PC2}: 1.8e-04, q_{PC3}: 6.7e-05). The score histogram and loading plot of PC3 can be seen in figure 3.28. The loading plot of PC2 can be seen in Appendix C.

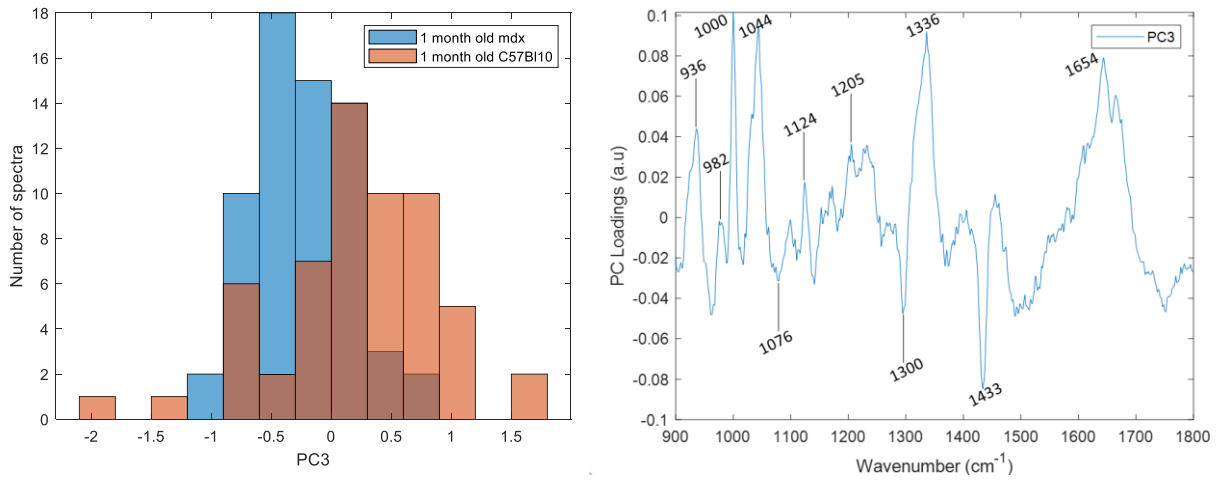


Figure 3.28: PC3 score histogram and loading plot for the one month old mdx and C57Bl/10 mice.

The results of the LDA analysis using PC2 and PC3 as input variables can be seen in figure 3.29.

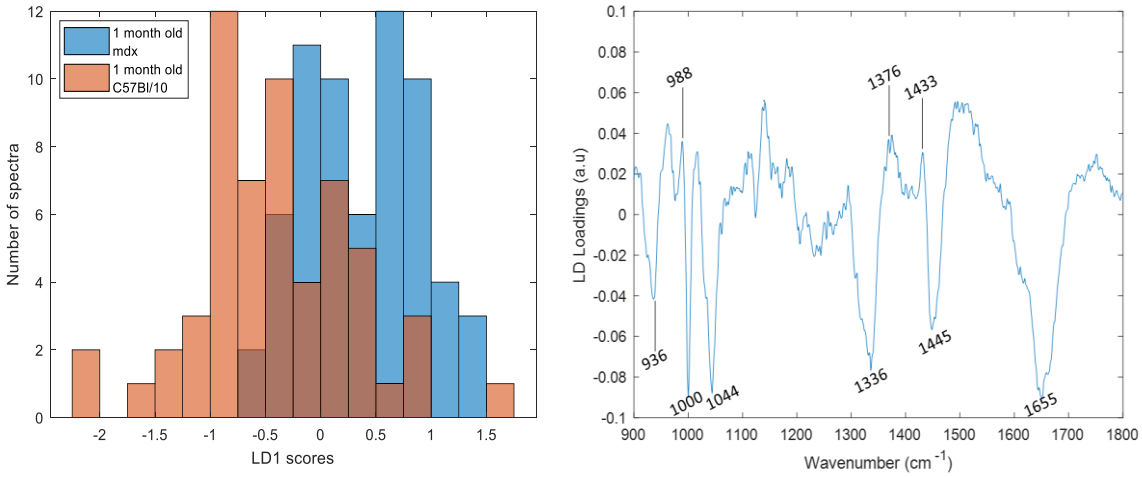


Figure 3.29: LD1 score histogram and LDF plot for the one month old mdx and C57Bl/10 mice.

The prominent bands found in PC and LDA loading plots were very similar with the bands discussed in the previous section (section 3.1.2.2.2) demonstrating that changes in muscle's biochemical composition are important for discrimination. Even though there was some overlap it can be seen from the histogram of the LD scores (figure 3.29) that positive LD1 values have a larger contribution from spectra acquired from the *mdx* mice whereas negative values from the spectra acquired from C57Bl/10 mice. Similarly, by looking at the histogram of PC3 scores (figure 3.28) spectra from C57Bl/10 mice have a larger contribution to the positive values whereas those from *mdx* mice have a larger contribution to the negative ones (peaks are upside down in PC3 and LD loading plots). This indicates that the peaks located around 936 cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1} were decreased in the spectra of the *mdx* mice. Smaller peaks around 980 cm^{-1}, 1076 cm^{-1}, 1376 cm^{-1} and 1433 cm^{-1} were found to be more prominent in the *mdx* spectra. The peaks around 1076 cm^{-1} and 1433 cm^{-1} have shown to have both protein and lipid contribution. Fat droplets exist in the muscle and so the lipid peaks might result from spectra acquired closer to muscle fat.

In the PLS-DA analysis the model generated using the first 4 components demonstrated the highest accuracy value. The spectra were not clearly separated around zero in any of the score histograms. Similar peaks as the ones from the PCA-LDA analysis were found in the PLS weight plots (Appendix C, Table 1) demonstrating that discrimination between the different groups using the different models was done on a similar biomolecular basis. The most prominent peaks in the loading/weight plots for all the PCs and PLS components and in the difference spectra and their tentative peak assignments are summarised in table 3.25.

Table 3.25: Summary table of the peaks associated with the one month old mdx and C57Bl/10 mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓		C57Bl/10
985	Proteins, Nucleic Acids	✓	✓	✓		*mdx*
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	C57Bl/10
1045	Proteins (Phenylalanine, Proline)	✓	✓	✓	✓	C57Bl/10
1075/1080	Lipids, Phospholipids	✓		✓		*mdx*
1124	Proteins	✓	✓			
1170	Proteins (Tyrosine)			✓		
1200	Proteins	✓				
1265	Amide III (proteins), Lipids	✓				
1300	Amide III (proteins), Lipids	✓	✓	✓		
1335	Proteins, Nucleic Acids	✓	✓	✓		C57Bl/10
1376	Proteins, Nucleic Acids	✓	✓	✓		
1433	Lipids	✓	✓	✓		*mdx*
1445	Proteins, Lipids	✓	✓	✓	✓	C57Bl/10
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	C57Bl/10

b) Three months old *mdx* vs. C57Bl/10 mice

Only PC2 was found to be significantly different between three months old *mdx* and C57Bl/10 mice ($q_{PC2}=8.2e^{-20}$). A sensitivity of 93.8% and a specificity of 76.6% was achieved prior cross-validation using PC2 as an input to the LDA. Since only PC2 was imported to LDA the PCA and LDA score and loading plots were identical and can be seen in figure 3.30.

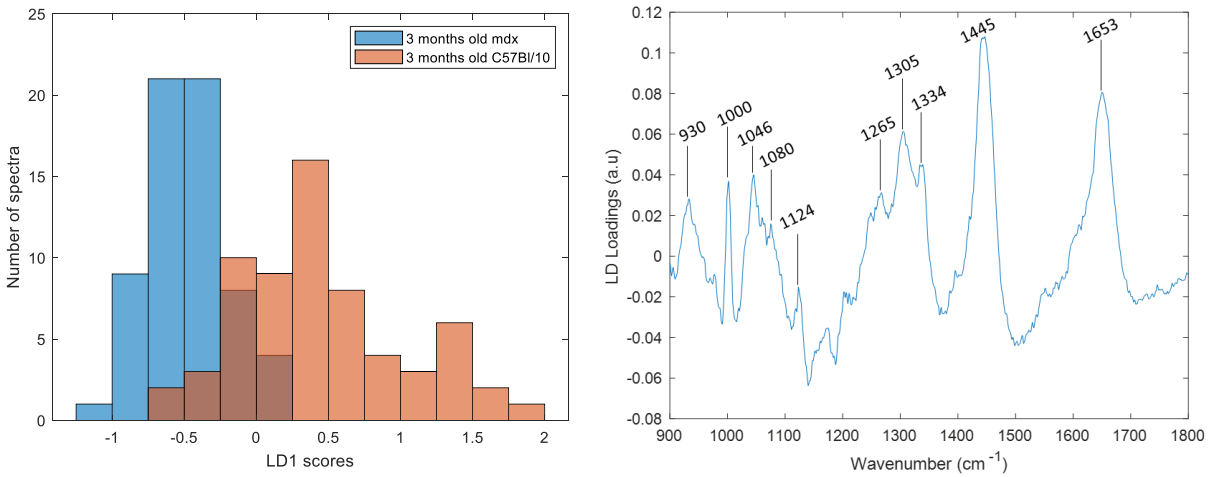


Figure 3.30: LD1 score histogram and LDF plot for the three months old mdx and C57Bl/10 mice.

In the PLS-DA analysis using the first five components generated the model with the highest accuracy value. The score and weight plot of the first component can be seen in figure 3.31. The weight plots of the rest of the components can be seen in Appendix C.

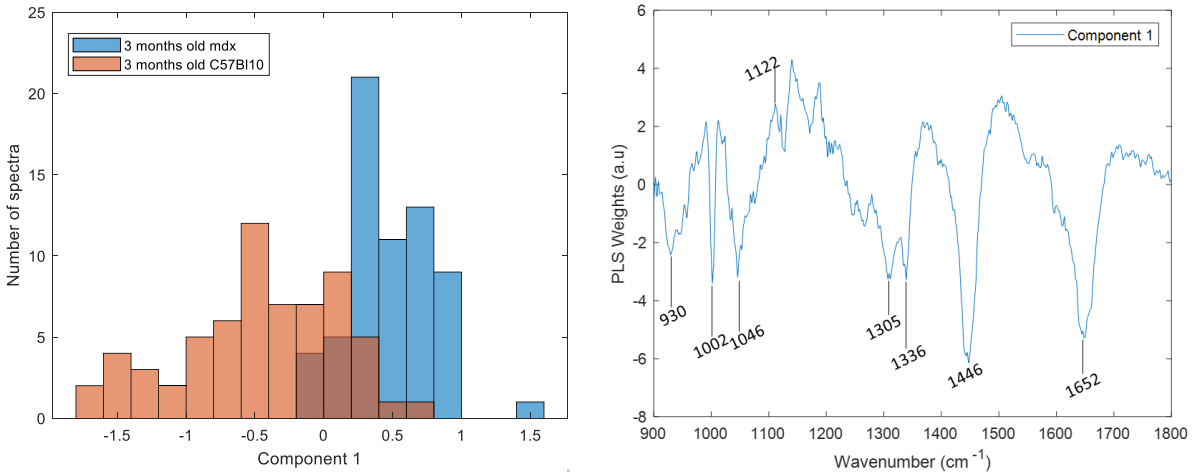


Figure 3.31: PLS Component 1 score histogram and weight plot for the three months old mdx and C57Bl/10 mice.

It can be seen in the score histograms in figures 3.29 and 3.30 that a better separation is achieved for the spectra acquired from three months old mice compared to the one month old mice. The LDA score histogram plot (figure 3.30) shows that positive values correspond to the spectra of the C57Bl/10 mice whereas negative values correspond to the spectra of the *mdx* mice whereas in the score histogram of the first pls component (figure 3.31) the spectra of the *mdx* mice were mostly accumulated in the positive values and the spectra of the C57Bl/10 mice in the negative

ones. The LDF and (component one) weight plots (figures 3.30 and 3.31 respectively) consisted of very similar peaks (in opposite directions) located around 935 cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1304 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1650 cm^{-1}. In both plots these peaks were not centred around zero, rendering associating them with one group more difficult. However, the direction of the peaks was in both cases the same as the direction that the contribution of the spectra from the C57Bl/10 mice was larger. This could indicate that these peaks decreased in the spectra of the *mdx* mice. This was consistent with the findings from the difference spectrum since most of these peaks were also present in that spectrum and were shown to have decreased for the *mdx* mice. Thus, similarly to the one month old mice, most of the major peaks (peaks around 935 cm^{-1}, 1002 cm^{-1}, 1046 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1}) seemed to have been reduced in the spectra of the three months old *mdx* mice when compared to the C57Bl/10 mice indicating that the changes found in the early disease stage are also seen in the more established one. The most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments are summarised in table 3.26.

Table 3.26: Summary table of the peaks associated with the three months old mdx and C57Bl/10 mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓		C57Bl/10
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	C57Bl/10
1045	Proteins (Phenylalanine, Proline)	✓	✓	✓	✓	C57Bl/10
1075/1080	Lipids, Phospholipids	✓	✓	✓		
1124	Proteins	✓	✓	✓		*mdx*
1265	Amide III (proteins), Lipids	✓	✓			
1300	Amide III (proteins), Lipids	✓	✓	✓	✓	C57Bl/10
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	C57Bl/10
1445	Proteins, Lipids	✓	✓	✓	✓	C57Bl/10
1653	Amide I (Proteins), Lipids	✓	✓	✓	✓	C57Bl/10

c) One vs. Three months old *mdx* mice

PCs 1 to 3 were found to be significantly different between one and three months old *mdx* mice. PC2 demonstrated the largest difference between the different groups followed by PC3. These PCs had a much smaller q-value than PC1 (q_{PC2}: 2.8e-14, q_{PC3}: 4.8e-04, q_{PC1}: 0.02). The loading plots of these PCs can be seen in Appendix C. The LD score histogram and LDF using PCs 1 to 3 can be seen in figure 3.32.

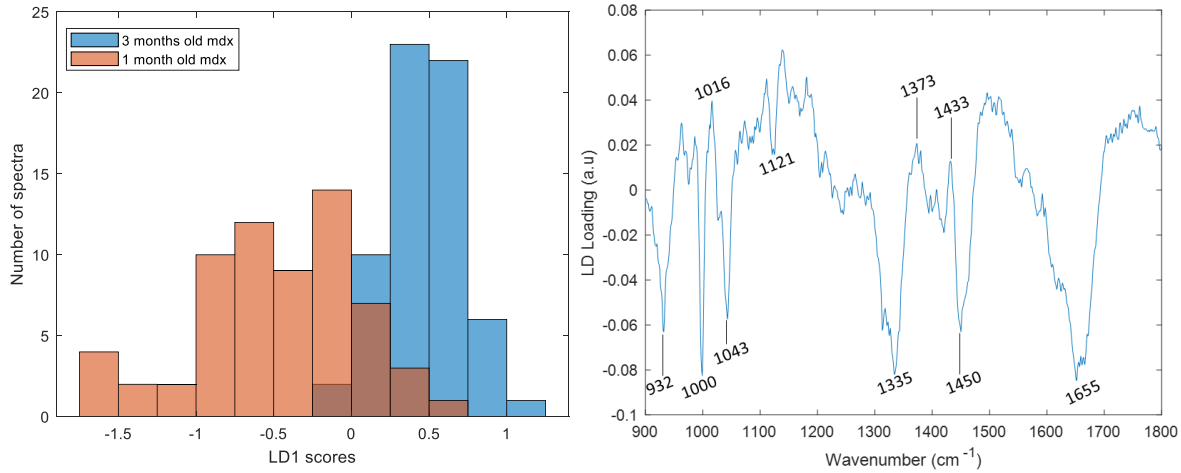


Figure 3.32: LD1 score histogram and LDF plot for the one and three months old mdx mice.

The PLS-DA model was built using the first three PLS components. The PLS score histogram and weights plot for component 1 can be seen in figure 3.33. The PLS weight plots for the rest of the components be found in Appendix C.

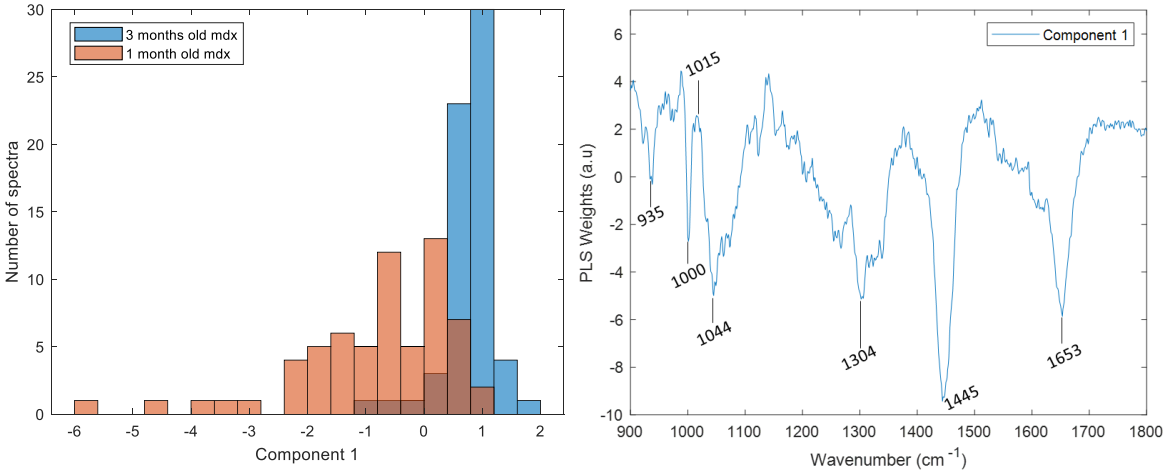


Figure 3.33: PLS Component 1 score histogram and weight plot for the one and three months old mdx mice.

As can be seen from the histogram of the LDA scores and of the scores from the first PLS component most of the spectra of the three months old mice have positive score values whereas the spectra of the one month old mice have negative score values (for both models). The respective loading plots show negative peaks around 935 cm^{-1}, 1002 cm^{-1}, 1046 cm^{-1}, 1305 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1}. The most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments are summarised in table 3.27.

Table 3.27: Summary table of the peaks associated with the one and three months old mdx mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
930	Proteins	✓	✓	✓		
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	One month old *mdx*
1016			✓	✓		
1045	Proteins (Phenylalanine, Proline)	✓	✓			One month old *mdx*
1080	Lipids, Phospholipids	✓				
1121	Proteins	✓	✓	✓		
1265	Amide III (Proteins), Lipids			✓		
1300	Amide III (proteins), Lipids	✓			✓	One month old *mdx*
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	One month old *mdx*
1375	Proteins, Nucleic Acids		✓	✓		
1435	Lipids	✓	✓			
1445/1450	Proteins, Lipids	✓	✓	✓	✓	One month old *mdx*
1655	Amide I (Proteins), Lipids	✓	✓	✓	✓	One month old *mdx*

d) Three months old $SOD1^{G93A}$ vs.C57Bl/6 mice

Using Student's t-tests followed by fdr correction PC2, PC3 PC6 and PC7 scores were found to be significantly different between the three months old $SOD1^{G93A}$ and C57Bl/6 mice. PC2 demonstrated the largest difference between the different groups followed by PC3 and PC7 and then PC6. These PCs had a much smaller q-value than PC2 (q_{PC2}: 1.3e-08, q_{PC3}: 0.006, q_{PC7}: 0.006, q_{PC7}: 0.02). The score histogram and loading plot of PC2 can be seen in figure 3.34, whereas the loading plots of PCs 3,6 and 7 can be seen in Appendix C.

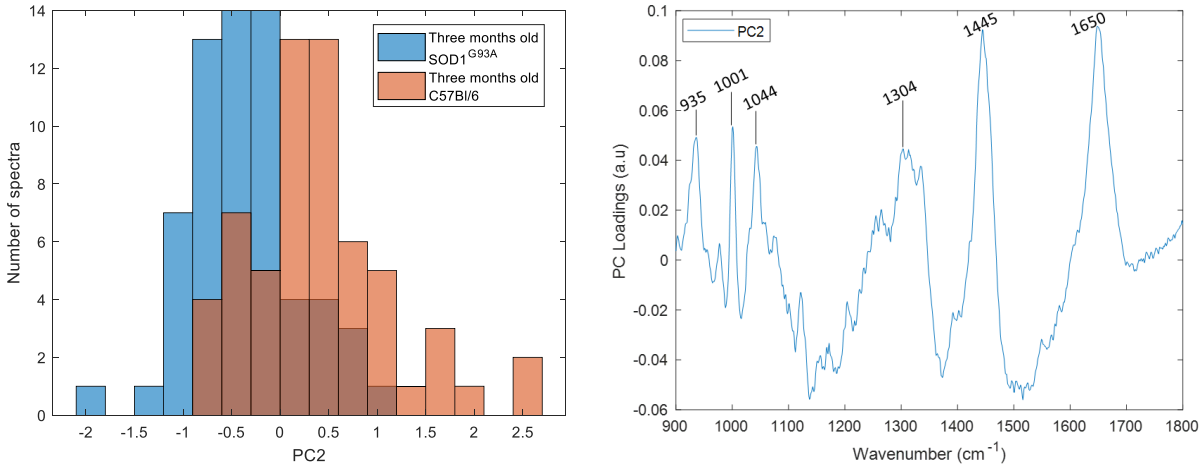


Figure 3.34: PC2 score histogram and loading plot for the three months old SOD1^{G93A} and C57Bl/6 mice.

The LDF and the LD score histogram calculated using the significant PCs, can be seen in figure 3.35. Using all significant PCs the sensitivity, specificity and accuracy of the PCA-LDA model prior cross-validation were 85.5%, 86.7%, 86.1%.

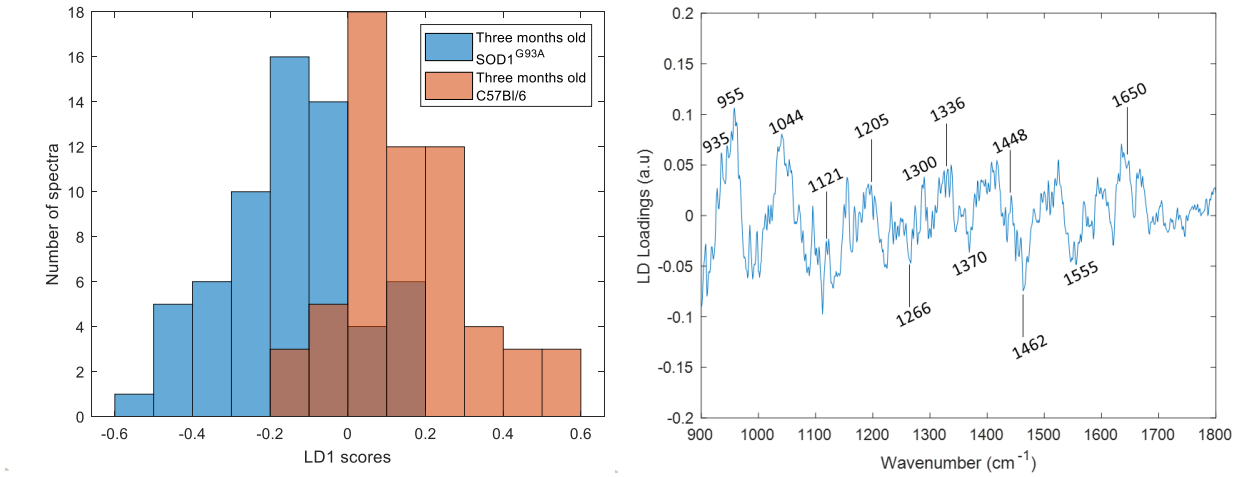


Figure 3.35: LD1 score histogram and LDF plot for the three months old SOD1^{G93A} and C57Bl/6 mice.

In the PLS-DA analysis the model generated using the first 3 components demonstrated the highest accuracy value. The PLS-DA score and weight plot of the first component can be seen in figure 3.36. The weight plots of the second and third components can be seen in Appendix C.

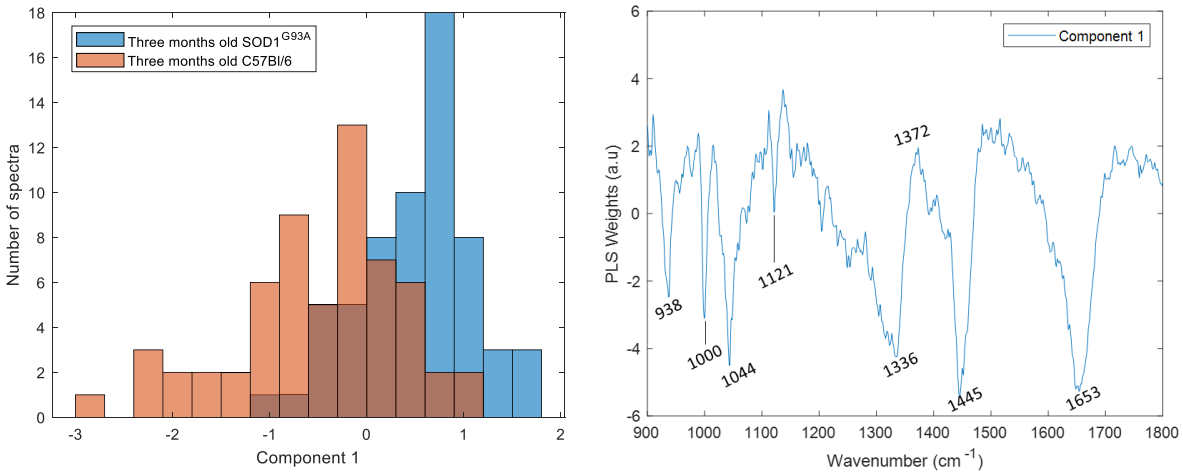


Figure 3.36: PLS Component 1 score histogram and weight plot for the three months old SOD1^{G93A} and C57Bl/6 mice.

The PCA and PLS weight plots for PC2 and component 1 were very similar (just in opposite directions) demonstrating that discrimination between three months old SOD1 and C57Bl/6 mice was due to similar peaks for both models. Despite some overlap, it can be seen in the score histogram of PC2 that in this PC negative values largely correspond to spectra from the SOD1^{G93A} mice whereas positive values have a larger contribution from C57Bl/6 mice. In the score histogram of component one spectra from SOD1^{G93A} have a larger contribution in the positive values whereas spectra from C57Bl/6 mice have a larger contribution in the negative values. The main peaks in the loading and weight plots were located around 935 cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1304 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1650 cm^{-1} and were all in the same direction as the direction that the contribution of the spectra from the C57Bl/6 mice was larger, possibly indicating that these peaks could have been decreased in the spectra of the SOD1^{G93A} mice. This was consistent with the findings from the difference spectrum since most of these peaks were also present in the spectrum and were shown to have decreased for the SOD1^{G93A} mice.

Most of these peaks were also found in the linear discriminant function. In the histogram of the LD scores positive LD1 values had a larger contribution from spectra acquired from the C57Bl/6 mice whereas negative values from the spectra acquired from SOD1^{G93A} mice. Positive peaks were located around 935 cm^{-1}, 1044 cm^{-1}, 1300 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1650 cm^{-1}. Negative peaks were located around 1266 cm^{-1}, 1370 cm^{-1}, 1462 cm^{-1} and 1555 cm^{-1}. The most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments are summarised in table 3.28.

Table 3.28: Summary table of the peaks associated with the three months old SOD1^{G93A} and C57Bl/6 mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓	✓	C57Bl/6
955	Proteins, Hydroxyapatite	✓	✓	✓		
1000	Proteins (Phenylalanine)	✓		✓	✓	C57Bl/6
1045	Proteins (Proline)	✓	✓	✓	✓	C57Bl/6
1080	Lipids, Phospholipids	✓		✓		
1121	Proteins	✓	✓	✓		
1151	Proteins, Carotenoids			✓		
1200	Proteins	✓	✓	✓		
1269	Amide III (Proteins), Lipids	✓		✓		SOD1^{G93A}
1300	Amide III (proteins), Lipids	✓	✓	✓		
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	C57Bl/6
1370	Proteins, Nucleic Acids		✓	✓		SOD1^{G93A}
1434	Lipids	✓		✓		
1445	Proteins, Lipids	✓	✓	✓	✓	C57Bl/6
1462	Nucleic Acids, Proteins, Palmitic Acid		✓			SOD1^{G93A}
1550	Proteins (Tryptophan)	✓	✓			SOD1^{G93A}
1615	Proteins (Tryptophan, Tyrosine)	✓				
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	C57Bl/6

e) One vs. Three months old SOD1^{G93A} mice

Scores from PCs 1,3,6 and 7 were found to be significantly different between one and three months old SOD1^{G93A} mice. When using all significant PCs, the PCA-LDA model demonstrated a sensitivity of 95.2%, specificity of 96.7% and accuracy of 95.9% prior cross-validation. PC1 (figure 3.37) was the most significant followed by PC7, PC3 and PC6. The latter PCs demonstrated a much smaller q-value thanPC1 (q_{PC1}: 1.2e-22, q_{PC7}: 0.02, q_{PC3}: 0.04, q_{PC6}: 0.04). The LD score histogram and the LDF generated using all the significant PCs are shown in figure 3.38. The loading plots of PCs 3,6 and 7 can be found in Appendix C.

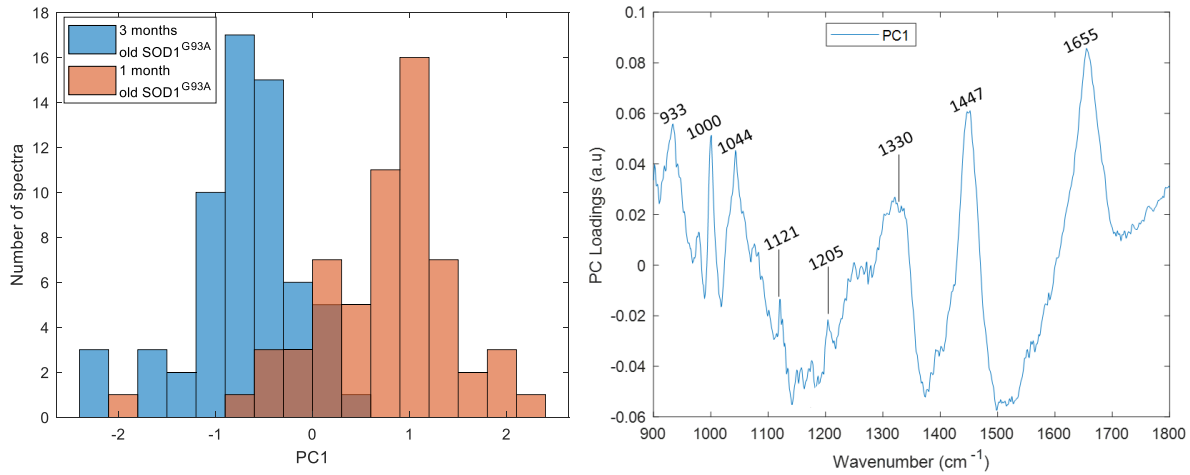


Figure 3.37: PC1 score histogram and loading plot for the one and three months old SOD1^{G93A} mice.

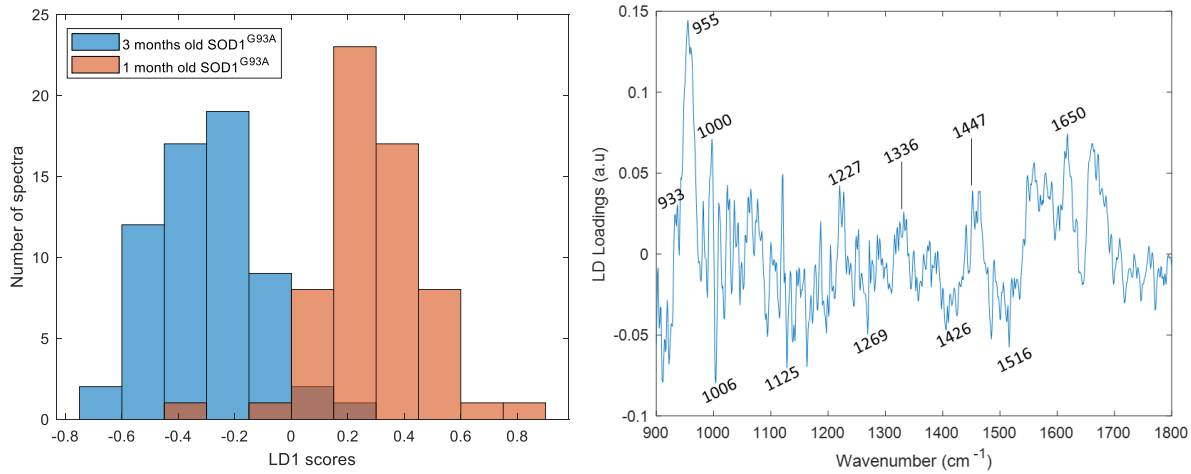


Figure 3.38: LD1 score histogram and LDF plot for the one and three months old SOD1^{G93A} mice.

The PLS-DA model was generated using the first four components. The score histogram and the weight plot of the first component can be seen in figure 3.39. The weight plots of components 2 to 4 can be found in Appendix C.

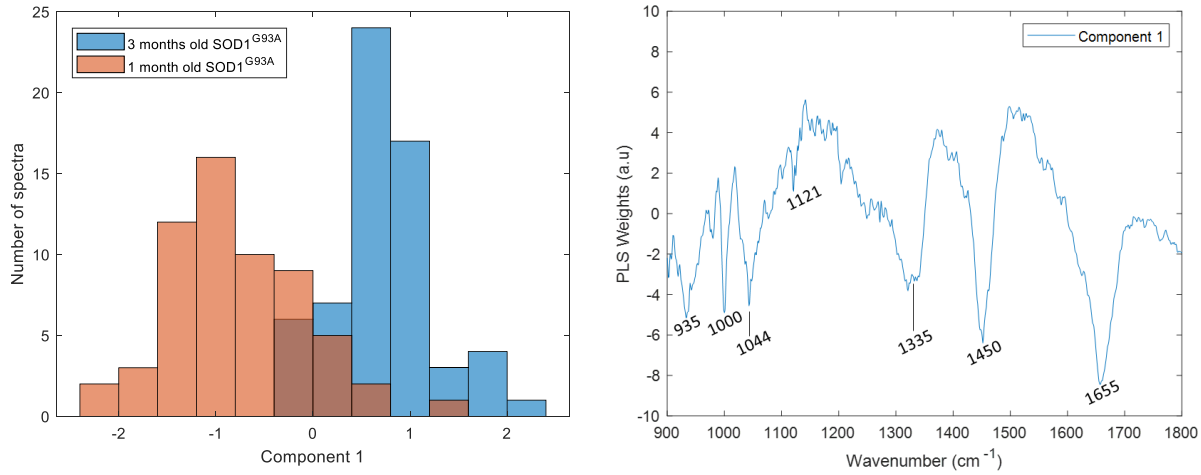


Figure 3.39: PLS Component 1 score histogram and weight plot for the one and three months old SOD1^{G93A} mice.

The loading and weight plots of PC1 and PLS component one were very similar with the most prominent peaks being in the direction where the contribution of the spectra acquired from younger SOD1^{G93A} mice was larger (positive score values for PC1 and negative score values for component 1). This was also observed in the respective difference spectrum. Thus, the bands located around 935 cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1650 cm^{-1} appeared to decrease as the disorder progressed. The LDF contained more noise. Hence, the identification of the Raman peaks responsible for the discrimination between the two groups was more difficult. The peaks around 933 cm^{-1}, 1000 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1650 cm^{-1} were also identified in the LDF. These were found to be increased in the spectra of the one month old SOD1^{G93A} mice (positive values in the LD score and loading plots). Additional peaks around 955 cm^{-1} and 1226 cm^{-1} were also found to be increased in these spectra. Negative peaks were located around 1125 cm^{-1}, 1269 cm^{-1}, 1426 cm^{-1} and 1516 cm^{-1} and were associated with the spectra acquired from the three months old SOD1 ^{G93A} mice (negative values in the LD score plot). The most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments are summarised in table 3.29.

Table 3.29: Summary table of the peaks associated with the one/three months old SOD1^{G93A} mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓	✓	One month old SOD1^{G93A}
950	Proteins, Hydroxyapatite	✓	✓	✓		
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	One month old SOD1^{G93A}
1045	Proteins (Proline)	✓		✓	✓	One month old SOD1^{G93A}
1080	Lipids, Phospholipids	✓		✓		
1121	Proteins	✓	✓	✓		Three months old SOD1^{G93A}
1151	Proteins, Carotenoids	✓				
1200	Proteins	✓				
1269	Amide III (Proteins), Lipids	✓	✓	✓		Three months old SOD1^{G93A}
1300	Amide III (proteins), Lipids	✓				
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	One month old SOD1^{G93A}
1370	Proteins, Nucleic Acids	✓				
1400	Proteins, Nucleic Acids			✓		
1430	Lipids	✓	✓			Three months old SOD1^{G93A}
1445/1450	Proteins, Lipids	✓	✓	✓	✓	One month old SOD1^{G93A}
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	One month old SOD1^{G93A}

f) One month old *mdx* vs. SOD1^{G93A} mice

PCs 1,2,3,5,6,7 and 9 were found to be significantly different between one month old *mdx* and SOD1^{G93A} mice. PCs 3 and 6 were found to be the most significant PCs (q_{PC3}: 1.2e-05, q_{PC6}: 7.6e-05) and demonstrated much smaller q-values than the rest of the PC scores. The score histogram and loading plot for PC3 can be seen in figure 3.40. The LD score histogram and the linear discriminant function generated using all the significant PCs are shown in figure 3.41.

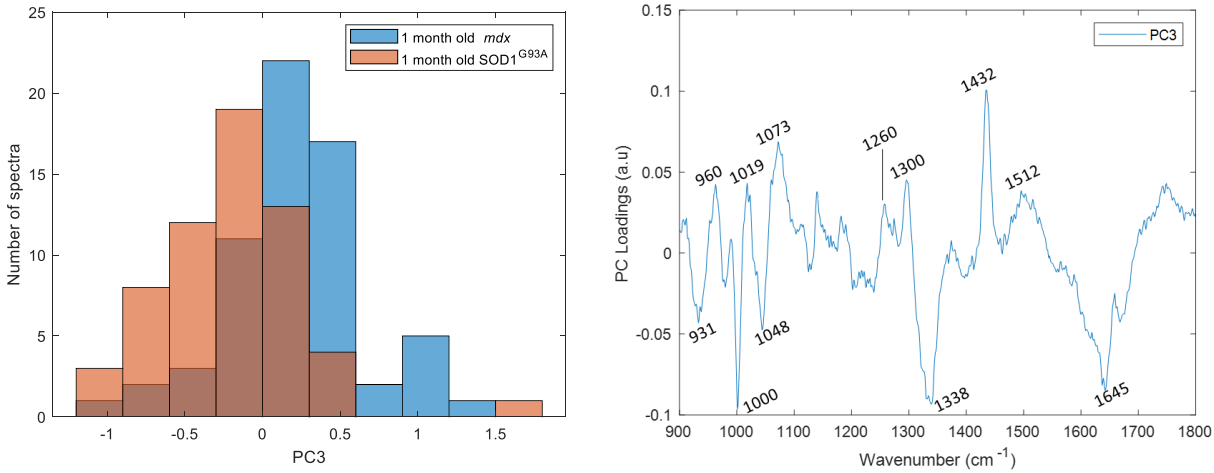


Figure 3.40: PC3 score histogram and loading plot for the one month old mdx and SOD1^{G93A} mice.

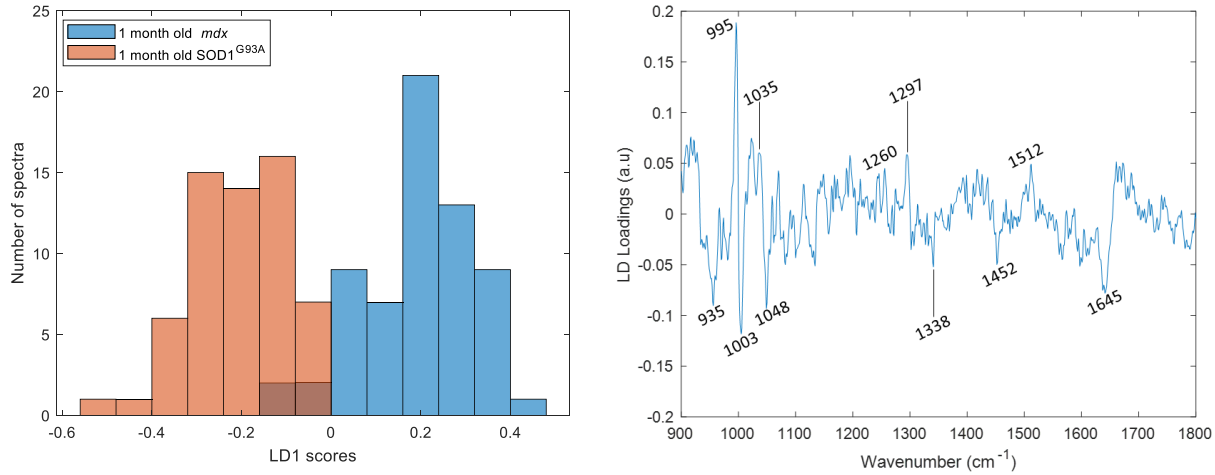


Figure 3.41: LD1 score histogram and LDF plot for the one month old mdx and SOD1^{G93A} mice.

The PLS-DA model was generated using the first four components. The score histogram and the weight plot of the first component can be seen in figure 3.42. The weight plots of components 2 to 4 can be found in Appendix C.

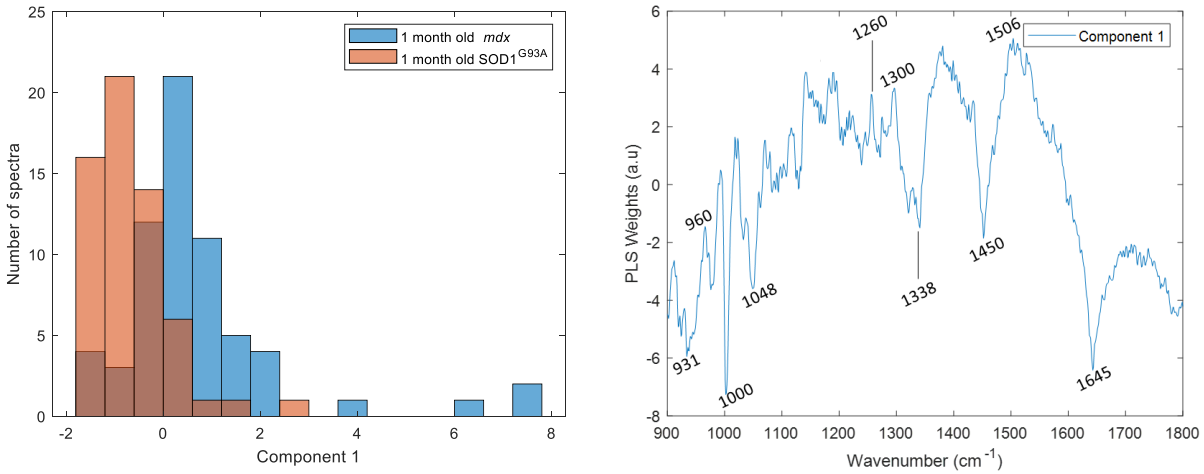


Figure 3.42: PLS Component 1 score histogram and weight plot for the one month old mdx and SOD1^{G93A} mice.

Despite some overlap in the score histogram of PC3, positive score values seemed to have a larger contribution from spectra acquired from one month old *mdx* mice whereas negative score values had a larger contribution from the spectra of the one month old SOD1^{G93A}. The negative peaks in the PC3 loading plot were located around 933 cm^{-1}, 1000 cm^{-1}, 1048 cm^{-1}, 1338 cm^{-1} and 1645 cm^{-1} and the positive peaks were around 963 cm^{-1}, 1019 cm^{-1}, 1073 cm^{-1}, 1267 cm^{-1}, 1300 cm^{-1}, 1432 cm^{-1} and 1510 cm^{-1}. Similar peaks were found in the weight plot of the first components from the PLS-DA analysis. The weight plot also contained some background. However, the peaks around 933 cm^{-1}, 1000 cm^{-1}, 1048 cm^{-1}, 1338 cm^{-1} and 1645 cm^{-1} were also in the direction that the contribution of the SOD1^{G93A} spectra were more prominent. Additionally, in the weight plot the peak around 1450 cm^{-1} also appeared and seemed to be increased in these spectra. Positive peaks were located around 965 cm^{-1}, 1267 cm^{-1}, 1300 cm^{-1} and 1506 cm^{-1}.

In the histogram of the LD scores the spectra of the different groups were very well separated around zero and positive LD1 values had a larger contribution from spectra acquired from the *mdx* mice whereas negative values from the spectra acquired from SOD1^{G93A} mice. Hence, the peaks in the associated LDF can be more robustly assigned to each group. Positive peaks were located around 995 cm^{-1}, 1020 cm^{-1}, 1260 cm^{-1}, 1297 cm^{-1} and 1506 cm^{-1}. Negative peaks were found around 940 cm^{-1}, 1003 cm^{-1}, 1048 cm^{-1} , 1338 cm^{-1}, 1452 cm^{-1} and 1645 cm^{-1}. Table 3.30 summarises the most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments.

Table 3.30: Summary table of the peaks associated with the one month old mdx/SOD1^{G93A} mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓	✓	SOD1^{G93A}
960	Proteins, Hydroxyapatite	✓		✓		
995	Uracil, Proline		✓			*mdx*
1003	Proteins (Phenylalanine)	✓	✓	✓	✓	SOD1^{G93A}
1035	Proteins (Proline, Valine, Phenylalanine)		✓			*mdx*
1045	Proteins (Proline)	✓		✓	✓	SOD1^{G93A}
1075	Lipids, Phospholipids	✓		✓		
1125	Proteins	✓		✓		
1265	Amide III (Proteins), Lipids	✓	✓	✓		*mdx*
1300	Amide III (proteins), Lipids	✓	✓			*mdx*
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	SOD1^{G93A}
1400	Proteins, Nucleic Acids			✓		
1432	Lipids	✓				
1445/1450	Proteins, Lipids	✓	✓	✓	✓	SOD1^{G93A}
1465	Nucleic Acids, Proteins, Palmitic Acid					
1512	Nucleotide Bases		✓			*mdx*
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	SOD1^{G93A}

g) Three months old *mdx* vs. SOD1^{G93A} mice

PCs 2 and 4 were found to be significantly different between one month old *mdx* and SOD1^{G93A} mice (q_{PC2}: 1.2e-07, q_{PC4}: 1.0e-08). The loading plots for PC2 and PC4 can be in Appendix C. The LD score histogram and the linear discriminant function generated using PCs 2 and 4 can be seen in figure 3.43. In the LDF negative peaks were located around 1000 cm^{-1}, 1267 cm^{-1}, 1300 cm^{-1}, 1445 cm^{-1} and 1645 cm^{-1}. Since, the spectra of the SOD1^{G93A} mice were more prominent in the negative values of the LD histogram, these peaks seemed to be increased in the spectra of these mice.

Spectra from both *mdx* and SOD1^{G93A} mice both contributed in the positive values of the LD histogram. However, the contribution of the *mdx* spectra appeared to be larger.

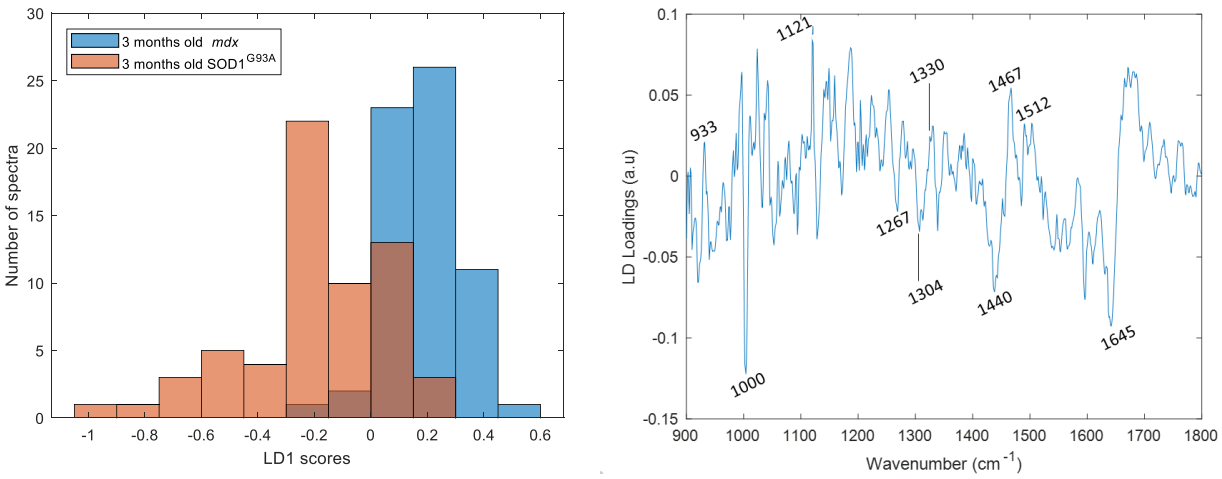


Figure 3.43: LD1 score histogram and LDF plot for the three months old mdx and SOD1^{G93A} mice.

The PLS-DA model was generated using the first five components. The score histogram and the weight plot of the first component can be seen in figure 3.44. The weight plots of components two to five can be found in Appendix C.

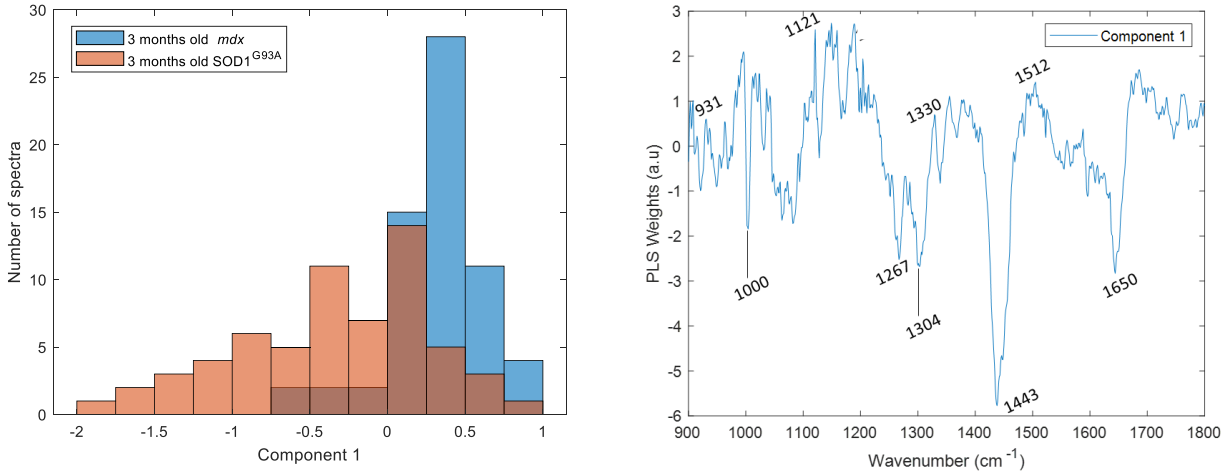


Figure 3.44: PLS Component 1 score histogram and weight plot for the three months old mdx and SOD1^{G93A} mice.

The peaks in the weight plot for component 1 were similar with the ones of the LDF plot. Negative peaks were located around 1000 cm^{-1}, 1267 cm^{-1}, 1304 cm^{-1}, 1443 cm^{-1} and 1650 cm^{-1} whereas positive peaks were located around 931 cm^{-1}, 1121 cm^{-1}, 1330 cm^{-1}, and 1512 cm^{-1}. Table 3.31 summarises the most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments.

Table 3.31: Summary table of the peaks associated with the three months old mdx and SOD1^{G93A}mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓		*mdx*
955	Proteins, Hydroxyapatite			✓		
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	SOD1^{G93A}
1045	Proteins (Proline)	✓		✓		
1075	Lipids, Phospholipids	✓		✓		
1125	Proteins	✓	✓	✓		*mdx*
1267	Amide III (Proteins), Lipids	✓	✓	✓		SOD1^{G93A}
1300	Amide III (proteins), Lipids	✓	✓	✓	✓	SOD1^{G93A}
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	
1400	Proteins, Nucleic Acids			✓		
1440/1455	Proteins, Lipids	✓	✓	✓	✓	SOD1^{G93A}
1465	Nucleic Acids, Proteins, Palmitic Acid		✓			*mdx*
1512	Nucleotide Bases		✓	✓		*mdx*
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	SOD1^{G93A}

3.1.2.3 Histology

Histological analysis was performed on muscle tissue sections of *mdx*, SOD1^{G93A}, C57Bl/6 and C57Bl/10 mice of both ages in order to investigate the morphological changes in muscle as the two neuromuscular disorders progress. H&E sections of one and three months old *mdx* and control muscles can be seen in figures 1 and 2 respectively. Sections of one and three months old SOD1^{G93A} and control mice can be seen in figures 3 and 4.

The histological assessment revealed that healthy muscle consisted of polygonal fibres, with peripherally located nuclei (figures 1c, 2c, 3c, 4c). Necrotic fibres with inflammatory cells and a small amount of early regeneration, indicated by small myofibres with centrall nuclei could be seen in the muscle sections of the dystrophic mice at one month of age (figures 1a and 1b). As the disorder progressed, the main histological findings were larger cells with centrally placed nuclei indicating active regeneration (figures 2a and 2b). Areas of inflammation were also present (figure 2b). There were no signs of pathology displayed in the tissue sections of the one month old SOD1^{G93A} mice (figure 3a). At three months of age the SOD1^{G93A} muscle showed signs of denervation in the form of grouped atrophy (figure 4a, b). Small angular, as well as hypertrophic,

fibres were present in the muscle. Finally, a small number of fibres with central nuclei indicated limited regeneration.

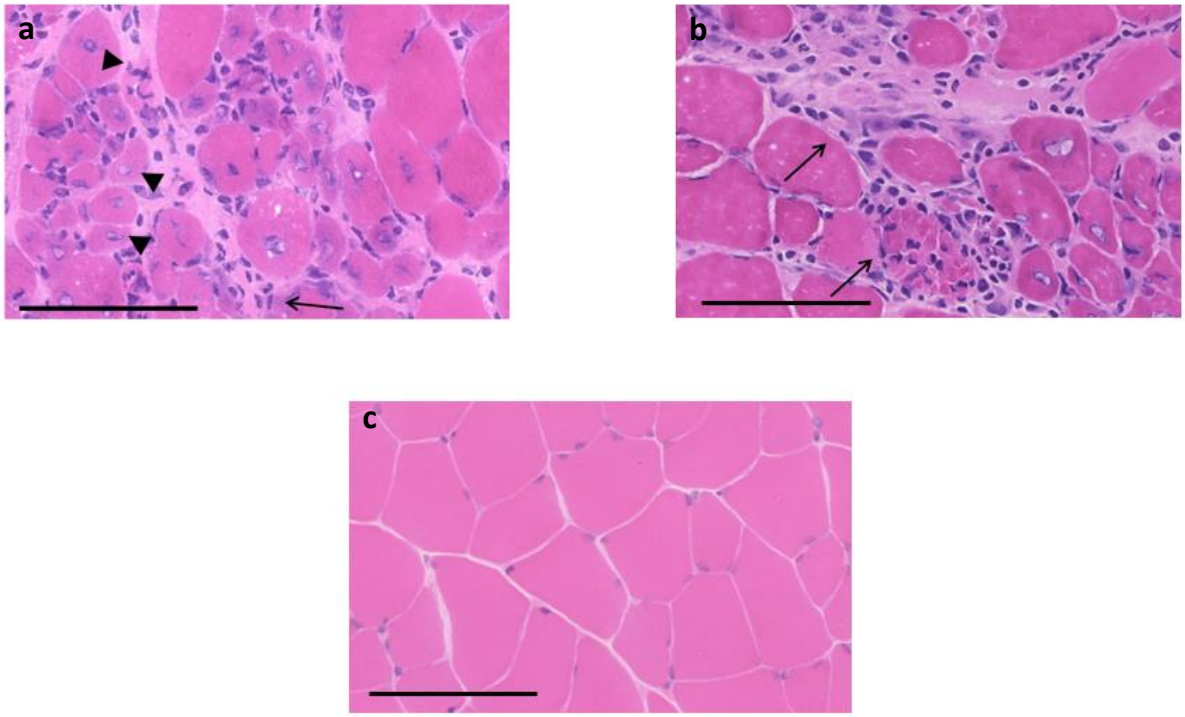


Figure 3.45: Gastrocnemius muscle sections from one month old mdx and C57Bl/10 mice stained with haematoxylin/eosin. *In one month old mdx mice sections necrotic fibres with inflammatory cells (a,b, arrows) and small myofibres with centrally placed nuclei (a, arrow heads) indicating early regeneration can be seen in the sections acquired from one month old mdx mice. Normal myofibers from control tissue can be seen in figure c. Magnification: x40, scale bars: 100µm.*

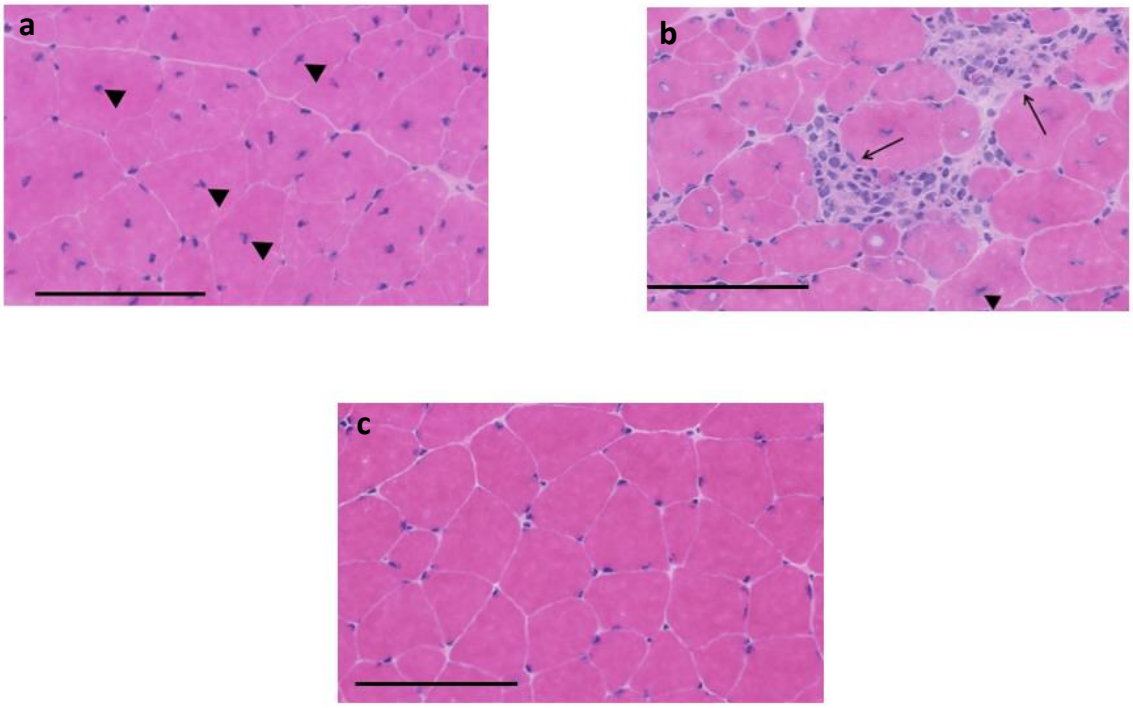


Figure 3.46: Gastrocnemius muscle sections from three months old mdx and C57Bl/10 mice stained with haematoxylin/eosin. *Larger muscle cells with centrally placed nuclei (regeneration) (a,b, arrowheads and inflammation (b, arrow) can be seen in the sections acquired from three months old mice. Normal myofibers from control tissue can be seen in figure c. Magnification: x40, scale bars: 100µm.*

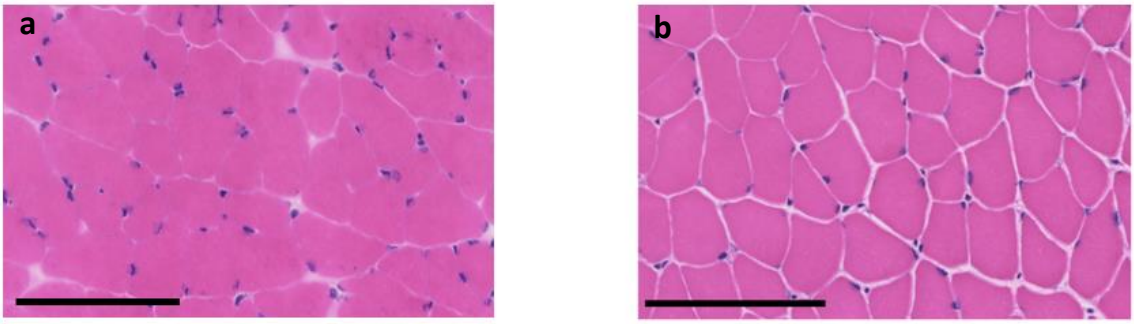


Figure 3.47: Gastrocnemius muscle sections from one month old SOD1^{G93A} and C57Bl/6 mice stained with haematoxylin/eosin. *One month old SOD1^{G93A} muscle displayed no evidence of pathology with normal myofibres present. No differences were observed between SOD1^{G93A} (a) and C57Bl/6 (b) muscle. Magnification: x40, scale bars: 100µm.*

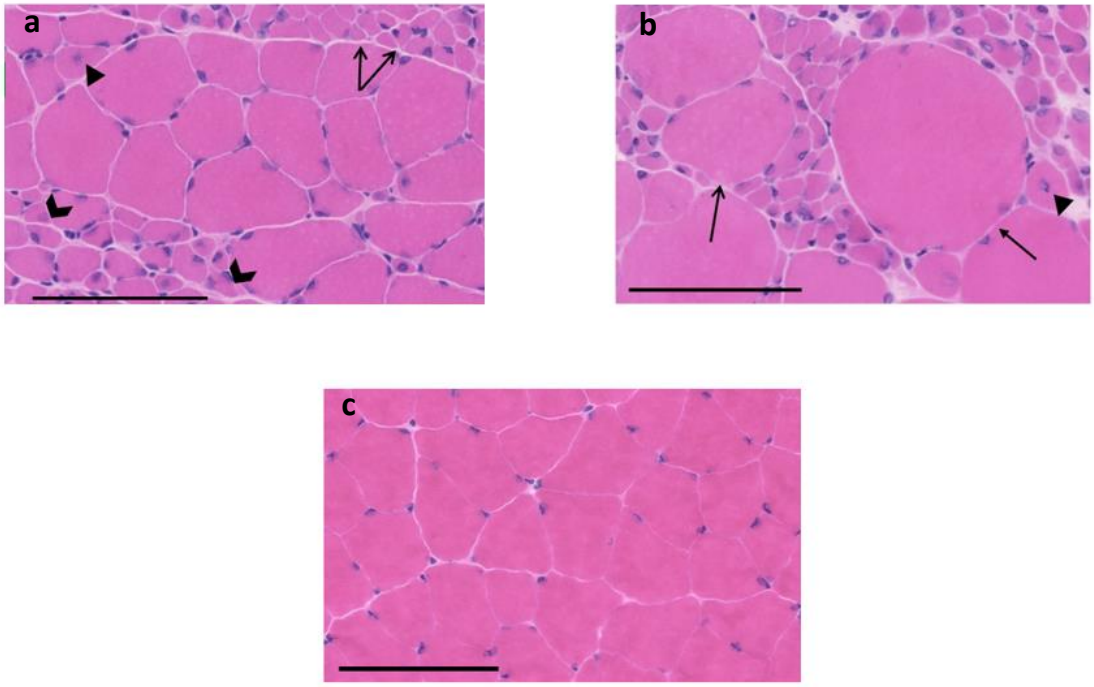


Figure 3.48: Gastrocnemius muscle sections from three months old SOD1^{G93A} and C57Bl/6 mice stained with haematoxylin/eosin. *Three months old SOD1^{G93A} showed signs of denervation in the form of grouped atrophy (a, double arrow), small angular fibres (a, chevrons), as well as hypertrophic fibres (b, arrow) and centrally placed nuclei (b, arrow head). Normal myofibers from control tissue can be seen in figure c. Magnification: x40, scale bars: 100µm.*

3.1.2.4 Post-experiment Motor Function Assessment

3.1.2.4.1 Rotarod test

Rotarod performance one week before (baseline recording) and one day after the experimental procedure was recorded for all mice in study 1 (see methods section 2.2.1) and can be seen in figures 3.49, 3.51, 3.54 and 3.56 for the different groups. Half of the mice were additionally tested two weeks post-experiment. Changes in the performance of these mice with the additional time point can be seen in graphs 3.50, 3.52, 3.55 and 3.57.

a) One month old SOD1^{G93A} and C57Bl/6 mice

The performance of one month old SOD1^{G93A} and control mice before and one day after they have undergone either the active or 'sham' types of recording can be seen in figure 3.49. There was no significant change in the rotarod performance of any of the groups. Baseline measurements for all four groups were mostly between 200 and 300s and were variable. After the experimental procedure the SOD1^{G93A} animals that had undergone both the active procedure and the 'sham' one performed slightly better the day after the recordings (insignificant difference). C57Bl/6 mice that had the active type of recording performed slightly worse one day after the experiment whereas the mice that had the 'sham' procedure performed slightly better. Thus, there was no consistent change in the performance of the mice or the variability of the performance across the four different groups.

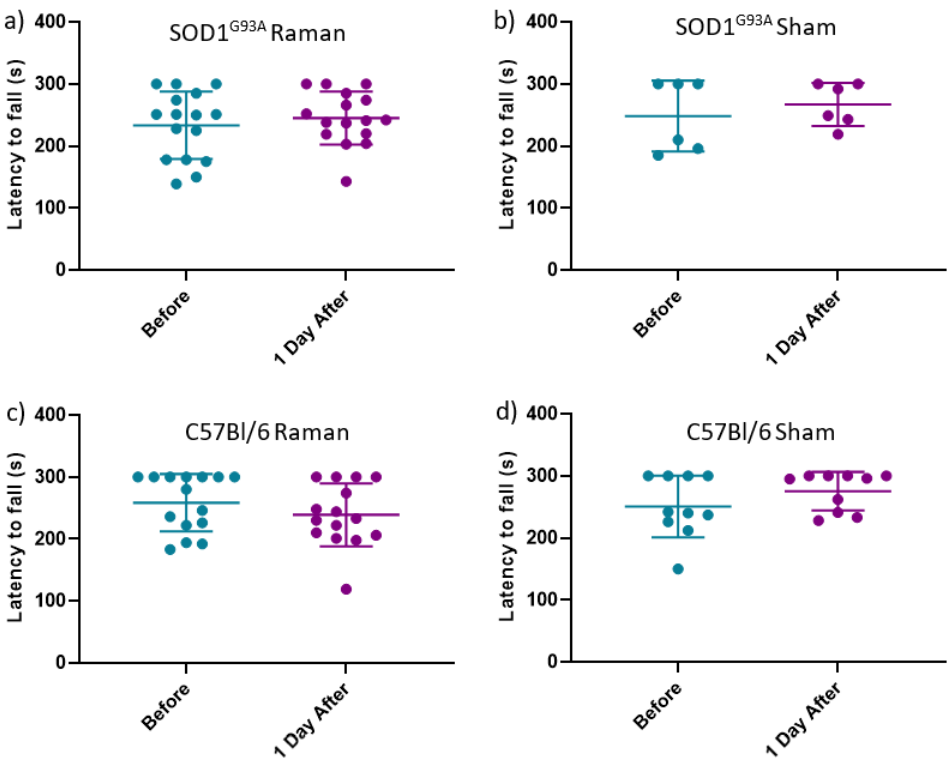


Figure 3.49: Rotarod performance for one month old SOD1^{G93A} and C57Bl/6 mice.
Rotarod performance measured as time to fall in seconds (y axis) at two time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Mean of each group +/- standard deviation are also shown.

Similarly, the mice underwent recordings at three time-points did not exhibit any significant change in their rotarod performance (figure 3.50). The performance of SOD1^{G93A} mice that had undergone the 'sham' procedure improved with time whereas that of the mice that had undergone the active recordings declined at the measurements taken two weeks after the recording. The control mice performed slightly better two weeks post-experiment from the previous time points.

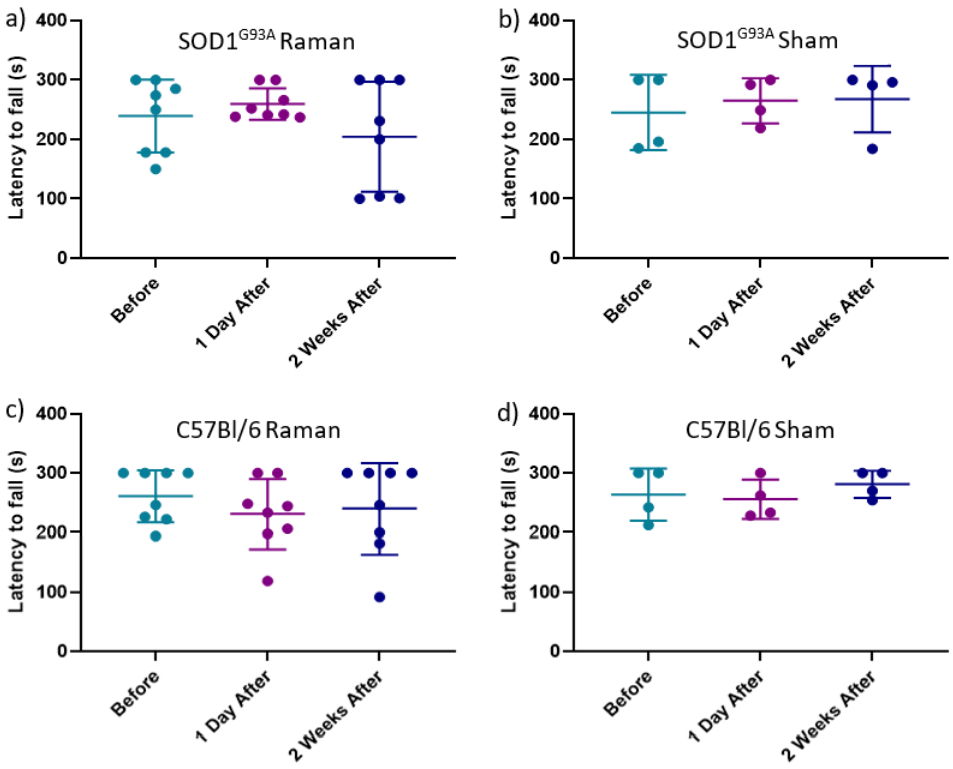


Figure 3.50: Rotarod performance for one month old SOD1^{G93A} and C57Bl/6 mice.
Rotarod performance measured as time to fall in seconds (y axis) at three time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Mean of each group +/- standard deviation are also shown.

b) Three months old SOD1^{G93A} and C57Bl/6 mice

Rotarod performance of all different sub-groups of three months old SOD1^{G93A} and C57Bl/6 mice at two time points is shown in figure 3.51. Baseline measurements for SOD1^{G93A} mice were mostly between 100s and 200s, for C57Bl/6 mice measurements were between 200s and 300s. The performance of C57Bl/6 animals before and one day after any of the experimental procedures remained almost unchanged, with most of the mice remaining on the rotarod for the whole five minute interval at both time points. Baseline measurements for SOD1^{G93A} animals were worse compared to C57Bl/6 mice as expected due to the disease stage. Like the control animals, the experimental procedure did not affect significantly the SOD1^{G93A} mice performance. However, the dispersion of the individual performances in both baseline and post-experiment recordings was much larger for SOD1^{G93A} animals than for NTg littermate controls.

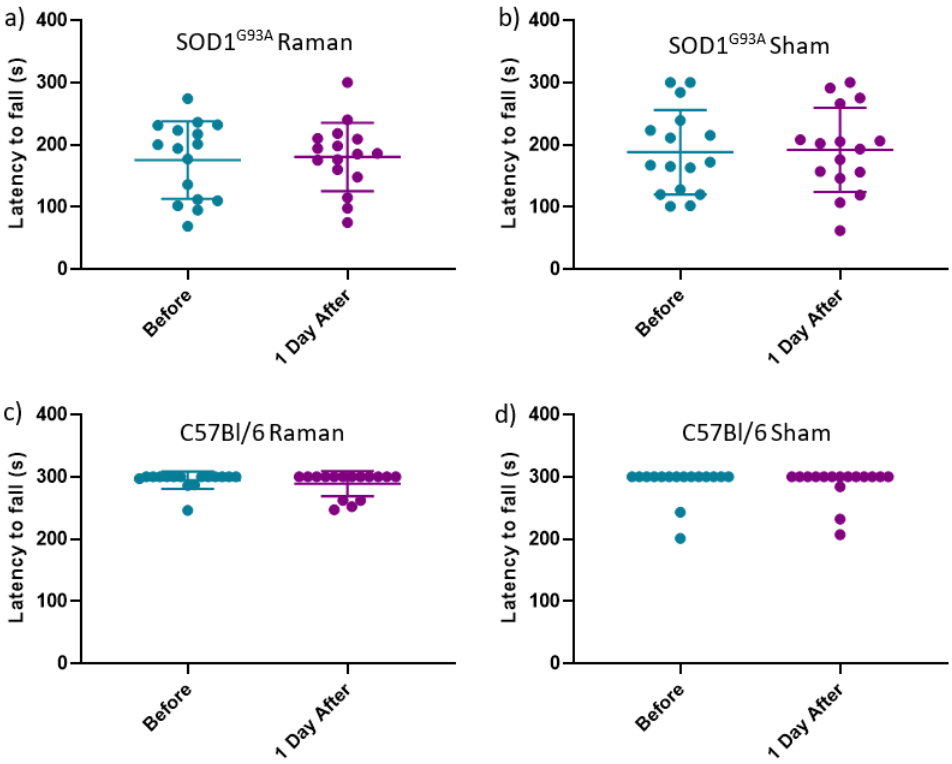


Figure 3.51: Rotarod performance for three months old SOD1^{G93A} and C57Bl/6 mice.
Rotarod performance measured as time to fall in seconds (y axis) at two time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Mean of each group +/- standard deviation are also shown in graphs (a)-(c). Median and interquartile range are shown in graph (d).

C57Bl/6 animals, whose rotarod performance was recorded at three time points (figure 3.52), exhibited a very similar behaviour before and after the experiment, with most of the mice staying on the rotarod for more than 280s at all three time points and both procedures. By contrast, performance of the SOD1^{G93A} mice exhibited a gradual decline after both experimental procedures, which was statistically significant in the active Raman group.

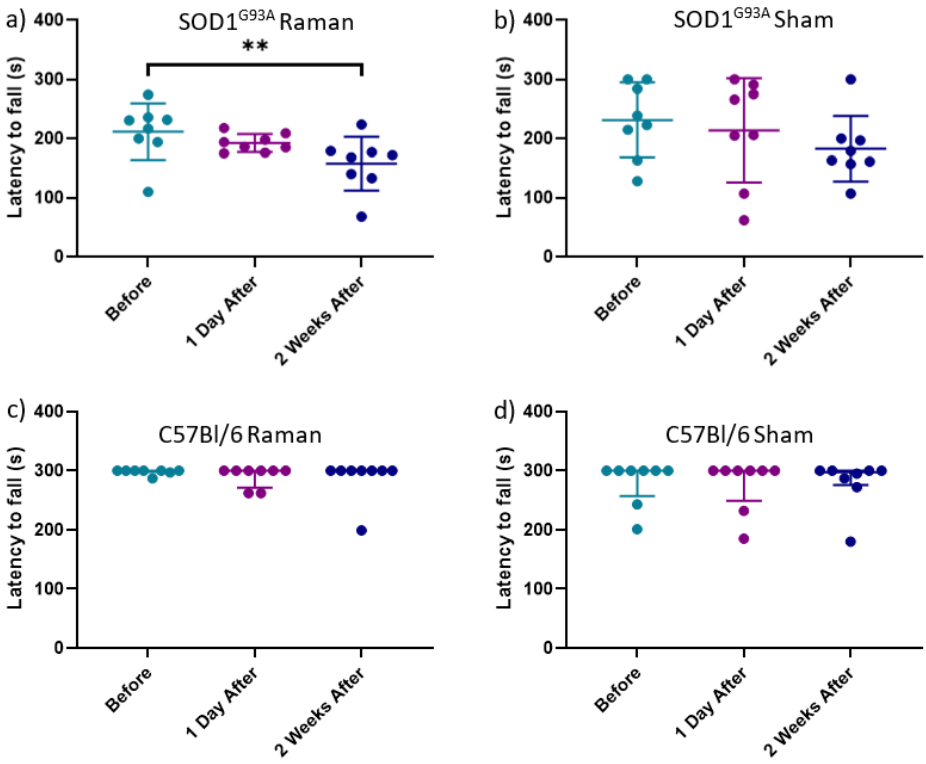


Figure 3.52: Rotarod performance for three months old SOD1^{G93A} and C57Bl/6 mice.
Rotarod performance measured as time to fall in seconds (y axis) at three time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Mean of each group +/- standard deviation are also shown in graphs (a)-(b). Median and interquartile range are shown in graphs (c)-(d). Asterisks indicate p < 0.01 (**).

Since the drop in the performance was not as significant in the SOD1^{G93A} that underwent the 'sham' procedure we attempted to get a better understanding of how these two procedures affect motor function at this late disease stage and how much of the observed change was due to disease progression by recording the rotarod performance of mice at 104 days of age (n=8) that did not have any procedure (figure 3.53).

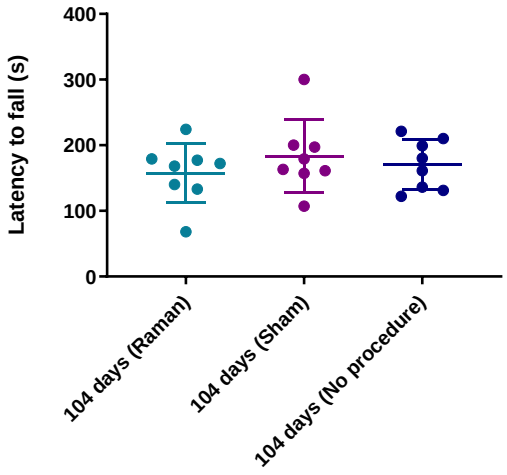


Figure 3.53: Rotarod performance for 104 days old SOD1^{G93A} mice.
Rotarod performance measured as time to fall in seconds (y axis) at three time points. Mean and (+/) standard deviation are shown for each group.

Ordinary one-way ANOVA did not identify a significant difference between the performances of the mice that had undergone any of the two procedures and the 'control' mice.

c) One month old *mdx* and C57Bl/10 mice.

Rotarod recordings of *mdx* and control one month old mice before and one day after they have undergone either the active or 'sham' type of recording are shown in figure 3.54. There was no significant change in the rotarod performance of any of the groups. With the exception of two control and six *mdx* mice, all animals were able to run the rotarod for five minutes before the experimental procedure. After the 'sham' procedure more animals were able to complete the rotarod test. Similarly, the mice that underwent the active type of Raman measurements perform either in the same way or even better after the experiment.

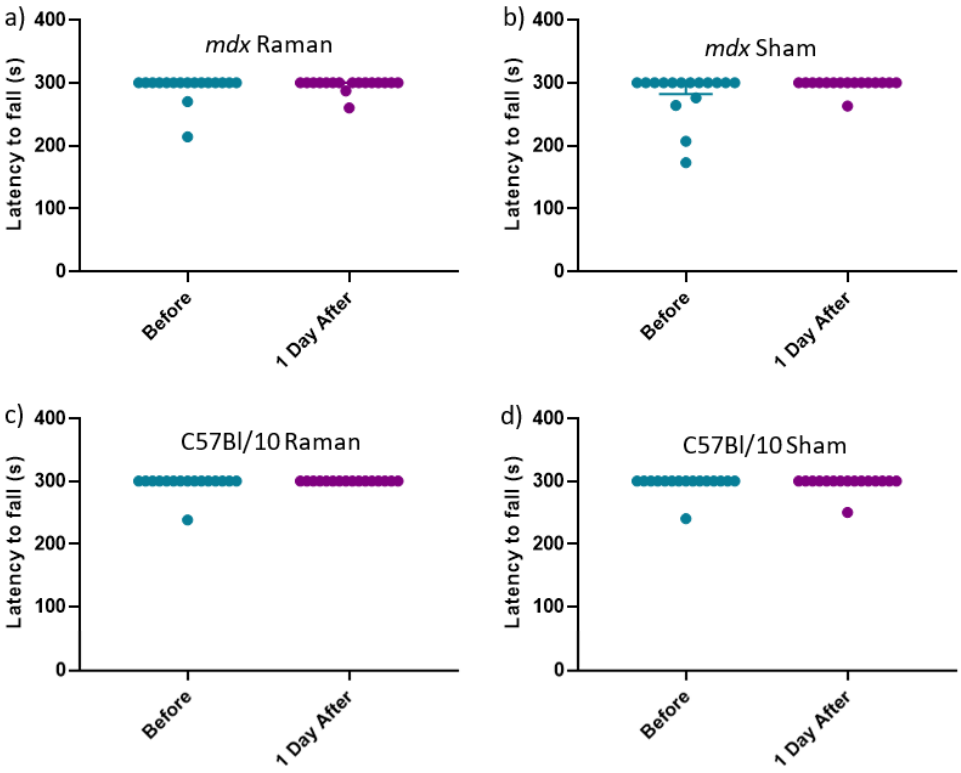


Figure 3.54: Rotarod performance for one month old mdx and C57Bl/10 mice.
Rotarod performance measured as time to fall in seconds (y axis) at two time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Median of each group and interquartile range are also shown.

The mice that underwent the additional rotarod assessment at two weeks after the experiment did not exhibit any significant impairment in their rotarod performance that could be attributed to the experimental procedure (figure 3.55). Most of the mice were able to stay on the rotarod for 300s at the third time point. Performance of *mdx* mice was more variable compared to control animals.

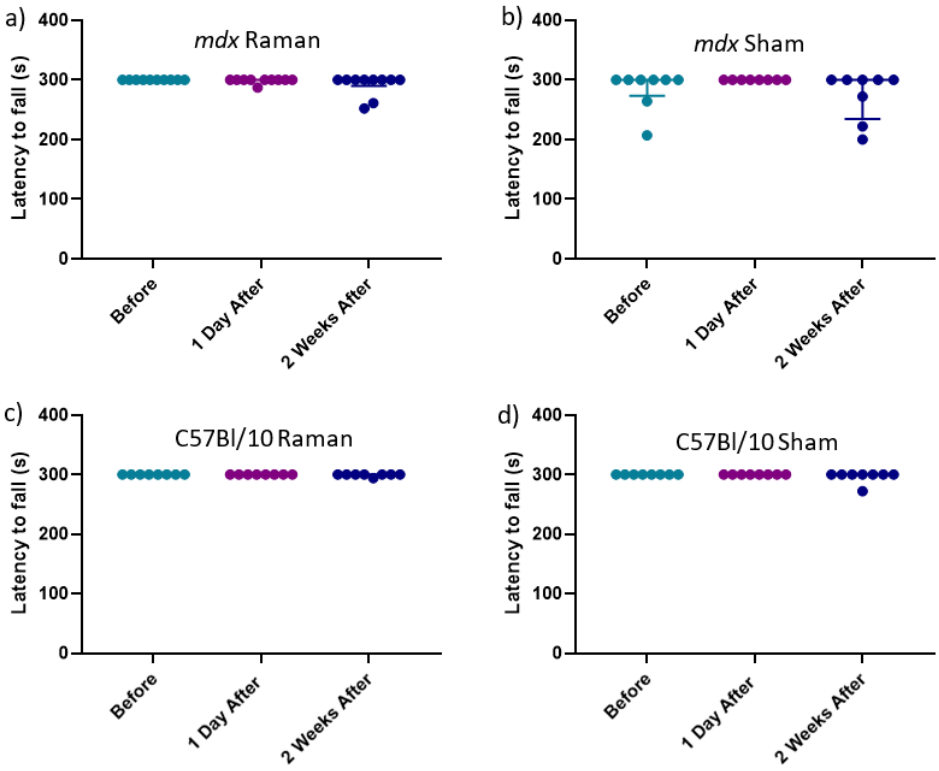


Figure 3.55: Rotarod performance for one month old mdx and C57Bl/10 mice.
Rotarod performance measured as time to fall in seconds (y axis) at three time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Median of each group and interquartile range are also shown in graphs.

d) Three months old *mdx* and C57Bl/10 mice

Rotarod performance of 3 months old *mdx* and C57Bl/10 mice are shown in figures 3.56 and 3.57. Almost all control animals were able to complete the rotarod test before and after the experimental procedure (both post-experiment time points) with only one mouse performing worse one day after the active Raman recording (fig. 12(c)). Thus, there was no significant change or variability in the rotarod performance of the control mice. Baseline rotarod recordings of *mdx* mice were more inconsistent, with the median value, however, being close to 300s for all different groups. After both types of procedure there was no significant or consistent impairment in the ability of the mice to perform the rotarod test with more mice being able to complete the test. For all *mdx* mice the consistency of the rotarod recordings increased at the second post-experiment time point.

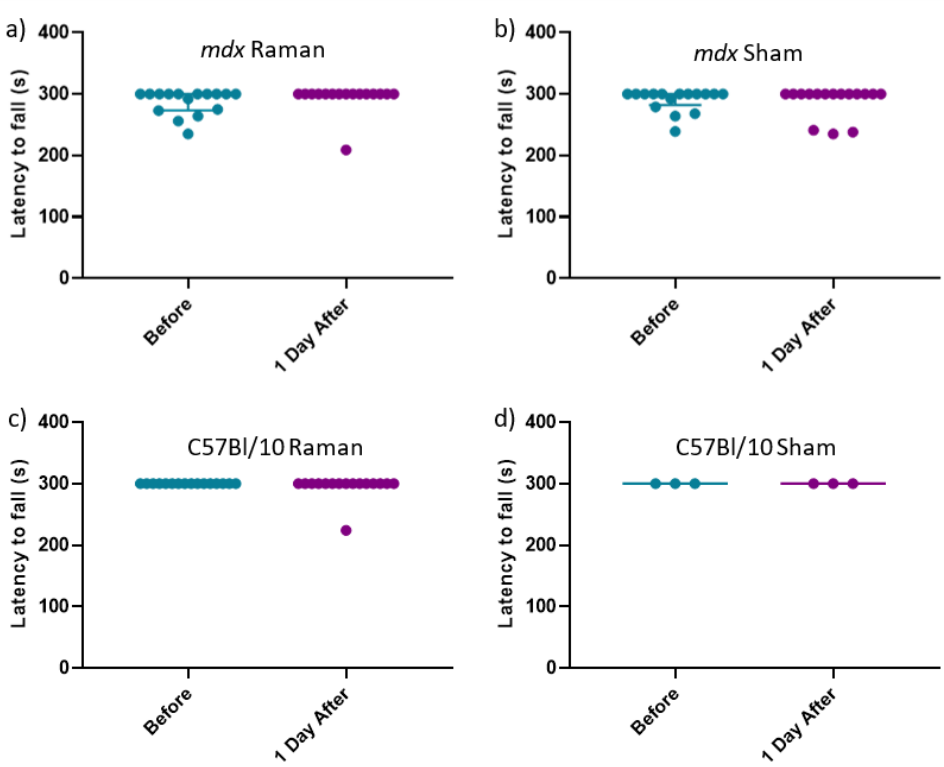


Figure 3.56: Rotarod performance for three months old mdx and C57Bl/10 mice.
Rotarod performance measured as time to fall in seconds (y axis) at two time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Median of each group and interquartile range are also shown.

