Statistical Modelling Of Mitochondrial Disease

John Patrick Grady

BSc (Hons) MSc MRes

This thesis is submitted for the degree of Doctor of Philosophy at Newcastle University

Wellcome Trust Centre for Mitochondrial Research

Institute of Neuroscience/Institute of Ageing and Health

November 2013

Author's declaration

This thesis is submitted for the degree of Doctor of Philosophy at Newcastle University. The research was conducted in the Wellcome Trust Centre for Mitochondrial Research under the supervision of Professor D.M. Turnbull, Dr Richard McNally, and Dr Claudia Racca. Unless otherwise stated all the work is my own, and was conducted between October 2010 and November 2013.

I certify that none of the material offered in this thesis has been previously submitted by me for a degree or any other qualification at this or any other university.

Abstract

Mitochondrial DNA mutations are a major cause of disease in the human population. Understanding the disease associated with these mutations is complicated by heteroplasmy, the mixture of wild-type and mutated mitochondrial DNA. Heteroplasmy can vary between cells, tissues, and organs, and the disease associated individual mutations is hugely varied on account of this. The mitochondrial genome encodes critical proteins of the oxidative phosphorylation system and mutation leads to energy deficits in cells and a wide range of secondary effects. The central and peripheral nervous system are commonly affected in mitochondrial disease and quality of life for patients is severely impaired.

Although pathogenic mitochondrial genetic mutations were first identified over twenty five years ago, little progress has been made in understanding the expected progression of disease in patients. The aim of this study was to use statistical modelling to further understanding of disease progression in mitochondrial DNA mutations. The Medical Research Council Mitochondrial Disease Cohort provided the majority of patient data. Patients had been assessed using the Newcastle Mitochondrial Disease Adult Scale, which facilitates quantitative research on mitochondrial disease burden.

This project comprises studies of two of the most common mitochondrial DNA mutations. The first study concerns patients with the m.3243A>G mutation, the most common pathogenic point mutation, and considers the effect of age and heteroplasmy on disease progression. Prediction models of both overall disease burden and specific phenotypic features were developed. Important features of the patient cohort were also examined, including heteroplasmy in different tissues and differences in disease expression between sexes. The second study looks at patients with single large-scale mitochondrial DNA mutations. The effect of deletion size, location of the deletion on the genome, and heteroplasmy were investigated, and all three predictors were found to be significant in understanding disease progression.

Acknowledgements

Firstly I would like to thank my supervisor Doug. He has been a great support over the past three years, particularly when I was ready to jack it all in and go and get a job as a gardener or audition for a part on The Apprentice. Yes, I really did that. I'd had enough of cell culture and pipettes and weird pulsating organisms growing in 96 well plates when all I wanted was a tidy ball of neurons that would show something straightforward and interesting. I'd really had enough. Just when I was ready to abort (did I say abort?) he managed to pull a project out of his head that I was really interested in, modelling disease progression. Much more my bag, and I'm very grateful to him. I eventually actually started to really enjoy my PhD, which was kind of weird, I'd resigned myself to sticking it out in the same way as I read Ulysses, which I found unfathomable but decided to read it word by word and get through it no matter what, just so I could say I'd done it. I managed to get through nineteen pages of that I chucked it in the bin. So grateful thanks for turning a torturous experience into an enjoyable and stimulating one.

Thanks also go to Claudia and Richard. To Richard, for putting up with me turning up and wittering about statistics for an hour every couple of weeks, then sending me off with some sage advice. To Claudia, for inspiring me to enjoy the idea of lab science even if I didn't like doing it very much. She was also a model of how to do things neatly and methodically, which I understand about as much as I understand Chinese but greatly respect in any case.

I'd like to thank those who collaborated on the various bits of research I've presented here, as I really didn't like doing the lab work and was so delighted to find that other people didn't mind it so much. So particular thanks go to Georgia, Helen, Julie, Tilly, and Yi.

I'd also like to thank Bob and Rob for keeping me on track (someone needed to), Bobby for his help with the cohort, and Simon Cockell for help with the technicalities of the database.

I would like to thank all the all patients registered in the cohort who have made this study possible, as well diagnostics team, who provided tons of stuff.

Huge thanks go to the various members current and past of the Mitochondrial Research Group, who have been a great source of supportive everything over the years. Afternoon tea time has been most entertaining over the last three years, disturbingly so. Thanks in particular to the poor souls who have had to sit next to me over the past few years and put up with my inane ramblings. Georgia, you've been very polite.

Thanks also go to my family, all my sisters and brother and mum and dad, and brothers in law, who have been very supportive over the last few years. They seemed to get more excited about my work than I did at times, which was very kind really. And to Chris, who didn't get excited about it at all, but gets some marks for tolerance.

And last of all, grateful thanks to the Wellcome Trust for my Studentship.

Table of Contents

List of Figures

List of Tables

List of Equations

Publications

Grady JP, Campbell G, Ratnaike T, Blakely EL, Falkous F, Nesbitt V, Schaefer AM, McNally RJ, Gorman GS, Taylor RW, McFarland R, Turnbull DM. Disease progression in patients with single, large-scale mitochondrial DNA deletions. Brain 2013.

Greaves LC, Elson JL, Nooteboom M, **Grady JP**, Taylor GA, Taylor RW, Mathers JC, Kirkwood TB, Turnbull DM. Comparison of mitochondrial mutation spectra in ageing human colonic epithelium and disease: absence of evidence for purifying selection in somatic mitochondrial DNA point mutations. PLoS Genetics 2012; 8(11).

Abbreviations

ADP	Adenosine diphosphate
AIC	Akaike's information criterion
ATP	Adenosine triphosphate
B2M	Beta-2-microglobulin gene
BIC	Bayes' information criterion
COX	Cytochrome c oxidase
CPEO	Chronic progressive external ophthalmoplegia
DNM1L	dynamin-1-like protein gene
ETC	Electron transport chain
FSGS	Focal segmental glomerular sclerosis
GLM	General linear model
HSP	Heavy strand promoter
KSS	Kearns-Sayre syndrome
LHON	Leber's hereditary optic neuropathy
LSP	Light strand promoter
MELAS	Mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes
MERRF	Myoclonic epilepsy with ragged-red fibres
ML	Maximum likelihood
MNGIE	Mitochondrial neurogastrointestinal encephalomyopathy
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
mtDNA	Mitochondrial DNA
MT-TL1	Mitochondrial tRNA leucine^{UUR} gene
mt-tRNA	Mitochondrial tRNA
NAD+	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
nDNA	Nuclear DNA
O_H	Origin of heavy strand synthesis
O_L	Origin of light strand synthesis
OPA1	Optic atrophy 1gene
OXPHOS	Oxidative phosphorylation
PARK2	Parkin gene
PINK1	PTEN-induced putative kinase 1 gene
PS	Pearson's syndrome
REML	Restricted (or residual) maximum likelihood
RITOLS	Ribonucleotide incorporation throughout the lagging strand
RNA	Ribonucleic acid

ROS	Reactive oxygen species
RRF	Ragged-red fibre
rRNA	Ribosomal ribonucleic acid
SDH	Succinate dehydrogenase
SKM	Skeletal muscle
tRNA	Transfer ribonucleic acid
TYMP	Thymidine phosphorylase
WT	Wild-type

Chapter 1

Chapter 1. **Introduction**

1.1. Mitochondria

Mitochondria are cytoplasmic organelles present in almost all eukaryotic cells which house machinery for ATP (adenosine triphosphate) production and perform critical roles in apoptosis (Wang and Youle, 2009), calcium homeostasis (Patergnani *et al.*, 2011), and iron-sulphur (Fe-S) cluster formation (Rouault and Tong, 2005), as well as numerous other functions. In humans, mitochondria are found in all cells other than mature red blood cells. Though the vast majority of proteins that make up mitochondria are encoded by nuclear DNA (nDNA), mitochondria also contain their own DNA, the only extra-nuclear source of DNA in humans, which encodes a small set of proteins vital for oxidative phosphorylation (OXPHOS). Mutations in mitochondrial DNA (mtDNA) are responsible for a wide range of disease that impact on mitochondrial and cellular function as a whole.