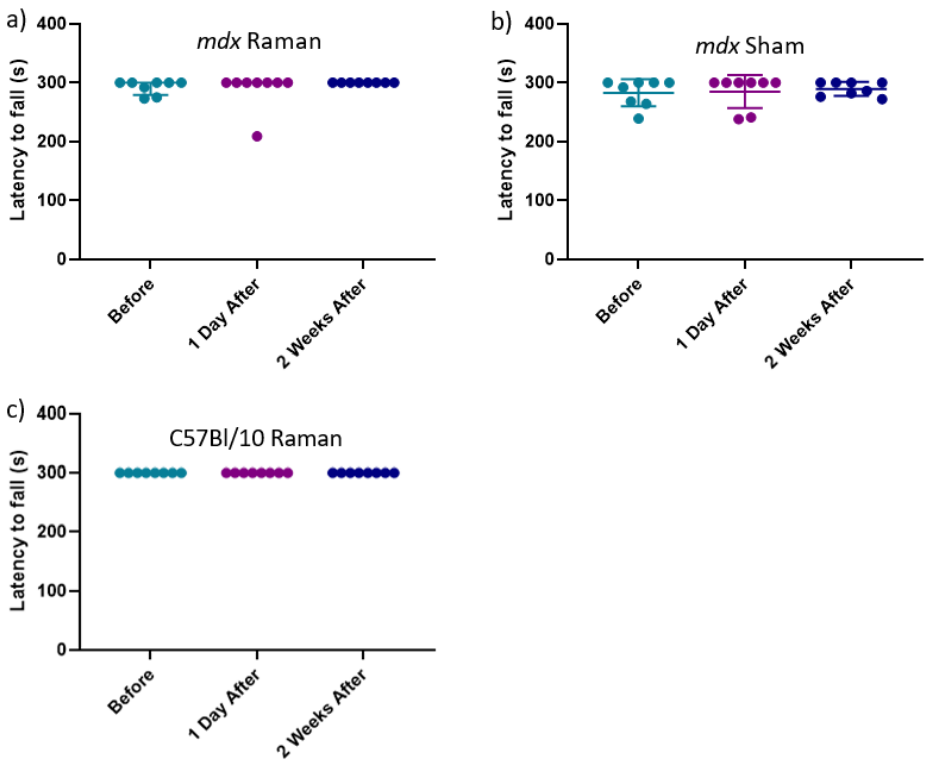


Figure 3.57: Rotarod performance for three months old mdx and C57Bl/10 mice.
Rotarod performance measured as time to fall in seconds (y axis) at three time points. Genotype and type of procedure are indicated in each graph. Dots indicate individual performances within each group and time point. Median of each group and interquartile range are also shown.

3.1.2.4.1 Catwalk

The Catwalk system was used to study the locomotor performance of all the mice in Study 1 (see methodology chapter 2.1.2). In order to assess if the injury caused by the Raman (active and 'sham') procedure led to any particular gait changes, all mice underwent the catwalk test a week before (baseline recording) and one day after the experimental procedure. A third recording, two weeks post-experiment, was acquired from half of the mice.

All statistics are displayed in tables 3.32-3.35. Only adjusted P-values (q-values) from the comparisons that yielded a significant result are shown in the tables. Results from the statistical tests between the baseline recordings and one day post-experiment recordings are shown in tables 3.32 and 3.33 for the Raman and 'sham' procedures respectively. Results between the baseline recordings and recordings acquired two weeks post-experiment are shown in tables 3.34 (Raman procedure) and 3.35 ('sham' procedure). Overall, there were no consistent gait changes observed across the different mice groups at any time point and for any procedure. Changes in a small number of gait parameters (either one day or two weeks after the procedure) were significant in some of the groups that underwent the active Raman procedure. There was no significant difference in locomotor performance for any of the groups that underwent the 'sham' procedure.

One month old *mdx* mice that underwent the Raman experimental procedure exhibited a smaller forelimb base of support and a decrease in the duty cycle of the LF/RH paw ratio a day after the experimental procedure compared to the baseline recording. The observed ratio difference could be due to a decrease in the duty cycle of the left front paw is on the floor, due to an increase in the time duty cycle of the right hind paw, or both. These parameters returned to pre-injury levels two weeks post-experiment and no other parameters changed significantly at this time point for this group.

The percentage of time spent walking with a diagonal step pattern increased significantly in three months old C57Bl/10 mice post-experiment at both time points (one day and two weeks) when compared to the baseline recordings.

Three months old *mdx* mice spent less time being supported on three paws two weeks post-experiment.

Table 3.32: Comparison between gait parameters determined a week before and a day after the Raman experimental procedure using the Catwalk system. *Adjusted P-values (q-values) following FDR correction (Q=0.05) shown when significant (q<0.05). (d): decreased with time, (i): increased with time.*

	1 month old SOD1^{G93A}	1 month old C57Bl/6	1 month old *mdx*	1 month old C57Bl/10	3 months old SOD1^{G93A}	3 months old C57Bl/6	3 months old *mdx*	3 months old C57Bl/10
Duration	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern #	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Ca	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Cb	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Aa	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Ab	NS	NS	NS	NS	NS	NS	NS	NS
Regularity index	NS	NS	NS	NS	NS	NS	NS	NS
BOS - front	NS	NS	***0.02 (d)***	NS	NS	NS	NS	NS
BOS - hind	NS	NS	NS	NS	NS	NS	NS	NS
Print position - right	NS	NS	NS	NS	NS	NS	NS	NS
Print position - left	NS	NS	NS	NS	NS	NS	NS	NS
Support - zero	NS	NS	NS	NS	NS	NS	NS	NS
Support - single	NS	NS	NS	NS	NS	NS	NS	NS
Support - diagonal	NS	NS	NS	NS	NS	NS	NS	***0.005(i)***
Support - lateral	NS	NS	NS	NS	NS	NS	NS	NS
Support - gridle	NS	NS	NS	NS	NS	NS	NS	NS
Support - three	NS	NS	NS	NS	NS	NS	NS	NS
Support - four	NS	NS	NS	NS	NS	NS	NS	NS
Initial contact - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Initial contact - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Max contact - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Max contact - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Max area - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Max area - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Intensity - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Intensity - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Print width - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Print width - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Print length - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Print length - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Print area - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Print area - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Stand time - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Stand time - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Swing - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Swing - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Stride length - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Stride length - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Duty cycle - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Duty cycle - LF/RH	NS	NS	***0.02 (d)***	NS	NS	NS	NS	NS
Swing speed - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Swing speed - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Stand index - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS

Stand index - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS

Table 3.33: Comparison between gait parameters determined a week before and a day after the 'sham' experimental procedure using the Catwalk system. *The t-tests were not carried out for one group due to small number of animals.*

	1 month old SOD1^{G93A}	**1 month old C57Bl/6**	**1 month old *mdx***	**1 month old C57Bl/10**	**3 months old SOD1^{G93A}**	**3 months old C57Bl/6**	**3 months old *mdx***	**3 months old C57Bl/10**
Duration	NS	NS	NS	NS	NS	NS	NS	
Step pattern Ca	NS	NS	NS	NS	NS	NS	NS	
Step pattern Cb	NS	NS	NS	NS	NS	NS	NS	
Step pattern Aa	NS	NS	NS	NS	NS	NS	NS	
Step pattern Ab	NS	NS	NS	NS	NS	NS	NS	
Regularity index	NS	NS	NS	NS	NS	NS	NS	
BOS - front	NS	NS	NS	NS	NS	NS	NS	
BOS - hind	NS	NS	NS	NS	NS	NS	NS	
Print position - right	NS	NS	NS	NS	NS	NS	NS	
Print position - left	NS	NS	NS	NS	NS	NS	NS	
Support - zero	NS	NS	NS	NS	NS	NS	NS	
Support - single	NS	NS	NS	NS	NS	NS	NS	
Support - diagonal	NS	NS	NS	NS	NS	NS	NS	
Support - lateral	NS	NS	NS	NS	NS	NS	NS	
Support - gridle	NS	NS	NS	NS	NS	NS	NS	
Support - three	NS	NS	NS	NS	NS	NS	NS	
Support - four	NS	NS	NS	NS	NS	NS	NS	
Initial contact - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Initial contact - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Max contact - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Max contact - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Max area - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Max area - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Intensity - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Intensity - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Print width - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Print width - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Print length - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Print length - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Print area - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Print area - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Stand time - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Stand time - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Swing - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Swing - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Stride length - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Stride length - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Duty cycle - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Duty cycle - LF/RH	NS	NS	NS	NS	NS	NS	NS	
Swing speed - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Swing speed - LF/RH	NS	NS	NS	NS	NS	NS	NS	

Stand index - RF/LH	NS	NS	NS	NS	NS	NS	NS	
Stand index - LF/RH	NS	NS	NS	NS	NS	NS	NS	

Table 3.34: Comparison between gait parameters determined a week before and two weeks after the Raman experimental procedure using the Catwalk system. *Adjusted P-values (q-values) following FDR correction (Q=0.05) shown when significant (q<0.05). (d): decreased with time, (i): increased with time.*

	1 month old SOD1^{G93A}	**1 month old C57Bl/6**	**1 month old *mdx***	**1 month old C57Bl/10**	**3 months old SOD1^{G93A}**	**3 months old C57Bl/6**	**3 months old *mdx***	**3 months old C57Bl/10**
Duration	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Ca	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Cb	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Aa	NS	NS	NS	NS	NS	NS	NS	NS
Step pattern Ab	NS	NS	NS	NS	NS	NS	NS	NS
Regularity index	NS	NS	NS	NS	NS	NS	NS	NS
BOS - front	NS	NS	NS	NS	NS	NS	NS	NS
BOS - hind	NS	NS	NS	NS	NS	NS	NS	NS
Print position - right	NS	NS	NS	NS	NS	NS	NS	NS
Print position - left	NS	NS	NS	NS	NS	NS	NS	NS
Support - zero	NS	NS	NS	NS	NS	NS	NS	NS
Support - single	NS	NS	NS	NS	NS	NS	NS	NS
Support - diagonal	NS	NS	NS	NS	NS	NS	NS	***0.021 (i)***
Support - lateral	NS	NS	NS	NS	NS	NS	NS	NS
Support - gridle	NS	NS	NS	NS	NS	NS	NS	NS
Support - three	NS	NS	NS	NS	NS	NS	***0.004 (d)***	NS
Support - four	NS	NS	NS	NS	NS	NS	NS	NS
Initial contact - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Initial contact - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Max contact - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Max contact - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Max area - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Max area - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Intensity - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Intensity - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Print width - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Print width - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Print length - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Print length - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Print area - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Print area - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Stand time - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Stand time - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Swing - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Swing - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Stride length - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Stride length - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Duty cycle - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Duty cycle - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Swing speed - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS

Swing speed - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS
Stand index - RF/LH	NS	NS	NS	NS	NS	NS	NS	NS
Stand index - LF/RH	NS	NS	NS	NS	NS	NS	NS	NS

Table 3.35: Comparison between gait parameters determined a week before and two weeks after the 'sham' experimental procedure using the Catwalk system. *Adjusted P-values (q-values) following FDR correction (Q=0.05) shown when significant (q<0.05).*

	1 month old SOD1^{G93A}	**1 month old C57Bl/6**	**1 month old *mdx***	**1 month old C57Bl/10**	**3 months old SOD1^{G93A}**	**3 months old C57Bl/6**	**3 months old *mdx***	**3 months old C57Bl/10**
Duration			NS	NS	NS	NS	NS	
Step pattern Ca			NS	NS	NS	NS	NS	
Step pattern Cb			NS	NS	NS	NS	NS	
Step pattern Aa			NS	NS	NS	NS	NS	
Step pattern Ab			NS	NS	NS	NS	NS	
Regularity index			NS	NS	NS	NS	NS	
BOS - front			NS	NS	NS	NS	NS	
BOS - hind			NS	NS	NS	NS	NS	
Print position - right			NS	NS	NS	NS	NS	
Print position - left			NS	NS	NS	NS	NS	
Support - zero			NS	NS	NS	NS	NS	
Support - single			NS	NS	NS	NS	NS	
Support - diagonal			NS	NS	NS	NS	NS	
Support - lateral			NS	NS	NS	NS	NS	
Support - gridle			NS	NS	NS	NS	NS	
Support - three			NS	NS	NS	NS	NS	
Support - four			NS	NS	NS	NS	NS	
Initial contact - RF/LH			NS	NS	NS	NS	NS	
Initial contact - LF/RH			NS	NS	NS	NS	NS	
Max contact - RF/LH			NS	NS	NS	NS	NS	
Max contact - LF/RH			NS	NS	NS	NS	NS	
Max area - RF/LH			NS	NS	NS	NS	NS	
Max area - LF/RH			NS	NS	NS	NS	NS	
Intensity - RF/LH			NS	NS	NS	NS	NS	
Intensity - LF/RH			NS	NS	NS	NS	NS	
Print width - RF/LH			NS	NS	NS	NS	NS	
Print width - LF/RH			NS	NS	NS	NS	NS	
Print length - RF/LH			NS	NS	NS	NS	NS	
Print length - LF/RH			NS	NS	NS	NS	NS	
Print area - RF/LH			NS	NS	NS	NS	NS	
Print area - LF/RH			NS	NS	NS	NS	NS	
Stand time - RF/LH			NS	NS	NS	NS	NS	
Stand time - LF/RH			NS	NS	NS	NS	NS	
Swing - RF/LH			NS	NS	NS	NS	NS	
Swing - LF/RH			NS	NS	NS	NS	NS	
Stride length - RF/LH			NS	NS	NS	NS	NS	
Stride length - LF/RH			NS	NS	NS	NS	NS	
Duty cycle - RF/LH			NS	NS	NS	NS	NS	
Duty cycle - LF/RH			NS	NS	NS	NS	NS	

Swing speed - RF/LH			NS	NS	NS	NS	NS	
Swing speed - LF/RH			NS	NS	NS	NS	NS	
Stand index - RF/LH			NS	NS	NS	NS	NS	
Stand index - LF/RH			NS	NS	NS	NS	NS	

3.1.2.5 Post-experiment Tissue Damage Assessment

In order to assess potential tissue damage caused by the *in vivo* Raman measurements the legs of a small number of mice were scanned after the experimental procedure. Optimisation of the geometry of the sample and the scanning parameters was necessary in order to be able to understand if hyper-intense regions detected in the MRI images could be attributed to the experimental procedure and differentiate them from artefacts.

3.1.2.5.1 Optimisation of the Sample Geometry

Testing scans were initially performed on dissected gastrocnemius muscles after the experimental procedure (figure 3.58 (a)). However, it was difficult to orientate the muscle and thus understand where the potential damage could be. To try and overcome this problem we next scanned whole legs. As shown in figure 3.58 (b) the anatomy of the legs in these scans was easier to understand in the axial MRI images and therefore identification of the region of interest was possible (region of potential insertion and thus muscle injury). Moreover, by removing the skin from the leg, the muscle looked much clearer (particularly in the axial planes) and the localisation of the area of interest was further facilitated (data not shown).

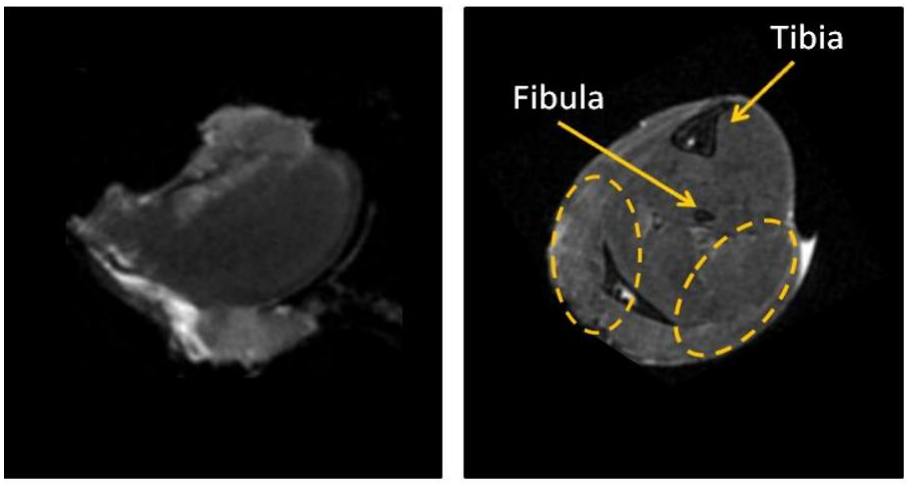


Figure 3.58: Different sample geometries.
MRI scan of dissected muscle (a) and axial plane image of the whole leg scan (b). Tibia and fibula helped orientation on the axial planes of the whole leg scans and thus identifying the regions of interest, shown in yellow circles.

3.1.2.5.2 Optimisation of Scanning Parameters

Since it was not clear if the procedure caused tissue damage and how significant it might be, multiple scans were required in order to optimise the scanning parameters and establish a

protocol. Legs from 5 mice were scanned after the Raman or 'sham' experimental procedure using T2 weighted spin echo based imaging techniques in order to establish the protocol. Multi-slice multi-echo (MSME) and RARE pulse sequences with multiple echo and repetition times were tested. Moreover, experimentation with resolution, matrix size and field of view was also required. Echo times of the multiple trial scans ranged from 6ms to 240ms whereas repetition times ranged from 1000ms to 14500ms. Field of view, spatial resolution and slice thickness varied between 2cm*2cm - 3cm*3cm, 0.3mm*0.3mm - 0.078mm*0.078mm, and 1mm-0.3mm respectively. Matrix sizes of 128*128 and 256*256 were tried. The legs were also scanned with and without fat suppression.

After the trial scans the protocol comprised an initial tri-plane FLASH sequence low-resolution scan for subject localisation, a fast RARE sequence scan for low signal to noise ratio (SNR) visualisation of the region of interest and planning of axial high resolution fast spin echo images (RARE sequence) and the final high resolution RARE scans. Fat suppression was used in order to avoid the intense fat signal. The final protocols are described in the methods section 2.1.5.1 in more detail.

3.1.2.3.3 MRI Scans

Post-mortem MRI scans were performed on gastrocnemius muscles (both legs) of nine mice that had undergone the active procedure and 4 mice that had undergone the 'sham' procedure for tissue damage assessment with the established protocol. To understand and assess the potential muscle injury and its evolution, mouse legs were scanned six hours, two days and two weeks post-experiment.

In figure 3.59 axial planes of the MRI scans at the three different time points from the mice that had undergone the active procedure are shown. In these images, hyper-intense (i.e. bright) regions that could be attributed to an inflammatory response caused by the recording can be seen in all three scans at six hours, in one scan two days after the experiment but not at two weeks (figure 3.59, yellow arrows). Inflammatory response after the muscle injury causes oedema in the tissue. This increases the T2 relaxation time of this tissue area, which therefore would look brighter in T2 weighted images than the surrounding tissue. The sites of the hyper-intense regions (about 2 mm distance from the edges) and the fact that in some cases they can be seen in both heads of the gastrocnemius muscle suggest that they could be attributed to muscle injury from the recordings. The area indicated by the yellow arrow in the third six-hours-post-injury scan could be a low signal lesion caused by the needle insertion and identified as an air bubble. Air has a very short T2 time and thus appears black in T2 images. Around the area with the dark signal ('air bubble') there is a hyper-intense region that could be due to oedema formation around the insertion site.

Axial planes of the MRI scans from the mice that had undergone the 'sham' procedure can be seen in figure 3.60 for two post-experiment time points (six hours and two days). There were no areas that could be attributed to oedema formation in any of the four scans.

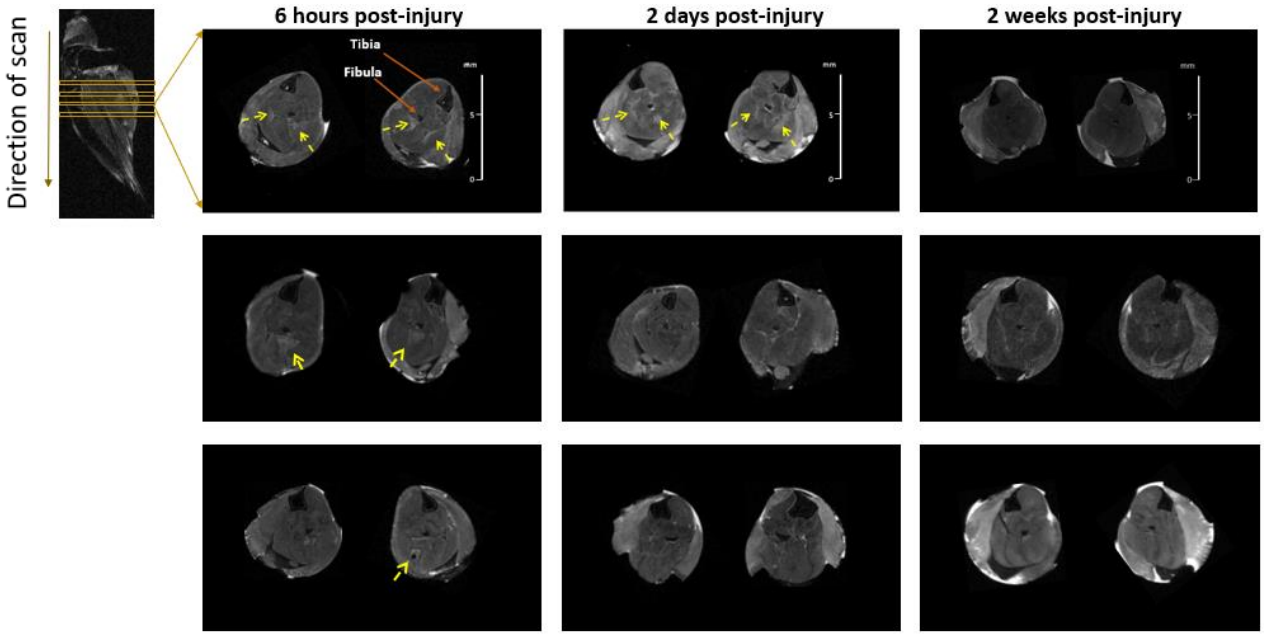


***Figure 3.59: Axial MRI images at three different time points after the active Raman procedure**.*
The yellow arrows indicate the hyper-intense regions that could be attributed to the experimental procedure. Such areas can be seen at six hours and two days but not at two weeks.

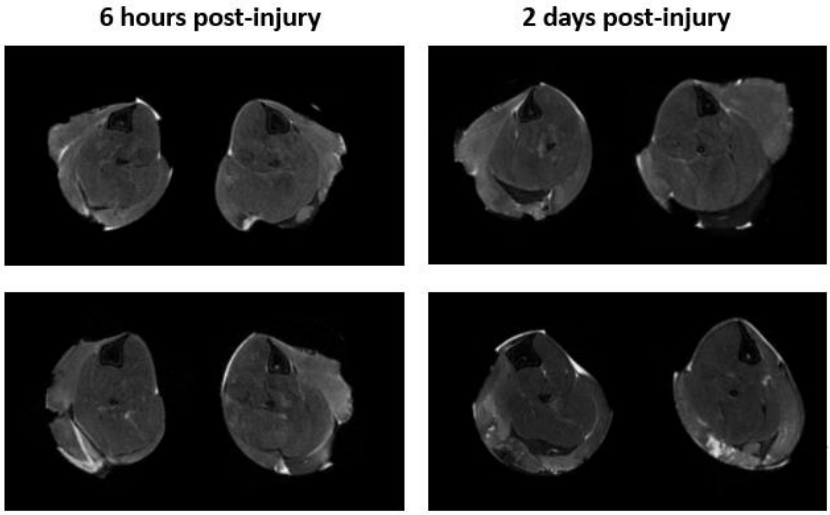


***Figure 3.60: Axial MRI images at two different time points after the sham Raman procedure**.*

3.1.3 Study 2: Intervention Study in *mdx* mice

3.1.3.1 Running Wheel Performance and Creatine Kinase (CK) Measurements

A single 48-hour epoch of voluntary running exercise was used to exacerbate muscle pathology in *mdx* mice. Two groups of three months old female *mdx* mice underwent the running wheel exercise (see methods section). As can be seen in figure 3.61 the two subgroups had a very similar overall performance, with animals in both groups running an average of 6.5km. However, the performance of subgroup 1 was more variable than that of subgroup 2. Total distance run by different animals in subgroup 1 ranged from 0.4 to 21.2 km, whereas in subgroup 2 ranged from 2.1 to 10.3.

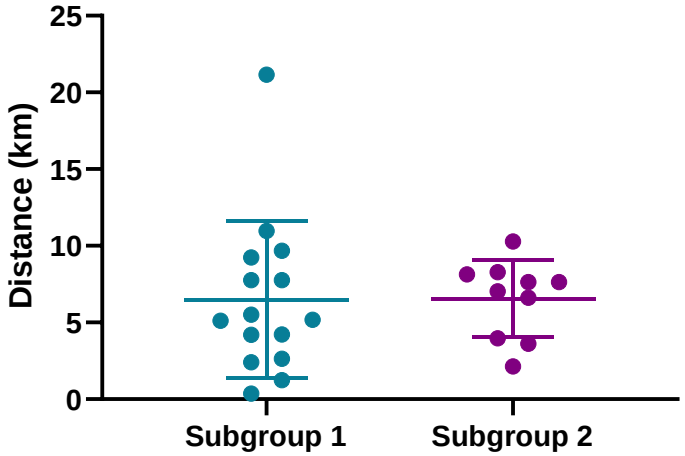


Figure 3.61: Running wheel performance of two different groups of three months old mdx mice. *The different groups ($n_{subgroup\ 1}$=15, subgroup 2 $n_{subgroup\ 2}$=10) had a similar average performance. Mean and standard deviation shown.*

After the running exercise serum CK activity was measured from mice in subgroup 2 (exercised mice) and a group of non-exercised *mdx* mice as a broad assessment of muscle damage. Creatine kinase levels of a group of C57Bl/10 mice was also measured for comparison. The results can be seen in figure 3.62.

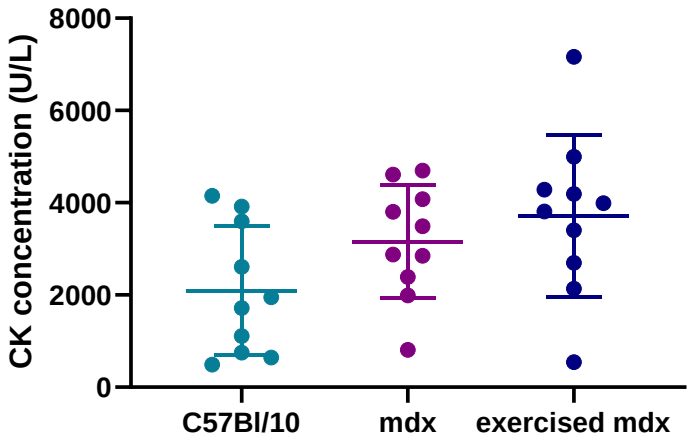


Figure 3.62: Serum CK levels in C57Bl/10 (n=10), non-exercised mdx (n=10) and exercised mdx (n=10) mice.

CK levels were variable for all different groups. As expected C57Bl/10 mice had lower serum CK levels from *mdx* mice and exercised *mdx* mice had the highest levels. However, CK levels of different groups were not significantly different.

There was no correlation between the distance run and measured CK levels of the exercised mice (data not shown).

3.1.3.2 In vivo recordings

PCA-LDA, PCA-QDA and PLS-DA were used to generate classification models. The performance of the models was validated using leave-one-spectrum out (LOOCV), leave-one-mouse-out (LOMOCV) and repeated leave-some-mice-out (RLSMOCV) cross-validation (CV). The results for the two group models cross-validated using LOSOCV and LOMOCV are presented in section 3.1.3.2.1. Repeated cross-validation is presented in section 3.1.3.2.2.

3.1.3.2.1 Two group model

a) Exercised vs. Non-exercised *mdx* mice

The model built using the spectra of three months old exercised and non-exercised *mdx* mice aimed to investigate the ability of the technique to monitor more subtle changes in muscle pathology as are those induced in *mdx* muscle by exercising. The predictive capability of the models using different analysis techniques and CV methods can be seen in Table 3.36. The PCA related models yielded much better specificity than sensitivity values when distinguishing exercised and non-exercised three months old *mdx* mice. This indicates that the models were able to identify more accurately the non-exercised *mdx* mice. The PLS-DA model achieved more balanced sensitivity and specificity values. Using LOMOCV all the indices of the PCA related models decreased considerably. The PLS-DA models achieved similar accuracy with the different CV methods, with an increased sensitivity and decreased specificity when LOMOCV was used. The PLS-DA models were considered to have demonstrated the best performance due to the more balanced sensitivity and specificity indices and the smallest change when different CV methods were used.

The ROC curves for the different techniques and CV methods can be seen in figure 3.63. The AUROC were 0.70 or above for all the models generated using PCA when LOSOCV was used but dropped to 0.66 (LDA) and 0.62 (QDA) for LOMOCV. Both PLS-DA models yielded higher AUC values for both CV methods.

Table 3.36: Two group PCA-LDA model classification model performance evaluated using different CV methods for the exercised and non-exercised mdx mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	61.4%	81.4%	71.6%
Leave-one-mouse-out CV	57.8%	72.9%	65.5%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	61.4%	77.9%	69.8%
Leave-one-mouse-out CV	54.4%	72.2%	62.9%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	71.9%	71.2%	71.6%
Leave-one-mouse-out CV	73.7%	66.1%	69.9

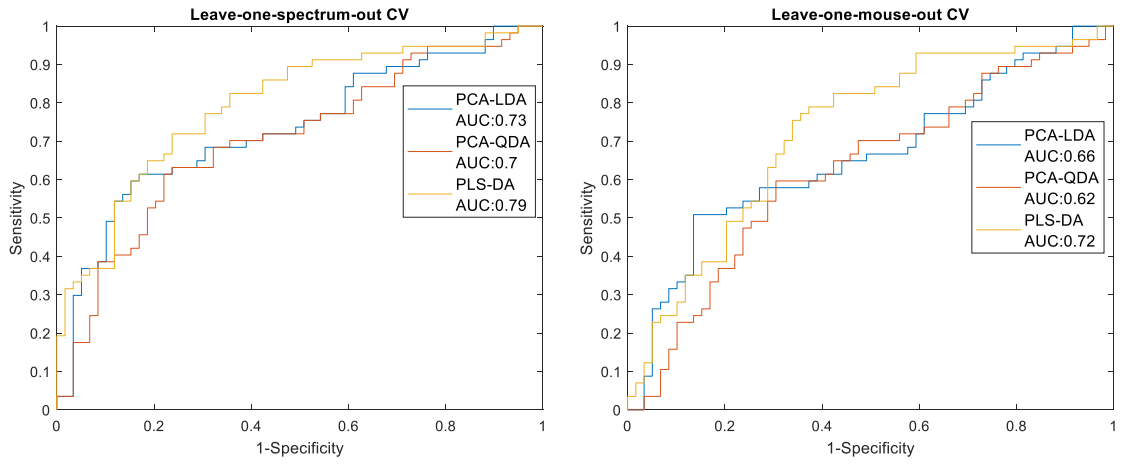


Figure 3.63: ROC curves for the cross validated classification models for the exercised and non-exercised mdx mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

3.1.3.2.2 Repeated cross-validation

a) Exercised vs. Non-exercised *mdx* mice

The classification performance of the RLSMOCV PCA-LDA, PCA-QDA and PLS-DA models for distinguishing between three months exercised and non-exercised mice can be seen in Table 3.37. The PLS-DA model demonstrated the highest sensitivity and accuracy values and was therefore considered to have achieved the best performance.

Table 3.37: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three months old exercised and non-exercised mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

Exercised/Non exercised mdx mice	***Sensitivity (+/- Standard Deviation, CoV)***	***Specificity (+/- Standard Deviation, CoV)***	***Accuracy (+/- Standard Deviation, CoV)***
PCA-LDA	60.1% (+/-2.8, CoV: 4.7%)	65.8% (+/-3.4, CoV: 5.2%)	62.5% (+/-2.0, CoV: 3.2%)
PCA-QDA	55.5% (+/-3.1, CoV: 5.6%)	72.0% (+/-2.9, CoV: 4.0%)	63.9% (+/-2.1, CoV: 3.3%)
PLS-DA	71.3 % (+/-2.0, CoV: 2.8%)	64.8% (+/-3.3, CoV: 5.1%)	68.4% (+/-1.8, CoV: 2.6%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model for the three months old *mdx* mice can be seen in figure 3.64. The PCA-LDA and QDA ROC curves had a mean AUC of 0.7 and 0.67 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.71.

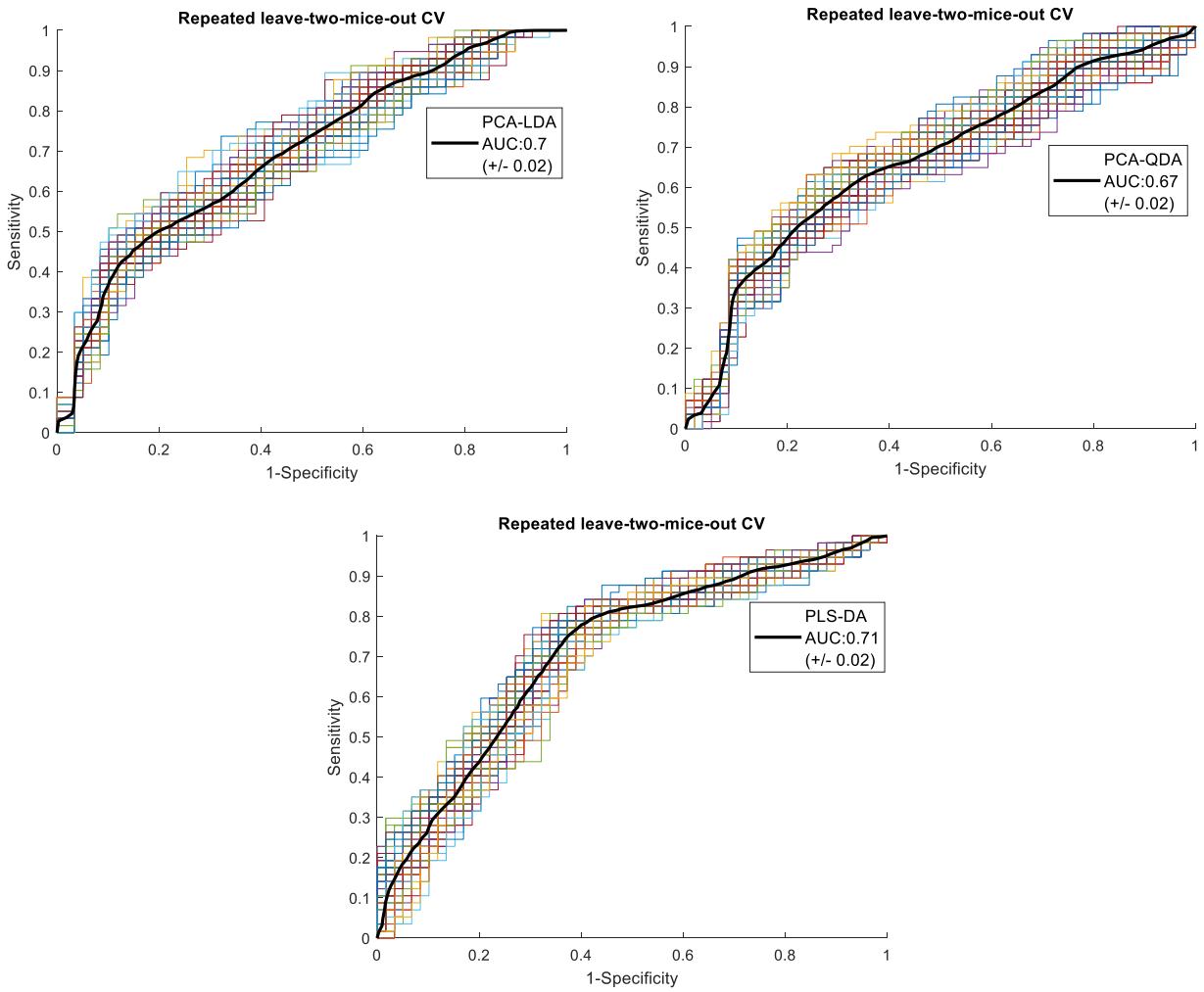


Figure 3.64: ROC curves generated from repeated cross-validation of exercised and non-exercised mdx mice. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

3.1.3.2.2 Basis of classification

3.1.3.2.2.1 Mean and Difference spectra

The mean and difference spectra of the exercised and non-exercised *mdx* mice were plotted to visually examine the most prominent peaks present in the spectra of the different groups. In figure 3.65 the mean spectra of each group (+/-) one standard deviation are shown.

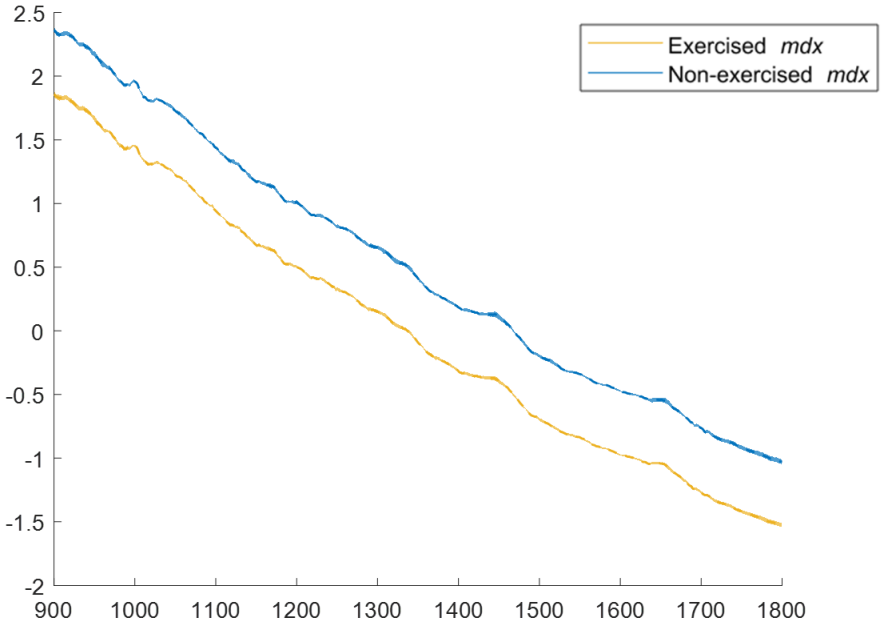


Figure 3.65: Mean Raman spectra for exercised and non-exercised mdx mice groups.
The mean spectra for exercised and non-exercised mdx mice are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

Figure 3.66 shows the mean and difference spectra of the two groups. The mean spectra of the different groups were very similar, with the same prominent peaks being present in the spectra of both groups. A small number of biochemically reliable peaks could be identified in the mean spectra and are summarised in table 3.38.

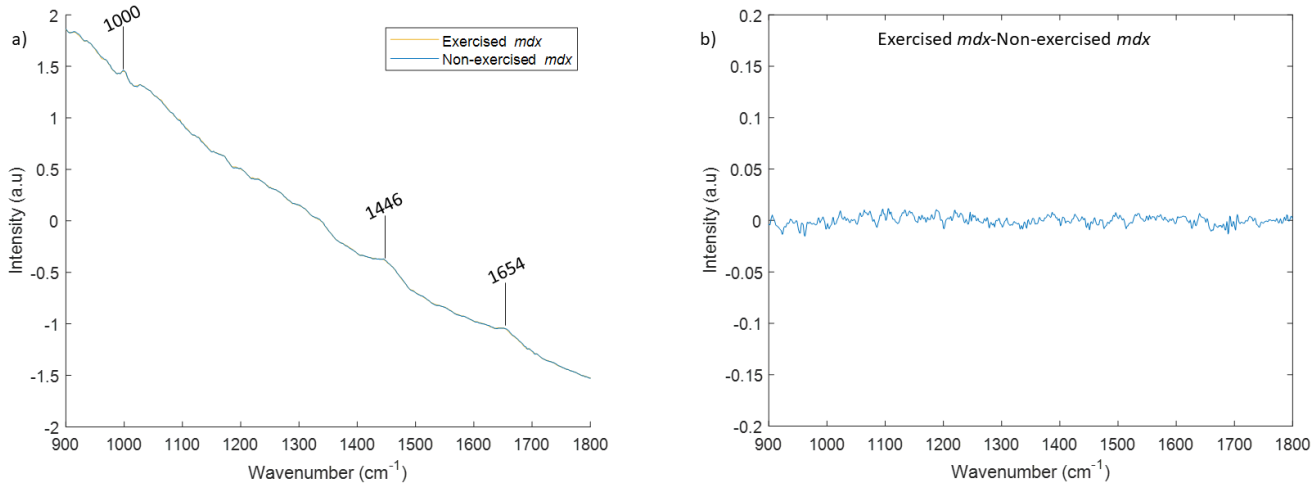


Figure 3.66: Mean and difference spectra of the exercised and non-exercised mdx mice.
a) Mean spectra for one and month old mdx and C57Bl/10 mice and b) difference spectrum. Prominent peaks are indicated in the mean spectrum.

Table 3.38: Prominent Raman peaks in mean spectra of the exercised and non-exercised mdx mice and tentative peak assignments.

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1444	CH_2 bending	Proteins, Lipids
1654	C=O stretching, C=C stretch	Amide I (proteins), Lipids

Background subtraction was used to aid identification of more spectral features present in the spectra. The background subtracted mean spectra of the different groups can be seen in figure 3.67. A larger number of peaks were identifiable in these mean spectra.

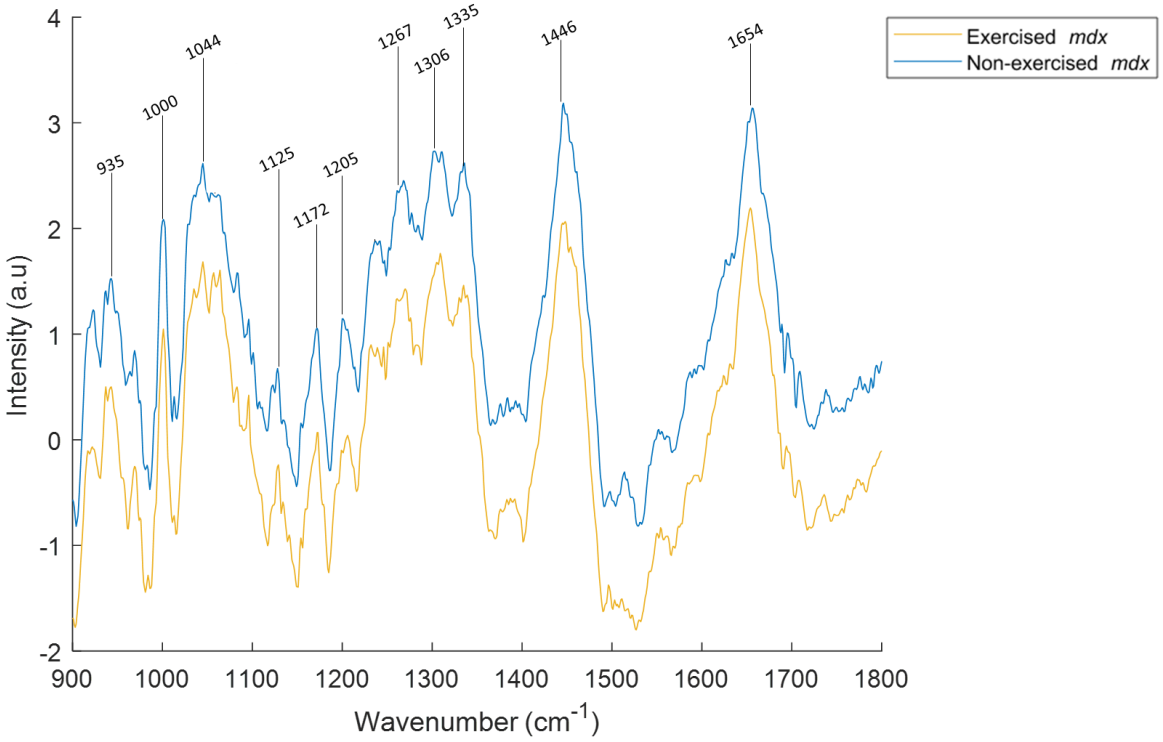


Figure 3.67: Background subtracted mean spectra of exercised and non-exercised mdx mice groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

Additional shoulder peaks can be seed around 950 cm^{-1}, 970 cm^{-1}, 1030 cm^{-1}, 1550 cm^{-1}, 1575 cm^{-1} and 1610 cm^{-1}. Tentative peak assignments for the peaks present in the background subtracted mean spectra are presented in table 3.39.

Table 3.39: Prominent Raman peaks in background subtracted spectra and tentative peak assignments. *Amino acids are specified when the peaks are largely associated with them in the literature.*

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
935	C-C stretching	Proteins
950		Proteins (Valine, Proline, Phenylalanine)
975	CH_2 deformation	Proteins, Nucleic Acids
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1030	C-H bending	Proteins (Phenylalanine, Proline)
1044		Proteins
1125	C-C stretching, C-N stretching	Proteins, Lipid
1170	C-H bending	Proteins (Tyrosine, Phenylalanine)
1205	$\nu(C\text{-}C_6H_5)$	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)
1260	C-N stretching	Amide III (proteins), Lipids
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1444	CH_2 bending	Proteins, Lipids
1550	$\nu(C\text{-}C)$	Proteins (Tryptohan)
1570	Ring breathing modes of the DNA/RNA bases	Nucleic Acids
1615	C=C stretching	Proteins (Tyrosine)
1654	C=O stretching, C=C stretching	Amide I (proteins), Lipids

3.1.3.2.2.2 Multivariate Analysis

As can be seen in the mean and difference figures in the previous section the most prominent peaks were present in the mean spectra of all the different groups and clear visual differences were not easily apparent. Hence, multivariate techniques were employed in order to utilise and further elucidate the biochemical features present in the spectra. The PCA loadings and the linear discriminant function (LDF), as well as the PLS weights, were plotted in order to illustrate the important peaks for spectral classification.

a) Exercised vs. Non-exercised *mdx* mice

Using Student's t-tests followed by fdr correction PC3 and PC6 and PC9 scores were found to be significantly different between the exercised and non-exercised mice. PC3 and PC6 scores were found to have the largest difference between the exercised and non-exercised mice (q_{PC3}:0.004 q_{PC6}: 0.004 and q_{PC9}: 0.02). PC3 and PC6 loading plots can be seen in figure 3.68. The spectra of the two groups were not clearly separated around zero in any of the score histograms. Thus, only the PCA loading plots are shown in this section.

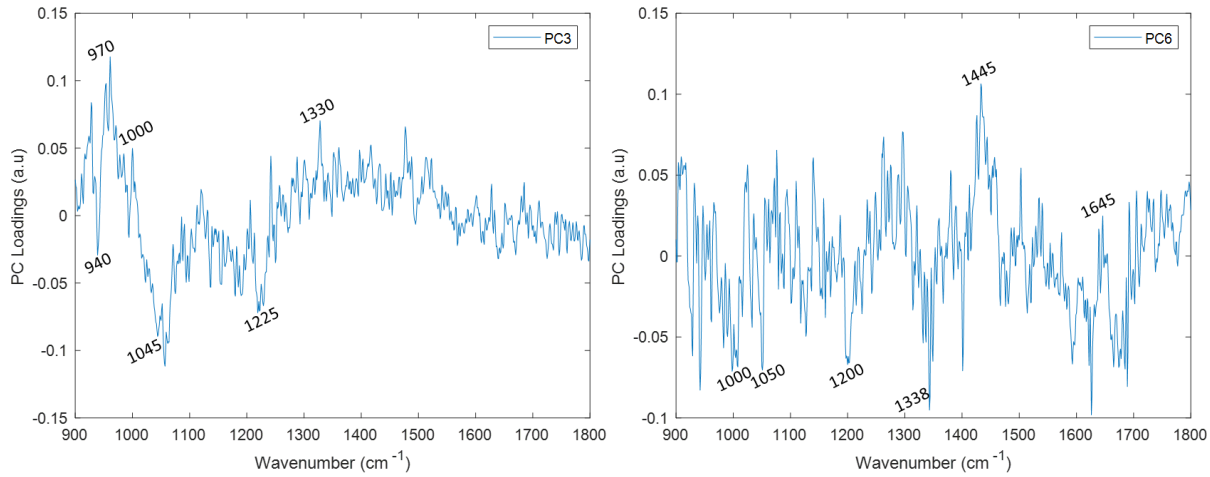


Figure 3.68: PC3 and PC6 and loading plots for the exercised and non-exercised mdx mice.

The LDA histogram and LDF plot using PC3, PC6 and PC9 as input variables can be seen in figure 3.69.

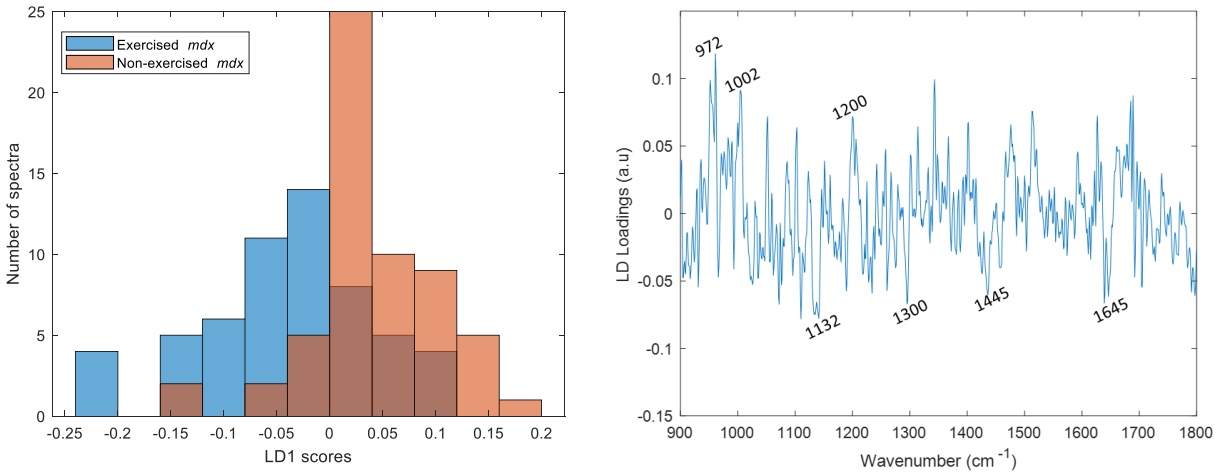


Figure 3.69: LD1 score histogram and LDF plot for the exercised and non-exercised mdx mice.

In the PLS-DA analysis the model generated using the first 2 components demonstrated the highest accuracy value. The spectra were not clearly separated around zero in any of the score histograms. The weight plots of the first two components can be seen in figure 3.70.

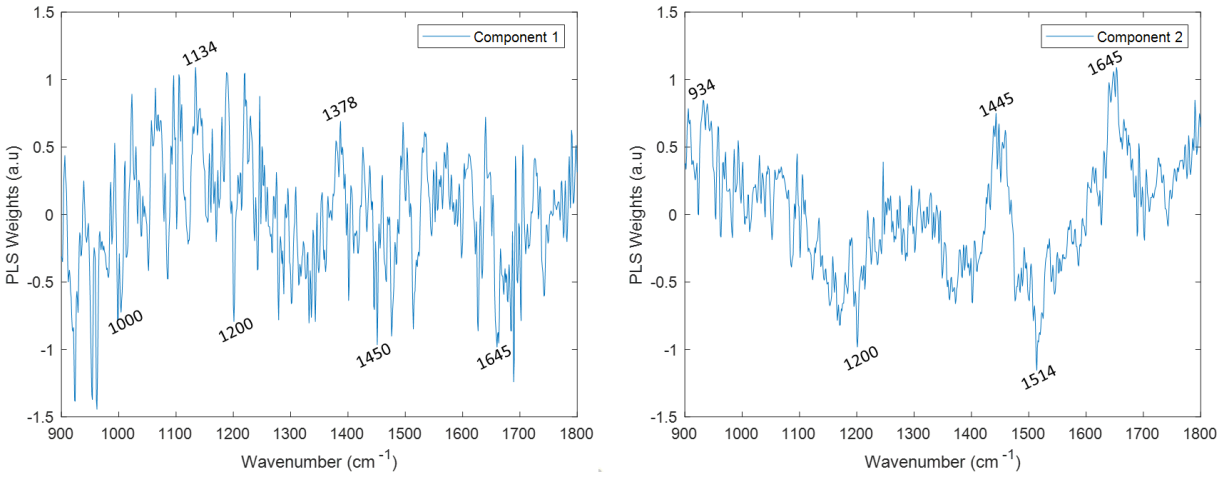


Figure 3.70: PLS component 1 and 2 weight plots for the exercised and non-exercised mdx mice.

Despite the noise present in the loading plots, biochemically relevant peaks could be identified. These peaks are summarised in table 3.40.

Table 3.40: Summary table of the peaks associated with the exercised and non-exercised mdx mice and tentative peak assignments.

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum
935/940	Proteins	✓	✓	✓	
970	Proteins, Nucleic Acids	✓	✓	✓	
1000	Proteins (Phenylalanine)	✓	✓	✓	✓
1045	Proteins (Phenylalanine, Proline)	✓			
1130	Proteins, Lipids	✓	✓	✓	
1200	Proteins	✓	✓	✓	
1300	Amide III (proteins), Lipids	✓	✓		
1335	Proteins, Nucleic Acids	✓	✓	✓	
1378	Proteins, Nucleic Acids	✓	✓	✓	
1445	Proteins, Lipids	✓	✓	✓	✓
1514	Nucleic Acids	✓	✓	✓	
1655	Amide I (proteins), Lipids	✓	✓	✓	✓

3.1.3.3 Ex vivo recordings

Similar to the *in vivo* recordings, PCA-LDA, PCA-QDA and PLS-DA were used to build classification models using the spectra recorded from different muscles (quadriceps, tibialis anterior, gastrocnemius, diaphragm) *ex vivo*. Leave one-spectrum-out and leave one-mouse-out cross-validations were employed to assess the performance of the models. The results are presented in the following sections.

3.1.3.3.1 Two group models

a) Quadriceps

The predictive capability of all the models generated using spectra acquired from quadriceps of exercised and non-exercised mice can be seen in table 3.41 and the respective ROC curves can be seen in figure 3.71.

Using LOOCV all models achieved much higher specificity than sensitivity, indicating that spectra acquired from non-exercised quadriceps were classified correctly more easily. Using LOMOCV, there was a big drop in specificity for all models leading thus to much lower accuracy values. PCA-LDA achieved the highest accuracy for LOOCV and PCA-QDA for LOMOCV. The AUROC was above 0.7 for all model when LOMOCV was employed but dropped significantly when LOOCV was used.

Table 3.41: Two group PCA-LDA model classification model performance evaluated using different CV methods for exercised and non-exercised quadriceps.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	53.3%	85.7%	71.3%
Leave-one-mouse-out CV	48.9%	64.3%	57.4%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	46.7%	87.5%	69.3%
Leave-one-mouse-out CV	51.1%	69.6%	61.4%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	51.1%	73.2%	63.4%
Leave-one-mouse-out CV	48.9%	48.2%	48.5%

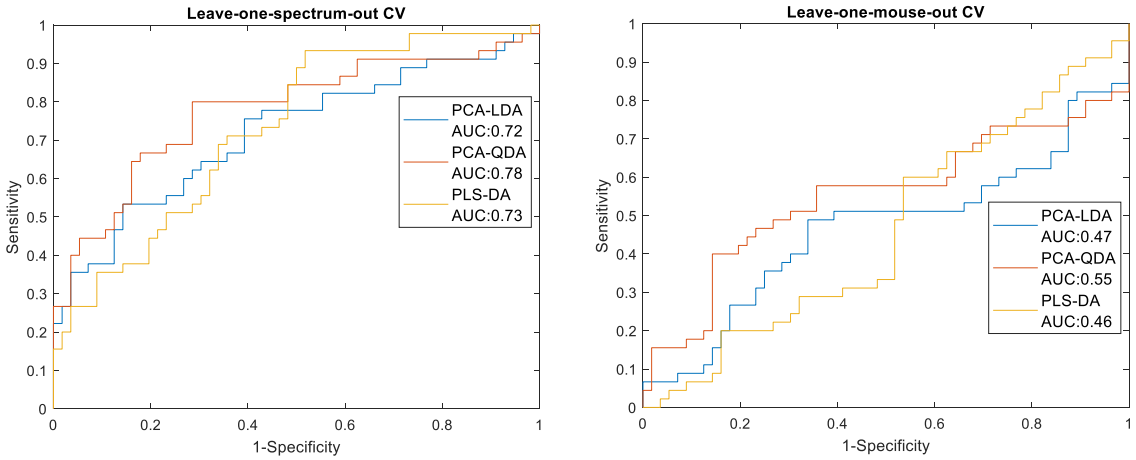


Figure 3.71: ROC curves for the cross validated classification models for exercised and non-exercised quadriceps. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

b) Tibialis Anterior (TA)

There were no PC scores significantly different between spectra acquired from TAs of exercised and non-exercised mice. Hence, LDA and QDA could not be performed. The predictive capability of the of the PLS-DA model using different CV methods can be seen in table 3.42 and the respective ROC curves can be seen in figure 3.72.

As can be seen from the accuracy values achieved with both CV methods and from the ROC curves the PLS-DA models were not able to discriminate between exercised and non-exercised mice using spectra acquired from TA. Using LOMOCV the sensitivity, specificity and accuracy values dropped significantly indicating that the technique did not detect differences in TA muscle that would allow it to correctly classify spectra in the different groups.

Table 3.42: Two group PLS-DA model classification model performance evaluated using different CV methods for exercised and non-exercised tibialis anterior.

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	29.3%	40.9%	35.3%
Leave-one-mouse-out CV	17.1%	31.8%	24.7%

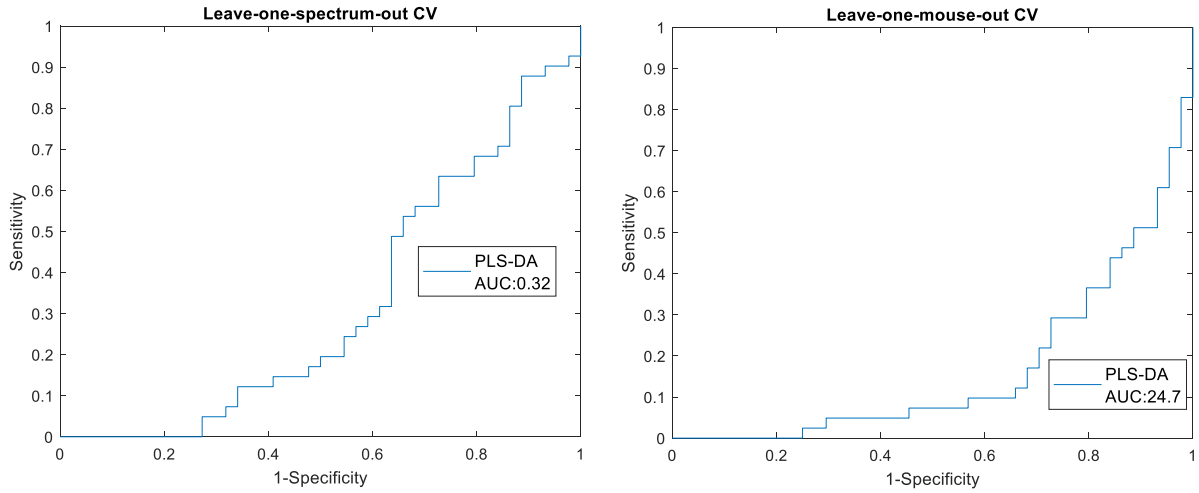


Figure 3.72: ROC curves for the cross validated PLS-DA classification models for exercised and non-exercised tibialis anterior. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

c) Diaphragm

Similarly, when using spectra acquired from the diaphragm of exercised and non-exercised mice there were no PC scores significantly different between the two groups. Hence, LDA and QDA could not be performed. The predictive capability of the of the PLS-DA model using different CV methods can be seen in table 3.43 and the respective ROC curves can be seen in figure 3.73.

The sensitivity, specificity and accuracy values were better compared to those achieved when using TA spectra. The AUROC remained unchanged and the accuracy of the model slightly increased when LOMOCV was used. However, the performance achieved with both CV approaches did not allow for discrimination between exercised and non-exercised muscle.

Table 3.43: Two group PLS-DA model classification model performance evaluated using different CV methods for exercised and non-exercised diaphragm.

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	44.8%	51.6%	48.3%
Leave-one-mouse-out CV	48.3%	54.8%	51.7%

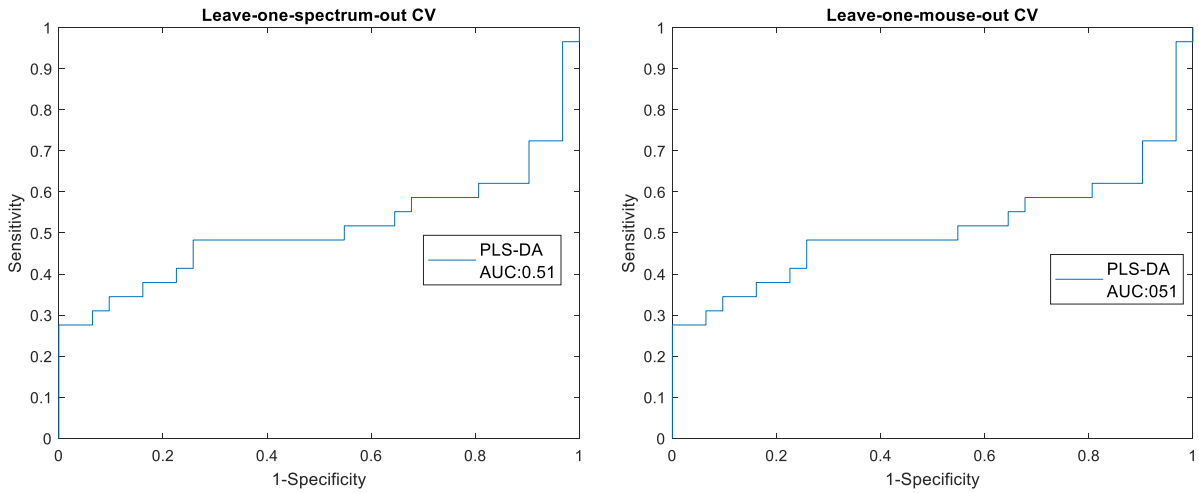


Figure 3.73: ROC curves for the cross validated PLS-DA classification models for exercised and non-exercised diaphragm. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

d) Gastrocnemius

The reliability of this classification performance is limited by the small numbers of samples in each group. The predictive capability of the models using different analysis techniques and CV methods can be seen in table 3.44 and the ROC curves can be seen in figure 3.74. Using LOMOCV the performance of the PCA-LDA model remained unchanged and there was a small drop in the sensitivity of the PCA-QDA model, with the accuracy of all models and CV approaches remaining above 80% and the AUROC being above 0.85. The ability of the PLS-DA model to correctly classify spectra was compromised considerably when LOMOCV was employed, as indicated by the big drop in accuracy and AUROC. The PCA-QDA model was demonstrated the best performance for LOMOCV.

Table 3.44: Two group PCA-LDA model classification model performance evaluated using different CV methods for exercised and non-exercised gastrocnemius.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	68.4%	95.5%	82.9%
Leave-one-mouse-out CV	68.4%	95.5%	82.9%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	78.9%	95.5%	87.8%
Leave-one-mouse-out CV	73.7%	95.5%	85.4%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	100%	86.4%	92.7%
Leave-one-mouse-out CV	68.4%	50%	58.5%

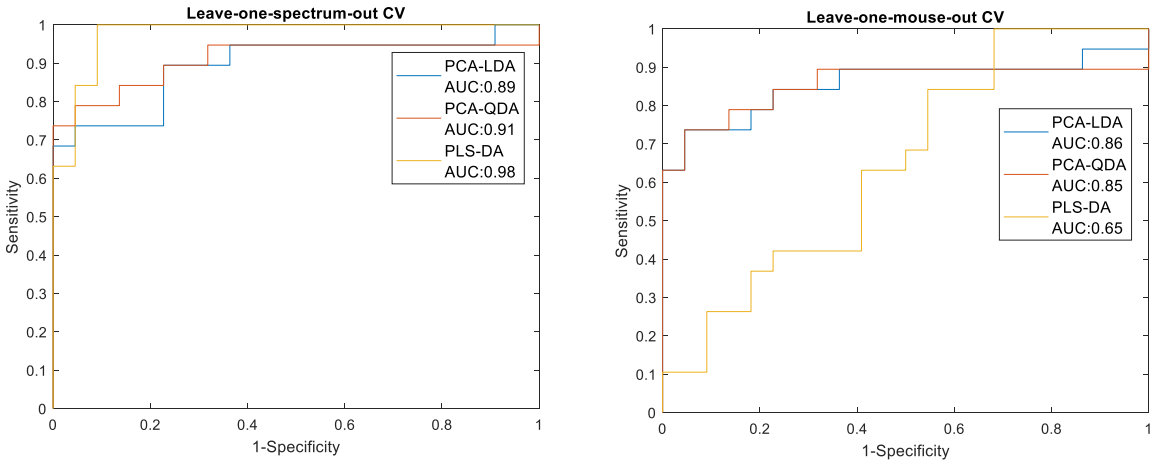


Figure 3.74: ROC curves for the cross validated classification models for exercised and non-exercised gastrocnemius. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

3.1.3.3.2 Different muscles comparison

a) Limb muscles vs. Diaphragm

The spectra acquired *ex vivo* from quadriceps, TA and gastrocnemius muscles were placed together in the limb muscles group. As can be seen in table 3.45, using LOMOCV decreased the ability of all models to correctly classify spectra due to a drop in sensitivity in the PCA related models and due to a drop in specificity in the PLS-DA model. The PLS-DA models yielded better sensitivity than specificity values. Using LOMOCV the PCA-LDA model was considered to have achieved the best performance as it achieved high accuracy value (93.5%) with the most balanced sensitivity and specificity values (both above 80%).

Table 3.45: Two group PCA-LDA model classification performance evaluated using different CV methods for limb muscles and diaphragm.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	100%	97.5%	98.1%
Leave-one-mouse-out CV	80.6%	96.7%	93.5%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	93.5%	99.2%	98%
Leave-one-mouse-out CV	77.4%	97.5%	93.5%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	100%	83.6%	86.9%
Leave-one-mouse-out CV	100%	74.6%	79.7%

3.1.4 Study 3: Preclinical MND Study

3.1.4.1 Classification Models Performance

PCA-LDA, PCA-QDA and PLS-DA were used to generate classification models. The performance of the models was validated using leave-one-spectrum out (LOOCV), leave-one-mouse-out (LOMOCV) and repeated leave-some-mice-out (RLSMOCV) cross-validation (CV). The results for the different two group models cross-validated using LOSOCV and LOMOCV are presented in section 3.6.1.1. Repeated cross-validation is presented in section 3.6.1.2.

3.1.4.1.1 Two Group Models

a) Three months old TDP-43 vs. TDP-43^{Q331K} mice

The model generated using the spectra from three months old TDP-43 and TDP-43^{Q331K} aimed to explore the ability of the technique to detect muscle pathology induced by the Q331K mutation in the TDP-43 gene at an established stage. The predictive capability of the models using different analysis and CV methods for the three months old TDP-43 mice can be seen in Table 3.46. The ROC curves, comparing the three different approaches for LOOCV and LOMOCV can be seen in figure 3.75.

Table 3.46: Two group PCA-LDA, PCA-QDA and PLS-DA model classification performance evaluated using different CV methods for the three months old TDP-43 and TDP-43^{Q331K} mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	84.4%	75%	79.9%
Leave-one-mouse-out CV	84.4%	71.2%	78.1%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	81.3%	75%	78.2%
Leave-one-mouse-out CV	78.1%	71.1%	75%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	93.8%	81.3%	87.5%
Leave-one-mouse-out CV	81.3%	78.1%	79.7%

Using LOMOCV the ability to correctly classify spectra decreased considerably for the PLS-DA model and moderately for the PCA-QDA model. The specificity of the PCA-LDA model also decreased when LOMOCV was used leading to a small drop in the accuracy. The area under the ROC curves also decreased for LOMOCV. However, the AUC values did not fall below 0.89. In general, the models attained better sensitivity than specificity values with most sensitivity values (for different methods and CVs) being above 80%.

Despite better sensitivity values achieved with the PCA related models when LOMO cross-validation was used, the PLS-DA model achieved more balanced sensitivity and specificity values and higher accuracy. Hence, for both CV approaches the PLS-DA model performed better than the PCA related models.

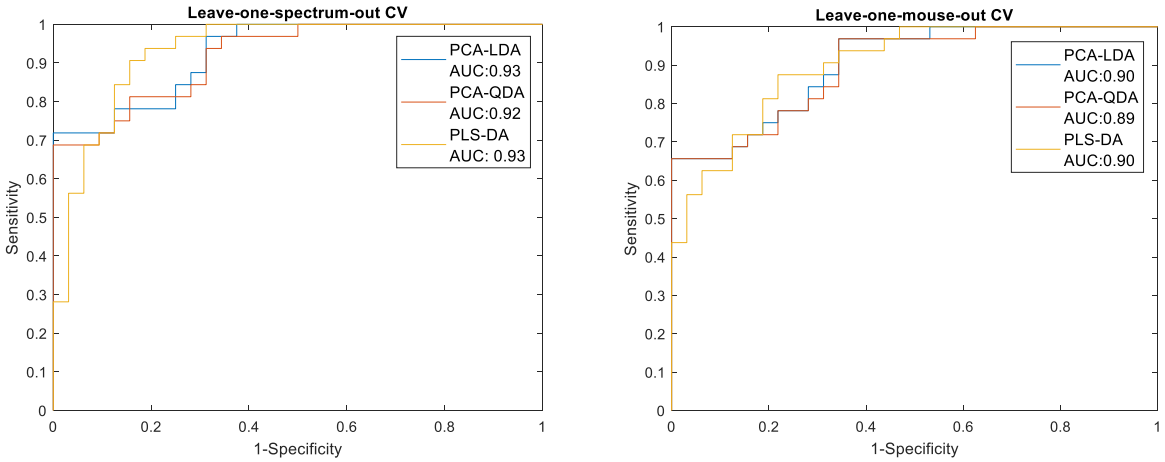


Figure 3.75: ROC curves for the cross validated classification models for the three months old TDP-43 and TDP-43^{Q331K} mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

b) Three months old SOD1^{G93A} vs. TDP-43^{Q331K} mice

The model generated using the spectra from three months old SOD1^{G93A} and TDP-43^{Q331K} mice aimed to explore the ability of the technique to detect MND related muscle pathology induced by mutations in different genes at a time when both models are manifesting a motor phenotype. The predictive capability of the models generated using different techniques and CV methods for the three months old *mdx* mice can be seen in Table 3.47.

Using LOMOCV the ability of both PCA related models to correctly classify spectra decreased. For both models there a drop in both sensitivity and specificity values was observed. The PLS-DA model demonstrated a slightly increased accuracy, due to an increase in specificity and a decrease in sensitivity, when using LOMOCV. The PCA-LDA model had the most balanced sensitivity and specificity values and the highest accuracy when LOOCV was used. PLS-DA demonstrated the highest accuracy for LOMOCV, with a much higher sensitivity than specificity.

Table 3.47: Two group PCA-LDA, PCA-QDA and PLS-DA model classification performance evaluated using different CV methods for the three months old SOD1^{G93A} and TDP-43^{Q331K} mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	79.3%	81.3%	80.3%
Leave-one-mouse-out CV	75.9%	75.0%	75.4%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	75.9%	81.3%	78.7%
Leave-one-mouse-out CV	72.4%	75.0%	73.8%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	96.6%	62.5%	78.7%
Leave-one-mouse-out CV	89.7%	70.0%	79.8%

The ROC curves for the different techniques and CV methods can be seen in figure 3.76. The AUROC was equal or above 0.77 for all the models generated using PCA. The PLS-DA model demonstrated the highest AUROC for both CV methods.

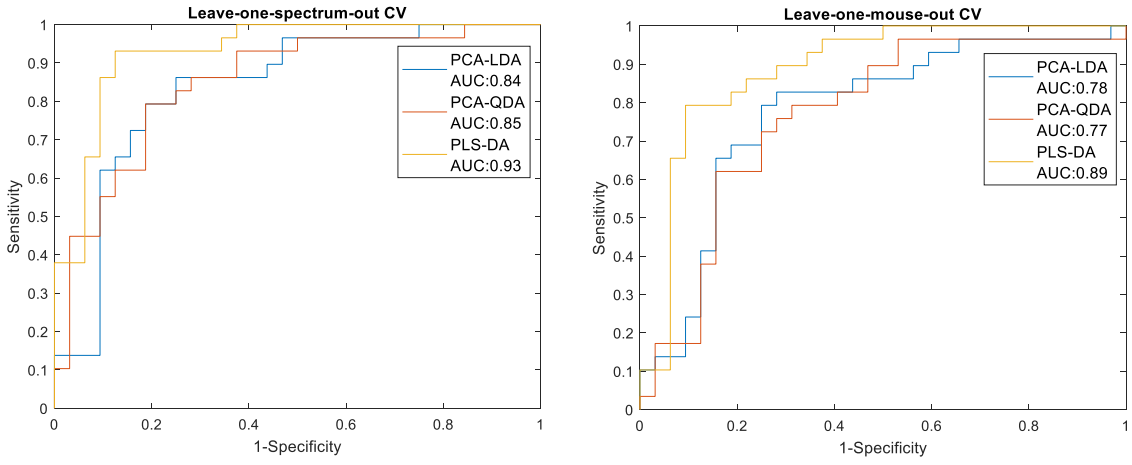


Figure 3.76: ROC curves for the cross validated classification models for the three months old SOD1^{G93A} and TDP-43^{Q331K} mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

3.1.4.1.2 Repeated Cross-Validation

a) TDP-43 vs. TDP-43^{Q331K} mice

Table 3.48 displays the classification performance of the cross-validated PCA-LDA, PCA-QDA and PLS-DA models for the three months old TDP-43 mice. The mean sensitivity, specificity and accuracy values (+/-) one standard deviation and the coefficients of variability (CoV) are shown. The PCA-QDA model achieved the best performance with an accuracy of 76.3%. The PLS-DA model had the larger coefficients of variability for all three indices and the PCA-LDA model had the smallest ones. For all three approaches RLSMOCV yielded a similar performance with LOMOCV.

Table 3.48: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three months old TDP-43 and TDP-43^{Q331K} mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

Three months old TDP-43/TDP-43^{Q331K} mice	***Sensitivity (+/- Standard Deviation, CoV)***	***Specificity (+/- Standard Deviation, CoV)***	***Accuracy (+/- Standard Deviation, CoV)***
PCA-LDA	84.9% (+/- 2.2, CoV:2.6%)	73.4% (+/- 1.7, CoV:2.3%)	79.2% (+/- 1.5, CoV:1.9%)
PCA-QDA	82.4% (+/- 2.1, CoV:2.5%)	76.1% (+/- 3.5, CoV:4.6%)	79.3% (+/- 2.1, CoV:2.6%)
PLS-DA	80.8% (+/- 5.4, CoV:6.7%)	76.8% (+/- 3.8, CoV:4.9%)	79.0% (+/- 3.4, CoV:4.3%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model can be seen in figure 3.77. The mean AUC value for each model (+/-) one standard deviation is also shown. The PCA-LDA and QDA ROC curves had mean AUCs of 0.92 and 0.91 respectively whereas the PLS-DA ROC curve had a mean AUC of 0.9.

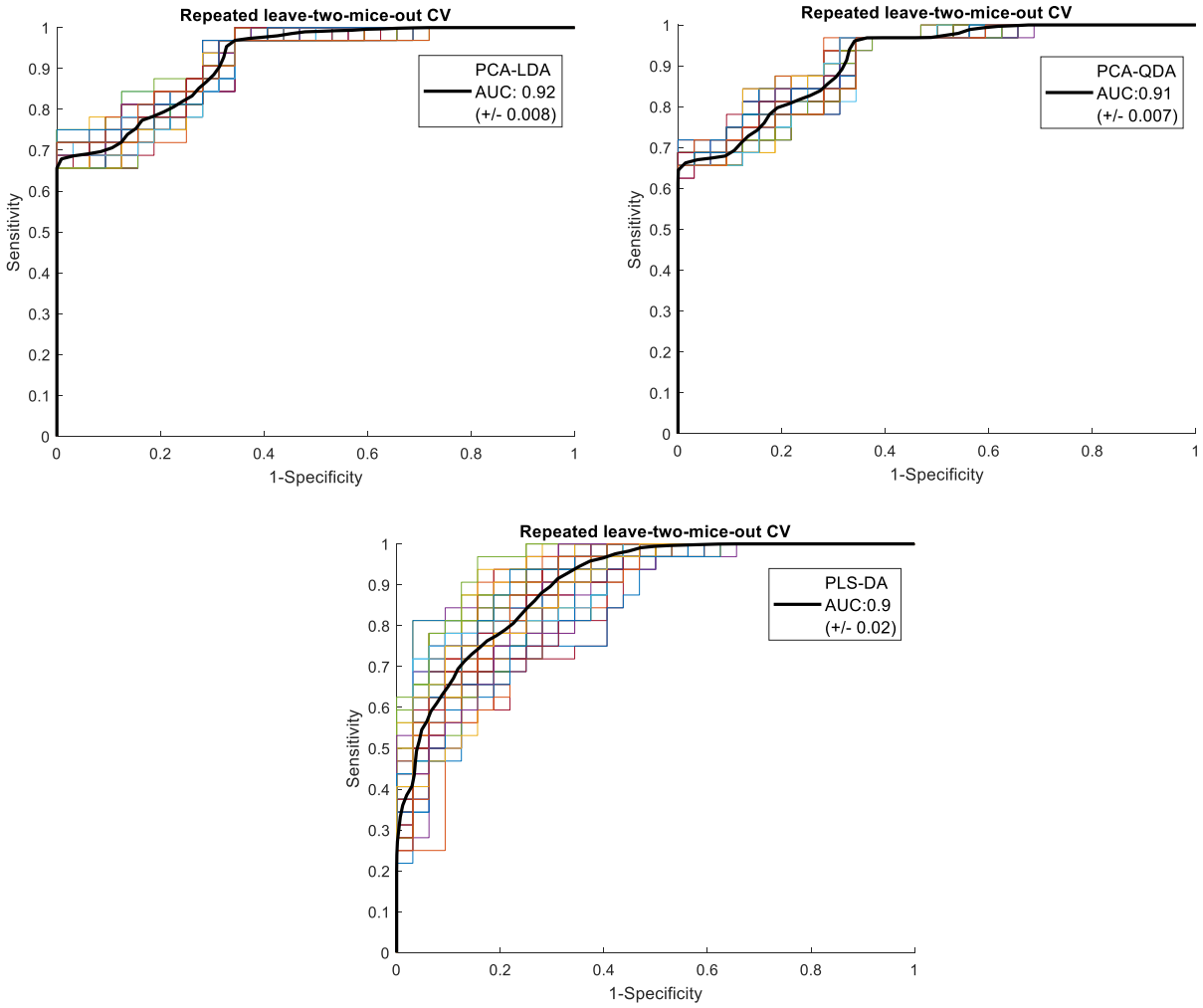


Figure 3.77: ROC curves generated from repeated cross-validation of three months old TDP-43 and TDP-43^{Q331K} mice models. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

b) SOD1^{G93A} vs. TDP-43^{Q331K} mice

Table 3.49 displays the classification performance of the cross-validated PCA-LDA, PCA-QDA and PLS-DA models for the three months old SOD1^{G93A} and TDP-43^{Q331K}mice. All the models attained similar accuracy values, with the LDA model having a more balanced sensitivity and specificity. Hence, the PCA-LDA model was considered to have demonstrated the best performance.

Table 3.49: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three months old SOD^{G93A} and TDP-43^{Q331K} mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

Three months old SOD1^{G93A}/ TDP-43^{Q331K} mice	***Sensitivity*** ***(+/- Standard Deviation, CoV)***	***Specificity*** ***(+/- Standard Deviation, CoV)***	***Accuracy*** ***(+/- Standard Deviation, CoV)***
PCA-LDA	77.7% (+/-3.5, CoV:4.5%)	77.2% (+/-3.7, CoV:4.8%)	77.5% (+/-2.7, CoV:3.5%)
PCA-QDA	76.2% (+/-3.3, CoV:6.5%)	76.6% (+/-2.7, CoV:4.9%)	76.4% (+/-1.9, CoV:4.2%)
PLS-DA	92.3% (+/-3.1, CoV:3.4%)	67.3% (+/-2.6, CoV:3.9%)	78.1% (+/-1.8, CoV:2.3%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model for the three months old SOD1^{G93A} and TDP-43^{Q331K} mice can be seen in figure 3.78. The PCA-LDA and QDA ROC curves had a mean AUC of 0.84 and 0.83 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.89.

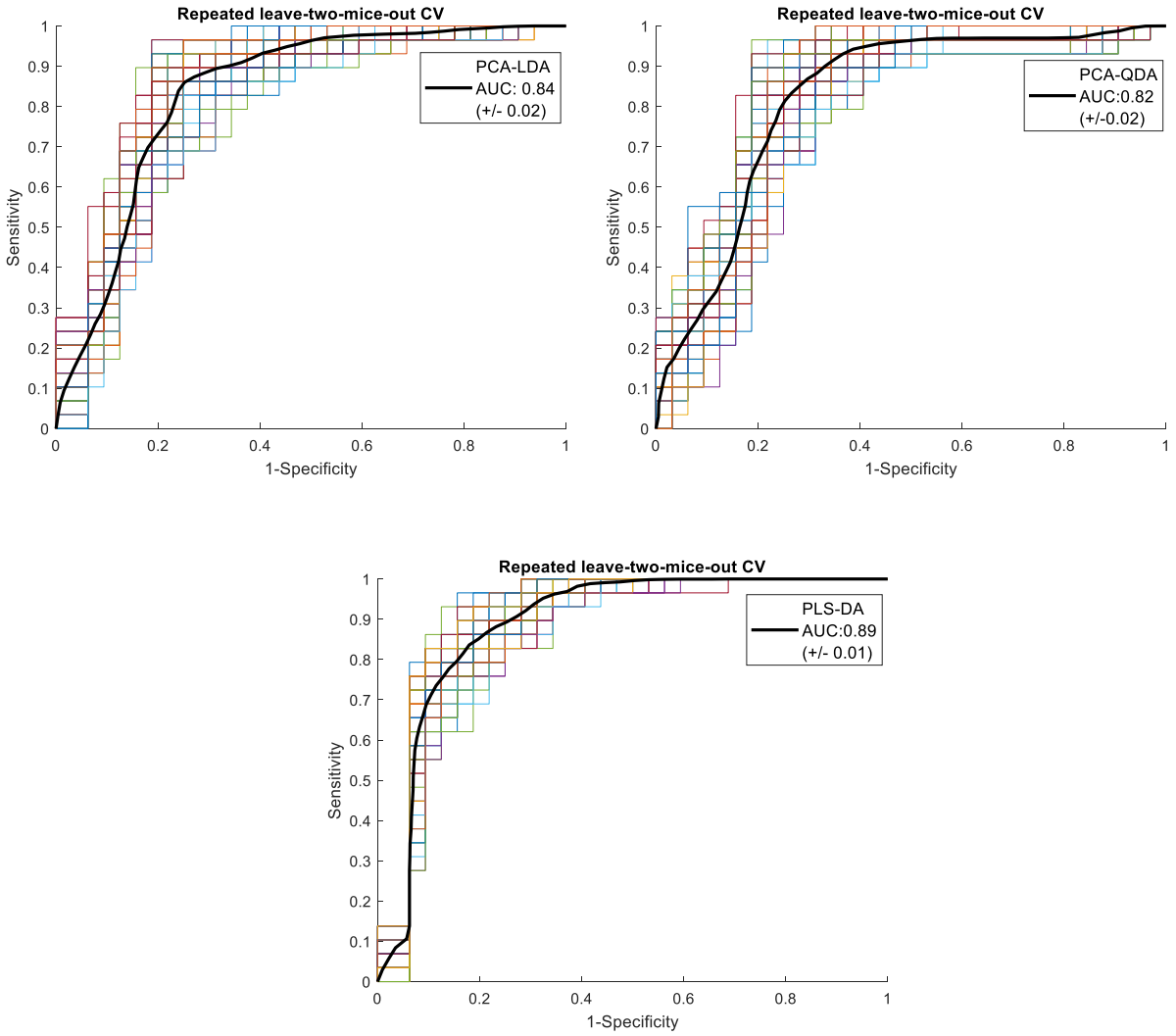


Figure 3.78: ROC curves generated from repeated cross-validation of three months old SOD^{G93A} and TDP-43^{Q331K} mice. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

3.1.4.2 Basis of classification

3.1.4.2.1 Mean and difference spectra

The mean and difference spectra of the different mice groups were plotted to visually examine the most prominent peaks and to identify the major differences between the groups of each model. In figure 3.79 the mean spectra of each group (+/-) one standard deviation are shown.

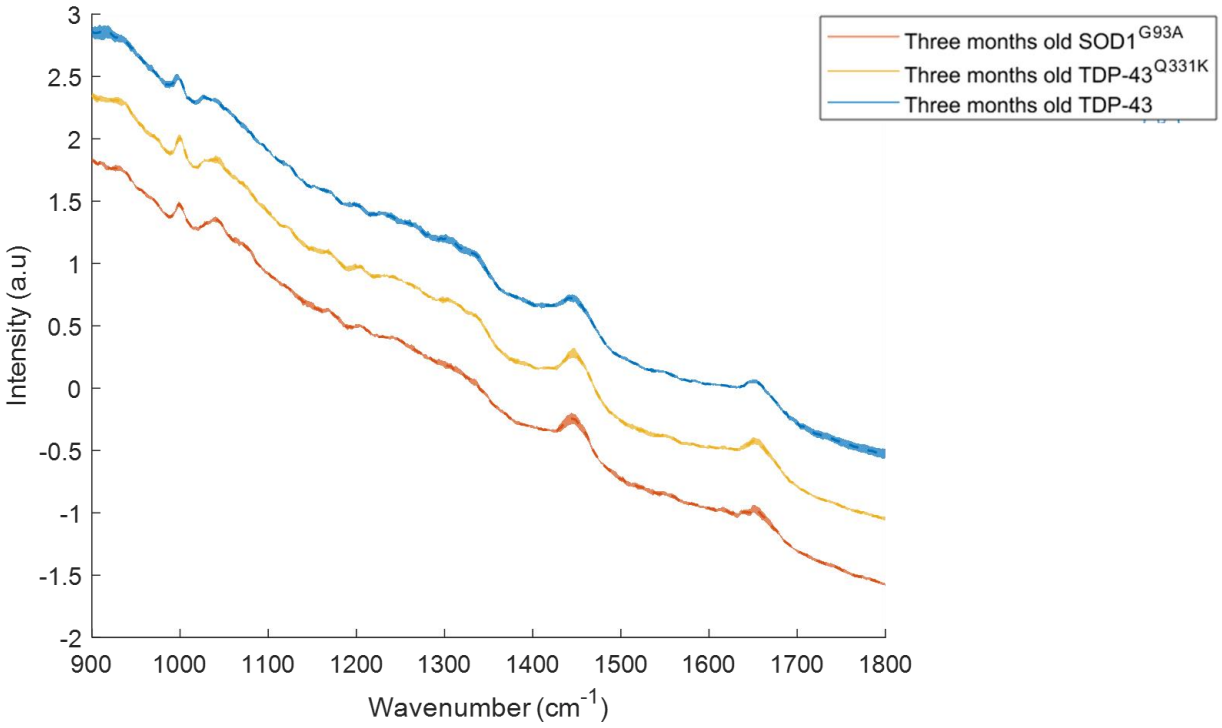


Figure 3.79: Mean Raman spectra for SOD1^{G93A}, TDP-43 and TDP-43^{Q331K} mice groups.
The mean spectra for three months old SOD1^{G93A}, TDP-43 and TDP-43^{Q331K} mice are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

Figures 3.80 and 3.81 show the overlaid mean and difference spectra of the combinations of groups used to build the two group models for this study. The mean spectra of the different groups are very similar, with the same prominent peaks being present in all of them. Biochemically reliable peaks could be identified despite the fluorescent background present in the spectra.

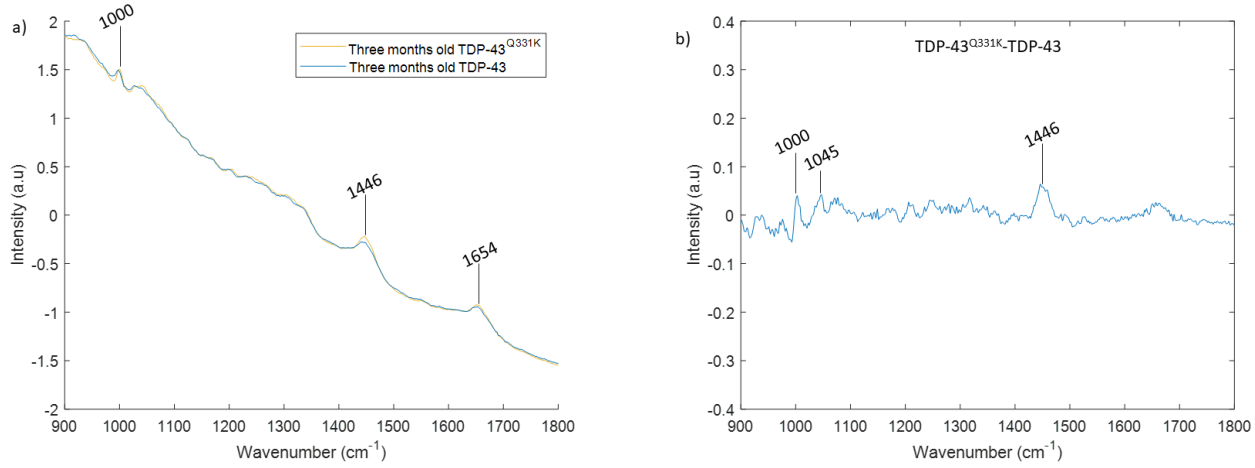


Figure 3.80: Mean and difference spectra of three months old TDP-43 and TDP-43^{Q331K} mice.
a) Overlaid mean spectra for three months old TDP-43 and TDP-43^{Q331K} mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

As can be seen in the difference spectrum above the peaks located around 1000 cm^{-1}, 1045 cm^{-1} and 1446 cm^{-1} were found to be increased in the spectra acquired from the diseased group.

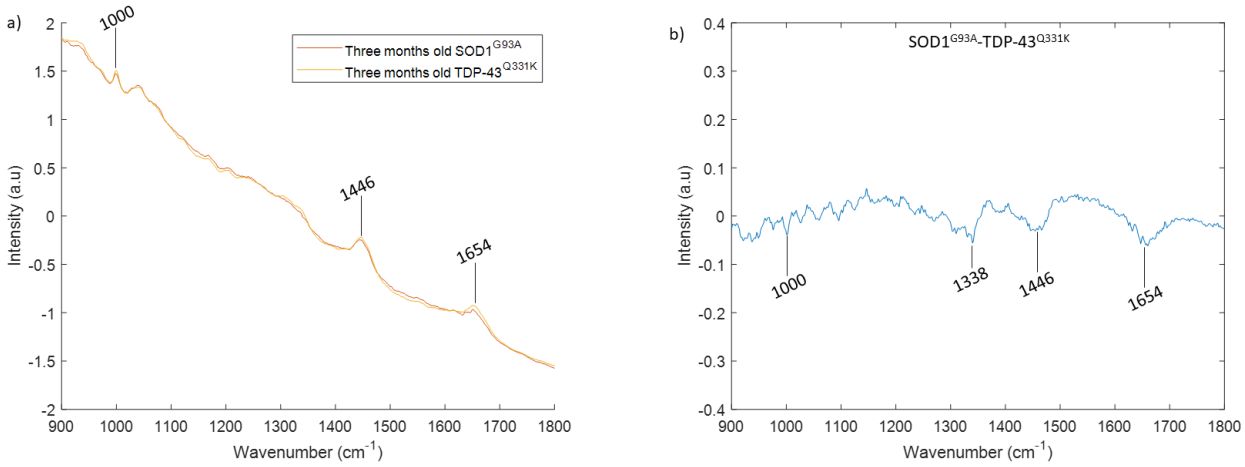


Figure 3.81: Mean and difference spectra of three months old SOD1^{G93A} and TDP-43^{Q331K} mice.
a) Mean spectra for three months old SOD1^{G93A} and TDP-43^{Q331K} mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

The peaks located around 1000 cm^{-1}, 1335 cm^{-1}, 1445 cm^{-1}, 1655 cm^{-1} were increased in the spectra of the TDP-43^{Q331K} when compared to the spectra of the SOD1^{G93A} mice.

The peaks that were identified in the mean and difference spectra of the mice groups used in this study and their tentative peak assignments are summarised in Table 3.50.

Table 3.50: Prominent Raman peaks in mean and difference spectra of the TDP-43, TDP-43^{Q331k} and SOD1^{G93A} mice and tentative peak assignments.

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1045		Proteins
1338	C-N stretching	Proteins, Nucleic Acids
1446	CH_2 bending, CH_2CH_3 bending	Proteins, Lipids
1655	C=O stretching, C=C stretch	Amide I (proteins), Lipids

Background subtraction was used to visually explore the spectral features present in the spectra further. The background subtracted mean spectra of the different groups can be seen in figure 3.82. A larger number of biochemically relevant peaks were identifiable in these mean spectra.

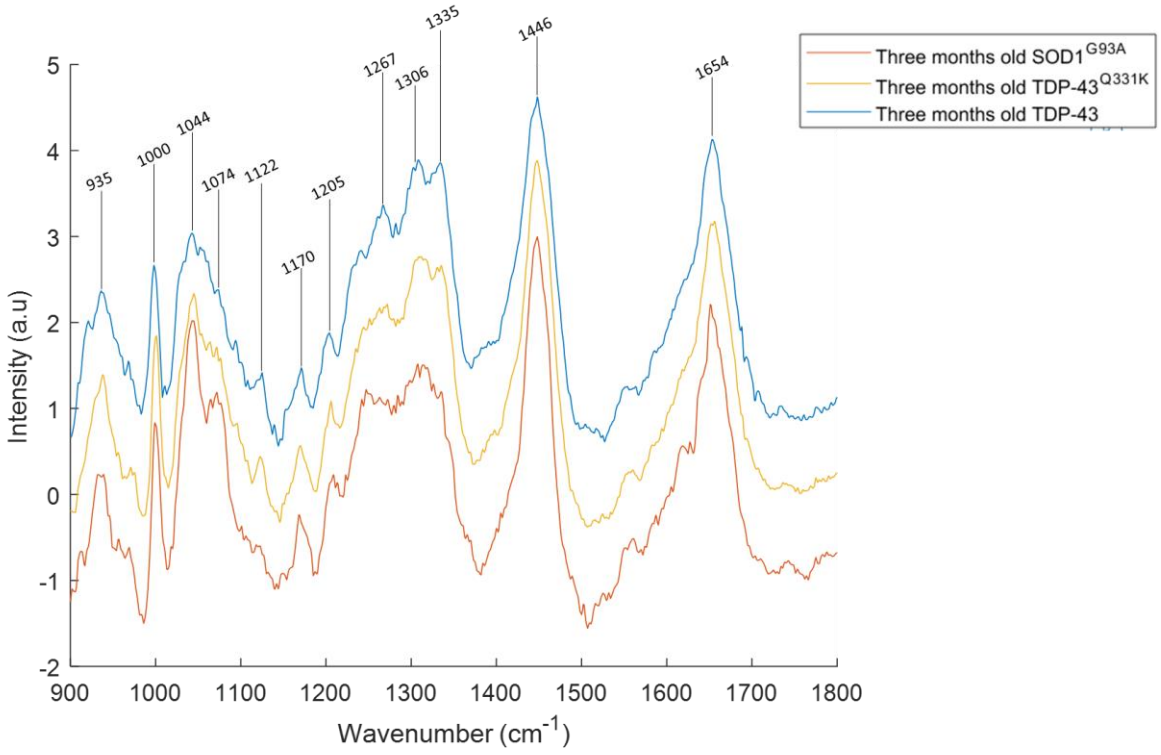


Figure 3.82: Background subtracted mean spectra of SOD1^{G93A}, TDP-43 and TDP-43^{Q331K} mice. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

Additional shoulder peaks can be seed around 975 cm^{-1} and 1550 cm^{-1}. The spectral features of the mean spectra varied in the spectral region between 1230 cm^{-1} and 1340 cm^{-1}, with the three peaks that were previously identified in this region (studies 1 and 2) being more clearly discernible in spectra of the TDP-43 mice. Tentative peak assignments for the peaks present in the background subtracted mean spectra are presented in table 3.51.

Table 3.51: Prominent Raman peaks in background subtracted spectra and tentative peak assignments. *Amino acids are specified when the peaks are largely associated with them in the literature.*

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
935	C-C stretching	Proteins
975	CH_2 deformation	Proteins, Nucleic Acids
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1074	C-C stretching, C-O stretching	Proteins (Tryptophan), Lipids
1121	C-C stretching, C-N stretching	Proteins, Lipid
1170	C-H bending	Proteins (Tyrosine, Phenylalanine)
1205	$\nu(C\text{-}C_6H_5)$	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)
1267	C-N stretching	Amide III (proteins), Lipids
1305	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1446	CH_2 bending, CH_2CH_3 bending	Proteins, Lipids
1550	$\nu(C\text{-}C)$	Proteins (Tryptohan)
1654	C=O stretching, C=C stretching	Amide I (proteins), Lipids

3.1.4.2.2 Multivariate Analysis

Multivariate techniques were employed in order to utilise and further elucidate the biochemical features present in the spectra. The PCA loadings and the linear discriminant function (LDF) as well as the PLS weights were plotted in order to illustrate the important peaks for spectral classification.

Assigning specific peaks to one of the two groups used in each model was not trivial since the two groups were not always clearly separated around zero in the corresponding score histograms or score plots. Additionally, the peaks in the loadings or weight plots were most often not centred around zero probably due to the presence of background in the spectra. Hence, a combination of the information found in the difference spectra and the loading/weight plots was used in an attempt to better understand the differences present in the spectra from the different groups. The score histograms and loading/weight plots in which the separation of the spectra from the different groups was best for each model are shown in the following sections. The score

histograms and loading/weight plots for the rest of the PCs and components used to build each model are shown in Appendix C. Tables 3.52 and 3.53 summarise and compare the most prominent peaks found in all the loading/weight plots (from all the components used for each model) and the ones found in the difference spectra of each two-group model.

a) Three months old TDP-43 vs. TDP-43^{Q331K} mice

Using Student's t-tests followed by fdr correction PC3 and PC6 scores were found to be significantly different between the three months old TDP-43/TDP-43^{Q331K} mice. PC1 scores were found to have the largest difference between the two groups (q_{PC3}= 9.7e^{-07} , q_{PC6}=0.004). The score histogram and loading plot of PC3 can be seen in figure 3.83. The histogram and loading plot of PC6 can be seen in Appendix C.

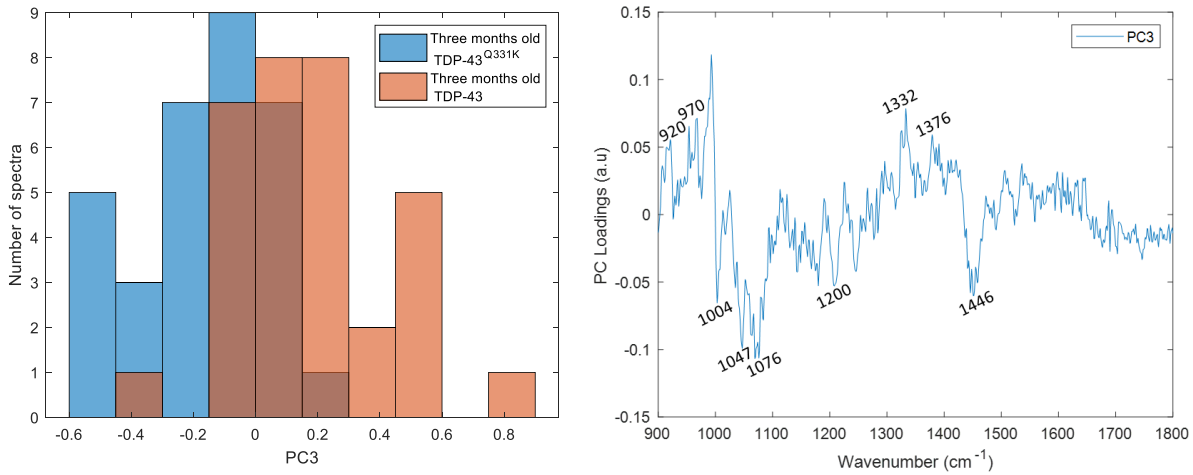


Figure 3.83: PC3 score histogram and loading plot for the three months old TDP-43 and TDP-43^{Q331K} mice.

The results of the LDA analysis using PC3 and PC6 as input variables can be seen in figure 3.84.

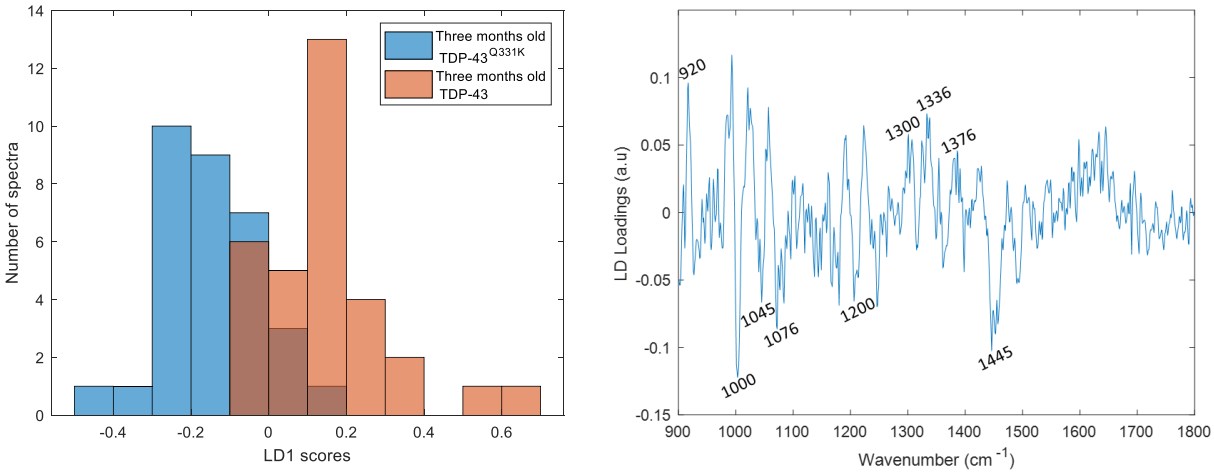


Figure 3.84: LD1 score histogram and LDF plot for the three months old TDP-43 and TDP-43^{Q331K} mice.

The most prominent bands that were identified in the previous section could also be seen in PC and LDA loading plots demonstrating that differences in muscle's biochemical composition were important for discrimination. There was an overlap in the score of the different groups in both PC and LDA score histograms, which made associating the prominent peaks present in the respective loading plots with one of the two groups more difficult. In both loading plots the spectra of the TDP-43^{Q331K} mice seemed to have a larger contribution in the negative scores. Thus, the negative peaks (1000 cm^{-1}, 1045 cm^{-1}, 1076 cm^{-1}, 1200 cm^{-1}, 1446 cm^{-1}) were considered to be increased in the spectra of the diseased mice whereas the positive peaks located around 970 cm^{-1}, 1300 cm^{-1}, 1335 cm^{-1} and 1376 cm^{-1} were considered to be increased in the spectra of the TDP-43 mice.

In the PLS-DA analysis the model generated using the first 4 components demonstrated the highest accuracy value. The spectra were not clearly separated around zero in any of the score histograms. Similar peaks as the ones from the PCA-LDA analysis were found in the PLS weight plots (Appendix C, Table 3) demonstrating that discrimination between the different groups using the different models was done on a similar biomolecular basis.

The most prominent biochemically relevant peaks found in the loading/weight plots for all the PCs and PLS components and in the difference spectra and their tentative peak assignments are summarised in table 3.52.

Table 3.52: Summary table of the peaks associated with the three months old TDP-43/TDP-43^{Q331K} mice and tentative peak assignments.

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum
920	Proteins (Proline)	✓	✓	✓	
970	Proteins, Nucleic Acids	✓	✓	✓	
1000	Proteins (Phenylalanine)	✓	✓	✓	✓
1045	Proteins (Phenylalanine, Proline)	✓	✓	✓	✓
1076	Lipids, Phospholipids	✓		✓	
1200	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)	✓			
1300	Amide III (proteins), Lipids	✓	✓	✓	
1335	Proteins, Nucleic Acids	✓	✓	✓	
1376	Proteins, Nucleic Acids	✓	✓	✓	
1445	Proteins, Lipids	✓	✓	✓	✓
1655	Amide I (proteins), Lipids			✓	

b) Three months old SOD1^{G93A} vs. TDP-43^{Q331K} mice

PC1, PC2 and PC5 scores were found to be significantly different between three months old SOD1^{G93A} and TDP-43^{Q331K}. PC1 scores were found to have the largest difference between the two groups (q_{PC1}= 4.4e^{-06}, q_{PC2}=0.04, q_{PC5}=0.04). The score histogram and loading plot of PC1 can be seen in figure 3.85. The loading plot of PC2 and PC5 can be found in Appendix C. The LD1 score histogram and LDF generated using PC1, PC2 and PC5 as inputs can be seen in figure 3.86.

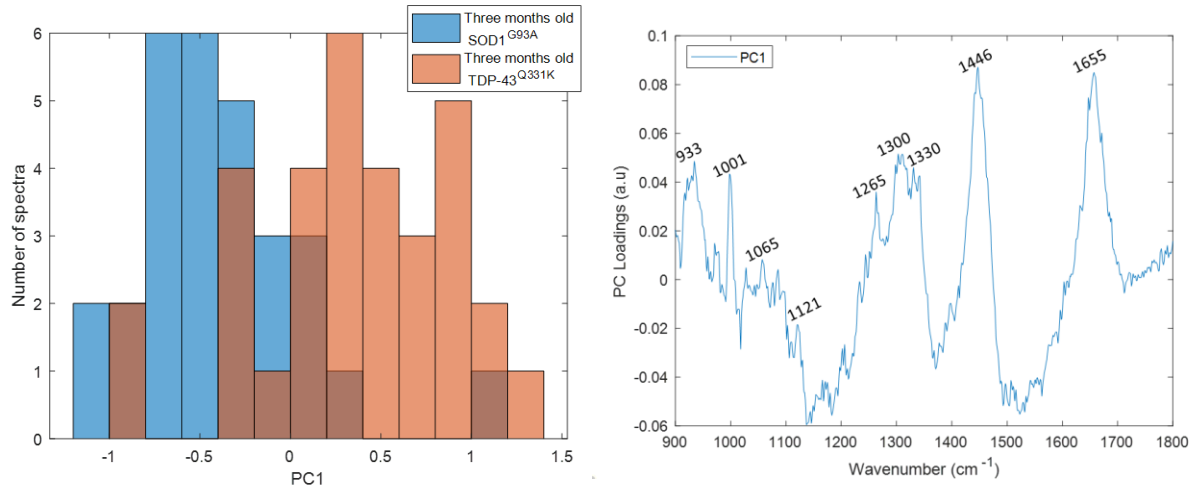


Figure 3.85: PC1 score histogram and loading plot for the three months old SOD1^{G93A} and TDP-43^{Q331K} mice.

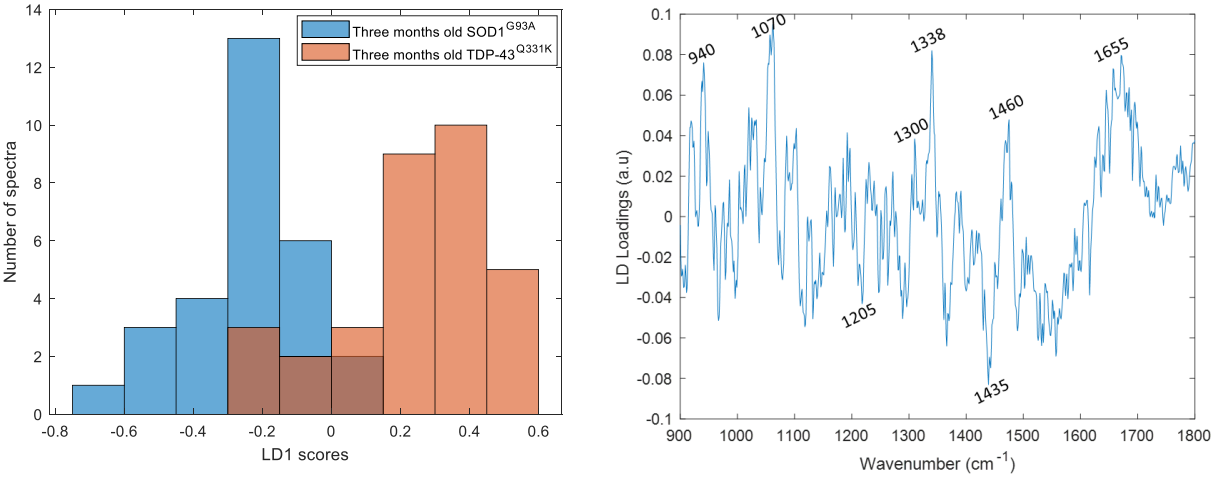


Figure 3.86: LD1 score histogram and LDF plot for the three months old SOD1^{G93A} and TDP-43^{Q331K} mice.

In the PLS-DA analysis using the first three components generated the model with the highest accuracy value. The score and weight plot of the first component can be seen in figure 3.87. weight plots of the rest of the components can be seen in Appendix C.

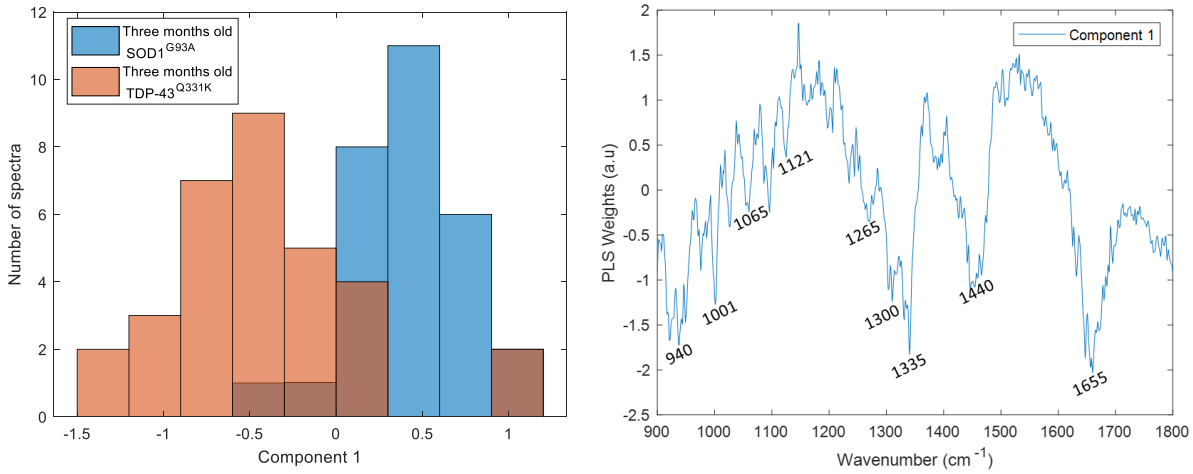


Figure 3.87: Component 1 score histogram and weight plot for the three months old SOD1^{G93A} and TDP-43^{Q331K} mice.

Despite some background present in PC1 and component one loading plots biochemically reliable peaks were identifiable in both loading plots. The scores of the TDP-43^{Q331K} mice contributed more in the positive values of the PC1 loading plot and in the negative values of the component one weight plot. Hence, the peaks in the loading/weight plots around 1000 cm^{-1}, 1265 cm^{-1}, 1300 cm^{-1}, 1335 cm^{-1}, 1440 cm^{-1} and 1655 cm^{-1} were considered to have increased in the spectra of this group. This was consistent with the findings of the difference spectrum since most of these peaks were also present in that spectrum and were shown to have decreased for the SOD1^{G93A} mice. The LDA score histogram plot showed that positive values corresponded to the spectra of the TDP-43^{Q331K} mice whereas negative values correspond to the spectra of the SOD1^{G93A} mice. The LDA loading plot consisted of similar peaks with PCA and PLS loading plots. However, the peak around 1446 cm^{-1} was split into two peaks, a negative one around 1434 cm^{-1} and a positive one around 1460 cm^{-1}. The most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments are summarised in table 3.53.

Table 3.53: Summary table of the peaks associated with the three months old SOD1^{G93A} and TDP-43^{Q331K} mice and tentative peak assignments.

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum
935/940	Proteins	✓	✓	✓	
1000	Proteins (Phenylalanine)	✓	✓	✓	✓
1065/1070	Lipids, Phospholipids	✓	✓	✓	
1121	Proteins	✓	✓	✓	
1205	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)	✓	✓	✓	
1265	Amide III (proteins), Lipids	✓		✓	
1300	Amide III (proteins), Lipids	✓	✓	✓	✓
1335	Proteins, Nucleic Acids	✓	✓	✓	✓
1435	Proteins, Lipids	✓	✓	✓	
1445	Proteins, Lipids	✓		✓	✓
1460	Proteins, Lipids	✓	✓		
1653	Amide I (Proteins), Lipids	✓	✓	✓	✓

3.1.5 Study 4: Male *mdx* Mice Study

3.1.5.1 Classification performance

PCA-LDA, PCA-QDA and PLS-DA were used to generate classification models. The performance of the models was validated using leave-one-spectrum out (LOOCV), leave-one-mouse-out (LOMOCV) and repeated leave-some-mice-out (RLSMOCV) cross-validation (CV). The results for the two group models cross-validated using LOSOCV and LOMOCV are presented in section 3.1.5.1.1. Repeated cross-validation is presented in section 3.1.5.1.2.

3.1.5.1.1 Two group models

a) One month old male *mdx* vs. C57Bl/10 mice

The predictive capability of the models using different analysis techniques and CV methods can be seen in Table 3.54. Using LOMOCV the specificity of both PCA related models has decreased leading to a drop in accuracy from 75% to 71.5% in both cases. The PLS-DA model achieved sensitivity and specificity values of 78.6% for LOOCV. The LOMO cross-validated PLS-DA model demonstrated lower sensitivity leading to an accuracy of 73.2%. The PLS-DA model achieved the best accuracy for both CV methods. The PCA related models, however, demonstrated higher sensitivity values with a similar accuracy when LOMOCV was used and might therefore be the models of choice where detecting the pathology is more important than the overall number of correctly classified spectra.

Table 3.54: Two group PCA-LDA model classification model performance evaluated using different CV methods for the one month old male mdx and C57Bl/10 mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	75.0%	75.0%	75.0%
Leave-one-mouse-out CV	75.0%	68.0%	71.5%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	78.6%	71.4%	75.0%
Leave-one-mouse-out CV	75.0%	68.0%	71.5%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	78.6%	78.6%	78.6%
Leave-one-mouse-out CV	68.0%	78.5%	73.2%

The ROC curves for the different techniques and CV methods can be seen in figure 3.88. The area under the ROC curves was around 0.70 for all the models generated using PCA. The PLS-DA models yielded an AUC of 0.91 and 0.87 for LOSOVC and LOSOVM respectively.

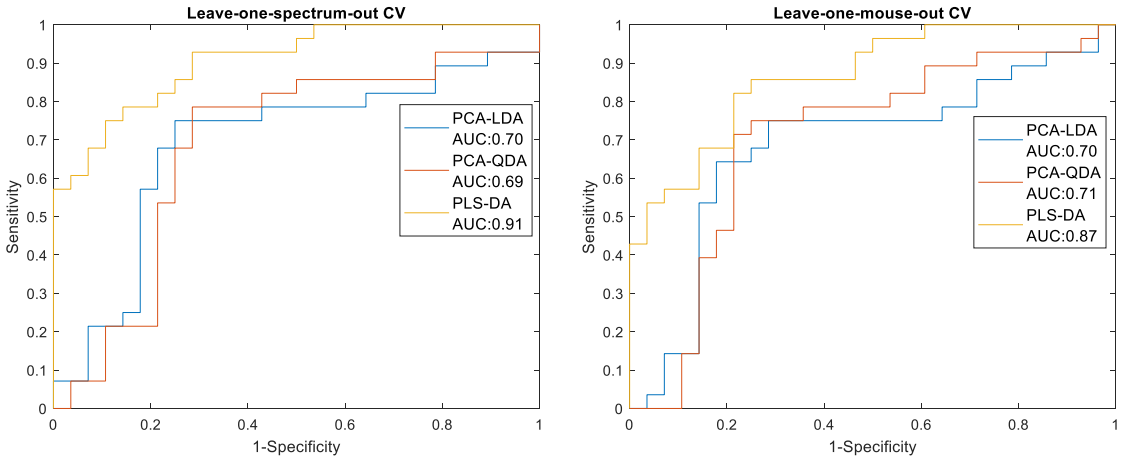


Figure 3.88: ROC curves for the cross validated classification models for the one month old male mdx and C57Bl/10 mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

b) Three months old male *mdx* vs. C57Bl/10 mice

This is a more established disease stage for the *mdx* mice. Thus, the model generated using the spectra from three months old *mdx* aimed to explore the ability of the technique to detect muscle pathology at a later disease stage. The predictive capability of all the models increased with age and can be seen in Table 3.55. The ROC curves, comparing the three different approaches for LOSOCV and LOMOCV can be seen in figure 3.89.

All models achieved sensitivity, specificity and accuracy values above 80%. The performance of the PCA-LDA and PLS-DA models did not change for the different cross-validation methods. The sensitivity of the PCA-QDA model decreased slightly when LOMOCV was used, leading to a small drop in accuracy. For both CV approaches the PLS-DA model performed better than the PCA related models. The area under the ROC curves slightly decreased for the PCA related models when LOMOCV was employed. However, the AUC values did not fall below 0.94. The PLS-DA models achieved AUROC of 0.99 for both CV methods.

Table 3.55: Two group PCA-LDA, PCA-QDA and PLS-DA model classification performance evaluated using different CV methods for the three months old male mdx and C57Bl/10 mice.

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	90.6%	84.4%	87.5%
Leave-one-mouse-out CV	90.6%	84.4%	87.5%

PCA-QDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	87.5%	90.6%	89.1%
Leave-one-mouse-out CV	84.4%	90.6%	87.5%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	100%	96.9%	98.4%
Leave-one-mouse-out CV	100%	96.9%	98.4%

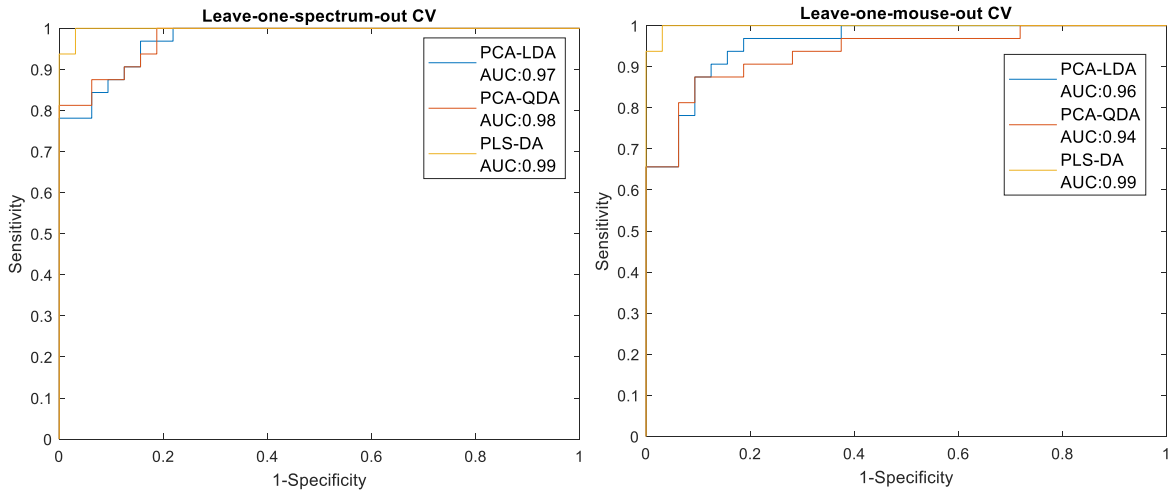


Figure 3.89: ROC curves for the cross validated classification models for the three months old male mdx and C57Bl/10 mice. *ROC curves for all models using leave-one-spectrum-out and leave-one-mouse-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

3.1.5.1.2 Repeated cross-validation

The classification performance of the RLSMOCV PCA-LDA, PCA-QDA and PLS-DA models for distinguishing between one month old male *mdx* and C57Bl/10 mice can be seen in Table 3.56. The PCA-QDA model demonstrated the highest sensitivity and accuracy values as well as the most balanced sensitivity and specificity and was therefore considered to have achieved the best performance.

Table 3.56: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for one month old mdx and C57Bl/10 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

One month old male mdx/C57Bl/10 mice	***Sensitivity (+/- Standard Deviation, CoV)***	***Specificity (+/- Standard Deviation, CoV)***	***Accuracy (+/- Standard Deviation, CoV)***
PCA-LDA	76.9% (+/-3.7, CoV:4.8%)	79.8% (+/-3.7, CoV:4.6%)	78.3% (+/-2.7, CoV:3.4%)
PCA-QDA	78.8% (+/-4.3, CoV:5.4%)	79.8% (+/-3.5, CoV:4.4%)	79.3% (+/-2.8, CoV:3.5%)
PLS-DA	74.1 % (+/-3.8, CoV:5.1%)	82.1% (+/-3.5, CoV:4.2%)	78.1% (+/-3.1, CoV:3.9%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model for the one month old male *mdx* mice can be seen in figure 3.90. The PCA-LDA and QDA ROC curves had a mean AUC of 0.80 and 0.81 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.84.

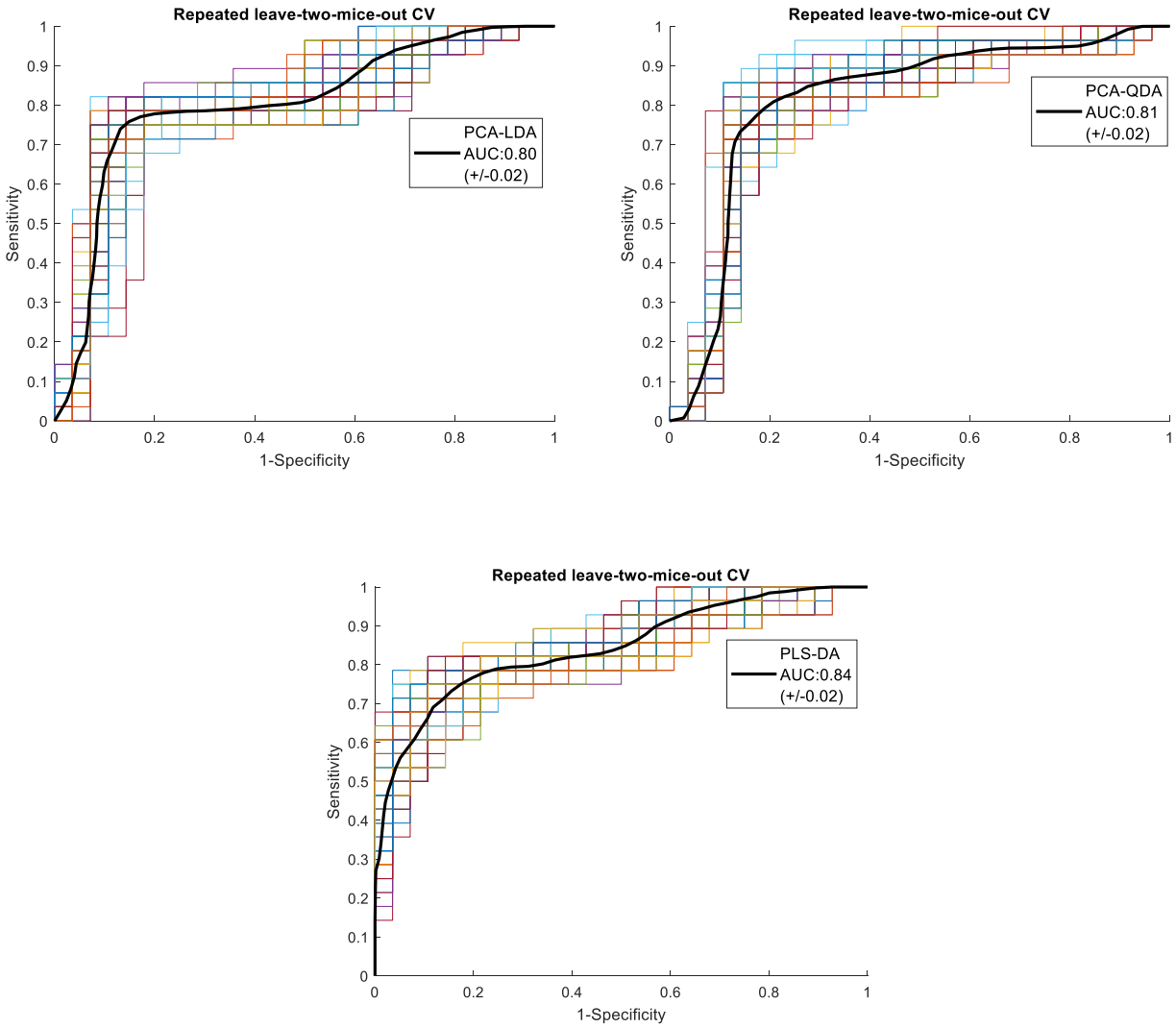


Figure 3.90: ROC curves generated from repeated cross-validation of one month old mdx and C57Bl/10 mice. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

Table 3.57 displays the classification performance of the cross-validated PCA-LDA, PCA-QDA and PLS-DA models for the three months male *mdx* and C57Bl/10 mice. The PCA-LDA and QDA models attained similar accuracy values, with the LDA model achieving a better sensitivity. The PLS-DA achieved the highest sensitivity and specificity values, having thus the better performance. The coefficients of variability were below 3% for all the performance indices in all the models.

Table 3.57: Repeated CV PCA-LDA, PCA-QDA and PLS-DA models classification performance for three months old mdx and C57Bl/10 mice. *The mean sensitivity, specificity and accuracy values from the 100 repetitions are shown. Standard deviation and coefficients of variability are also displayed. CoV: Coefficient of Variability.*

Three months old male mdx/C57Bl/10 mice	***Sensitivity*** ***(+/- Standard Deviation, CoV)***	***Specificity*** ***(+/- Standard Deviation, CoV)***	***Accuracy*** ***(+/- Standard Deviation, CoV)***
PCA-LDA	92.4 (+/-1.6, CoV:1.7%)	88.5 (+/-2.5, CoV:2.8%)	90.5 (+/-1.4, CoV:1.5%)
PCA-QDA	87.8 (+/-1.9, CoV:2.2%)	91.3 (+/-2.4, CoV:2.6%)	89.5 (+/-1.5, CoV:1.7%)
PLS-DA	100% (+/-0, CoV:0%)	97.7% (+/-1.4, CoV:1.4%)	98.9% (+/-0.7, CoV:0.7%)

The ROC curves generated in each of the repetitions and the mean ROC curve for each model for the three months old male *mdx* mice can be seen in figure 3.91. The PCA-LDA and QDA ROC curves had a mean AUC of 0.97 and 0.96 respectively whereas the PLS-DA curves demonstrated a mean AUC of 0.99.

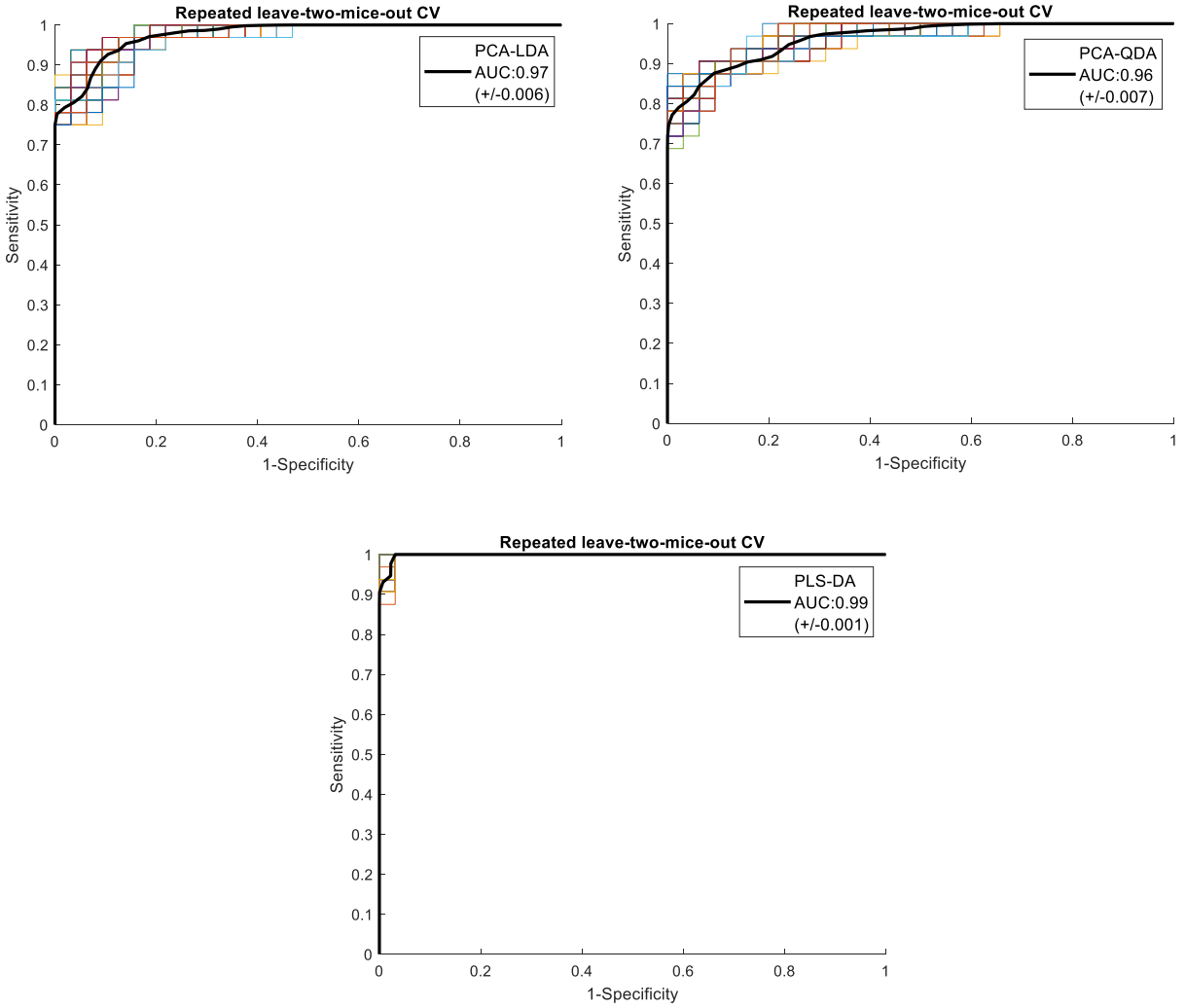


Figure 3.91: ROC curves generated from repeated cross-validation of three months old mdx and C57Bl/10 mice. *ROC curves are shown for each of the hundred repetitions during cross-validation. The mean ROC curve (black line) is also shown. The mean AUC value for each model (+/-) one standard deviation also displayed.*

3.1.5.2 Basis of Classification

3.1.5.2.1 Mean and Difference Spectra

The mean and difference spectra of the different mice groups were plotted to visually examine the most prominent peaks and to identify the major differences between the groups of each model. In figures 3.92 and 3.93 the mean spectra of each group (+/-) standard deviation are shown.

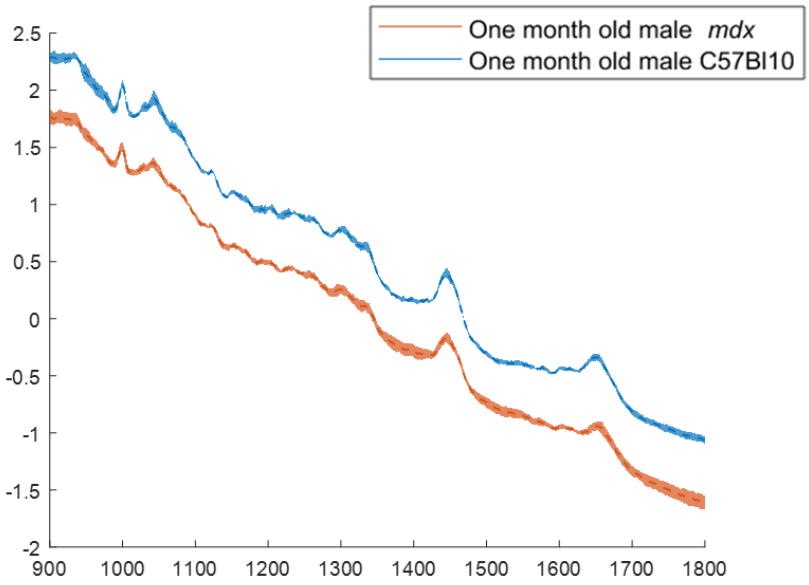


Figure 3.92: Mean Raman spectra for one month old male mdx and C57Bl/10 mice groups.
The mean spectra for mdx and control mice of one month of age are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

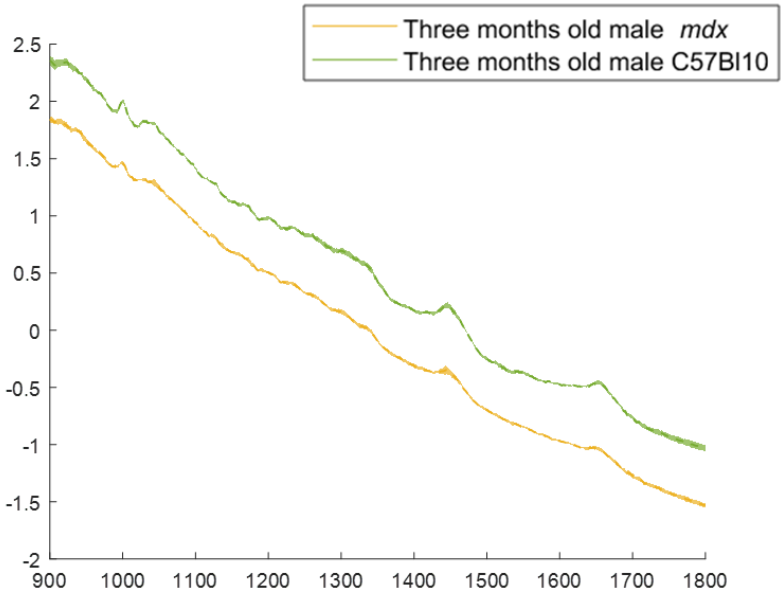


Figure 3.93: Mean Raman spectra for three months old male mdx and C57Bl/10 mice groups.
The mean spectra for mdx and control mice of three months of age are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

Figures 3.94 and 3.95 show the mean and difference spectra of the combinations of groups used to build the two group models. The mean spectra of the different groups consisted of similar peaks. Despite the signal to noise ratio of the spectra acquired with the probe being limited by the fluorescent background, biochemically reliable peaks could be seen in the mean and difference spectra.

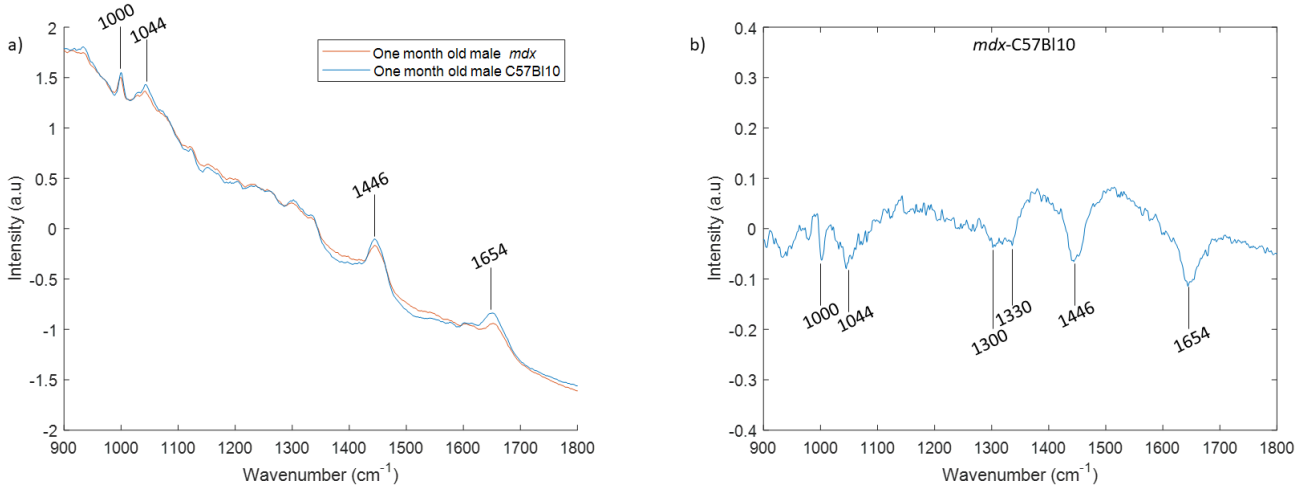


Figure 3.94: Mean and difference spectra of one month old mdx and C57Bl/10 mice.
a) Mean spectra for one month old mdx and C57Bl/10 mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

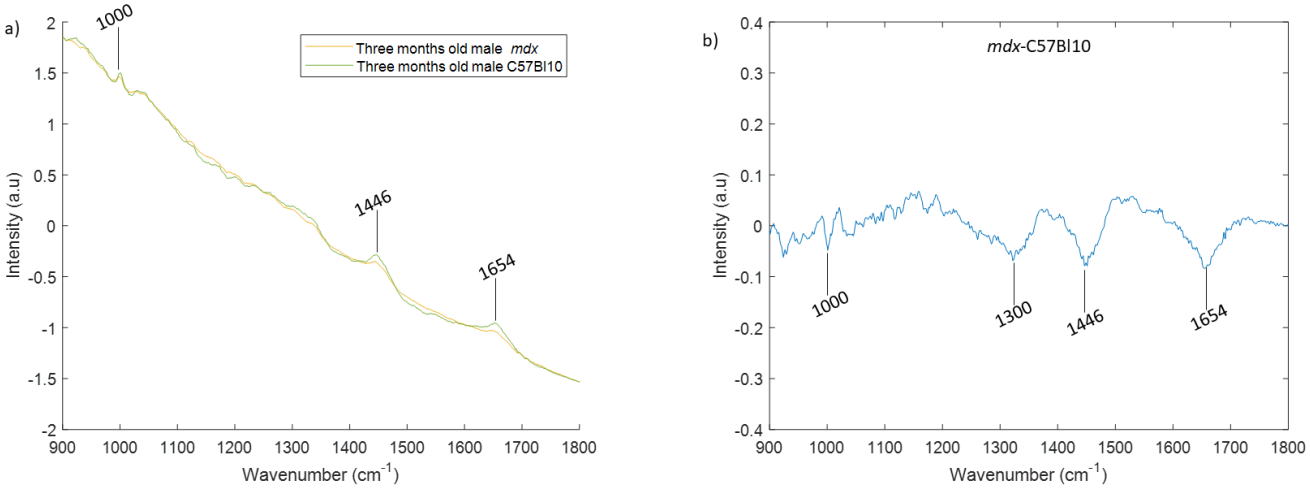


Figure 3.95: Mean and difference spectra of three months old mdx and C57Bl/10 mice.
a) Mean spectra for three months old mdx and C57Bl/10 mice and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

A decrease in the prominent peaks located around 1000 cm^{-1}, 1044 cm^{-1}, 1300 cm^{-1}, 1330 cm^{-1}, 1445 cm^{-1}, 1655 cm^{-1} as the disorder progresses was observed in the difference spectra of the *mdx* mice. Tentative peak assignments for these peaks can be seen in table 3.58.

Table 3.58: Prominent Raman peaks in mean and difference spectra of the mdx and C57Bl/10 mice and tentative peak assignments. *A decrease in these peaks was observed as the disorder progressed for the mdx mice.*

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1444	CH_2 bending	Proteins, Lipids
1654	C=O stretching, C=C stretch	Amide I (proteins), Lipids

Background subtraction was used to aid identification of more spectral features present in the spectra. The background subtracted mean spectra of the different groups can be seen in figures 3.96 and 3.97. A larger number of peaks were identifiable in these mean spectra. The prominent peaks were very similar for the different groups. The peaks between 1240 cm^{-1} and 1340 cm^{-1} were less clearly identifiable in the mean spectrum of the three months old wild-type mice compared to the mean spectra of the other groups.

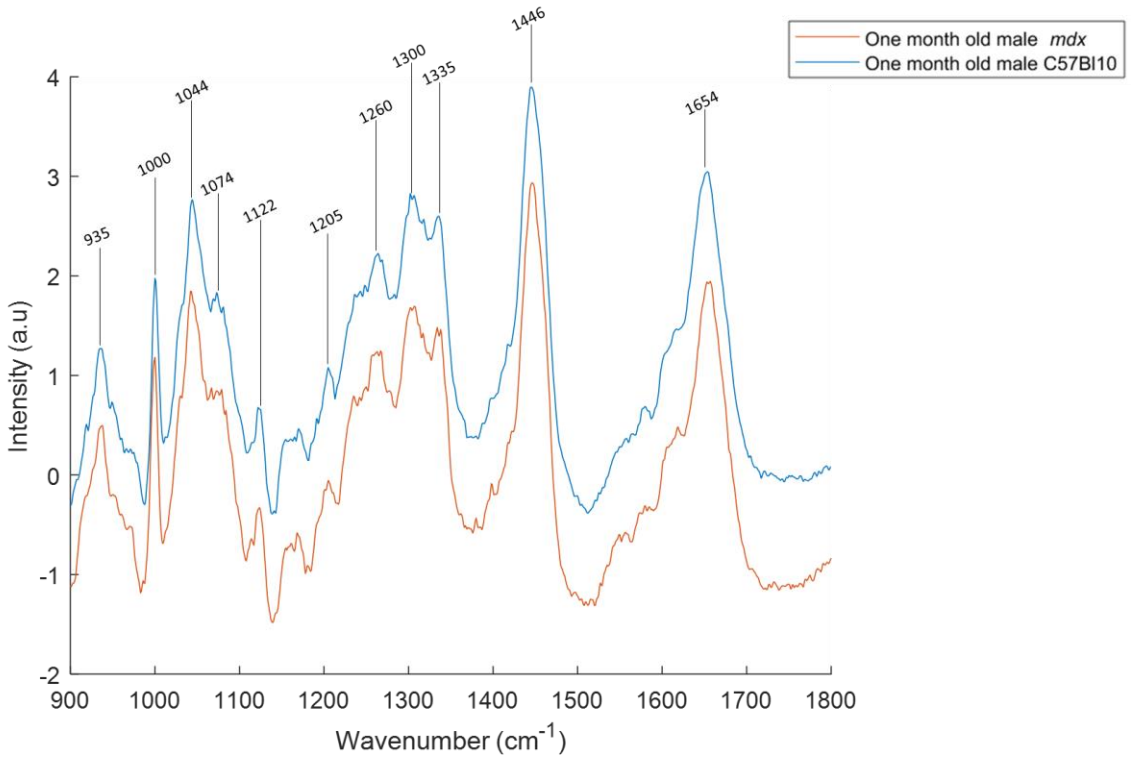


Figure 3.96: Background subtracted mean spectra of one month old male mdx and C57Bl/10 mice groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

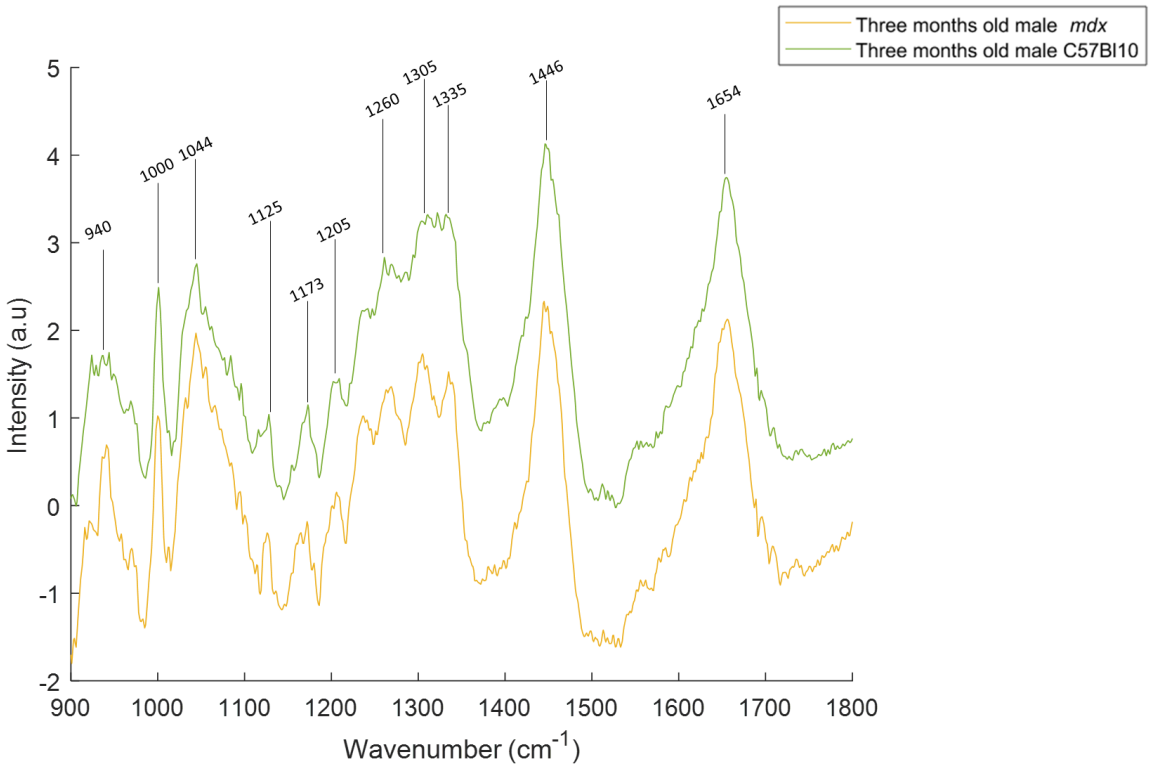


Figure 3.97: Background subtracted mean spectra of three months old male mdx and C57Bl/10 mice groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

Additional shoulder peaks can be seed around 970 cm^{-1}, 1030 cm^{-1}, 1555 cm^{-1}, 1575 cm^{-1} and 1615 cm^{-1}. Tentative peak assignments for the peaks present in the background subtracted mean spectra are presented in table 3.59.

Table 3.59: Prominent Raman peaks in background subtracted spectra and tentative peak assignments. *Amino acids are specified when the peaks are largely associated with them in the literature.*

Wavenumber (cm^{-1})	Vibrational Modes	Assignment
935	C-C stretching	Proteins
950		Proteins (Valine, Proline, Phenylalanine)
975	CH_2 deformation	Proteins, Nucleic Acids
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1030	C-H bending	Proteins (Phenylalanine, Proline)
1044		Proteins
1076	C-C stretching, C-O stretching	Proteins (Tryptophan), Lipids
1121	C-C stretching, C-N stretching	Proteins, Lipid
1170	C-H bending	Proteins (Tyrosine, Phenylalanine)
1205	$\nu(C\text{-}C_6H_5)$	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)
1260	C-N stretching	Amide III (proteins), Lipids
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1444	CH_2 bending	Proteins, Lipids
1550	$\nu(C\text{-}C)$	Proteins (Tryptohan)
1570	Ring breathing modes of the DNA/RNA bases	Nucleic Acids
1615	C=C stretching	Proteins (Tyrosine)
1654	C=O stretching, C=C stretching	Amide I (proteins), Lipids

3.1.5.2.2 Multivariate Analysis

The PCA loadings and the linear discriminant function (LDF) as well as the PLS weights were plotted in order to illustrate the important peaks for spectral classification in the two group models generated using the spectra of one and three months old male *mdx* and wild-type mice.

Similar to the previous studies, a combination of the information found in the difference spectra and the loading/weight plots was used in an attempt to better understand the differences present in the spectra from the different groups. The score histograms and loading/weight plots in which the separation of the spectra from the different groups was best for each model are shown in the following sections. The loading/weight plots for the rest of the PCs and components used to build each model are shown in Appendix C. Tables 3.60 and 3.61 summarise and compare the most prominent peaks found in all the loading/weight plots (from all the components used for each model) and the ones found in the difference spectra of each two-group model.

a) One month old *mdx* vs. C57Bl/10 mice

Only PC2 scores were found to be significantly different between one month old male *mdx* and C57Bl/10 mice (q_{PC2}: $5.4e^{-06}$). Since only PC2 was imported to LDA the PCA and LDA score and loading plots were identical and can be seen in figure 3.98.

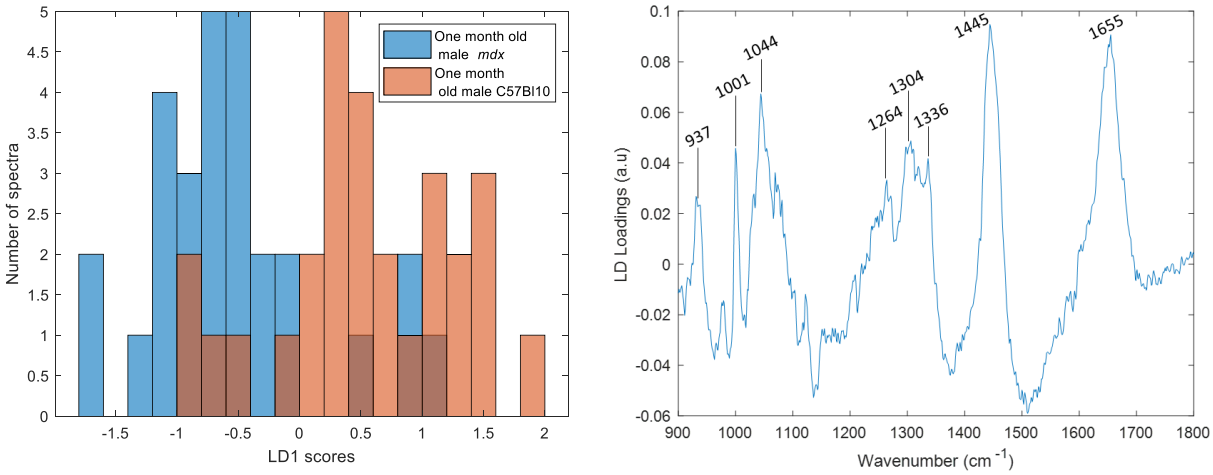


Figure 3.98: LD1 score histogram and LDF plot for the one month old mdx and C57Bl/10 mice.

In the PLS-DA analysis the model generated using the first component demonstrated the highest accuracy value. The score histogram and weights plot of component one can be seen in figure 3.99

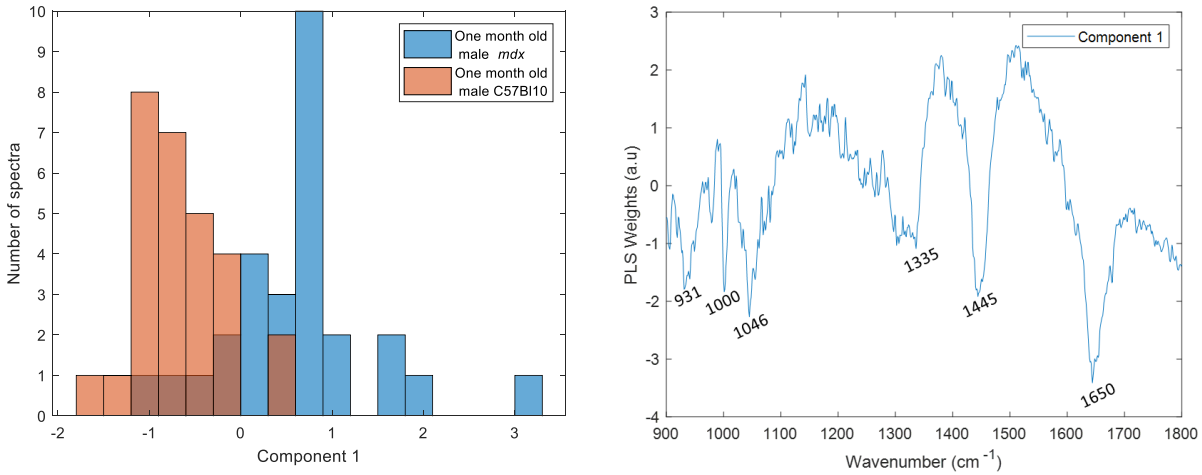


Figure 3.99: Component 1 score histogram and weight plot for the one month old mdx and C57Bl/10 mice.

Even though there was some overlap it can be seen from the histogram of the LD scores that positive LD1 values had a larger contribution from spectra acquired from the C57Bl/10 mice. This could indicate that the peaks located around 936 cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1304 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1} were decreased in the spectra of the *mdx* mice. Similar peaks in the opposite direction were found in the PLS weight plot. These peaks were very similar with the bands discussed in the previous section demonstrating that changes in muscle's biochemical composition were important for discrimination. The peaks found in the loading plots of PC1 and component 1 as well as in the difference spectrum along with their tentative peak assignments are summarised in table 3.60.

Table 3.60: Summary table of the peaks associated with the one month old male mdx/C57Bl/10 mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
935	Proteins	✓	✓	✓	✓	C57Bl/10
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	C57Bl/10
1045	Proteins (Phenylalanine, Proline)	✓	✓	✓	✓	C57Bl/10
1265	Amide III (proteins), Lipids	✓	✓			
1300	Amide III (proteins), Lipids	✓	✓		✓	
1335	Proteins, Nucleic Acids	✓	✓	✓	✓	C57Bl/10
1445	Proteins, Lipids	✓	✓	✓	✓	C57Bl/10
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	C57Bl/10

b) Three months old male *mdx* vs. C57Bl/10 mice

PCs 1 and 2 were found to be significantly different between the three months old *mdx* and C57Bl/10 mice. PC1 demonstrated the largest difference between the different groups followed by PC2 (q_{PC1}: $7.6e^{-09}$, q_{PC2}: 0.005). The LDA histogram and LDF plot using PC1 and PC2 as input variables can be seen in figure 3.101.

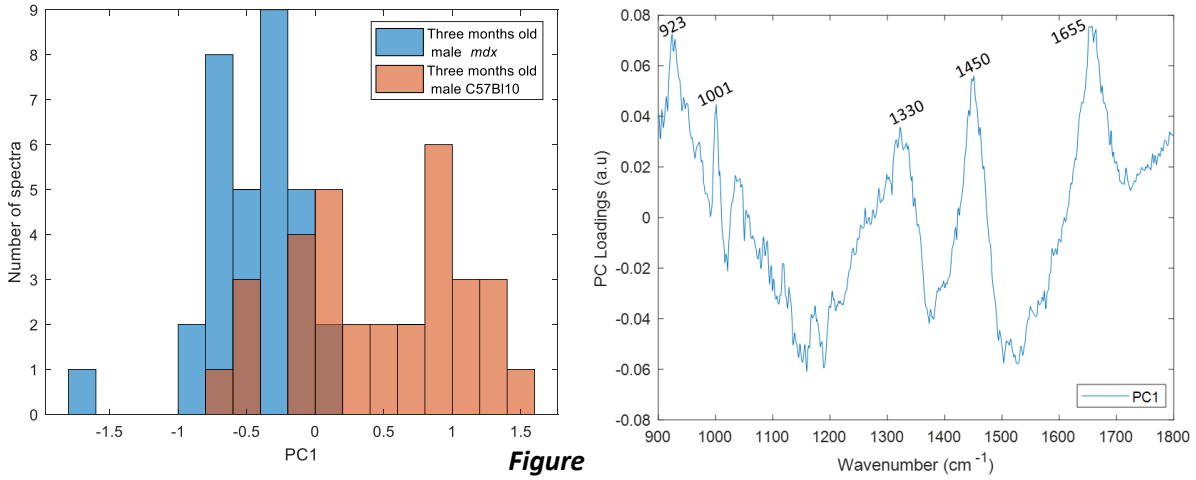


Figure 3.100: PC1 score histogram and loading plot for the three months old male mdx and C57Bl/10 mice.

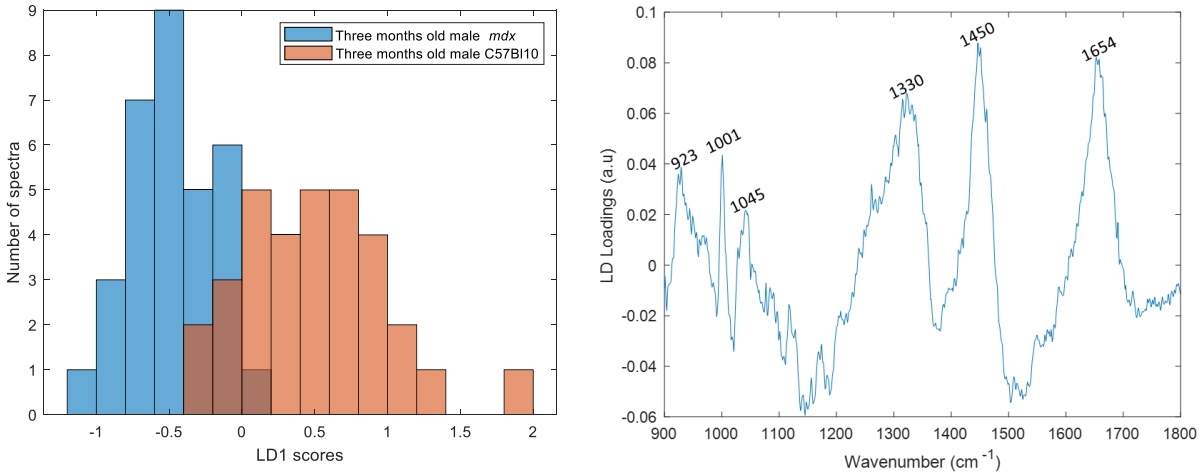


Figure 3.101: LD1 score histogram and LDF plot for the three months old male mdx and C57Bl/10 mice.

The PLS-DA model was built using the first four PLS components. The PLS weights plot for component 1 can be seen in figure 3.102. The PLS score histograms and weight plots for the rest of the components can be found in Appendix C.

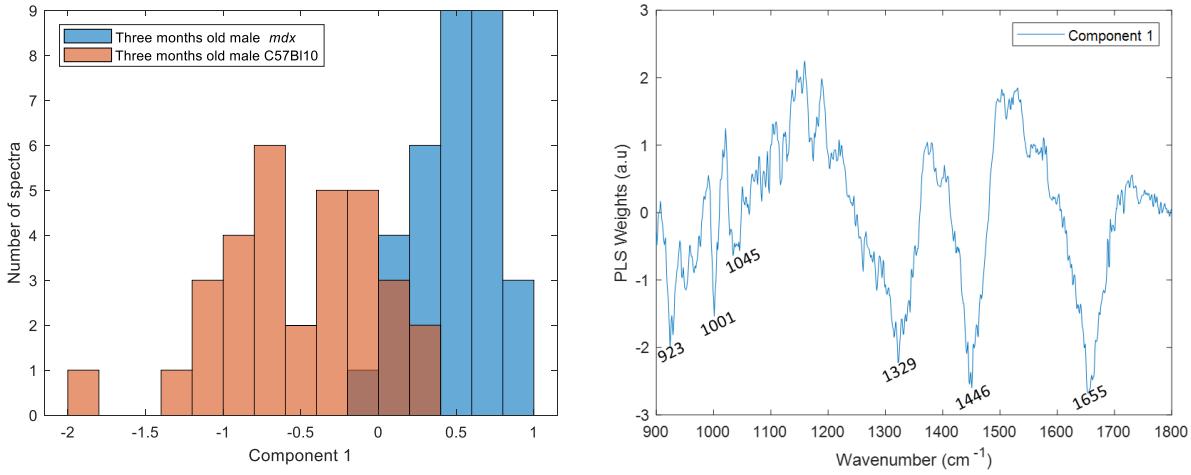


Figure 3.102: Component 1 score histogram and weight plot for the three months old male mdx and C57Bl/10 mice.

Despite some background present in PC1 and PC2 loading plots biochemically reliable peaks were identifiable in both loading plots. Interestingly, less background was present in the LDF plot. Spectra from C57Bl/10 mice had a larger contribution in the positive values in the LD1 histogram and in the negative values in the weight plot of PLS component one. Hence, the peaks around 923 cm^{-1}, 1001 cm^{-1}, 1045 cm^{-1}, 1330 cm^{-1}, 1446 cm^{-1} and 1655 cm^{-1} that were found in the LDF and weight plots in the respective directions were considered as increased in the spectra of the wild-type mice. As can be seen from the histogram of the LDA scores and of the scores from the first PLS component most of the spectra of the three months old mice have positive score values whereas the spectra of the one month old mice have negative score values (for both models). Similar to the findings from study 1, where spectra acquired from female *mdx* and C57Bl/10 mice

were used, the most prominent peaks in the spectra decreased as the disorder progressed. The most prominent peaks in the loading/weight plots of the models and in the difference spectra and their tentative peak assignments are summarised in table 3.61.

Table 3.61: Summary table of the peaks associated with three months old mdx mice and tentative peak assignments. *In the group column, the group that the peaks were more prominent is displayed.*

Wavenumber (cm^{-1})	Tentative Assignment	PCA loading plots	LDF	PLS weight plots	Difference Spectrum	Group
923	Proteins (Proline)	✓	✓	✓	✓	C57Bl/10
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	C57Bl/10
1045	Proteins (Phenylalanine, Proline)	✓	✓	✓		C57Bl/10
1125	Proteins			✓		
1300	Amide III (proteins), Lipids				✓	
1335	Proteins, Nucleic Acids	✓	✓	✓		C57Bl/10
1445	Proteins, Lipids	✓	✓	✓	✓	C57Bl/10
1478	Nucleic Acids			✓		
1655	Amide I (Proteins), Lipids	✓	✓	✓	✓	C57Bl/10

3.1.6 Study 5: Combined Probe Study

3.1.6.1 Electrophysiological Recordings

CMAP amplitude was measured in the hindlimb of $SOD1^{G93A}$ and C57Bl/6 mice with the in-house built combined probe. In order to compare the performance of the probe relative to standard methods, whenever possible CMAP measurements were also undertaken with the standard methods used for electrophysiological recordings (ring electrodes, concentric needle electrode). Figure 3.103 shows the measurements acquired with the different methods for both mice groups. Recorded CMAP amplitude showed no differences between the combined probe and any of the standard methods for the $SOD1^{G93A}$ mice and only differed significantly between the combined probe and ring electrodes for the C57Bl/6 mice. In order to assess the performance of the combined probe the CMAP amplitudes from the measurements acquired with the combined probe for $SOD1^{G93A}$ and the C57Bl/6 mice were compared. As can be seen in figure 3.104 the recordings from the $SOD1^{G93A}$ mice were significantly different ($p<0.001$) from the C57Bl/6 mice. In figure 3.105 the results for the same comparison using the standard methods are shown.

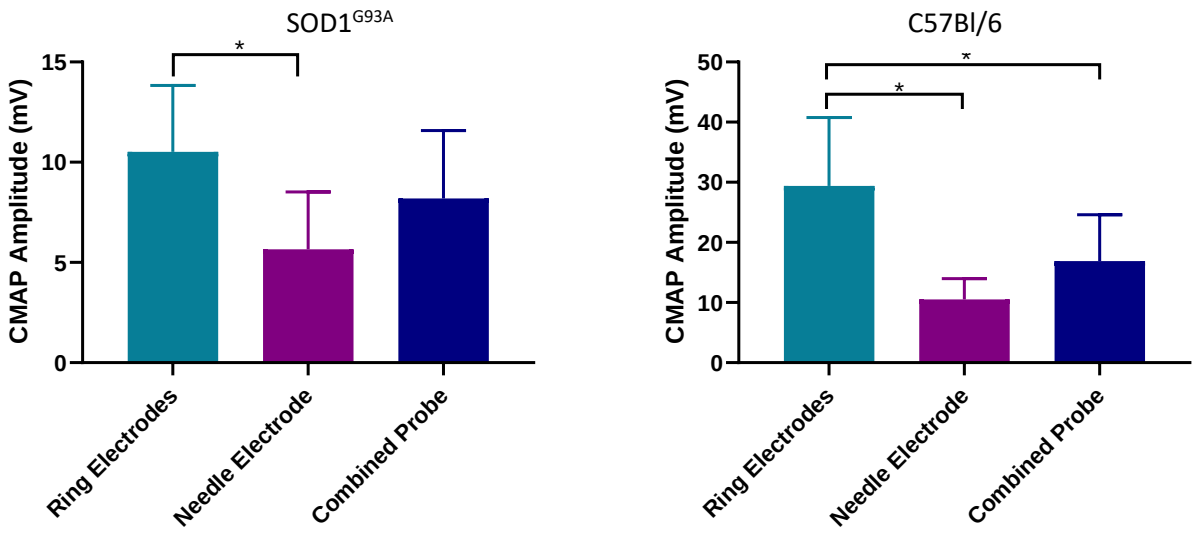


Figure 3.103: CMAP amplitudes for the SOD1^{G93A} and C57Bl/6 mice recorded using three different methods. *Mean and standard deviation shown. Asterisks indicate p<0.05 using Turkey post-test following mixed effect model repeated measures ANOVA.*

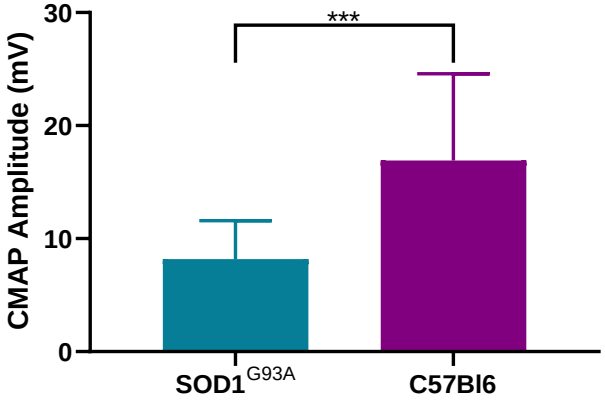


Figure 3.104: CMAP amplitudes for the SOD1^{G93A} and C57Bl/6 mice recorded using the combined probe. *Mean and standard deviation shown. Asterisks indicate p<0.001 using student's t-test.*

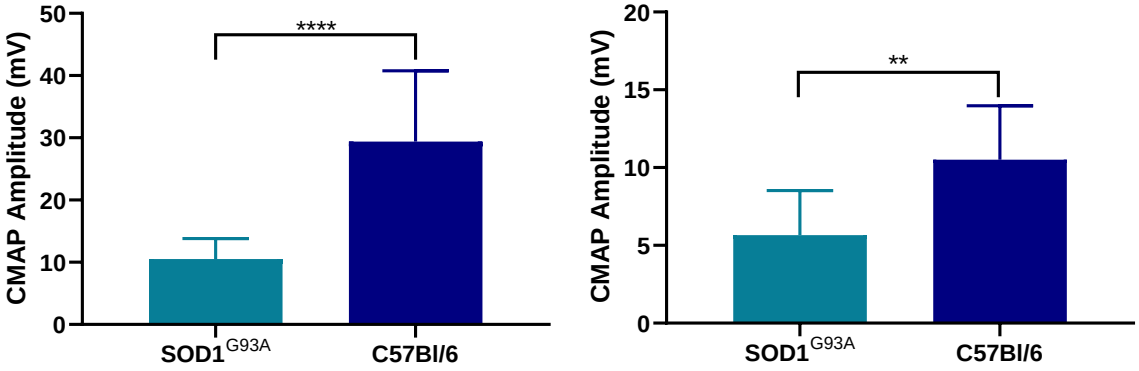


Figure 3.105: CMAP amplitudes for the SOD1^{G93A} and C57Bl/6 mice recorded using standard methods. *CMAP amplitudes recorded with the ring electrodes (left) and the needle electrode (right). Mean and standard deviation shown. Two asterisks indicate p<0.01, four asterisks indicate p<0.0001 using student's t-test.*

3.1.6.2 Raman recordings

Raman recordings were also acquired using the combined probe in order to assess if electrophysiological recordings affected the ability of the probe to detect the Raman signal and if the recorded spectra were different from the spectra recorded using the Raman probe. Background subtracted mean spectra of two groups of three months old SOD1^{G93A} acquired with the combined and Raman probes are shown in figure 3.106. As can be seen in the background subtracted mean spectra the prominent peaks are the same for the recordings acquired with the different probes.

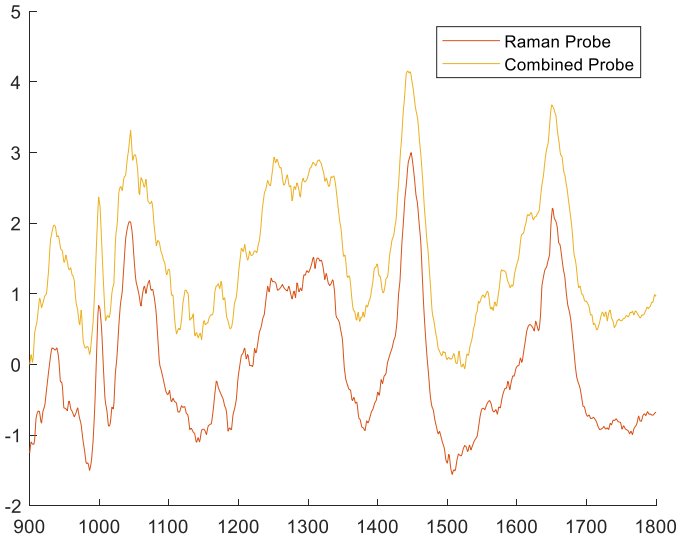


Figure 3.106: Background subtracted mean Raman spectra of SOD1^{G93A} mice acquired using the Raman and combined probes.

Additionally, CMAPs were recorded with the combined probe before and after the Raman measurements in order to investigate the effect of the Raman measurements on the electrophysiological recordings. As can be seen in figure 3.107 the CMAP amplitudes were not significantly different before and after the Raman recordings for any of the groups.

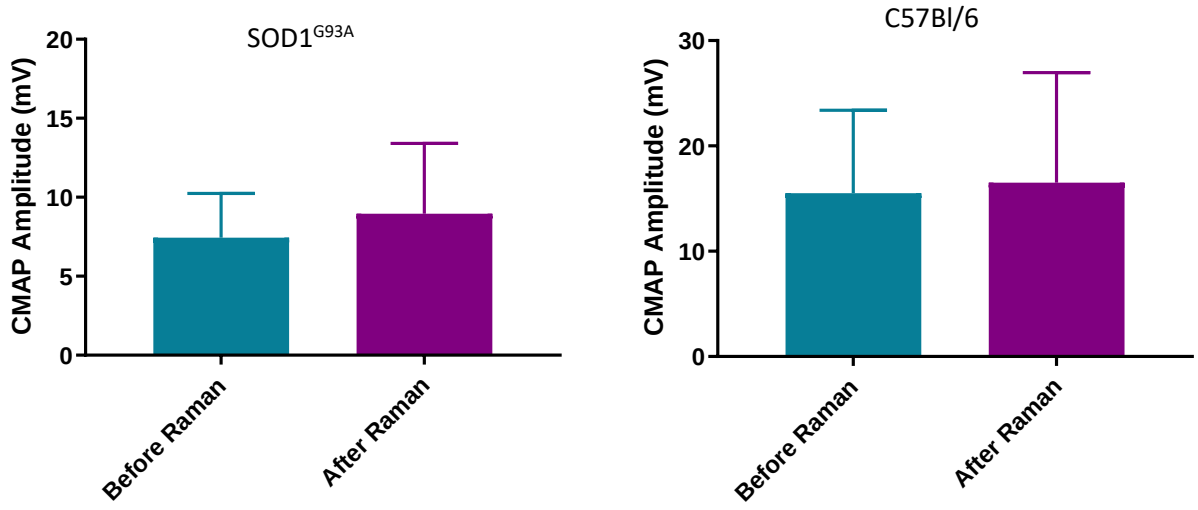


Figure 3.107: CMAP amplitude for SOD1^{G93A} and C57Bl/6 mice before and after the Raman recordings.