1.1.1. Origins of mitochondria

Mitochondria have long been believed to have evolved from free Eubacteria that became integrated into a primitive eukaryote through endosymbiosis (Margulis, 1971), a theory which proposed that the fusion occurred after the emergence of the cell nucleus. A competing theory, known as the 'hydrogen hypothesis' proposes that the mitochondrion and the nucleus formed contemporaneously, after fusion of a hydrogen-dependent Archaebacterium and a hydrogen-producing Eubacterium (Martin and Muller, 1998). In both theories, however, there is a single endosymbiotic event, and a subsequent transfer of the majority of the genetic material from the proto-mitochondrion to the nucleus.

1.1.2. Structure

Traditionally the mitochondria has been viewed as a rod shaped or ovoid organelle thought to be approximately 2μm long and 0.5μm in diameter, encapsulated by a double membrane (Palade, 1953). The outer membrane contains an abundance of the voltage dependent anion channel (VDAC) otherwise known as porin, which when open allows for free movement of molecules and ions of low molecular weight (under 10kDa) between the cytoplasm and the inter-membrane space (Alberts *et al.*, 2002).

The inner mitochondrial membrane (IMM) is essentially impermeable, but contains a large number of transport proteins that regulate the flow of material into the

mitochondrial matrix contained within the membrane, and is rich in cardiolipin. The IMM was traditionally though to be highly folded creating invaginations (known as cristae) into the matrix, clearly visible using electron microscopy; however, it is now known that the cristae are formed from a distinct membrane connected to the inter-membrane space through cristae junctions, and thus the IMM can be subdivided into two parts, the inner boundary membrane (IBM) and the cristae membrane (CM) (Frey and Mannella, 2000). The CM is enriched in proteins involved in oxidative phosphorylation, iron-sulphur cluster biogenesis, protein synthesis and mtDNA-encoded protein transport, whilst the IBM is rich in proteins responsible for mitochondrial fusion and nDNA-encoded protein transport (Vogel *et al.*, 2006). The invaginations of the CM create a huge surface area over which OXPHOS can be conducted.

The mitochondrial matrix contains multiple mtDNA molecules, transcriptional and translational machinery, the various proteins responsible for the tricarboxylic acid (TCA) cycle, and is the site of Fe-S cluster formation.

1.1.3. Oxidative phosphorylation

Mitochondria are often referred to as the powerhouses of the cell, on account of their major role in the production of cellular ATP. OXPHOS is a highly efficient process for the production of ATP that utilises several intermediate products of the TCA cycle (also known as Kreb's cycle or the citric acid cycle) (Hatefi, 1985).

Cellular respiration begins with glycolysis in the cytosol, where glucose is broken down into pyruvate, producing two molecules of ATP (Equation 1.1).

$$Glucose + 2NAD^+ + 2ADP + 2P_i \rightarrow 2Pyruvate + 2NADH + 2H^+ + 2ATP + 2H_2O$$

Equation 1.1. Glycolysis.

The pyruvate is transported across the double mitochondrial membrane into the matrix, where the enzyme pyruvate dehydrogenase converts it into acetyl CoA (Equation 1.2).

$$Pyruvate + CoA + NAD^+ \rightarrow Acetyl\ CoA + NADH + H^+ + CO_2$$

Equation 1.2. Pyruvarte decarboxylation.

Acetyl CoA is one of the major substrates of the TCA cycle, along with NAD+ and FADH, and provides the carbon atoms within the acetyl group to be oxidised whilst the other two substrates are reduced, becoming electron carriers. The overall TCA cycle is

shown in Equation 1.3,

$$Acetyl\ CoA + 3NAD^+ + FAD + ADP + P_i + 2H_2O \rightarrow CoA + 3NADH + 2H^+ + FADH_2 + ATP + CO_2$$

Equation 1.3 Outcome of the TCA cycle.

NADH and $FADH_2$ are essential components of the final stages of respiration, OXPHOS, in that they act as electron carriers, donating electrons to complex I and II respectively. These complexes in turn transfer electrons to complexes III and IV, releasing energy which is used to pump H^+ into the intermembrane space from the matrix. This creates an electro-chemical gradient which is then used by complex V (ATP synthase) to generate ATP from ADP.

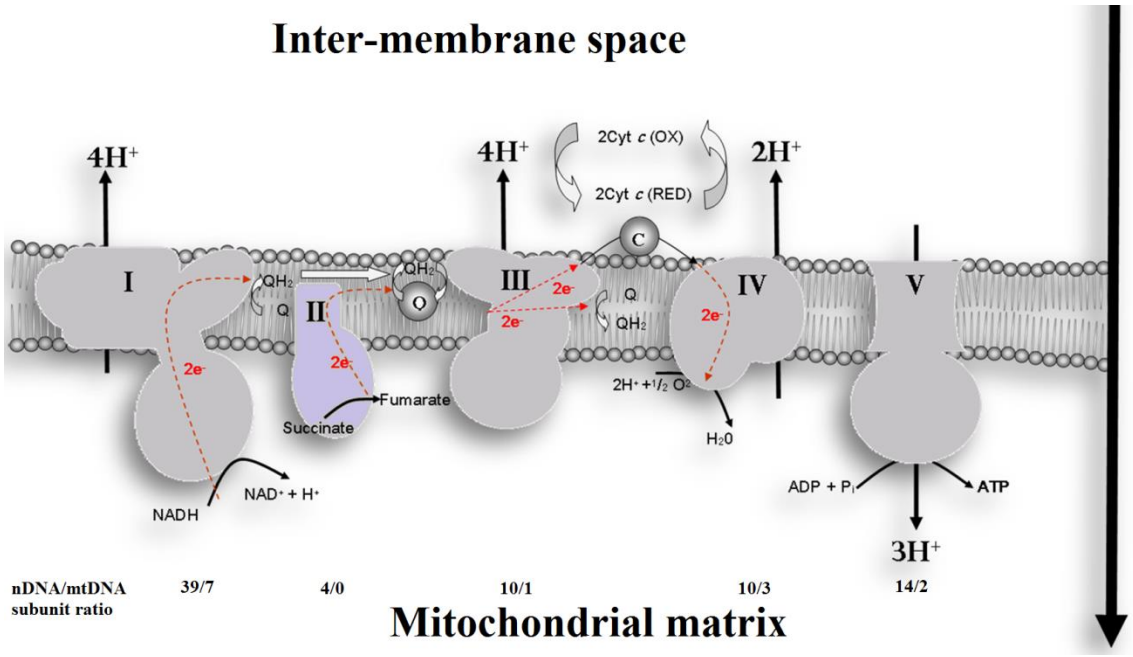


Figure 1.1 Oxidative phosphorylation.

The protein complexes are embedded in the cristae membrane of the inner mitochondrial membrane. Electrons (in red) enter the electron transport chain (ETC) via complexes I and II and are transported to complex III via reduction of ubiquinone (Q), which is soluble in the membrane. Complex III re-oxidises the ubiquinol (QH_2) back to ubiquinone, and the electrons pass via cytochrome *c* (C) in a further redox reaction to complex IV. During this process complexes I, III, and IV pump electrons from the matrix to the inter-membrane space, creating an electro-chemical gradient, which drives the ATP synthase (Complex V) to produce ATP. Below each subunit the ratio of nuclear encoded to mitochondrially encoded subunits is shown. Complex II is the only complex that is entirely encoded by nuclear DNA. Image (adapted) courtesy of Eve Simcox.

1.1.3.1. Complex I

Complex I (NADH dehydrogenase (ubiquionone)) couples transfer of two electrons from NADH to ubiquinone with translocation of four protons across the inner membrane. It is the largest complex of the OXPHOS system, made up of 46 subunits of which 7 are mtDNA encoded (Ugalde *et al.*, 2004). The overall reaction is shown below (Equation 1.4).

$$NADH + Q + 5H^{+}_{matrix} \rightarrow NAD^{+} + QH_2 + 4H^{+}_{ims}$$

Equation 1.4 Complex I reaction.

Complex I dysfunction is the most frequently observed OXPHOS defect in humans, caused by mutations of both nuclear and mitochondrial origin, and related both to mutations in the structure itself and the assembly of the complex (Mimaki *et al.*, 2012). Leber's hereditary optic neuropathy (LHON), for example, is a disease generally caused by mutations in genes encoding complex I subunits (Wallace *et al.*, 1988).