3.1.7 Comparing spectra from different tissues

A small number of Raman spectra were recorded from femur and tibia bones as well as from blood samples acquired from three months old *mdx* mice in order to examine their Raman signature using the fibre optic probe and compare it with that of muscle. The mean spectra can be seen in figure 3.108.

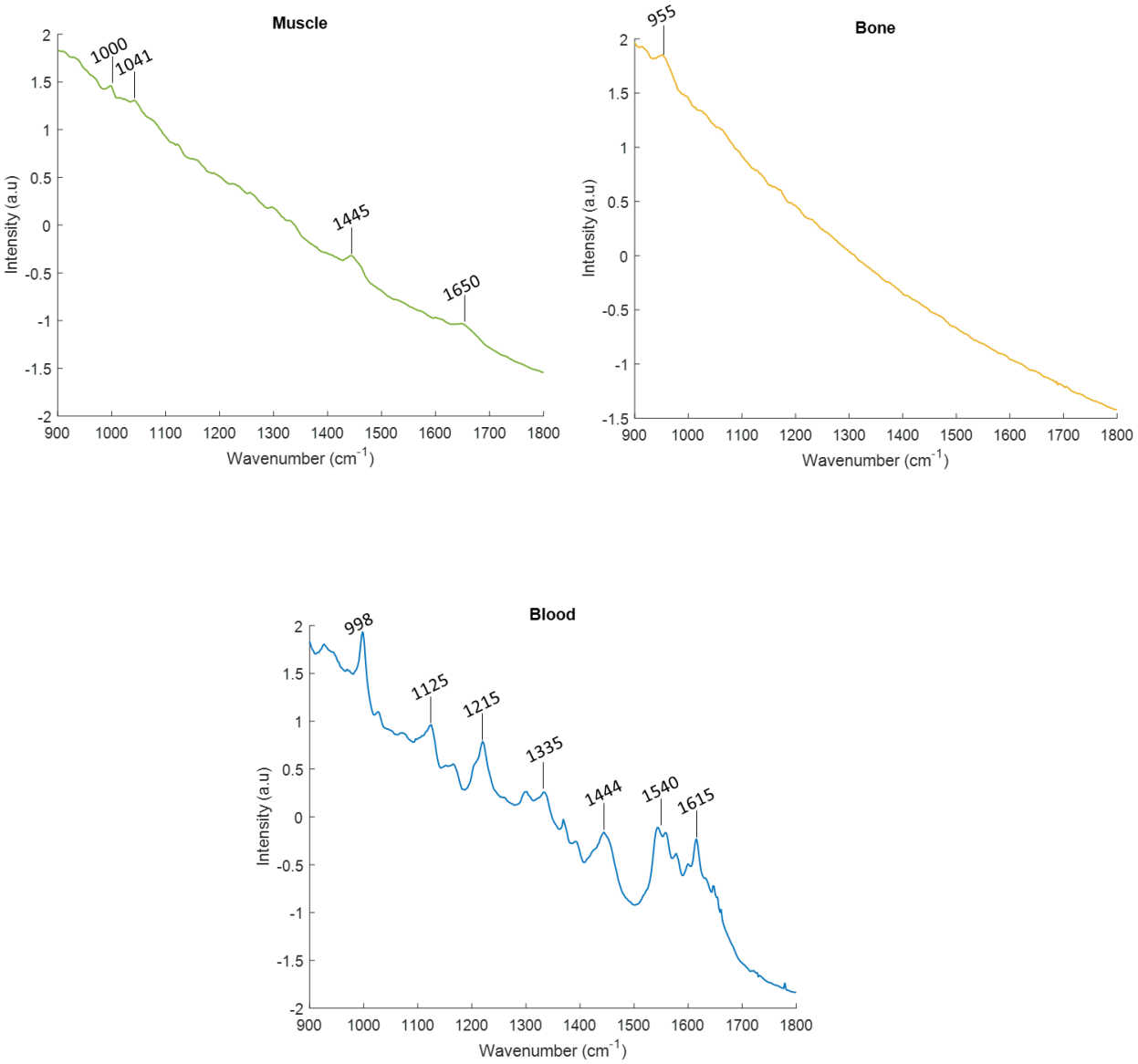


Figure 3.108: Mean Raman spectra from muscle, bone and blood of three months old mdx mice. *The most prominent peaks in the mean spectra of the different organs and blood are indicated.*

Despite the signal to noise ratio of the spectra acquired with the probe being limited by the fluorescent background, biochemically reliable peaks could be seen in the mean spectra. Additionally, clear visual differences were apparent in the mean spectra of the different components. The main peaks found in the different mean spectra and their tentative assignments are summarised in table 3.62. The references can be found in Appendix A.

Table 3.62: Prominent peaks of spectra acquired from muscle, bone and blood and tentative peak assignments.

Wavenumber (cm^{-1})			Tentative Assignment
Muscle	**Bone**	**Blood**	
	955		Hydroxyapatite
		998	Phenylalanine
1000			Phenylalanine
1041			Proline
		1125	Proteins
		1215	Proteins
		1335	Proteins, Nucleic Acids
1445		1445	Proteins, Lipids
		1545	Tryptophan
		1615	Tyrosine, Tryptophan, Phenylalanine
1650			Amide I (proteins), Lipids

The spectra recorded from bone (positive score values) were clearly separated from the spectra acquired from muscle (negative score values) in the PC1 histogram as can be seen in figure 3.109. The peaks identified in the loading plot (figure 3.109) were the same peaks with the ones found in the mean spectra of the different tissues, with the positive peak around 955 cm^{-1} being associated with the bone spectra and negative peaks around 1000 cm^{-1}, 1044 cm^{-1}, 1445 cm^{-1} and 1652 cm^{-1} being present in the muscle spectra.

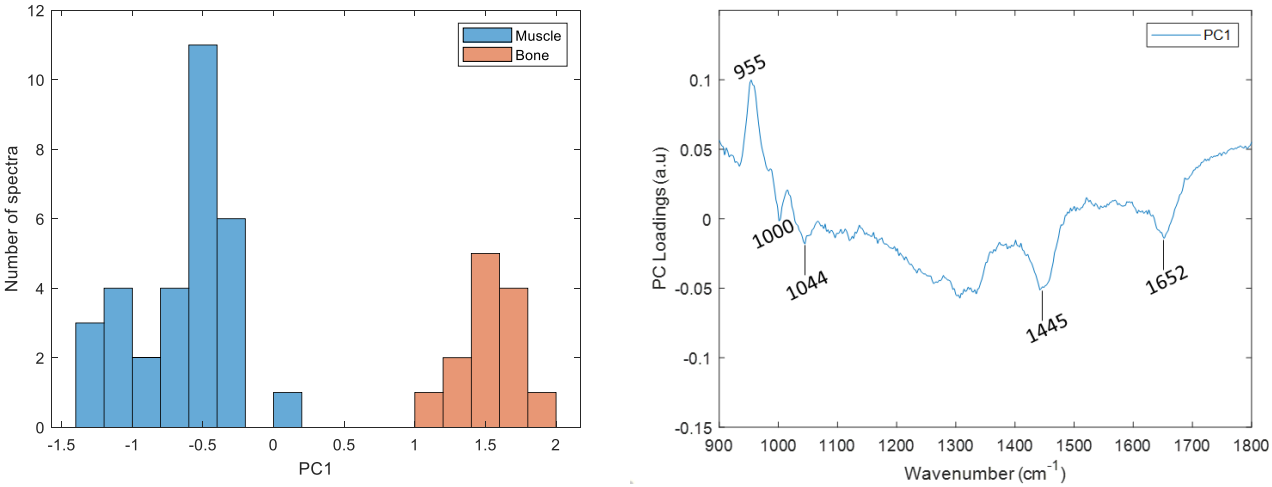


Figure 3.109: PC1 score histogram and loading plot following PCA on the spectra measured from muscle and bone.

Similarly, blood spectra were clearly separated from the muscle spectra in the histogram plot of PC1 scores (figure 3.110). The loading plot was dominated by the much larger peaks found in the blood spectra and was thus very similar with the mean blood spectrum.

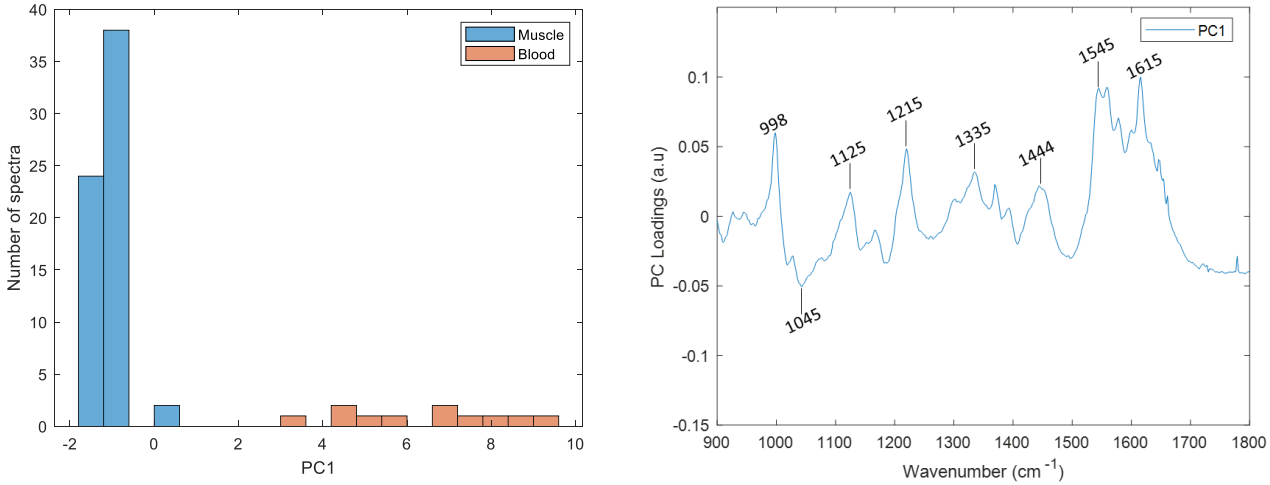


Figure 3.110: PC1 score histogram and loading plot following PCA on the spectra measured from muscle and blood.

Both leave one-spectrum-out cross-validated PCA-LDA models (muscle/bone, muscle/blood) achieved sensitivity, specificity and accuracy values of 100%.

3.2 Human Tissue Recordings

This study aimed to test the ability of the fibre-optic probe to record Raman spectra from *ex vivo* human muscle and to explore the potential of the technique to separate between different neuromuscular disorders. Additionally, Raman recordings were also obtained using a Raman microscope system in order to compare the performance of the fibre-optic probe with a more widely used Raman system. In section 3.1.2.1 the demographic and clinical characteristics of the subjects from which the muscle biopsies were acquired as well as the numbers of the muscle samples are summarised. A summary of the Raman spectra acquired in this study is presented in section 3.1.2.2.

3.2.1 Summary of experiments

3.2.1.1 Demographic and Clinical Characteristics

A total of 77 muscle samples were used in this study. The demographic data of the patients with mitochondrial disorders, MND and healthy volunteers can be seen in tables 3.63 to 3.65 respectively. In table 3.66 the demographic and clinical characteristics of the patients that participated in the prospective muscle collection are presented. In that table muscle biopsies are placed into groups based on the biopsy findings. The 'myopathy' group in the prospective muscle collection consisted of muscle samples from patients with different types of myopathies. The demographic data of the patients with the different types of myopathies are presented in table 3.67.

Table 3.63: Demographic and clinical characteristics of the patients with mitochondrial disorders. *The muscle that was sampled from each patient is also shown.*

Variables	**Whole Group (n=15)**	**M.3243A>G (n=11)**	**POLG-related (n=3)**	**Single mtDNA deletion (n=1)**
Sex				
Male	9 (60%)	6	2	1
Female	6 (40%)	5	1	-
Age				
Mean (yrs)	49	46	66.7	29
Range (yrs)	25-80	25-62	56-80	-
Muscle				
Quadriceps	1	1	-	-
Tibialis Anterior	11	8	2	1
Vastus Lateralis	2	2	-	-
Hamstrings	-	-	-	-
Deltoid	1	-	1	-
Heteroplasmy (%)	65.2	65.2	N/A	N/A

Table 3.64: Demographic characteristics and muscle sampled from healthy volunteers.

Variables	**Healthy (n=10)**
Sex	
Male	6 (60%)
Female	4 (40%)
Age	
Mean (yrs)	33
Range (yrs)	17-54
Muscle	
Quadriceps	-
Tibialis Anterior	-
Vastus Lateralis	-
Hamstrings	10

Table 3.65: Demographic and clinical characteristics of the patients with MND. *The muscle that was sampled from each patient is also shown.*

Variables	**MND** ($n_{patients}$=14, $n_{samples}$=16)
Sex	
Male	6 (42%)
Female	6 (42%)
Unknown	2
Age	
Mean (yrs)	63.7
Range (yrs)	39-81
Unknown	2
Disease Duration	
Mean (mo)	28.9
Range (mo)	10-121
Unknown	2
Muscle	
Quadriceps	9
Biceps	7

Table 3.66: Demographic and clinical characteristics of the patients that participated in the prospective muscle collection in Sheffield. The patients are grouped using the biopsy findings. *The muscle that was sampled from each patient is also shown.*

Variables	**Whole Group (n=38)**	**Myopathic (n=15)**	**Age-related (n=4)**	**Non-specific (n=12)**	**Normal (n=7)**
Sex					
Male	19 (50%)	10	1	3	5
Female	19 (50%)	5	3	9	2
Age					
Mean (yrs)	53	51.5	69.3	55.5	42.6
Range (yrs)	20-80	22-73	59-80	31-77	20-76
Muscle					
Quadriceps	13	7	2	1	3
Biceps	10	2	-	6	2
Deltoid	15	6	2	5	2
Main Symptoms					
Proximal weakness	17	13	1	3	-
Ataxia	9	-	1	6	2
Myalgia	4	1	-	2	1
Incidental CK	3	-	-	-	3
Ptosis	1	-	1	-	-
Dysarthria/Dysphagia	2	-	1	-	1
Muscle stiffness/spasms	1	-	-	1	-
Muscle hypertrophy	1	1	-	-	-
Cognitive decline	1	-	1	-	-

Table 3.67: Demographic and clinical characteristics of the patients with myopathy that participated in the prospective muscle collection in Sheffield. The patients are grouped using the biopsy findings. *The muscle that was sampled from each patient is also shown.*

Variables	Myopathy (n=15)	Muscular dystrophy (n=4)	Metabolic myopathy (n=1)	Inclusion body myositis (n=1)	Statin induced myopathy (n=1)	Inflammatory myopathy (n=3)	Unspecified myopathic findings (n=5)
Sex							
Male	10 (66.7%)	3	1	1	0	2	2
Female	5 (33.3%)	1	0	0	1	1	3
Age							
Mean (yrs)	51.5	45	45	56	55	50.3	60.4
Range (yrs)	22-73	22-72	-	-	-	25-67	35-73
Muscle							
Quadriceps	7	2	-	1	1	2	1
Biceps	2	-	1	-	-	-	1
Deltoid	6	2	-	-	-	1	3

The samples from the prospective muscle biopsy without neuromuscular disease pathology (table 3.67: age-related, non-specific changes and normal) were collectively grouped as 'not myopathy' for the following data analysis. Similarly, the different types of myopathies were all grouped as a single 'myopathy' group. Finally, the samples acquired from healthy volunteers will be referred to as 'healthy' in the following sections.

3.2.1.2 Raman recordings

A total of 2500 spectra were recorded from the human muscle samples *ex vivo* using the fibre-optic Raman probe. Ten spectra were acquired in each site and were then averaged prior the analysis. Thus, 250 spectra were used to generate the PCA-LDA and PLS-DA models. A total of 286 spectra were acquired using the Raman microscope. A summary of the spectra acquired for each

of the different pathological groups, as these were grouped for the subsequent data analysis can be seen in Table 3.68.
It is also important to note that for the MND group the spectra acquired with the probe and the microscope were from different samples. Four samples were examined only using the fibre-optic probe. Recordings from the other ten samples were obtained with both the probe and microscope. However, the probe spectra were all saturated due to fluorescent signal and it was therefore not possible to use them.

Table 3.68: Summary table of ex vivo Raman recordings acquired from human muscle samples.

Group	**Number of samples**	**Number of Probe Spectra**	**Number of Microscope Spectra**
1. Myopathy	15	56	54
2. Not myopathy	20	59	77
3. Samples with biopsy result and final diagnosis not in agreement (Grouped as 'Not Myopathy' in 3.2.2 sections / Grouped as 'Myopathy' in 3.2.3 sections)	3	8	10
4. Mitochondrial Disorders	15	41	53
5. MND	Probe Samples: 4 Microscope Samples: 10	87	18
6. Healthy	10	35	38

3.2.2 Biopsy findings

In the samples acquired from the open muscle biopsies in Sheffield, three samples were from patients with a final clinical diagnosis of myopathy but who had a negative biopsy. In the following sections these samples are labelled as 'not myopathy', according to the biopsy findings.

3.2.2.1 Classification models performance

The two group models generated using the spectra of one pathological group (Myopathy, Mitochondrial Disorders, MND) and the spectra acquired from the 'healthy' muscle samples aimed to explore the ability of the technique to detect muscle pathology of different origins and separate diseased from healthy muscle. The classification performance of these two group models generated using the probe and microscope spectra are presented in sections a-c. Subsequently, the potential of the technique to be used as a screening test for neuromuscular disorders was investigated with multi-group models. The five-group models presented in section d aimed to test the ability of the *ex vivo* recordings to accurately classify samples in the three different pathological groups, the group with no neuromuscular condition ('not myopathy' group) and the 'healthy' group. Finally, the models presented in the last two sections aimed to explore the potential of the technique to separate patients with neuromuscular conditions who would necessitate a biopsy to further investigate the type of their disease (mitochondrial disorders, myopathies) from patients that would not need a biopsy (MND, 'not myopathy' group), reducing, thus, the number of unnecessary invasive procedures.

a) Myopathy vs. Healthy

The predictive capability of the models generated using the probe and microscope spectra can be seen in tables 3.69 and 3.70 respectively. The ROC curves comparing the performance of the two different analysis and CV methods can be seen in figure 3.111 for the probe models and in figure 3.112 for the microscope ones.

As can be seen in table 3.69, when using LOSOCV the ability of both the PCA-LDA and PLS-DA probe models to correctly classify spectra decreased with both accuracy values decreasing by approximately 10%. For both CV approaches the PLS-DA model performed better than the PCA-LDA ones, mainly due to much better specificity values.

Table 3.69: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'myopathy' and 'healthy' samples. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	77.8%	52.6%	67.4%
Leave-one-sample-out CV	66.7%	44.7%	57.6%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	81.5%	97.4%	88.0%
Leave-one-sample-out CV	70.4%	86.8%	77.2%

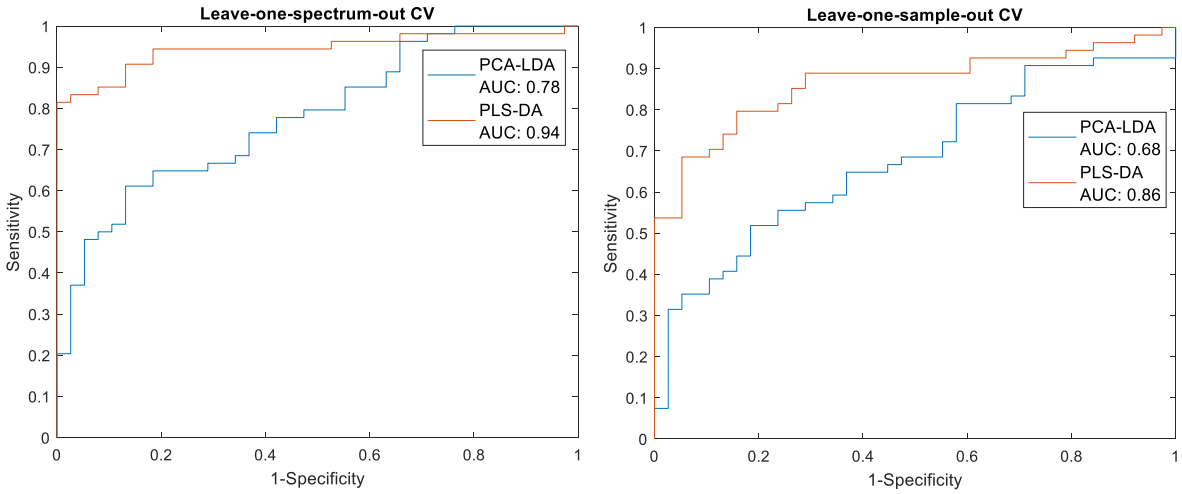


Figure 3.111: ROC curves for the cross validated classification models for the 'myopathy' and 'healthy' samples. *ROC curves for all models generated using the probe spectra using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

For the spectra acquired using the microscope Raman system only the PLS-DA models were generated since there were no PCs with significantly different scores between the two groups and hence PCA-LDA was not undertaken. As can be seen in table 3.70 using LOSOCV the PLS-DA model attained decreased accuracy mainly due a drop in the sensitivity. The AUC was 0.75 for both CV approaches.

Table 3.70: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'myopathy' and 'healthy' samples. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	No significant PCs	No significant PCs	No significant PCs
Leave-one-sample-out CV	No significant PCs	No significant PCs	No significant PCs

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	66.1%	68.6%	67.0%
Leave-one-sample-out CV	55.5%	65.7%	60.0%

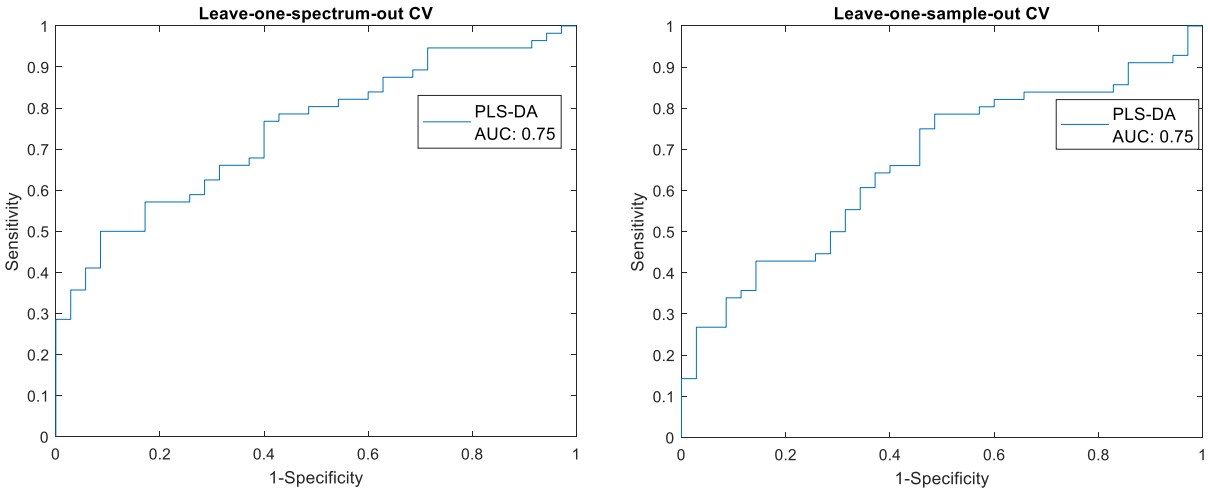


Figure 3.112: ROC curves for the cross validated classification models for the 'myopathy' and 'healthy' samples. *ROC curves for the PLS-DA models generated using the microscope spectra using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

In general, the models generated using the probe spectra performed better than the ones generated using the microscope recordings. The LOMO cross-validated PLS-DA probe model demonstrated a good performance for separating between 'myopathy' from 'healthy' samples, achieving an accuracy of 77% whereas the microscope model demonstrated an accuracy of 60%, with a sensitivity of only 55.5%.

b) Mitochondrial Disorders vs. Healthy

The predictive capability of the models generated using the probe and microscope spectra can be seen in tables 3.71 and 3.72 respectively. The ROC curves comparing the performance of the two different analysis and CV methods can be seen in figure 3.113 for the probe models and in figure 3.114 for the microscope ones.

All the models generated using the probe spectra (table 3.71) attained much better sensitivity than specificity values, with all sensitivity values being above 85%, indicating that it was easier for the models to correctly classify the diseased spectra. The different cross-validation approaches did not have a big impact on the performance of the PCA-LDA model. The LOSO cross-validated PLS -DA model demonstrated a decreased ability to correctly classify spectra mainly due to a big drop in specificity. The LOSOCV PCA-LDA and PLS-DA models achieved similar accuracy values, with the PLS-DA model having slightly more balanced sensitivity and specificity values and a slightly better AUC.

Table 3.71: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'mitochondrial disorders' and 'healthy' samples. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	88.7%	63.2%	78.0%
Leave-one-sample-out CV	88.7%	60.5%	76.9%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	85.9%	76.3%	81.3%
Leave-one-sample-out CV	85.0%	63.2%	75.8%

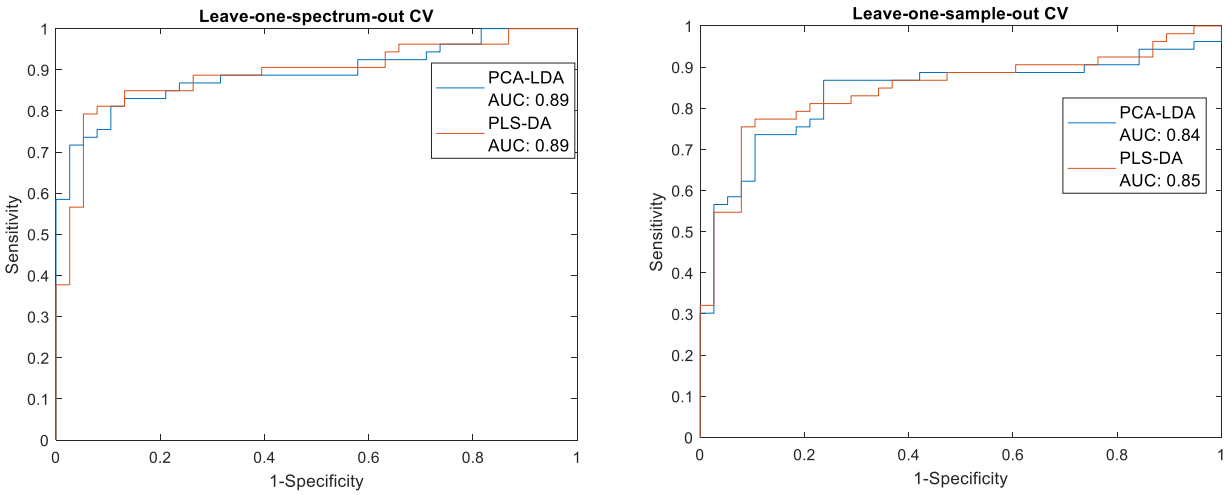


Figure 3.113: ROC curves for the cross validated classification models for the 'mitochondrial disorders' and 'healthy' samples. *ROC curves for all probe models using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

Using LOSOCV did not have a big impact on the performance of the PCA-LDA model, which for both methods achieved much better specificity than sensitivity values. The different CV methods did have a larger impact on the PLS-DA model, with the leaving-one-sample-out approach yielding a considerably lower sensitivity value. Both LOSO cross-validated models attained an accuracy of 75% but the PCA-LDA model achieved that due to much higher sensitivity value (85.4%) whereas the PLS-DA model demonstrated much more balanced sensitivity and specificity values (both around 75%). The AUC was 0.83 and 0.81 for the LOSOCV PCA-LDA and PLS-DA models respectively.

Table 3.72: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'mitochondrial disorders' and 'healthy' samples. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	85.4%	65.7%	76.3%
Leave-one-sample-out CV	82.9%	65.7%	75.0%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	85.4%	77.1%	81.6%
Leave-one-sample-out CV	75.6%	74.3%	75.0%

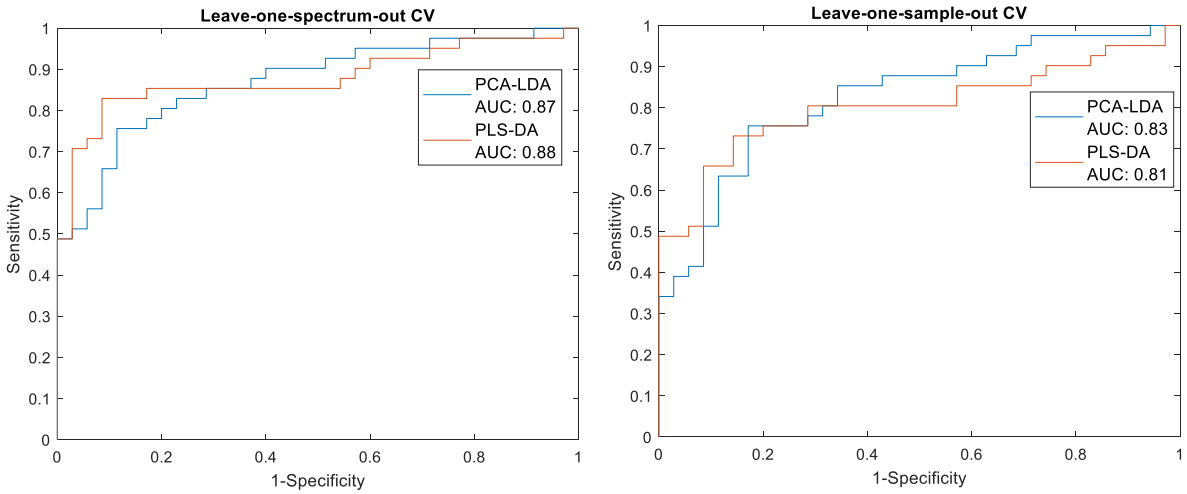


Figure 3.114: ROC curves for the cross validated classification models for the 'mitochondrial disorders' and 'healthy' samples. *ROC curves for all microscope models using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

The leave-one-sample-out PCA-LDA probe and microscope models demonstrated similar performances, with both models achieving much higher sensitivity (above 80%) than specificity values (between 60% and 66%). The LOSO cross-validated PLS-DA probe and microscope models achieved similar accuracy values, around 75%. However, the microscope model demonstrated more balanced sensitivity and specificity values (both around 75%).

c) MND vs. Healthy

The predictive capability of the models generated using the probe and microscope spectra can be seen in tables 3.73 and 3.74 respectively. The ROC curves comparing the performance of the two different analysis and CV methods can be seen in figure 3.115 for the probe models and in figure 3.116 for the microscope ones.

All the probe models yielded sensitivity, specificity and accuracy values of 100% and an AUC value of 1 when separating between 'MND' and 'healthy' samples.

Table 3.73: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'MND' and 'healthy' samples. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	100%	100%	100%
Leave-one-sample-out CV	100%	100%	100%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	100%	100%	100%
Leave-one-sample-out CV	100%	100%	100%

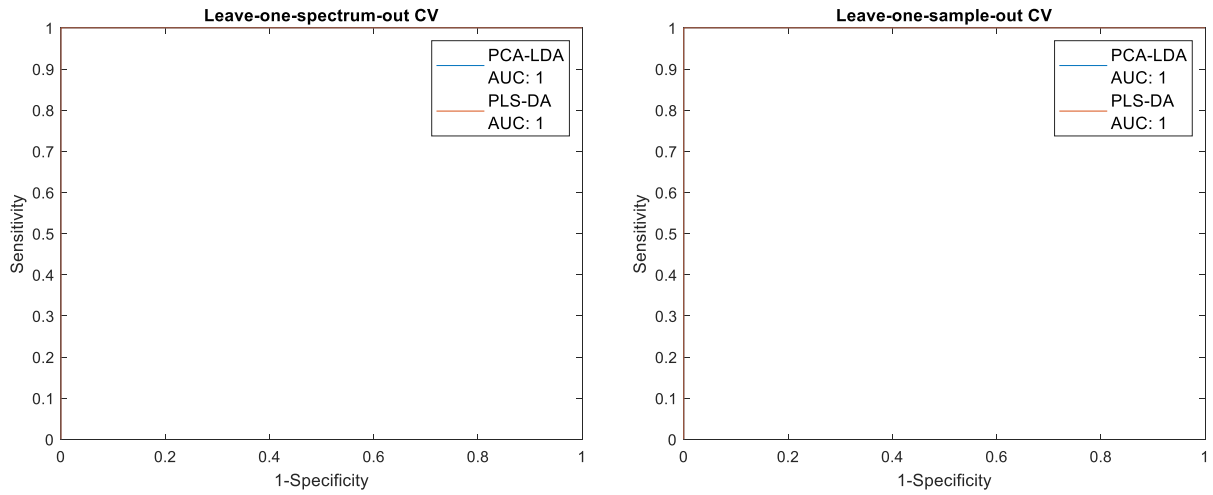


Figure 3.115: ROC curves for the cross validated classification models for the 'MND' and 'healthy' samples. *ROC curves for all probe models using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

When LOSOCV was used the PCA-LDA microscope model demonstrated a decreased ability to correctly classify spectra mainly due to a drop in specificity. The sensitivity of the PLS-DA model decreased slightly but the specificity increased, leading, thus, to an unchanged accuracy value. All models maintained a sensitivity value above 93%. For LOSO cross-validation the PLS-DA model achieved the best performance, with an accuracy value of 95.1% and an AUC value of 0.99.

Table 3.74: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'MND' and 'healthy' samples. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	98.9%	88.6%	95.9%
Leave-one-sample-out CV	95.4%	74.3%	89.3%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	94.3%	97.1%	95.1%
Leave-one-sample-out CV	93.1%	100%	95.1%

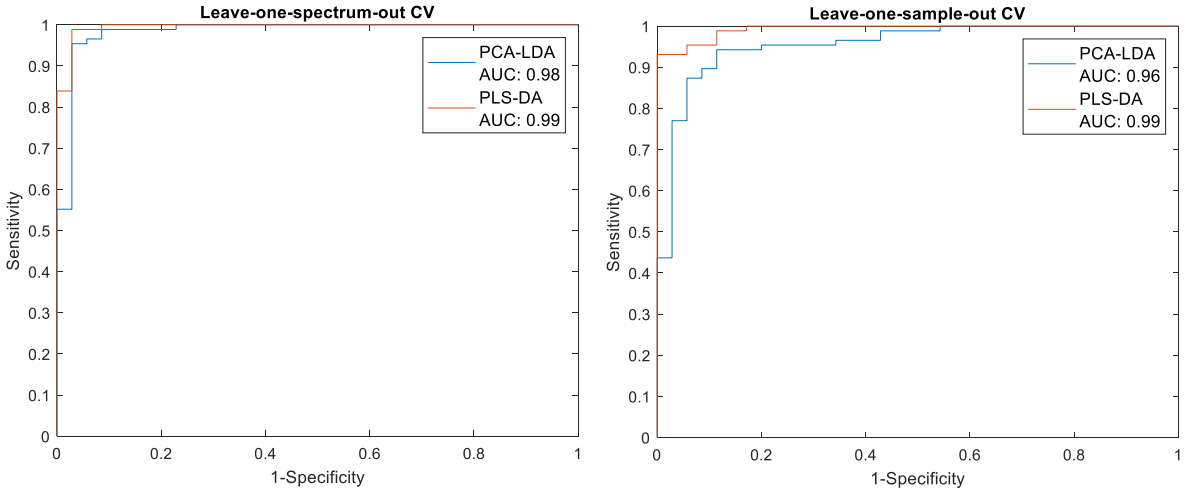


Figure 3.116: ROC curves for the cross validated classification models for the 'MND' and 'healthy' samples. *ROC curves for all microscope models using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

d) Myopathy vs. Mitochondrial Disorders vs. MND vs. Not myopathy vs. Healthy

The PCA-LDA and PLS-DA five group models generated using all the spectra acquired from the *ex vivo* recordings using the probe and the microscope can be seen in tables 3.75 and 3.76 respectively. As can be seen in table 3.75 leave-one-sample-out CV caused a decrease in the ability of the probe models to correctly classify the spectra (decreased accuracy values for most of the groups). The LOSO cross-validated PCA-LDA model achieved much higher specificity than sensitivity values for the 'myopathy' 'not myopathy' and 'healthy' groups, with sensitivity values below 50% for these three groups. The model demonstrated a better ability to correctly classify the spectra acquired from 'mitochondrial disease' samples (66% sensitivity) and achieved a sensitivity and specificity of 100% for the 'MND' group. The PLS-DA LOSO model demonstrated similar performance, with the 'mitochondrial disorders' and 'MND' groups also yielding the highest accuracy values and much lower sensitivity than specificity values for the rest of the groups.

Table 3.75: Five group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 33.3%	LOOCV: 88.8%	LOOCV: 76.8%
	LOSOCV: 20.4%	LOSOCV: 83.2%	LOSOCV: 69.6%
Mitochondrial Disorders	LOOCV: 73.6%	LOOCV: 87.8%	LOOCV: 84.8%
	LOSOCV: 66.0%	LOSOCV: 80.7%	LOSOCV: 77.6%
MND	LOOCV: 100%	LOOCV: 100%	LOOCV: 100%
	LOSOCV: 100%	LOSOCV: 100%	LOSOCV: 100%
Not myopathy	LOOCV: 69.0%	LOOCV: 74.2%	LOOCV: 72.4%
	LOSOCV: 47.1%	LOSOCV: 71.2%	LOSOCV: 62.8%
Healthy	LOOCV: 42.1%	LOOCV: 94.8%	LOOCV: 86.8%
	LOSOCV: 34.2%	LOSOCV: 93.4%	LOSOCV: 84.4%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 22.3%	LOOCV: 87.6%	LOOCV: 73.6%
	LOSOCV: 38.9%	LOSOCV: 83.2%	LOSOCV: 73.6%
Mitochondrial Disorders	LOOCV: 62.3%	LOOCV: 87.8%	LOOCV: 80.8%
	LOSOCV: 75.5%	LOSOCV: 79.7%	LOSOCV: 78.8%
MND	LOOCV: 100%	LOOCV: 100%	LOOCV: 100%

	LOSOCV: 88.9%	LOSOCV: 100%	LOSOCV: 99.2%
Not myopathy	LOOCV: 55.1%	LOOCV: 74.2%	LOOCV: 67.6%
	LOSOCV: 54.0%	LOSOCV: 85.3%	LOSOCV: 74.4%
Healthy	LOOCV: 52.6%	LOOCV: 88.2%	LOOCV: 82.8%
	LOSOCV: 31.6%	LOSOCV: 92.0%	LOSOCV: 82.8%

The microscope models yielded very unbalanced sensitivity and specificity values with much higher specificity than sensitivity values for most of the groups. Since the one-versus-all approach was used to build these models the very high specificity values led to high accuracy. The leave-one-sample out cross-validated microscope models achieved very low sensitivity values (between 8.6% and 45.7%) for all the groups apart from the 'MND' group, for which sensitivity values above 95% were attained. In general, the LOSO models generated using the microscope spectra achieved lower sensitivity but higher specificity values than the models generated using the probe spectra for most of the groups.

Table 3.76: Five group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 28.6%	LOOCV: 93.5%	LOOCV: 80.8%
	LOSOCV: 8.9%	LOSOCV: 90.9%	LOSOCV: 74.8%
Mitochondrial Disorders	LOOCV: 34.1%	LOOCV: 93.1%	LOOCV: 84.6%
	LOSOCV: 29.3%	LOSOCV: 90.6%	LOSOCV: 81.8.6%
MND	LOOCV: 100%	LOOCV: 84.4%	LOOCV: 89.2%
	LOSOCV: 96.6%	LOSOCV: 82.4%	LOSOCV: 86.7%
No disease	LOOCV: 52.2%	LOOCV: 79.4%	LOOCV: 73.1%
	LOSOCV: 37.3%	LOSOCV: 74.0%	LOSOCV: 65.4%
Healthy	LOOCV: 25.7%	LOOCV: 93.2%	LOOCV: 85.0%
	LOSOCV: 8.6%	LOSOCV: 91.6%	LOSOCV: 81.5%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 23.2%	LOOCV: 92.2%	LOOCV: 78.7%
	LOSOCV: 12.5%	LOSOCV: 92.6%	LOSOCV: 76.9%
Mitochondrial Disorders	LOOCV: 43.9%	LOOCV: 83.7%	LOOCV: 78.0%
	LOSOCV: 41.5%	LOSOCV: 79.6%	LOSOCV: 74.1%
MND	LOOCV: 97.7%	LOOCV: 88.4%	LOOCV: 91.3%
	LOSOCV: 95.4%	LOSOCV: 85.9%	LOSOCV: 88.8%
No disease	LOOCV: 34.3%	LOOCV: 85.8%	LOOCV: 73.8%
	LOSOCV: 25.4%	LOSOCV: 83.6%	LOSOCV: 69.9%
Healthy	LOOCV: 48.6%	LOOCV: 92.8%	LOOCV: 87.4%
	LOSOCV: 45.7%	LOSOCV: 94.0%	LOSOCV: 88.1%

e) Myopathy vs. Mitochondrial vs. No need for biopsy vs. Healthy

The predictive capability of the four group models generated using the probe and microscope spectra can be seen in tables 3.77 and 3.78 respectively. In these models the 'MND' and 'not myopathy' groups were placed together in the 'no need for biopsy' group. Thus, these models were generated in order to explore the ability of the technique to distinguish between myopathies, mitochondrial disorders, muscle with pathological findings (more severe ones in case of MND and more subtle changes in the 'not myopathy' group) for which a biopsy would not be necessary and healthy muscle. Using LOSO decreased the ability of all the models to correctly classify spectra. For both PCA-LDA and PLS-DA LOSO cross-validated probe models only the sensitivity achieved for the mitochondrial disorders was above 60%, with the sensitivities demonstrated for all the other groups being below 50%. Specificity values were much higher than sensitivity values.

Table 3.77: Four group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 25.5%	LOOCV: 89.3%	LOOCV: 75.6%
	LOSOCV: 14.8%	LOSOCV: 81.1%	LOSOCV: 66.8%
Mitochondrial Disorders	LOOCV: 75.5%	LOOCV: 84.8%	LOOCV: 82.8%
	LOSOCV: 62.9%	LOSOCV: 81.7%	LOSOCV: 77.6%
No biopsy	LOOCV: 70.5%	LOOCV: 71.0%	LOOCV: 70.8%
	LOSOCV: 49.5%	LOSOCV: 66.9%	LOSOCV: 59.6%
Healthy	LOOCV: 39.5%	LOOCV: 93.4%	LOOCV: 85.2%
	LOSOCV: 39.5%	LOSOCV: 90.1%	LOSOCV: 82.4%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 44.4%	LOOCV: 89.8%	LOOCV: 80%
	LOSOCV: 33.3%	LOSOCV: 83.7%	LOSOCV: 72.8%
Mitochondrial Disorders	LOOCV: 79.7%	LOOCV: 82.7%	LOOCV: 82.0%
	LOSOCV: 69.8%	LOSOCV: 78.7%	LOSOCV: 76.8%
No biopsy	LOOCV: 62.5%	LOOCV: 82.1%	LOOCV: 74.0%
	LOSOCV: 48.6%	LOSOCV: 80.0%	LOSOCV: 66.8%
Healthy	LOOCV: 60.5%	LOOCV: 92.9%	LOOCV:88.0 %
	LOSOCV: 47.4%	LOSOCV: 89.2%	LOSOCV: 82.8%

The models generated using the microscope spectra were also characterised by very unbalanced sensitivity and specificity values for most of the groups and a moderate decrease in the accuracy values when LOSO cross-validation was employed. In the LOSO cross-validated PLS-DA model the 'no need for biopsy' and 'healthy' groups achieved the highest and most balanced sensitivities and specificities, whereas the sensitivities for the other two groups were particularly low for the other two groups independently of the analysis and CV methods.

Table 3.78: Four group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 26.8%	LOOCV: 94.5%	LOOCV: 80.4%
	LOSOCV: 17.9%	LOSOCV: 92.2%	LOSOCV: 77.6%
Mitochondrial Disorders	LOOCV: 17.1%	LOOCV: 95.9%	LOOCV: 84.6%
	LOSOCV: 9.8%	LOSOCV: 94.7%	LOSOCV: 82.5%
No biopsy	LOOCV: 86.4%	LOOCV: 37.1%	LOOCV: 63.6%
	LOSOCV: 84.4%	LOSOCV: 31.1%	LOSOCV: 59.8%
Healthy	LOOCV: 17.1%	LOOCV: 93.2%	LOOCV: 83.9%
	LOSOCV: 5.7%	LOSOCV: 92.8%	LOSOCV: 82.2%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 19.6%	LOOCV: 93.0%	LOOCV: 78.7%
	LOSOCV: 8.9%	LOSOCV:93.0 %	LOSOCV: 76.6%
Mitochondrial Disorders	LOOCV: 26.8%	LOOCV: 93.1%	LOOCV: 83.6%
	LOSOCV: 24.4%	LOSOCV: 91.4%	LOSOCV: 81.8%
No need for biopsy	LOOCV: 70.1%	LOOCV: 68.2%	LOOCV: 69.2%
	LOSOCV: 68.8%	LOSOCV: 65.9%	LOSOCV: 67.5%
Healthy	LOOCV: 74.3%	LOOCV: 78.2%	LOOCV: 77.2%
	LOSOCV: 71.4%	LOSOCV: 76.9%	LOSOCV: 76.2%

f) Need for biopsy vs. No need for biopsy vs. Healthy

The predictive capability of the three group models generated using the probe and microscope spectra can be seen in tables 3.79 and 3.80 respectively. In these models the 'myopathy' and 'mitochondrial disorders' were grouped together as for these patients a biopsy would facilitate the diagnosis of their neuromuscular conditions. The 'no need for biopsy' group consisted of the 'MND' and 'not myopathy' spectra as in the previous section. Using LOSO decreased the ability of the probe models to correctly classify spectra in all different groups. Both PCA-LDA and PLS-DA LOSO cross-validated probe models did not yield very good sensitivity values, with no sensitivity value being above 50%. However, this group demonstrated the lowest accuracy values due to

lower specificity. For the 'no need for biopsy' and 'healthy' groups the specificity values achieved with both models were much higher than the sensitivity ones, leading, thus, to higher accuracy values.

Table 3.79: Three group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Need for biopsy	LOOCV: 72.9%	LOOCV: 55.9%	LOOCV: 63.2%
	LOSOCV: 59.8%	LOSOCV: 42.0%	LOSOCV: 49.6%
No need for biopsy	LOOCV: 57.1%	LOOCV: 82.8%	LOOCV: 72%
	LOSOCV: 41.9%	LOSOCV: 74.5%	LOSOCV: 60.8%
Healthy	LOOCV: 31.6%	LOOCV: 94.3%	LOOCV: 84.8%
	LOSOCV: 21.1%	LOSOCV: 93.3%	LOSOCV: 82.4%
PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Biopsy	LOOCV: 64.5%	LOOCV: 69.3%	LOOCV: 67.6%
	LOSOCV: 57.0%	LOSOCV: 54.5%	LOSOCV: 55.6%
No biopsy	LOOCV: 68.6%	LOOCV: 78.6%	LOOCV: 74.4%
	LOSOCV: 47.6%	LOSOCV: 75.2%	LOSOCV: 63.6%
Healthy	LOOCV: 55.3%	LOOCV: 93.4%	LOOCV: 87.6%
	LOSOCV: 34.2%	LOSOCV: 88.0%	LOSOCV: 80.0%

Similar to the probe models, using LOSOCV also decreased most of the accuracy values of the microscope models. Both LOSO cross-validated microscope models also demonstrated very unbalanced sensitivity and specificity values.

Table 3.80: Three group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Need for biopsy	LOOCV: 43.3%	LOOCV: 82.0%	LOOCV: 68.9%
	LOSOCV: 40.2%	LOSOCV: 81.5%	LOSOCV: 67.5%
No need for biopsy	LOOCV: 83.8%	LOOCV: 50.0%	LOOCV: 68.1%
	LOSOCV: 82.5%	LOSOCV: 47.0%	LOSOCV: 66.1%
Healthy	LOOCV: 2.9%	LOOCV: 94.4%	LOOCV: 83.2%
	LOSOCV: 2.9%	LOSOCV: 94.4%	LOSOCV: 83.2%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Need for biopsy	LOOCV: 55.7%	LOOCV: 80.4%	LOOCV: 72.1%
	LOSOCV: 47.4%	LOSOCV: 75.1%	LOSOCV: 65.7%
No need for biopsy	LOOCV: 67.5%	LOOCV: 65.2%	LOOCV: 66.4%
	LOSOCV: 62.9%	LOSOCV: 64.4%	LOSOCV: 63.1%
Healthy	LOOCV: 54.3%	LOOCV: 89.6%	LOOCV: 85.3%
	LOSOCV: 42.9%	LOSOCV: 86.5%	LOSOCV: 81.1%

3.2.2.2 Basis of classification

3.2.2.2.1 Mean and Difference spectra

The mean and difference spectra of the different groups were plotted to visually examine the most prominent peaks and to identify the major differences between the groups of each of the two group models. The mean probe spectrum of each group (+/-) one standard deviation can be seen in figure 3.117 and the mean microscope spectrum of each group (+/-) one standard deviation is shown in figure 3.118.

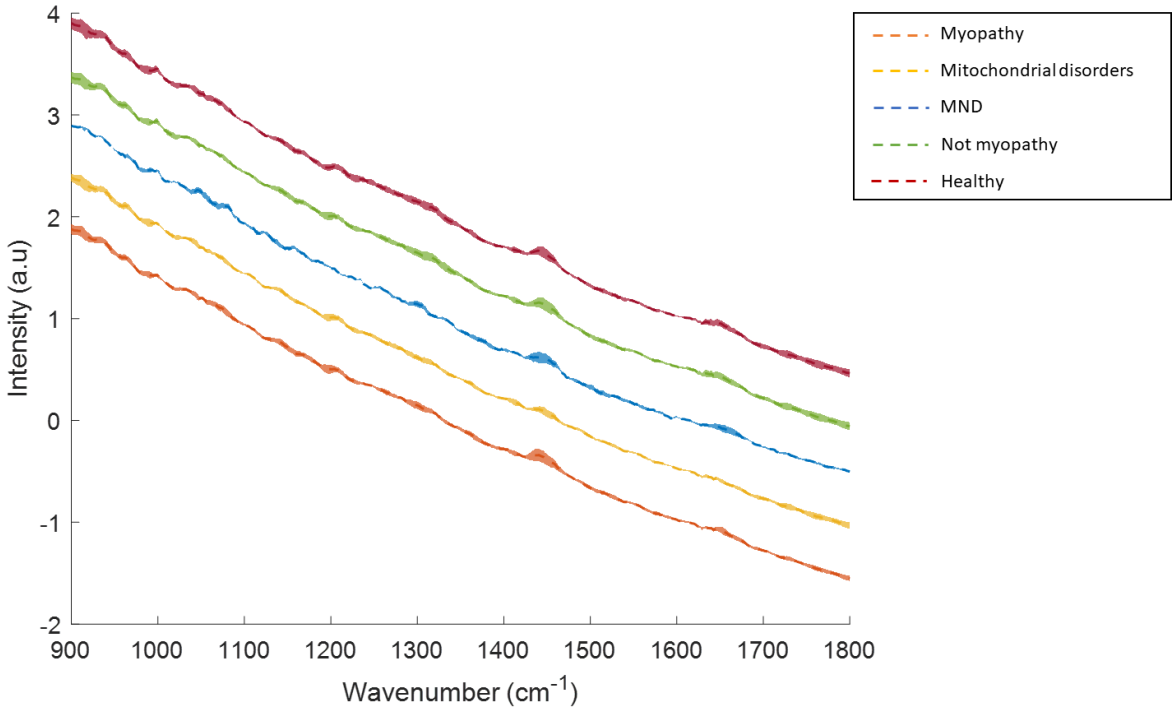


Figure 3.117: Mean probe Raman spectra for the different muscle biopsy groups.
The mean spectra for the different muscle biopsy groups are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

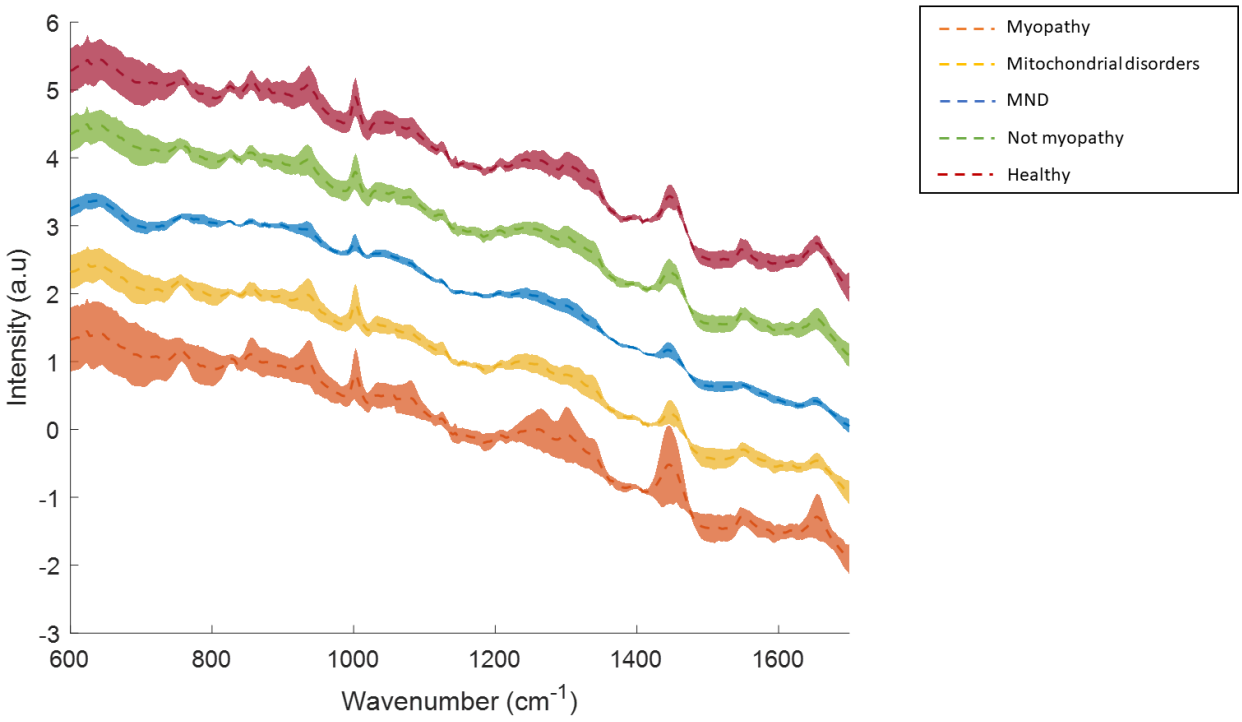


Figure 3.118: Mean microscope Raman spectra for the different muscle biopsy groups.
The mean spectra for the different groups are shown with the dotted lines. The shaded areas represent (+/-) one standard deviation from the mean spectrum. The spectra have been offset for clarity.

Figures 3.119 to 3.121 show the mean and difference probe spectra of the combinations of groups used to build the two group models. The mean spectra of the different groups consisted of similar peaks. A small number of biochemically reliable peaks could be seen in the mean and difference probe spectra (table 3.81).

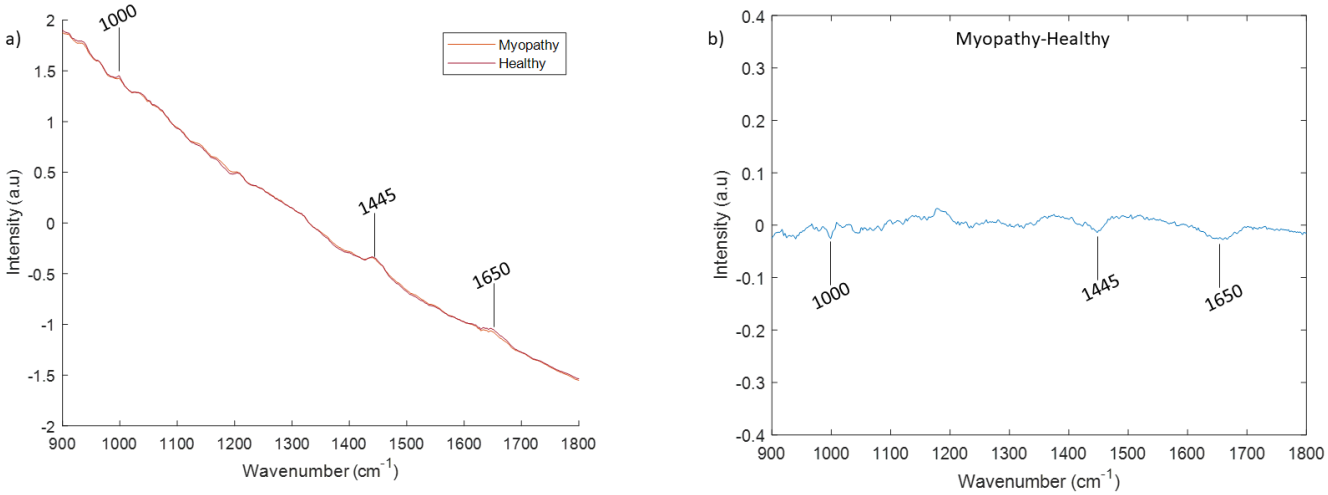


Figure 3.119: Mean and difference probe spectra of the 'myopathy' and 'healthy' samples.
a) Mean spectra for the 'myopathy' and 'healthy' samples and b) difference spectrum. Prominent peaks are indicated in both graphs.

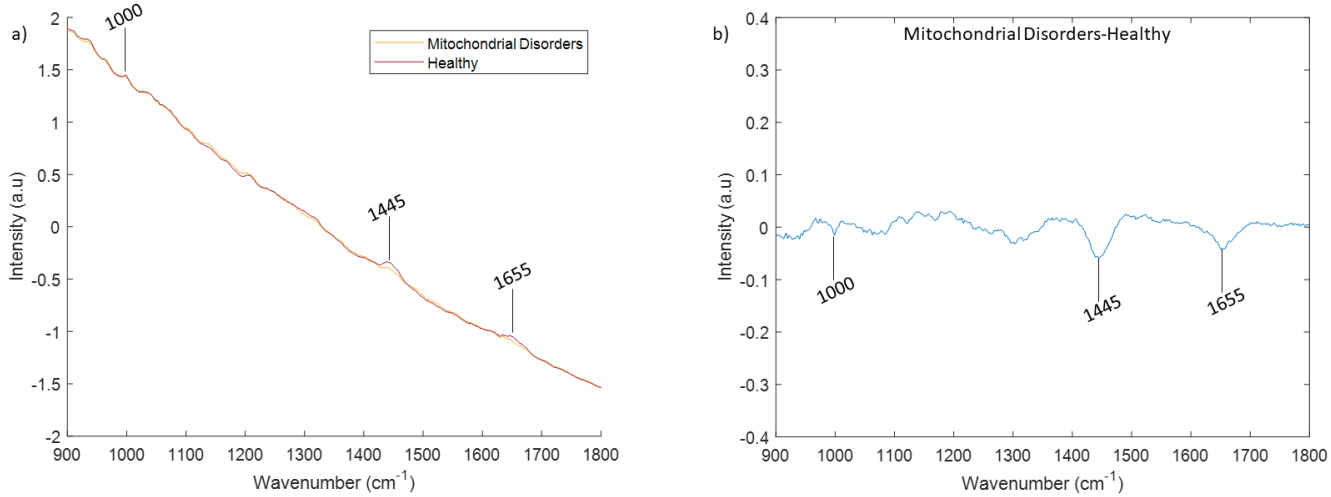


Figure 3.120: Mean and difference probe spectra of the 'mitochondrial disorders' and 'healthy' samples.
a) Mean spectra for the 'mitochondrial disorders' and 'healthy' samples and b) difference spectrum. Prominent peaks are indicated in both graphs.

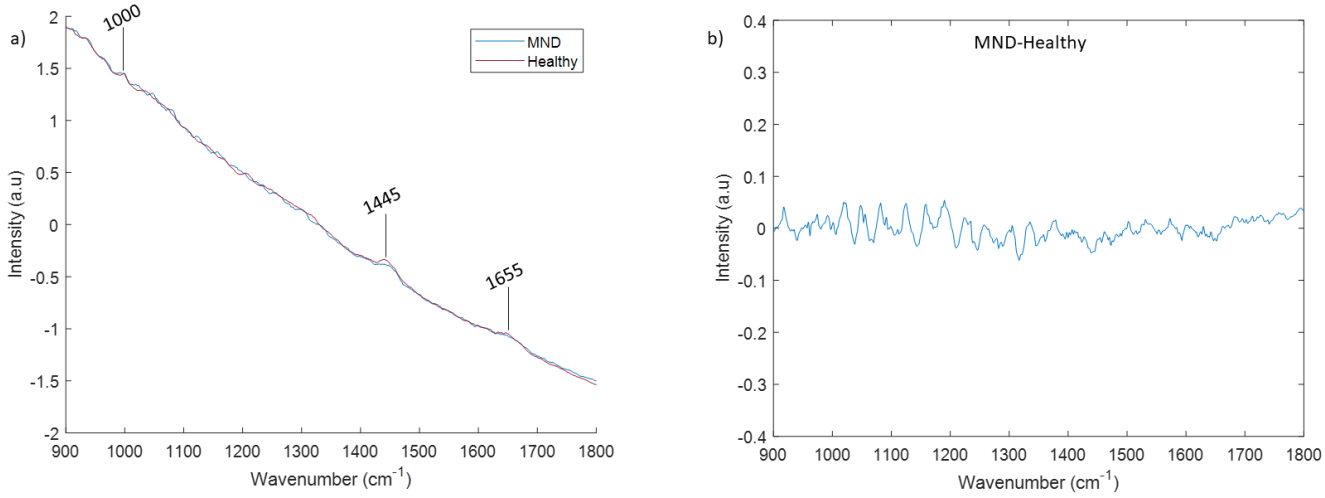


Figure 3.121: Mean and difference probe spectra of the 'MND' and 'healthy' samples.
a) Mean spectra for the MND and healthy samples and b) difference spectrum. Prominent peaks are indicated on the mean spectra.

Table 3.81: Prominent Raman peaks in mean and difference spectra of the different muscle biopsy groups and tentative peak assignments.

Wavenumber (cm^{-1})	Vibrational Mode *	Assignment
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1445	CH_2 bending, CH_2CH_3 bending	Proteins, Lipids
1655	C=O stretching, C=C stretching	Amide I (proteins), Lipids

Figures 3.122 to 3.124 show the mean and difference microscope spectra of the combinations of groups used to build the two group models. Similar to the probe spectra, the mean spectra of the different groups consisted of similar peaks. A larger number of biochemically reliable peaks could be identified in the spectra acquired with the microscope (table 3.82).

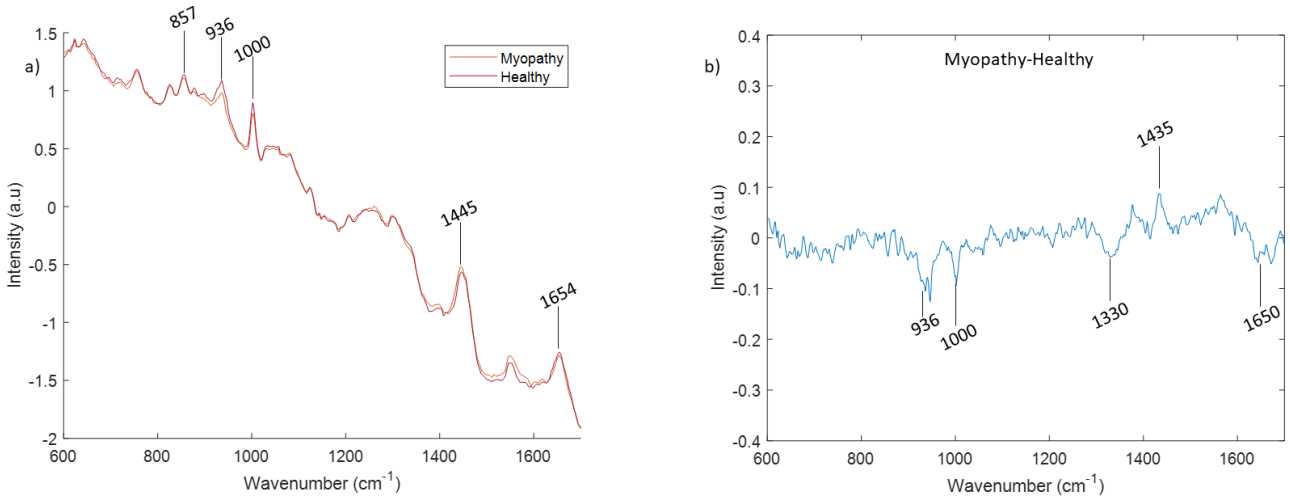


Figure 3.122: Mean and difference microscope spectra of the 'myopathy' and 'healthy' samples.
a) Mean spectra for the myopathic and healthy samples and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

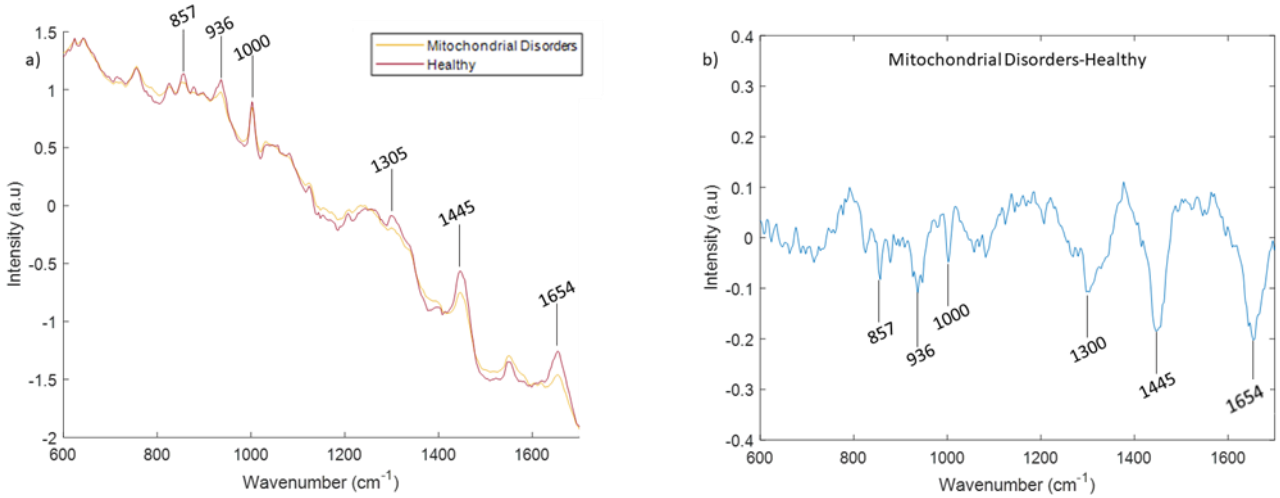


Figure 3.123: Mean and difference spectra microscope of the 'mitochondrial disorders' and 'healthy' samples.
a) Mean spectra for the 'mitochondrial disorders' and 'healthy' samples and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.

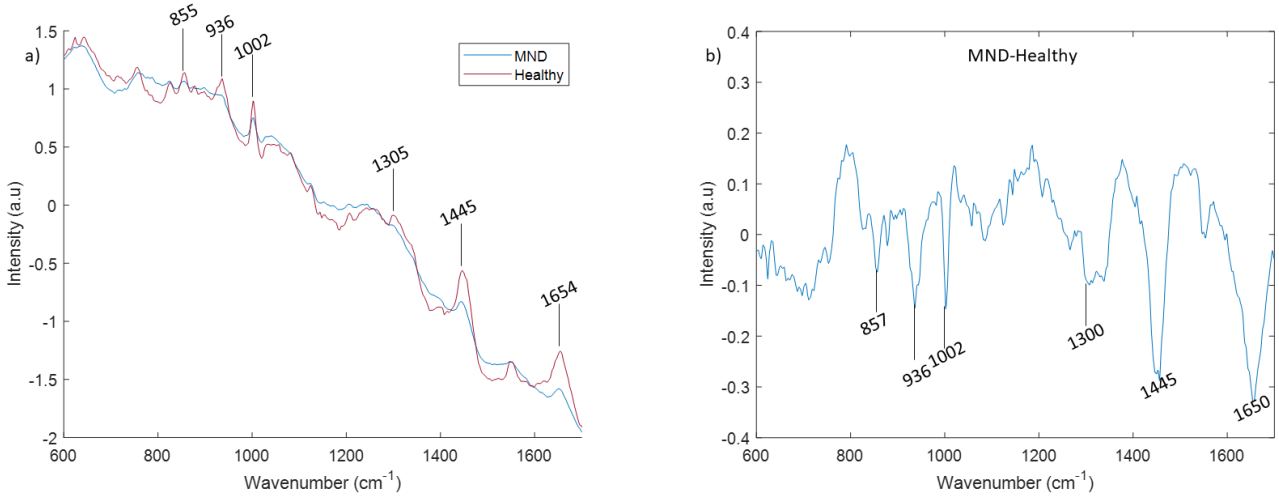


Figure 3.124: Mean and difference microscope spectra of the 'MND' and 'healthy' samples. *a) Mean spectra for the 'MND' and 'healthy' samples and b) difference spectrum. Prominent peaks that differ between the two groups are indicated in both graphs.*

Table 3.82: Prominent Raman peaks in mean and difference spectra of the different muscle biopsy groups and tentative peak assignments. *These peaks were found to be decreased in the 'mitochondrial disorders' and 'MND' spectra when these were compared with spectra acquired from healthy muscle.*

Wavenumber (cm^{-1})	Vibrational Mode *	Assignment
827	Ring breathing mode	Proteins (Tyrosine)
855	C-C stretching	Proteins (Tyrosine, Proline Glycogen)
935	C-C stretching	Proteins
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1300	CH_2 twisting	Amide III (proteins), Lipids
1335	CH_3CH_2 wagging	Proteins, Nucleic Acids
1445	CH_2 bending, CH_2CH_3 bending	Proteins, Lipids
1654	C=O stretching, C=C stretch	Amide I (proteins), Lipids

Background subtraction was used to aid identification of more spectral features present in the spectra. The background subtracted mean spectra of the different groups can be seen in figures 3.125 and 3.126 for the probe and microscope spectra respectively. A larger number of peaks

were identifiable in these mean spectra. The prominent peaks were very similar for the different groups. The peaks present in the probe and microscope spectra and their tentative peak assignments can be seen in tables 3.83 and 3.84 respectively.

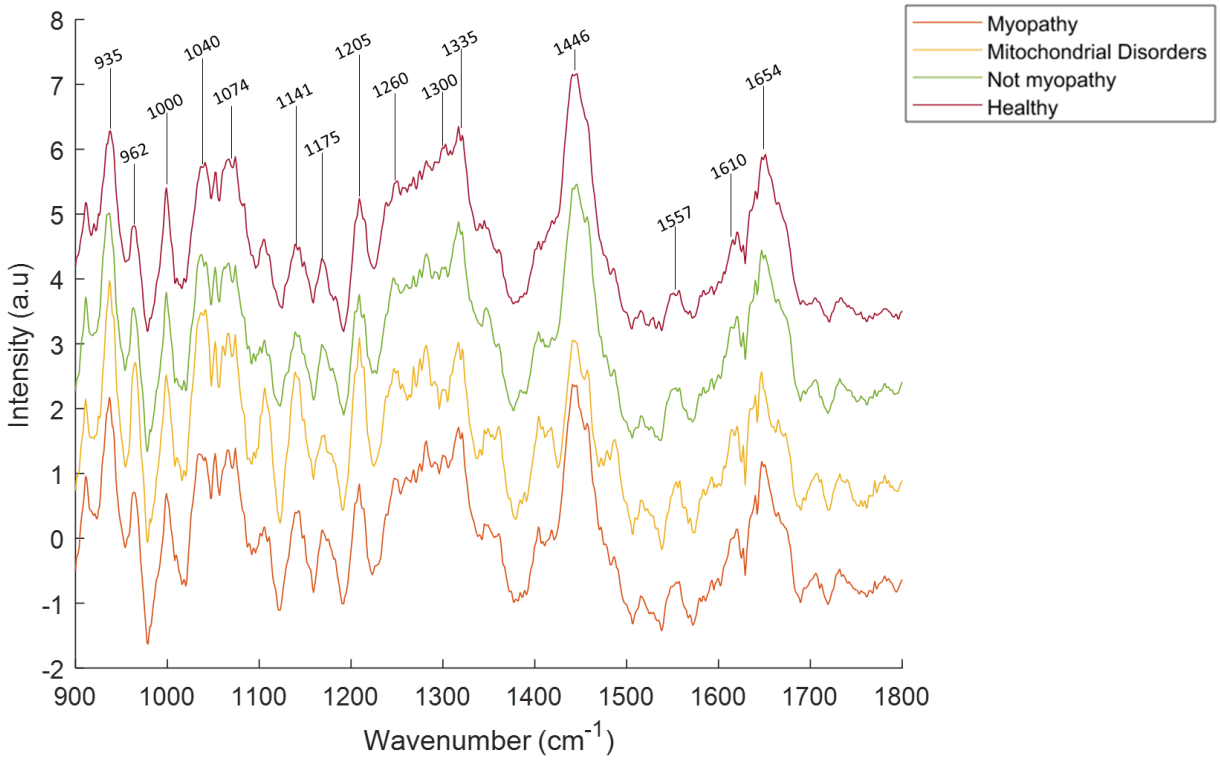


Figure 3.125: Background subtracted mean probe spectra of the different muscle biopsy groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

Table 3.83: Prominent Raman peaks in the background subtracted spectra of the different muscle biopsy groups and tentative peak assignments.

Wavenumber (cm^{-1})	Vibrational Mode *	Assignment
935	C-C stretching	Proteins
1000	Phenyl ring breathing mode	Proteins (Phenylalanine)
1040		Proteins
1075	C-C stretching, C-O stretching	Proteins (Tryptophan), Lipids
1141		Proteins (Valine)
1175	C-H bending	Proteins (Tyrosine, Phenylalanine)
1205	$\nu(C\text{-}C_6H_5)$	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)
1260	C-N stretching	Amide III (proteins), Lipids
1300	CH_2 twisting	Amide III (proteins), Lipids
1330	C-N stretching C-C Stretching	Proteins
1410		Proteins, Nucleic acids
1444	CH_2 bending	Proteins, Lipids
1557	$\nu(C\text{-}C)$	Proteins (Tryptophan)
1615	C=C stretching	Proteins (Tyrosine)
1650	C=O stretching, C=C stretching	Amide I (proteins), Lipids

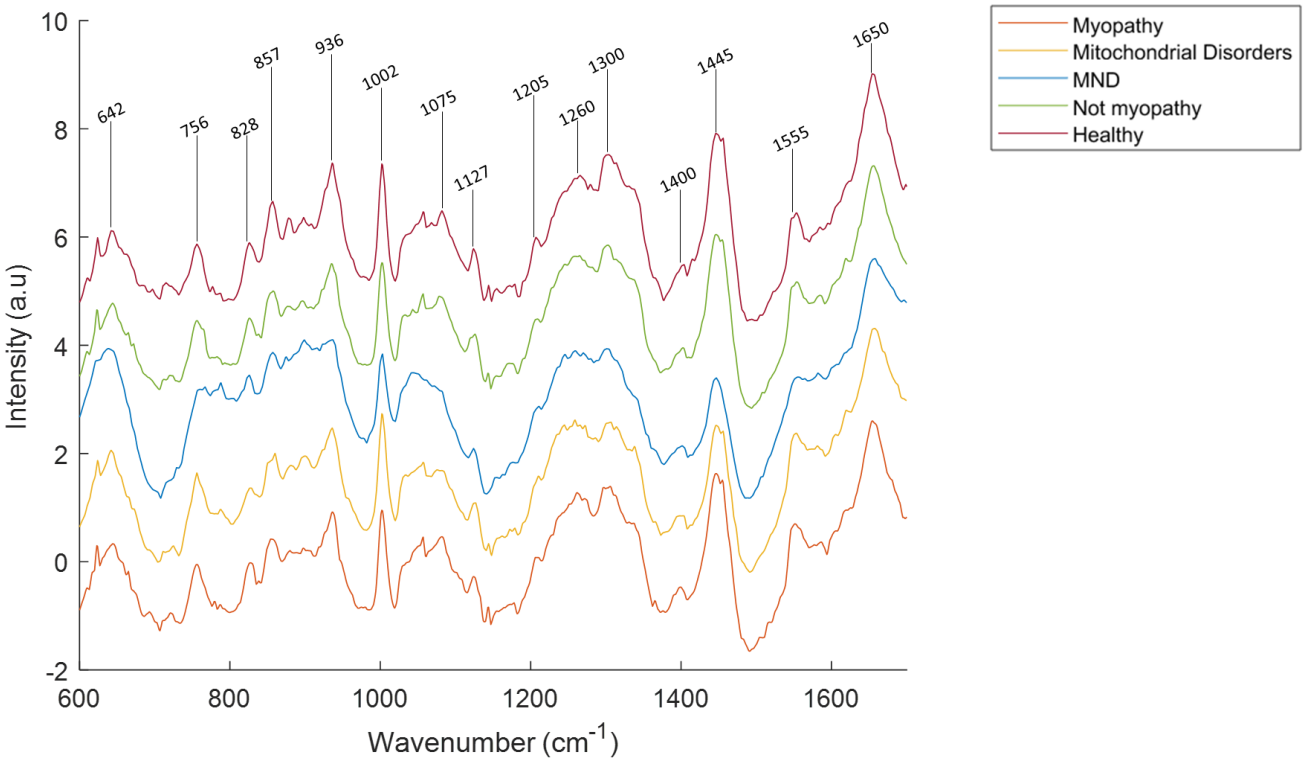


Figure 3.126: Background subtracted mean microscope spectra of the different muscle biopsy groups. *The spectra have been offset for clarity and the most prominent peaks have been highlighted.*

Table 3.84: Prominent Raman peaks in the background subtracted mean microscope spectra of the different muscle biopsy groups and tentative peak assignments.

Wavenumber (cm^{-1})	Vibrational Mode *	Assignment
642	C-C twisting	Nucleic acids (Bases), Proteins (Tyrosine)
756	Symmetric breathing	Proteins (Tryptophan)
825	Ring breathing	Proteins (Tyrosine)
855	C-C stretching	Proteins (Tyrosine, Proline, Glycogen)
935	C-C stretching	Proteins
1002	Phenyl ring breathing mode	Proteins (Phenylalanine)
1044		Proteins
1076	C-C stretching, C-O stretching	Proteins (Tryptophan), Lipids
1125	C-C stretching, C-N stretching	Proteins, Lipid
1205	$\nu(C\text{-}C_6H_5)$	Proteins (Phenylalanine, Tyrosine, Hydroxyproline)
1260	C-N stretching	Amide III (proteins), Lipids
1300	CH_2 twisting	Amide III (proteins), Lipids
1330	C-N stretching C-C Stretching	Proteins
1400	Symmetric bending	Proteins
1444	CH_2 bending	Proteins, Lipids
1555	$\nu(C\text{-}C)$	Proteins (Tryptohan)
1654	C=O stretching, C=C stretching	Amide I (proteins), Lipids

The most prominent peaks found in the *ex vivo* spectra from human tissue were similar with the ones found in the spectra acquired in the *in vivo* experiments and those of major muscle components. The peaks between 1260 and 1340 were less clearly identifiable in probe spectra compared to both the microscope spectra and most of the spectra acquired *in vivo* from mice.

3.2.2.2.2 Multivariate statistics

Multivariate techniques were employed in order to utilise and further elucidate the biochemical features present in the spectra. The PCA loadings and the linear discriminant function (LDF) as

well as the PLS weights were plotted in order to illustrate the important peaks for spectral classification.

Assigning specific peaks to one of the two groups used in each model was not trivial since the two groups were not always clearly separated around zero in the corresponding score histograms or score plots. Additionally, the peaks in the loadings or weight plots were most often not centred around zero probably due to the presence of background in the spectra. Hence, a combination of the information found in the difference spectra and the loading/weight plots was used in an attempt to better understand the differences present in the spectra from the different groups. The score histograms and loading/weight plots in which the separation of the spectra from the different groups was best for each model are shown in the following sections. The score histograms and loading/weight plots for the rest of the PCs and components used to build each model are shown in Appendix C. Tables 3 and 4 summarise and compare the most prominent peaks found in all the loading/weight plots (from all the components used for each model) and the ones found in the difference spectra of each two-group model.

a) Myopathy vs. Healthy

Using Student's t-tests followed by fdr correction PC1 and PC4 scores were found to be significantly different between the myopathy samples and the samples from healthy volunteers (q_{PC1}:0.0014, q_{PC4}: 0.0009). The spectra of the two groups were not clearly separated around zero in any of the score histograms. Thus, only the PCA loading plots are shown in this section (figure 3.127).

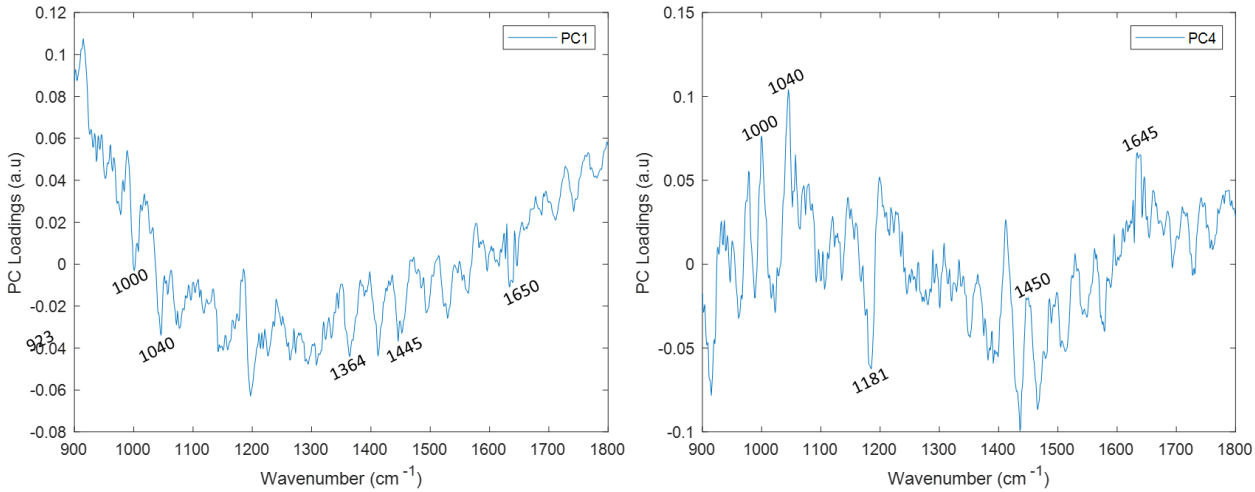


Figure 3.127: PC1 and PC4 and loading plots for the 'myopathy' and 'healthy' samples.

The LDA histogram and LDF plot using PC1 and PC4 as input variables can be seen in figure 3.128.

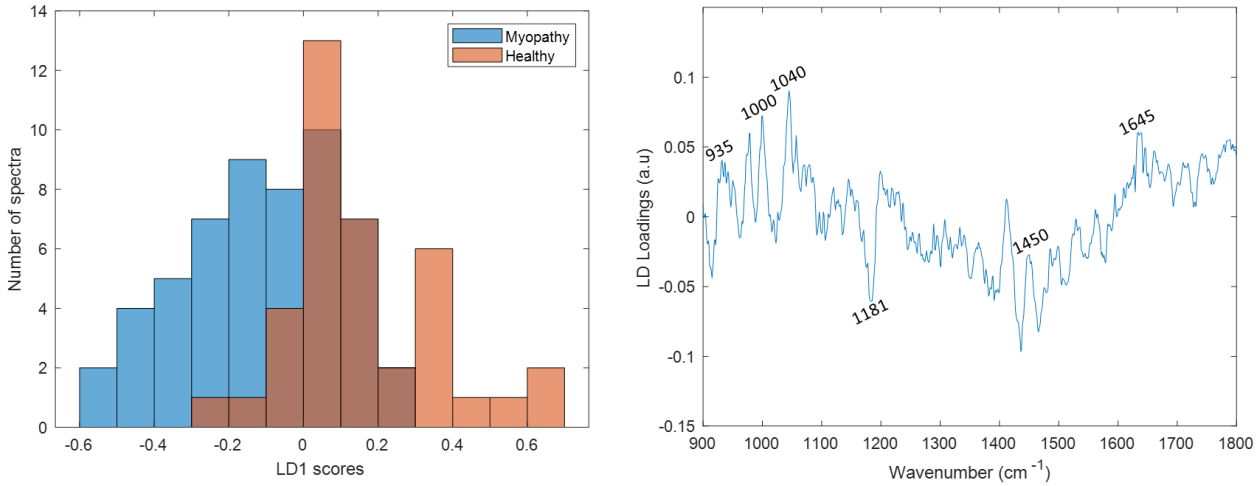


Figure 3.128: LD1 score histogram and LDF plot for the 'myopathy' and 'healthy' samples.

In the PLS-DA analysis the model was generated using the first 5 components. Similar to the PCA analysis, the spectra were not clearly separated around zero in any of the score histograms (Appendix C). The weight plots of the first four components can be seen in figure 3.129. The weight plot of the fifth component can be seen in Appendix C. Despite some background present in the weight plot of component 1, biochemically relevant peaks were apparent.