1.1.3.2. Complex II

Complex II, also known as succinate dehydrogenase (SDH) or succinate ubiquionone oxioreductase, is the only complex of the OXPHOS system that is entirely nuclear encoded, and is also the smallest, comprising four subunits. It does not translocate protons as the other four complexes do, however it plays the crucial role of catalysing the conversion of succinate to fumarate whilst generating $FADH_2$, which is then oxidised to FAD and the electrons used to reduce ubiquinone to ubiquinol (Hagerhall, 1997). The overall reaction is shown in Equation 1.5.

$$succinate + Q \rightarrow fumarate + QH_2$$

Equation 1.5 Complex II reaction.

Complex II mutations are associated with tumorigenic phenotypes, thought to be through caused by excessive generation of reactive oxygen species (ROS) or stabilization of hypoxia inducible factor 1 (HIF1) during normoxia (Hoekstra and Bayley, 2013).

1.1.3.3. Complex III

Complex III, or ubiquinol:cytochrome *c* oxioreductase comprises 11 subunits, only one of which is mitochondrially encoded, by the *MT-CYB* gene (cytochrome *b*). The

complex transfers two electrons to cytochrome *c* from ubiquinol, and also pumps two protons into the inter membrane space. The reaction is shown in Equation 1.6.

$$2QH_2 + Q + 2Cytc_{oxidised} + 2H^+_{matrix} \rightarrow 2Q + QH_2 + 2Cytc_{reduced} + 4H^+_{ims}$$

Equation 1.6 Complex III reaction, the Q cycle.

Complex III abnormalities are rare in humans (Benit *et al.*, 2009). The phenotype of Complex III abnormalities are similar to other OXPHOS defects, in particular LHON is also associated with Complex III mutation (Brown *et al.*, 1992).

1.1.3.4. Cytochrome *c*

Cytochrome *c* is a highly conserved nuclear encoded haem-containing protein found in the inter membrane space, and has a critical role in the ETC, apoptosis, and also acts as a ROS scavenger (Huttemann *et al.*, 2011). In OXPHOS it carries a single electron from complex III to IV.

1.1.3.5. Complex IV

Complex IV, or cytochrome *c* oxidase (COX) is the terminal complex of the ETC, and comprises 13 subunits, three of which are mitochondrially encoded. It reduces O_2 to H_2O using electrons provided by cytochrome *c*, and transfers 4 protons per oxidised molecule from the matrix to the inter membrane space, as shown in Equation 1.7. Four electrons are required for the reduction of one molecule of O_2 to H_2O, and this step is thought to be rate-limiting in the mammalian ETC (Huttemann *et al.*, 2011).

$$4Cytc_{reduced} + 8H^+_{matrix} + O_2 \rightarrow 4Cytc_{oxidised} + 2H_2O + 4H^+_{ims}$$

Equation 1.7 Complex IV reaction.

1.1.3.6. Complex V

Complex V, the F_OF_1 ATP synthase, is the final step of OXPHOS, which catalyses the conversion of ADP to ATP. The complex is made up of two domains, F_O located in the membrane, and F_1 which extends into the matrix. Each unit comprises multiple copies of several subunits, the majority of which (14) are nDNA encoded, with two mitochondrially encoded subunits, both in the F_O domain. Proton movement through the F_O domain generates rotary torque, which powers the formation of ATP from ADP and phosphate. Each molecule of ATP costs 2.7 protons in production (Ferguson, 2010).

Mutations in Complex V often lead to very severe disease, characterised by lesions in the brain and particularly the striatum, including NARP (neurogenic muscle weakness, ataxia, and retinitis pigmentosa) and Leigh Syndrome, a severe infantile neurological disorder resulting in infant death, commonly through respiratory failure (Schon *et al.*, 2001).

1.1.3.7. Supercomplexes

It has been shown that the ETC complexes tend to group together into conglomerates called supercomplexes; for instance, almost all mammalian Complex I is found aggregated with Complexes III and IV (Schagger and Pfeiffer, 2000). It is thought that Complexes III and IV are assembled independently, but that Complex I assembly is multi-stage and the final stage, the addition of the NADH dehydrogenase catalytic module, occurs after supercomplex formation (Moreno-Lastres *et al.*, 2012; Winge, 2012). From the perspective of disease and pathology this is of fundamental importance, as inter-dependence of the complexes is potentially explanatory of genetic defects in one complex affecting functional operation or assembly of other complexes, particularly as regards Complex III or IV defects affecting Complex I assembly.

1.1.4. Other functions of mitochondria

1.1.4.1. Iron homeostasis and iron-sulphur (Fe-S) cluster biogenesis

The production of Fe-S clusters is the sole conserved function of both mitochondria and primitive mitochondria (mitosomes) across all eukaryotes, underlying the importance of this process. Fe-S clusters are essential in OXPHOS as they facilitate the electron transfer by repeated redox changes from Fe^{2+} to Fe^{3+}, and form part of complexes I, II, and III (Schultz and Chan, 2001).

Iron within mitochondria is used for haem or Fe-S cluster synthesis or stored in mitochondrial ferritin. Import of iron into mitochondria through the inner membrane is regulated by mitoferrins (Paradkar *et al.*, 2009), though the transport mechanism from the cytosolic iron pool and across the outer mitochondrial membrane remains to be elucidated (Richardson *et al.*, 2010).

1.1.4.2. Ca^{2+} signalling and homeostasis

Mitochondria have long been recognised for their role in cytosolic Ca^{2+} buffering (Carafoli, 2003) and by extension extracellular Ca^{2+} regulation (Cohen and Fields, 2004), and are able to absorb up to 1000nmol Ca^{2+} per mg of mitochondrial protein, imported via the outer membrane VDAC pore and the calcium uniporter located in the inner mitochondrial membrane (Kirichok *et al.*, 2004). The importance of this role cannot be understated; Ca^{2+} signalling is an essential regulator of cellular function (Clapham, 2007) and has a fundamental role in intercellular communication (Hofer *et al.*, 2000), induction of apoptosis (Orrenius *et al.*, 2003), regulation of ATP production (Griffiths and Rutter, 2009), and regulation of pre-synaptic transmission in neurons (Kostyuk, 2007).

1.1.4.3. Apoptosis

Programmed cell death, or apoptosis, is a vital function of an organism to eliminate unwanted or damaged cells, and is vital for proper embryonic development (Danial and Korsmeyer, 2004). It is inducible by both an extrinsic cell receptor mediated pathway, and an intrinsic cell damage mediated pathway, contingent upon the release of cytochrome *c* from the mitochondria and subsequent activation of caspase-9 (Wang and Youle, 2009).

1.1.4.4. Reactive oxygen species

Though mitochondrially produced reactive oxygen species (ROS) were historically thought to lack a physiological role and were only associated with cell damage, there is growing evidence that ROS have a critical physiological role. It is hypothesised that ROS levels at a basal level maintains homeostatic function in the cell, but fluctuations in ROS alter signalling pathways; ROS is known to have roles in cell differentiation, autophagy, immune cell activity, and metabolic adaptation (Sena and Chandel, 2012). ROS have been shown to induce reversible posttranslational modifications of several proteins within important signalling cascades (Finkel, 1998; Rhee *et al.*, 2000).

Complex I and III are both major sources of ROS (Turrens and Boveris, 1980; Sugioka *et al.*, 1988), though Complex I is thought to be more a more proliferative producer of superoxide O_2^-, predominantly when ATP is not being produced and there is a high proton motive force (the sum of the membrane potential and pH gradient), or there is a high $NADH/NAD^+$ ratio in the matrix (Murphy, 2009). However, Complex III produces O_2^- on both sides of the inner membrane, but Complex I produces it only matrix-side.

Overall, though mitochondria are a significant source of cellular ROS, it is not by any means certain that they are the major source of ROS within cells, as the endoplasmic reticulum and peroxisomes are also known sources (Brown and Borutaite, 2012).

1.1.5. Biogenesis regulation

Mitochondrial biogenesis is tightly regulated by cells in order to tailor ATP production to suit cellular energy requirements. The peroxisome proliferator-activated receptor γ (PPARγ) co-activator 1α (PGC-1α) is the primary controller of mitochondrial biogenesis, and has a critical role in signalling cascades involving AMP activated protein kinase (AMPK) and the Sirtuin class of proteins that sense energy imbalances in cells and regulate mitochondrial biogenesis accordingly (Hardie *et al.*, 2012; Andreux *et al.*, 2013).