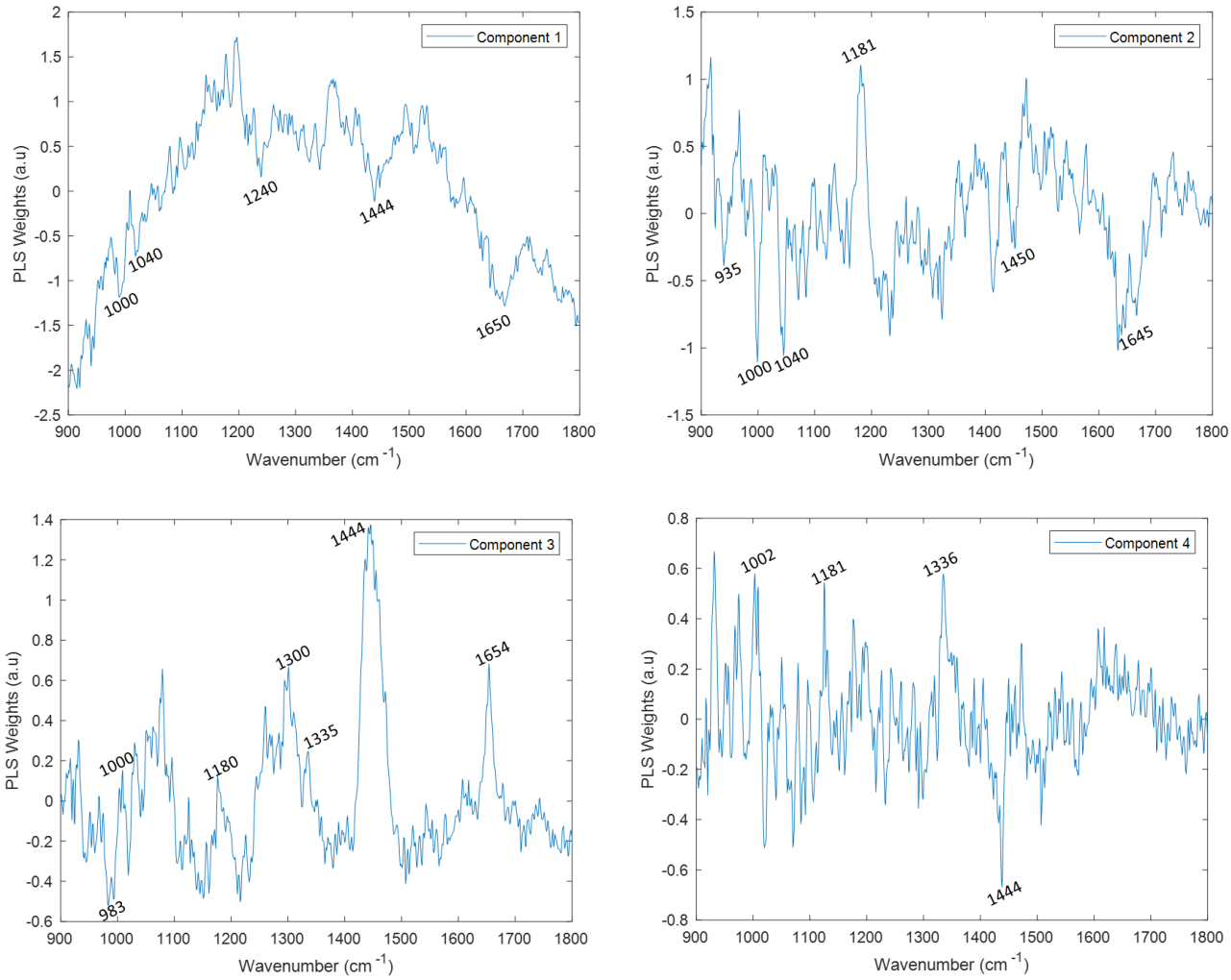


Figure 3.129: PLS component 1,2,3 and 4 weights plots for the 'myopathy' and 'healthy' samples.

Using the spectra acquired with the microscope there were no significantly different PC scores between the two groups. The PLS-DA model was generated using the first 6 components. The PLS weight plots of the first 4 components can be seen in figure 4. The weight plots of components 5 and 6 can be found in Appendix C.

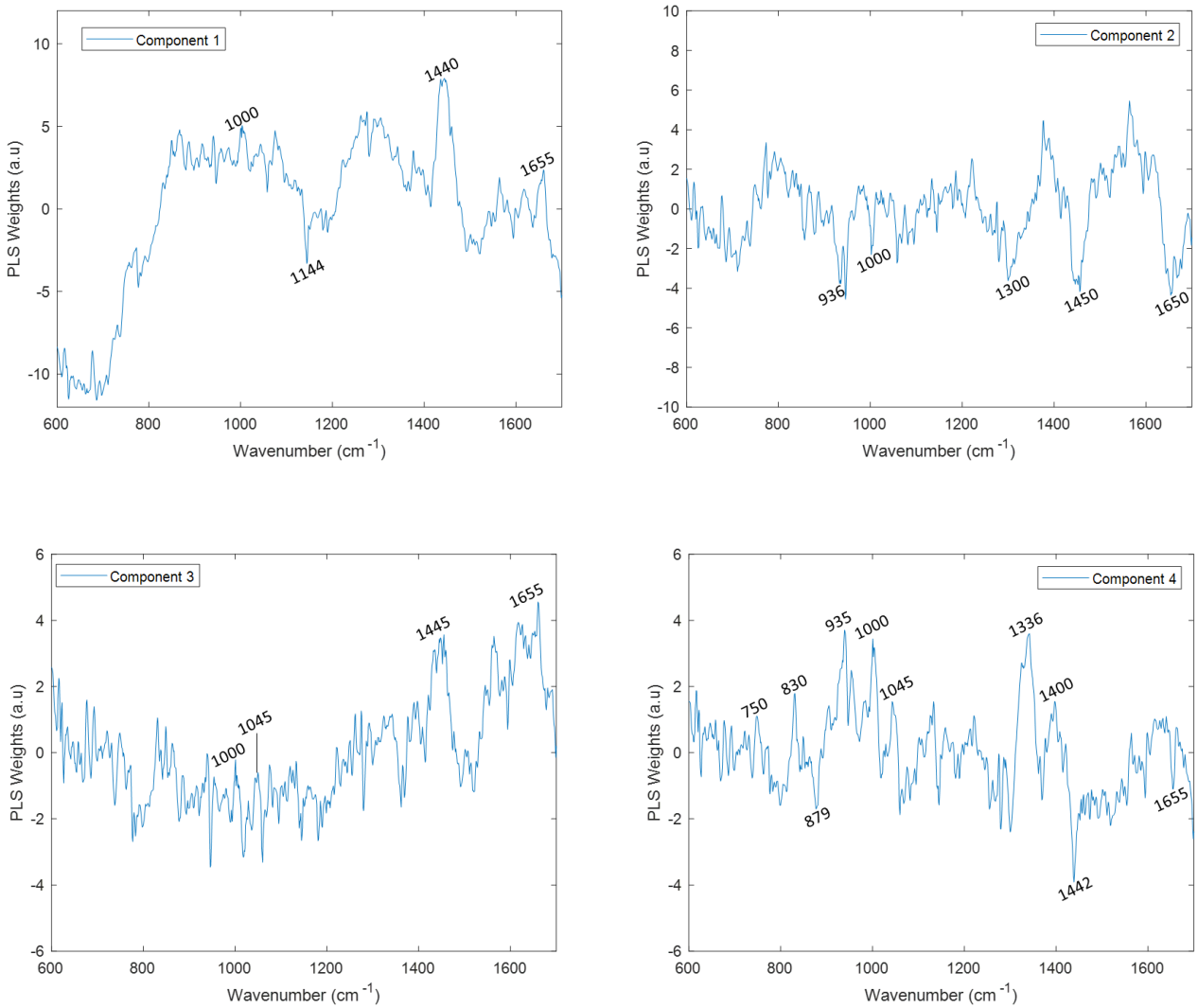


Figure 3.130: PLS component 1,2,3 and 4 weights plots for the 'myopathy' and 'healthy' samples.

The most prominent peaks found in the probe and microscope spectra and in the loading/weight plots of the different analyses are summarised and compared in table 3.85. Similar to the results of the previous sections (*in vivo* recordings) the peaks located around 935 cm^{-1}, 1000 cm^{-1}, 1040 cm^{-1}, 1300 cm^{-1}, 1330 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1} were also found in these spectra and loading plots, indicating, thus, that similar peaks were important for classification. These peaks were prominent in loading/weight plots of both the probe' and microscope spectra.

Table 3.85: Summary table of the peaks associated with the 'myopathy' and 'healthy' muscle samples and tentative peak assignments. *The common peaks in the probe and microscope spectra are indicated with bold font.*

		Probe			Microscope		
Wavenumber (cm^{-1})	**Tentative Assignment**	**PCA-LDA**	**PLS-DA**	**Difference Spectrum**	**PCA-LDA**	**PLS-DA**	**Difference Spectrum**
750	Proteins (Tryptophan)					✓	
830	Proteins (Tyrosine, Tryptophan)					✓	
879	Proteins (Hydroxiproline, Tryptophan)					✓	
935	Proteins	✓				✓	✓
985	Proteins, Nucleic Acids		✓				
1000	Proteins (Phenylalanine)	✓	✓	✓		✓	✓
1040	Proteins (Phenylalanine, Proline)		✓			✓	
1181	Proteins (Tyrosine)		✓	✓			
1300	Amide III (proteins), Lipids	✓	✓			✓	
1335	Proteins, Nucleic Acids		✓			✓	✓
1400	Proteins					✓	
1445	Proteins, Lipids	✓	✓	✓		✓	✓
1655	Amide I (proteins), Lipids	✓	✓	✓		✓	✓

b) Mitochondrial Disorders vs Healthy

Using the probe spectra PC2 scores were found to be significantly different between mitochondrial disorders and healthy samples ($q_{PC2}=2.2e^{-10}$). The LD1 score histogram and LDF plot (identical with PC2 score and loading plots) are shown in figure 3.131.

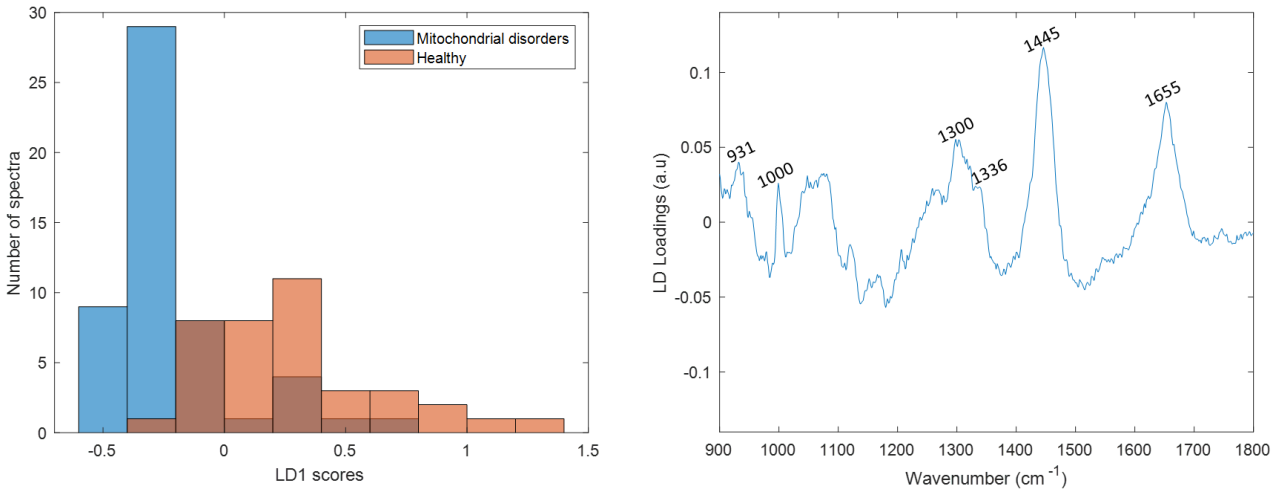


Figure 3.131: LD1 score histogram and LDF plot for the 'mitochondrial disorders' and 'healthy' samples.

In the PLS-DA analysis the model generated using only the first component demonstrated the highest accuracy value. The score and weight plot of the first component can be seen in figure 3.132.

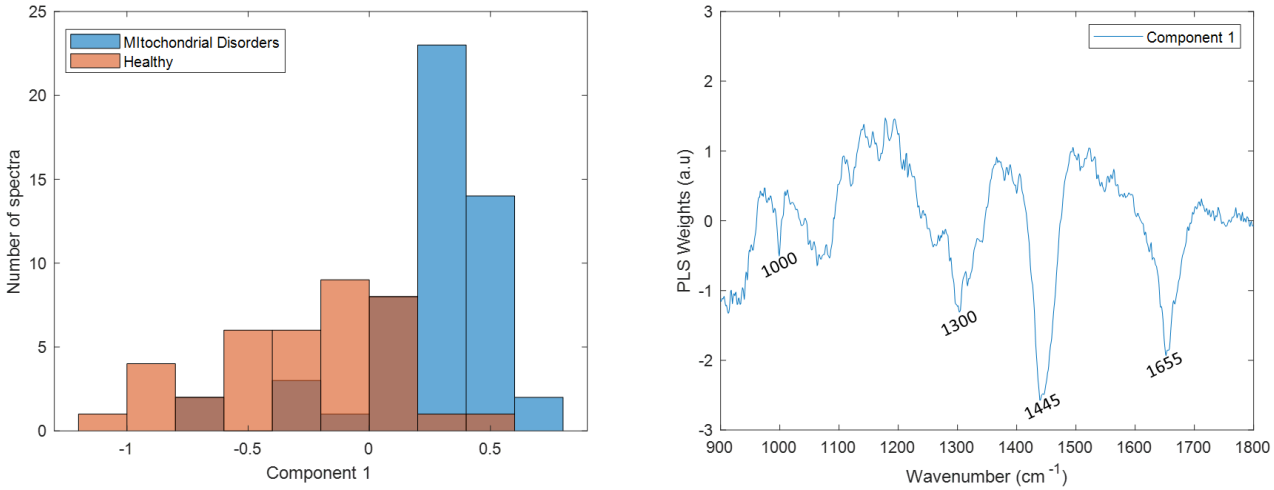


Figure 3.132: PLS component 1 score histogram and weights plot for the 'mitochondrial disorders' and 'healthy' samples.

Using the spectra acquired with the microscope, PC2 was also the only significant PC between mitochondrial disorders samples and muscle samples from healthy volunteers ($q^{PC2}=5.6e^{-09}$) (figure 3.133). In the PLS-DA analysis the model generated using the three first components demonstrated the highest accuracy value. The score and weight plot of the first component can be seen in figure 3.134.

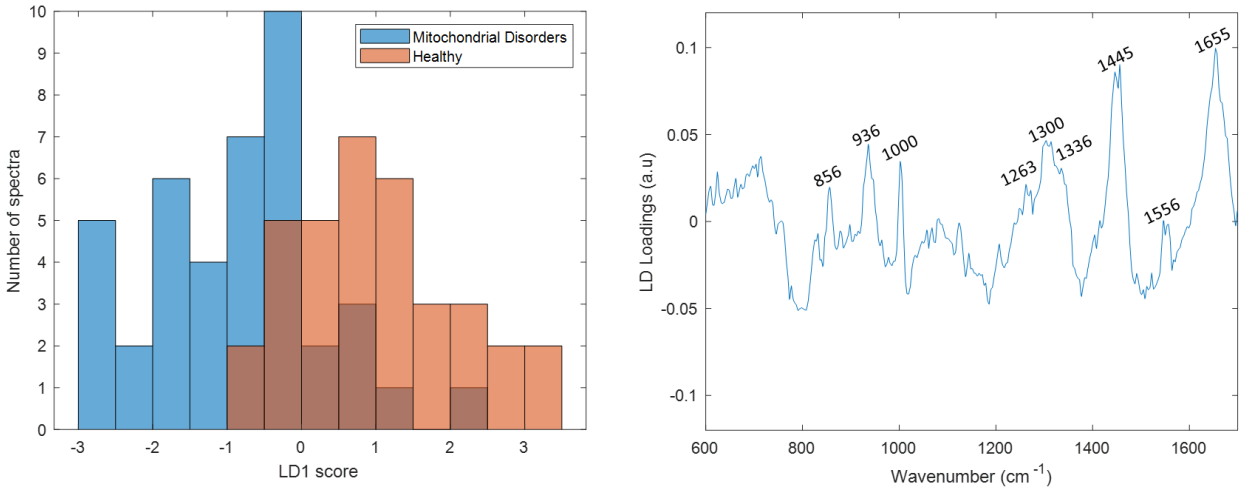


Figure 3.133: LD1 score histogram and LDF plot for the 'mitochondrial disorders' and 'healthy' samples.

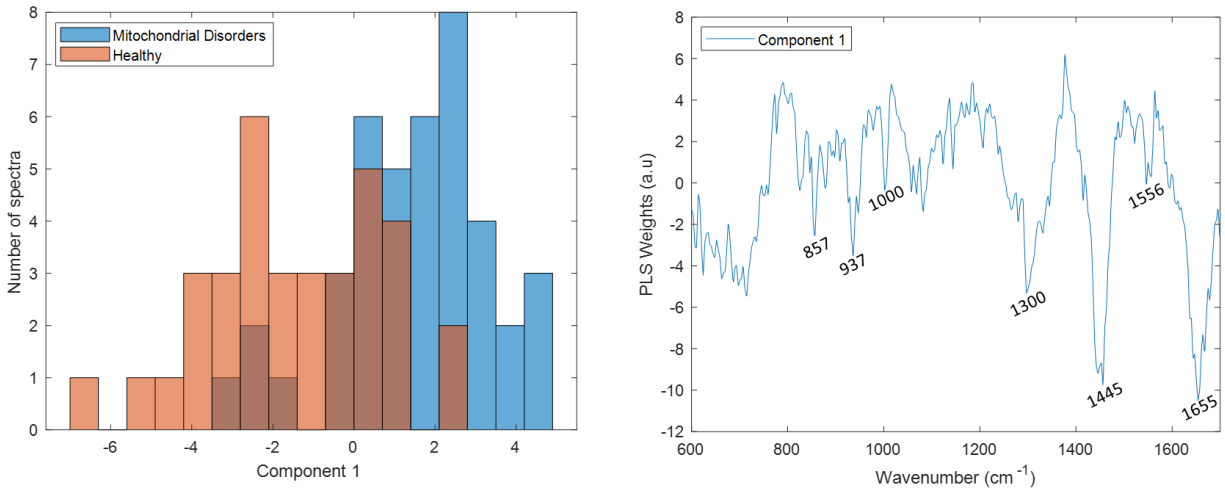


Figure 3.134: PLS component 1 score histogram and weights plot for the 'mitochondrial disorders' and 'healthy' samples.

Even though there was some overlap it can be seen from the histogram of the LD scores that positive LD1 values had a larger contribution from spectra acquired from the 'healthy' group for both the probe and microscope models. This could indicate that the peaks located around 933 cm^{-1}, 1000 cm^{-1}, 1044 cm^{-1}, 1300 cm^{-1}, 1336 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1} were decreased in the spectra acquired from the 'mitochondrial disorders' samples. Similar peaks in the opposite direction were found in the PLS weight plots. These peaks were very similar with the bands discussed in the previous sections demonstrating that changes in muscle's biochemical

composition were important for discrimination. The most prominent peaks found in the loading plots and weight plots of the probe and microscope models as well as in the difference spectra along with their tentative peak assignments are summarised in table 3.86.

Table 3.86: Summary table of the peaks associated with the 'mitochondrial disorders' and 'healthy' muscle samples and tentative peak assignments. *The common peaks in the probe and microscope spectra are indicated with bold font. In the group column, the group that the peaks were more prominent is displayed.*

		Probe			Microscope			
Wavenumber (cm^{-1})	**Tentative Assignment**	**PCA-LDA**	**PLS-DA**	**Difference Spectrum**	**PCA-LDA**	**PLS-DA**	**Difference Spectrum**	**Group**
857	Proteins (Tyrosine, Proline, Glycogen)				✓	✓	✓	Healthy
935	Proteins	✓			✓	✓	✓	Healthy
1000	Proteins (Phenylalanine)	✓	✓	✓	✓	✓	✓	Healthy
1045	Proteins (Phenylalanine, Proline)					✓		
1128	Proteins					✓		
1265	Amide III (proteins), Lipids		✓		✓			
1300	Amide III (proteins), Lipids	✓	✓		✓	✓	✓	Healthy
1335	Proteins, Nucleic Acids	✓			✓			
1445	Proteins, Lipids	✓	✓	✓	✓	✓	✓	Healthy
1555	Proteins (Tryptophan)				✓	✓		
1655	Amide I (proteins), Lipids	✓	✓	✓	✓	✓	✓	Healthy

c) MND vs Healthy

Using the spectra acquired with the Raman probe PC3 scores were significantly different between the two groups (q_{PC3}=5.1e^{-19}). Thus, the LDA score histogram and LDF plot were identical with the PCA score and loading plots and can be seen in figure 3.135. In the PLS-DA analysis 2 components were used to generate the model. The score histograms and weight plots for the first 2 PLS components are displayed in figure 3.136.

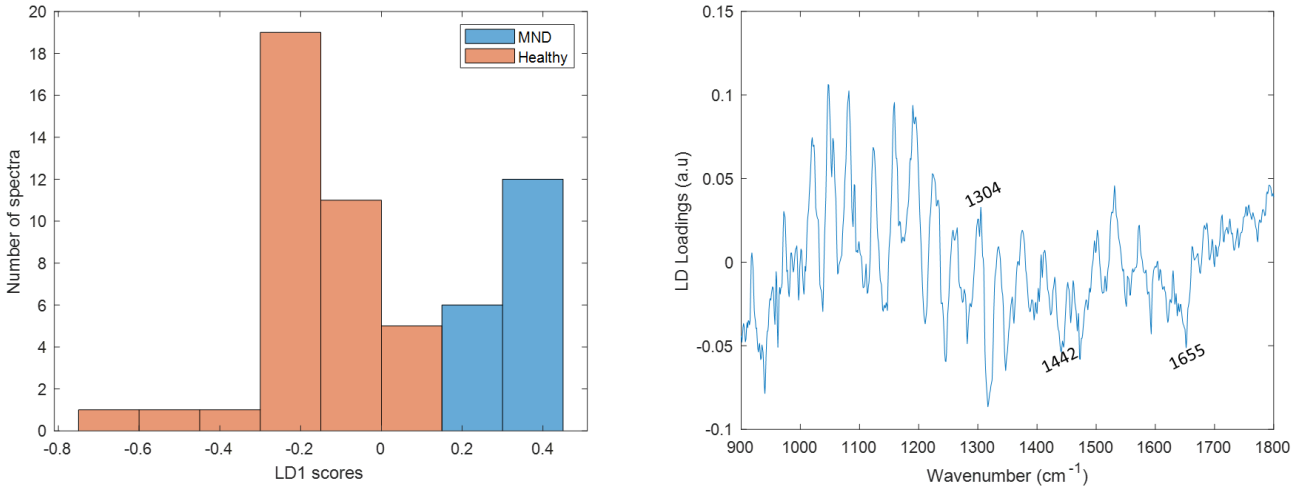


Figure 3.135: LD1 score histogram and LDF plot for the 'MND' and 'healthy' samples.

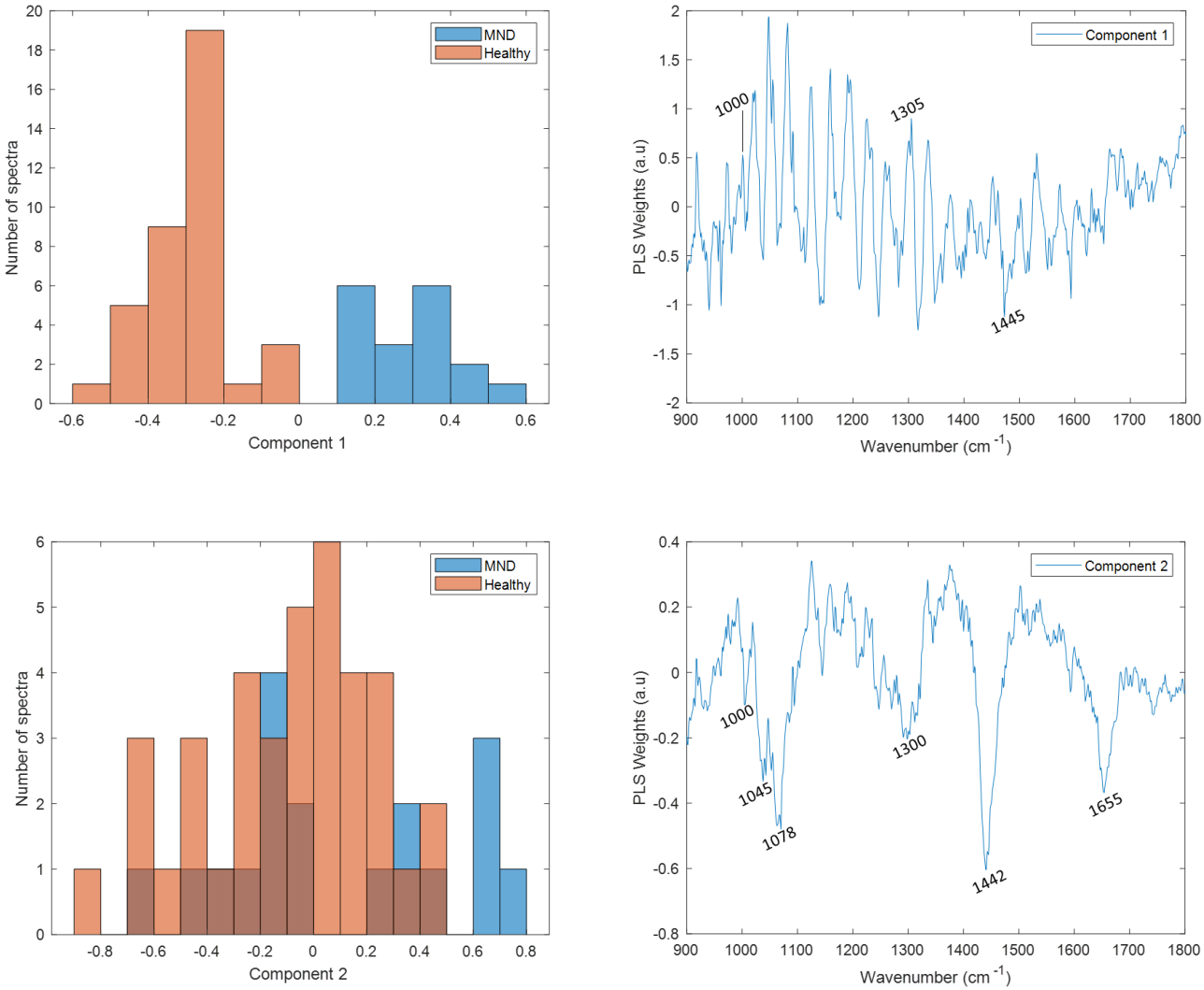


Figure 3.136: PLS component 1 and 2 score histograms and weights plots for the 'MND' and 'healthy' samples.

PCs 1, 2 and 5 were found to be significantly different between 'MND' and 'healthy' muscle using the microscope spectra. PC2 (figure 3.137) demonstrated the largest difference between the different groups followed by PC1. These PCs had a much smaller q-value than PC5 (q_{PC2}: 2.3e-14, q_{PC1}: 5.4e-08, q_{PC5}: 0.0021). The score histograms and loading plots of PC1 and PC5 can be seen in Appendix C. The LD score histogram and LDF generated using PCs 1,2 and 5 can be seen in figure 3.138.

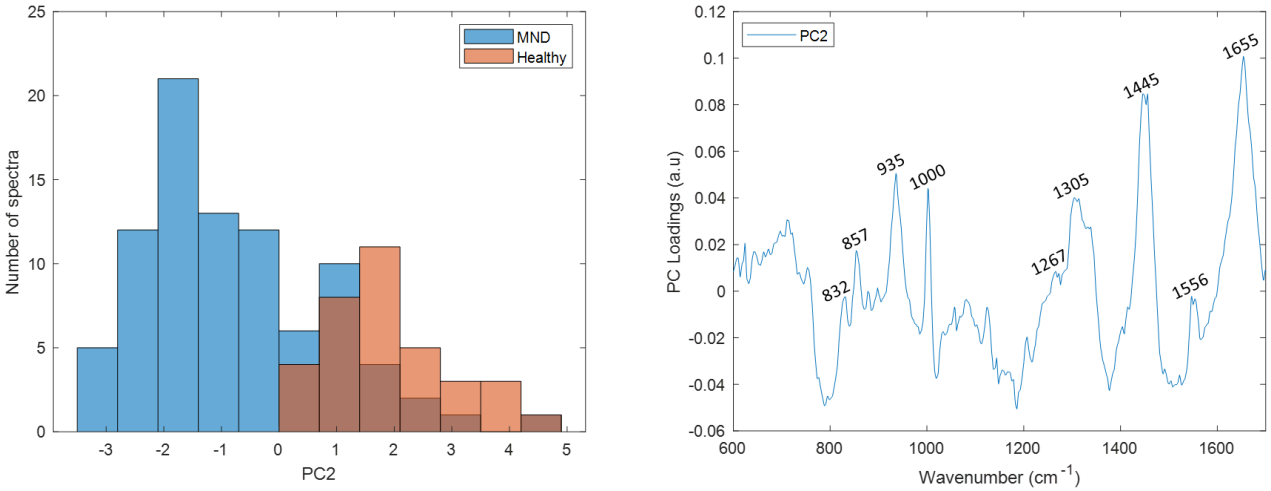


Figure 3.137: PC2 score histogram and loadings plot for the 'MND' and 'healthy' samples.

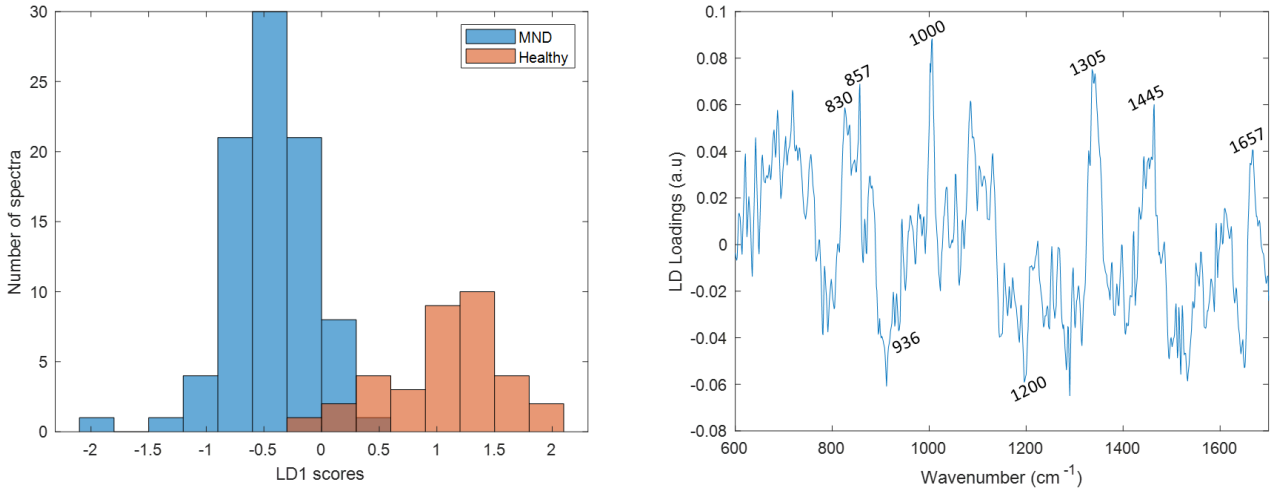


Figure 3.138: LD1 score histogram and LDF plot for the 'MND' and 'healthy' samples.

In the PLS-DA analysis the model generated using the first 6 components demonstrated the highest accuracy. The score histogram and PLS weight plot for the component 1 are displayed in figure 3.139. The weight plots for components 2 to 6 can be found in Appendix C.

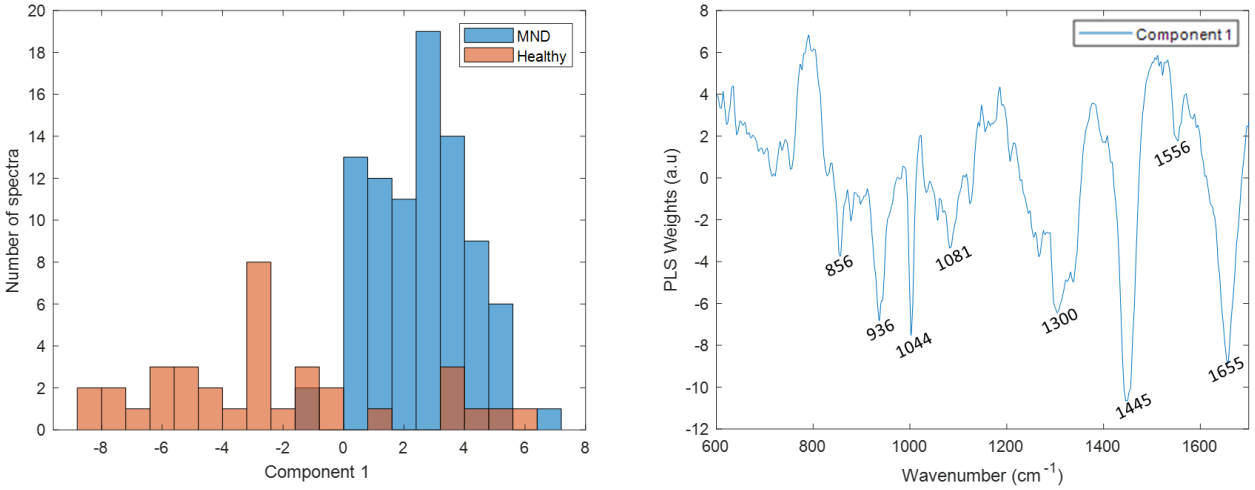


Figure 3.139: PLS component 1 score histogram and weights plot for the 'MND' and 'healthy' samples.

As can be seen in the histogram of PC2 and the LD scores positive values had a larger contribution from spectra acquired from the 'healthy' group for the microscope model. Thus, the peaks located around 855 cm^{-1}, 1000 cm^{-1}, 1300 cm^{-1}, 1445 cm^{-1} and 1655 cm^{-1} were considered to have been decreased in the spectra acquired from the 'MND' samples. Similar peaks in the opposite direction were found in the PLS weight plot for component one. However, due to background present in the weight plot, the peaks could not be associated easily to one of the two groups. The peaks found in the loading plots and weight plots of the probe and microscope models as well as in the difference spectra along with their tentative peak assignments are summarised in table 3.87.

Table 3.87: Summary table of the peaks associated with the 'MND' and 'healthy' muscle samples and tentative peak assignments. *The common peaks in the probe and microscope spectra are indicated with bold font. In the group column, the group that the peaks were more prominent is displayed.*

		Probe			Microscope			
Wavenumber (cm^{-1})	**Tentative Assignment**	**PCA-LDA**	**PLS-DA**	**Difference Spectrum**	**PCA-LDA**	**PLS-DA**	**Difference Spectrum**	**Group**
830	Proteins (Tyrosine)				✓			
857	Proteins (Tyrosine, Proline, Glycogen)				✓	✓	✓	Healthy
935	Proteins				✓	✓	✓	
1000	Proteins (Phenylalanine)		✓		✓	✓	✓	Healthy
1045	Proteins (Phenylalanine, Proline)		✓			✓		
	Proteins, Lipids, Phospholipids		✓			✓		
1200	Proteins				✓			
1265	Amide III (proteins), Lipids				✓			
1300	Amide III (proteins), Lipids	✓	✓		✓	✓	✓	Healthy
1335	Proteins, Nucleic Acids					✓		
1445	Proteins, Lipids	✓	✓		✓	✓	✓	Healthy
1555	Proteins, Tryptophan				✓	✓		
1655	Amide I (proteins), Lipids	✓	✓		✓	✓	✓	Healthy

3.2.3 Final diagnosis

In the samples acquired from the open muscle biopsies in Sheffield, three samples were from patients with a final diagnosis of myopathy but who had a negative biopsy. To explore how this may have altered the analyses in the following sections these samples are labelled as 'myopathy', according to the final diagnosis in this chapter. In 3.2.3.1 the main investigations undertaken for these samples are presented. In section 3.2.3.2 the results of the analysis with groups based on final diagnosis are presented.

3.2.3.1 Main Investigations Report

From the 38 patients that had open muscle biopsies in Sheffield (table 3.66 section 3.1.2.1) 22 people also underwent an EMG examination. Concordance of the biopsy and EMG findings with the final diagnosis was examined for the 'myopathy' and 'not myopathy' groups and is presented in the table below. In this table only the patients that had both a biopsy and an EMG were included, and the final diagnosis is the working diagnosis at the time of data analysis.

Table 3.88: Concordance between EMG and biopsy findings in the 'myopathy' and 'not myopathy' groups.

Group	Concordant: EMG-in agreement with final diagnosis, Biopsy-in agreement with final diagnosis	Discordant: EMG-in agreement with final diagnosis, Biopsy-not in agreement with final diagnosis	Discordant: EMG-not in agreement with final diagnosis, Biopsy-in agreement with final diagnosis	Concordant: EMG-not in agreement with final diagnosis, Biopsy-not in agreement with final diagnosis
Final diagnosis: Myopathy	7	1	4	2
Final diagnosis: Not myopathy	7	0	1	0

In order to evaluate the performance of muscle biopsy and EMG the sensitivity, specificity and accuracy of the techniques for detecting myopathies were calculated using the final diagnosis. The confusion matrices of the techniques for the 'myopathy' and 'not myopathy' groups can be seen in tables 3.89 and 3.90.

Table 3.89: EMG confusion matrix for 'myopathy' and 'not myopathy' groups.

		EMG	
		Myopathy	**Not myopathy**
Final Diagnosis	**Myopathy**	8	6
	Not myopathy	1	7

Table 3.90: Biopsy confusion matrix for 'myopathy' and 'not myopathy' groups.

		Biopsy	
		Myopathy	**Not myopathy**
Final Diagnosis	**Myopathy**	15	3
	Not myopathy	0	20

The sensitivity, specificity and accuracy of EMG and muscle biopsy are shown in table 3.91. In this table Raman data are also shown, the sensitivity, specificity and accuracy of the PLS-DA cross-validated (LOSO) model generated using the spectra acquired with the fibre-optic probe from the 'myopathy' and 'not myopathy' samples are also shown in order to compare the new technique with the already existing diagnostic tests.

Table 3.91: Sensitivity, specificity and accuracy values for detecting myopathies using different methods.

	Muscle Biopsy	EMG	Raman Spectroscopy (PLS-DA)
Sensitivity (%)	83.3	58.0	48.4
Specificity (%)	100	87.5	41.6
Accuracy (%)	92.1	68.1	44.7

It is important to note that the diagnostic indices for the EMG are calculated using less samples since only 22 patients underwent the EMG examination. Muscle biopsy and Raman spectroscopy accuracy values are calculated using data acquired from all 38 patients (see section 3.2.1.1 table 3.66 for details).

The muscle biopsy demonstrated the best sensitivity, specificity and accuracy values. EMG achieved considerably better specificity than sensitivity, demonstrating thus a better ability to correctly classify non-myopathic cases. Finally, Raman spectroscopy was not able to correctly classify the samples using the Raman spectra.

3.2.3.2 Classification models performance when using the final clinical diagnosis

a) Myopathy vs. Healthy

The predictive capability of the models generated using the probe and microscope spectra can be seen in tables 3.92 and 3.93 respectively. The ROC curves comparing the performance of the two different analysis and CV methods can be seen in figure 3.140 for the probe models and in figure 3.141 for the microscope ones.

As can be seen in table 3.92, when using LOSOCV the ability of both the PCA-LDA probe models to correctly classify spectra decreased slightly, whereas the accuracy value increased for the PLS-DA model, due to an increase in the sensitivity. For both CV approaches the PLS-DA model had lower sensitivity but higher specificity values that led to higher accuracy values. The sensitivity and specificity values were also more balanced for the PLS-DA models.

Table 3.92: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'myopathy' and 'healthy' samples. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	81.3%	44.7%	67.6%
Leave-one-sample-out CV	78.1%	42.1%	64.7%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	71.9%	97.4%	81.4%
Leave-one-sample-out CV	76.6%	92.1%	82.4%

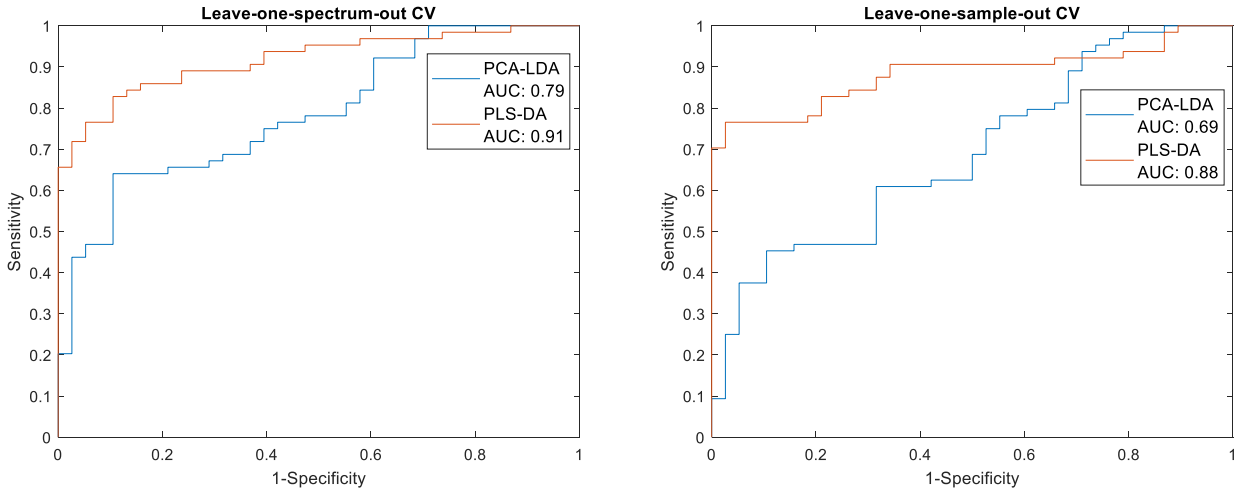


Figure 3.140: ROC curves for the cross validated classification models for the 'myopathy' and 'healthy' samples. *ROC curves for all models generated using the probe spectra using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

For the spectra acquired using the microscope Raman system only the PLS-DA models were generated since there were no PCs with significantly different scores between the two groups. As can be seen in table 3.93 using LOSOCV the PLS-DA model attained decreased accuracy due a drop in the sensitivity. The AUC was above 0.70 for both CV approaches.

Table 3.93: Two group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods for the 'myopathy' and 'healthy' samples. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	No significant PCs	No significant PCs	No significant PCs
Leave-one-sample-out CV	No significant PCs	No significant PCs	No significant PCs

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Leave-one-spectrum-out CV	68.8%	74.3%	70.7%
Leave-one-sample-out CV	59.4%	74.3%	64.6%

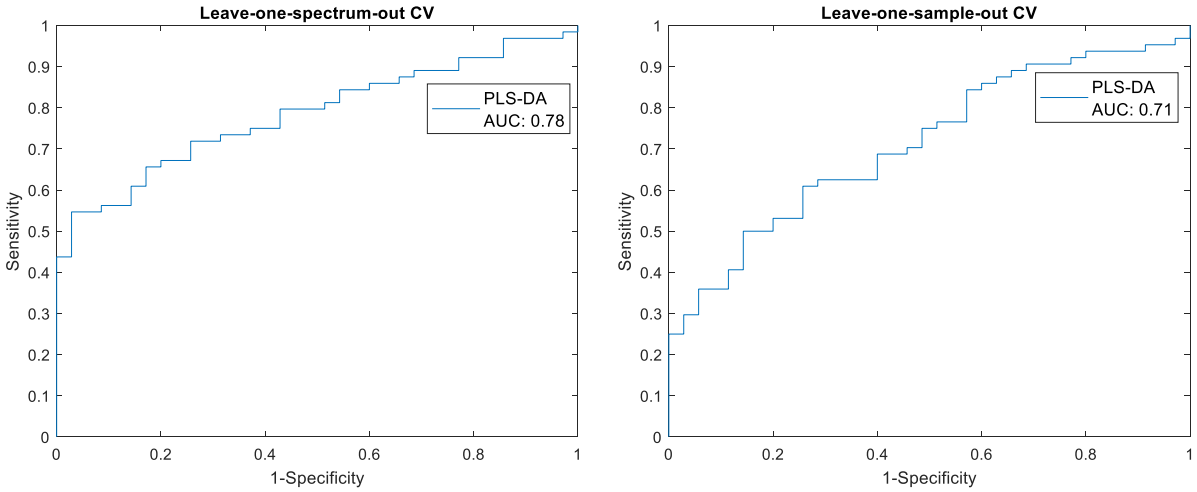


Figure 3.141: ROC curves for the cross validated classification models for the 'myopathy' and 'healthy' samples. *ROC curves for all models generated using the microscope spectra using leave-one-spectrum-out and leave-one-sample-out CV are shown. The area under the ROC curve (AUC) for the different models is also displayed.*

In general, the models generated using the probe spectra performed better than the ones generated using the microscope recordings. The LOMO cross-validated PLS-DA probe model demonstrated a good performance for separating between 'myopathy' from 'healthy' samples,

achieving an accuracy of 82.4% whereas the microscope model demonstrated an accuracy of 64.6%.

b) Myopathy vs. Mitochondrial Disorders vs. No need for biopsy vs. Healthy

The predictive capability of the four group models generated using the probe and microscope spectra can be seen in tables 3.94 and 3.95 respectively. In these models the 'MND' and 'not myopathy' groups were placed together in the 'no need for biopsy' group. Thus, these models were generated in order to explore the ability of the technique to distinguish between myopathies, mitochondrial disorders, muscle with pathological findings (more severe ones in case of MND and more subtle changes in the 'not myopathy' group) for which a biopsy would not be necessary and healthy muscle. Using LOSO decreased the ability of all the models to correctly classify spectra. For both PCA-LDA and PLS-DA LOSO cross-validated probe models only the sensitivity achieved for the mitochondrial disorders was above 60%, with the sensitivities demonstrated for all the other groups being below 50%. Specificity values were much higher than sensitivity values.

Table 3.94: Four group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 32.8%	LOOCV: 80.6%	LOOCV: 68.4%
	LOSOCV: 18.7%	LOSOCV: 77.4%	LOSOCV: 62.4%
Mitochondrial Disorders	LOOCV: 66.0%	LOOCV: 85.3%	LOOCV: 81.2%
	LOSOCV: 64.1%	LOSOCV: 84.8%	LOSOCV: 80.4%
No need for biopsy	LOOCV: 58.9%	LOOCV: 73.5%	LOOCV: 68%
	LOSOCV: 49.5%	LOSOCV: 68.4%	LOSOCV: 61.2%
Healthy	LOOCV: 39.5%	LOOCV: 92.0%	LOOCV: 84.0%
	LOSOCV: 39.5%	LOSOCV: 92.0%	LOSOCV: 84.0%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 31.1%	LOOCV: 84.9%	LOOCV: 71.2%
	LOSOCV: 26.6%	LOSOCV: 72.6%	LOSOCV: 60.8%
Mitochondrial Disorders	LOOCV: 83.0%	LOOCV: 79.7%	LOOCV: 80.4%
	LOSOCV: 69.8%	LOSOCV: 79.2%	LOSOCV: 77.2%
No need for biopsy	LOOCV: 60.0%	LOOCV: 81.9%	LOOCV: 73.6%
	LOSOCV: 32.6%	LOSOCV: 76.1%	LOSOCV: 59.6%
Healthy	LOOCV: 52.6%	LOOCV: 93.0%	LOOCV: 87.6%
	LOSOCV: 31.6%	LOSOCV: 88.7%	LOSOCV: 80.0%

The models generated using the microscope spectra were also characterised by very unbalanced sensitivity and specificity values for most of the groups and a moderate decrease in the accuracy values when LOSO cross-validation was employed. In the LOSO cross-validated PLS-DA model the 'no need for biopsy' and 'healthy' groups achieved the highest and most balanced sensitivities and specificities, whereas the sensitivities for the other two groups were particularly low for the other two groups independently of the analysis and CV methods.

Table 3.95: Four group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 21.9%	LOOCV: 87.9%	LOOCV: 73.1%
	LOSOCV: 21.9%	LOSOCV: 85.6%	LOSOCV: 71.3%
Mitochondrial Disorders	LOOCV: 21.9%	LOOCV: 95.9%	LOOCV: 85.3%
	LOSOCV: 14.6%	LOSOCV: 95.9%	LOSOCV: 84.3%
No need for biopsy	LOOCV: 86.9%	LOOCV: 46.4%	LOOCV: 67.1%
	LOSOCV: 84.2%	LOSOCV: 42.1%	LOSOCV: 63.6%
Healthy	LOOCV: 11.4%	LOOCV: 92.1%	LOOCV: 82.2%
	LOSOCV: 5.7%	LOSOCV: 92.8%	LOSOCV: 82.2%

PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Myopathy	LOOCV: 18.9%	LOOCV: 93.7%	LOOCV: 76.9%
	LOSOCV: 4.7%	LOSOCV: 91.9%	LOSOCV: 72.4%
Mitochondrial Disorders	LOOCV: 34.1%	LOOCV: 89.8%	LOOCV: 81.8%
	LOSOCV: 24.4%	LOSOCV: 86.1%	LOSOCV: 77.4%
No need for biopsy	LOOCV: 74.0%	LOOCV: 66.4%	LOOCV: 70.3%
	LOSOCV: 66.4%	LOSOCV: 65.0%	LOSOCV: 65.7%
Healthy	LOOCV: 57.1%	LOOCV: 81.7%	LOOCV: 78.7%
	LOSOCV: 62.9%	LOSOCV: 78.9%	LOSOCV: 76.9%

c) Need for biopsy vs. No need for biopsy vs. Healthy

The predictive capability of the three group models generated using the probe and microscope spectra can be seen in tables 3.96 and 3.97 respectively. In these models the 'myopathy' and 'mitochondrial disorders' were grouped together as for these patients a biopsy would facilitate the diagnosis of the neuromuscular conditions. The 'no need for biopsy' group consisted of the 'MND' and 'not myopathy' spectra as in the previous section. Using LOSO decreased the ability

of the probe models to correctly classify spectra in all different groups. Both PCA-LDA and PLS-DA LOSO cross-validated probe models did not yield very good sensitivity values, with only the sensitivity of the 'need for biopsy' group being above 50%. However, this group demonstrated the lowest accuracy values due to lower specificity. For the 'no need for biopsy' and 'healthy' groups the specificity values achieved with both models were much higher than the sensitivity ones, leading, thus, to higher accuracy values.

Table 3.96: Three group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the probe spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Biopsy	LOOCV: 74.4%	LOOCV: 45.9%	LOOCV: 59.2%
	LOSOCV: 70.1%	LOSOCV: 40.6%	LOSOCV: 54.4%
No biopsy	LOOCV: 43.2%	LOOCV: 81.3%	LOOCV: 66.8%
	LOSOCV: 37.9%	LOSOCV: 79.4%	LOSOCV: 63.6%
Healthy	LOOCV: 31.6%	LOOCV: 95.8%	LOOCV: 86.0%
	LOSOCV: 26.3%	LOSOCV: 94.8%	LOSOCV: 84.4%
PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Biopsy	LOOCV: 64.1%	LOOCV: 66.9%	LOOCV: 65.6%
	LOSOCV: 59.8%	LOSOCV: 53.4%	LOSOCV: 56.4%
No biopsy	LOOCV: 55.3%	LOOCV: 79.4%	LOOCV: 74.0%
	LOSOCV: 43.2%	LOSOCV: 78.1%	LOSOCV: 64.8%
Healthy	LOOCV: 52.6%	LOOCV: 92.0%	LOOCV: 86.0%
	LOSOCV: 42.1%	LOSOCV: 87.3%	LOSOCV: 80.4%

Table 3.97: Three group PCA-LDA and PLS-DA classification model performance evaluated using different CV methods. *The models were generated using the microscope spectra.*

PCA-LDA	***Sensitivity***	***Specificity***	***Accuracy***
Biopsy	LOOCV: 30.5%	LOOCV: 84.5%	LOOCV: 64.7%
	LOSOCV: 24.8%	LOSOCV: 80.7%	LOSOCV: 60.1%
No biopsy	LOOCV: 86.9%	LOOCV: 30.7%	LOOCV: 59.4%
	LOSOCV: 82.2%	LOSOCV: 27.1%	LOSOCV: 55.2%
Healthy	LOOCV: 2.9%	LOOCV: 99.6%	LOOCV: 87.8%
	LOSOCV: 5.7%	LOSOCV: 90.4%	LOSOCV: 88.1%
PLS-DA	***Sensitivity***	***Specificity***	***Accuracy***
Biopsy	LOOCV: 56.2%	LOOCV: 79.0%	LOOCV: 70.6%
	LOSOCV: 51.4%	LOSOCV: 73.5%	LOSOCV: 65.4%
No biopsy	LOOCV: 73.3%	LOOCV: 64.3%	LOOCV: 68.9%
	LOSOCV: 65.7%	LOSOCV: 62.9%	LOSOCV: 64.3%
Healthy	LOOCV: 34.3%	LOOCV: 92.0%	LOOCV: 85.0%
	LOSOCV: 34.3%	LOSOCV: 90.4%	LOSOCV: 83.6%

4. Discussion

4.1 Animal Experiments

In general, the results of the *in vivo* studies demonstrate that the technique was able to discriminate between healthy and diseased muscle for different neuromuscular disorders. In order to better investigate the potential of fibre-optic Raman spectroscopy to detect muscle disease *in vivo*, different analysis and cross-validation methods that have been successfully used in previous Raman spectroscopy studies, were employed. The results of the *in vivo* work are discussed in the following sections. The first hypothesis, presented in section 1.3.2, stated that in vivo Raman spectroscopy of muscle can distinguish between neuropathic and myogenic disorders. The second hypothesis was that intra-muscular *in vivo* Raman spectroscopy can detect changes over time in the two neuromuscular disorders mouse models used in this study. The classification performance of the models generated using spectra acquired from *mdx* (myopathic) and $SOD1^{G93A}$ (neurogenic) mice at two different disease stages are summarized and discussed in more detail in section 4.1.1.6. The classification performance of the models using the spectra recorded from the one and three months old mice are summarized and discussed in more detail in sections 4.1.1.1-4.1.1.2 (*mdx* mice) and 4.1.1.4-4.1.1.5 ($SOD1^{G93A}$ mice). In general, all the above-mentioned models achieved high accuracy values, demonstrating that the technique was able to distinguish between neurogenic and myopathic disorders and between different disease stages.

The third hypothesis stated that *in vivo* Raman spectroscopy of muscle will not cause significant muscle injury and impair motor performance in mice. The results addressing this hypothesis are discussed in section 4.1.2. Catwalk, rotarod and MRI data show that the technique did not cause any extended damage in muscle tissue and did not lead to consistent or extended impairment of motor function across the different groups.

4.1.1 Neurogenic and Myopathic Neuromuscular Disorders: Studies 1 and 4

4.1.1.1 Summary of the classification performance: mdx mice

All models generated using the one month old *mdx* and C57Bl/10 mice displayed accuracy values around or above 70%. The PLS-DA models performed better than the PCA-related models, yielding accuracy values above 76% with sensitivity values above 80%, specificity values above 70% and AUROC above 0.85 for all CV methods. Using spectra acquired from three months old *mdx* and C57Bl/10 generated models that yielded accuracy values above 80%. The PLS-DA models achieved the best performance with accuracy values above 90% and AUROC above 0.96 for all different CV methods and more balanced sensitivity and specificity values. The two-group models generated using spectra of one and three months old *mdx* mice in general yielded higher sensitivity than specificity values indicating that the models were able to correctly classify more spectra obtained from the three months old mice. The PLS-DA models were considered to have exhibited the best performance due to accuracy values above 88% for all CV methods and higher specificity values (above 80%) than the PCA-related models. All coefficients of variability were below 5.4%.

The models generated using the spectra obtained from male *mdx* mice also achieved better accuracy values for the three months old mice compared to one month old mice. When comparing one month old *mdx* to C57Bl/10 mice, all the models achieved accuracy values above 71%. Cross-validation methods had a big impact on the ability of the models to correctly classify

spectra and using LOSO cross-validation yielded the lowest accuracy values for all three analysis methods. Coefficients of variability varied between 3.4% and 5.4%. Unlike the one month old mice related models the performance of the models generated using the three months old mice spectra did not change considerably for different CV methods for any analysis. The PLS-DA models demonstrated the best performance with sensitivity of 100% and specificity around 97% for all different CV methods.

A drop in sensitivity and specificity values when leave- one-mouse-out (or more mice) instead of leave-one-spectrum-out was used could indicate that the model does not capture sufficiently inter-subject variability within the same pathology group. Hence, it can classify correctly the spectra acquired from one mouse less easily when only spectra from other mice are used. In general, there was a moderate drop in the accuracy values of most *mdx* mice models when either LOMO or RLSMO cross-validations were used when compared to LOO for all different analysis methods. This was larger for the models and indices related to the one month old mice (for both male and female mice), with the models generated using the spectra of the three months old mice remaining almost unaffected by the different CV methods. Additionally, the coefficients of variability of indices associated with the three months old mice were lower than the ones related to the one month old mice. Taken together, these findings could indicate that there was a larger variability present in the *mdx* muscle at one month of age than at three months of age.

4.1.1.2 Classification performance and mdx muscle pathology

There is agreement in the literature that a sudden onset of necrosis takes place in muscle of young *mdx* mice at approximately 3 weeks of age (Beilharz *et al.*, 1992; Grounds and Torrisi, 2004; Roig, Roma and Fargas, 2004; Stupka *et al.*, 2004; Radley-Crabb *et al.*, 2014). By approximately 25 days there is a peak in necrosis after which it decreases significantly (Grounds *et al.*, 2008).

At approximately 4 weeks 20%-80% of the *mdx* limb muscles have been reported to be affected and inflammation, phagocytosis of necrotic tissue and early myogenesis are the main processes taking place in the dystrophic muscle, with the rest of the muscle fibres being unaffected (Grounds and Torrisi, 2004; Grounds *et al.*, 2008; Radley-Crabb *et al.*, 2014) (also in agreement with histological analysis in the present study-section 3.1.2.3, figure 3.45). This description of disease progression is not specific for gastrocnemius muscle or *mdx* mice from our colony but indicates that there is severe damage in the tissue that might affect a large number of myofibres and can explain the good classification at this stage (better than $SOD1^{G93A}$ vs. C57Bl/6 mice). However, since the percentage of intact muscle could be up to 80% a large number of spectra might be taken from unaffected tissue and thus it might be more difficult to separate from C57Bl/10 mice. This abundance of relatively normal tissue may explain why the classification performance is lower than that achieved for the older *mdx* mice. The differences in the spectra acquired from the different regions (e.g. regions with inflammation, phagocytosis, early myogenesis, unaffected tissue) in the 'diseased' muscle, along with potential variability in the amount of affected muscle could complicate the process of building a model and lead to the observed decrease in the accuracy values when LOSOCV was used compared to LOOCV at this age.

Around the 8^{th} week of age, necrosis stabilises at a low level of approximately 6% and the cycle of necrosis and regeneration continues throughout life, although it is further reduced by 12 months

of age (Grounds *et al.*, 2008; Radley-Crabb *et al.*, 2014). However, even if only 3% of *mdx* muscle fibres undergo necrosis every day, approximately 20% of muscle tissue will undergo necrosis in the course of a week (Grounds, 2014; Radley-Crabb *et al.*, 2014). The cycle of necrosis and active regeneration lasts approximately 3 weeks, hence, about at least 60% of myofibres will be actively affected by these processes at any time for adult mice, despite the low level of necrosis (Radley-Crabb *et al.*, 2014). Although necrosis and degeneration are still present to a smaller degree at about three months of age, the main histological findings at this disease stage are centrally nucleated muscle fibres indicating significantly higher muscle regeneration present in the *mdx* muscle compared to healthy muscle (Bulfield *et al.*, 1984b; Roig, Roma and Fargas, 2004; Grounds, 2014; Gutpell, Hrinivich and Hoffman, 2015) (section 3.1.2.3, figure 3.46). Between 2 and 3 months of age more than 60% of the gastrocnemius muscle myofibres have been shown to have centrally located nuclei and percentages of completely unaffected myofibres of hind limb muscles as low as 4.53% have been reported (Bulfield *et al.*, 1984b; Roig, Roma and Fargas, 2004; Radley, Davies and Grounds, 2008; Gutpell, Hrinivich and Hoffman, 2015). Thus, at this disease stage there is mainly regeneration present in the muscle, making it less heterogeneous than the earlier stage studied. In addition, there is also less completely unaffected muscle. All of the above could lead to better separation between the spectra acquired from healthy and diseased mice (due to less amount of unaffected muscle) and the fact that less processes take place in muscle could result in models with better ability to capture the inter-spectral and inter-subject variability. This is also indicated by the three months old mice models performance being less affected by different CV methods and having less variable performance (i.e. smaller coefficients of variability for both male and female three months old models compared to one month old mice models). Moreover, the fact that different processes are predominant in the muscle at the two different disease stages (necrosis and inflammation vs. regeneration) could partially explain the good separation between one and three months old mice.

4.1.1.3 Male vs. female mdx mice comparison

Using the RLSMO cross-validated models to compare female and male performance, the PCA-related models generated using the male spectra exhibited a better ability to distinguish between healthy and diseased tissue at the early disease stage, mainly due to higher specificity values. The PLS-DA models achieved similar accuracy values but with the female mice model achieving better sensitivity than specificity. When using RLSMOCV to compare the performance of the models generated using male three months old mice spectra compared to that of the female models, it's obvious that male mice models performed better with higher sensitivity and specificity values.

It has been previously shown that there are differences in the dystropathology between male and female *mdx* mice, with male mice experiencing more severe muscle damage at the early disease stage (Salimena et al. 2004). At three months of age muscle from male mice presents more inflammation and necrosis and less regeneration than female muscle (Salimena et al. 2004). When comparing the models generated from female and male mice it is important to acknowledge that 16 female mice were used per group and only eight male mice per group. Using a larger number of male (for example 16 mice) mice could affect the performance of the models and their difference compared to the performance generated using the spectra from the female mice.

4.1.1.4 Summary of the classification performance: SOD1^{G93A} mice

The two-group models generated using the spectra from the one month old SOD1^{G93A} and control mice were not able to distinguish between the two groups. There were no PCs with significantly different scores between the two groups and the accuracy achieved using PLS-DA was below 50% for LOMO cross-validation. Using the spectra acquired from three months old SOD1^{G93A} and C57Bl/6 mice, the PCA-LDA models achieved the best performance, with accuracy values above 77% and sensitivity values around or above 80% for all different CV methods. The coefficients of variability were below 5% for all indices. All the two-groups models generated using the spectra of one and three months old SOD1^{G93A} demonstrated a very good performance with sensitivity, specificity and accuracy values above 82% for all different CV values. The PLS-DA models demonstrated the best performance with sensitivity, specificity values about 94% for the different CV methods. The AUROC values were above 0.90 for all different analysis and CV methods and coefficients of variability remained below 4.5%. Different CV methods did not have a big influence on the any of the diagnostic indices.

4.1.1.5 Classification performance and SOD1^{G93A} muscle pathology

One month of age is a pre-symptomatic disease stage for this mouse model (Mead *et al.*, 2011; Bennett *et al.*, 2014). Thus, the inability of the technique to separate between the two groups could be due to biochemical changes in the muscle of the SOD1^{G93A} not yet being present or being too subtle. Consistent with this, there were no changes observed in the histological analysis of one month old SOD1^{G93A} muscle (section 3.1.2.3, figure 3.46).

As can be seen from the classification performance of the models described above, the changes present in the muscle of three months old SOD1^{G93A} were sufficient in order for the technique to be able to separate between SOD1^{G93A} and control mice as well as between one and three months old mice. This is an established disease stage for this mouse model and pathology is readily apparent in the histological analysis (section 3.1.2.3, figure 3.47) ; it was therefore expected that it would be easier to detect differences in the muscle when compared to the pre-disease onset stage (Hegedus, Putman and Gordon, 2007; Mead *et al.*, 2011; Bennett *et al.*, 2014; Chen *et al.*, 2020).

4.1.1.6 mdx vs. SOD1^{G93A} classification performance summary and limitations

When comparing one month old *mdx* and SOD1^{G93A} mice different CV methods did not have a big impact on the performance of the models for any of the analyses. All models achieved accuracy values above 90% with generally balanced sensitivities and specificities. All the diagnostic indices had very low coefficients of variability (below 2.5%). As already mentioned above, the one month old SOD1^{G93A} mice are at a pre-disease onset stage whereas the *mdx* mice are at an early disease stage with pathology evident in the muscle. This could also have made the separation between the two groups easier.

When using spectra acquired from the three months old mice the PCA-related models demonstrated accuracies above 82%, sensitivities above 87% and specificities above 73%. The

PLS-DA model was thought to have demonstrated the best performance due to similar accuracy values and more balanced sensitivity and specificity values. Coefficients of variability remained below 4.5%. Thus, as the two neuromuscular disorders progressed the ability of the technique to separate between the two different groups decreased, as expected. This could be due to similar disease-related changes taking place in the muscle, for example regeneration related processes, since centrally-nucleated myofibres were found in the histological analysis of both three months old *mdx* and SOD1^{G93A} mice (section 3.1.2.3, figure 3.46 and 3.48).

For these comparisons it is important to acknowledge the fact that mice with different genetic backgrounds were compared and this might have played a role in the ability of the model to separate the two groups. Furthermore, the SOD1^{G93A} mouse model is a transgenic model and the existence of the transgene might also cause artefacts that could lead to a better separation.

Notwithstanding the above mentioned limitations, the separation between neurogenic and myopathic pathology is a very important finding and possibly the most important one regarding clinical translation of the technique. Understanding the origin of the neuromuscular disorder could aid and accelerate the diagnostic process, by indicating more helpful diagnostic follow up tests (for example in many neurogenic conditions a muscle biopsy would not be necessary). In the clinic, neurogenic and myopathic disorders can sometimes be clinically indistinguishable; for example, lower motor neurone predominant MND and chronic myopathy can often present with similar symptoms and clinical examination findings. Electromoygraphy can usually distinguish between the two categories but sometimes difficulties arise as, for example large motor unit potentials and late recruitment, that are normally characteristic of denervation, might be also present in chronic myopathy (Paganoni and Amato, 2013; Sogawa *et al.*, 2017). Hence, additional information on the underlying aetiology of the neuromuscular condition could be very helpful.

4.1.2 Post *in vivo* Raman spectroscopy recording effects

Evaluating the extent of muscle damage caused by the needle insertion and the laser light during the *in vivo* Raman recordings is essential in order to investigate the potential of the technique as a clinical test. Furthermore, understanding the effect of the recordings on both the locomotion and motor performance of the mice and on tissue structure will be of use in determining whether *in vivo*, intra-muscular Raman spectroscopy can be used alongside other biomarkers of disease in pre-clinical murine studies.

The evolution of skeletal muscle injury after a surgically-induced needle injury has been studied in rats by Contreras-Muñoz *et al* (Contreras-Muñoz *et al.*, 2016). In this study it was shown that due to degeneration and inflammation following the injury, oedema was apparent in the first 24h after the experimental procedure and was then slowly re-absorbed over the following 3 weeks (Contreras-Muñoz *et al.*, 2016). To assess short- and long-term effects of the procedure on gait and motor function in the present study, rotarod and catwalk recordings were acquired one week before the (active and 'sham') Raman recordings one day and two weeks post-experiment. Post-mortem MRI was performed on a small number of mice in order to assess the potential tissue damage and its evolution six hours, two days and two weeks post-experiment.

For three groups (one month old SOD1^{G93A} and C57Bl/6, three months old C57Bl10 mice) the 'sham' recordings were not completed; less mice underwent the 'sham' procedure than initially

planned. Therefore, the catwalk and rotarod data are also undertaken on fewer mice for these groups. This decision was reached after taking into consideration that the mice that had undergone the active procedure had not exhibited any consistent change in the rotarod performance. Additionally, no sub-group that had undergone the 'sham' procedure showed a decreased rotarod performance.

The results from the motor function assessment tests and the imaging studies are discussed in the following sections.

4.1.2.1 Motor Function Assessment

4.1.2.1.2 Rotarod Test and post-experiment performance

The rotarod test is a difficult task that probes muscle strength and condition, as well as motor co-ordination and balance. It is often used in animal studies to assess motor function and measure motor performance (Mead *et al.*, 2011). Mice can be tested on the rotarod at a constant set speed, or at an accelerating mode. Advantages of rotarod testing include the fact that it is an automated procedure and that there is no need for training in order to be able to administer the test (Bennett *et al.*, 2014). The main disadvantage of the test is that it based on a forced exercise (Bennett *et al.*, 2014). Moreover, the performance can be quite variable which leads to large animal numbers being required for reliable results.

Initially, it was used by Dunham and Miya for neurological function studies (Dunham and Miya, 1957). Since then it has been employed in animal studies using various mouse and rat strains. Examples of such studies include investigation of the effects of drug administration on animal behaviour and motor function, recovery from stroke and traumatic brain injury (Hamm *et al.*, 1994; Chen *et al.*, 2001; Cenci and Lundblad, 2005). The test has been previously used in both mouse models we have employed in this study (SOD1^{G93A} and *mdx*) in order to assess changes in motor function and muscle weakness due to disease progression (Mead *et al.*, 2011; McDonald *et al.*, 2015; Oliván *et al.*, 2015), or due to an intervention such as drug administration in treatment studies (Mead *et al.*, 2011; Sali *et al.*, 2012; Potenza *et al.*, 2016; Ito *et al.*, 2017). It has, therefore, been used in this study as a gross assessment of the motor function and muscle strength following the *in vivo* Raman procedure injury.

From the rotarod data of all the groups, with the exception of the 3 months old SOD1^{G93A} mice, there is no indication that the *in vivo* Raman technique causes any impairment of the motor function of the mice. Additionally, the active and 'sham' procedures do not seem to have a different effect on the performance of these mice. Generally, both *mdx* and control mice seem to perform better with less variability in this test than SOD1^{G93A} and C57Bl/6 mice (at all time points of rotarod measurements). In keeping with these observations, rotarod performance of SOD1^{G93A} mice has been previously shown to vary among different groups from an early age (40 days) and most of the mice (SOD1^{G93A} and C57Bl/6) in this study were not able to stay on the rod for 300s even at this early disease stage (Mead *et al.*, 2011). Three months old C57Bl/6 mice exhibited a more consistent behaviour than one month old control mice. Since C57Bl/6 animals are not affected by disease progression this may be attributed to age related behavioural changes, for example, getting accustomed to handling.

The only group that exhibited a significant change in the rotarod performance after the experiment is the three months old SOD1^{G93A} mice that have undergone the active Raman

experimental procedure. The data acquired from the 'sham' mice sub-group followed a similar downward trend, but the decrease was not significant. The 'active Raman' group performed significantly worse only two weeks after the Raman measurements but not in the rotarod measurement the day after the experiment. This could suggest that the acute injury was not extensive and did not have a bigger impact on muscle function in these mice (as compared to the *mdx* mice, for example). The fact that only the three months old $SOD1^{G93A}$ mice exhibited a significantly decreased ability to perform the test two weeks post-experiment could be attributed to the fact that the diseased muscle is less able to cope with the long-term effects of the injury due to compromised structure and function and worse regenerative capacity. Since the active procedure is potentially more severe than the 'sham', the long-term injury of this procedure could be more extensive and therefore have a larger impact on the rotarod performance of the mice.

However, it has been previously shown that there is an impairment in the motor function that leads to a similar decrease in the rotarod performance of the $SOD1^{G93A}$ mice between 83 days (age of our mice at the baseline recording) and 104 days (age of the mice at the two weeks post experimental recording) of age due to disease progression (Mead *et al.*, 2011). Hence, it is perhaps more likely that the observed change in the rotarod data (or at least part of the change) was due to the disease progression in the three weeks interval between the two rotarod measurements rather than the experimental procedure. To further investigate the significant change in the performance of only one of the two groups and get a better understanding of how these two procedures affect motor function at this late disease stage and how much of the observed change was due to disease progression by recording the rotarod performance of mice at 104 days of age (n=8) that did not have any procedure (section 3.1.2.4.1b, figure 3.53).

Ordinary one-way ANOVA did not identify a significant difference between the performances of the mice that had undergone any of the two procedures and the 'control' mice. This along with the varying results coming from the two different three months old $SOD1^{G93A}$ sub-groups (active and 'sham' procedure groups) might reflect the inherent variability of the test, and the fact that a larger number of mice is required to understand better if the technique has a more severe long-term effect on the $SOD1^{G93A}$ mice. It has been shown before that the rotarod is a variable readout for this mouse strain (as also seen in the present data) and large n numbers are required (n=14 per group suggested by the protocol described in (Mead *et al.*, 2011). Thus, the lack of consistent findings for this strain at the third time point might be due to the small number of mice (n=8) used in the second post experimental time-point.

In conclusion, the performance of the mice was not significantly affected by any of the procedures one day post-experiment, indicating that the technique does not severely affect the motor function of the mice in the first few days post-procedure. As can be seen from the $SOD1^{G93A}$ and C57Bl/6 mice the performance of these mice on the rotarod can be quite variable. This renders drawing conclusions from the data acquired from the second post-experiment time point more difficult due to the smaller animal numbers used at this time point. The rotarod performance of the *mdx* and C57Bl/10 mice was much more consistent and most of the mice completed the full test at all three time points, indicating that the procedure did not affect the ability of the dystrophic and control mice to perform the rotarod test at any time point.

4.1.2.1.3 Catwalk Gait Analysis

The Catwalk is an automated gait analysis system that can quantitively assess gait and locomotion in rodents. The static and dynamic parameters generated by the catwalk system have therefore been used in numerous animal studies to assess motor performance and coordination. The main advantages of the Catwalk include objectivity, sensitivity and that it is an unforced activity. Unforced activities, for example, letting animals walk freely are less stressful for the animals and might prove more useful in detecting gait abnormalities, that might be more easily detectable during a normal activity. However, the large number of parameters generated can complicate choosing the most important ones and the subsequent data analysis. Additionally, the test and the manual labelling of the footprints can be time-consuming in studies with large animal numbers.

Catwalk has been extensively employed to characterize the motor phenotype and assess motor function in mouse and rat models of neurological, muscular and skeletal disorders that affect movement such as Huntington's Disease (Vandeputte *et al.*, 2010; Abada *et al.*, 2013), Parkinson's Disease (Vandeputte *et al.*, 2010), Multiple Sclerosis (Bernardes, Leite and Oliveira, 2017), Amyotrophic Lateral Sclerosis (Mancuso, Oliván, *et al.*, 2011; Mead *et al.*, 2011), Limb Girdle muscular dystrophy 2i (Maricelli *et al.*, 2016) and arthritis (Parvathy and Masocha, 2013b). Catwalk was also used to assess coordination and balance of *mdx* and control mice by Prigonine et al., in a study that investigated the existence of cerebellum dysfunction in Duchenne muscular dystrophy (Prigogine *et al.*, 2012). It has also been used in studies involving various rodent models to assess gait disturbances and the recovery process (with or without treatment) from stroke (Parkkinen *et al.*, 2013), peripheral nerve injury (with emphasis on sciatic nerve injury) (Deumens *et al.*, 2007; Crowley, Kataoka and Itaka, 2018), skeletal muscle injury (Ninagawa *et al.*, 2013; Vieira, Kenzo-kagawa and Cogo, 2016) and knee injury (Gabriel *et al.*, 2007). Viera et al. demonstrated a significant change in various gait analysis parameters like stand duration, maximum intensity, swing duration, swing speed and stride length post skeletal muscle injury caused by injection of snake venom (Vieira, Kenzo-kagawa and Cogo, 2016). Changes in the maximum contact area and stance phase duration were reported by et al after a severe crush injury of the TA (Ninagawa *et al.*, 2013).

Pain is expected to cause behavioural adaptations (Gabriel *et al.*, 2007). In the case of pain in the limbs these will involve changes in the gait. Moreover, injury of the limb muscles might lead to motor function impairment. Catwalk has therefore been used in this study, as in the previous studies, involving different types of injuries mentioned above, to investigate gait disturbances and motor impairment caused by the *in vivo* Raman injury.

The temporal, spatial and kinetic parameters are calculated separately for each paw. Most of these parameters are associated with the pressure that the animal puts on the paw and are therefore dependent on the animal's weight (Gabriel *et al.*, 2009). Since the Catwalk recordings in this study were acquired over the course of more than three weeks changes in the weight were expected especially for the younger mice. Hence, in order to prevent changes in the animal's weight from affecting the results the ratios of the diagonal paw parameters (LF/RH, RF/LH), instead of the absolute values (individual paw readings), were used (Liu *et al.*, 2013).

The analysis of the Catwalk data was done for each group separately as inter-strain or age differences can affect locomotion and thus Catwalk performance. Moreover, it is essential to

understand whether the experimental procedure affects healthy and diseased muscle tissue differently for future preclinical and clinical studies. Tissue undergoing pathological changes due to an underlying neuromuscular disorder might have worse regenerative capacity which could lead to the procedure having a more severe effect on diseased muscle.

Motor performance of the mice does not seem to be severely affected by any of the experimental procedures for any of the different strains and ages. There are no obvious trends observed in any of the parameters and there are no parameters consistently affected across the various groups. Mice that underwent the 'sham' procedure did not exhibit any significant changes in gait at any time point. Changes in a small number of gait parameters (either one day or two weeks after the procedure) were observed only in some of the groups that underwent the active Raman procedure.

Forelimb base of support (BOS) of one month old *mdx* mice was decreased a day after the experimental procedure. BOS is a parameter related to interlimb coordination. Larger BOS allows for more stable gait and an increase in the BOS could make up for an unstable gait (Mead *et al.*, 2011). The fact that the BOS of the forelimb decreased and the hindlimb BOS did not change significantly shows that the procedure did not cause instability. Since the procedure was not performed on the front limbs (only in both hind limbs) the observed change in the front limb BOS might be due to hindlimb rather than forelimb compensation (e.g. pain when exerting pressure on the hind limbs that affects posture and gait, changing, thus, the placement position of the fore limbs).

The LF/RH duty cycle ratio of the one month old *mdx* mice decreased one day post-experiment. This could be due to a decrease in LF duty cycle, an increase in RH duty cycle or both. In this instance the duty cycle ratio decrease was due to a post-experiment increase in the RH duty cycle. Duty cycle (percentage of stance in the step cycle) could be affected by the stand (duration of time the paw is in contact with the glass floor) and swing (duration of time the paw in not in contact with the glass floor). In the present data it was an increase in the stand phase of the RH step cycle (swing + stand) that led to an increase in the duty cycle. The change in the duration of the stand phase was not significant per se (i.e. stand parameter was not significantly larger post-experiment) but was significantly altered in the context of its percentage in the step cycle. Gastrocnemius muscles of the hind limbs (the muscles on which the procedure was performed) are related to the end of the stance phase (Krouchev, Kalaska and Drew, 2006; Mancuso *et al.*, 2011) and have been shown to be important muscles for ankle plantarflexion in mice (Krouchev, Kalaska and Drew, 2006; Charles, Cappellari and Hutchinson, 2018). The increase in duty cycle may therefore arise from a decreased ability of injured gastrocnemius muscle to push the hindlimb off the ground ending, thus, the stance phase and initiating the swing phase (Mancuso *et al.*, 2011). This could suggest that putting weight on the injured paw for a longer time is less painful than the movement required to lift the foot off the floor.

The percentage of time spent walking on a diagonal step pattern was found to be significantly increased in three months old C57Bl/10 mice at both post-experiment time points in comparison with the baseline recordings. It is easier to understand this change when changes in the other support stepping patterns are also taken under consideration even if they were not found to be significant. There was an upsurge in all step patterns related to support on two paws and support on three and four paws had the largest fall at both post-experiment time points (data not shown).

Similarly, three months old *mdx* mice used a significantly reduced three paw stepping pattern two weeks after the experimental procedure. Support on four paws was also decreased, although not significantly. These changes were compensated mostly by a non-significant increase in the diagonal step pattern and less so by an increase in single, griddle and lateral step patterns.
Normally mice spent approximately 70%-80% of their time walking on a diagonal stepping pattern and increased support on more than two paws (three or four) is often associated with instability; the mice attempt to make up for an unstable gait using more than two paws (Gabriel *et al.*, 2009; Mead *et al.*, 2011a; Maricelli *et al.*, 2016). The decrease in the three and four paw stepping patterns in both three months old *mdx* and C57Bl/10 mice indicates that the experimental procedure does not cause reduced gait stability that would urge the mice to rely on more than two paws. The increase in the diagonal pattern could result from habituation to the gait analysis test (since it was the second and third time that the mice were undertaking it), which could lead in the mice walking more normally in subsequent.
In conclusion, the 'sham' procedure appears to be harmless to the motor activity of the mice and does not seem to cause any gait disturbances at any time point after the experimental procedure. Only inconsistent changes, that were not similar with those reported in other injury paradigms, were observed in a small number of gait parameters in three of the groups that underwent the active experimental procedure. Hence, the technique was not thought to cause a severe impairment of the motor function consistently across the different groups for a long period of time.

4.1.2.2 Tissue Damage Assessment

MRI can effectively detect oedema and haemorrhage and is therefore used to assess muscle injury in preclinical and clinical studies (Lovering, Mcmillan and Gullapalli, 2009; Contreras-Muñoz *et al.*, 2016). It has been previously used in various murine models including $SOD1^{G93A}$ and *mdx* mice. These studies include monitoring the natural course of the disease in $SOD1^{G93A}$ mice (Mead *et al.*, 2011) and the effect of gene therapy in *mdx* mice (Park *et al.*, 2015). Moreover, MRI studies have been previously employed to successfully assess induced muscle injury in various murine models (Pratt *et al.*, 2012; Contreras-Muñoz *et al.*, 2016). It was therefore used for the evaluation of the potential skeletal muscle damage caused by the Raman procedure in this study.
Due to the small size of the animals' gastrocnemius muscle, the even smaller potential injury site and the lack of certainty that our recordings injure the muscle substantially, experimentation with the sample geometry and the scanning parameters was required. T_2 weighting was selected as it increases the contrast between healthy and injured (or diseased) tissue. This arises from the fact that tissue damage (for example muscle injury) results in an inflammatory response which includes the formation of oedema (Theodorou, Theodorou and Kakitsubata, 2012). Since water has long T_2 relaxation time, the tissue site with a higher water content (oedema) appears brighter in T_2-weighted images than the surrounding normal tissue (Allisy-Roberts and Williams, 2008; Pratt *et al.*, 2012). Different imaging techniques have different advantages. The main advantage of RARE imaging, which was chosen for this study, is the significant decrease of the time required to acquire a complete image. However, in RARE T_2 weighted images, T_2 relaxation time of fat increases resulting thus in a very bright fat signal (Allisy-Roberts and Williams, 2008). Therefore, fat suppression was used.

Parameters that define high resolution and high contrast MRI images are the following: TE (echo time), TR (repetition time), signal to noise ratio (SNR), spatial resolution and slice thickness. Echo and repetition times describe specific time intervals during the image acquisition and determine the image contrast. Figure 4.1 shows the T2 contrast in the images of two tissues with different T2 times (e.g. healthy and injured muscle tissue). Although the signal intensity is higher for both tissues for short TEs (TE_1 in figure 4.1), the contrast (i.e. the difference of the two signals) is not maximum, which in this work might have led to a difficulty in identifying the oedema. As can be seen in figure 4.1, the TE for maximum contrast (TE_2) is at a region of lower signal intensities (longer TE). However, TE must not be too long as this will lead to a very small signal and thus a poor SNR (Allisy-Roberts and Williams, 2008). Therefore, identifying and using the optimum TE between two specific tissue areas (healthy and injured muscle), required multiple test scans and was essential.

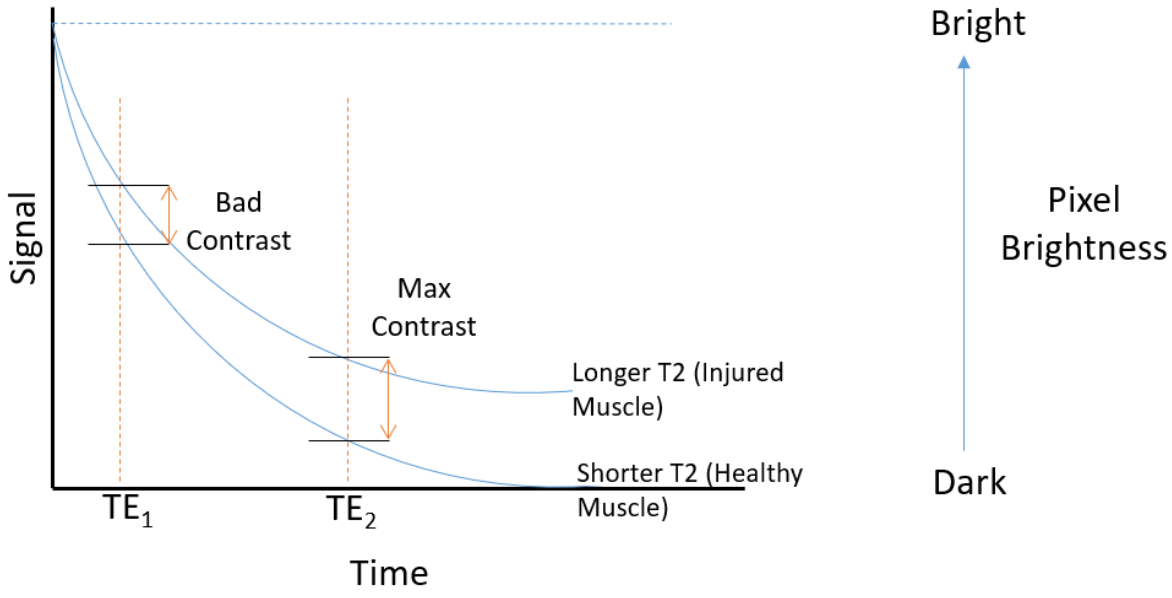


Figure 4.1**: **T_2 Contrast.
Maximum contrast between the tissues with different T_2 times is at TE_2. Shorter echo time leads to higher signals from both tissues but worse contrast. Longer echo time leads to very low signals.

Longer TR also increases the image contrast, but it also increases the overall scan time. Therefore, finding an appropriate TR was also very important. Since all noise decreases the contrast between tissues, SNR is quite an important factor that influences the image contrast as well (Allisy-Roberts and Williams, 2008). Signal averaging is a commonly used signal processing technique that increases the strength of the signal in comparison to the noise. In MRI, by increasing the number of excitations (Nex) or number of signal averages, each slice is acquired more than once and averaged in order to increase the SNR and thus its contrast. However, this again increases the scan time and optimisation is necessary.
Due to the small size of the sample very high resolution images were required in order to be able to spot and identify the tissue damage. Reduction of the pixel size (given by the ratio of the FOV and the matrix size), which for a 2-D image essentially gives the spatial resolution, and the slice

thickness, can only be achieved at the expense of imaging time. However, due to the use of the fast spin echo imaging modality, which largely reduced the imaging time, high spatial resolution images with an increased SNR (up to 50 image averages) were acquired with an imaging time of approximately 2 hours.

Bright areas in the MRI images that could be attributed to muscle damage due to our active Raman procedure were observed six hours post-injury. However, the tissue damage seems to be limited. The oedema, although probably still present in one of the mice scanned two days post-experiment, seemed to have decreased at the second time point and was undetectable in the mice scanned at two weeks. No bright areas due to inflammation were identified in the scans of the mice that had undergone the 'sham' procedure at six hours and two days after the injury, although only 2 mice were scanned at each time point. Hence, no mice were scanned at the third time point for this procedure.

4.1.3 *mdx* Intervention and preclinical MND studies

Studies 2 and 3 were performed in an attempt to assess the potential of the technique to detect more subtle differences or changes in the muscle within each neuromuscular disorder (not healthy vs. diseased and neurogenic vs. myopathic). The paradigms used were an intervention (study 2; running wheel in *mdx* mice) or different mouse models of the same disorder (study 3; $SOD1^{G93A}$ vs TDP-43).

4.1.3.1 mdx Intervention Study

4.1.3.1.1 CK measurements

Elevated levels of serum CK have been previously reported in both DMD patients and in the *mdx* mice and creatine kinase levels in serum are often used to measure muscle damage and the effects of an intervention in studies involving models of DMD (Anderson, Mcintosh and Poettcker, 1996; Granchelli, Pollina and Hudecki, 2000; Nguyen *et al.*, 2002; De Luca *et al.*, 2005; Spurney *et al.*, 2009). Despite the relationship between muscle damage and serum CK not being straightforward, CK levels have been previously shown to increase with exercise due to the exercise-induced exacerbation of the pathology (De Luca *et al.*, 2005; Hodgetts *et al.*, 2006; Radley and Grounds, 2006; Grounds *et al.*, 2008; Spurney *et al.*, 2009). Thus, serum CK was measured as a broad indicator of exercised induced muscle damage in the intervention study. As can be seen in figure 3.62 (section 3.1.3.1) there was an increase in the serum CK levels of the *mdx* mice after exercise, but the change was not significant. Hodgetts et al. also reported a similar insignificant increase in serum CK levels, despite a significant increase in necrotic fibres in limb muscles (Hodgetts *et al.*, 2006). Moreover, in our study we did not find a correlation between the individual distance run by each mouse and their serum CK levels. This could be due to the fact that we did not measure the change in the CK levels due to the exercise but the absolute values after the end of the exercise. Interestingly, we also did not observe a significant difference in the serum CK level between C57Bl/10 and *mdx* mice at three months of age. This was probably due to the large biological variation of the serum CK levels of individual mice, which a well-known problem of this biomarker (Radley and Grounds, 2006; Grounds *et al.*, 2008; Carlson, 2014).

4.1.3.1.2 Running Wheel Exercise, Pathology exacerbation and Classification Performance

It is well established that exercise intensifies myofibril necrosis and deteriorates muscle strength in *mdx* mice (Grounds *et al.*, 2008; De Luca, 2014; Manning and O'Malley, 2015). Since *mdx* mice normally develop a mild phenotype, exercise induced tissue damage that leads to exacerbation of muscle pathology and tissue damage is often employed in order to evaluate the efficacy of potential new treatments (Radley and Grounds, 2006; Brunelli *et al.*, 2007; Grounds *et al.*, 2008; Radley, Davies and Grounds, 2008). Voluntary wheel running and treadmill running are the most widely used *in vivo* running exercises. The main advantages of the wheel running exercise are that it is a voluntary activity and that there is no need for extensive training or human intervention. Additionally, each individual animal's voluntary capacity for running can be recorded (De Luca, 2014). One major disadvantage of that type of exercise is that since it is a voluntary exercise, the distance run by individual mice can differ considerably (De Luca, 2014). Increased damage in *mdx* muscle has been reported following as little as 24 and 48 hours of voluntary exercise, although several studies employ much longer periods of exercise, such as 4 weeks and 60 days(Carter *et al.*, 1995; Archer, Vargas and Anderson, 2006; Hodgetts *et al.*, 2006; Landisch, Kosir and Nelson, 2008).

In this study we used a single 48-hour epoch of wheel running exercise since it has been previously shown to exacerbate the disease in order to assess the potential of the technique to separate exercised from non-exercised mice (Radley and Grounds, 2006) . The 48-hour running wheel exercise is a quick and relatively easy intervention and allows to study of the induced necrosis before the development of new tissue/regeneration (Grounds *et al.*, 2008) . Exercise has a different impact on various limb muscles (depending on which muscles are recruited for the specific exercise regimes) and this needs to be taken into account. Using a single 24 hour bout of exercise Archer et.al showed that damaged myofibres significantly increased in quadriceps ($P<0.01$) and that quadriceps were more damaged than TA, diaphragm and gastrocnemius (Archer, Vargas and Anderson, 2006). A significant increase in damaged myofibres in gastrocnemius muscle was also observed, albeit less significant than quadriceps ($P<0.05$). Quadriceps were also shown to be more severely affected during running wheel exercise than TA after a single 48-hours exercise epoch, with the extent of exercised induced damage being insignificant for TA and the percentage of necrotic fibres significantly increasing from approximately 8% to approximately 13% for quadriceps (Radley and Grounds, 2006). In the same study gastrocnemius muscle was also reported to have reasonable amounts of damage after the running wheel exercise (10%-15%) (Radley and Grounds, 2006).

Since it has been shown that different muscles are affected differently by exercise a combination of *in vivo* and *ex vivo* measurements were obtained for this study in order to record from as many muscles as possible and better evaluate the ability of the technique to detect changes in muscle pathology.

In agreement with the literature, the models generated using the spectra from exercised and non-exercised TA muscle demonstrated very low sensitivity and specificity values (below 35%), indicating that the very limited amount of exercise-induced damage and alterations in muscle pathology, previously reported for this muscle, were not enough for separation between exercised and non-exercised TA. This was also similar for diaphragm.

The models generated using spectra acquired *ex vivo* from quadriceps achieved accuracy values between 63.4% and 71.3% for the different models, suggesting that the more severe changes in muscle pathology, reported in the literature, were more easily detectable by the technique. However, when LOSOCV was used there was a big drop in the all the diagnostic indices leading to accuracy values ranging between 48.5% and 61.4% for the different analysis methods. The sample size of quadriceps allowed for six spectra to be recorded from each sample. This in combination with the relatively small number of samples (8 exercised and 10 non-exercised samples) and the big variability in the performance of the mice (and thus probably the variability in the exacerbation of the disease/subject) can probably explain the decrease in the ability of the models to correctly classify spectra when LOSO was used. As a large number of spectra were acquired per sample, taking out only one spectrum would still leave a lot of information from that muscle in the training set, that would allow to correctly classify the left out spectrum. Moreover, since the running wheel performance of the mice was quite variable (see figure), the changes in the muscle due to exercise will also be variable and would vary considerably between different subjects. Thus, the number of samples used might have not been enough to capture this variability.

The models generated using the spectra acquired from exercised and non-exercised *mdx* mice *in vivo* from gastrocnemius muscle achieved accuracy values around 70% when LOO was used but there was a drop (62%-65.5%) when LOMOCV and RLSMOCV were used. These models also demonstrated very unbalanced sensitivity (above 54%) and specificity (above 70%). This indicates that the models were able to identify more accurately the spectra from the non-exercised mice. The drop in accuracies when LOMOCV was used instead of LOOCV suggests that the model could not capture the inter-subject variability sufficiently. The PLS-DA models demonstrated better performance with similar accuracy values but more balanced sensitivity and specificity values.

Since it has been previously reported that gastrocnemius muscle is moderately affected during the running wheel exercise, sensitivity, specificity and accuracy values lower than those achieved for separating between *mdx* and C57Bl/10 mice, where a clear difference in muscle pathology is present, were expected. As can be seen in figure 3.61 (section 3.1.3.1) the mice that underwent the wheel running exercise and active Raman recordings run an average of 6.5km with the total distance run by different animals ranging from 0.4 to 21.1 km. Similar total and average run distances and large variability in the performance of individual mice have previously been reported and as already mentioned above are a big limitation of this intervention (Archer, Vargas and Anderson, 2006; Hodgetts *et al.*, 2006; Radley and Grounds, 2006; Radley, Davies and Grounds, 2008; De Luca, 2014). For this study the big variability in the individual performances of these mice could lead to two problems. Firstly, some of the mice run very small distances and will, thus, have limited changes in muscle pathology and may therefore be classified as non-exercised mice. Secondly, since the running wheel performance of the mice was quite variable, the changes in the muscle due to exercise will also be variable. This in turn will make it more difficult to build a model that will be able to sufficiently capture this variability due to exercise will also be variable. This, in turn, will make it more difficult to build a model that will be able to sufficiently capture this variability and separate the two groups.

Altogether, the results of this study showed some promise in detecting potentially subtle pathological changes, but further histological analysis would be required to understand better how variable the pathological changes in the muscles actually were. A better understanding of

the correlation between run distance and exacerbation of the pathology could then better facilitate a larger study with more homogeneous groups (for example, mice that run above a set distance that seemed to be causing clear changes, mice than run below that distance, mice that did not exercise).

4.1.3.1.3 Comparing Different Muscles

Finally, the spectra acquired *ex vivo* from non-exercised quadriceps, TA and gastrocnemius muscles were placed together in the limb muscles group and compared to spectra acquired from diaphragm in order to explore the separation achieved by the technique between different muscles. Using LOMOCV decreased the ability of all models to correctly classify spectra due to a drop in sensitivity in the PCA related models and due to a drop in specificity in the PLS-DA model. However, all the models were able to separate between limb muscles and diaphragm with high accuracy values. The PLS-DA models yielded better sensitivity than specificity values. Using LOMOCV the PCA-LDA model was considered to have achieved the best performance as it achieved high accuracy value (93.5%) with the most balanced sensitivity and specificity values (both above 80%). Diaphragm has been shown to be more severely affected than the limb muscles with severity of pathology that resembles that of the DMD patients (Stedman *et al.*, 1991; Niebroj-Dobosz, Fidzianska and Glinka, 1997; Grounds *et al.*, 2008). This alongside structural differences in the different types of muscle could have led to the very good separation achieved between the different muscles.

4.1.3.2 Preclinical MND Study

4.1.3.2.1 Summary of Classification Performance

All models generated using the spectra from TDP-43 and TDP-43^{Q331K} mice achieved accuracy values above 75% with higher sensitivities (above 75%) than specificities (above 71%). The PLS-DA models demonstrated the best performance with accuracy values around above 80%. Different CV methods had a larger impact on the PLS-DA model compared to the PCA-related ones.

When comparing SOD1^{G93A} mice to TDP-43^{Q331K}, the RLSMO cross-validated PCA-related models yielded accuracy values around 77% with very balanced sensitivity and specificity values. The PLS-DA models achieved much higher sensitivity values (89.7%-96.6%) than the PCA-related ones but with much lower specificity values (62%-70%). Hence, despite similar accuracy values achieved with all the different analyses, PCA-LDA was considered to have demonstrated the best performance.

4.1.3.2.2 SOD1^{G93A} and TDP-43 mouse models: classification performance comparisons and limitations

In general, the technique was able to separate between TDP-43 mice with and without the Q331K gene mutation with similar sensitivity and specificity values to those achieved by the three months old SOD1^{G93A} vs. C57Bl/6 mice models (particularly the LOMO and RLSMO cross-validated models demonstrated very similar performances) at a disease stage, where phenotypic changes are apparent for both mouse models (Mead *et al.*, 2011a; Arnold *et al.*, 2013). The main difference between these comparisons is that for the SOD1^{G93A} mice the control mice were the non-

transgenic C57Bl/6 littermates, whereas for the TDP-43^{Q331K} mice the control mice used were the TDP-43 mice, which are transgenic mice in which the human TARDBP gene without the Q331K mutation has been inserted. Ideally in each animal study both types of control mice should be used as this would give an opportunity to better understand the effect of the artefacts present in the data due to the insertion of the transgene in the mouse genome. However, the overall classification performance for the TDP mice at a symptomatic disease stage is encouraging.

As described above good classification performances were achieved by the models generated using SOD1^{G93A} and TDP-43^{Q331K} mice spectra. Similar to the comparison between SOD1^{G93A} and *mdx* mice, it also needs to be stressed here that these are different mouse strains. Thus, the interpretation of the performance of these models needs to be done with caution as the above mentioned factors might cause differences that could facilitate a better ability to classify between the two models not due to disease related reasons. Moreover, not much is known yet about the TDP-43 mouse model and how well it resembles the human condition, as this is a relatively new and not very intensively studied model. It has been shown by Arnold et al. that there is denervation present in the TDP-43^{Q331K} muscle and previously undertaken unpublished work from my supervisors' groups has also revealed that CMAP amplitude is significantly lower in the TDP-43^{Q331K} mice compared to TDP-43 mice (Arnold et al., 2013; Stephenson, Alix, Mead unpublihsed data).

4.1.4 Comparison of Raman Spectroscopy performance with existing *in vivo* biomarkers

In this section established *in vivo* methods for studying the natural course of disease in MND and DMD preclinical models are compared to Raman spectroscopy.

4.1.4.1 mdx Mouse Model

Established *in vivo* biomarkers for monitoring disease progression and treatment efficacy for the *mdx* mouse model include various functional tests, histological analysis and imaging and have been presented in section 1.1.3.2.1. In comparison to Raman spectroscopy functional tests do not require anaesthesia and are completely non-invasive. However, Raman spectroscopy is a more objective technique since it is not confounded by behavioural and weight changes. Additionally, the mild phenotype may limit the ability of these techniques to be used as biomarker for this mouse model. Despite the limited phenotype, it has been shown that there are histological and biochemical changes present in *mdx* muscle, which could make techniques like Raman spectroscopy more suitable to monitor disease progression and alterations due to intervention. As already mentioned, the main disadvantage of MRI is that it is a very expensive technique and that long acquisition times, and, thus, prolonged anaesthesia, is required in order to get high resolution images. This can lead to difficulties in the maintenance of anaesthesia and very careful monitoring of physiological parameters is required for prolonged periods of time. This problem is not encountered with Raman recordings for which the maintenance of anaesthesia was required for much shorter period of time. Histological analysis is the gold standard assessment but requires sacrificing the animals and hence cannot be performed longitudinally.

In the current study, histological analysis showed pathology present in *mdx* muscle at both one and three months of age (section 3.1.2.3, figure 3.45 and 3.46). Raman spectroscopy achieved

good separation between the two mice groups (*mdx* and C57Bl/10 mice) at one month of age, better separation at three months of age and also very good separation between one and three months old *mdx* mice. The performance of the *mdx* mice on the rotarod test was not significantly different between one month old *mdx* and C57Bl/10 mice and between one and three months old *mdx* mice. The rotarod running time did, however, differ significantly between three months old *mdx* and C57Bl/10 mice ($p<0.01$). Interestingly, one day after the Raman procedure the three months old *mdx* mice performed better on the rotarod than before and there was no significant difference between *mdx* and control mice at any time point. Thus, in our study *in vivo* recordings utilising biochemical changes present in the muscle were more sensitive in detecting disease at an early stage and disease progression between different stages than rotarod.

4.1.4.2 SOD1^{G93A} mouse model

Similar tests as for the *mdx* mice are being used as *in vivo* biomarkers for the SOD1^{G93A} mouse model. Hence, the advantages and disadvantages compared to Raman spectroscopy won't be repeated in this section. A previously published work from the group of one of my supervisors will be summarised because it gives an insight on how different sensitive biomarkers are for the particular SOD1^{G93A} mouse model we used in the current study. This work on the SOD1^{G93A} model has shown that the rotarod is the most sensitive readout of motor function decline (Mead *et al.*, 2011). The report demonstrated that different cohorts of mice all experienced a decline in rotarod performance from as early as 45 days, a 20% reduction in performance from baseline by day 50-60 (for the different cohorts), and with a further decline in performance until approximately 130 days (Mead *et al.*, 2011). In this study, both neurological scoring systems and gait analysis were shown to be less sensitive, with disease onset, determined as the first change in the neurological scale (from 0 to 0.5), being detected around the 70 days of age (for the different cohorts). In addition, most of the gait analysis parameters did not show inconsistent differences between SOD1 and C57Bl6 mice until after the 75-days time point (Mead *et al.*, 2011). Some of the gait parameters, like diagonal and three limbs support and duty cycle were consistently disturbed for all the time points after 75 days, whereas others like stride length and stand time were significantly different between the two groups only after 100 days and only at specific time points (Mead *et al.*, 2011). However, it was also shown that the rotarod was the most variable measure and so larger animal numbers are required for reliable measurements (n=14) (Mead *et al.*, 2011). Finally, T2 MRI scans were shown to be able to detect disease progression due to muscle atrophy from 60 to 120 days, as lower limb volume normalised to animal weight was significantly between Tg and NTg mice at all time-points (60, 90, 120 days) (Mead *et al.*, 2011). For the same mouse model, it was shown that running wheel exercise can also detect alterations in motor function from an early age (approximately around 45 days) with much less variability than rotarod, making it, thus, easier to use smaller animal numbers (Bennett *et al.*, 2014).

In the current study, histological analysis showed pathology present in SOD1^{G93A} muscle at three months of age (section 3.1.2.3, figure 3.47) and Raman spectroscopy achieved very good separation between the two mice groups (SOD1^{G93A} and C57Bl/6 mice) at three months of age as well as between SOD1^{G93A} of one and three months of age. The performance of the SOD1^{G93A} mice on the rotarod test was not significantly different between one month old SOD1^{G93A} and C57Bl/6

mice. The rotarod running time did, however, differ significantly between three months old $SOD1^{G93A}$ and C57Bl/6 mice ($p<0.0001$) and between one and three months old $SOD1^{G93A}$ mice ($p<0.01$). Additionally, CMAP amplitudes were also shown to differ significantly between Tg and control mice at three months of age (section 3.1.6.1, figures 3.104, 3.105). We did not record any CMAPs from one month old $SOD1^{G93A}$ mice. Hence, *in vivo* Raman recordings, rotarod and electrophysiological recordings were all able to separate between $SOD1^{G93A}$ and C57Bl/6. Raman spectroscopy and rotarod were also able to detect changes between one and three months old $SOD1^{G93A}$ mice and did not detect any changes that would allow a good separation between the one month old mice. It would be interesting to further evaluate the potential of *in vivo* intra-muscular Raman spectroscopy by undertaking Raman recordings in other time points. For example, around the time point that the first rotarod performance changes were shown for that model (between 45 and 60 days) at the above mentioned study (Mead *et al.*, 2011).

4.1.5 Combined Probe Study

Electrophysiology is one of the mainstays of the diagnostic pathway for neuromuscular disorders. Hence, combining Raman spectroscopy with electrophysiological measurements in one probe could facilitate incorporation of the new technique in the current diagnostic pathway. Moreover, having a combined probe allows for both types of information to be used. Having the electrophysiological output may also allow the Raman measurement to be targeted to particular parts of the muscle. For example, it could be targeted to electrophysiologically abnormal and normal areas of a given muscle. In addition, the EMG signal could be used to ensure the probe is in the muscle of interest, which can be difficult to ascertain from visual inspection/muscle palpation alone.

In order to compare the performance of the combined probe relative to standard methods, CMAP measurements were also undertaken with ring electrodes and a concentric needle electrode from three months old $SOD1^{G93A}$ and C57Bl/6 mice. As expected, the CMAP amplitudes recorded with the combined probe were more similar to the CMAP amplitudes recorded with the needle electrode than the ring electrodes for each of the mice groups. This can be attributed to the fact that the ring electrodes are recording from all hindlimb muscles, whereas the needle electrode and the combined probe electrode are inserted in the gastrocnemius muscle and, hence only record the from that muscle. Similar to the recordings acquired using the standard methods (section 3.1.6.1, figure 3.105), the recordings obtained with the combined probe from the $SOD1^{G93A}$ mice (section 3.1.6.1, figures 3.104) were significantly different ($p<0.001$) from the C57Bl/6 mice.

Moreover, the adjustments done to the Raman probe in order to be able to acquire electrophysiological recordings did not seem to affect the ability of the probe to acquire Raman spectra, as can be seen in figure 3.106 (section 3.1.6.2) and the Raman recordings did not cause any significant changes in the electrophysiological recordings (section 3.1.6.2, figure 3.107). Of course, more rigorous testing is required in order to ensure the probe is acquiring Raman and electrophysiological recordings properly. For example, in order to better evaluate the potential and sensitivity of the electrophysiological component of the probe, it would be worth attempting to record spontaneous EMG activity, as such recordings would have much smaller amplitude (in

the microvolt range), and it would, thus, be more difficult to record. This study, does, however, demonstrate the potential of a combined EMG/Raman probe.

4.1.6. Data Analysis Considerations

Detection of cosmic rays, saturation of the detector, or very poor SNR can often affect the quality of the acquired spectra. Saturation of the detector due to fluorescence, often induced by tissue bleeding, and cosmic rays were intermittently encountered during spectral acquisition. Since these are easily detectable, when this was the case the acquisition was aborted, and the recording was repeated. The spectra were also visually examined prior the analysis and a small number of spectra was excluded due to remaining saturation and cosmic rays or due to a very different background shape. A more 'objective' approach to remove outliers could be developed by incorporating an outlier detecting algorithm (for example Q-test) as the first pre-processing step of the analysis (Butler *et al.*, 2016).

Raman spectra often require pre-processing before being analysed. The most commonly performed preparatory steps are noise and background removal and normalization (Butler *et al.*, 2016). Fluctuations of the number of detected photons due to their discrete nature (shot noise), dark current present in the detector, background fluctuations due to laser excitation source and background instabilities as well as temperature fluctuations of the detector and the laser emitting source (1/f component), all contribute to the noise present in the Raman spectra (Smulko & Wróbel, 2017; Smulko, Wrobel, & Barman, 2015). In order to reduce spectral noise and enhance the quality of the signal, acquisition settings can be altered as a first step. For example longer acquisition times and higher laser power can be used or multiple acquisitions for data averaging can be implemented (Smulko, Wrobel and Barman, 2015; Butler *et al.*, 2016; Smulko and Wróbel, 2017). In case of biological specimens and *in vivo* applications though, as is the case in the current study, an increase of the laser power or the acquisition times is limited by the need to avoid tissue damage and long examination times. In this study ten spectra were acquired per insertion and were subsequently averaged prior the data analysis in order to reduce the noise present in the spectra. As a next step, if changes in the acquisition settings are not enough, the spectra are often manipulated using techniques like the Savitzky-Golay filter and wavelet denoising (WDN) techniques (Trevisan *et al.*, 2012; Butler *et al.*, 2016). Although sometimes necessary, these methods often introduce artifacts (such as changes in the peak shape) that can affect classification (Smulko, Wrobel and Barman, 2015). Smoothing was not applied in the data prior the analysis in order to avoid spectral features degrading (Butler *et al.*, 2016).

Fluorescence generated in the sample or in the fibres, as well as unfiltered Raman scattering generated in the silica optical fibres, and CCD thermal fluctuations can affect the baseline of the spectrum and cause an overall background shape (Day and Stone, 2013; Stevens *et al.*, 2014; Smulko, Wrobel and Barman, 2015). In order to remove the baseline, a number of techniques have been proposed, such as first- or second-order differentiation and polynomial baseline fitting (Lieber and Mahadevan-Jansen, 2003). Although often useful in order to remove this big source of variation and help focus on the more subtle changes of interest , all of these methods have limitations and can lead to the introduction of systematic or non-systematic errors and unintended artifacts in the data set (Almond, 2012; Butler *et al.*, 2016). First- or second- order

differentiation introduces severe distortions in the Raman band shapes (Lieber and Mahadevan-Jansen, 2003). Additionally, since the first or second order derivatives do not have a regular Raman appearance, complex mathematical fitting algorithms have to be used to get back to a conventional Raman form (Lieber and Mahadevan-Jansen, 2003; Butler *et al.*, 2016). Estimation and subtraction of the background using polynomial fitting does not necessarily lead to a reproduction of the true spectral features and has been previously shown to introduce a large false variance in the data set which could result to multivariate techniques not taking into account more subtle real variance or to reduced classification performance (Shaver, 2001; Hutchings *et al.*, 2009; Almond, 2012; Byrne *et al.*, 2016). Hence, these techniques should be used with caution and validation is necessary whenever applied (Shaver, 2001; Hutchings *et al.*, 2009; Almond, 2012; Byrne *et al.*, 2016).

In the current data analysis, mean centering was applied prior to PCA and PLS-DA to scale the data and aid comparison. When the baseline is similar between spectra, the mean centering could remove the background sufficiently, since the average baseline is contained in the mean spectrum. The normalised spectra acquired for Study 1 before and after mean centering can be seen in Appendix D (figure D-1). After mean centering a large amount of the background has been subtracted from the spectra. No further baseline correction technique was used in the analysis in an attempt to better explore the uncorrected data, avoiding the possibility of introducing artifacts, and the ability of the multivariate analysis to extract biochemically important information despite the presence of the background. The spectra were also normalised prior to the multivariate statistical analysis in order to ensure that fluctuations in the intensity will not be mistaken as interesting features in the spectrum.

In general, background is usually represented by the first PCs in PCA-LDA. In this work, the loading plots of PC1 displayed a strong background contribution for most of the two-group models. However, the PC1 scores were rarely found to be significantly different between the two groups and so PC1 was most often excluded from the subsequent analysis steps. In the cases that it was found significant, for example in the model generated using the spectra of three months old *mdx* mice, biochemically relevant peaks could be identified despite the background contribution. Some background was also present in some of the models in PC2 (section 3.1.2.2.2, figure 3.34), whose scores were found to differ significantly between the different groups in most of the models. Some of the PLS weight plots, usually from components 2 and 3 (Appendix C, figure C-5, C-3) also displayed some background fluorescent signal. Hence, the variation in the fluorescence signal across the dataset could play some role in classification. However, biochemically relevant peaks, similar to those found in the difference spectra, in the background subtracted mean spectra and in loadings/weights plot of the other components are clearly identifiable in these loadings/weights plot too and so it cannot be assumed that classification is mainly done due the background component of the signal. Background subtraction, by fitting a second order polynomial, was attempted in a part of the data set (spectra acquired for study one) and was not found to improve the classification performance (data not shown) (Lieber and Mahadevan-Jansen, 2003). The above findings suggest that multivariate statistical models were able to extract and use the subtle spectral features that are important for classification between the different groups despite the fluorescent background present in the spectra. However, a more rigorous analysis using the above mentioned background subtraction technique and examination of the

loading plots could help to further understand the role of the background in spectral classification. Additionally, other methods for elimination of the background could also be tested, like green glass correction or subtraction of a stainless steel probe spectrum.

There was some noise contribution present in the loading plots of the later PCs and LVs, for example in PLS components (for example components 4 and 5 in C-13 and PC9 in C-10, Appendix C). In most of the 'noisier' loadings/weights plots relevant peaks could also be identified. Optimising the process of component selection could help in better understanding the contribution of each one of the components and could lead to a 'stricter' selection approach, that could denoise the data set further. For example, since during the cross-validation process the component selection takes place for each new training set, the times that each component is selected over all iterations of a cross-validation can be recorded. Then, only components selected in, for example more than 95% of the cross-validation iterations, could be used to build the 'final' model. This would be very interesting, particularly in combination with the repeated cross-validation, as in that type of CV there is a very large number of iterations and training sets and could result in a more robust assessment of the 'most important' components. Of course, if both noise and important information are present in a component, as seems to be the case in some of the later PCs and LVs, then some noise will still remain in the model if the component is selected and some important information might be lost if the component is excluded. Increasing the number of animals or samples used per group could also lead in some cases in denoising as the loadings and weights plots from the studies that were conducted with eight mice instead of 16 displayed a larger contribution of noise.

PCA-LDA and PCA-QDA yielded similar results and PLS-DA demonstrated a better performance from the PCA-related algorithms for most of the models. The difference in the performance could be due to the fact that PLS analysis finds the variance covariate with the group labels, whereas PCA does not and so some predictive information might be missed. In general, no more than 7 PCs and 8 LVs were used for the generation of the models.

4.1.7 Basis of Classification

The main peaks found in the spectra recorded from murine muscle in our study were located around 935 cm^{-1}, 975 cm^{-1}, 1000 cm^{-1}, 1030 cm^{-1}, 1044 cm^{-1}, 1076 cm^{-1}, 1121 cm^{-1}, 1170 cm^{-1}, 1205 cm^{-1}, 1260 cm^{-1}, 1300 cm^{-1}, 1335 cm^{-1}, 1444 cm^{-1}, 1550 cm^{-1}, 1570 cm^{-1}, 1610 cm^{-1}, 1654 cm^{-1}. Similar peaks, located around 934 cm^{-1}, 1002 cm^{-1}, 1179 cm^{-1}, 1264 cm^{-1}, 1301 cm^{-1}, 1336 cm^{-1}, 1447 cm^{-1}, 1576 cm^{-1}, 1655 cm^{-1} and 1737 cm^{-1}, were reported by Al-Rifai et al. in Raman spectra acquired from hindlimb muscles of BALB/c mice (Al-Rifai *et al.*, 2019). Huang et. al measured Raman spectra of different organs acquired from C3H/HeN mice (Huang *et al.*, 2011). The peaks that they identified as most prominent in skeletal muscle were: 851 cm^{-1}, 962 cm^{-1}, 1065 cm^{-1}, 1258 cm^{-1}, 1297 cm^{-1}, 1437 cm^{-1}, 1542 cm^{-1}, 1653 cm^{-1} and 1737 cm^{-1} (Huang *et al.*, 2011). Minamikawa et. al also acquired Raman spectra of skeletal muscle and other components around peripheral nerves from Wistar rats (Minamikawa, Harada and Takamatsu, 2015). The main skeletal muscle peaks were found to be located at 746 cm^{-1}, 1000 cm^{-1}, 1124 cm^{-1}, 1305 cm^{-1}, 1333 cm^{-1}, 1450 cm^{-1}, 1550 cm^{-1}, 1581 cm^{-1}, 1650 cm^{-1} (Minamikawa, Harada and Takamatsu, 2015). Pesolet et al. measured the Raman signal of isolated intact muscle fibres from the depressor

muscle of the giant barnacle and found the main Raman peaks located around 900, 937, 980, 1002, 1045, 1128, 1158, 1207, 1300, 1328, 1417, 1450, 1520, 1556, 1586, 1650 cm^{-1} (Pézolet *et al.*, 1980). The prominent Raman bands in the spectra of important muscle components like myosin, tropomyosin, actin and collagen can be found in Appendix B. A more detailed comparison between the peaks present in the recorded Raman spectra and the spectra of skeletal muscle and muscle components found in the literature can be seen in Table 3.24 in section 3.1.2.2.1.

In recordings acquired in this study the muscle spectrum was also shown to differ considerably from that of adjacent tissues like bone and blood (section 3.1.7). In the bone mean spectrum, only one peak around 955 cm^{-1} was clearly identifiable. This strong peak in the region between 950 cm^{-1} and 970 cm^{-1} has been previously shown to be a characteristic peak of hard tissue like bone and teeth and has been extensively associated with hydroxyapatite (Appendix A), a major component of these tissues (Huang *et al.*, 2011; Buchwald *et al.*, 2012). The blood mean spectrum contained the most clearly identifiable peaks. The main peaks were located around 1000 cm^{-1}, 1125 cm^{-1}, 1215 cm^{-1}, 1335 cm^{-1}, 1445 cm^{-1}, 1545 cm^{-1} and 1615 cm^{-1}. All these peaks are associated with amino acids and proteins and have been previously shown to be present in whole blood and red blood cells spectra (Deng *et al.*, 2005; Huang *et al.*, 2011).

Altogether, the peaks identified in the Raman muscle spectra were consistent with previously reported peaks in the literature and differed from the peaks present in different tissues. As expected, most of the peaks present in the muscle spectrum were associated with protein vibrational modes (for example table 3.23 in section 3.1.2.2.1) due to the high protein content of the muscle and many of them were also shown to be present in the spectra of myofibrillar and extracellular matrix proteins like myosin, actin, tropomyosin and collagen. As highlighted by Pesolet et al., the peak around 935 cm^{-1} and the strong amide III band located around 1650 cm^{-1} indicate an α-helical conformation of the proteins present in the muscle (Pézolet *et al.*, 1980). Amide III bands above 1240 cm^{-1} have also been associated with α-helix structure and the absence of a band below 1240 cm^{-1} in the Amide III (1225-1320) region of the muscle spectra also indicates very little β-sheet structure (Pézolet *et al.*, 1988) .

In general, the Raman spectrum of skeletal muscle is complex due to several overlapping bands (different protein bands as well as protein and lipid or nucleic acids overlap). Band overlapping often leads to broadening of the Raman peaks, as is often the case with the peaks in the Amide I region and was apparent in the mean spectra acquired in the study. Thus, identification of individual constituents from Raman spectra acquired from muscle tissue is quite complex. In addition to this, multiple molecular pathways and biochemical changes have been implicated in the progression of neuromuscular disorders. Hence, an unequivocal understanding of the changes occurring in the muscle and, thus, in the recorded spectra could not be formed.

The most consistent finding across the different two-group models was a decrease of the most prominent peaks of the spectra (935 cm^{-1}, 1000 cm^{-1}, 1300-5 cm^{-1}, 1330-5 cm^{-1}, 1445 cm^{-1} and 1650) as the disorders progressed, as demonstrated in the difference spectra of the *mdx* and control mice (both ages) and one and three months old *mdx* mice as well as three months old $SOD1^{G93A}$ and control and one and three months old $SOD1^{G93A}$ mice. In general, peak assignments on the PC and PLS loads are difficult. As already mentioned in the results section, using the score histograms and loadings plots of the PCA and PLS-DA analyses to understand the changes that occur in the muscle as the disorders progress is not trivial, due to the scores of the different

groups not being perfectly separated around zero in score histograms and/or the peaks not being centred around zero in the loading plots. However, biologically relevant peaks were present in the loading plots (as these were seen and explained in the background subtracted spectra). Additionally, the above mentioned peaks were in the loading plots in the direction, in which the spectra of the control (diseased vs. control models) or younger mice (different disease stages models) were predominant.

The peak located around 935 cm^{-1} has been previously assigned to C-C stretching vibrations of protein segments with an α-helical conformation (Yu, Lippert and Peticolas, 1973; Pézolet *et al.*, 1988; Stone *et al.*, 2002). This peak has also been highly associated with glycogen (Gautam *et al.*, 2015). The sharp band around 1000 cm^{-1} is one of the few peaks that is widely attributed to only one molecule and is related to the breathing mode of the phenyl ring in phenylalanine. The peaks located around 1300cm^{-1} have been attributed to Amide III mode of α -helical proteins as well as various bending modes of lipids and nucleic acid bases (Pézolet *et al.*, 1988; Stone *et al.*, 2004; Sato *et al.*, 2014). The band near 1335 cm^{-1} has been attributed to CH_3/CH_2 wagging modes of proteins and nucleic acids (Stone *et al.*, 2002; Al-Rifai *et al.*, 2019). The peak located around 1445 cm^{-1} has been assigned to the CH_2 bending mode arising from both proteins and lipids (Al-Rifai *et al.*, 2019). Finally, a band around 1650 cm^{-1} is assigned to an amide I mode of α-helical proteins and it also includes a contribution from a lipids C=C stretching mode (Sato *et al.*, 2014).

Hence, the most consistent differences between the spectra appeared to be related to changes in the α -helical proteins and lipids present in muscle. As can be seen in the histological images (section 3.1.2.3) both neuromuscular disorders there is damage in the diseased muscle structure, with inflammation, necrosis, regeneration, myofibre atrophy and hypertrophy being present in muscle at different disease stages. Since all the decreased peaks are present in the spectra of myofibrillar proteins, like myosin and tropomyosin, a decrease in these peaks could be indicative of the various processes and muscle fibre conformations present in the muscle that could result in a decreased amount of completely unaffected muscle fibres and myofibrillar proteins. Al-Rifai et al. also demonstrated a reduction in the protein related peaks of spectra acquired from BALB/c mice hindlimb muscles during the course of ischemic progression, associating, the decrease in the protein content with myofibre atrophy and inflammation present in the muscle due to ischaemia A large number of atrophic myofibres were also present in the three months old $SOD1^{G93A}$ muscle. Dystrophic muscle has also been shown to have higher hydration due to inflammation-induced oedema, myofibre necrosis and higher water content of regenerating myofibres (Dunn and Zaim-Wadghiri, 1999; Radley-Crabb *et al.*, 2014). Since water causes little Raman scattering, it would not have a large contribution to the signal (new peaks) but the increased water content could act to dilute the signal coming from muscle and, thus, contribute to decreased intensity in the more prominent Raman peaks. The peak intensity of the Raman signal has been previously shown to depend on water content present in porcine skin samples with a significant decrease of the intensities being observed for increasing water content (Kim, Byun and Lee, 2017). Finally, the decrease in the bands associated with α -helical conformation of proteins in Amide I and III regions as well as around 935 cm^{-1} was also observed by Gautam et al. in the spectra of fly models of different human muscle diseases. Interestingly, oxidative stress, that has been shown to take place in the muscle of both mouse models, has been previously demonstrated to cause similar structural changes (Dobrowolny *et al.*, 2008; Schill *et al.*, 2016; Rivas-Arancibia *et al.*, 2017).

Many more peaks were present in the spectra and loading plots of the PCA and PLS-DA analyses. Hence, there are more biochemical features that could contribute to spectral classification and further work could potentially elucidate more significant classifying features. A better understanding of the changes occurring in the Raman spectra as the different neuromuscular disorders progress could allow for a more targeted approach to spectral acquisition, for example acquisition of spectra truncated in a limited wavenumber range of interest that could lead to decreased acquisition times (Almond, 2012). It is, however, not necessary in order to use the technique as a biomarker for neuromuscular disease, so long as the classification depends on biochemically relevant differences and can, thus, be reproduced consistently.

4.2 Human Tissue Recordings

Ex vivo measurements were acquired using the fibre-optic probe and a commercial Raman microscope system from healthy and diseased muscle tissue and the results are summarised and discussed in the following sections. The last hypothesis presented in section 1.3.2 stated that Raman spectroscopy of human muscle samples can detect muscle pathology and distinguish between neuromuscular diseases. The classification performance of the two group models generated using the spectra acquired from healthy and pathological (myopathy, MND and mitochondrial disorders) human tissue are summarised and discussed in section 4.2.1.1. Despite the limitations of this data set pathological findings were observed in the tissue from patients with MND and mitochondrial myopathies and these samples were well separated from healthy tissue, supporting, thus, the first half of the hypothesis. The classification performance of the multi-group models was hampered by very low sensitivity values for all the groups apart from the tissue acquired from MND patients. Thus, as discussed in section 4.2.1.2 in more detail, more tissue samples are required to further investigate whereas the technique is able to separate effectively between different neuromuscular disorders.

4.2.1 Dataset Limitations

There were some limitations associated with the human samples, that render the data analysis and interpretation of this dataset challenging:

- For the samples acquired in biopsies (‘myopathy’, ‘not myopathy’ and ‘mitochondrial disorders’ groups) a big limitation was that the Raman spectra were not acquired from the same samples that underwent histological examination. Instead an extra muscle sample was acquired for the purposes of the study. Hence, there is a possibility that slightly different pathological features, or indeed, not pathological features could be present in the samples that underwent histological examination and the ones from which the Raman spectra were acquired.
- The samples were acquired from different muscles and were harvested in different ways (section 2.2.1). Additionally, healthy volunteers and patients were not age-matched. Healthy tissue is very difficult to acquire since muscle biopsy is an invasive procedure. Finding an adequate number of samples from the same muscles or from age-matched volunteers would be quite difficult for this early phase research.
- The numbers of samples in each group differed considerably and this could introduce bias in the models.

Further limitations related to each one of the groups are discussed in the next sections in the context of the classification performance of the models.

4.2.2 Biopsy findings: Classification: Summary and Limitations

4.2.1.1 Two-group models

4.2.1.1.1 Myopathy vs. Healthy

The LOSO cross-validated models generated using the probe spectra from 'myopathy' and 'healthy' samples achieved accuracy values around 58% for the PCA analysis and 77% for the PLS-DA analysis. Using LOSOCV instead of LOOCV caused a big drop in both sensitivity and specificity values of these models indicating that the variability between different samples (due to biological or pathology variance) was not sufficiently captured by the models. The big variability could also be the reason for the big difference in the performance in the PCA-LDA and PLS-DA models. Interestingly, the loadings plot of the most significant PC and the weights plot of the first PLS component were very different for that comparison (which was almost never the case with the *in vivo* and the rest of the *ex vivo* data), with the first PLS component displaying more biochemically relevant features. Since the two models performed so differently it is not easy to draw any certain conclusions on the ability of the technique to separate 'myopathy' from 'healthy' samples but the fact that the PLS-DA model demonstrated a good ability to classify between the two groups with known biochemical features present in the weights plot is encouraging.

For the 'microscope' data, there were no PCs with significantly different scores between the two groups identified, whereas the LOSO cross-validated PLS-DA model achieved an accuracy of 60%, which was also considerably decreased compared to LOOCV.

The main limitation of the 'myopathy' group is that it was a very heterogenous group as it included different types of myopathy. Since different types of myopathies can cause different histological and biochemical changes in the samples, the spectra of this group could have been quite variable. For example, while muscle fibre size variation and non-specific signs of regeneration (e.g. internal nuclei) are common to myopathies, specific changes are also seen. For example, in the samples collected, several were of dystrophic myopathies (n=4) which can also show differing features depending upon the exact cause, for example, protein aggregates, fibre type predominance and amount of necrosis. Inclusion body myositis (n=1 in the cohort) displays rimmed vacuoles, positive aggregates and an inflammatory response which is primarily around the myofibres (Pestronk, no date). As, there was only a small number of samples from each type of myopathy, that could hinder the identification of the spectral characteristics of each myopathy type. Thus, to improve the statistical validity of the models, and render the analysis more representative of the classification of new spectra, the models would ideally include a larger number of samples from each type of myopathy (Kendall, 2002; Almond, 2012). A larger dataset could lead to a better representation of the spectral characteristics of each type of myopathy in the model, in turn capturing more accurately the variation in the spectra coming from different types of myopathies. A larger data set would also allow for the variability within each patient/sample to be better accounted for (Kendall, 2002; Almond, 2012). Thus, further investigation is required in order to determine the potential of the technique.

4.2.1.1.2 Mitochondrial Disorders vs. Healthy

A good ability to discriminate between 'healthy' and 'mitochondrial disorders' samples was achieved by all the probe and microscope models with accuracies about 75% for all the different analysis and CV methods and the microscope and probe models demonstrated similar performances. Higher sensitivities (mostly above 80%) than specificities (above 60%) were achieved for almost all the models with only the LOSO cross-validated PLS-DA model displaying balanced sensitivity and specificity values around 75%. The different CV methods did not affect considerably the ability of most models to correctly classify the spectra. The 'mitochondrial disorders' groups also consisted of samples from patients with different types of mitochondrial diseases but in this case the majority of the samples were from patients with m.3243A>G mitochondrial disease (table 3.63, section 3.2.1.1) and this could have potentially led to a 'less variable' data set. Similar to the results of the *in vivo* work, the difference spectra and loading plots of the probe and microscope models also demonstrated a decrease in the most prominent peaks for the diseased samples.

More samples from different types of mitochondrial disorders could help understand better the ability of the technique to discriminate between mitochondrial disorders from healthy tissue and between different mitochondrial disorders.

4.2.1.1.3 MND vs. Healthy

Different samples were probed with the fibre-optic probe (n=4) and the microscope (n=10). The probe spectra were collected prior the initiation of this project as preliminary data and there was no access to the microscope at that time. The ten 'MND' samples that were used in this study and were measured using the microscope did not produce reliable spectra using the probe due to an excess of fluorescence, that caused saturation of the spectrometer.

As already mentioned in the results section spectra from only four 'MND' samples were acquired using the fibre-optic probe, limiting the reliability of the classification performance. Additionally, the loadings plots revealed that the classification could have largely relied on periodic features present in the spectra of the 'MND' samples that could be the result of etaloning. Such features were not seen in the spectra or loadings of any other samples/ comparisons.

Using the microscope spectra accuracy values above 90% were demonstrated for all the models. PLS-DA yielded a better performance, with an accuracy of 95.1% and more balanced sensitivity and specificity values. This could be partially due to a larger number of components used to build the PLS-DA (optimal number of components:8), PCA-LDA (3 significant PCs) that could have led a small amount of important information being 'left out' by the PCA-LDA model. Similar to the results of the *in vivo* work in the $SOD1^{G93A}$ mice, the difference spectra and loading plots of the microscope models also demonstrated a decrease in the most prominent peaks for the diseased samples, possibly indicating the translational potential of the technique.

The main limitation with this comparison is that the 'MND' tissue was acquired post-mortem whereas the tissue from healthy volunteers was acquired during a surgical procedure. Hence, more 'MND' samples are required that would ideally be harvested in the same way as the 'healthy' tissue and would allow for a 'fairer' comparison. Additionally, it is important to test the technique on muscle from MND patients at an earlier disease stage (when the technique would

be useful in clinic), as tissue acquired post-mortem may manifest severe disease could thus lead to a much better classification performance (when compared to healthy tissue). The very good classification performance achieved with the 'microscope' models is, however, very encouraging and demonstrates that it is worth to further investigate the potential of the technique to detect muscle pathology originating from neurological conditions.

4.2.1.2 Multi-group models

The potential of the technique to be used as a screening tool for neuromuscular disorders was investigated with multi-group models. The five-group models ('myopathy', 'mitochondrial disorders', 'MND', 'not myopathy', 'healthy') aimed to test the ability of the *ex vivo* recordings to accurately classify samples in the three different pathological groups, the group with no neuromuscular condition ('not myopathy' group) and the 'healthy' group. The four and three-group models aimed to explore the potential of the technique to separate patients with neuromuscular conditions who would necessitate a biopsy to further investigate the type of their disease (mitochondrial disorders, myopathies) from patients that would not need a biopsy (MND, 'not myopathy' group), reducing, thus, the number of unnecessary expensive and invasive procedures.

The common characteristics in the performance of all the different multi-group models (probe/microscope, PCA-LDA/PLS-DA) was that most of the groups demonstrated high specificity values. For some groups in each model good sensitivity values were also achieved whereas others demonstrated very low sensitivities. CV method had a big impact on the ability of the models to correctly classify spectra in most of the groups, with most values decreasing considerably when LOSO was used. In general, no models managed to achieve sensitivity, specificity and accuracy values above 50% for all the groups, mainly due to very low sensitivities achieved for certain ones. All together this could suggest that the technique might not be ideal as a screening tool, for which a very good sensitivity is required. However, more investigation is required to draw any certain conclusions. As already mentioned above the dataset had several limitations, which if overcome, might lead to a different result. Groups of larger and equal numbers of samples with well-defined pathology could lead to more robust models, that could correctly classify new samples more easily. Larger sample numbers could also result in the model being able to better classify new samples even if different pathologies are grouped together (for example as 'biopsy' group) since the model might be able to account better for the variability between the different disorders and the different patients. If the models still achieve much better sensitivity that specificity values for all the different groups, the classification thresholds could be adjusted to achieve a more balanced performance.

4.2.3 Final Diagnosis: Raman Spectroscopy, EMG and Biopsy Comparison

As already described in the introduction, two of the mainstays of the current diagnostic pathway of neuromuscular disorders are electromyography (EMG) and muscle biopsy. The main limitations of these techniques are discussed in section 1.1.1. Raman spectroscopy could potentially be used in clinic, alongside EMG, using, for example, a combined Raman/EMG probe, as one of the first steps in the diagnostic investigations. If Raman spectroscopy manages to be more specific than

EMG and can separate disorders within each of the two main groups of neuromuscular disorders (myopathic and neurogenic disorders), a combined probe could enhance and specify the information currently acquired with EMG. Additionally, the two techniques combined might be able to better indicate if biopsy is required for a patient. For this to be achieved models generated using larger numbers of Raman spectra with high sensitivity and specificity values for separating samples with pathological findings that necessitate a muscle biopsy from the ones that do not would be required. Finally, identification of optimum muscle biopsy sites to ensure that muscle pathology is properly sampled could also be achieved.

Diagnostic indices of EMG and muscle biopsy for detecting myopathies EMG were determined for the samples acquired from open biopsies in Sheffield (see table 3.66). As can be seen in table 3.90 in section 3.2.3.1 muscle biopsy demonstrated sensitivity, specificity and accuracy values above 83%. EMG achieved an accuracy value of 68% with a sensitivity below 60% (table 3.89). Interestingly, both tests demonstrated a better ability to correctly classify the non-myopathic cases. Finally, Raman spectroscopy was not able to correctly classify the samples using the Raman spectra with accuracy values below 50%. This finding is, of course, disappointing, it could, however, be due to the heterogeneity of the samples included in the two groups. As already mentioned above the myopathy group consisted of different types of myopathies, with only a few samples in each type. Similarly, as can be seen in table 3.66 (section 3.2.1.1), the 'not myopathy' groups consisted of limited numbers of samples with age-related and non-specific pathological changes and 'normal' ones. The heterogeneity of both groups, and, hence the possibly large variability, present in the dataset might have hindered the process of creating a classification model with good predictive ability. Larger numbers of samples in each one of the subgroups could allow for the variability in the dataset to be more sufficiently capture and, thus, allow for a better classification performance. Hence, further investigation is required to understand if and how Raman spectroscopy could be used in the clinical setting.

4.2.4 Raman Spectroscopy in the Clinical Setting: Further Considerations

As already mentioned above the desired performance of a test highly depends on how it will be used in the clinical setting. If the aim is to use the developed technique as a screening tool for neuromuscular disorders in order to identify which patients show signs of neuromuscular disease and would need further examinations (or what type of examinations they might need, for example 'biopsy' and 'not biopsy' groups) then high sensitivity values are required, and ROC curves analysis could be employed to optimise sensitivity. If the models cannot simultaneously achieve high sensitivity and specificity values and the increase in sensitivity leads to a decrease in specificity, this approach could lead to increased numbers of examination tests being undertaken as more people would be falsely diagnosed as having a disease.

On the contrary if the aim is to use the test in order to reach a final diagnosis, high specificity values are required in order to ensure that a healthy individual or someone with less severe muscle problems (for example with age-related changes) will not be misdiagnosed as having a neuromuscular disorder. As many neuromuscular disorders are very severe and not treatable (for example MND) false positive results could have serious adverse consequences for the patient and his family.

As already stressed in the previous sections further testing is required to understand the potential and limitations of the technique to probe muscle health *in vivo*. A larger study with more samples from patients with different neuromuscular disorders and at different disease stages could be the next step to further explore the ability of the technique and the role it could fulfil in the clinical setting.

4.3 Future Work

4.3.1 Preclinical Setting

The present work has demonstrated the ability of *in vivo* intramuscular Raman spectroscopy to detect muscle pathology of different origins at different disease stages in preclinical murine models. The project, however, had limitations that could be overcome with future work. Re-analysing the collected data using a baseline correction method could provide a more thorough understanding of the role the background changes in the classification performance of the models. Additionally, histological study of tissue acquired from the exercised *mdx* mice and the TDP-43 and TDP-43^{Q331K} could offer a better understanding of the effect of exercise on the muscle of individual *mdx* mice and the changes present in the muscle of the TDP-43 mice. This could lead to a better evaluation of the outcome of the respective studies the correlation between existent pathology and classification performance.

In order to better evaluate the technique, a similar study to that undertaken in study 1 can be conducted in other disease stages of the *mdx* and SOD1^{G93A} mouse models. Similarly, a comparison of the performance of the Raman recordings with already existing biomarkers could be included. Of particular interest would be to use SOD1^{G93A} mice at a young age but when a phenotypic change can already be detected (for example around 45 days old mice) and compare them to the one three months old mice recordings. Additionally, a study on the effect of longitudinal Raman recordings on muscle function could be conducted in order to investigate if the technique can be used for repeated measurements on the same subjects. These investigations could further elucidate the ability of the technique to detect muscle pathology at different disease stages and its potential as a biomarker for monitoring disease progression. Finally, the developed technique could be used in other mouse models of these or other neuromuscular disorders.

4.3.2 Clinical Setting

Ex vivo recordings from human muscle

The present work has demonstrated the ability of the probe to record spectra from human muscle *ex vivo*. It was clear from the Raman classification models that population and pathology variance were not sufficiently captured in the current data sets. In order to better evaluate the performance of the technique on human tissue, further analysis of the collected spectra could be undertaken (for example background subtraction and different analysis techniques). If these models also fail to capture the variance in the data set further work can focus on acquiring more 'myopathy' samples from patients with disorders such as inflammatory myopathy, muscular dystrophies and inclusion body myositis in order to build on the already existing samples. Subsequently, the developed technique could be used in muscle samples acquired from patients with other neuromuscular disorders in order to get a better understanding of the areas that the developed technique would be more useful in a clinical setting.

Translation into clinics

If the *ex vivo* results further demonstrate the ability of the probe to detect pathology and separate neuromuscular disorders, the next step towards translation will involve building and testing the combined Raman/EMG probe. The EMG component of the probe will need to be tested to ensure

that recording spontaneous EMG activity (which is in the microvolt range) can be done reliably. Tests to ensure the acquisition of reproducible and reliable Raman spectra with good signal to noise ratio will also need to be undertaken. In order to be able to use the probe as a clinical *in vivo* tool work will need to be undertaken to ensure reproducible spectral measurements for different probes and different systems (Almond, 2012). This step could involve developing correction algorithms in order to remove systematic differences in the measurements due to different spectrometers being used.

After developing the probe, a clinical trial will need to be undertaken in order to evaluate the performance of the probe *in vivo* in human muscle, assess the safety of the probe and build a dataset of Raman spectra. A feasibility study could be the first step in order to demonstrate the performance of the probe *in vivo* and also test the safety of the probe on patients with different neuromuscular disorders before using the probe on a larger scale. A significant *in vivo* database from spectra acquired from healthy muscle and pathological muscle from patients with neuromuscular disorders will need to be built in order to develop robust classification models. Building this database *in vivo* ensures that recordings are acquired under the exact same conditions that the probe will be used in clinic. It could, however, mean that it will take much longer to build a substantially large database that will be able to sufficiently capture population and pathology variance. After demonstrating the probe performance *in vivo* and the safety of the probe a larger, randomised trial could take place in order to further evaluate the performance of the technique. Being able to compare spectra acquired with different probes and different spectrometers would allow for multi-centre clinical trials to take place and larger datasets of Raman spectra from patients with different neuromuscular disorders to be created, regularly updated and, finally, be shared between different centres.

5. Overall Conclusions

In this project a method for recording from muscle *in vivo* using a fibre-optic Raman probe was developed. The results demonstrate the ability of intra-muscular *in vivo* Raman spectroscopy to probe muscle health *in vivo* in preclinical models of neuromuscular disorders. The developed technique was able to detect muscle pathology and accurately distinguish between muscle disease originating from different neuromuscular disorders. Moreover, it was shown that the recordings did not impair motor function. Further investigation is required in order to better understand how sensitive the technique is at subtle changes in muscle pathology due to disease progression or an intervention. However, the results highlight the potential of the technique to be used in the preclinical setting (for example in preclinical drug studies).

Parallel to animal studies recordings from human muscle specimens with different types of disease were acquired *ex vivo*. This work demonstrated the ability of the probe to record spectra from human muscle and distinguish with good accuracy values between healthy and diseased tissue for certain neuromuscular disorders. The spectra acquired from human muscle displayed very similar characteristic features as the spectra acquired from animals *in vivo* highlighting the translational potential of the technique. Further work is required in order to better explore the ability of the technique to detect different types of muscle disease at different disease stages and finally better understand the role that Raman spectroscopy could potentially fulfil in the clinical setting.

Bibliography

Aartsma-Rus, A., Ginjaar, I. B. and Bushby, K. (2016) 'The importance of genetic diagnosis for Duchenne muscular dystrophy.', *Journal of medical genetics*, 53(3), pp. 145–51. doi: 10.1136/jmedgenet-2015-103387.

Abada, Y. K. *et al.* (2013) 'Assessment of Motor Function , Sensory Motor Gating and Recognition Memory in a Novel BACHD Transgenic Rat Model for Huntington Disease', 8(7), pp. 1–13. doi: 10.1371/journal.pone.0068584.

Al-Chalabi, A. and Hardiman, O. (2013) 'The epidemiology of ALS: a conspiracy of genes, environment and time', *Nature Reviews Neurology*. Nature Publishing Group, 9(11), pp. 617–628. doi: 10.1038/nrneurol.2013.203.

Al-Rifai, R. *et al.* (2019) 'Subcutaneous and transcutaneous monitoring of murine hindlimb ischemia by in vivo Raman spectroscopy.', *The Analyst*. Royal Society of Chemistry, 144(15), pp. 4677–4686. doi: 10.1039/c8an02449a.

Alfano, R. . *et al.* (1991) 'Human breast tissues studied by infrared fourier transform Raman spectroscopy', *Lasers in Life Sciences*, 4(1), pp. 23–28.

Alimova, A. *et al.* (2009) 'In vivo molecular evaluation of guinea pig skin incisions healing after surgical suture and laser tissue welding using Raman spectroscopy.', *Journal of photochemistry and photobiology. B, Biology*, 96(3), pp. 178–83. doi: 10.1016/j.jphotobiol.2009.06.004.

Allisy-Roberts, P. and Williams, J. (2008) *Farr's Physics for Medical Imaging*. 2nd edn. Elsevier.

Almond, M. L. (2012) *Towards Objective Endoscopic Diagnosis of Barrett's Associated Early Neoplasia Using Fibre-Optic Raman Spectroscopy*. Cranfield University.

Amthor, H. *et al.* (2004) 'Albumin targeting of damaged muscle fibres in the mdx mouse can be monitored by MRI', *Neuromuscular Disorders*, 14(12), pp. 791–796. doi: 10.1016/j.nmd.2004.08.004.

Andersen, J. L. *et al.* (2000) 'The 2B myosin heavy chain gene is expressed in human skeletal muscle.', *J Physiol*, 539, p. 29P–30P.

Anderson, J. E., Mcintosh, L. M. and Poettcker, R. (1996) 'Deflazacort but not prednisone improves both muscle repair and fiber growth in diaphragm and limb muscle in vivo in the mdx dystrophic mouse', *Muscle and Nerve*, 19(12), pp. 1576–1585. doi: 10.1002/(SICI)1097-4598(199612)19:12<1576::AID-MUS7>3.0.CO;2-7.

Archer, J. D., Vargas, C. C. and Anderson, J. E. (2006) 'Persistent and improved functional gain in mdx dystrophic mice after treatment with L-arginine and deflazacort', *The FASEB Journal*, 24, pp. 1–25.

Arnold, E. S. *et al.* (2013) 'ALS-linked TDP-43 mutations produce aberrant RNA

splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43', *Proceedings of the National Academy of Sciences of the United States of America*, 110(8). doi: 10.1073/pnas.1222809110.

Balaban, B. *et al.* (2005) 'Corticosteroid treatment and functional improvement in Duchenne muscular dystrophy: long-term effect.', *American journal of physical medicine & rehabilitation*, 84(11), pp. 843–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16244521.

Ballabio, D. and Consonni, V. (2013) 'Classification tools in chemistry. Part 1: Linear models. PLS-DA', *Analytical Methods*, 5(16), pp. 3790–3798. doi: 10.1039/c3ay40582f.

Bame, M. *et al.* (2014) 'Amino Acids as biomarkers in the SOD1G93A mouse model of ALS', *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Elsevier B.V., 1842(1), pp. 79–87. doi: 10.1016/j.bbadis.2013.10.004.

Baraga, J. J., Feld, M. S. and Rava, R. P. (1992) 'In situ optical histochemistry of human artery using near infrared Fourier transform Raman spectroscopy.', *Proceedings of the National Academy of Sciences*, 89(8), pp. 3473–3477. doi: 10.1073/pnas.89.8.3473.

Baraga, J. J., Feld, M. S. and Rava, R. P. (1992) 'Rapid Near-Infrared Raman Spectroscopy of Human Tissue with a Spectrograph and CCD Detector', *Applied Spectroscopy*, 46(2), pp. 187–190. doi: 10.1366/0003702924125555.

Barber, S. C. and Shaw, P. J. (2010) 'Oxidative stress in ALS: Key role in motor neuron injury and therapeutic target', *Free Radical Biology and Medicine*. Elsevier Inc., 48(5), pp. 629–641. doi: 10.1016/j.freeradbiomed.2009.11.018.

Bareja, A. *et al.* (2014) 'Human and Mouse Skeletal Muscle Stem Cells: Convergent and Divergent Mechanisms of Myogenesis', *PLoS ONE*. Edited by S. E. Alway, 9(2), p. e90398. doi: 10.1371/journal.pone.0090398.

Barnes, R. J., Dhanoa, M. S. and Lister, S. J. (1989) 'Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra', *Applied Spectroscopy*, 43(5), pp. 772–777. doi: 10.1366/0003702894202201.

Bäumer, D., Talbot, K. and Turner, M. R. (2014) 'Advances in motor neurone disease', *Journal of the Royal Society of Medicine*, 107(1), pp. 14–21. doi: 10.1177/0141076813511451.

Beilharz, M. W. *et al.* (1992) 'Quantitation of Muscle Precursor Cell Activity in Skeletal Muscle by Northern Analysis of MyoD and Myogenin Expression : Application to Dystrophic (mdx) Mouse Muscle', 331, pp. 326–331.

Benatar, M. (2007) 'Lost in translation: Treatment trials in the SOD1 mouse and in human ALS', *Neurobiology of Disease*, 26(1), pp. 1–13. doi: 10.1016/j.nbd.2006.12.015.

Benatar, M. *et al.* (2016) 'ALS biomarkers for therapy development: State of the field and future directions', *Muscle & Nerve*, 53(2), pp. 169–182. doi: 10.1002/mus.24979.

Bennett, E. J. *et al.* (2014) 'Early detection of motor dysfunction in the SOD1G93A mouse model of amyotrophic lateral sclerosis (ALS) using home cage running wheels', *PLoS ONE*, 9(9). doi: 10.1371/journal.pone.0107918.

Bensimon, G., Lacomblez, L. and Meininger, V. (1994) 'A Controlled Trial of Riluzole in Amyotrophic Lateral Sclerosis', *New England Journal of Medicine*, 330(9), pp. 585–591. doi: 10.1056/NEJM199403033300901.

Bernardes, D., Leite, A. and Oliveira, R. (2017) 'Comprehensive catwalk gait analysis in a chronic model of multiple sclerosis subjected to treadmill exercise training'. BMC Neurology, pp. 1–14. doi: 10.1186/s12883-017-0941-z.

Biewener, A. A. (1990) 'Biomechanics of mammalian terrestrial locomotion.', *Science (New York, N.Y.)*, 250(4984), pp. 1097–103. doi: 10.1126/science.2251499.

Biewener, A. A. (2005) 'Biomechanical consequences of scaling', *Journal of Experimental Biology*, 208(9), pp. 1665–1676. doi: 10.1242/jeb.01520.

Biggar, W. D. (2006) 'Duchenne Muscular Dystrophy', *Pediatrics in Review*, 27(3), pp. 83–88. doi: 10.1542/pir.27-3-83.

Bilsland, L. G. *et al.* (2010) 'Deficits in axonal transport precede ALS symptoms in vivo.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(47), pp. 20523–8. doi: 10.1073/pnas.1006869107.

Bloemberg, D. and Quadrilatero, J. (2012) 'Rapid determination of myosin heavy chain expression in rat, mouse, and human skeletal muscle using multicolor immunofluorescence analysis.', *PloS one*, 7(4), p. e35273. doi: 10.1371/journal.pone.0035273.

Boillée, S., Vande Velde, C. and Cleveland, D. W. (2006) 'ALS: a disease of motor neurons and their nonneuronal neighbors.', *Neuron*, 52(1), pp. 39–59. doi: 10.1016/j.neuron.2006.09.018.

Bonilla, E., Tanji, K. and Minetti, C. (1999) 'Muscle Pathonlogy', in Younger, S. D. (ed.) *Motor Disorders*. Philadelphia.

Bonnier, F. and Byrne, H. J. (2012) 'Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems', *The Analyst*, 137(2), pp. 322–32. doi: 10.1039/C1AN15821J.

Boustany, N. N. *et al.* (1999) 'Analysis of nucleotides and aromatic amino acids in normal and neoplastic colon mucosa by ultraviolet resonance raman spectroscopy.', *Laboratory investigation; a journal of technical methods and pathology*, 79(10), pp. 1201–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10532584.

Bozzo, F., Mirra, A. and Carrì, M. T. (2017) 'Oxidative stress and mitochondrial

damage in the pathogenesis of ALS: New perspectives', *Neuroscience Letters*. Elsevier Ireland Ltd, 636, pp. 3–8. doi: 10.1016/j.neulet.2016.04.065.

Brauchle, E. and Schenke-Layland, K. (2013) 'Raman spectroscopy in biomedicine - non-invasive in vitro analysis of cells and extracellular matrix components in tissues', *Biotechnology Journal*, 8(3), pp. 288–297. doi: 10.1002/biot.201200163.

Brereton, R. G. and Lloyd, G. R. (2014) 'Partial least squares discriminant analysis: Taking the magic away', *Journal of Chemometrics*, 28(4), pp. 213–225. doi: 10.1002/cem.2609.

Brooks, B. R. *et al.* (2000) 'El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis', *Amyotrophic Lateral Sclerosis*, 1(5), pp. 293–299. doi: 10.1080/146608200300079536.

Brooks, S. V, Zerba, E. and Faulkner, J. A. (1995) 'Injury to muscle fibres after single stretches of passive and maximally stimulated muscles in mice.', *The Journal of physiology*, 488 (Pt 2, pp. 459–69. doi: 10.1113/jphysiol.1995.sp020980.

Browne, E. C. and Abbott, B. M. (2016) 'Recent progress towards an effective treatment of amyotrophic lateral sclerosis using the SOD1 mouse model in a preclinical setting', *European Journal of Medicinal Chemistry*, 121, pp. 918–925. doi: 10.1016/j.ejmech.2016.02.048.

Brunelli, S. *et al.* (2007) 'Nitric oxide release combined with nonsteroidal antiinflammatory activity prevents muscular dystrophy pathology and enhances stem cell therapy', 104(1).

Buchwald, T. *et al.* (2012) 'Identifying compositional and structural changes in spongy and subchondral bone from the hip joints of patients with osteoarthritis using Raman spectroscopy', *Journal of Biomedical Optics*, 17(1), p. 17007. doi: 10.1117/1.JBO.17.1.017007.

Bulfield, G. *et al.* (1984a) 'X chromosome-linked muscular dystrophy (mdx) in the mouse.', *Proceedings of the National Academy of Sciences of the United States of America*, 81(4), pp. 1189–92. doi: 10.1073/pnas.81.4.1189.

Bulfield, G. *et al.* (1984b) 'X chromosome-linked muscular dystrophy (mdx) in the mouse.', *Proceedings of the National Academy of Sciences of the United States of America*, 81(4), pp. 1189–92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6583703.

Burgess, R. W., Cox, G. A. and Seburn, K. L. (2016) 'Neuromuscular Disease Models and Analysis.', *Methods in molecular biology (Clifton, N.J.)*, 1438, pp. 349–94. doi: 10.1007/978-1-4939-3661-8_19.

Butler, H. J. *et al.* (2016) 'Using Raman spectroscopy to characterise biological materials', *Nature Protocols*. Nature Publishing Group, 11(4), pp. 664–687. doi: 10.1038/nprot.2016.036.

Byrne, H. J. *et al.* (2015) 'Spectropathology for the next generation: Quo vadis?', *The Analyst*, 140(7), pp. 2066–2073. doi: 10.1039/C4AN02036G.

Byrne, H. J. *et al.* (2016) 'Spectral pre and post processing for infrared and Raman spectroscopy of biological tissues and cells', *Chemical Society Reviews*. Royal Society of Chemistry, 45(7), pp. 1865–1878. doi: 10.1039/c5cs00440c.

Caballero-Garrido, E. *et al.* (2017) 'Characterization of long-term gait deficits in mouse dMCAO, using the CatWalk system', *Behavioural Brain Research*. Elsevier, 331(November 2016), pp. 282–296. doi: 10.1016/j.bbr.2017.05.042.

Cacchiarelli, D. *et al.* (2010) 'MicroRNAs Involved in Molecular Circuitries Relevant for the Duchenne Muscular Dystrophy Pathogenesis Are Controlled by the Dystrophin/nNOS Pathway', *Cell Metabolism*, 12(4), pp. 341–351. doi: 10.1016/j.cmet.2010.07.008.

Caiozzo, V. J. *et al.* (2003) 'Single-fiber myosin heavy chain polymorphism: how many patterns and what proportions?', *American journal of physiology. Regulatory, integrative and comparative physiology*, 285(3), pp. R570-80. doi: 10.1152/ajpregu.00646.2002.

Capitanio, D. *et al.* (2012) 'Molecular Signatures of Amyotrophic Lateral Sclerosis Disease Progression in Hind and Forelimb Muscles of an SOD1 ^{G93A} Mouse Model', *Antioxidants & Redox Signaling*, 17(10), pp. 1333–1350. doi: 10.1089/ars.2012.4524.

Cárcamo, J. J. *et al.* (2012) 'Raman and surface-enhanced Raman scattering in the study of human rotator cuff tissues after shock wave treatment', *Journal of Raman Spectroscopy*, 43(2), pp. 248–254. doi: 10.1002/jrs.3019.

Cardy, C. M. and Potter, T. (2007) 'The predictive value of creatine kinase, EMG and MRI in diagnosing muscle disease', *Rheumatology*, 46(10), pp. 1617–1618. doi: 10.1093/rheumatology/kem211.

Carew, E., Asher, I. and Stanley, H. (1975) 'Laser raman spectroscopy--new probe of myosin substructure', *Science*, 188(4191), pp. 933–936. doi: 10.1126/science.1138362.

Carlson, G. (2014) 'Serum Creatine Kinase analysis in mouse models of muscular', (Id), pp. 2–7.

Carr, S. J. *et al.* (2017) *Mass spectrometry-based protein analysis to unravel the tissue pathophysiology in Duchenne muscular dystrophy*, *Proteomics - Clinical Applications*. doi: 10.1002/prca.201700071.

Carter, G. *et al.* (1995) 'Effect of voluntary wheel-running exercise on muscles of the mdx mouse', *Neuromuscular Disordersrsers*, 5(4), pp. 323–332. doi: 10.1016/0960-8966(94)00063-F.

de Carvalho, M. *et al.* (2008) 'Electrodiagnostic criteria for diagnosis of ALS', *Clinical*

Neurophysiology, 119(3), pp. 497–503. doi: 10.1016/j.clinph.2007.09.143.

Cenci, M. A. and Lundblad, M. (2005) 'Utility of 6-hydroxydopamine lesioned rats in the preclinical screening of novel treatments for parkinson disease', *Movement Disorders*, pp. 193–208. doi: 10.1016/B978-012088382-0/50016-5.

Chan, J. W. *et al.* (2006) 'Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells.', *Biophysical journal*, 90(2), pp. 648–56. doi: 10.1529/biophysj.105.066761.

Chang, Y. *et al.* (2008) 'Messenger RNA Oxidation Occurs Early in Disease Pathogenesis and Promotes Motor Neuron Degeneration in ALS', *PLoS ONE*. Edited by M. R. Cookson, 3(8), p. e2849. doi: 10.1371/journal.pone.0002849.

Charles, J. P., Cappellari, O. and Hutchinson, J. R. (2018) 'A dynamic simulation of musculoskeletal function in the mouse hindlimb during trotting locomotion', *Frontiers in Bioengineering and Biotechnology*, 6(MAY), pp. 1–19. doi: 10.3389/fbioe.2018.00061.

Chen, J. *et al.* (2001) 'Intravenous Administration of Human Umbilical Cord Blood Reduces Behavioral Deficits After Stroke in Rats', *Stroke*, 32(11), pp. 2682–2688. doi: 10.1161/hs1101.098367.

Chen, X. *et al.* (2020) 'Microendoscopy detects altered muscular contractile dynamics in a mouse model of amyotrophic lateral sclerosis', *Scientific Reports*. Springer US, 10(1), pp. 1–10. doi: 10.1038/s41598-019-56555-z.

Chen, Y. *et al.* (2014) 'Raman spectroscopy analysis of the biochemical characteristics of molecules associated with the malignant transformation of gastric mucosa.', *PloS one*, 9(4), p. e93906. doi: 10.1371/journal.pone.0093906.

Cheng, W.-T. *et al.* (2005) 'Micro-Raman spectroscopy used to identify and grade human skin pilomatrixoma.', *Microscopy research and technique*, 68(2), pp. 75–9. doi: 10.1002/jemt.20229.

Chiang, H. K. *et al.* (2009) 'In situ Raman spectroscopic monitoring of hydroxyapatite as human mesenchymal stem cells differentiate into osteoblasts', *Journal of Raman Spectroscopy*, 40(5), pp. 546–549. doi: 10.1002/jrs.2161.

Chowdary, M. V. P. *et al.* (2009) 'Biochemical correlation of Raman spectra of normal, benign and malignant breast tissues: a spectral deconvolution study.', *Biopolymers*, 91(7), pp. 539–46. doi: 10.1002/bip.21171.

Chung, M. J. and Suh, Y.-L. (2002) 'Ultrastructural Changes of Mitochondria in the Skeletal Muscle of Patients with Amyotrophic Lateral Sclerosis', *Ultrastructural Pathology*, 26(1), pp. 3–7. doi: 10.1080/01913120252934260.

Clarke, R. H. *et al.* (1987) 'Laser Raman spectroscopy of calcified atherosclerotic lesions in cardiovascular tissue', *Applied Optics*, 26(16), p. 3175. doi:

10.1364/AO.26.003175.

Cohn, E. J. *et al.* (1934) 'Studies in the physical chemistry of amino acids, peptides and related substances. I. The apparent molal volume and the electrostriction of the solvent', *J. Am. Chem. Soc.*, 56, p. 784.

Conrad, A. O. and Bonello, P. (2016) 'Application of Infrared and Raman Spectroscopy for the Identification of Disease Resistant Trees', *Frontiers in Plant Science*, 6(January), pp. 1–8. doi: 10.3389/fpls.2015.01152.

Conroy, J. *et al.* (2005) 'Qualitative and quantitative analysis of chlorinated solvents using Raman spectroscopy and machine learning', *Proceedings of SPIE - The International Society for Optical Engineering*, 5826, pp. 131–142. doi: 10.1117/12.605056.

Consolino, C. M. and Brooks, S. V (2004) 'Susceptibility to sarcomere injury induced by single stretches of maximally activated muscles of mdx mice.', *Journal of applied physiology (Bethesda, Md. : 1985)*, 96(2), pp. 633–8. doi: 10.1152/japplphysiol.00587.2003.

Constantin, B. (2014) 'Dystrophin complex functions as a scaffold for signalling proteins', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(2), pp. 635–642. doi: 10.1016/j.bbamem.2013.08.023.

Contreras-Muñoz, P. *et al.* (2016) 'A New Surgical Model of Skeletal Muscle Injuries in Rats Reproduces Human Sports Lesions', *International Journal of Sports Medicine*, 37(3), pp. 183–190. doi: 10.1055/s-0035-1555933.

Cookson, M. R. *et al.* (2002) 'Cu/Zn superoxide dismutase (SOD1) mutations associated with familial amyotrophic lateral sclerosis (ALS) affect cellular free radical release in the presence of oxidative stress', *Amyotrophic Lateral Sclerosis*, 3(2), pp. 75–85. doi: 10.1080/146608202760196048.

Cooper-Knock, J., Jenkins, T. and Shaw, P. J. (2013) *Clinical and molecular aspects of motor neuron disease*, *Colloquium Series on …*. Available at: http://www.morganclaypool.com/doi/abs/10.4199/C00093ED1V01Y201309GMM004%5Cnpapers3://publication/doi/10.4199/C00093ED1V01Y201309GMM004.

Copeland, R. A. and Spiro, T. G. (1985) 'Ultraviolet resonance Raman spectra of cytochrome c conformational states', *Biochemistry*, 24(18), pp. 4960–4968. doi: 10.1021/bi00339a035.

Cossu, G. and Sampaolesi, M. (2007) 'New therapies for Duchenne muscular dystrophy: challenges, prospects and clinical trials', *Trends in Molecular Medicine*, 13(12), pp. 520–526. doi: 10.1016/j.molmed.2007.10.003.

Cothlup, N. B. and Daly, L. H. (1975) *Introduction to Infrared and Raman Spectroscopy*. First Edit. London: Academic Press.

Coulton, G. R. *et al.* (1988) 'The mdx mouse skeletal muscle myopathy: II. Contractile properties.', *Neuropathology and applied neurobiology*, 14(4), pp. 299–314. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3221977.

Couratier, P. *et al.* (2016) 'Epidemiology of amyotrophic lateral sclerosis: A review of literature', *Revue Neurologique*. Elsevier Masson SAS, 172(1), pp. 37–45. doi: 10.1016/j.neurol.2015.11.002.

Cozzolino, M. and Carrì, M. T. (2012) 'Mitochondrial dysfunction in ALS', *Progress in Neurobiology*, 97(2), pp. 54–66. doi: 10.1016/j.pneurobio.2011.06.003.

Crowley, S. T., Kataoka, K. and Itaka, K. (2018) 'Combined CatWalk Index : an improved method to measure mouse motor function using the automated gait analysis system', *BMC Research Notes*. BioMed Central, pp. 1–8. doi: 10.1186/s13104-018-3374-x.

Da Cruz, S. *et al.* (2012) 'Elevated PGC-1α Activity Sustains Mitochondrial Biogenesis and Muscle Function without Extending Survival in a Mouse Model of Inherited ALS', *Cell Metabolism*, 15(5), pp. 778–786. doi: 10.1016/j.cmet.2012.03.019.

Cyrulnik, S. E. *et al.* (2007) 'Delayed Developmental Language Milestones in Children with Duchenne's Muscular Dystrophy', *The Journal of Pediatrics*, 150(5), pp. 474–478. doi: 10.1016/j.jpeds.2006.12.045.

D'Amico, E. *et al.* (2013) 'Clinical perspective on oxidative stress in sporadic amyotrophic lateral sclerosis', *Free Radical Biology and Medicine*, 65, pp. 509–527. doi: 10.1016/j.freeradbiomed.2013.06.029.

Dahlbom, K., Lindberg, C. and Oldfors, A. (2002) 'Inclusion body myositis: morphological clues to correct diagnosis.', *Neuromuscular disorders : NMD*, 12(9), pp. 853–7. doi: 10.1016/s0960-8966(02)00098-6.

Das, K. *et al.* (2006) 'Raman spectroscopy of parathyroid tissue pathology', *Lasers in Medical Science*, 21(4), pp. 192–197. doi: 10.1007/s10103-006-0397-7.

Day, J. C. C. and Stone, N. (2013) 'A subcutaneous Raman needle probe', *Applied Spectroscopy*, 67(3), pp. 349–354. doi: 10.1366/12-06651.

Deconinck, N. and Dan, B. (2007) 'Pathophysiology of Duchenne Muscular Dystrophy: Current Hypotheses', *Pediatric Neurology*, 36(1), pp. 1–7. doi: 10.1016/j.pediatrneurol.2006.09.016.

Delhaye, M. and Dhamelincourt, P. (1975) 'Raman Microprobe and Microscope with Laser Excitation', *Journal of Raman Spectroscopy*, 3.

Deng, J. L. *et al.* (2005) 'Study of the effect of alcohol on single human red blood cells using near-infrared laser tweezers Raman spectroscopy', *Journal of Raman Spectroscopy*, 36(3), pp. 257–261. doi: 10.1002/jrs.1301.

Desroches, J. *et al.* (2015) 'Characterization of a Raman spectroscopy probe system

for intraoperative brain tissue classification.', *Biomedical optics express*, 6(7), pp. 2380–97. doi: 10.1364/BOE.6.002380.

Deumens, R. *et al.* (2007) 'The CatWalk gait analysis in assessment of both dynamic and static gait changes after adult rat sciatic nerve resection', *Journal of Neuroscience Methods*, 164(1), pp. 120–130. doi: 10.1016/j.jneumeth.2007.04.009.

Diem, M. (2015) 'Vibrational Microspectroscopy (MSP)', in *Modern Vibrational Spectroscopy and Micro-Spectroscopy*. First Edit. Chichester, UK: John Wiley & Sons, Ltd, pp. 235–250. doi: 10.1002/9781118824924.ch11.

Diem, M., Griffiths, P. R. and Chalmers, J. M. (2008) *Vibrational Spectroscopy for Medical Diagnosis*. London: John Wiley & Sons, Ltd.

Dobrowolny, G. *et al.* (2008) 'Skeletal Muscle Is a Primary Target of SOD1G93A-Mediated Toxicity', *Cell Metabolism*, 8(5), pp. 425–436. doi: 10.1016/j.cmet.2008.09.002.

Dunham, N. W. and Miya, T. S. (1957) 'A note on a simple apparatus for detecting neurological deficit in rats and mice.', *Journal of the American Pharmaceutical Association. American Pharmaceutical Association*, 46(3), pp. 208–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/13502156.

Dunn, J. F. and Radda, G. K. (1991) 'Total ion content of skeletal and cardiac muscle in the mdx mouse dystrophy: Ca2+ is elevated at all ages.', *Journal of the neurological sciences*, 103(2), pp. 226–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1880541.

Dunn, J. F. and Zaim-Wadghiri, Y. (1999) 'Quantitative magnetic resonance imaging of the mdx mouse model of Duchenne muscular dystrophy.', *Muscle & nerve*, 22(10), pp. 1367–71. doi: 10.1002/(sici)1097-4598(199910)22:10<1367::aid-mus5>3.0.co;2-h.

Dupuis, L. and Loeffler, J.-P. (2009) 'Neuromuscular junction destruction during amyotrophic lateral sclerosis: insights from transgenic models', *Current Opinion in Pharmacology*, 9(3), pp. 341–346. doi: 10.1016/j.coph.2009.03.007.

Duraipandian, S. (2012) 'Real-time Raman spectroscopy for <italic>in vivo</italic>, online gastric cancer diagnosis during clinical endoscopic examination', *Journal of Biomedical Optics*, 17(8), p. 81418. doi: 10.1117/1.jbo.17.8.081418.

Duraipandian, S. *et al.* (2012) 'Simultaneous Fingerprint and High-Wavenumber Confocal Raman Spectroscopy Enhances Early Detection of Cervical Precancer In Vivo', *Analytical Chemistry*, 84(14), pp. 5913–5919. doi: 10.1021/ac300394f.

Eberhardt, K. *et al.* (2018) 'Raman and infrared spectroscopy reveal that proliferating and quiescent human fibroblast cells age by biochemically similar but not identical processes', *PLoS ONE*, 13(12), pp. 1–22. doi: 10.1371/journal.pone.0207380.

Emery, A. E. (2002) 'The muscular dystrophies', *The Lancet*, 359(9307), pp. 687–695. doi: 10.1016/S0140-6736(02)07815-7.

Emery, A. E. ., Muntoni, F. and Quinlivan, R. (2015) *Duchenne muscular dystrophy*. 4th edn. Oxford: Oxford University Press.

Eng, C. M. *et al.* (2008) 'Scaling of muscle architecture and fiber types in the rat hindlimb.', *The Journal of experimental biology*, 211(Pt 14), pp. 2336–45. doi: 10.1242/jeb.017640.

Ennion, S. *et al.* (1995) 'Characterization of human skeletal muscle fibres according to the myosin heavy chains they express', *Journal of Muscle Research and Cell Motility*, 16(1), pp. 35–43. doi: 10.1007/BF00125308.

Erckens, R. J. *et al.* (1997) 'Raman Spectroscopy for Non-Invasive Characterization of Ocular Tissue: Potential for Detection of Biological Molecules', *Journal of Raman Spectroscopy*, 28(5), pp. 293–299. doi: 10.1002/(SICI)1097-4555(199705)28:5<293::AID-JRS47>3.0.CO;2-0.

Esmonde-White, K. a *et al.* (2011) 'Fiber-optic Raman spectroscopy of joint tissues.', *The Analyst*, 136(8), pp. 1675–85. doi: 10.1039/c0an00824a.

Feld, M. S. *et al.* (1995) 'Detection and characterization of human tissue lesions with near infrared Raman Spectroscopy', in Lakowicz, J. R. (ed.), pp. 99–104. doi: 10.1117/12.208468.

Fendel, S. and Schrader, B. (1998) 'Investigation of skin and skin lesions by NIR-FT-Raman spectroscopy', *Fresenius' Journal of Analytical Chemistry*, 360(5), pp. 609–613. doi: 10.1007/s002160050767.

Ferraiuolo, L. *et al.* (2011) 'Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis', *Nature Reviews Neurology*. Nature Publishing Group, 7(11), pp. 616–630. doi: 10.1038/nrneurol.2011.152.

Ferraro, J. R., Nakamoto, K. and Brown, C. W. (2002) *Introductory Raman Spectroscopy*. Second Edi. London: Academic Press.

Flanigan, K. M. (2014) 'Duchenne and becker muscular dystrophies', *Neurologic Clinics*. Elsevier Inc, 32(3), pp. 671–688. doi: 10.1016/j.ncl.2014.05.002.

Fong, P. Y. *et al.* (1990) 'Increased activity of calcium leak channels in myotubes of Duchenne human and mdx mouse origin.', *Science (New York, N.Y.)*, 250(4981), pp. 673–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2173137.

Frame, L. *et al.* (2018) 'Development of a label-free Raman imaging technique for differentiation of malaria parasite infected from non-infected tissue', *Analyst*. Royal Society of Chemistry, 143(1), pp. 157–163. doi: 10.1039/c7an01760j.

Frank, C. J. *et al.* (1994) 'Characterization of human breast biopsy specimens with near-IR Raman spectroscopy.', *Analytical chemistry*, 66(3), pp. 319–26. doi:

10.1021/ac00075a002.

Frank, C. J., McCreery, R. L. and Redd, D. C. B. (1995) 'Raman spectroscopy of normal and diseased human breast tissues.', *Analytical chemistry*, 67(5), pp. 777–83. doi: 10.1021/ac00101a001.

Frushour, B. G. and Koenig, J. L. (1974) 'Raman spectroscopic study of tropomyosin denaturation.', *Biopolymers*, 13(9), pp. 1809–19. doi: 10.1002/bip.1974.360130913.

Fuglsang-Frederiksen, A. (2006) 'The role of different EMG methods in evaluating myopathy', *Clinical Neurophysiology*, 117(6), pp. 1173–1189. doi: 10.1016/j.clinph.2005.12.018.

Gabriel, A. F. *et al.* (2007) 'The CatWalk method: A detailed analysis of behavioral changes after acute inflammatory pain in the rat', *Journal of Neuroscience Methods*, 163(1), pp. 9–16. doi: 10.1016/j.jneumeth.2007.02.003.

Gabriel, A. F. *et al.* (2009) 'The CatWalk method : Assessment of mechanical allodynia in experimental chronic pain', 198(2006), pp. 477–480. doi: 10.1016/j.bbr.2008.12.018.

Gasior-Głogowska, M. *et al.* (2010) 'Structural alteration of collagen fibres - spectroscopic and mechanical studies', *Acta of Bioengineering and Biomechanics*, 12(4), pp. 53–60.

Gautam, R. *et al.* (2015) 'Raman spectroscopic studies on screening of myopathies.', *Analytical chemistry*, 87(4), pp. 2187–94. doi: 10.1021/ac503647x.

Ge, Y. *et al.* (2004) 'Differential expression of the skeletal muscle proteome in mdx mice at different ages', *Electrophoresis*, 25(15), pp. 2576–2585. doi: 10.1002/elps.200406013.

De Gelder, J. *et al.* (2007) 'Reference database of Raman spectra of biological molecules', *Journal of Raman Spectroscopy*, 38(9), pp. 1133–1147. doi: 10.1002/jrs.1734.

Germond, A. *et al.* (2018) 'Raman spectral signature reflects transcriptomic features of antibiotic resistance in Escherichia coli', *Communications Biology*. Springer US, 1(1). doi: 10.1038/s42003-018-0093-8.