1.1.6. Dynamics

Though traditionally thought of as discrete organelles, mitochondria in fact form a dynamic network that undergoes constant fission and fusion (Westermann, 2010). The importance of this dynamism is illustrated by the fact that mutations in several nDNA encoded proteins with critical roles in fission and fusion are associated with disease, including mitofusins 1 and 2 (MFN1 and MFN2), optic atrophy 1 (OPA1), and dynamin-1-like protein (DNM1L) (Liesa *et al.*, 2009). OXPHOS generates an electro-chemical gradient across the inner membrane of the mitochondria, which is maintained at around -140mV (Gerencser *et al.*, 2012), and a healthy membrane potential is a requirement for fusion to occur (Legros *et al.*, 2002; Meeusen *et al.*, 2004).

As well as fission and fusion, mitochondria undergo autophagic degredation known as mitophagy, which has been shown to selectively target impaired mitochondria (Kim *et al.*, 2007; Twig *et al.*, 2008). This process is intimately related to human disease, for example in the mutations of the *PINK1* and *PARK2* genes that are linked to inherited early-onset Parkinson's disease and which are involved in the mitophagic pathway (Chen and Chan, 2009).

The umbrella term 'mitochondrial dynamics' also includes consideration of the migration and movement of mitochondria within the cytoplasm to areas of high energy need, for instance neuronal growth cones, and pre- and post- synaptic sites (Morris and Hollenbeck, 1993; Li *et al.*, 2004; Miller and Sheetz, 2004; Chang *et al.*, 2006).

1.1.7. The mitochondrial genome

The human mitochondrial genome (mtDNA) is a 16,569bp double-stranded circular intron-less genome located in the mitochondrial matrix. It codes 37 genes; 13 hydrophobic proteins essential for OXPHOS, 22 tRNAs, and 2 rRNAs. The structure is shown in Figure 1.2. The majority of the genes (all but one of the protein encoding

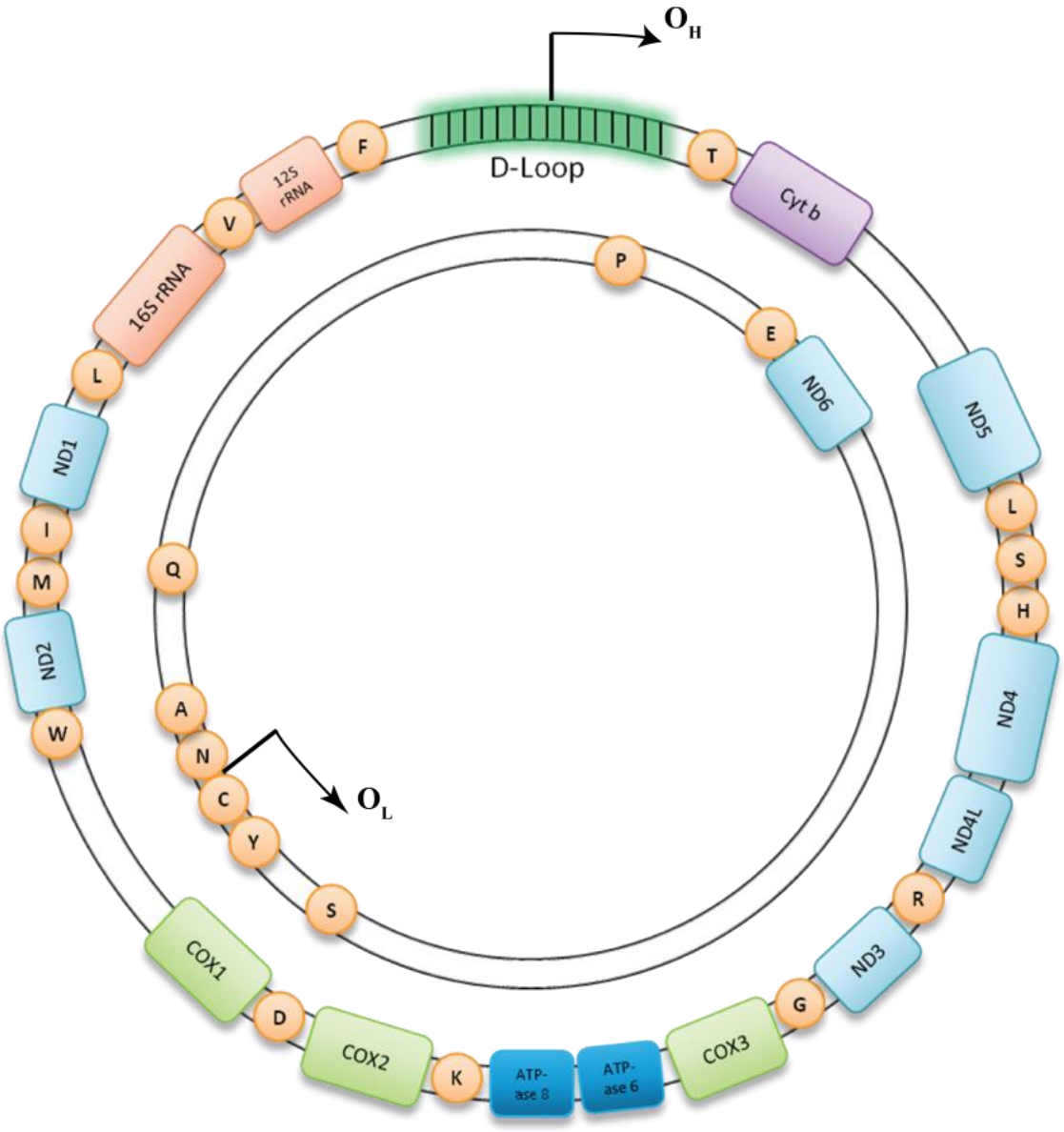


Figure 1.2 The human mitochondrial genome.

The genome encodes 37 genes, including 13 polypeptides, 22 tRNAs, and 2 rRNAs. All but one of the protein encoding genes are found on the outer heavy strand of the genome; 6 subunits of Complex I (*MT-ND1*-5 and ND4L) , one Complex III subunit (Cytochrome B), three subunits of Complex IV (COX1-3), and two subunits of Complex V (ATPase 6 & 8). The final protein encoding gene is the ND6 subunit of Complex I on the inner light strand. The tRNA genes are dispersed around both strands. The D-loop is the only major non-coding region of the genome. The origins of heavy (O_H) and light (O_L) strand origin are also shown. Image (adapted) courtesy of Dr Casey Wilson.

genes, 14 of the tRNAs, and both rRNAs) are transcribed from the outer (heavy) strand. The other 8 tRNAs and the protein encoding gene *MT-ND6* are transcribed from the inner (light) strand. The only substantial non-coding region, the displacement-loop or d-loop, contains the major control elements for transcription and replication, and extends approximately 1kb in human mtDNA.

Multiple copies of the mtDNA molecule are found within an individual cell; in general mammalian cells house tens to hundreds of mitochondria, and each mitochondrion contains several copies of the mitochondrial genome (Wiesner *et al.*, 1992), with an estimated 5,000 to 10,000 copies of the genome per cell (Lightowlers *et al.*, 1997). mtDNA is packaged into discrete nucleoids composed of mtDNA with some of the machinery necessary for replication and transcription of the DNA (Brown *et al.*, 2011). The number of DNA molecules per nucleoid is debated, but has recently been measured as generally a single copy per nucleoid (Kukat *et al.*, 2011). In this study mitochondrial transcription factor A (TFAM) was reported as the bulk constituent of the nucleoid, though other proteins are also present, including single-stranded DNA binding protein (mtSSB).

1.1.8. Transcription and translation

There are three sites of transcription initiation within the D-loop, the L-strand promoter (LSP) and the two H-strand promoters (HSP1 and 2); transcription generates long polycistronic molecules, as depicted in Figure 1.3. The LSP generates a single transcript containing the all the genes on the L-strand, as HSP2 does for the H-strand; HSP1, on the other hand, generates a transcript containing only the two rRNA genes and two tRNA genes (Montoya *et al.*, 1982; Chang and Clayton, 1984; Zollo *et al.*, 2012). Transcription is bi-directional and conducted by the mitochondrial RNA polymerase (POLMRT) in conjunction with other proteins including TFAM and mitochondrial transcription factor B2 (Rebelo *et al.*, 2011). Termination of transcription is implemented by mitochondrial termination factors (MTERF), four of which have so far been identified. The roles of these factors are still very much under debate however; it has been recently shown, contrary to long standing belief, that MTERF1 does not couple rRNA gene transcription initiation and termination but instead appears to block transcription to avoid transcriptional interference at the L-strand promoter (Terzioglu *et al.*, 2013).

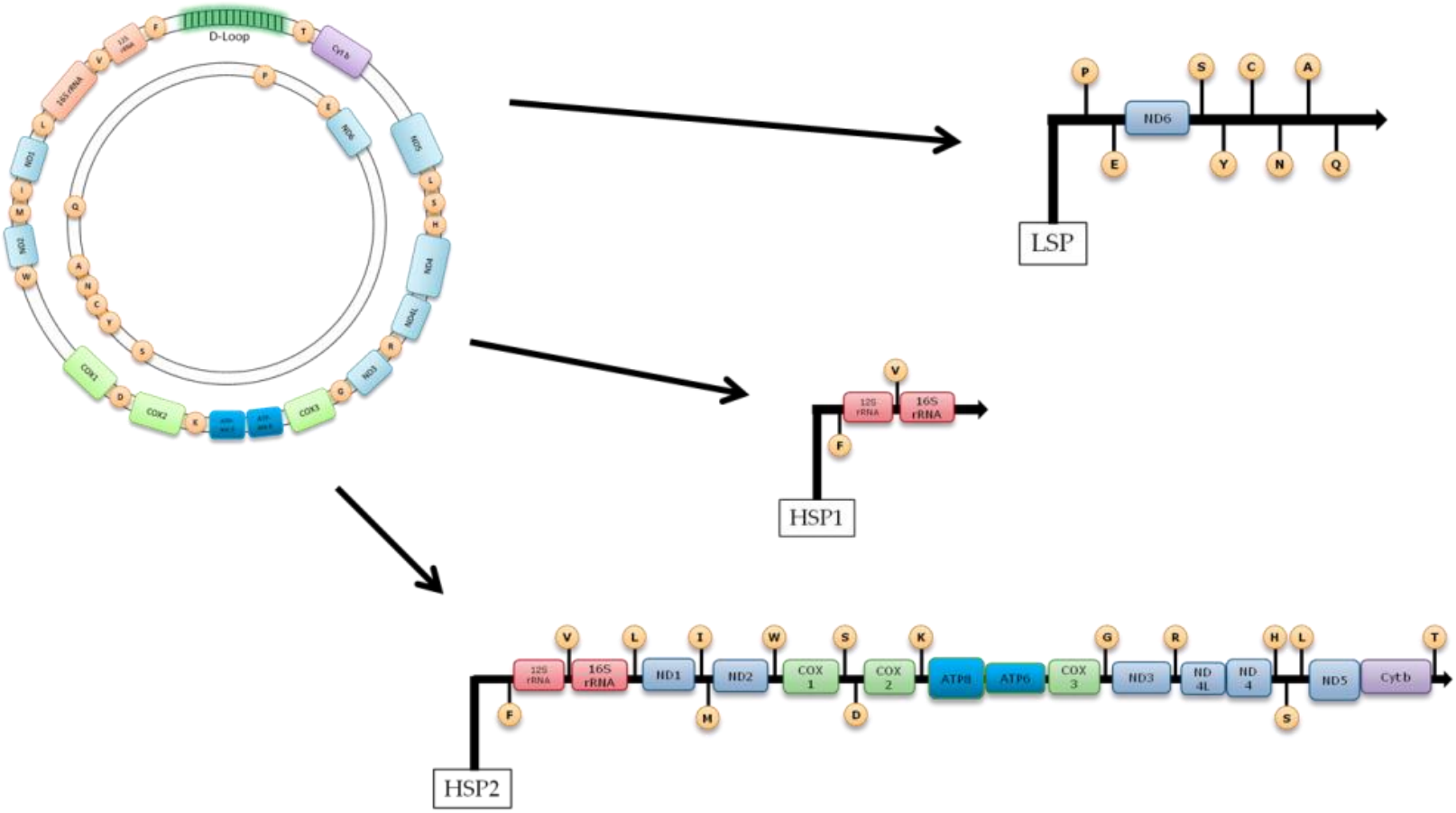


Figure 1.3 Mitochondrial DNA transcription.

The D-loop contains one promoter site for the light strand (LSP) and two for the heavy strand (HSP1 and 2). LSP and HSP2 generate polycistronic transcripts of all the genes on each strand respectively, whereas HSP1 generates a shorter polycistronic transcript containing only the two rRNA genes and two tRNAs. Image courtesy of Dr Casey Wilson.

Post-transcriptional processing of the polycistronic transcripts is proposed to occur by the 'tRNA punctuation' model, whereby the cloverleaf-like folding of the tRNAs act as a substrate for the RNA processing enzymes, (Ojala *et al.*, 1981; Rorbach and Minczuk, 2012) though some questions remain unanswered by this, for instance the processing of mRNAs not separated by tRNAs (Temperley *et al.*, 2010b).

Translation of mitochondrial mRNA occurs in three phases; initiation, elongation, and termination, and occur in the mitochondrial matrix in mitoribosomes, which are constructed of an estimated 80-100 nuclear encoded proteins and the two mitochondrial rRNA species, one in each ribosome subunit (O'Brien, 2003; Smits *et al.*, 2010). Mitochondrial ribosomes differ from both bacterial and eukaryotic cytoplasmic ribosomes in their high protein and low rRNA content. Mitochondrial mRNA translation utilises and requires only 22 tRNAs, as compared to the 31 required for the nuclear genome (Barrell *et al.*, 1980); 8 of the tRNAs can recognise codons with any base in the third position, and 14 of them discriminate between pyrimidine and purines at the third position, thus recognising 60 codons in total; two of the remaining four codons (UAA and UAG) are stop codons recognised by the mitochondrial translational release factor 1a (mtRF1a), and the remaining two (AGA and AGG) are proposed to cause a -1 frameshift which moves the ribosome to a recognised UAG stop codon (Temperley *et al.*, 2010a).

1.1.9. Replication

mtDNA is replicated independently of nDNA and replicates throughout the cell cycle (Bogenhagen and Clayton, 1977), including post-mitotic cells such as skeletal muscle and neurons (Reeve *et al.*, 2009), though the turnover rate in such cells is thought to be very slow compared to mitotically active cells (Wang *et al.*, 1997). Replication and transcription have been shown to be highly co-ordinated with the cell cycle, at least in cells synchronised by serum depletion; mtDNA replication peaks towards the end of the G_1 phase preceding nDNA replication in the S phase, and has a second peak towards the end of the S phase after nDNA replication preceding mitosis (Chatre and Ricchetti, 2013) The replication process is reported to take 60 (Clayton, 1982) to 75 (Korr *et al.*, 1998) minutes. Mitochondria are entirely dependent on nuclear encoded proteins for DNA replication and maintenance machinery (Shadel, 2008).

There are several major competing theories regarding the process of mtDNA replication. The first and more entrenched theory is known as the 'asynchronous' or

'strand displacement' model (Clayton, 1982), which suggests that replication initiates at the origin of heavy strand replication (O_H) and two-thirds of the H-strand is replicated before O_L (the L-strand origin of replication) is exposed and light strand replication initiates. Once the strands are completed they are circularised, superhelical turns are introduced, and finally the D-loop is replicated. More recently, a second method known as the 'synchronous' model has been proposed (Holt *et al.*, 2000), based on evidence of replication intermediates that are resistant to single-strand nucleases. This model was refined based on further experiments that demonstrated replication occurs bi-directionally, initiates from a broad range of the genome incorporating *MT-ND5*, *MT-ND6* and *MT-CYB* and terminates at rather than initiates from O_H (Bowmaker *et al.*, 2003). A third mechanism, was proposed that is similar to the asynchronous model but involves simultaneous binding of RNA to the lagging strand (the as yet unreplicated L-strand) as the H-strand is replicated, which is then subsequently converted to DNA; this method is known as RITOLS (ribonucleotide incorporation throughout the lagging strand), and was evidenced by the activity of RNase H on replication intermediates, since this enzyme acts only on RNA hybridised to DNA (Yasukawa *et al.*, 2006).

The currently known essential machinery for mtDNA replication are the mitochondrial DNA polymerase gamma (POLG) comprising a catalytic unit POLG and accessory subunit POLG2, the mitochondrial helicase TWINKLE, the mitochondrial single-stranded DNA binding protein (mtSSB) and the mitochondrial RNA polymerase (POLRMT) (McKinney and Oliveira, 2013).