Gilson, T. . and Hendra, P. . (1970) *Laser Raman Spectroscopy*. First Edit. London: Wiley Interscience.

Gniadecka, M. *et al.* (1997) 'Diagnosis of basal cell carcinoma by Raman spectroscopy', *Journal of Raman Spectroscopy*, 28(2–3), pp. 125–129. doi: 10.1002/(SICI)1097-4555(199702)28:2/310.1002/(SICI)1097-4555(199702)28:2/3<125::AID-JRS65>3.0.CO.

Granchelli, J. A. *et al.* (2013) 'The mdx mouse model as a surrogate for Duchenne muscular dystrophy', *FEBS Journal*, 280(17), pp. 4177–4186. doi:

10.1111/febs.12267.The.

Granchelli, J. A., Pollina, C. and Hudecki, M. S. (2000) 'Pre-clinical screening of drugs using the mdx mouse', *Neuromuscular Disorders*, 10(4–5), pp. 235–239. doi: 10.1016/S0960-8966(99)00126-1.

Grounds, M. (2014) 'Quantification of histopathology in Haemotoxylin and Eosin stained muscle sections Official reviewer', (Id), pp. 1–13.

Grounds, M. D. *et al.* (2008) 'Towards developing standard operating procedures for pre-clinical testing in the mdx mouse model of Duchenne muscular dystrophy', *Neurobiology of Disease*, 31(1), pp. 1–19. doi: 10.1016/j.nbd.2008.03.008.

Grounds, M. D. and Torrisi, J. O. (2004) 'Anti-TNFa (Remicade) therapy protects dystrophic skeletal muscle from necrosis.', *The FASEB Journal*, 18, pp. 676–682. doi: 10.1096/fj.03-1024com.

Guiraud, S. *et al.* (2015) 'The Pathogenesis and Therapy of Muscular Dystrophies', *Annual Review of Genomics and Human Genetics*, 16(1), pp. 281–308. doi: 10.1146/annurev-genom-090314-025003.

Gupta, P. K. *et al.* (2012) 'A predictive model for amyotrophic lateral sclerosis (ALS) diagnosis', *Journal of the Neurological Sciences*. Elsevier B.V., 312(1–2), pp. 68–72. doi: 10.1016/j.jns.2011.08.021.

Gurney, M. E. *et al.* (1994) 'Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation.', *Science (New York, N.Y.)*, 264(5166), pp. 1772–5. doi: 10.1126/science.8209258.

Gutpell, K. M., Hrinivich, W. T. and Hoffman, L. M. (2015) 'Skeletal Muscle Fibrosis in the mdx / utrn + / - Mouse Validates Its Suitability as a Murine Model of Duchenne Muscular Dystrophy', 3, pp. 1–13. doi: 10.1371/journal.pone.0117306.

Halter, B. *et al.* (2010) 'Oxidative stress in skeletal muscle stimulates early expression of Rad in a mouse model of amyotrophic lateral sclerosis', *Free Radical Biology and Medicine*, 48(7), pp. 915–923. doi: 10.1016/j.freeradbiomed.2010.01.014.

Hamm, R. J. *et al.* (1994) 'The rotarod test: an evaluation of its effectiveness in assessing motor deficits following traumatic brain injury.', *Journal of neurotrauma*, 11(2), pp. 187–96. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7932797%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/7932797.

Hanlon, E. B. *et al.* (2000) 'Prospects for in vivo Raman spectroscopy.', *Physics in medicine and biology*, 45(2), pp. R1-59. doi: 10.1088/0031-9155/45/2/201.

Harraz, M. M. *et al.* (2008) 'SOD1 mutations disrupt redox-sensitive Rac regulation of NADPH oxidase in a familial ALS model', *Journal of Clinical Investigation*. doi: 10.1172/JCI34060.

Hartman, K. A., Clayton, N. and Thomas, G. J. (1973) 'Studies of virus structure by Raman spectroscopy I. R17 virus and R17 RNA', *Biochemical and Biophysical Research Communications*, 50(3), pp. 942–949. doi: 10.1016/0006-291X(73)91336-3.

Hattori, Y. *et al.* (2007) 'In vivo Raman study of the living rat esophagus and stomach using a micro-Raman probe under an endoscope', *Applied Spectroscopy*, 61(6), pp. 579–584. doi: 10.1366/000370207781269747.

Hatzipetros, T. *et al.* (2015) 'A Quick Phenotypic Neurological Scoring System for Evaluating Disease Progression in the SOD1-G93A Mouse Model of ALS', *Journal of Visualized Experiments*, (104). doi: 10.3791/53257.

Hegedus, J., Putman, C. T. and Gordon, T. (2007) 'Time course of preferential motor unit loss in the SOD1G93A mouse model of amyotrophic lateral sclerosis', *Neurobiology of Disease*, 28(2), pp. 154–164. doi: 10.1016/j.nbd.2007.07.003.

Higgins, C. M. J. *et al.* (2002) 'Mutant Cu, Zn superoxide dismutase that causes motoneuron degeneration is present in mitochondria in the CNS.', *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(6), p. RC215. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11886899.

Hilber, K. *et al.* (1999) 'Kinetic properties of myosin heavy chain isoforms in single fibers from human skeletal muscle.', *FEBS letters*, 455(3), pp. 267–70. doi: 10.1016/s0014-5793(99)00903-5.

Hilton, J. B., White, A. R. and Crouch, P. J. (2015) 'Metal-deficient SOD1 in amyotrophic lateral sclerosis', *Journal of Molecular Medicine*, 93(5), pp. 481–487. doi: 10.1007/s00109-015-1273-3.

Hobro, A. J. *et al.* (2013) 'Raman spectroscopic analysis of malaria disease progression via blood and plasma samples', *The Analyst*, 138(14), p. 3927. doi: 10.1039/c3an00255a.

Hobson, E. V. *et al.* (2016) 'Clinical aspects of motor neurone disease', *Medicine*. Elsevier Ltd, 44(9), pp. 552–556. doi: 10.1016/j.mpmed.2016.06.004.

Hodgetts, S. *et al.* (2006) 'Reduced necrosis of dystrophic muscle by depletion of host neutrophils , or blocking TNF a function with Etanercept in mdx mice', 16, pp. 591–602. doi: 10.1016/j.nmd.2006.06.011.

Hsieh-Li, H. M. *et al.* (2000) 'A mouse model for spinal muscular atrophy.', *Nature genetics*, 24(1), pp. 66–70. doi: 10.1038/71709.

Hu, X. *et al.* (2017) 'Are mice good models for human neuromuscular disease? Comparing muscle excursions in walking between mice and humans', *Skeletal Muscle*. Skeletal Muscle, 7(1), p. 26. doi: 10.1186/s13395-017-0143-9.

Huang, N. *et al.* (2011) 'Full range characterization of the Raman spectra of organs in a murine model', *Optics Express*, 19(23), pp. 22892–909. doi: 10.1364/OE.19.022892.

Huang, Z. *et al.* (2001) 'Rapid near-infrared Raman spectroscopy system for real-time in vivo skin measurements.', *Optics letters*, 26(22), pp. 1782–1784. doi: 10.1364/OL.26.001782.

Huang, Z., McWilliams, A., Lam, S., *et al.* (2003) 'Effect of formalin fixation on the near-infrared Raman spectroscopy of normal and cancerous human bronchial tissues', *International Journal of Oncology*, pp. 649–655. doi: 10.3892/ijo.23.3.649.

Huang, Z., McWilliams, A., Lui, H., *et al.* (2003) 'Near-infrared Raman spectroscopy for optical diagnosis of lung cancer.', *International journal of cancer*, 107(6), pp. 1047–52. doi: 10.1002/ijc.11500.

Huang, Z., Teh, S. K., *et al.* (2010) 'In vivo detection of epithelial neoplasia in the stomach using image-guided Raman endoscopy', *Biosensors and Bioelectronics*. Elsevier B.V., 26(2), pp. 383–389. doi: 10.1016/j.bios.2010.07.125.

Huang, Z., Bergholt, M. S., *et al.* (2010) 'In vivo early diagnosis of gastric dysplasia using narrow-band image-guided Raman endoscopy.', *Journal of biomedical optics*, 15(3), p. 37017. doi: 10.1117/1.3420115.

Hutchings, J. *et al.* (2009) 'The potential for histological screening using a combination of rapid Raman mapping and principal component analysis', *Journal of Biophotonics*, 2(1–2), pp. 91–103. doi: 10.1002/jbio.200810070.

Ito, M. *et al.* (2017) 'Protein-anchoring therapy of biglycan for mdx mouse model of duchenne muscular dystrophy', *Human Gene Therapy*, 28(5), pp. 428–436. doi: 10.1089/hum.2015.088.

Jenkins, T. M. *et al.* (2016) 'The role of cranial and thoracic electromyography within diagnostic criteria for amyotrophic lateral sclerosis.', *Muscle & nerve*, 54(3), pp. 378–85. doi: 10.1002/mus.25062.

Jermyn, M. *et al.* (2016) 'A review of Raman spectroscopy advances with an emphasis on clinical translation challenges in oncology', *Physics in Medicine & Biology*. IOP Publishing, 61, pp. R370–R400. doi: 10.1088/0031-9155/61/23/R370.

Jones, R. A. *et al.* (2017) 'Cellular and Molecular Anatomy of the Human Neuromuscular Junction', *Cell reports*, 21(9), pp. 2348–2356. doi: 10.1016/j.celrep.2017.11.008.

Jungbluth, H. *et al.* (2004) 'Magnetic resonance imaging of muscle in nemaline myopathy.', *Neuromuscular disorders : NMD*, 14(12), pp. 779–84. doi: 10.1016/j.nmd.2004.08.005.

Justice, M. J. and Dhillon, P. (2016) 'Using the mouse to model human disease: increasing validity and reproducibility', *Disease Models & Mechanisms*, 9(2), pp. 101–103. doi: 10.1242/dmm.024547.

Jyothi Lakshmi, R. *et al.* (2002) 'Tissue Raman spectroscopy for the study of radiation

damage: Brain irradiation of mice.', *Radiation research*, 157(2), pp. 175–82. doi: 10.1667/0033-7587(2002)157[0175:trsfts]2.0.co;2.

Kaczor, J. J. *et al.* (2007) 'Low intensity training decreases markers of oxidative stress in skeletal muscle of mdx mice', *Free Radical Biology and Medicine*, 43(1), pp. 145–154. doi: 10.1016/j.freeradbiomed.2007.04.003.

Kamemoto, L. E. *et al.* (2010) 'Near-Infrared Micro-Raman Spectroscopy for in Vitro Detection of Cervical Cancer', *Applied Spectroscopy*, 64(3), pp. 255–261. doi: 10.1366/000370210790918364.

Kaminaka, S. *et al.* (2001) 'Near-infrared Raman spectroscopy of human lung tissues: possibility of molecular-level cancer diagnosis', *Journal of Raman Spectroscopy*, 32(2), pp. 139–141. doi: 10.1002/jrs.680.

Kaur, S. J., McKeown, S. R. and Rashid, S. (2016) 'Mutant SOD1 mediated pathogenesis of Amyotrophic Lateral Sclerosis', *Gene*. Elsevier B.V., 577(2), pp. 109–118. doi: 10.1016/j.gene.2015.11.049.

Keller, S. *et al.* (1994) 'Application of near-infrared-Fourier transform Raman spectroscopy in medical research', *Journal of Raman Spectroscopy*, 25(7–8), pp. 663–671. doi: 10.1002/jrs.1250250729.

Kendall, C. (2002) *A Study of Raman Spectroscopy for the Early Detection and Classification of Malignacy in Oesophageal Tissue*. Cranfield University.

Kengne-Momo, R. P. *et al.* (2012) 'Protein Interactions Investigated by the Raman Spectroscopy for Biosensor Applications', *International Journal of Spectroscopy*, 2012(i), pp. 1–7. doi: 10.1155/2012/462901.

Kieran, D. *et al.* (2005) 'A mutation in dynein rescues axonal transport defects and extends the life span of ALS mice.', *The Journal of cell biology*, 169(4), pp. 561–7. doi: 10.1083/jcb.200501085.

Kiernan, M. C. *et al.* (2011) 'Amyotrophic lateral sclerosis', *The Lancet*. Elsevier Ltd, 377(9769), pp. 942–955. doi: 10.1016/S0140-6736(10)61156-7.

Kim, R. B. *et al.* (2016) 'State of the field: An informatics-based systematic review of the SOD1-G93A amyotrophic lateral sclerosis transgenic mouse model', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 17(1–2), pp. 1–14. doi: 10.3109/21678421.2015.1047455.

Kim, S., Byun, K. M. and Lee, S. Y. (2017) 'Influence of water content on Raman spectroscopy characterization of skin sample', *Biomedical Optics Express*, 8(2), p. 1130. doi: 10.1364/BOE.8.001130.

King, W. M. *et al.* (2007) 'Orthopedic outcomes of long-term daily corticosteroid treatment in Duchenne muscular dystrophy', *Neurology*, 68(19), pp. 1607–1613. doi: 10.1212/01.wnl.0000260974.41514.83.

Kishp, K. and Noda, H. (1983) ‘Laser Raman Studies on Myosin, C-Protein, and Myosin-C-Protein Complex1’, *The Journal of Biochemistry*, 94(2), pp. 353–359. doi: 10.1093/oxfordjournals.jbchem.a134363.

Koenig, J. L. and Sutton, P. L. (1969) ‘Raman spectrum of the right-handed α-helix of poly-L-alanine’, *Biopolymers*, 8(2), pp. 167–171. doi: 10.1002/bip.1969.360080203.

Koljenovic, S. *et al.* (2007) ‘Raman spectroscopic characterization of porcine brain tissue using a single fiber-optic probe’, *Anal. Chem*, 79(2), pp. 557–564. doi: 10.1021/ac0616512.

Koljenović, S. *et al.* (2005) ‘Detection of meningioma in dura mater by Raman spectroscopy’, *Analytical Chemistry*, 77(24), pp. 7958–7965. doi: 10.1021/ac0512599.

Koljenović, S. *et al.* (2005) ‘Tissue characterization using high wave number Raman spectroscopy’, *Journal of Biomedical Optics*, 10(3), p. 31116. doi: 10.1117/1.1922307.

Komachi, Y. *et al.* (2005) ‘Micro-optical fiber probe for use in an intravascular Raman endoscope.’, *Applied optics*, 44(22), pp. 4722–4732. doi: 10.1364/AO.44.004722.

Kong, K. *et al.* (2015) ‘Raman spectroscopy for medical diagnostics - From in-vitro biofluid assays to in-vivo cancer detection’, *Advanced Drug Delivery Reviews*. The Authors, 89, pp. 121–134. doi: 10.1016/j.addr.2015.03.009.

Kong Chong, C. *et al.* (1992) ‘Raman spectroscopy with a fiber-optic probe’, *Vibrational Spectroscopy*, 3(1), pp. 35–45. doi: 10.1016/0924-2031(92)85022-S.

Kraemer, M., Buerger, M. and Berlit, P. (2010) ‘Diagnostic problems and delay of diagnosis in amyotrophic lateral sclerosis’, *Clinical Neurology and Neurosurgery*, 112(2), pp. 103–105. doi: 10.1016/j.clineuro.2009.10.014.

Krafft, C. *et al.* (2005) ‘Near infrared Raman spectra of human brain lipids.’, *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, 61(7), pp. 1529–35. doi: 10.1016/j.saa.2004.11.017.

Krafft, C. *et al.* (2009) ‘Disease recognition by infrared and Raman spectroscopy’, *Journal of Biophotonics*, 2(1–2), pp. 13–28. doi: 10.1002/jbio.200810024.

Krafft, C., Dietzek, B. and Popp, J. (2009) ‘Raman and CARS microspectroscopy of cells and tissues’, *Analyst*, 134(6), pp. 1046–1057. doi: 10.1039/b822354h.

Krarup, C. (2011) ‘Lower motor neuron involvement examined by quantitative electromyography in amyotrophic lateral sclerosis’, *Clinical Neurophysiology*, 122(2), pp. 414–422. doi: 10.1016/j.clinph.2010.06.027.

Krouchev, N., Kalaska, J. F. and Drew, T. (2006) ‘Sequential activation of muscle synergies during locomotion in the intact cat as revealed by cluster analysis and direct decomposition’, *Journal of Neurophysiology*, 96(4), pp. 1991–2010. doi: 10.1152/jn.00241.2006.

Kumar, S. T. (2013) 'Physical and Chemical Characterization of Biomaterials', in *Characterization of Biomaterials*. Elsevier, pp. 11–47. doi: 10.1016/B978-0-12-415800-9.00002-4.

Lai, C.-H. *et al.* (2010) 'Open muscle biopsy in suspected myopathy: diagnostic yield and clinical utility', *European Journal of Neurology*, 17(1), pp. 136–142. doi: 10.1111/j.1468-1331.2009.02765.x.

Landisch, R. M., Kosir, A. M. and Nelson, S. A. (2008) 'ADAPTIVE AND NONADAPTIVE RESPONSES TO VOLUNTARY WHEEL RUNNING BY mdx MICE', (October), pp. 1290–1303. doi: 10.1002/mus.21141.

Lau, D. P. *et al.* (2003) 'Raman spectroscopy for optical diagnosis in normal and cancerous tissue of the nasopharynx - Preliminary findings', *Lasers in Surgery and Medicine*, 32(3), pp. 210–214. doi: 10.1002/lsm.10084.

Lau, D. P. *et al.* (2005) 'Raman spectroscopy for optical diagnosis in the larynx: Preliminary findings.', *Lasers in surgery and medicine*, 37(3), pp. 192–200. doi: 10.1002/lsm.20226.

Law, A. W. L. *et al.* (2017) 'In situ cellular level Raman spectroscopy of the thyroid', *Biomedical Optics Express*, 8(2), p. 670. doi: 10.1364/boe.8.000670.

Lee, L. C., Liong, C. Y. and Jemain, A. A. (2018) 'Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps', *Analyst*, 143(15), pp. 3526–3539. doi: 10.1039/c8an00599k.

Lieber, C. A. and Mahadevan-Jansen, A. (2003) 'Automated Method for Subtraction of Fluorescence from Biological Raman Spectra', *Applied Spectroscopy*, 57(11), pp. 1363–1367. doi: 10.1016/0019-1035(74)90189-4.

Liguori, R. *et al.* (1997) 'Electromyography in myopathy', *Neurophysiologie Clinique*, 27(3), pp. 200–203. doi: 10.1016/S0987-7053(97)83775-6.

Lin, D. *et al.* (2011) 'Investigation on the interactions of lymphoma cells with paclitaxel by Raman spectroscopy', *Spectroscopy*, 25(1), pp. 23–32. doi: 10.3233/SPE-2011-0497.

Lin, K. *et al.* (2016) 'Rapid fiber-optic raman spectroscopy for real-time in vivo detection of gastric intestinal metaplasia during clinical gastroscopy', *Cancer Prevention Research*, 9(6), pp. 476–483. doi: 10.1158/1940-6207.CAPR-15-0213.

Lindberg, M. J. *et al.* (2005) 'Systematically perturbed folding patterns of amyotrophic lateral sclerosis (ALS)-associated SOD1 mutants', *Proceedings of the National Academy of Sciences*, 102(28), pp. 9754–9759. doi: 10.1073/pnas.0501957102.

Lippert, J. L., Tyminski, D. and Desmeules, P. J. (1976) 'Determination of the

Secondary Structure of Proteins by Laser Raman Spectroscopy', *Journal of the American Chemical Society*, 98(22), pp. 7075–7080. doi: 10.1021/ja00438a057.

Liu, C.-H. *et al.* (1992) 'Raman, fluorescence, and time-resolved light scattering as optical diagnostic techniques to separate diseased and normal biomedical media', *Journal of Photochemistry and Photobiology B: Biology*, 16(2), pp. 187–209. doi: 10.1016/1011-1344(92)80008-J.

Liu, Y. *et al.* (2013) 'Quantitative gait analysis of long-term locomotion deficits in classical unilateral striatal intracerebral hemorrhage rat model', *Behavioural Brain Research*. Elsevier B.V., 257, pp. 166–177. doi: 10.1016/j.bbr.2013.10.007.

Lloyd, G. R. *et al.* (2012) 'Assessing the performance of spectroscopic models for cancer diagnostics using cross-validation and permutation testing', *Biomedical Vibrational Spectroscopy V: Advances in Research and Industry*, 8219, p. 82190C. doi: 10.1117/12.919864.

Lloyd, G. R. *et al.* (2013) 'Discrimination between benign, primary and secondary malignancies in lymph nodes from the head and neck utilising Raman spectroscopy and multivariate analysis', *The Analyst*, 138(14), p. 3900. doi: 10.1039/c2an36579k.

Loeffler, J.-P. *et al.* (2016) 'The Role of Skeletal Muscle in Amyotrophic Lateral Sclerosis', *Brain Pathology*, p. n/a–n/a. doi: 10.1111/bpa.12350.

Lord, R. . and Yu, N. (1970) 'Laser-excited Raman spectroscopy of biomolecules 50', *J. Mol. Biol.*, 50, pp. 509–24.

Löscher, W. N. and Feldman, E. L. (2014) 'Motor Neuron Diseases', in *Atlas of Neuromuscular Diseases- A practical Guideline*. Vienna: Springer, Vienna, pp. 283–290. doi: 10.1007/978-3-7091-1605-0.

Lovering, R. M., Mcmillan, A. B. and Gullapalli, R. P. (2009) 'Location of myofiber damage in skeletal muscle after lengthening contractions', *Muscle and Nerve*, 40(4), pp. 589–594. doi: 10.1002/mus.21389.

De Luca, A. *et al.* (2005) 'A multidisciplinary evaluation of the effectiveness of cyclosporine A in dystrophic Mdx mice', *American Journal of Pathology*, 166(2), pp. 477–489. doi: 10.1016/S0002-9440(10)62270-5.

De Luca, A. (2014) 'Use of treadmill and wheel exercise for impact on mdx mice phenotype', (Id), pp. 1–9.

Lui, H. *et al.* (2012) 'Real-time raman spectroscopy for in vivo skin cancer diagnosis', *Cancer Research*, 72(10), pp. 2491–2500. doi: 10.1158/0008-5472.CAN-11-4061.

Luo, G. *et al.* (2013) 'Defective mitochondrial dynamics is an early event in skeletal muscle of an Amyotrophic lateral sclerosis mouse model', *PLoS ONE*, 8(12), pp. 1–10. doi: 10.1371/journal.pone.0082112.

Lykina, A. A. *et al.* (2018) 'Raman spectroscopy for kidney tissue and its neoplasms

research', *Journal of Physics: Conference Series*, 1096(1), p. 12116. doi: 10.1088/1742-6596/1096/1/012116.

Mah, J. K. *et al.* (2014) 'A systematic review and meta-analysis on the epidemiology of Duchenne and Becker muscular dystrophy', *Neuromuscular Disorders*, 24(6), pp. 482–491. doi: 10.1016/j.nmd.2014.03.008.

Mahadevan-Jansen, A. (1996) 'Raman spectroscopy for the detection of cancers and precancers', *Journal of Biomedical Optics*, 1(1), p. 31. doi: 10.1117/12.227815.

Mahadevan-Jansen, A. *et al.* (1998) 'Near-infrared Raman spectroscopy for in vitro detection of cervical precancers.', *Photochemistry and photobiology*, 68(1), pp. 123–32. doi: 10.1562/0031-8655(1998)068<0123:nirsfv>2.3.co;2.

Mahadevan-Jansen, A. and Richards-Kortum, R. (1997) 'Raman spectroscopy for cancer detection: a review', in *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 2722–2728. doi: 10.1109/IEMBS.1997.756895.

Malini, R. *et al.* (2006) 'Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: a Raman spectroscopy study.', *Biopolymers*, 81(3), pp. 179–93. doi: 10.1002/bip.20398.

Mancuso, R. *et al.* (2011) 'Evolution of gait abnormalities in SOD1 G93A transgenic mice', *Brain Research*. Elsevier B.V., 1406, pp. 65–73. doi: 10.1016/j.brainres.2011.06.033.

Manning, J. and O'Malley, D. (2015) 'What has the mdx mouse model of duchenne muscular dystrophy contributed to our understanding of this disease?', *Journal of Muscle Research and Cell Motility*. Springer International Publishing, 36(2), pp. 155–167. doi: 10.1007/s10974-015-9406-4.

Manoharan, R. *et al.* (1992) 'Quantitative histochemical analysis of human artery using Raman spectroscopy', *Journal of Photochemistry and Photobiology B: Biology*, 16(2), pp. 211–233. doi: 10.1016/1011-1344(92)80009-K.

Manoharan, R. *et al.* (1998) 'Raman spectroscopy and fluorescence photon migration for breast cancer diagnosis and imaging.', *Photochemistry and photobiology*, 67(1), pp. 15–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9477761.

Maricelli, J. W. *et al.* (2016) 'Trendelenburg-Like Gait , Instability and Altered Step Patterns in a Mouse Model for Limb Girdle Muscular Dystrophy 2i', pp. 1–19. doi: 10.1371/journal.pone.0161984.

Marshall, P. A., Williams, P. E. and Goldspink, G. (1989) 'Accumulation of collagen and altered fiber-type ratios as indicators of abnormal muscle gene expression in the mdx dystrophic mouse', *Muscle & Nerve*, 12(7), pp. 528–537. doi: 10.1002/mus.880120703.

Marx, J. O., Olsson, M. C. and Larsson, L. (2006) 'Scaling of skeletal muscle shortening velocity in mammals representing a 100,000-fold difference in body size.', *Pflugers Archiv : European journal of physiology*, 452(2), pp. 222–30. doi: 10.1007/s00424-005-0017-6.

McDonald, A. A. *et al.* (2015) 'Disease course in mdx:utrophin+/- mice: comparison of three mouse models of Duchenne muscular dystrophy.', *Physiological reports*, 3(4), p. e12391. doi: 10.14814/phy2.12391.

McGreevy, J. W. *et al.* (2015) 'Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy', *Disease Models & Mechanisms*, 8(3), pp. 195–213. doi: 10.1242/dmm.018424.

McGregor, H. C. *et al.* (2016) 'Real-time endoscopic Raman spectroscopy for *in vivo* early lung cancer detection', *Journal of Biophotonics*, 13, p. n/a-n/a. doi: 10.1002/jbio.201500204.

Mead, R. J. *et al.* (2011) 'Optimised and rapid pre-clinical screening in the SOD1(G93A) transgenic mouse model of amyotrophic lateral sclerosis (ALS).', *PloS one*, 6(8), p. e23244. doi: 10.1371/journal.pone.0023244.

Mierzejewski, B. *et al.* (2020) 'Human and mouse skeletal muscle stem and progenitor cells in health and disease.', *Seminars in cell & developmental biology*. Elsevier, 104(January), pp. 93–104. doi: 10.1016/j.semcdb.2020.01.004.

Miller, R. G. *et al.* (2007) 'Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND)', in Miller, R. G. (ed.) *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/14651858.CD001447.pub2.

Miller, T. M. *et al.* (2006) 'Gene transfer demonstrates that muscle is not a primary target for non-cell-autonomous toxicity in familial amyotrophic lateral sclerosis.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(51), pp. 19546–51. doi: 10.1073/pnas.0609411103.

Min, Y. K. *et al.* (2005) '1064 nm near-infrared multichannel Raman spectroscopy of fresh human lung tissues', *Journal of Raman Spectroscopy*, 36(1), pp. 73–76. doi: 10.1002/jrs.1280.

Minamikawa, T., Harada, Y. and Takamatsu, T. (2015) 'Ex vivo peripheral nerve detection of rats by spontaneous Raman spectroscopy.', *Scientific reports*. Nature Publishing Group, 5(1), p. 17165. doi: 10.1038/srep17165.

Ming, L. C. *et al.* (2017) 'Real time near-infrared Raman spectroscopy for the diagnosis of nasopharyngeal cancer', *Oncotarget*, 8(30), pp. 49443–49450. doi: 10.18632/oncotarget.17703.

Miura, T. and Thomas, G. J. (1995) 'Raman Spectroscopy of Proteins and Their Assemblies', in Biswas, B. . and Roy, S. (eds) *Proteins: Structure, Function, and*

Engineering. Subcellular Biochemistry, vol 24. Boston, MA, pp. 55–99. doi: 10.1007/978-1-4899-1727-0_3.

Mizuno, A. *et al.* (1994) 'Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours', *Journal of Raman Spectroscopy*, 25(1), pp. 25–29. doi: 10.1002/jrs.1250250105.

Mo, J. *et al.* (2009) 'High wavenumber Raman spectroscopy for in vivo detection of cervical dysplasia', *Anal. Chem*, 81(21), pp. 8908–8915. doi: 10.1021/ac9015159.

Moloney, E. B., de Winter, F. and Verhaagen, J. (2014) 'ALS as a distal axonopathy: Molecular mechanisms affecting neuromuscular junction stability in the presymptomatic stages of the disease', *Frontiers in Neuroscience*, 8(8 JUL), pp. 1–18. doi: 10.3389/fnins.2014.00252.

Motz, J. T. *et al.* (2005) 'Real-time Raman system for in vivo disease diagnosis.', *Journal of biomedical optics*, 10(3), p. 31113. doi: 10.1117/1.1920247.

Munch, C., O'Brien, J. and Bertolotti, A. (2011) 'Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells', *Proceedings of the National Academy of Sciences*, 108(9), pp. 3548–3553. doi: 10.1073/pnas.1017275108.

Musarò, A. (2010) 'State of the art and the dark side of amyotrophic lateral sclerosis', *World Journal of Biological Chemistry*, 1(5), p. 62. doi: 10.4331/wjbc.v1.i5.62.

Nawaz, H. *et al.* (2011) 'Comparison of subcellular responses for the evaluation and prediction of the chemotherapeutic response to cisplatin in lung adenocarcinoma using Raman spectroscopy', *Analyst*, 136(12), pp. 2450–2463. doi: 10.1039/c1an15104e.

Nguyen, H. H. *et al.* (2002) 'Overexpression of the cytotoxic T cell GalNAc transferase in skeletal muscle inhibits muscular dystrophy in mdx mice', 99(8).

Nguyen, T. T. *et al.* (2012) 'Characterization of type I and IV collagens by Raman microspectroscopy: Identification of spectral markers of the dermo-epidermal junction', *Spectroscopy: An International Journal*, 27(5–6), pp. 421–427. doi: 10.1155/2012/686183.

Niebroj-Dobosz, I., Fidzianska, A. and Glinka, Z. (1997) 'Comparative studies of hind limb and diaphragm muscles of mdx mice', *Bam-Padova-*, 7(1), pp. 381–386.

Ninagawa, N. T. *et al.* (2013) 'Transplantated Mesenchymal Stem Cells Derived from Embryonic Stem Cells Promote Muscle Regeneration and Accelerate', 2(4). doi: 10.1089/biores.2013.0012.

Noothalapati, H. N., Iwasaki, K. I. and Yamamoto, T. Y. (2017) 'Biological and Medical Applications of Multivariate Curve Resolution Assisted Raman Spectroscopy', *Analytical Sciences*, 33(January).

Notingher, I. *et al.* (2004) 'In situ non-invasive spectral discrimination between bone

cell phenotypes used in tissue engineering', *Journal of Cellular Biochemistry*, 92(6), pp. 1180–1192. doi: 10.1002/jcb.20136.

Noto, Y.-I. *et al.* (2014) 'Contrasting echogenicity in flexor digitorum profundus-flexor carpi ulnaris: a diagnostic ultrasound pattern in sporadic inclusion body myositis.', *Muscle & nerve*, 49(5), pp. 745–8. doi: 10.1002/mus.24056.

Nowak, K. J. and Davies, K. E. (2004) 'Duchenne muscular dystrophy and dystrophin: pathogenesis and opportunities for treatment', *EMBO reports*, 5(9), pp. 872–876. doi: 10.1038/sj.embor.7400221.

Nzwalo, H. *et al.* (2014) 'Delayed diagnosis in ALS: The problem continues', *Journal of the Neurological Sciences*. Elsevier B.V., 343(1–2), pp. 173–175. doi: 10.1016/j.jns.2014.06.003.

Ó Faoláin, E. *et al.* (2005) 'A study examining the effects of tissue processing on human tissue sections using vibrational spectroscopy', *Vibrational Spectroscopy*, 38(1–2), pp. 121–127. doi: 10.1016/j.vibspec.2005.02.013.

Okada, M. *et al.* (2012) 'Label-free Raman observation of cytochrome c dynamics during apoptosis', *Proceedings of the National Academy of Sciences of the United States of America*, 109(1), pp. 28–32. doi: 10.1073/pnas.1107524108.

Oliván, S. *et al.* (2015) 'Comparative study of behavioural tests in the SOD1G93A mouse model of amyotrophic lateral sclerosis.', *Experimental animals / Japanese Association for Laboratory Animal Science*, 64(2), pp. 147–53. doi: 10.1538/expanim.14-0077.

Oshima, J. *et al.* (2009) 'Regional genomic instability predisposes to complex dystrophin gene rearrangements', *Human Genetics*, 126(3), pp. 411–423. doi: 10.1007/s00439-009-0679-9.

Paganoni, S. *et al.* (2014) 'Diagnostic timelines and delays in diagnosing amyotrophic lateral sclerosis (ALS)', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 15(5–6), pp. 453–456. doi: 10.3109/21678421.2014.903974.

Paganoni, S. and Amato, A. (2013) 'Electrodiagnostic Evaluation of Myopathies', *Physical Medicine and Rehabilitation Clinics of North America*, 24(1), pp. 193–207. doi: 10.1016/j.pmr.2012.08.017.

Pansarasa, O. *et al.* (2014) 'Amyotrophic lateral sclerosis and skeletal muscle: An update', *Molecular Neurobiology*, 49(2), pp. 984–990. doi: 10.1007/s12035-013-8578-4.

Park, J. *et al.* (2015) 'Multi-parametric MRI at 14T for muscular dystrophy mice treated with AAV vector-mediated gene therapy', *PLoS ONE*, 10(4), pp. 1–24. doi: 10.1371/journal.pone.0124914.

Parkkinen, S. *et al.* (2013) 'Gait impairment in a rat model of focal cerebral ischemia',

Stroke Research and Treatment, (May 2017). doi: 10.1155/2013/410972.

Parvathy, S. S. and Masocha, W. (2013a) 'Gait analysis of C57BL/6 mice with complete Freund's adjuvant-induced arthritis using the CatWalk system', *BMC Musculoskeletal Disorders*, 14(1), p. 14. doi: 10.1186/1471-2474-14-14.

Parvathy, S. S. and Masocha, W. (2013b) 'Gait analysis of C57BL/6 mice with complete Freund's adjuvant-induced arthritis using the CatWalk system', *BMC Musculoskeletal Disorders*. BMC Musculoskeletal Disorders, 14(1), p. 1. doi: 10.1186/1471-2474-14-14.

Pasinelli, P. and Brown, R. H. (2006) 'Molecular biology of amyotrophic lateral sclerosis: insights from genetics', *Nature Reviews Neuroscience*, 7(9), pp. 710–723. doi: 10.1038/nrn1971.

Pellegrino, M. A. *et al.* (2003) 'Orthologous myosin isoforms and scaling of shortening velocity with body size in mouse, rat, rabbit and human muscles.', *The Journal of physiology*, 546(Pt 3), pp. 677–89. doi: 10.1113/jphysiol.2002.027375.

Pence, I. and Mahadevan-Jansen, A. (2016) 'Clinical instrumentation and applications of Raman spectroscopy', *Chem. Soc. Rev.*, 45(7), pp. 1958–1979. doi: 10.1039/C5CS00581G.

Perlman, R. L. (2016) 'Mouse models of human disease: An evolutionary perspective.', *Evolution, medicine, and public health*, 2016(1), pp. 170–6. doi: 10.1093/emph/eow014.

Pestronk, A. (no date) *Neuromuscular Disease Center*. Available at: https://neuromuscular.wustl.edu (Accessed: 15 October 2020).

Petrof, B. J. *et al.* (1993) 'Dystrophin protects the sarcolemma from stresses developed during muscle contraction.', *Proceedings of the National Academy of Sciences of the United States of America*, 90(8), pp. 3710–4. doi: 10.1073/pnas.90.8.3710.

Pette, D., Peuker, H. and Staron, R. S. (1999) 'The impact of biochemical methods for single muscle fibre analysis.', *Acta physiologica Scandinavica*, 166(4), pp. 261–77. doi: 10.1046/j.1365-201x.1999.00568.x.

Pette, D. and Staron, R. S. (1997) 'Mammalian skeletal muscle fiber type transitions.', *International review of cytology*, 170, pp. 143–223. doi: 10.1016/s0074-7696(08)61622-8.

Pézolet, M. *et al.* (1980) 'Laser Raman scattering. A molecular probe of the contractile state of intact single muscle fibers.', *Biophysical journal*, 31(1), pp. 1–8. doi: 10.1016/S0006-3495(80)85036-3.

Pézolet, M. *et al.* (1988) 'Raman spectroscopy of cytoplasmic muscle fiber proteins. Orientational order', *Biophysical Journal*, 53(3), pp. 319–325. doi: 10.1016/S0006-

3495(88)83109-6.

Plesia, M. *et al.* (2021) 'In Vivo Fiber Optic Raman Spectroscopy of Muscle in Preclinical Models of Amyotrophic Lateral Sclerosis and Duchenne Muscular Dystrophy.', *ACS chemical neuroscience*, 12(10), pp. 1768–1776. doi: 10.1021/acschemneuro.0c00794.

Potenza, R. L. *et al.* (2016) 'Fingolimod: A Disease-Modifier Drug in a Mouse Model of Amyotrophic Lateral Sclerosis', *Neurotherapeutics*. Neurotherapeutics, 13(4), pp. 918–927. doi: 10.1007/s13311-016-0462-2.

Pradat, P.-F. and Dib, M. (2009) 'Biomarkers in Amyotrophic Lateral Sclerosis', *Molecular Diagnosis & Therapy*, 13(2), pp. 115–125. doi: 10.1007/BF03256320.

Pratt, S. J. P. *et al.* (2012) 'An in vivo rodent model of contraction-induced injury in the quadriceps muscle', *Injury*. Elsevier Ltd, 43(6), pp. 788–793. doi: 10.1016/j.injury.2011.09.015.

Pratt, S. J. P. *et al.* (2013) 'Temporal changes in magnetic resonance imaging in the mdx mouse', pp. 3–7. doi: 10.1186/1756-0500-6-262.

Prayson, R. A. (2006) 'Diagnostic Yield Associated With Multiple Simultaneous Skeletal Muscle Biopsies', *American Journal of Clinical Pathology*, 126(6), pp. 843–848. doi: 10.1309/78B3M0TGJYT4RUUM.

Prelle, A. *et al.* (2002) 'Retrospective study of a large population of patients with asymptomatic or minimally symptomatic raised serum creatine kinase levels.', *Journal of neurology*, 249(3), pp. 305–11. doi: 10.1007/s004150200010.

Prigogine, C. *et al.* (2012) 'Cerebellar dysfunction in the mdx mouse: An electrophysiological and behavioral study', *Neuromuscular Disorders*, 22(9–10), pp. 837–838. doi: 10.1016/j.nmd.2012.06.119.

Puppels, G. J. *et al.* (1991) 'Laser irradiation and Raman spectroscopy of single living cells and chromosomes: sample degradation occurs with 514.5 nm but not with 660 nm laser light.', *Experimental cell research*, 195(2), pp. 361–7. doi: 10.1016/0014-4827(91)90385-8.

Puppels, G. J. (1999) 'Confocal Raman microspectroscopy', in Mason, W. T. (ed.) *Fluorescent and luminescent probes*, pp. 377–406.

Radley-Crabb, H. G. *et al.* (2014) 'Dystropathology increases energy expenditure and protein turnover in the Mdx mouse model of Duchenne muscular dystrophy', *PLoS ONE*, 9(2). doi: 10.1371/journal.pone.0089277.

Radley, H. G., Davies, M. J. and Grounds, M. D. (2008) 'Reduced muscle necrosis and long-term benefits in dystrophic mdx mice after cV1q (blockade of TNF) treatment', 18, pp. 227–238. doi: 10.1016/j.nmd.2007.11.002.

Radley, H. G. and Grounds, M. D. (2006) 'Cromolyn administration (to block mast cell

degranulation) reduces necrosis of dystrophic muscle in mdx mice', 23, pp. 387–397. doi: 10.1016/j.nbd.2006.03.016.

Rakhit, R. and Chakrabartty, A. (2006) 'Structure, folding, and misfolding of Cu,Zn superoxide dismutase in amyotrophic lateral sclerosis', *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1762(11–12), pp. 1025–1037. doi: 10.1016/j.bbadis.2006.05.004.

Raman, C. . (1928) 'A new radiation', *Indian J. Phys*, 2, p. 387.

Rao, M. V. and Nixon, R. A. (2003) 'Defective Neurofilament Transport in Mouse Models of Amyotrophic Lateral Sclerosis: A Review', *Neurochemical Research*, 28(7), pp. 1041–1047. doi: 10.1023/A:1023259207015.

Rau, J. V. *et al.* (2016) 'RAMAN spectroscopy imaging improves the diagnosis of papillary thyroid carcinoma.', *Scientific reports*. Nature Publishing Group, 6(September), p. 35117. doi: 10.1038/srep35117.

Reaume, A. G. *et al.* (1996) 'Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury', *Nature Genetics*, 13(1), pp. 43–47. doi: 10.1038/ng0596-43.

Redd, D. C. B., Feng, Z. C., *et al.* (1993) 'Raman Spectroscopic Characterization of Human Breast Tissues: Implications for Breast Cancer Diagnosis', *Applied Spectroscopy*, 47(6), pp. 787–791. doi: 10.1366/0003702934067072.

Redd, D. C. B., Frank, C. J., *et al.* (1993) 'Raman spectrosopic characterization of human malignant tissues', in Cubeddu, R., Svanberg, S., and van den Bergh, H. (eds), pp. 185–190. doi: 10.1117/12.166824.

Régal, L. *et al.* (2006) 'The G93C Mutation in Superoxide Dismutase 1', *Archives of Neurology*, 63(2), p. 262. doi: 10.1001/archneur.63.2.262.

Rehman, I., Movagashi, Z. and Rehman, S. (2012) *Vibrational Spectroscopy for tissue analysis*. First Edit. Boca Raton: CRC Press.

Rivas-Arancibia, S. *et al.* (2017) 'Structural changes of amyloid beta in hippocampus of rats exposed to ozone: A raman spectroscopy study', *Frontiers in Molecular Neuroscience*, 10(May), pp. 1–11. doi: 10.3389/fnmol.2017.00137.

Robberecht, W. and Philips, T. (2013) 'The changing scene of amyotrophic lateral sclerosis', *Nature Reviews Neuroscience*, 14(4), pp. 248–264. doi: 10.1038/nrn3430.

Roig, M., Roma, J. and Fargas, A. (2004) 'Longitudinal pathologic study of the gastrocnemius muscle group in mdx mice', pp. 27–34. doi: 10.1007/s00401-003-0773-3.

Rome, L. C., Sosnicki, A. A. and Goble, D. O. (1990) 'Maximum velocity of shortening of three fibre types from horse soleus muscle: implications for scaling with body size.', *The Journal of physiology*, 431, pp. 173–85. doi:

10.1113/jphysiol.1990.sp018325.

Rosen, D. R. *et al.* (1993) 'Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis', *Nature*, 362(6415), pp. 59–62. doi: 10.1038/362059a0.

Rosenthal, N. and Brown, S. (2007) 'The mouse ascending: perspectives for human-disease models', *Nature Cell Biology*, 9(9), pp. 993–999. doi: 10.1038/ncb437.

Ruiten, H. Van, Bushby, K. and Guglieri, M. (2017) 'State-of-the-Art Advances in Duchenne Muscular Dystrophy', (March), pp. 90–99. Available at: http://emjreviews.com/wp-content/uploads/State-of-the-Art-Advances-in-Duchenne-Muscular-Dystrophy.pdf.

Ruiz-Chica, A. J. *et al.* (2004) 'Characterization by Raman spectroscopy of conformational changes on guanine–cytosine and adenine–thymine oligonucleotides induced by aminooxy analogues of spermidine', *Journal of Raman Spectroscopy*, 35(2), pp. 93–100. doi: 10.1002/jrs.1107.

Ruiz-Perez, D. *et al.* (2018) 'So you think you can PLS-DA?', in *2018 IEEE 8th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)*. IEEE, pp. 1–1. doi: 10.1109/ICCABS.2018.8542038.

Saavedra, R. *et al.* (2014) 'Raman measurements in silica glasses irradiated with energetic ions', in *AIP Conference Proceedings*, pp. 118–124. doi: 10.1063/1.4900466.

Sali, A. *et al.* (2012) 'Glucocorticoid-treated mice are an inappropriate positive control for long-term preclinical studies in the mdx mouse', *PLoS ONE*, 7(4), pp. 1–9. doi: 10.1371/journal.pone.0034204.

Salimena, C. M. and Lagrota-Candido, Jussara, Quirico-Santos, T. (2004) 'Gender dimorphism influences extracellular matrix expression and regeneration of muscular tissue in mdx dystrophic mice', *Histochem Cell Biol*, 122, pp. 435–444. doi: 10.1007/s00418-004-0707-8.

Sasaki, S. and Iwata, M. (1996) 'Impairment of fast axonal transport in the proximal axons of anterior horn neurons in amyotrophic lateral sclerosis.', *Neurology*, 47(2), pp. 535–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8757033.

Sato, H. *et al.* (2014) 'Biomedical Applications of Raman Spectroscopy', in *Encyclopedia of Analytical Chemistry*. Chichester, UK: John Wiley & Sons, Ltd, pp. 1–12. doi: 10.1002/9780470027318.a9281.

Sato, T. *et al.* (2005) 'Rapid disease progression correlates with instability of mutant SOD1 in familial ALS.', *Neurology*, 65(12), pp. 1954–7. doi: 10.1212/01.wnl.0000188760.53922.05.

Schiaffino, S. (2010) 'Fibre types in skeletal muscle: a personal account.', *Acta*

physiologica (Oxford, England), 199(4), pp. 451–63. doi: 10.1111/j.1748-1716.2010.02130.x.

Schiaffino, S. and Reggiani, C. (2011) 'Fiber Types in Mammalian Skeletal Muscles', *Physiological Reviews*, 91(4), pp. 1447–1531. doi: 10.1152/physrev.00031.2010.

Schiffenbauer, A. (2014) 'Imaging: Seeing Muscle in New Ways', *Current Opinion in Rheumatology*, 26(6), pp. 712–716. doi: 10.1097/BOR.0000000000000105.

Schill, K. E. *et al.* (2016) 'Muscle damage, metabolism, and oxidative stress in mdx mice: Impact of aerobic running.', *Muscle & nerve*, 54(1), pp. 110–7. doi: 10.1002/mus.25015.

Scott, S. *et al.* (2008) 'Design, power, and interpretation of studies in the standard murine model of ALS', *Amyotrophic Lateral Sclerosis*, 9(1), pp. 4–15. doi: 10.1080/17482960701856300.

Scott, W., Stevens, J. and Binder-Macleod, S. A. (2001) 'Human skeletal muscle fiber type classifications.', *Physical therapy*, 81(11), pp. 1810–6. doi: 10.1093/ptj/81.11.1810.

Seow, C. Y. and Ford, L. E. (1991) 'Shortening velocity and power output of skinned muscle fibers from mammals having a 25,000-fold range of body mass.', *The Journal of general physiology*, 97(3), pp. 541–60. doi: 10.1085/jgp.97.3.541.

Shafer-Peltier, K. E. *et al.* (2003) 'Toward a Glucose Biosensor Based on Surface-Enhanced Raman Scattering', *Journal of the American Chemical Society*, 125(2), pp. 588–593. doi: 10.1021/ja028255v.

Shaibani, A. *et al.* (2015) 'Diagnostic outcome of muscle biopsy', *Muscle and Nerve*, 51(5), pp. 662–668. doi: 10.1002/mus.24447.

Shaver, J. . (2001) 'Chemometrics for Raman Spectroscopy', in *Handbook of Raman Spectroscopy*, pp. 292–298.

Shetty, G. *et al.* (2006) 'Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus.', *British journal of cancer*, 94(10), pp. 1460–1464. doi: 10.1038/sj.bjc.6603102.

Shi, P. *et al.* (2010) 'Mitochondrial dysfunction in amyotrophic lateral sclerosis', *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Elsevier B.V., 1802(1), pp. 45–51. doi: 10.1016/j.bbadis.2009.08.012.

Shibata, N. *et al.* (2001) 'Morphological evidence for lipid peroxidation and protein glycoxidation in spinal cords from sporadic amyotrophic lateral sclerosis patients.', *Brain research*, 917(1), pp. 97–104. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11602233.

Shim, M. G. *et al.* (1999) 'Study of fiber-optic probes for in vivo medical Raman spectroscopy', *Applied Spectroscopy*, 53(6), pp. 619–627. doi:

10.1366/0003702991947225.

Si, Y. *et al.* (2014) 'Smads as muscle biomarkers in amyotrophic lateral sclerosis.', *Annals of clinical and translational neurology*, 1(10), pp. 778–87. doi: 10.1002/acn3.117.

Silveira, L. *et al.* (2002) 'Correlation between near-infrared Raman spectroscopy and the histopathological analysis of atherosclerosis in human coronary arteries.', *Lasers in surgery and medicine*, 30(4), pp. 290–7. doi: 10.1002/lsm.10053.

Smerdu, V. *et al.* (1994) 'Type IIx myosin heavy chain transcripts are expressed in type IIb fibers of human skeletal muscle.', *The American journal of physiology*, 267(6 Pt 1), pp. C1723-8. doi: 10.1152/ajpcell.1994.267.6.C1723.

Smerdu, V. and Cvetko, E. (2013) 'Myosin heavy chain-2b transcripts and isoform are expressed in human laryngeal muscles.', *Cells, tissues, organs*, 198(1), pp. 75–86. doi: 10.1159/000351293.

Smith, E. and Dent, G. (2005) *Modern Raman Spectroscopy- A Practical Approach*. First Edit. Chichester: John Wiley & Sons, Ltd.

Smith, L. R. and Barton, E. R. (2014) 'Collagen content does not alter the passive mechanical properties of fibrotic skeletal muscle in mdx mice', *AJP: Cell Physiology*, 306(10), pp. C889–C898. doi: 10.1152/ajpcell.00383.2013.

Smulko, J. and Wróbel, M. S. (2017) 'Noise sources in Raman spectroscopy of biological objects', *Dynamics and Fluctuations in Biomedical Photonics XIV*, 10063, p. 100630Q. doi: 10.1117/12.2254807.

Smulko, J., Wrobel, M. S. and Barman, I. (2015) 'Noise in biological Raman spectroscopy', in *2015 International Conference on Noise and Fluctuations (ICNF)*. IEEE, pp. 1–6. doi: 10.1109/ICNF.2015.7288562.

Sogawa, K. *et al.* (2017) 'Neurogenic and Myogenic Diseases: Quantitative Texture Analysis of Muscle US Data for Differentiation', *Radiology*, 283(2), pp. 492–498. doi: 10.1148/radiol.2016160826.

Spurney, C. F. *et al.* (2009) 'Preclinical drug trials in the mdx mouse: Assessment of reliable and sensitive outcome measures', *Muscle & Nerve*, 39(5), pp. 591–602. doi: 10.1002/mus.21211.

Stedman, H. H. *et al.* (1991) 'The mdx mouse diaphragm reproduces the degenerative changes of Duchenne muscular dystrophy.', *Nature*, 352(6335), pp. 536–9. doi: 10.1038/352536a0.

Stevens, O. *et al.* (2016) 'Developing fibre optic Raman probes for applications in clinical spectroscopy', *Chem. Soc. Rev.* Royal Society of Chemistry, 45, pp. 1919–1934. doi: 10.1039/C5CS00850F.

Stevens, O. A. C. *et al.* (2014) 'A low background Raman probe for optical biopsy of

brain tissue', *Biomedical Vibrational Spectroscopy VI: Advances in Research and Industry*, 8939(0), p. 89390W. doi: 10.1117/12.2044139.

Stone, N. *et al.* (2002) 'Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers', *Journal of Raman Spectroscopy*, 33(7), pp. 564–573. doi: 10.1002/jrs.882.

Stone, N. *et al.* (2004) 'Raman spectroscopy for identification of epithelial cancers.', *Faraday discussions*, 126(1), pp. 141–57. doi: 10.1039/b304992b.

Stone, N. and Matousek, P. (2008) 'Advanced transmission Raman spectroscopy: A promising tool for breast disease diagnosis', *Cancer Research*, 68(11), pp. 4424–4430. doi: 10.1158/0008-5472.CAN-07-6557.

Straughan, B. . and Walker, S. (eds) (1976) *Spectroscopy, Volume 2*. Second Edi. London.

Stupka, N. *et al.* (2004) 'The calcineurin signal transduction pathway is essential for successful muscle regeneration in mdx dystrophic mice', pp. 299–310. doi: 10.1007/s00401-003-0807-x.

Su, L. *et al.* (2012) 'Raman spectral properties of squamous cell carcinoma of oral tissues and cells', *Laser Physics*, 22(1), pp. 311–316. doi: 10.1134/s1054660x12010185.

Al Sultan, A. *et al.* (2016) 'The genetics of amyotrophic lateral sclerosis: current insights', *Degenerative Neurological and Neuromuscular Disease*, 43(6), p. 49. doi: 10.2147/DNND.S84956.

Surmacki, J. M. *et al.* (2017) 'Label-free monitoring of tissue biochemistry following traumatic brain injury using Raman spectroscopy', *Analyst*. Royal Society of Chemistry, 142(1), pp. 132–139. doi: 10.1039/c6an02238c.

Talari, A. C. S. *et al.* (2015) 'Raman Spectroscopy of Biological Tissues', *Applied Spectroscopy Reviews*, 50(1), pp. 46–111. doi: 10.1080/05704928.2014.923902.

Talbot, J. and Maves, L. (2016) 'Skeletal muscle fiber type: using insights from muscle developmental biology to dissect targets for susceptibility and resistance to muscle disease.', *Wiley interdisciplinary reviews. Developmental biology*, 5(4), pp. 518–34. doi: 10.1002/wdev.230.

Tanabe, Y., Esaki, K. and Nomura, T. (1986) 'Skeletal muscle pathology in X chromosome-linked muscular dystrophy (mdx) mouse', *Acta Neuropathologica*, 69(1–2), pp. 91–95. doi: 10.1007/BF00687043.

Tang, H. *et al.* (2007) 'NIR Raman spectroscopic investigation of single mitochondria trapped by optical tweezers', *Optics Express*, 15(20), p. 12708. doi: 10.1364/oe.15.012708.

Taylor, J. P., Brown, R. H. and Cleveland, D. W. (2016) 'Decoding ALS: from genes to

mechanism', *Nature*, 539(7628), pp. 197–206. doi: 10.1038/nature20413.

Teixeira, C. S. B. *et al.* (2009) 'Thyroid tissue analysis through Raman spectroscopy', *Analyst*, 134(11), pp. 2361–2370. doi: 10.1039/b822578h.

Theodorou, D. J., Theodorou, S. J. and Kakitsubata, Y. (2012) 'Skeletal muscle disease: patterns of MRI appearances', *The British Journal of Radiology*, 85(1020), pp. e1298–e1308. doi: 10.1259/bjr/14063641.

Thomas, G. J. *et al.* (1995) 'Polarized Raman spectra of oriented fibers of A DNA and B DNA: anisotropic and isotropic local Raman tensors of base and backbone vibrations', *Biophysical Journal*. Elsevier, 68(3), pp. 1073–1088. doi: 10.1016/S0006-3495(95)80282-1.

Tobisawa, S. *et al.* (2003) 'Mutant SOD1 linked to familial amyotrophic lateral sclerosis, but not wild-type SOD1, induces ER stress in COS7 cells and transgenic mice.', *Biochemical and biophysical research communications*, 303(2), pp. 496–503. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12659845.

Tomasová Studynková, J. *et al.* (2007) 'The role of MRI in the assessment of polymyositis and dermatomyositis.', *Rheumatology (Oxford, England)*, 46(7), pp. 1174–9. doi: 10.1093/rheumatology/kem088.

Trevisan, J. *et al.* (2012) 'Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives.', *The Analyst*, 137(14), pp. 3202–15. doi: 10.1039/c2an16300d.

Turner, M. *et al.* (2013) 'Mechanisms, models and biomarkers in amyotrophic lateral sclerosis Martin', *Amyotroph Lateral Scler Frontotemporal Degener*, 14(0–1), pp. 19–32. doi: 10.3109/21678421.2013.778554.Mechanisms.

Turner, M. R. *et al.* (2009) 'Biomarkers in amyotrophic lateral sclerosis', *The Lancet Neurology*. Elsevier Ltd, 8(1), pp. 94–109. doi: 10.1016/S1474-4422(08)70293-X.

Urushitani, M. *et al.* (2006) 'Chromogranin-mediated secretion of mutant superoxide dismutase proteins linked to amyotrophic lateral sclerosis.', *Nature neuroscience*, 9(1), pp. 108–18. doi: 10.1038/nn1603.

Vandeputte, C. *et al.* (2010) 'Automated quantitative gait analysis in animal models of movement disorders'.

Varmuza, K. and Filzmoser, P. (2009) *Introduction to Multivariate Statistical Analysis in Chemometrics*. 1st edn. Boca Raton: CRC Press.

Vieira, W. F., Kenzo-kagawa, B. and Cogo, J. C. (2016) 'Low-Level Laser Therapy (904 nm) Counteracts Motor Deficit of Mice Hind Limb following Skeletal Muscle Injury Caused by Snakebite- Mimicking Intramuscular Venom Injection'. doi: 10.1371/journal.pone.0158980.

De Vos, K. J. *et al.* (2007) 'Familial amyotrophic lateral sclerosis-linked SOD1 mutants perturb fast axonal transport to reduce axonal mitochondria content.', *Human molecular genetics*, 16(22), pp. 2720–2728. doi: 10.1093/hmg/ddm226.

Vucic, S. (2016) 'Motor Unit Number Index (MUNIX): A novel biomarker for ALS?', *Clinical Neurophysiology*. International Federation of Clinical Neurophysiology, 127(4), pp. 1938–1939. doi: 10.1016/j.clinph.2016.01.012.

Vucic, S., Rothstein, J. D. and Kiernan, M. C. (2014) 'Advances in treating amyotrophic lateral sclerosis: insights from pathophysiological studies.', *Trends in neurosciences*, 37(8), pp. 433–42. doi: 10.1016/j.tins.2014.05.006.

Wachsmann-Hogiu, S., Weeks, T. and Huser, T. (2009) 'Chemical analysis in vivo and in vitro by Raman spectroscopy-from single cells to humans', *Current Opinion in Biotechnology*, 20(1), pp. 63–73. doi: 10.1016/j.copbio.2009.02.006.

Walter, G. *et al.* (2005) 'Noninvasive monitoring of gene correction in dystrophic muscle', *Magnetic Resonance in Medicine*, 54(6), pp. 1369–1376. doi: 10.1002/mrm.20721.

Wang, H. *et al.* (2011) 'Depth-resolved in vivo micro-Raman spectroscopy of a murine skin tumor model reveals cancer-specific spectral biomarkers', *Journal of Raman Spectroscopy*, 42(2), pp. 160–166. doi: 10.1002/jrs.2677.

Wang, J. *et al.* (2003) 'Copper-binding-site-null SOD1 causes ALS in transgenic mice: aggregates of non-native SOD1 delineate a common feature.', *Human molecular genetics*, 12(21), pp. 2753–64. doi: 10.1093/hmg/ddg312.

Wang, J. *et al.* (2015) 'Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy improves in vivo diagnosis of esophageal squamous cell carcinoma at endoscopy', *Scientific Reports*. Nature Publishing Group, 5(1), p. 12957. doi: 10.1038/srep12957.

Wei, R. *et al.* (2012) 'Protein misfolding, mitochondrial dysfunction and muscle loss are not directly dependent on soluble and aggregation state of mSOD1 protein in skeletal muscle of ALS', *Biochemical and Biophysical Research Communications*. Elsevier Inc., 417(4), pp. 1275–1279. doi: 10.1016/j.bbrc.2011.12.126.

Wei, R. *et al.* (2013) 'Differential effects of mutant SOD1 on protein structure of skeletal muscle and spinal cord of familial amyotrophic lateral sclerosis: Role of chaperone network', *Biochemical and Biophysical Research Communications*. Elsevier Inc., 438(1), pp. 218–223. doi: 10.1016/j.bbrc.2013.07.060.

Weng, Y.-M. *et al.* (2003) 'Structural analysis of triacylglycerols and edible oils by near-infrared Fourier transform Raman spectroscopy.', *Applied spectroscopy*, 57(4), pp. 413–8. doi: 10.1366/00037020360625952.

Weydt, P. *et al.* (2003) 'Assessing disease onset and progression in the SOD1 mouse model of ALS', *NeuroReport*, 14(7), pp. 1051–1054. doi:

10.1097/01.wnr.0000073685.00308.89.

White, S. J. *et al.* (2006) 'Duplications in the DMD gene.', *Human mutation*, 27(9), pp. 938–45. doi: 10.1002/humu.20367.

Wong, M. and Martin, L. J. (2010) 'Skeletal muscle-restricted expression of human SOD1 causes motor neuron degeneration in transgenic mice', *Human Molecular Genetics*, 19(11), pp. 2284–2302. doi: 10.1093/hmg/ddq106.

Wong, P. C. *et al.* (2000) 'Copper chaperone for superoxide dismutase is essential to activate mammalian Cu/Zn superoxide dismutase.', *Proceedings of the National Academy of Sciences of the United States of America*, 97(6), pp. 2886–91. doi: 10.1073/pnas.040461197.

Yiu, E. M. and Kornberg, A. J. (2015) 'Duchenne muscular dystrophy', *Journal of Paediatrics and Child Health*, 51(8), pp. 759–764. doi: 10.1111/jpc.12868.

Yu, N. ., Li, X. . and Kuck, J. (1996) 'Biomedical applications of Raman Spectroscopy: eye lens research and cardiovascular disease', in Clark, R. J. . and Hester, R. . (eds) *Biomedical applications of Raman Spectroscopy*. John Wiley and Sons.

Yu, T. -J, Lippert, J. L. and Peticolas, W. L. (1973) 'Laser Raman studies of conformational variations of poly-L-lysine', *Biopolymers*, 12(9), pp. 2161–2176. doi: 10.1002/bip.1973.360120919.

Zammit, P. S. *et al.* (2004) 'Muscle satellite cells adopt divergent fates: a mechanism for self-renewal?', *The Journal of cell biology*, 166(3), pp. 347–57. doi: 10.1083/jcb.200312007.

Zhu, G. *et al.* (2011) 'Raman spectra of amino acids and their aqueous solutions.', *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*. Elsevier B.V., 78(3), pp. 1187–95. doi: 10.1016/j.saa.2010.12.079.

Appendix A

Raman Shifts Table

Raman Shift (cm^{-1})	Assignment	Tissue/Compound
617	Thymine	Thymine (De Gelder *et al.*, 2007)
618	Phenylalanine	Collagen (Frank *et al.*, 1994), Phenylalanine (Zhu *et al.*, 2011)
	Proteins	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
620	Phenylalanine	Diseased breast tissue (Frank *et al.*, 1994)
621	Phenylalanine	Oesophagus (Stone *et al.*, 2002),
622	Phenylalanine	Phenylalanine (Zhu *et al.*, 2011)
623	Adenine	Adenine (De Gelder *et al.*, 2007)
662	Cystine	Hard mass pilomaxitroma (probably collagen shifted peak from collagen peak at 667) (Cheng *et al.*, 2005)
664	Valine	Valine (De Gelder *et al.*, 2007)
665	Valine	Valine (Zhu *et al.*, 2011)
666	Guanine, Thymine (DNA)	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
667	Cystine	Collagen type I (Cheng *et al.*, 2005)
668	Guanine	DNA, RNA (Mahadevan-Jansen, 1996)
669	Cytosine	Oesophagus (Stone *et al.*, 2002)
	Thymine	(Puppels *et al.*, 1991)
740	Thymine	Thymine (De Gelder *et al.*, 2007)
741	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
742	DNA, Tryptophan	Murine blood cells (Huang *et al.*, 2011)
902 (903 measured)	Valine	Valine (De Gelder *et al.*, 2007)
913	Glucose	Glucose (Talari *et al.*, 2015)
914	Glucose	Glucose (De Gelder *et al.*, 2007)

915	Ribose	Protein (ribose-5-phospate isomerase) (Hartman, Clayton and Thomas, 1973)
	Phenylalanine	Phenylalanine solid (Zhu *et al.*, 2011)
916	Proline	L-Proline (De Gelder *et al.*, 2007)
917	Deoxyribose	DNA (Mahadevan-Jansen, 1996)
	Glutathione	Glutathione (De Gelder *et al.*, 2007)
918	Proline, hydroxyproline	Collagen Type I, patient normal skin dermis (Talari *et al.*, 2015)
920	Proline ring/glucose/lactic acid	(Stone *et al.*, 2004)
	Proline ring (collagen assignment)	Collagen (Bonnier and Byrne, 2012)
	Glucose	Glucose (Mahadevan-Jansen *et al.*, 1998)
	Elastin	Elastin (Kendall, 2002)
	Proline ring	Ductal carcinoma (breast), Collagen Type I measured (Frank, McCreery and Redd, 1995)
930	Proline	Proline (De Gelder *et al.*, 2007)
931	Carbohydrates	Glycogen solution, solid (Talari *et al.*, 2015)
	Glutathione	Glutathione (De Gelder *et al.*, 2007)
932	Skeletal, α-helix	DNA (Puppels *et al.*, 1991), (Stone *et al.*, 2004)
933	Proline, hydroxyproline, collagen	Hard mass pilomaxitroma, shifted from 937 in collagen and skin dermis (Cheng *et al.*, 2005)
934	Proline	Type I collagen (Frank, McCreery and Redd, 1995)
	Collagen	Collagen (Bonnier and Byrne, 2012)
935	Collagen	Fibrocystic human breast tissue (Kendall, 2002) (Redd 1993)
	Proline, valine and protein backbone (α-helix conformation)/glycogen	Oesophagus (Stone *et al.*, 2002)
936	Collagen	Collagen (Kendall, 2002)
937	Glycogen	Glycogen (Kendall, 2002)
	α-helix	Intact muscle fibre (Pézolet *et al.*, 1980)

	Proline	Infiltrating Duct Carcinoma (Frank, McCreery and Redd, 1995)
	Proline, hydroxyproline, skeletal of collagen backbone	Collagen Type I, Skin dermis (Cheng *et al.*, 2005)
	Collagen type I	Skin dermis (Wang *et al.*, 2011)
	Collagen	Collagen Type I (Nguyen *et al.*, 2012)
938	C-C backbone	Collagen (Kendall, 2002), (Mahadevan-Jansen, 1996)
940	Proline, Valine	Normal human skin (Gniadecka *et al.*, 1997)
	Triple helix vibration	Collagen Type I, dermis (Fendel and Schrader, 1998)
941	Skeletal modes (polysaccharides, amylose)	Oesophagus (Shetty *et al.*, 2006)
	Adenine	Adenine (De Gelder *et al.*, 2007)
948	Valine (m)	Valine solution (Zhu *et al.*, 2011)
	Valine (s)	Valine (De Gelder *et al.*, 2007)
950	4-hydroxyproline	Increased in malignant breast tissue (Feld *et al.*, 1995)
	Most probably due to single bond stretching vibrations for the amino acids proline and valine and polysaccharides	Normal skin (decreased in BCC) (Gniadecka *et al.*, 1997)
	Calcium-phosphate	Murine skull and teeth (Huang *et al.*, 2011)
	Deoxyribose	Gastric cancer DNA changes (Chen *et al.*, 2014)
	Cholesterol	Cell cultures Escheria coli (Germond *et al.*, 2018)
951	Protein α-helix	Mouse brain tissue (release of protective factors after irradiation) (Jyothi Lakshmi *et al.*, 2002)
	Proline (m)	Proline (De Gelder *et al.*, 2007)
	Phenylalanine (w)	Phenylalanine (De Gelder *et al.*, 2007)
952	Proline (w)	Proline solid (Zhu *et al.*, 2011)
955	Hydroxyapatite	(Manoharan *et al.*, 1992)
956	Carotenoids	Human brain tissue (Mizuno *et al.*, 1994)

	Carotenoids	Brain tissue, Absent in normal tissue (Mahadevan-Jansen and Richards-Kortum, 1997)
957	Hydroxyapatite, carotenoid, cholesterol	Colon tissue (Stone *et al.*, 2004)
	Carotenoid, Cholesterol	Thyroid gland tissue (absent in normal tissue) (Rau *et al.*, 2016)
960	Calcification-Hydroxyapatite	Human brain tissue, central neurocytoma (Mizuno *et al.*, 1994)
	Calcification-Hydroxyapatite	Coronary artery segments (Clarke *et al.*, 1987)
	Calcification	Calcified plaque (Yu, Li and Kuck, 1996)
	Phosphate of Hydroxyapatite	Hydroxyapatite and hard mass pilomaxitroma (Cheng *et al.*, 2005)
965	Valine (m)	Valine (Zhu *et al.*, 2011)
966	Hydroxyapatite	Calcified plaque (Clarke *et al.*, 1987)
	Triple helix vibrations	Collagen type I, normal skin dermis (Fendel and Schrader, 1998)
	Hydroxyapatite	Breast tissue (Stone *et al.*, 2004)
950-970	Apatite peak region	Osteogenic differentiation of human mesenchymal stem cells (Brauchle and Schenke-Layland, 2013), (Chiang *et al.*, 2009)
971	Cytosine	Cytosine (De Gelder *et al.*, 2007)
972	Collagen	Collagen, normal tissue assignments (Frank, McCreery and Redd, 1995)
	Glutathione	Glutathione (De Gelder *et al.*, 2007)
975	Ribose	Protein (Hartman, Clayton and Thomas, 1973)
	Deoxyribose	DNA (Mahadevan-Jansen, 1996)
978	Phosphate ion	Phospholipids, human cervix, human cervical biopsies (Mahadevan-Jansen, 1996)
983	Tyrosine	Tyrosine (De Gelder *et al.*, 2007)
984	Thymine, Uracil	Thymine, Uracil (De Gelder *et al.*, 2007)
987	Proline	Proline (De Gelder *et al.*, 2007)

988	Tryptophan, Glutathione	Tryptophan, Glutathione (De Gelder *et al.*, 2007)
991	Cytosine	Cytosine (De Gelder *et al.*, 2007)
	Phenylalanine, NADH	Murine stomach, small intestine, colon, bladder, lung, brain (Huang *et al.*, 2011)
992	Proline	Proline (Zhu *et al.*, 2011)
994	Proline	Proline (De Gelder *et al.*, 2007)
995	Uracil	Uracil (De Gelder *et al.*, 2007)
1000	Phenylalanine	Oral tissue (increased in malignant tissue) (Malini *et al.*, 2006)
	NADH free and bound	Weak in normal oral tissue (Malini *et al.*, 2006)
1001	Phenylalanine	Oesophagus (Stone *et al.*, 2004), (Stone *et al.*, 2002)
	Phenylalanine	Phenylalanine (Kendall, 2002) (Kendall unpublished)
	Phenylalanine	Increased in malignant breast tissue (Feld *et al.*, 1995)
1002	Phenylalanine	Skin dermis, hard mass pilomaxitroma (Cheng *et al.*, 2005)
	Phenylalanine	Collagen (Frank *et al.*, 1994)
	Hydroxyproline, Tyrosine	Type I collagen, placenta (Frank, McCreery and Redd, 1995)
1004, 1005	Phenylalanine (vs)	Phenylalanine (De Gelder *et al.*, 2007), (Zhu *et al.*, 2011)
1011	Cytosine	Cytosine (De Gelder *et al.*, 2007)
1014	Tryptophan	(Mahadevan-Jansen *et al.*, 1998)
1015	Glutathione	Glutathione (De Gelder *et al.*, 2007)
	Tryptophan	Protein (Hartman, Clayton and Thomas, 1973)
1016	Tryptophan	DNA in water (Kendall, 2002) (Manoharan 1995)
1029	Valine	Valine (Zhu *et al.*, 2011)
1030	Phenylalanine (collagen)	Collagen Type I, hard mass pilomaxitroma (Cheng *et al.*, 2005)

	Collagen	Collagen (Kendall, 2002)
1031	Proline	Proline (Yu, Li and Kuck, 1996)
	Phenylalanine	Normal (increased) and diseased bronchial tissue (Huang, McWilliams, Lui, *et al.*, 2003)
	Phenylalanine	Prostate (Stone *et al.*, 2004)
1032	Phenylalanine	Rabbit and human cornea (Erckens *et al.*, 1997)
	Proline	Type I collagen (Frank, McCreery and Redd, 1995)
	Phenylalanine (collagen)	Skin dermis (Cheng *et al.*, 2005)
	Phenylalanine	Phenylalanine (Zhu *et al.*, 2011)
1033	Phenylalanine	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
1034	Phenylalanine (collagen)	Soft mass pilomaxitroma (Cheng *et al.*, 2005)
1035	Valine, proline	Valine, proline (De Gelder *et al.*, 2007)
1037	Phenylalanine	Phenylalanine (De Gelder *et al.*, 2007)
1041	Proline	Proline (Zhu *et al.*, 2011)
1043	Proline	Type I collagen (placenta) (Frank, McCreery and Redd, 1995)
1044	Proline	Murine serum, colon, bladder, stomach, small intestine, lung, brain 17
1045	Proline	Proline (De Gelder *et al.*, 2007)
1046	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
1048	Guanine	Guanine (De Gelder *et al.*, 2007)
1057	DNA	(Puppels, 1999)
1061	Lipids	(Kendall, 2002)
1070	Collagen, elastin	(Kendall, 2002)
	Triglycerides, fatty acids	Normal iliac artery (Silveira *et al.*, 2002)
1071	Apatite	Human tooth enamel (Kendall, 2002)
	Glucose	Glucose (Krafft *et al.*, 2005)
1073	Calcium carbonate apatite	Calcified atheromatous plaques (arterial fragments (Silveira *et al.*, 2002)

1074	Triglycerides	Adipose tissue from human aorta (Alfano *et al.*, 1991)
	Adipose tissue	(Kendall, 2002)
	Glucose, triglycerides	Measured in Murine Blood, small intestine, Colon (Huang *et al.*, 2011), assignment from (Krafft *et al.*, 2005), (Silveira *et al.*, 2002)
	Glutathione	Glutathione (De Gelder *et al.*, 2007)
1076	Lipids	Normal oral tissue
	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
1078	Phospholipids	Increased in normal lung tissue (Huang, McWilliams, Lui, *et al.*, 2003)
	Lipids, Nucleic acid	Measured in human colon and breast tissue (diff between cancerous and healthy) (Stone *et al.*, 2004)
	Triglyceride	Subcutaneous fat in stomach (Duraipandian, 2012)
	Tryptophan	Tryptophan solid (Zhu *et al.*, 2011)
1080	Phospholipids	Human brain tissue (Mizuno *et al.*, 1994)
	Typical phospholipids	Normal oral tissue (Malini *et al.*, 2006)
1082	Phospholipids, Nucleic acids	Normal human skin (Gniadecka *et al.*, 1997)
	Lipids	Normal human breast tissue, Lipid (Kendall, 2002) (Redd 1993)
1083	Proteins (and lipid mode to lesser degree)	Oesophagus (Stone *et al.*, 2002)
	Glycogen	Glycogen (Kendall, 2002) (Kendall unpublished)
	Proline (m)	Proline (De Gelder *et al.*, 2007)
1084	Lipids	Thyroid carcinoma →1086 in normal tissue (Rau *et al.*, 2016)
1095	Phosphate	DNA (Mahadevan-Jansen *et al.*, 1998)
1099	PO^-_2	DNA (Mahadevan-Jansen *et al.*, 1998)
	Palmitic Acid, tyrosine	Palmitic Acid (De Gelder *et al.*, 2007), Tyrosine (Zhu *et al.*, 2011)
1100	Uracil, collagen	Uracil (De Gelder *et al.*, 2007),

		Collagen solution (Zhu *et al.*, 2011)
	Lipid	(Gniadecka *et al.*, 1997)
1101	DNA	Normal/cancerous cervical tissue (Kamemoto *et al.*, 2010)
1102	Collagen	Collagen (Zhu *et al.*, 2011)
1104	Lipids	(Kendall, 2002) Yu 1996
1105	Tryptophan	Tryptophan (Zhu *et al.*, 2011)
1120	Tryptophan	Tryptophan solid (Zhu *et al.*, 2011)
	Carotene	Measured in murine blood (Huang *et al.*, 2011)
1123	Glycogen	(Kendall, 2002) Kendall (unpublished)
	Glucose	Glucose (Mahadevan-Jansen *et al.*, 1998)
	Lipids, proteins, glucose	Oesophagus (Stone *et al.*, 2004)
	Proteins (protein assignment)	Small difference between normal lung tissue and adenocarcinoma (more intense in cancer]) (Huang, Bergholt, *et al.*, 2010)
1124	Skeletal muscle main Raman band	Human muscle (Minamikawa, Harada and Takamatsu, 2015)
1126	Proteins	Protein (Puppels, 1999)
	Phospholipids	Human brain tissue (Mizuno *et al.*, 1994)
	Valine	Valine (Zhu *et al.*, 2011)
	Proteins	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
1131	Fatty acids	Cholesterol ester, cholesteryl palmitate, glyceryl tripalmitate (Krafft *et al.*, 2005)
1134	Adenine	Adenine (De Gelder *et al.*, 2007)
1143	Glutathione	Glutathione (De Gelder *et al.*, 2007)
1144	Valine	Valine (Zhu *et al.*, 2011)
1155	Proteins, carotenoids	Oesophagus (Stone *et al.*, 2002)
1156	Carotenoid	Human breast carcinoma, Beta carotene (Redd 1993)
	Carotenoids	Brain tissue present in neurinomas (absent in normal tissue) (Mahadevan-Jansen and Richards-Kortum, 1997)

	Proteins	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
	Carotenoids	Thyroid gland tissue (present in normal/abnormal tissue) (Rau *et al.*, 2016), (Talari *et al.*, 2015)
	Proteins	Thyroid gland tissue (present in normal/abnormal tissue) (Rau *et al.*, 2016)
		(Talari *et al.*, 2015)
	Thymine	Thymine (De Gelder *et al.*, 2007)
	β- carotene	β-carotene (De Gelder *et al.*, 2007)
	Tyrosine	Tyrosine (Zhu *et al.*, 2011)
1157	Carotenoid	Normal human breast tissue (Kendall, 2002) (Redd 1993)
	Carotenoid	Human breast tissue, acoustic neurinoma, beta carotene (Mizuno *et al.*, 1994)
	β-carotene accumulation	Atheroma in human coronary arteries (Silveira *et al.*, 2002)
	Phenylalanine	Phenylalanine (De Gelder *et al.*, 2007)
1158	Carotenoid	Carotid artery (Feld *et al.*, 1995)
	Carotenoid	Normal colon (Kendall, 2002) (Redd 1993)
	Guanine	Guanine (De Gelder *et al.*, 2007)
	Phenylalanine	Phenylalanine solid (Zhu *et al.*, 2011)
1160	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
	Protein	Murine blood, colon, small intestine (Huang *et al.*, 2011)
1163	Tyrosine	Collagen (Type I), (Cheng *et al.*, 2005)
1166	Collagen type I	Collagen type I (human placenta) (Frank, McCreery and Redd, 1995)
1169	Tyrosine	Measured in hard mass pilomaxitroma (probably shifted collagen peak) (Cheng *et al.*, 2005)
1170	Tyrosine	Oesophagus (Stone *et al.*, 2002)
1171	Tyrosine	Measured in soft mass PMX prob shifted collagen peak (Cheng *et al.*, 2005)

	Phenylalanine, tyrosine	Proteins appearing after irradiation (brain tissue) (Jyothi Lakshmi *et al.*, 2002)
1172	Tyrosine	Adenocarcinoma (bronchial tissue) (Huang, McWilliams, Lui, *et al.*, 2003)
1173	Cytosine, guanine	Guanine–cytosine oligonucleotides (Ruiz-Chica *et al.*, 2004)
	Proline	Proline solid (Zhu *et al.*, 2011)
1174	Palmitic Acid	Palmitic Acid (De Gelder *et al.*, 2007)
1175	Proline	Proline (De Gelder *et al.*, 2007)
	Tyrosine, phenylalanine	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
1180, 1180-1184	Cytosine, guanine, adenine	Guanine–cytosine oligonucleotides (Ruiz-Chica *et al.*, 2004)
	Tyrosine	(Miura and Thomas, 1995)
1190	b-carotene	b-carotene (De Gelder *et al.*, 2007)
1191	Valine	Valine (De Gelder *et al.*, 2007)
1194	Proline	Proline (De Gelder *et al.*, 2007)
1197	Carotenoid	(Frank *et al.*, 1994)
1199	Tryptophan	Control and extracted from A549 cells protein differences (cisplatin-lung adenocarcinoma) (Nawaz *et al.*, 2011)
	Arginine	Arginine (Zhu *et al.*, 2011)
1200	Tyrosine	Tyrosine (De Gelder *et al.*, 2007)
1201 (1202)	Tyrosine	Tyrosine solid (Zhu *et al.*, 2011)
1205	Tyrosine, phenylalanine	Bovine Albumin (Kendall, 2002) (Kendall unpublished)
	Tyrosine, phenylalanine	Amino acid spectra (HORIBA presentation)
1206	Hydroxyproline	Collagen (Frank *et al.*, 1994)
	Hydroxyproline, tyrosine	Breast tissue (Stone *et al.*, 2004)
	Hydroxyproline, tyrosine	Collagen type I (human placenta) and infiltrating duct carcinoma (Frank, McCreery and Redd, 1995)
1207	Tyrosine, phenylalanine	Bovine serum albumin (Kendall, 2002)

1208	Tryptophan, phenylalanine	Bronchial tissue (Huang, McWilliams, Lui, *et al.*, 2003)
	Ring breathing modes of the DNA/RNA bases,	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
	amide III (protein)	
1209	Tryptophan, phenylalanine	Oesophagus (Stone *et al.*, 2002)
1210	Hydroxyproline	Chicken leg bone (Keller *et al.*, 1994)
	Phenylalanine	Phenylalanine solution (Zhu *et al.*, 2011)
1214	Phenylalanine, tyrosine	Phenylalanine (Zhu *et al.*, 2011),
		Tyrosine (De Gelder *et al.*, 2007)
1215	Tyrosine	Tyrosine (Zhu *et al.*, 2011)
1216 (1618 measured)	Proline, phenylalanine	Proline, phenylalanine (De Gelder et al., 2007)
1220, 1221	Amide III (β-sheet)	Oesophagus, prostate (Stone *et al.*, 2004), (Stone *et al.*, 2002)
1224	Amide III (β-sheet)	Peak after irradiation-brain tissue (Jyothi Lakshmi *et al.*, 2002)
1237 (1236 measured too)	Amide III	Human eye lens (Keller *et al.*, 1994)
1238	Amide III	Elastin (Kendall, 2002)
1239	Amide III	Thyroid tissue, several hormones (Rau *et al.*, 2016),
		Papillary carcinoma (Teixeira *et al.*, 2009)
	Thymine	DNA, RNA (Kendall, 2002)
1240	One of the two most distinct peaks for RNA (with 813)	Slightly elevated concentration of RNA in the transformed cells versus the normal cells (Chan *et al.*, 2006)
	Amide III region of β-sheet conformations of polypeptides→ very intense peak (1235-1240)	Poly-lysine (Lippert, Tyminski and Desmeules, 1976)
	Proline	Proline (De Gelder *et al.*, 2007), (Zhu *et al.*, 2011)
1243	Amide III (collagen)	Prostate tissue (Stone *et al.*, 2004)
	Amide III	Liver (Keller *et al.*, 1994),

		Collagen (Kendall, 2002)
	Amide III	Chicken leg bone
	Amide III	Rabbit lens (Erckens *et al.*, 1997)
1245	Amide III (random coil)	Human brain tissue, glioma grade III (Mizuno *et al.*, 1994)
1246	Amide III	Protein (Gniadecka *et al.*, 1997)
1247	Amide III	Diseased breast tissue, collagen (Frank *et al.*, 1994)
		Collagen type I, infiltrating ductal carcinoma (Frank, McCreery and Redd, 1995)
	Collagen	Fibrocystic human breast tissue (Kendall, 2002) (Redd 1993)
	Amide III (collagen)	Guinea pig skin incisions (Alimova *et al.*, 2009)
	Amide III	Normal squamous cells (human cervical tissue) (Kamemoto *et al.*, 2010)
	Collagen	Collagen (Zhu *et al.*, 2011)
	Thymine, Tyrosine	Thymine, Tyrosine (De Gelder *et al.*, 2007)
1248	Amide III	Collagen (Mahadevan-Jansen *et al.*, 1998)
	Collagen	Collagen (Zhu *et al.*, 2011)
	Tyrosine	Tyrosine (Zhu *et al.*, 2011)
	Amide III (collagen)	Tendon (Gasior-Głogowska *et al.*, 2010)
1255	Cytosine	(Puppels, 1999)
	Uracil	Uracil (De Gelder *et al.*, 2007)
1256	Glycogen	Glycogen (Kendall, 2002) (Kendall unpublished)
1257	Adenine, thymine, RNA bases, amide III	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
1258	Adenine, cytosine, Amide III, Lipids	Murine muscle and adipose tissue (strong peak) (Huang *et al.*, 2011),
	Adenine, cytosine, Amide III,	Oesophagus (Stone *et al.*, 2004)
1259	Amide III	Silicone gel (Frank *et al.*, 1994)

	Guanine, cytosine	guanine–cytosine oligonucleotides (Ruiz-Chica *et al.*, 2004)
1260	Tyrosine	Protein secondary structure (Miura and Thomas, 1995)
	Amide III	Chicken leg bone (Keller *et al.*, 1994)
		Malignant breast tissue (Feld *et al.*, 1995)
	Amide III	Colon tissue (Stone *et al.*, 2004)
	Structural protein modes of tumors	Breast tissue (Feld *et al.*, 1995),
1261	Tryptophan	Tryptophan (Zhu *et al.*, 2011)
	Thymine	Thymine (De Gelder *et al.*, 2007)
	1261-1269: Amide III	Found in different states of horse heart cytochrome c (Copeland and Spiro, 1985)
1262	Amide III	Benign or normal cervix, benign or normal uterus, benign or normal endometrium (Liu *et al.*, 1992)
1263	Amide III	Collagen (Kendall, 2002)
	Tyrosine	Tyrosine (Zhu *et al.*, 2011)
	Thymine, adenine	Differences between normal cells and transformed cell lines (Chan *et al.*, 2006)
1264	Fatty acids	Normal iliac artery (Silveira *et al.*, 2002)
1265	Amide III	Normal breast tissue (Frank, McCreery and Redd, 1995), (Chen *et al.*, 2014)
	Amide III (collagen)	Bronchial tissue (fresh and fixated) (Huang, McWilliams, Lam, *et al.*, 2003)
	Amide III α-helix, collagen, tryptophan	Normal and cancer bronchial tissue (more prominent in normal) (Huang, McWilliams, Lui, *et al.*, 2003)
	Tyrosine (m)	Tyrosine (De Gelder *et al.*, 2007), (Chen *et al.*, 2014)
	Amide III	Thyroid peak (Law *et al.*, 2017)
	Amide III	Healthy and dysplastic gastric tissue *in vivo* (Huang, Bergholt, *et al.*, 2010)
1266	Proline	Proline (Zhu *et al.*, 2011)
	Amide III	Healthy gastric tissue (Chen *et al.*, 2014)

	Triacylglycerols,	Purchased compounds (Weng *et al.*, 2003)
	free fatty acids	
	α-helix of histones	Measured in normal gastric mucosa tissue (Chen *et al.*, 2014)
	Amide III	Found in intact single mitochondria (Tang *et al.*, 2007)
1267	Amide III	Lymph node, diseased breast tissue, collagen (Frank *et al.*, 1994),
		Infiltrating ductal carcinoma (breast) (Frank, McCreery and Redd, 1995)
	Phospholipid membrane	(Mahadevan-Jansen *et al.*, 1998)
	Lipids	Normal brain tissue (Malini *et al.*, 2006)
	Proline,	Proline (De Gelder *et al.*, 2007), Tyrosine solid (Zhu *et al.*, 2011)
1268	Amide III	In bovine albumin (Kendall, 2002) (Kendall unpublished)
	Phospholipids	Normal murine brain tissue (Jyothi Lakshmi *et al.*, 2002)
1269	Amide III (α-helix)	Type I collagen (human placenta) (Frank, McCreery and Redd, 1995),
		Human brain tissue (Mizuno *et al.*, 1994)
	Amide III	Cancerous gastric tissue (Chen *et al.*, 2014)
	α-helix of histones	Shifted in cancerous gastric tissue from 1266(Chen *et al.*, 2014)
1270-1310	Amide III (α-helix)	Intact muscle fibres (Pézolet *et al.*, 1980)(Pézolet *et al.*, 1988),
		(Sato *et al.*, 2014)
1300	Amide III	(Sato *et al.*, 2014)
	Phospholipids	Human brain tissue (Mizuno *et al.*, 1994)
	Lipids-fatty acids	Breast (Manoharan *et al.*, 1998)
	Fatty acids	Healthy breast tissue (Hanlon *et al.*, 2000)
	Lipids	Difference between normal skin tissue and BCC (Gniadecka *et al.*, 1997)
	Lipids	Peak found in normal brain tissue (Jyothi Lakshmi *et al.*, 2002)

	Lipids	Meningioma (Koljenović *et al.*, 2005)
1301	Lipids	Lipid (TPE), Normal human breast tissue (Kendall, 2002)
	Lipids	Present 24h after stress in brain tissue (Jyothi Lakshmi *et al.*, 2002)
	Lipids, phospholipid-intralipid, cholesterol derivatives (due to fatty acid chains)	Brain after traumatic brain injury, also in compounds (Surmacki *et al.*, 2017)
	Cholesterol, Fatty acids	Necrotic core of the atheromatous plaque (arteries), normal iliac artery (Silveira *et al.*, 2002)
	Lipids	Strong band in normal breast tissue (Chowdary *et al.*, 2009) (assignment from other papers
	Carboxylic acid salts	Thyroid peak (Law *et al.*, 2017)
1302	Collagen	Collagen Type I (Cheng *et al.*, 2005)
	Collagen, phospholipids	Normal bronchial tissue (Huang, McWilliams, Lui, *et al.*, 2003)
	Lipids, collagen	Murine stomach, small intestine, colon, bladder, lung (Huang *et al.*, 2011)
	Amide III	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
1303	Triglycerides	Adipose tissue (J. J. Baraga, Feld and Rava, 1992; Joseph J. Baraga, Feld and Rava, 1992)
	Phospholipid membrane	Phospholipid membrane (Mahadevan-Jansen *et al.*, 1998)
1304	Adenine, cytosine, lipids	Colon small peak (Stone *et al.*, 2004)
1306	Lipids	Human colon (Redd, Frank, *et al.*, 1993)
1307	Adenine	Adenine (De Gelder *et al.*, 2007)
	Lipids, collagen	Thyroid healthy and carcinoma tissues (Rau *et al.*, 2016), (Talari *et al.*, 2015)
	Lipids	Malaria infected and non-infected spleen (Frame *et al.*, 2018)
	Collagen, lipids	Soft mass pilomaxitroma (probably shifted collagen peak) (Cheng *et al.*, 2005)
1308	Phenylalanine	Phenylalanine (De Gelder *et al.*, 2007)

	Alanine	Alanine (Zhu *et al.*, 2011)
1309	Lipids	Normal human skin (Gniadecka *et al.*, 1997)
	Tryptophan, glutathione	Tryptophan, glutathione (De Gelder *et al.*, 2007)
	Collagen, lipids	Hard mass pilomaxitroma (probably shifted collagen peak) (Cheng *et al.*, 2005)
1310	Phenylalanine	Phenylalanine (Zhu *et al.*, 2011)
1311	Amide III	Staphylococcal protein A (Kengne-Momo *et al.*, 2012)
1314	Collagen	Difference between normal and cancerous nasopharyngeal tissue (Lau *et al.*, 2003)
	Cytochrome c	Cytosol (Okada *et al.*, 2012)
	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
1315	Tryptophan	Tryptophan (Zhu *et al.*, 2011)
1316	Guanine	DNA, RNA (Mahadevan-Jansen *et al.*, 1998)
1317	Proline	Proline (De Gelder *et al.*, 2007)
	Valine	Valine (De Gelder *et al.*, 2007)
	Typical Phospholipids	Normal oral tissue (Malini *et al.*, 2006)
1330	Tryptophan	Cervix, uterus, ovary (Liu *et al.*, 1992)
1331	Valine	Valine solid (Zhu *et al.*, 2011)
	α-helix	Poly-L-alanine (Koenig and Sutton, 1969)
1332	Adenine	Adenine (De Gelder *et al.*, 2007)
1333	Glycogen	Glycogen (Kendall, 2002) (Kendall unpublished)
	Guanine	DNA (Mahadevan-Jansen *et al.*, 1998)
	Proline	Proline (De Gelder *et al.*, 2007)
	Guanine	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
1334	Glutathione	Glutathione (De Gelder *et al.*, 2007)
1335	Adenine	DNA, RNA (Mahadevan-Jansen *et al.*, 1998)
	Collagen, polynucleotide chain	Oesophagus (Stone *et al.*, 2002)

	Collagen, nucleic acids	Present only in abnormal bronchial tissue (adenocarcinoma)
	Collagen, nucleic acids	Murine whole blood, blood pellete, serum, stomach, small intestine colon, lung brain, kidney, liver, spleen, teeth skull (Huang *et al.*, 2011)
	Proline	Proline (Zhu *et al.*, 2011)
	Collagen, nucleic acids	Adenocarcinoma in pig kidney tissue (Lykina *et al.*, 2018), (Talari *et al.*, 2015)
	No assignment	Found in human cancerous nasopharyngeal tissue (*in vivo*) (Ming *et al.*, 2017)
1336	Polynucleotide chain	Colon (Stone *et al.*, 2004)
	Phenylalanine	Phenylalanine (De Gelder *et al.*, 2007)
	Valine	Valine solution (Zhu *et al.*, 2011)
	Adenine, guanine (ring breathing modes in the DNA bases), proteins	Differences between normal cells and transformed cell lines (neoplastic) (Chan *et al.*, 2006)
	Collagen, nucleic acid, tryptophan	Hard mass pilomaxitroma (absent in normal skin dermis) (Cheng *et al.*, 2005)
1338	Tryptophan	Amino acid (Mahadevan-Jansen *et al.*, 1998)
	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
	Amide III (α-helix)	Staphylococcal protein A (Kengne-Momo *et al.*, 2012)
	Adenine	Lymphoma cells (Lin *et al.*, 2011)
1339	Adenine	DNA (Mahadevan-Jansen *et al.*, 1998)
	Collagen, nucleic acid, tryptophan	Soft mass PMX (absent in normal skin dermis) (Cheng *et al.*, 2005)
	Albumin	Albumin, human brain tissue spectra (brain tissue injury) (Surmacki *et al.*, 2017)
	Adenine	B DNA fibre (Thomas *et al.*, 1995)
1340	Tryptophan	Amino acid (Erckens *et al.*, 1997)
	Nucleic acid	Increased in colon cancerous tissue (Mahadevan-Jansen and Richards-Kortum, 1997)
	Tryptophan	Tryptophan solid (Zhu *et al.*, 2011)

1341	Adenine, guanine	Stronger in malignant oral tissue (Su *et al.*, 2012)
1343	Collagen	Type I collagen (Frank, McCreery and Redd, 1995)
1365	Guanine, tryptophan	(Hartman, Clayton and Thomas, 1973)
	Tryptophan	Hard mass PMX (Cheng *et al.*, 2005)
1366	Tyrosine	Tyrosine (Zhu *et al.*, 2011)
1367	Phospholipids	Brain tissue before irradiation (Jyothi Lakshmi *et al.*, 2002)
1368	Glutathione	Glutathione (De Gelder *et al.*, 2007)
1369	Thymine	Thymine (De Gelder *et al.*, 2007)
1370	Saccharide band	Brain tissue (Krafft *et al.*, 2005)
1371	Adenine, palmitic acid	Adenine, palmitic acid (De Gelder *et al.*, 2007)
1373	Thymine, adenine, guanine	Differences between normal and neoplastic cells (Chan *et al.*, 2006)
1396	Valine	Valine (Zhu *et al.*, 2011)
1398	Valine	Valine (De Gelder *et al.*, 2007)
1400	Uracil, adenine	(Hartman, Clayton and Thomas, 1973)
1401	Symmetric bending in proteins	(Kendall, 2002)
1403	Glutathione	Glutathione (De Gelder *et al.*, 2007)
1408	Thymine	Thymine (De Gelder *et al.*, 2007)
	Histidine	Histidine
	Proteins	Quiescent cells cultivated for 14 to 100 days (Eberhardt *et al.*, 2018)
1409	Alanine	Alanine (De Gelder *et al.*, 2007)
1410	Proline	Proline (De Gelder *et al.*, 2007)
1413	Phenylalanine	Phenylalanine solid (Zhu *et al.*, 2011)
1418 (1417 measured too)	Uracil	Uracil (De Gelder *et al.*, 2007)
	Proline	Proline (De Gelder *et al.*, 2007)
1419	Adenine	Adenine (De Gelder *et al.*, 2007)
1420-1450	Lipids	Difference between normal and cancerous skin (Gniadecka *et al.*, 1997)

1421	Palmitic acid	Palmitic acid (De Gelder *et al.*, 2007)
	Adenine, guanine breathing modes of DNA, RNA bases	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
	Adenine, guanine	(Puppels *et al.*, 1991)
1422	Deoxyribose	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
	Guanine	Guanine (De Gelder *et al.*, 2007)
1423	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
1424	Lipids, deoxyribose	Only in healthy thyroid tissue (Rau *et al.*, 2016)
1439	Lipids, proteins	Human brain tissue (Mizuno *et al.*, 1994)
1440	Cholesterol, phospholipids fatty acid	Measured in cholesterol, glyceryl tripalmitate (Krafft *et al.*, 2005)
1441	Lipids	Predicting laryngeal tissue pathology (Lau *et al.*, 2005)
	Cholesterol	atheromatous plaque in coronary artery (Silveira *et al.*, 2002)
1442	Lipids	Normal human breast tissue (Redd, Feng, *et al.*, 1993)
	Lipids, fatty acids	Normal breast tissue (Manoharan *et al.*, 1998)
	Fatty acids	Normal breast tissue (Hanlon *et al.*, 2000)
	Fatty acids, lipids, proteins	Murine skin, stomach, small intestine, colon, bladder, lung (strong peaks) (Huang *et al.*, 2011)
	Triglycerides, fatty acids	Triglycerides, fatty acids, brain tissue (Krafft *et al.*, 2005)
1443	Elastin	Elastin (Kendall, 2002) (Kendall unpublished)
	Lipids, proteins	Colon, prostate tissue (Stone *et al.*, 2004)
	Proline	Proline (De Gelder *et al.*, 2007)
1444	Lipid	Lipid (Redd, Feng, *et al.*, 1993)
	Triglycerides	Adipose tissue (human aorta) (J. J. Baraga, Feld and Rava, 1992; Joseph J. Baraga, Feld and Rava, 1992)

1445	Deformation in Collagen	Collagen (Kendall, 2002) (Kendall unpublished
	Collagen, phospholipids	Present in healthy bronchial tissue (increased), SCC and adenocarcinoma (Huang, McWilliams, Lui, *et al.*, 2003)
	Protein, lipids	Increased in cancerous nasopharyngeal tissue (Lau *et al.*, 2003)
	Collagen, phospholipids	Fresh and fixed bronchial tissue (Huang, McWilliams, Lam, *et al.*, 2003)
	Collagen, phospholipids	Found in healthy thyroid tissue (Rau *et al.*, 2016), (Talari *et al.*, 2015)
1446	Proteins, lipids	Oesophagus tissue (Stone *et al.*, 2002)
1447	Proteins, lipids	Decreased in frozen parenchymal placenta tissue (Ó Faoláin *et al.*, 2005)
	Phenylalanine	Phenylalanine (De Gelder *et al.*, 2007)
1448	Deoxyribose	DNA (Kendall, 2002) (Mahadevan-Janson 1998)
	Collagen	More prominent in human bronchial cancerous tissue (Kaminaka *et al.*, 2001)
	Phenylalanine	Phenylalanine (Zhu *et al.*, 2011)
1460	Lipids, collagen	Collagen Type I (Cheng *et al.*, 2005)
	Proline, cytosine	Proline (Zhu *et al.*, 2011), Cytosine (De Gelder *et al.*, 2007)
	Deoxyribose	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
1462	Deoxyribose	DNA (Mahadevan-Jansen *et al.*, 1998)
	Adenine	Adenine (De Gelder *et al.*, 2007)
1463	Tryptophan	Tryptophan (Zhu *et al.*, 2011)
1465	Palmitic acid	Palmitic acid (De Gelder *et al.*, 2007)
1468	Guanine	Guanine (Zhu *et al.*, 2011)
1480	DNA, vibration of purine bases	(Kendall, 2002)
1481	Palmitic acid	Palmitic acid (De Gelder *et al.*, 2007)
1482	Adenine	Adenine (De Gelder *et al.*, 2007)
1485	Nucleic acids	Colon (Boustany *et al.*, 1999)

	(Adenine, guanine)	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
1486	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
1487	Nucleic acid	Colon (Kendall, 2002)
1510	Cytosine	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
	Adenine, guanine	(Puppels *et al.*, 1991)
1548	Tryptophan	Oesophagus (Stone *et al.*, 2004)
1570	Vibration of purine bases	DNA (Mahadevan-Jansen, 1996)
	Nucleotides	Breast cancer (mouse) (Kendall, 2002) (Schrader 1995)
1573	Guanine, adenine, TRP protein	Oesophagus (Stone *et al.*, 2004)
1575	Guanine, adenine	Differences between normal cells and neoplastic cell lines (Chan *et al.*, 2006)
1576	Nucleic acids	Increased in colon carcinoma (Feld *et al.*, 1995), (Mahadevan-Jansen and Richards-Kortum, 1997)
	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
	Guanine	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
1577	Bound and free NADH	NADH (Malini *et al.*, 2006)
	Guanine, adenine	(Puppels *et al.*, 1991)
1578	Guanine	Oligonucleotide solution (Ruiz-Chica *et al.*, 2004)
1578	Phenylalanine	Cell lines [49]
	Nucleotides, DNA	Breast cancer (mouse, cell culture) (Kendall, 2002)
1579	Tryptophan	Tryptophan (Zhu *et al.*, 2011)
1602	Phenylalanine	Increased in abnormal bronchial tissue (Huang, McWilliams, Lui, *et al.*, 2003)
	Phenylalanine	Phenylalanine (De Gelder *et al.*, 2007)
1603	Phenylalanine, tyrosine	(Stone *et al.*, 2004)
1610	Tyrosine	Colon mucosa (Boustany *et al.*, 1999)

	Cytosine	Oligonucleotide solutions (Ruiz-Chica *et al.*, 2004)
1612	Adenine	Adenine (De Gelder *et al.*, 2007)
1615	Tyrosine	Tyrosine (Zhu *et al.*, 2011)
	Tyrosine, tryptophan, protein	Differences between normal cells and transformed cell lines (neoplastic) (Chan *et al.*, 2006)
1616	Tyrosine	Bovine Serum Albumin (Kendall, 2002)
	Tyrosine, tryptophan	Oesophagus (Stone *et al.*, 2002)
	Tryptophan	Tryptophan (De Gelder *et al.*, 2007)
1617	Tyrosine, phenylalanine	(Puppels *et al.*, 1991)
1618	Tryptophan	Present only in abnormal bronchial tissue (Huang, McWilliams, Lui, *et al.*, 2003)
	Free and bound NADH	Normal oral tissue (Malini *et al.*, 2006)
1619	Valine	Valine (De Gelder *et al.*, 2007)
	Tryptophan	Tryptophan (Zhu *et al.*, 2011)
1620	Tryptophan, tyrosine, phenylalanine, uracil	(Hartman, Clayton and Thomas, 1973)
	Tryptophan	Colon mucosa (Kendall, 2002)
1633	Valine	Valine (De Gelder *et al.*, 2007)
1637	Amide I	Collagen, Diseased breast tissue (Frank *et al.*, 1994)
	Amide I (α-helix, β-sheet)	Brain tissue (Jyothi Lakshmi *et al.*, 2002)
1650	Amide I (α-helix)	Normal human aorta, bovine insulin, human artery (Kendall, 2002)
	Lipids	Colon mucosa (Kendall, 2002)
	Typical Phospholipids, Amide I	Normal, Malignant oral tissue (Malini *et al.*, 2006)
1651	Amide I	Breast tissue and malignant tumors, healthy human skin, collagen, ovarian cancer (Kendall, 2002)
1652	Lipids	Breast tissue, Lipid (TPE) (Kendall, 2002)
	Amide I (α-helix)	a-Poly-L-glutamate, human strateum corneum (Kendall, 2002)

	1652-1653 Lipid	Breast (Stone *et al.*, 2004)
1653	Cytosine	Cytosine (De Gelder *et al.*, 2007)
	Lipids	Normal human breast tissue, Human Colon (Redd 1993)
	Amide I	Bovine Albumin (Kendall unpublished)
1654	Amide I	Normal breast tissue (Frank, McCreery and Redd, 1995)
1655	Thymine	Thymine (De Gelder *et al.*, 2007)
	Amide I (α-helix)	Coronary artery, lymph node, endometrial cancer, human strateum corneum (Kendall, 2002)
	Amide I (Collagen)	Collagen type I (Cheng *et al.*, 2005)
	Lipids	Normal oral tissue (Malini *et al.*, 2006)
	Collagen, Elastin,	Abnormal bronchial tissue (Huang, McWilliams, Lui, *et al.*, 2003)
1656	Lipids, amide proteins, phospholipids	Control molecules, brain tissue (Jyothi Lakshmi *et al.*, 2002)
1657	Amide I	Malignant breast tissue (Feld *et al.*, 1995), diseased breast tissue, infiltrating duct carcinoma (Frank, McCreery and Redd, 1995)
	Lipids-Fatty acids	Colon (Manoharan *et al.*, 1998),
		Normal breast tissue (Hanlon *et al.*, 2000)
	Triglycerides	Atherosclerotic samples (Silveira *et al.*, 2002)
1658	Amide I	Aorta (J. J. Baraga, Feld and Rava, 1992; Joseph J. Baraga, Feld and Rava, 1992)
	Amide I (α-helix)	Brain tissue (Jyothi Lakshmi *et al.*, 2002)
1659	Amide I, collagen	Fresh bronchial tissue, increased in cancerous tissue (Min *et al.*, 2005)
	Cholesterol	Atherosclerotic samples (Silveira *et al.*, 2002)
	Amide I	Benign breast tumor, benign or normal ovary, cervix, uterus, uterus cancer, human brain tissue, bovine serum albumin (Kendall, 2002)

1661	Amide I	Elastin (Kendall unpublished), Normal human skin (Gniadecka *et al.*, 1997)
1662	Nucleic acids	Colon adenocarcinoma (Feld *et al.*, 1995)
	Triglycerides	Adipose tissue from human aorta (J. J. Baraga, Feld and Rava, 1992; Joseph J. Baraga, Feld and Rava, 1992)
1664	Amide I	Collagen (Huang, McWilliams, Lui, *et al.*, 2003)
1665	Amide I	Skin dermis, soft PMX (Huang, McWilliams, Lui, *et al.*, 2003)
1666	Collagen	Bronchial tissue, strongest in cancerous tissue (Kaminaka *et al.*, 2001)
1668	Collagen	Collagen (Zhu *et al.*, 2011)

Appendix B

Raman spectra associated with skeletal muscle tissue.

Raman bands in cm^{-1}

Myosin: 622, 645, 704, 755, 780, 830, 855, 901, 940, 962 (shoulder), 1004, 1033 (1044), 1081, 1104, 1128, 1160, 1175, 1209, 1244, 1265 (shoulder), 1304, 1320, 1342, 1402, 1423, 1451, 1587, 1607, 1650 (Carew, Asher and Stanley, 1975)

Myosin: 620, 640, 750, 828, 903, 938, 1003, 1035, 1077, 1101, 1128, 1155, 1172, 1205, 1246, 1318, 1340, 1405, 1427, 1449, 1555, 1650 (Kishp and Noda, 1983)

Tropomyosin: 620, 644, 716, 750, 906, 940, 983, 1006, 1054, 1087, 1106, 1121, 1156, 1172, 1211, 1253, 1293, 1325, 1342, 1425, 1451, 1460, 1625, 1655 (Frushour and Koenig, 1974)

Actin: 1000, 1070, 1445, 1650 (acquired from figure) (Huang, Teh, *et al.*, 2010)

Type I Collagen: 762, 814, 856, 872, 920, 936, 1004, 1030,1240, 1264, 1335, 1445, 1651 (Kendall, 2002)

Other Collagen measurements:

Type I Collagen: 922, 1004, 1032, 1247, 1273, 1325, 1410, 1455, 1668 (Zhu *et al.*, 2011)

Type I Collagen: 644, 763, 822, 859, 879, 925, 942, 1009, 1084, 1171, 1250, 1278, 1459, 1682 (Cárcamo *et al.*, 2012)

Type III Collagen: 762, 819, 858, 896, 924, 941, 1008, 1166, 1251, 1276, 1307, 1451, 1674 (Cárcamo *et al.*, 2012)

Glycogen: 709, 761, 853, 937, 1048, 1083,1123, 1258, 1333, 1377,1455 (Kendall, 2002)

Intact Muscle Fibres: 623, 642, 759, 828, 854, 900,937, 980, 1002, 1045, 1128, 1158, 1207, 1328, 1417, 1450, 1520, 1556, 1586, 1650 (Pézolet *et al.*, 1980)

Skeletal Muscle: 851, 962, 1065, 1258, 1297, 1437, 1542, 1653, 1737 (Huang *et al.*, 2011)

Skeletal Muscle: 746, 1000, 1124, 1309, 1333, 1450, 1550, 1581, 1650 (Minamikawa, Harada and Takamatsu, 2015)

Skeletal Muscle: 934, 1002, 1179, 1264, 1301, 1336, 1447, 1576, 1655, 1737(Al-Rifai *et al.*, 2019)

Appendix C

Animal Experiments

Study 1

a) One month old *mdx* vs. C57Bl/10

PCA analysis

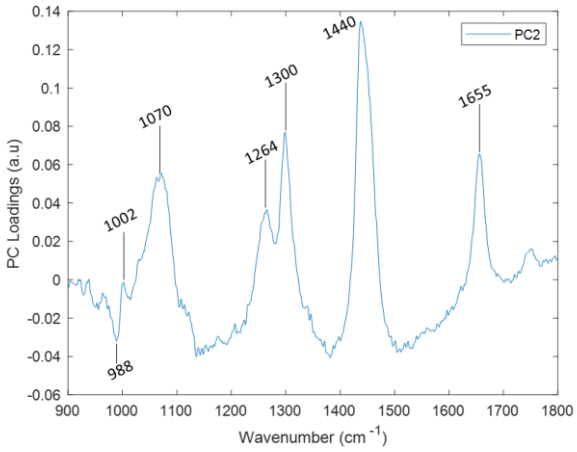


Figure C-1: PC2 loading plot.

PLS-DA analysis

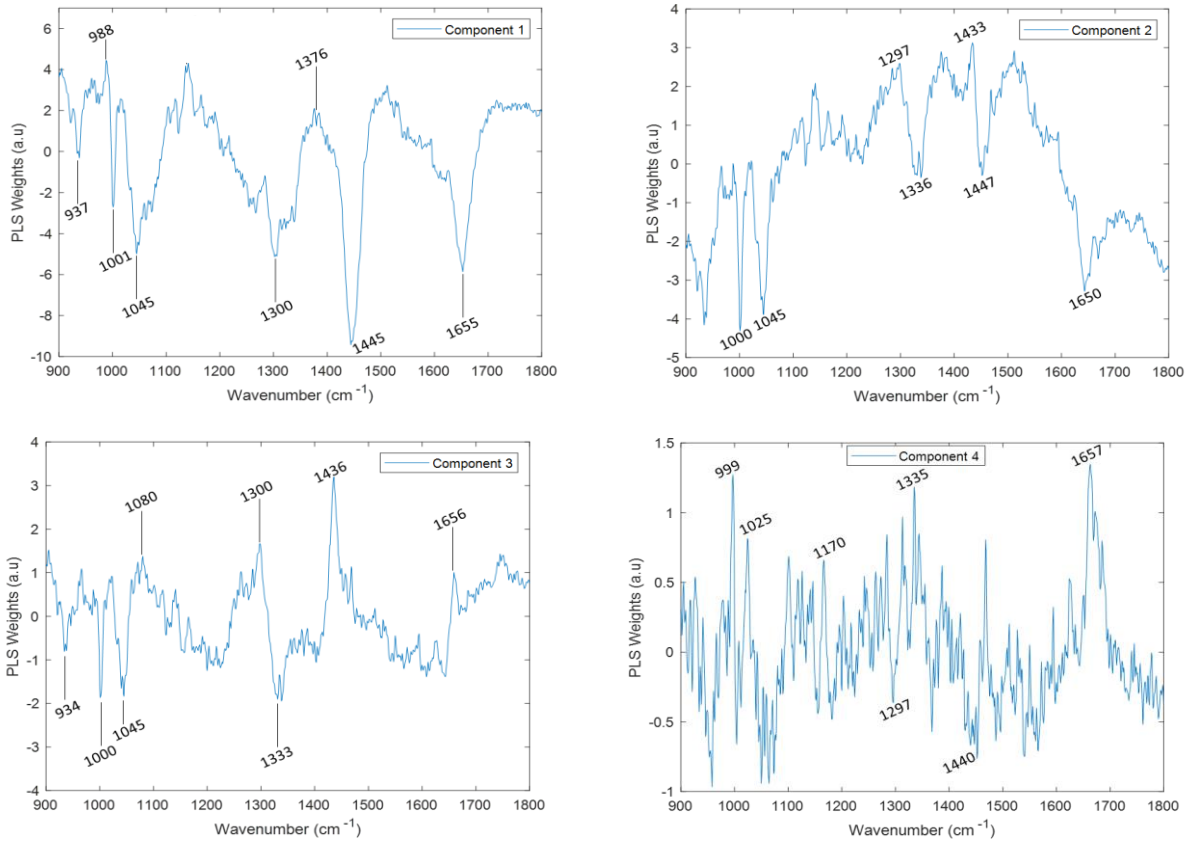


Figure C-2: PLS Component 1,2,3 and 4 weight plots.

b) 3 months old *mdx*/C57Bl/10

PLS-DA analysis

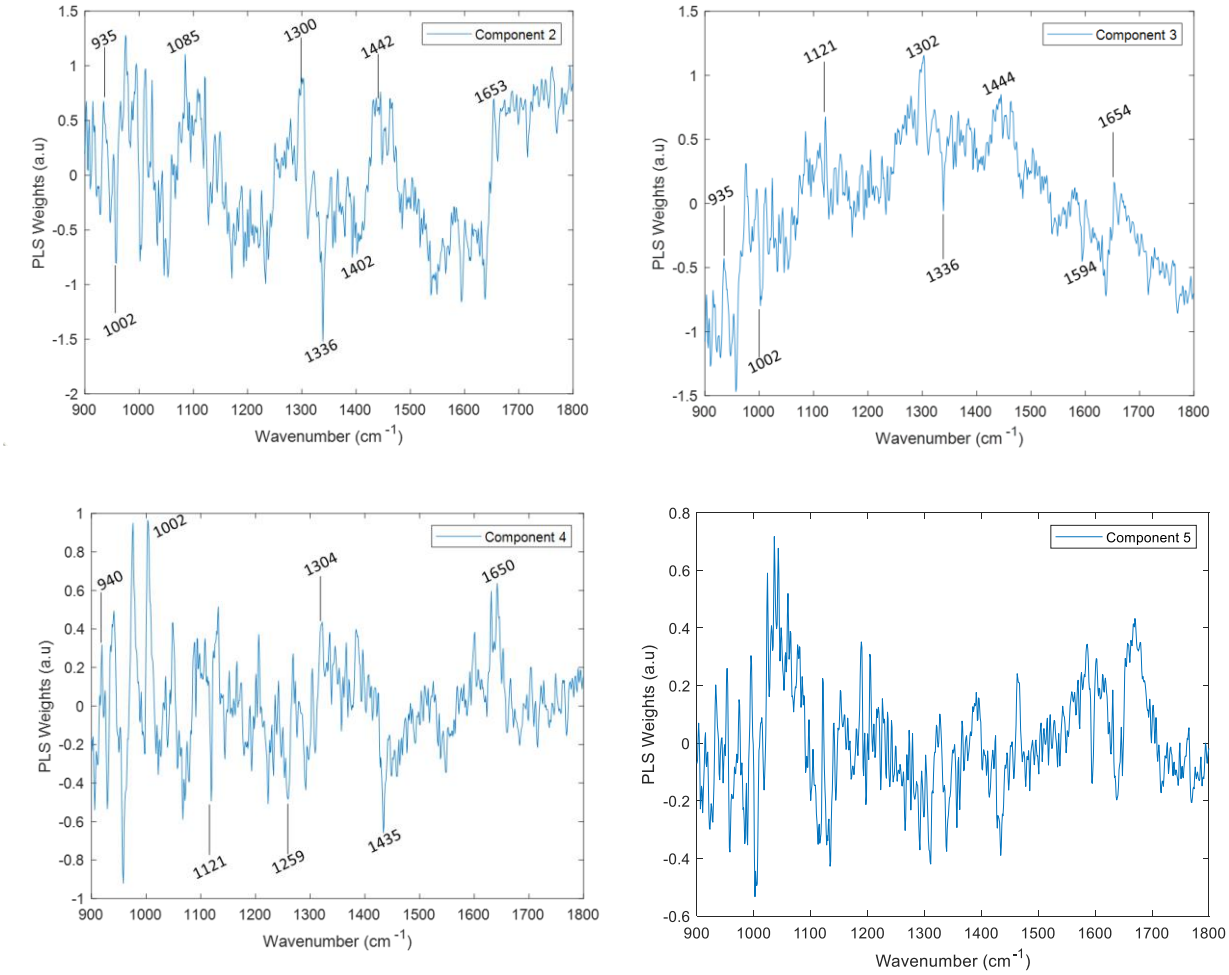


Figure C-3: PLS Component 2,3,4 and 5 weight plots.

c) One/Three months old *mdx* mice

PCA analysis

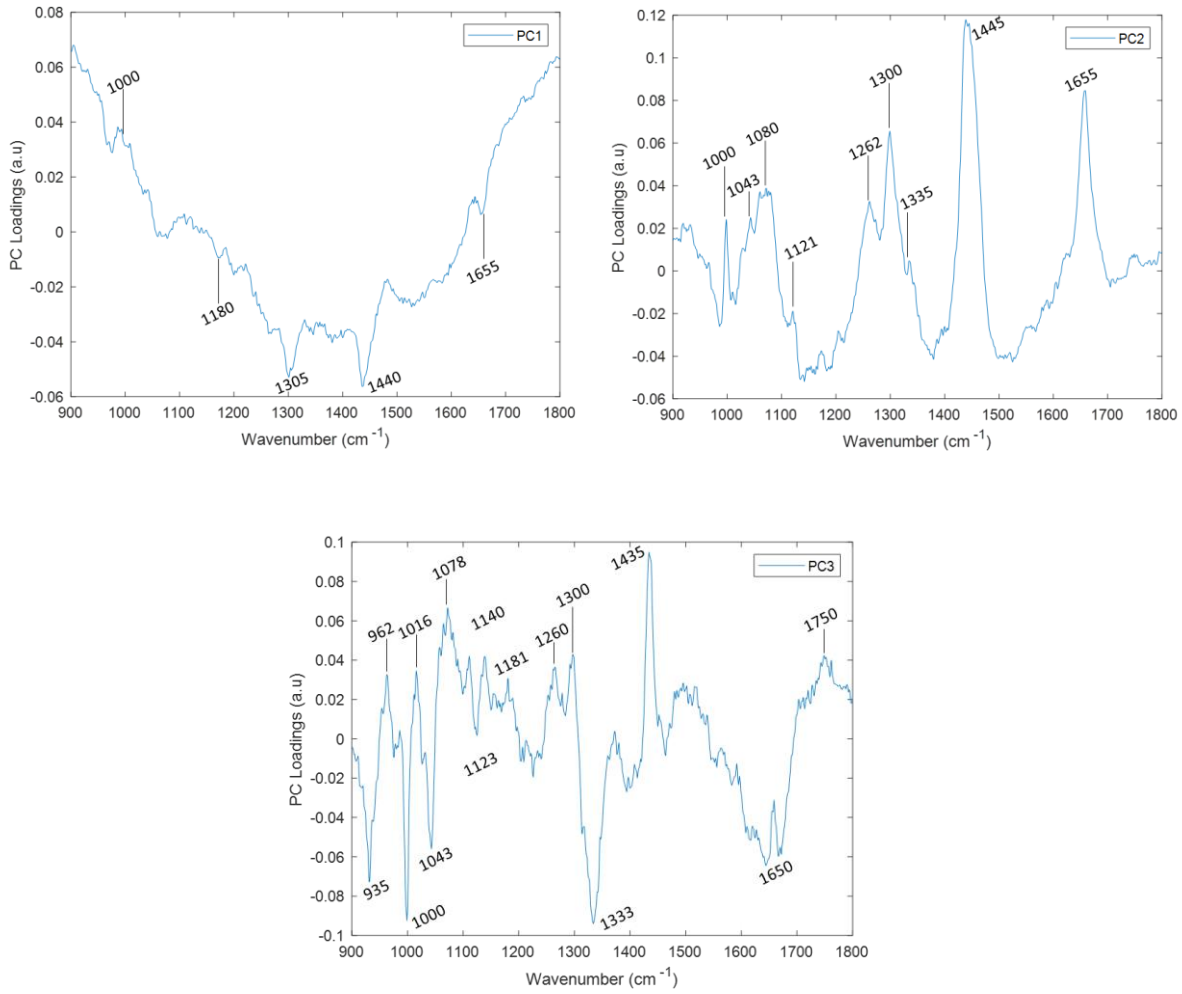


Figure C-4: PC1, PC2 and PC3 loading plots.

PLS-DA analysis

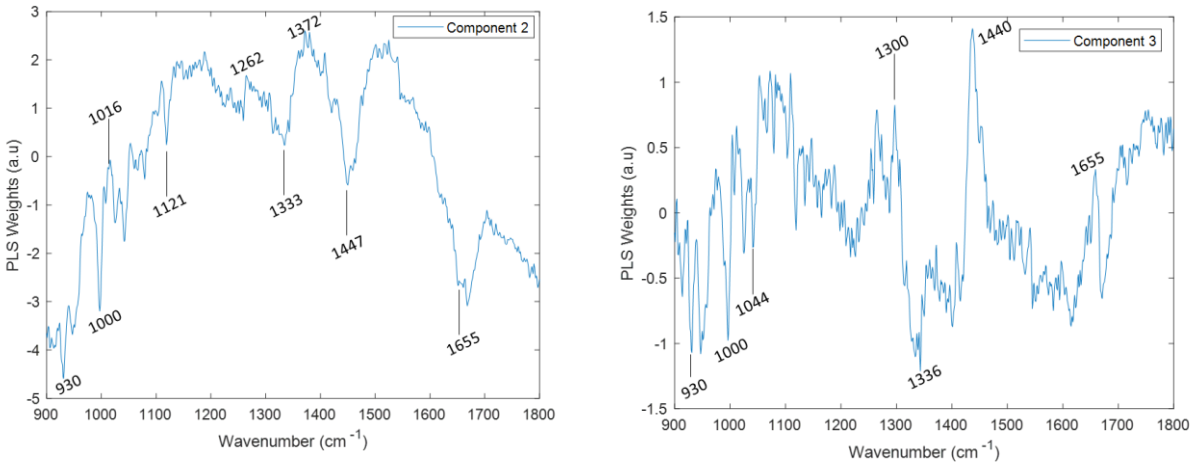


Figure C-5: PLS Component 2 and 3 weight plots.

d) Three months old SOD1^{G93A}/C57Bl/6

PCA Analysis

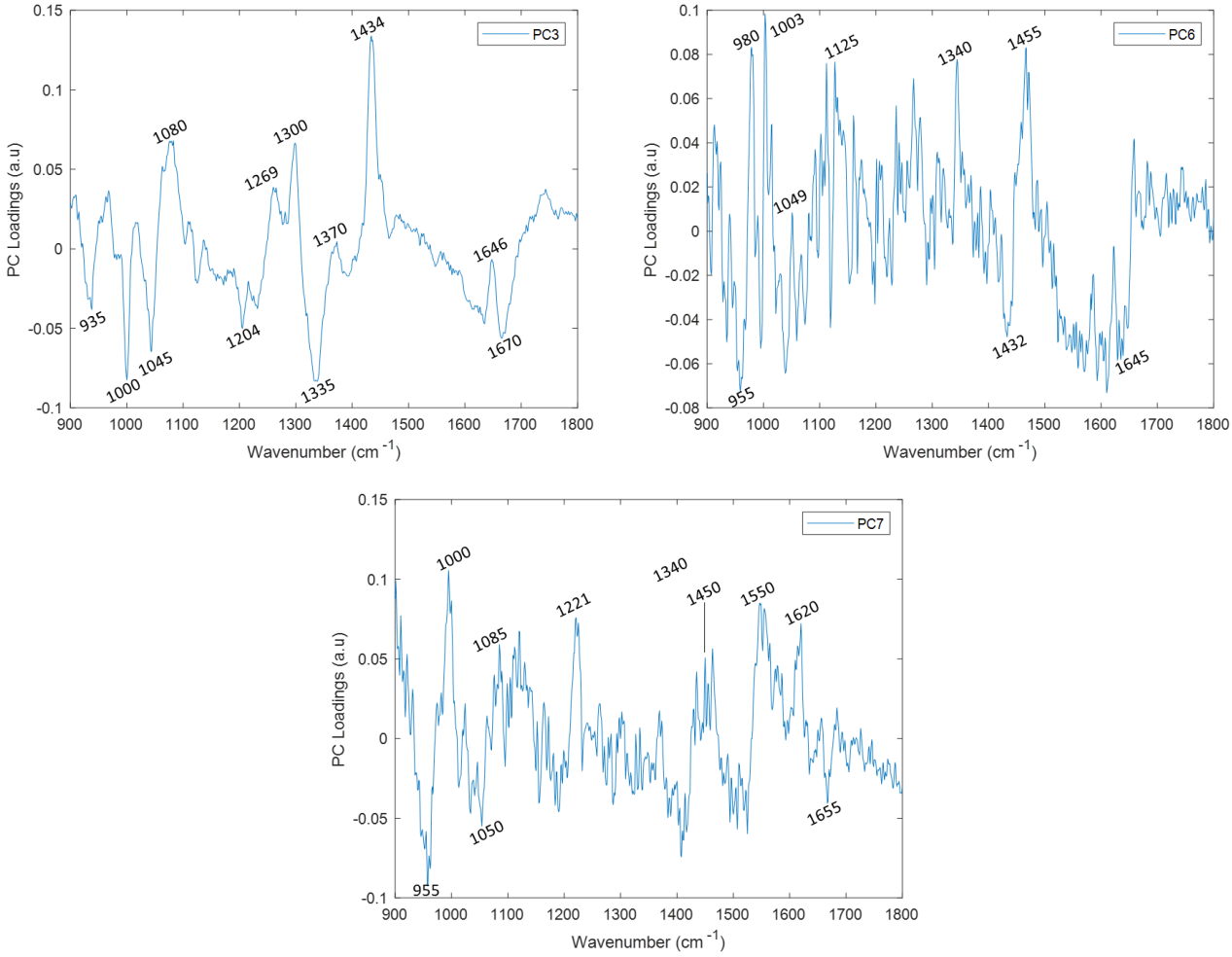


Figure C-6: PC3, PC6 and PC7 loading plots.

PLS-DA Analysis

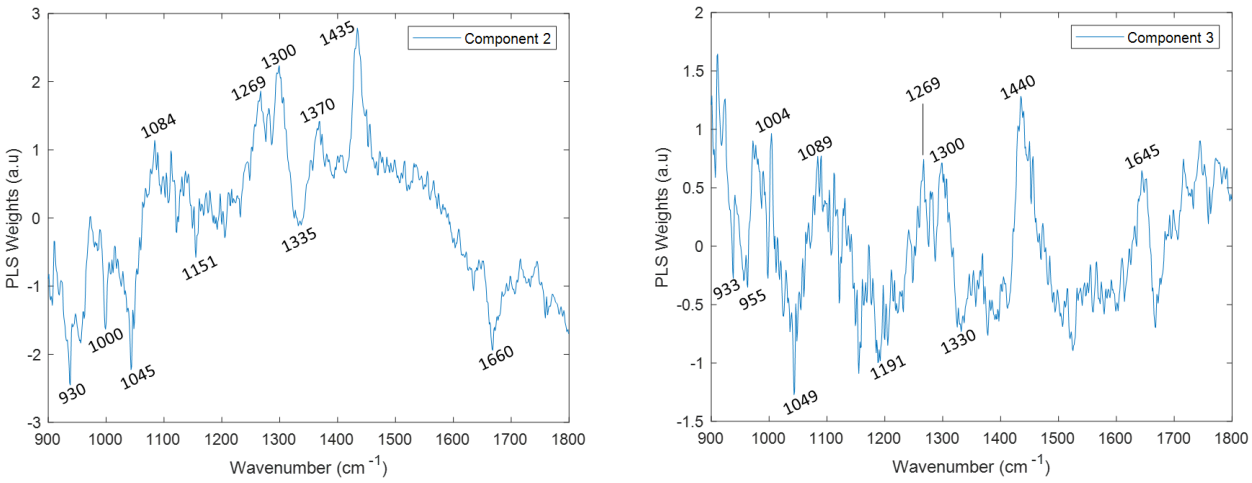


Figure C-7: PLS Component 2 and 3 weight plots.

e) One/Three months SOD1^{G93A} mice

PCA Analysis

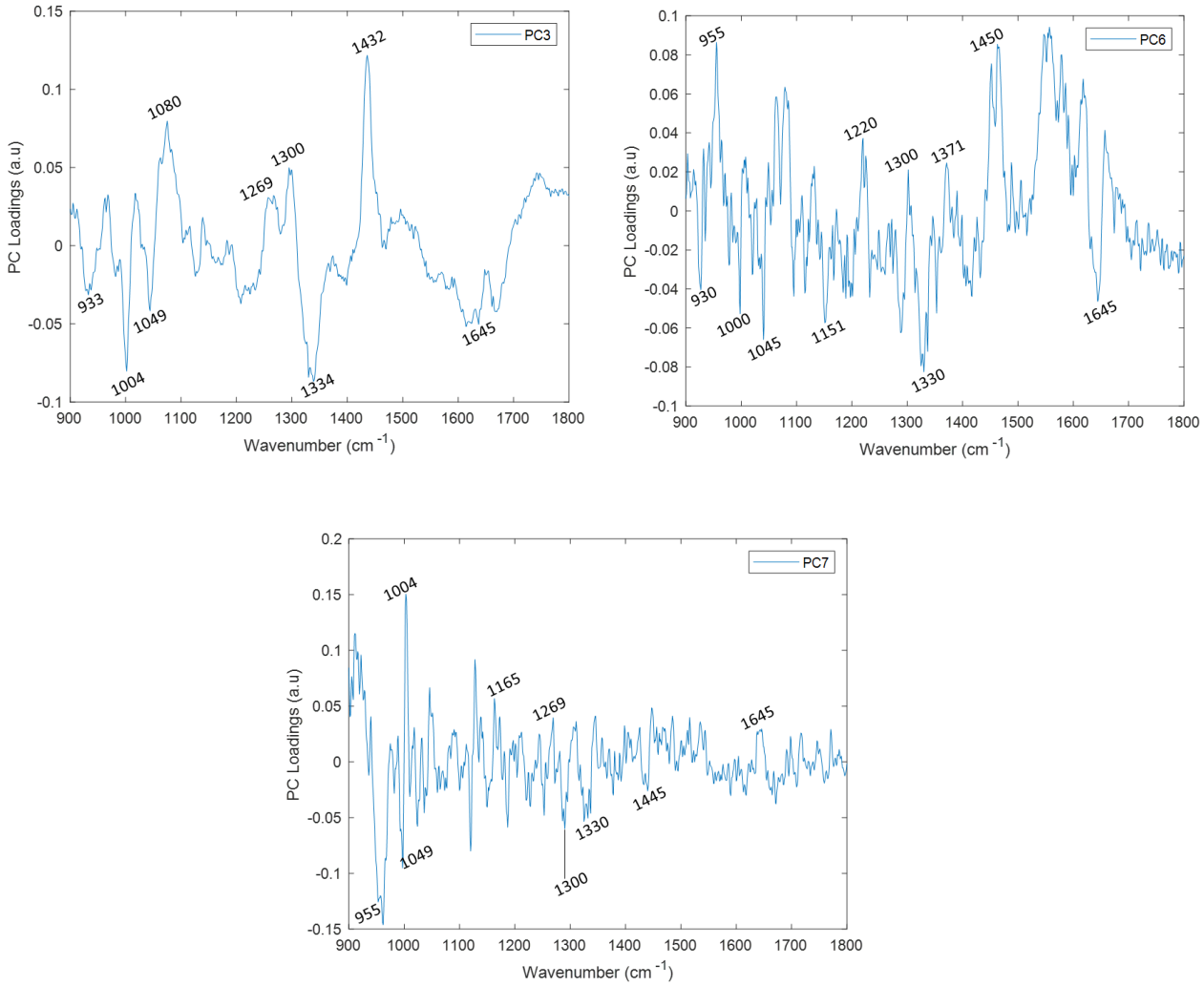


Figure C-8: PC3, PC6 and PC7 loading plots.

PLS-DA Analysis

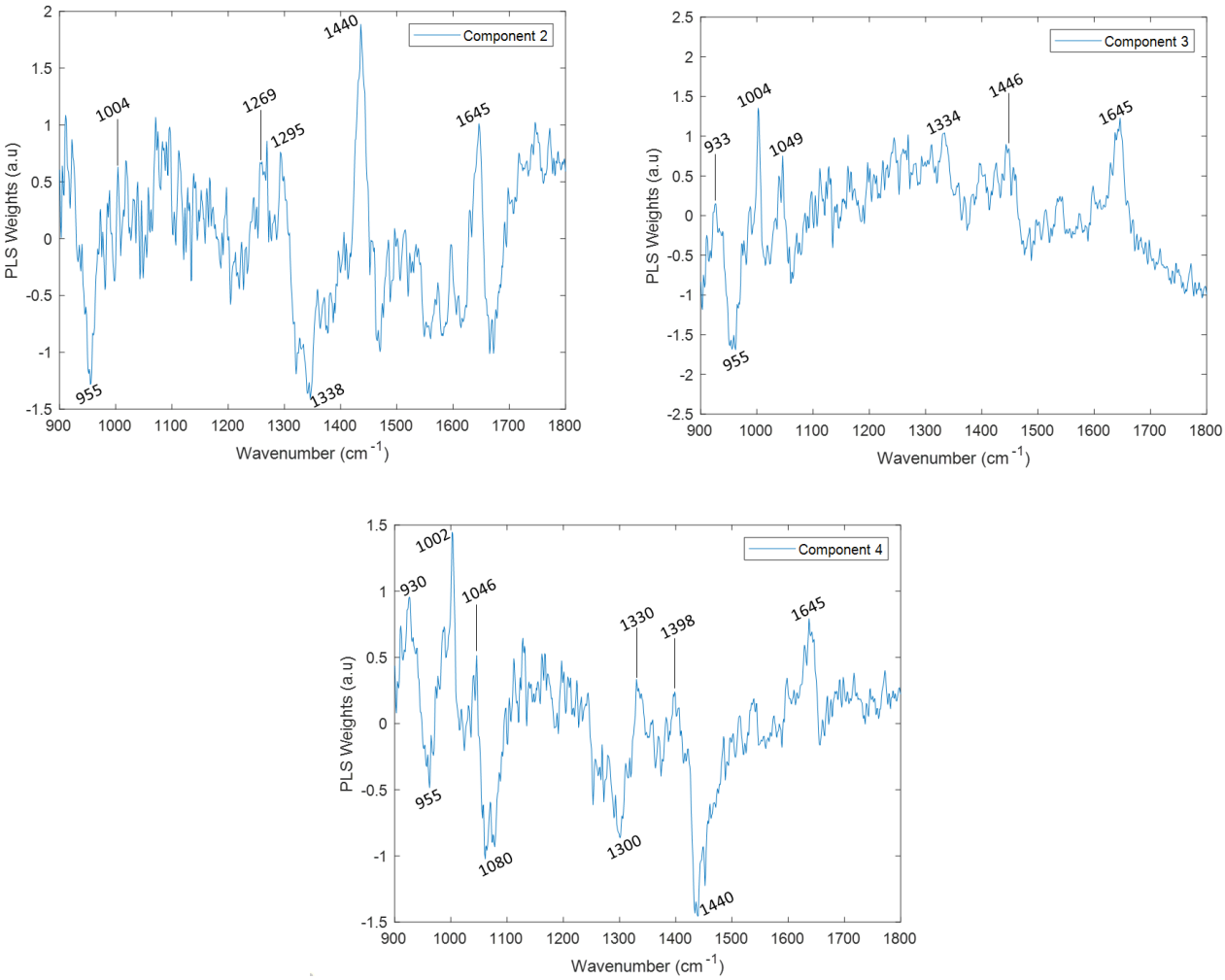


Figure C-9: PLS Component 2 and 3 weight plots.

f) One month old *mdx*/SOD1^{G93A}

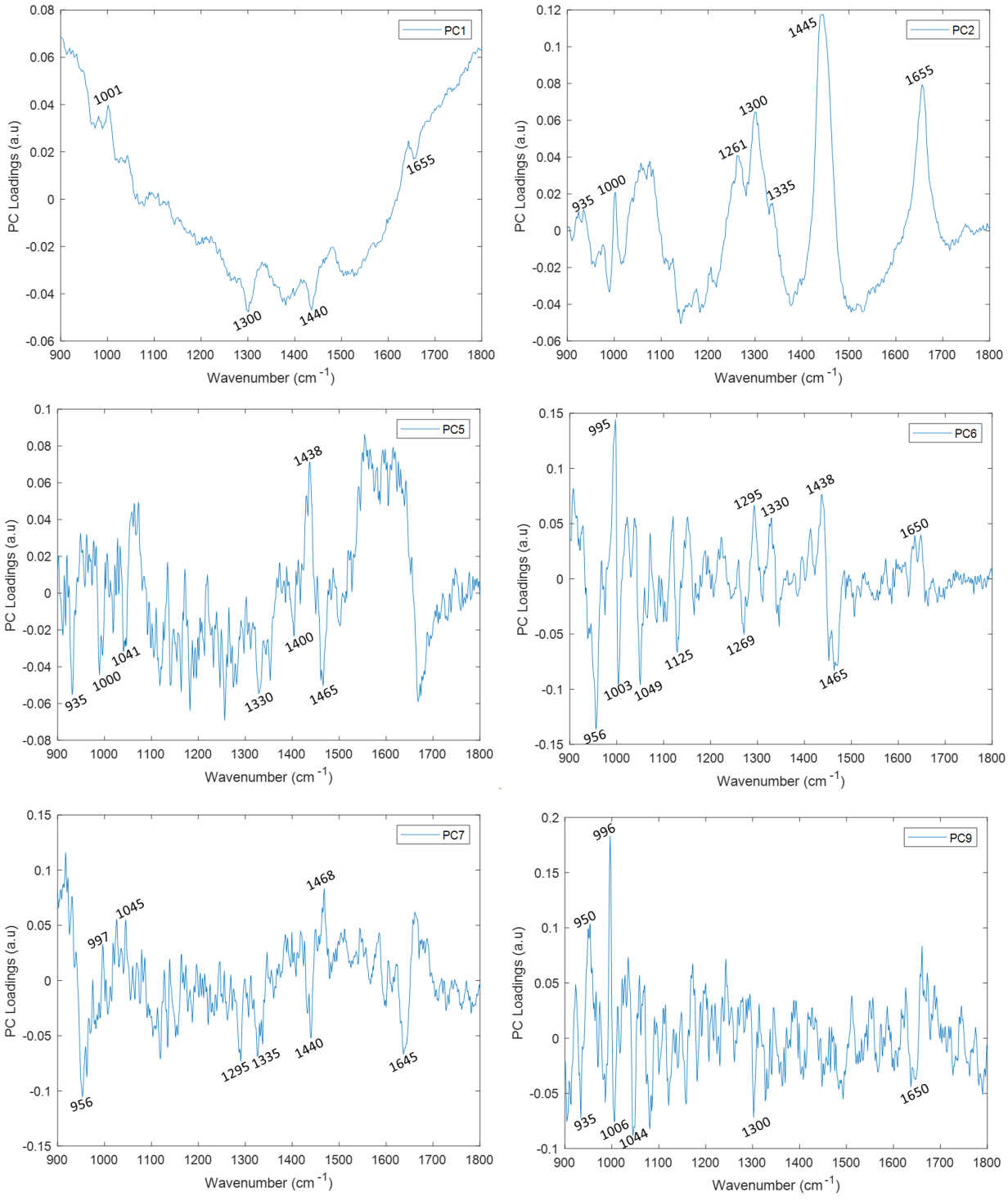


Figure C-10: PC1, PC3, PC5, PC6, PC7 and PC9 loading plots

PLS-DA Analysis

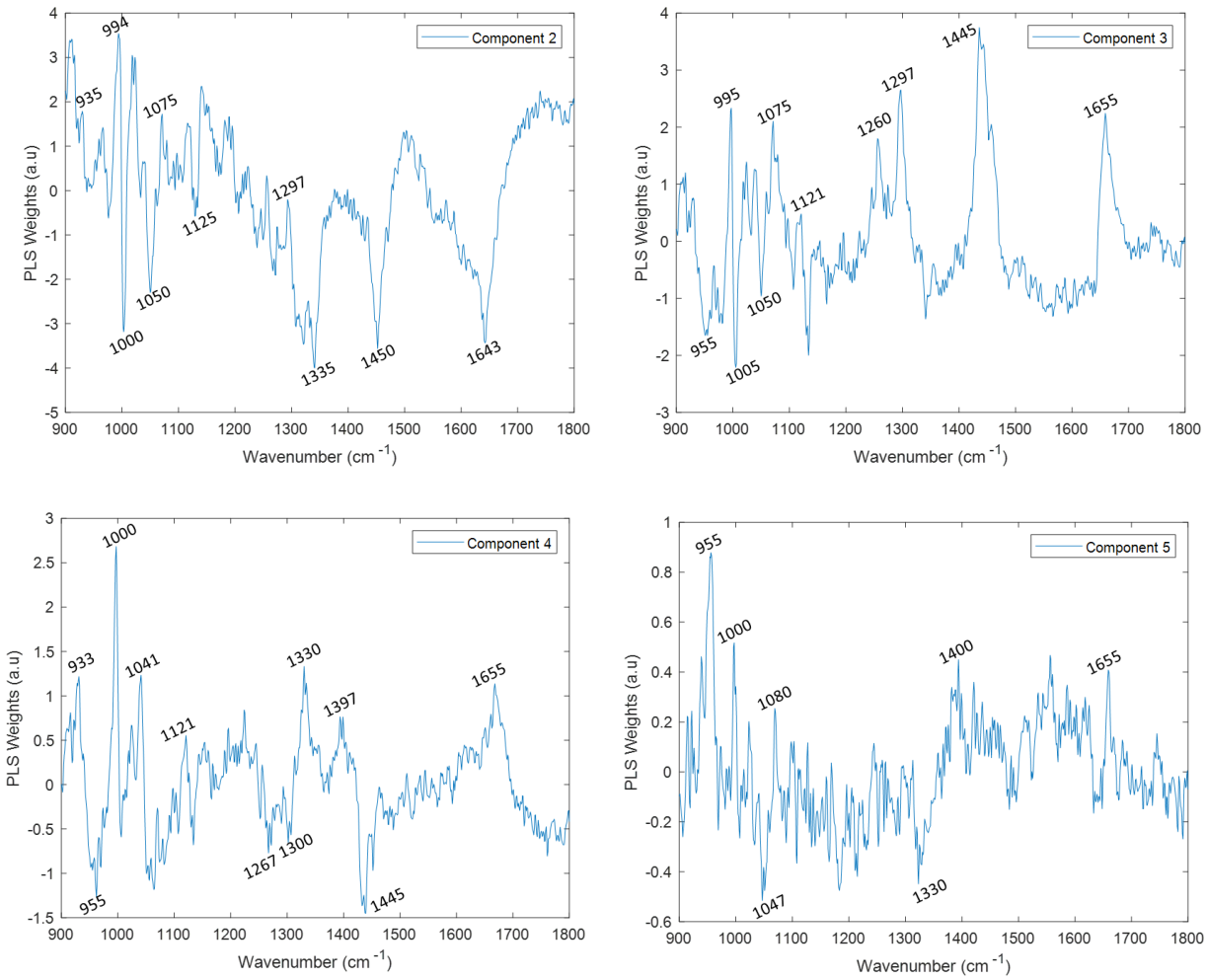


Figure C-11: PLS Component 2,3,4 and 5 weight plots.

g) Three months old *mdx*/SOD1^{G93A}

PCA Analysis

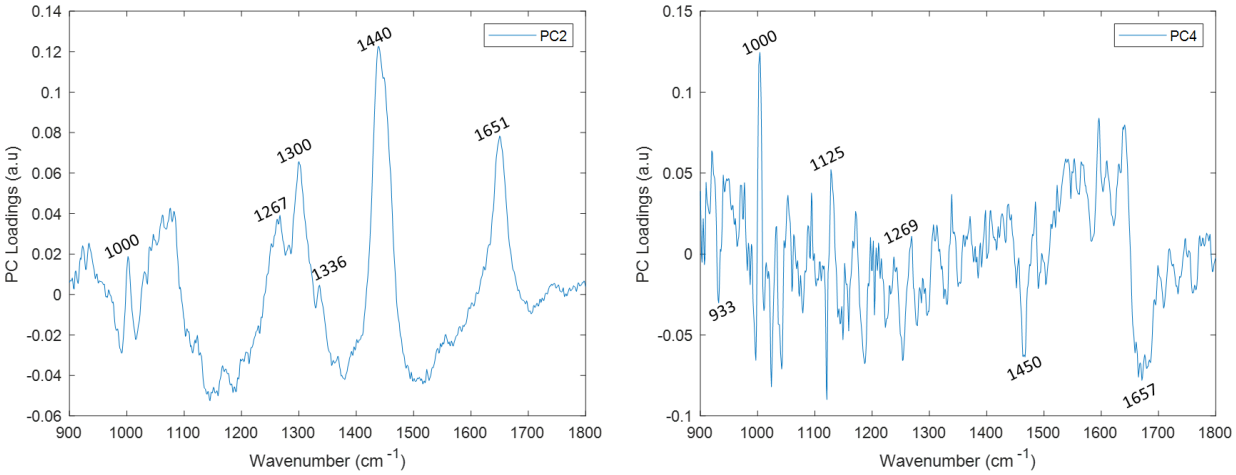


Figure C-12: PC2 and PC4 loading plots.

PLS-DA Analysis

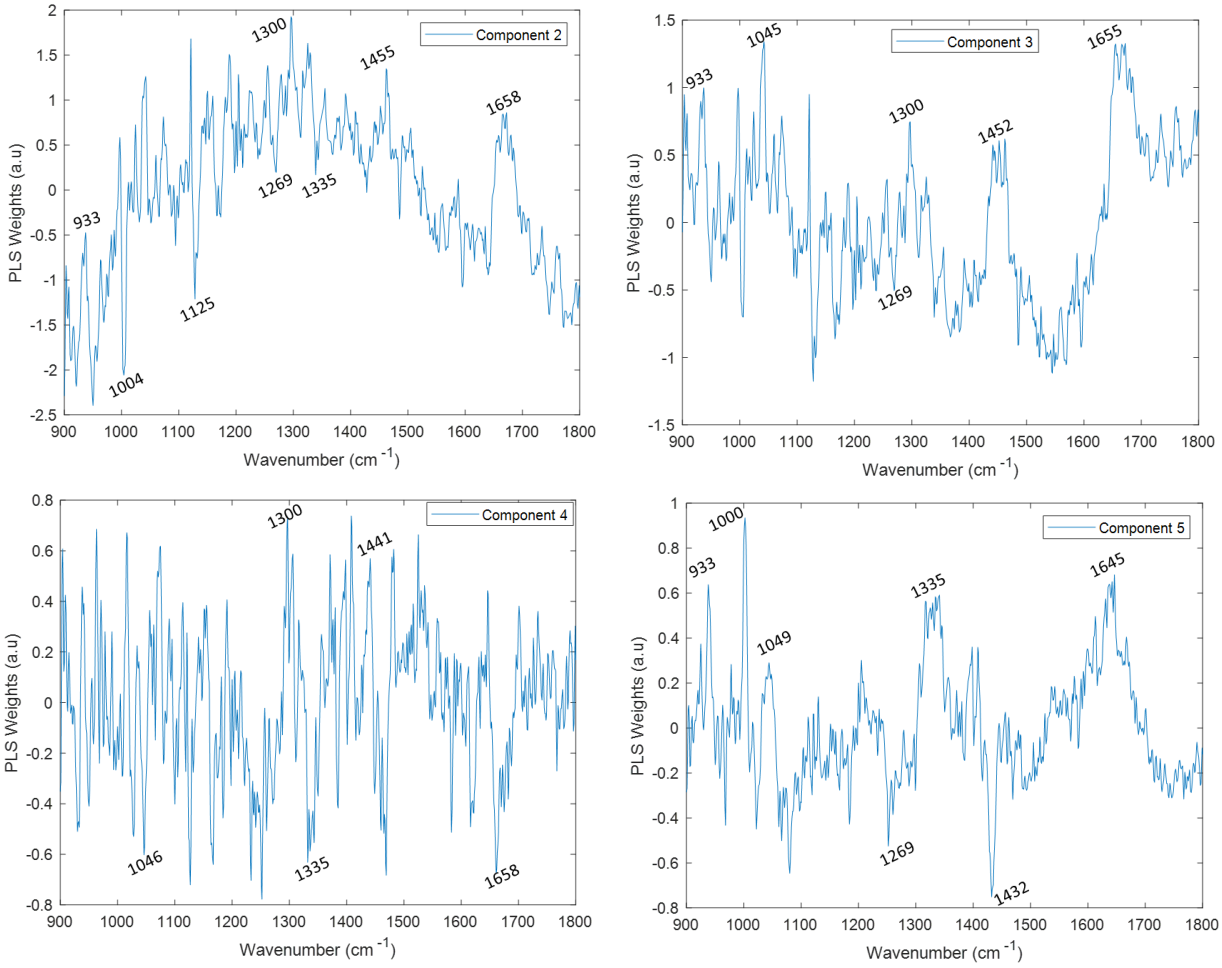


Figure C-13: PLS Component 2,3,4 and 5 weight plots

Study 3

a) TDP-43 vs. TDP-43^{Q331K}

PCA analysis

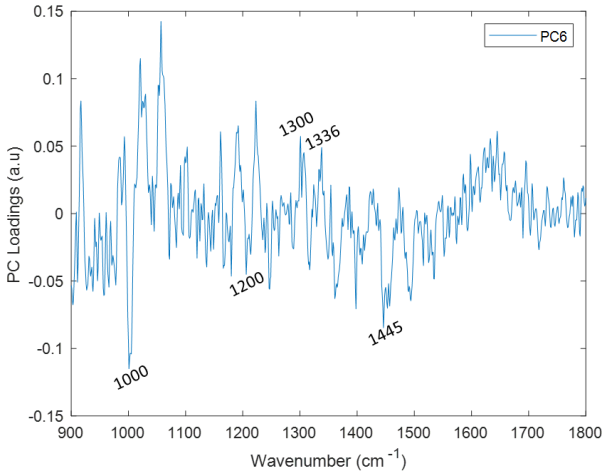


Figure C-14: PC6 loading plot.

PLS-DA analysis

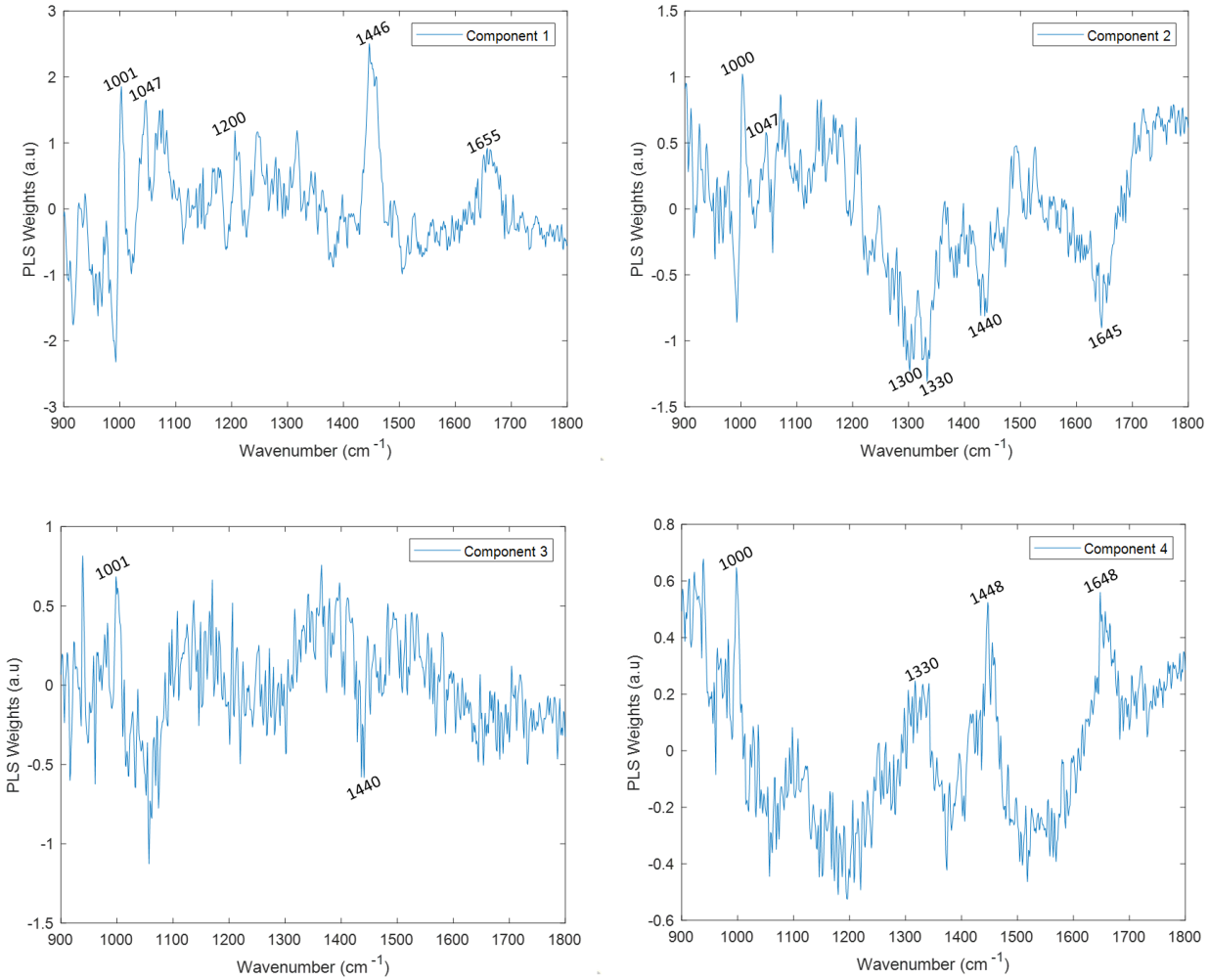


Figure C-15: PLS Component 1,2,3 and 4 weight plots.

b) SOD1^{G93A} vs. TDP-43^{Q331K}

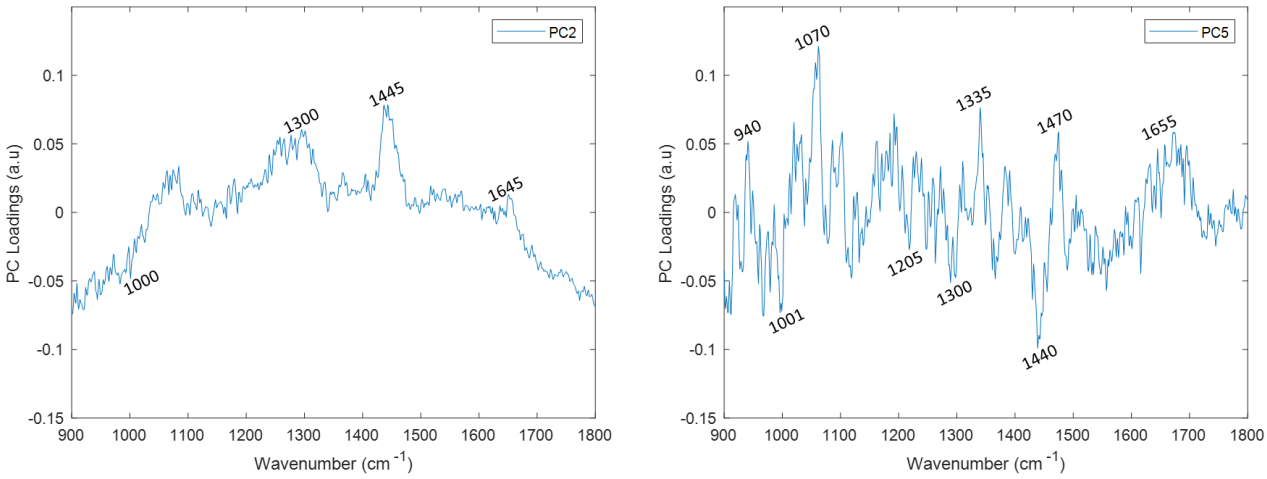


Figure C-16: PC2 and PC5 loading plots.

PLS-DA Analysis

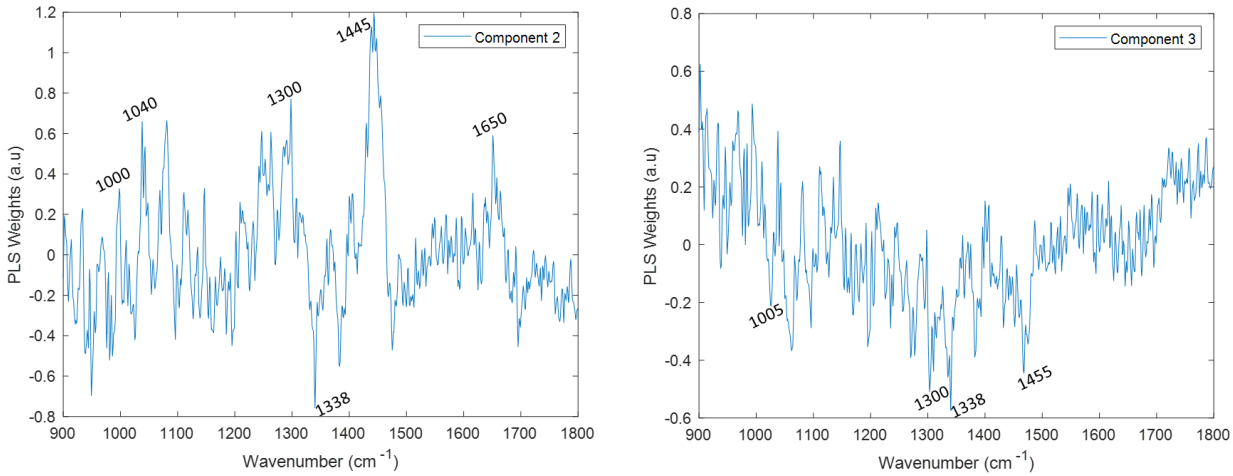


Figure C-17: PLS Component 1,2,3 and 4 weight plots.

Study 4

a) three months old *mdx* mice

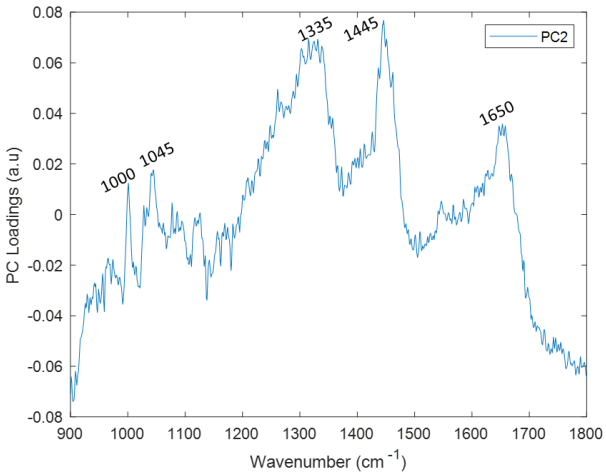
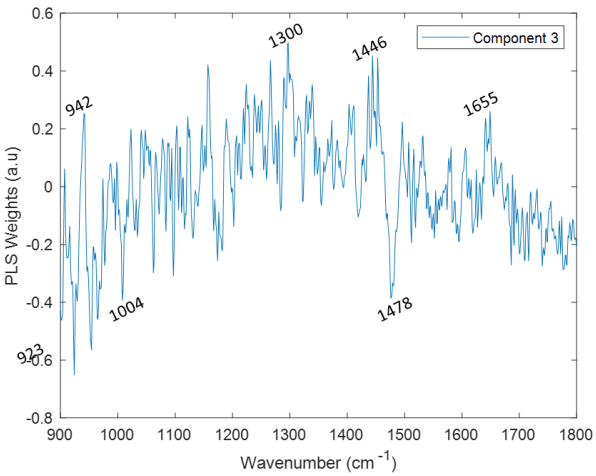


Figure C-1: PC2 loading plot.



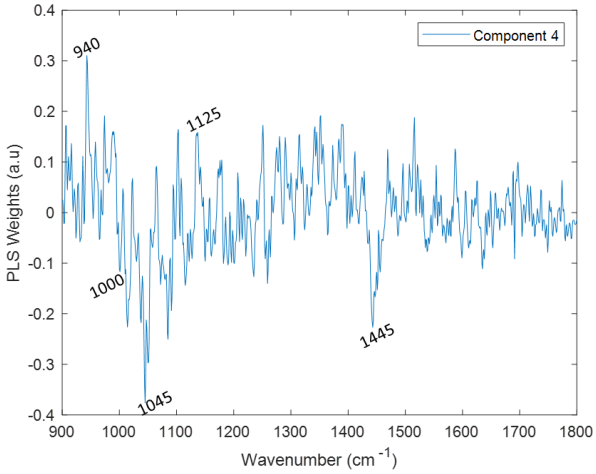


Figure C-18: PLS Component 2 and 3 weight plots.

Human Tissue Recordings

‘Mitochondrial Disorders’ vs. ‘Healthy’ samples

Microscope

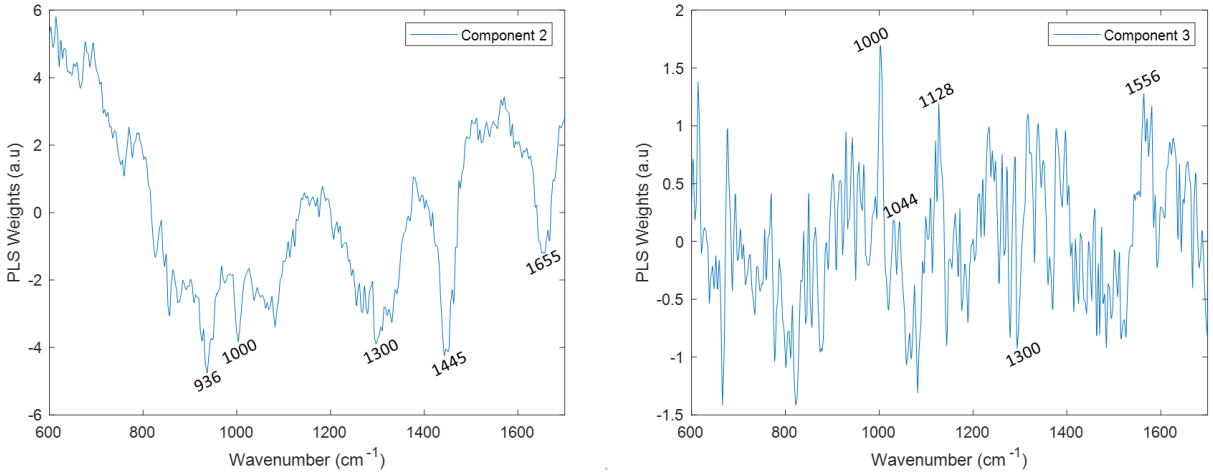


Figure C-19: PLS Component 2 and 3 weight plots.

‘MND’ vs. ‘Healthy’ samples

Microscope

PCA Analysis

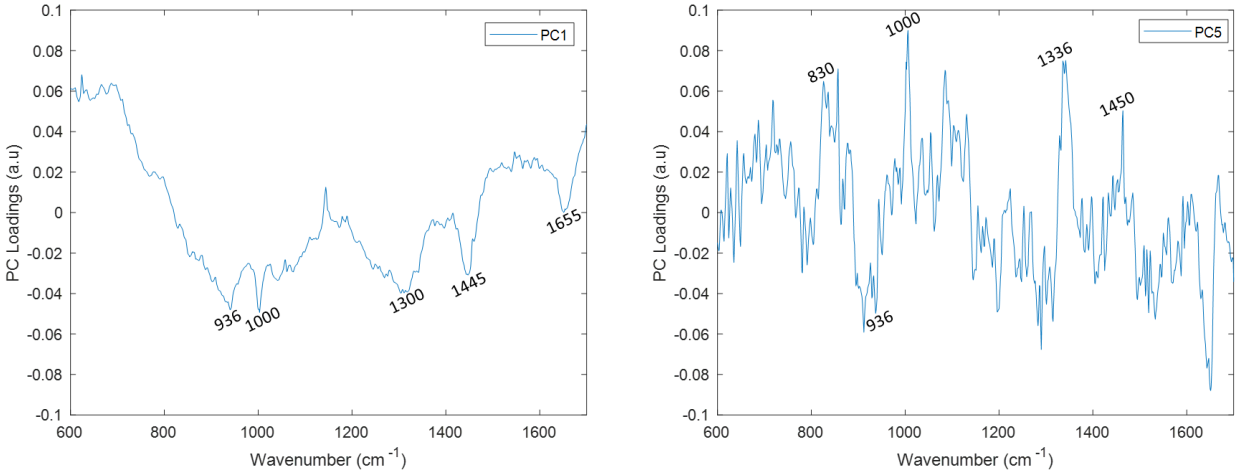


Figure C-20: PC1 and PC5 loading plots.

PLS-DA

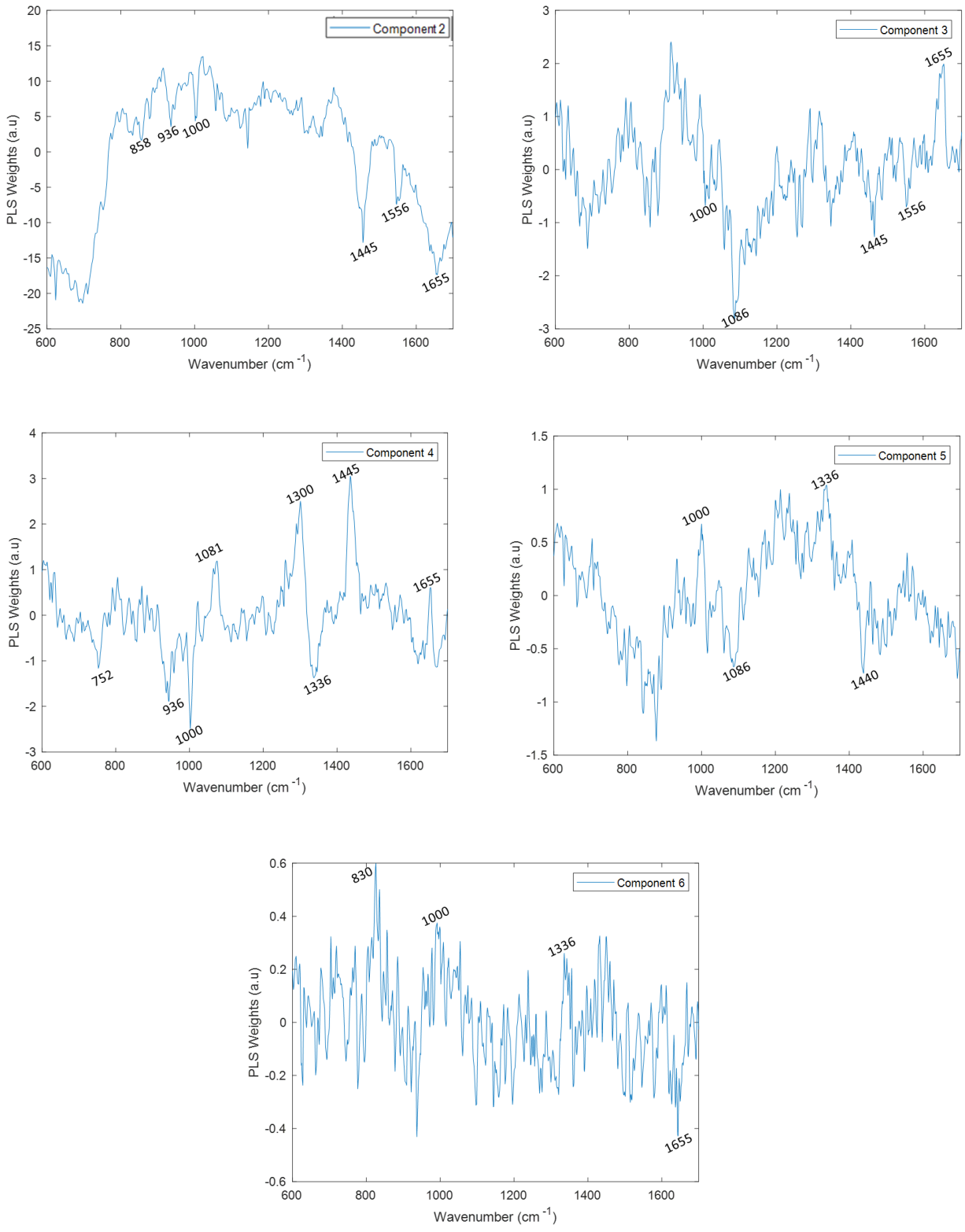


Figure C-21: PLS Component 2,3,4,5 and 6 weight plots.

Appendix D

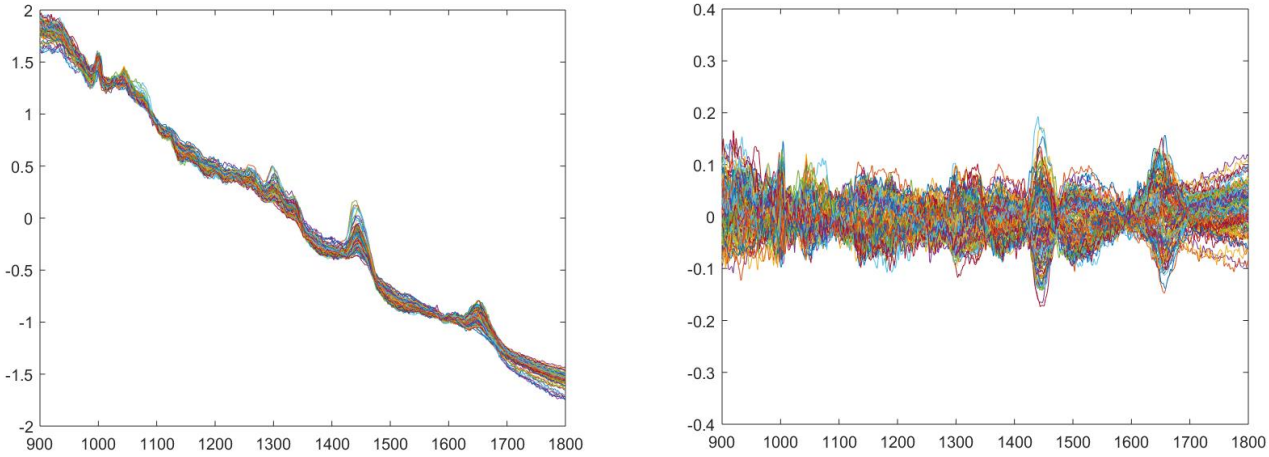


Figure D-1: Normalised Raman spectra (left) and normalised and mean centred Raman spectra (right).

Appendix E

Project Main Outputs and Achievements

- Plesia M, Stevens OA, Lloyd GR, Kendall CA, Coldicott I, Kennerley AJ, Miller G, Shaw PJ, Mead RJ, Day JCC, Alix JJP. *In Vivo* Fiber Optic Raman Spectroscopy of Muscle in Preclinical Models of Amyotrophic Lateral Sclerosis and Duchenne Muscular Dystrophy. ACS Chem Neurosci. 2021 May 19;12(10):1768-1776. doi: 10.1021/acschemneuro.0c00794.
- 2017 Winter Science Meeting, The Academy of Medical Sciences, London, November 2017 (Poster and flash presentation)
- Conference: SPEC 2018, The International Conference on Clinical Vibrational Spectroscopy, Glasgow, United Kingdom, June 2018 (Poster and flash presentation)
- Conference: North East Postgraduate Conference 2018, Newcastle, November 2018 (Oral presenation)
- Conference 29th International Symposium on ALS/MND 2018, Glasgow, United Kingdom, December 2018 (Oral presentation)