1.2. Mitochondrial genetics

1.2.1. Heteroplasmy

As cells contain multiple copies of the mitochondrial genome, it is possible for cells to harbour mtDNA with different polymorphic or pathogenic variations. This mixture of different mitochondrial genomes is termed heteroplasmy. The co-existence of wild-type and mutant mtDNA species within individual cells is a fundamental aspect of mitochondrial genetics and the phenotypic expression of mutation.

Quantitatively, the word heteroplasmy is generally used to refer to the proportion of pathogenic mutant mtDNA molecules within a cell or tissue, expressed as a percentage. The proportion of wild-type is therefore 100% minus the heteroplasmy level in the case of a single pathogenic variant.

1.2.2. Clonal expansion

Clonal expansion refers to the process of a mutated mtDNA molecule expanding in population within a cell, increasing the level of heteroplasmy. Thus a single mutated mtDNA molecule may expand in influence until a cell or tissue (generated through mitosis) contains a majority of the mutated molecule or even become homoplasmic.

The mathematics and biology underlying clonal expansion are not well understood. It was initially thought that a mutation such as a deletion would lead to a smaller mtDNA molecule that would replicate faster and therefore tend to accumulate (Wallace, 1992); though one group does find evidence to support this (Diaz *et al.*, 2002; Fukui and Moraes, 2009), this is in non-physiological conditions, and other work in trans-mitochondrial cell lines it has been shown not to be the case (Tang *et al.*, 2000). It is argued that the time between replications is far greater than the replication time (the half-life of mtDNA is between 8 and 23 days whereas the time for replication is little over an hour (Korr *et al.*, 1998)) and thus faster replication would confer no advantage. However, it has also been shown that most mtDNA replication occurs in peri-nuclear mitochondria (Davis and Clayton, 1996) which could mean only a small population of cellular mtDNA undergo replication at a more rapid rate to maintain the overall population; yet this finding is also contradicted by a more recent study that demonstrated replication throughout the cytoplasm (Magnusson *et al.*, 2003).

Other theories that also explain expansion for non-deletion mutations proposed that mitochondria with mutations would proliferate in order to overcome respiratory chain deficiencies, causing clonal expansion (Yoneda *et al.*, 1992), or that mutated mtDNA leads to slower OXPHOS and less ROS than wild-type mitochondria which are degraded at a higher rate as a result (de Grey, 1997). It has also been shown, however, that even without a selective advantage random intracellular drift leads can lead to mutations expanding to high levels (Elson *et al.*, 2001).

Clonal expansion has been shown to occur with ageing in many tissues, though there are tissue specific differences in the mutations that accumulate, for instance point mutations but not deletions in the colon (Taylor *et al.*, 2003), whereas deletions are frequently found in muscle (Fayet *et al.*, 2002; Yu-Wai-Man *et al.*, 2010a) and the brain (Kraytsberg *et al.*, 2006). Though the reasons for this tissue specificity are unproven, it is postulated that mitotic tissues may exert negative selective pressure on deletions.

1.2.3. Threshold effect

Pathogenic mutations at a low level of heteroplasmy are thought not to exhibit an observable phenotype; OXPHOS impairment and other primary or secondary phenotypic effects are observed only when heteroplasmy reaches a critical threshold (Rossignol *et al.*, 2003). This has been documented in several tissues, for instance muscle fibres from MERRF (myoclonic epilepsy with ragged red fibres) and MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes) patients, where a cellular threshold level of around 80% (Taylor and Turnbull, 2005) or 90% (Shoffner *et al.*, 1990) have been reported, though it a threshold for a severe defect has also been reported at around 60% (Miyabayashi *et al.*, 1992). Large-scale single deletions are generally thought to have a lower threshold at around 60% (Hayashi *et al.*, 1991; Rossignol *et al.*, 2003). Measurements and estimates vary, but it is thought that the threshold is likely to vary according to tissue type, dependent on the level of energy requirement or tissue specific nuclear gene expression; it has been recently shown that for mtDNA deletions the extra-ocular muscles have a significantly lower threshold for expression of an OXPHOS defect than skeletal muscle (Greaves *et al.*, 2010), though no difference in threshold for point mutations.

This issue will be discussed in greater detail with regard to m.3243A>G in Chapter 4 and single-large scale mtDNA deletions in Chapter 5.

1.2.4. Segregation

As a result of continuous replication, segregation of mtDNA mutations is possible in both mitotic and post-mitotic tissues. Random segregation was proposed to explain the observed tissue specific segregation in many mutations (Macmillan *et al.*, 1993), though there is evidence that segregation is not random (Raap *et al.*, 2012), and specific tissues demonstrate consistent alteration from the average heteroplasmy level of an individual, for instance in m.3243A>G mutation tends to be highest in muscle and urine and lower in buccal mucosa, hair, and blood (Chinnery *et al.*, 1999). Specific nuclear encoded mitochondrial proteins that affect this segregation are beginning to be identified, for example the GIMAP3, an outer mitochondrial membrane GTPase (Jokinen *et al.*, 2010). This issue will be considered in more depth in Chapter 4 in relation to the m.3243A>G mutation, particularly with regard to the dynamic change of heteroplasmy through the life of patients in different tissues.

1.2.5. Inheritance

Inheritance of mtDNA is thought to be strictly maternal, though there is one documented case of an inherited paternal mtDNA deletion (Schwartz and Vissing, 2002). Destruction of sperm mitochondria is thought to occur post-fertilisation by proteasomal degradation (Sutovsky *et al.*, 2000) though autophagy has also been demonstrated (Al Rawi *et al.*, 2011), and more recently pre-fertilisation degradation has been proposed (Luo *et al.*, 2013).

The mtDNA in oocytes arise from primordial germ cells, which have been shown to have a very low number of mtDNA molecules, around 200 (Jenuth *et al.*, 1996; Cree *et al.*, 2008) . This is thought to cause what is termed a 'genetic bottleneck'; the restriction and then re-amplification of mtDNA leads to random shifts of heteroplasmy from one generation to the next, observable in the variability of heteroplasmy in the children of mothers with pathogenic mtDNA mutations (Taylor and Turnbull, 2005).

1.3. Mitochondrial disease

1.3.1. mtDNA mutations

mtDNA mutations are prevalent, causing disease in at least 1 in 5000 people and estimated to be present asymptomatically in at least 1 in 200 live births, whilst *de novo* mutations are thought to occur in at least 1 in 1000 live births (Elliott *et al.*, 2008). Current treatment options are limited (Pfeffer *et al.*, 2012) and clinical care is generally focussed on management of complications (Horvath *et al.*, 2008).

The mitochondrial genome is thought to be significantly more vulnerable to mutation than the nuclear genome for a number of reasons, including the proximity to ROS produced from the ETC, the lack of protective histones for the DNA, and fewer repair mechanisms than for nuclear DNA. It is estimated that the mutation rate of mtDNA is ten times that of nuclear DNA (Brown *et al.*, 1979).

Recent research into mtDNA mutation have shown that levels of mutation are often surprisingly low even in aged individuals, indicating that repair or degradation mechanisms of damaged mtDNA are more robust than previously thought (Shokolenko *et al.*, 2009). These mechanisms are entirely nuclear in origin, and include base excision repair, single-strand break repair, and mismatch repair (Kazak *et al.*, 2012), and there is also evidence of homologous recombination after double-strand breaks (Bacman *et al.*, 2009; Fukui and Moraes, 2009) (though the possibility of mtDNA recombination is hotly disputed (Stewart *et al.*, 2008a)). In spite of this, however, the human

mitochondrial genome still has a high mutation rate, and deleterious mutations associated with disease are the focus of this particular work. Other mutations, such as somatic mutations accumulated through life and associated with ageing, or those encountered as ancient regional variations in the human population, are also frequent and important areas of research (Wallace, 2010).

1.3.1.1. mtDNA point mutations

Point mutations are single base pair substitutions in mtDNA, and were first identified as a cause of clinical disease in 1988 when Wallace *et al.* associated a mtDNA point mutation with Leber's hereditary optic neuropathy (LHON) (Wallace *et al.*, 1988). The m.3243A>G related to MELAS was identified shortly afterwards (Goto *et al.*, 1990b; Kobayashi *et al.*, 1990), along with the m.8344A>G mutation associated with MERRF. Epidemiological studies have estimated the minimum prevalence of the most common point mutation, m.3243A>G, to be 3.65/100,000 (Schaefer *et al.*, 2008) or 16.3 (Majamaa *et al.*, 1998), though the first of these studies was based on attendance at mitochondrial clinic and likely to be a significant underestimate, and the second based on identification of patients with moderate to severe neurological symptoms. The most recent studies in large groups not *a priori* identified by symptoms reported prevalence of 236/100,000 (Manwaring *et al.*, 2007) and 140/100,000 (Elliott *et al.*, 2008), considerably higher than previous estimates.

Mutations in protein coding genes may be synonymous (resulting in an unchanged amino acid coding) or non-synonymous. However, the majority of point mutations are identified in tRNA genes, and most inherited mutations are tRNA mutations. Studies on inheritance of mtDNA mutations in mice using POLG mutator mice (which cause frequent sporadic mutations) followed by backcrossing to wild-type males to remove the POLG mutation whilst leaving the existing mtDNA mutations demonstrated that most non-synonymous changes in protein coding genes were lost within two generations, suggesting strong selection against deleterious mutations (Stewart *et al.*, 2008b). Other studies have also demonstrated similar findings (Fan *et al.*, 2008). Elson *et al.* speculated that the most pathogenic mutations (with the most pronounced biochemical defect) would be sporadic rather than inherited, and found that though the mutation threshold for respiratory chain deficiency in muscle was not associated with inheritance, there was a clear difference in blood, where sporadic mutations were generally undetectable (Elson *et al.*, 2009).

The most common point mutation is the m.3243A>G mutation; however this mutation is the focus of Chapter 4 and will not be discussed further in this introduction. However, several other mutations will be discussed briefly. Figure 1.4 depicts the position of several known mutations on the mitochondrial genome. Though rare, multiple pathogenic point mutations in patients have been reported (Nakamura *et al.*, 2010).

The m.8344A>G mutation in $tRNA^{Lys}$ is most commonly associated with MERRF syndrome (myoclonic epilepsy with ragged red fibres) and was first identified in 1990 in association with this syndrome (Shoffner *et al.*, 1990; Yoneda *et al.*, 1990). The mutation was demonstrated to cause a severe reduction in protein synthesis (Chomyn *et al.*, 1991; Yoneda *et al.*, 1994).The mutation is in the TΨC loop of the tRNA, and has been shown to cause a loss of the usual taurine modification of the uridine wobble-position anticodon; this is thought to weaken the codon-anticodon binding and thereby impede protein translation, as it has been shown to lead to the tRNA being unable to translate it's cognate codon and leads to ribosomal stalling (Yasukawa *et al.*, 2000a; Yasukawa *et al.*, 2001). It does not appear to cause mistranslation of non-cognate codons. MERRF is a multi-systemic disorder, characterised by myoclonus, but patients demonstrate a wide range of neurological defects such as ataxia, generalised epilepsy, weakness, and dementia. 80% of patients with MERR F syndrome carry the m.8344A>G mutation. It has recently been suggested that the m.8344A>G mutation is more aptly named myoclonic ataxia than myoclonic epilepsy as this seems to be the more common phenotypic presentation (Mancuso *et al.*, 2013a).

Though tRNA mutations are more commonly observed, there are some relatively common mutations in protein encoding genes, for instance in the ATPase subunit 6 gene (*MT-ATP6*). Mutations in this gene include the m.8993T>G mutation usually associated with NARP (neuropathy, ataxia, and retinitis pigmentosa) (Schon *et al.*, 2001), though it has also been associated with Leigh's syndrome, a devastating infant onset progressive neurological disorder resulting in childhood mortality, with necrotic lesions of the brain stem, basal ganglia and thalamus observed *post mortem* (Leigh and Thompson, 1951). Other mutations in this gene associated with Leigh's syndrome are m.9176T>G (Carrozzo *et al.*, 2001) and m.9176T>C (Thyagarajan *et al.*, 1995).

Most (> 95%) of mutations that lead to Leber's hereditary optic neuropathy (LHON) occur in Complex I subunit genes, for instance *MT-ND1* (m.3460G>A), *MT-ND4* (m.11778G>A) and *MT-ND6* (m.14484T>C), which all cause loss of retinal ganglion

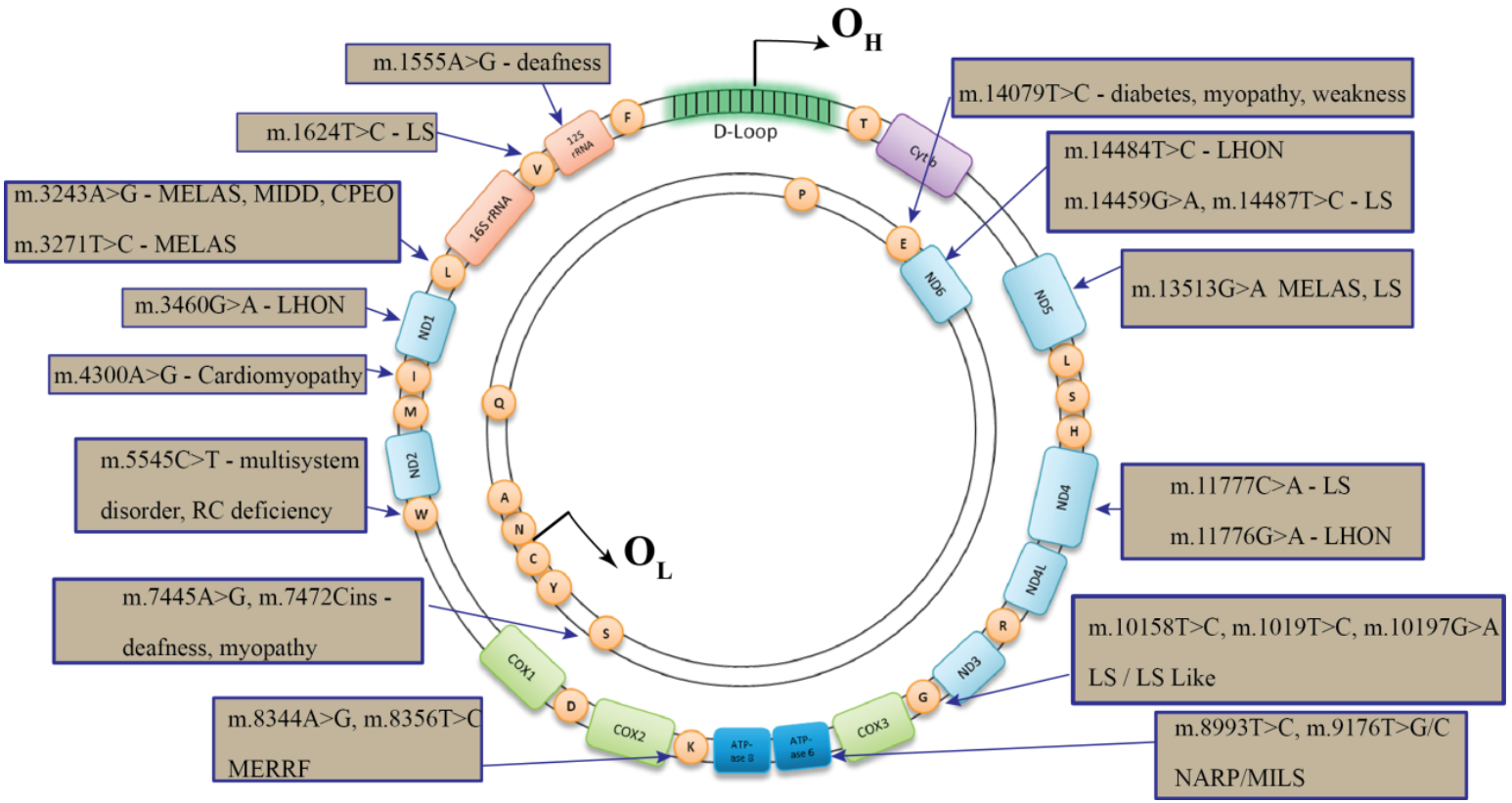


Figure 1.4 Point mutations associated with disease.

Point mutations and the associated disease phenotypes are shown, connected by an arrow to their associated gene. Abbreviations: LHON – Leber's hereditary optic neuropathy. LS - Leigh syndrome. MELAS - mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes. MERRF – myoclonic epilepsy and ragged red fibres. MIDD – maternally inherited diabetes and deafness. MILS – maternally inherited Leigh syndrome. NARP – neurogenic muscle weakness, ataxia, retinitis pigmentosa. PS – Pearson's syndrome.

cells in the optic nerve, leading to sudden acute or sub-acute visual loss (Chinnery *et al.*, 2000b).

1.3.1.2. mtDNA rearrangements

There are two major forms of mtDNA rearrangement that have been considered in a disease context, mtDNA deletions (molecules with missing sections of genetic material), and duplications (molecules with extra repeated genetic information).

1.3.2. Single large-scale mtDNA deletions

Single-large scale mtDNA deletions are a common cause of mitochondrial disease, and are generally found at heteroplasmic levels in multiple tissues of an affected individual. Such deletions are found throughout the mitochondrial genome though predominantly within the major arc (between the two origins of replication, as shown in Figure 1.2). Deletions vary in size, but the most frequently reported is the so called 'common deletion' of 4,977bp extending from base 8,470 in *MT-ATP8* to base 13,477 in *MT-ND5*, flanked by a 13bp direct repeat (Zeviani *et al.*, 1988).

It was long thought that deletions were most likely to occur through errors in mtDNA replication (Shoffner *et al.*, 1989), though the proposed mechanism requires unprotected single-stranded DNA and, as discussed in section 1.1.9, this is not thought to occur. Additionally, mitotic tissues, where mtDNA replication proceeds at a faster rate than post-mitotic, should be expected to demonstrate faster accumulation of deleted species, however deletions are rarely found in mitotic tissue such as the colon (Taylor *et al.*, 2003) but are much more commonly found in post-mitotic tissue such as muscle and the brain, and have been shown to accumulate with age (Cortopassi *et al.*, 1992; Melov *et al.*, 1995; Kraytsberg *et al.*, 2006). More recently a mechanism based on errors during repair of double-stranded breaks has been proposed, caused by direct homologous repeats on either side of the break, which is consistent with the pattern of deletions found in tissues and ageing (Krishnan *et al.*, 2008) and also with the flanking of most deletions by direct repeats (Samuels *et al.*, 2004).

Though common, deletions are generally thought not to be inherited through the germline, though there is an isolated report (Shanske *et al.*, 2002).

Disease associated with single large-scale mtDNA deletions are discussed in depth in Chapter 5.

1.3.3. Multiple deletions

Multiple mtDNA deletions are also found, as a secondary effect of nDNA mutation, and are discussed in section 1.3.5.

1.3.4. Duplications and dimers

A second major class of rearrangement are mtDNA duplications, which have been associated with disease in patients with myopathy and multisystemic features (Poulton *et al.*, 1989; Rotig *et al.*, 1992; Poulton and Holt, 1994; Martin Negrier *et al.*, 1998). However, it has been shown that the disease state of an individual is generally governed by the presence of mtDNA species with deletions and that duplications are most likely to not exhibit pathological effects (Manfredi *et al.*, 1997). Duplication dimers, consisting of repeated deleted mtDNA species, are also reported (Brockington *et al.*, 1995; Jacobs *et al.*, 2004), though the pathogenicity of these are thought to be the same as duplication monomers.

1.3.5. Nuclear DNA mutations

Of the estimated 1500 proteins that functionally make up the mitochondrion, the mitochondrion itself encodes merely thirteen; the nuclear genome encodes the remainder (Zhu *et al.*, 2009). Mutations in these genes are responsible for a great number of disease phenotypes. There are over 80 genes in the OXPHOS complexes alone, and many others with a critical role in assembly, catalytic regulation, stability, and maintenance of the complexes and supercomplexes (Vartak *et al.*, 2013) with novel contributors being recognised at a progressive pace (Ikeda *et al.*, 2013).

The nuclear encoded mitochondrial replication enzymes are a critical breakpoint in mitochondrial function. Mutations in the *POLG* and *PEO1* genes encoding the POLG cayaltyic subunit and TWINKLE helicase respectively are associated with a spectrum of disease caused by transcriptional defects including multiple deletions and point mutations that are found in affected tissues, which commonly present as chronic progressive external ophthalmoplegia (CPEO) (Hudson and Chinnery, 2006; Fratter *et al.*, 2010; Wallace, 2010). Pathogenically low levels of mtDNA are observed in patients with *POLG* mutations that lead to Alpers-Huttenlocher syndrome (Davidzon *et al.*, 2005) or mitochondrial neurogastrointestinal encephalomyopathy (MNGIE), caused by mutation in the *TYMP* gene for thymidine phosphorylase (Nishino *et al.*, 2001). This is known as mtDNA depletion. *TWINKLE* mutations are also associated with mtDNA depletion (Sarzi *et al.*, 2007), though they are more commonly associated with the CPEO phenotype.

Mutations in genes related to mitochondrial dynamics are also a known cause of disease, for instance *OPA1* which encodes a protein with a critical role in mitochondrial fusion is named from its association with autosomal dominant optic atrophy (Delettre *et al.*, 2000).

This is a very brief overview of just a small number of known nDNA encoded mitochondrial proteins that have known associations with disease; though this thesis is focussed on mtDNA mutation, it is important to acknowledge that the nuclear genome is a deeper and broader source of mitochondrial disease.

1.3.6. Phenotype and progression of mtDNA disease

The phenotypic presentation of mtDNA disease is highly variable; part of this variability has long been attributed heteroplasmy (Ciafaloni *et al.*, 1991), though specific mtDNA mutations are also associated with certain characteristic clinical phenotypes, such as the m.3243A>G mutation with MELAS or MIDD (maternally inherited diabetes and deafness), m.8344A>G with MERRF, and single large-scale deletions with CPEO, Kearns-Sayre syndrome (KSS) and Pearson's syndrome (PS). However, the links between genotype and phenotype are not exclusive and frequently overlap; MELAS is associated with a host of other mutations, as indeed are MERRF, MIDD, and CPEO, which in particular is associated with a variety of mutations, both deletions, point, mutations, and nuclear defects.

The variety of clinical symptoms reported in mitochondrial disease is vast; myopathy and cardiomyopathy are common, as are neurological features such as cerebellar ataxia and epilepsy, but there are few if any organs or systems that are not reported as affected by mtDNA disease in some form. This heterogeneity is a problem for clinical diagnosis (McFarland and Turnbull, 2009).

Though universally recognised as clinically progressive (Zwirner and Wilichowski, 2001; Arpa *et al.*, 2003; Taylor and Turnbull, 2005; Majamaa-Voltti *et al.*, 2006; Whittaker *et al.*, 2007; Horvath *et al.*, 2008; Coku *et al.*, 2010; Chen *et al.*, 2012), disease progression in patients with mtDNA mutation is little understood, both in terms of the likelihood of development of specific system involvement, or the rate of progression of the burden of disease on patients.

Discussion of the phenotypic presentation of the mutations studied in this thesis will be considered in detail in chapters devoted to each mutation.

1.4. Data sources

1.4.1. The MRC Mitochondrial Diseases Patient Cohort Study UK

The Mitochondrial Diseases Patient Cohort Study UK is a large cohort of living patients with genetically and/or biochemically confirmed mitochondrial disease. It is the largest such cohort globally, and is funded by the Medical Research Council (MRC) Centre for Translational Research in Neuromuscular Diseases. Symptomatic adults and children comprise the majority of the cohort, but there are additional asymptomatic individuals who have requested genotyping (usually due to a family history) and have proved positive. All individuals are phenotypically characterised in out-patient clinics, on the basis of examination, clinical history, and detailed investigation.

This patient cohort forms the basis of the studies conducted for this thesis.

1.4.2. The NHS Specialised Services for Rare Mitochondrial Diseases

The NHS Specialised Services for Rare Mitochondrial Diseases is situated in Newcastle upon Tyne. A large part of the diagnostic tissue, samples, and genetic information used in this work have been provided by the service.

1.5. The NMDAS

The Newcastle Mitochondrial Disease Adult Scale (NMDAS) was published in 2006 as a semi-quantitative rating scale to monitor mitochondrial disease (Schaefer *et al.*, 2006). It is a clinically validated tool, that has been extensively used both at our centre (Apabhai *et al.*, 2011; Bates *et al.*, 2012b; Lax *et al.*, 2012) and other specialist mitochondrial centres (de Laat *et al.*, 2012; Enns *et al.*, 2012; Orsucci *et al.*, 2012; Yatsuga *et al.*, 2012; Kornblum *et al.*, 2013; Mancuso *et al.*, 2013b). The NMDAS permits quantitative analysis of both general and system-specific disease progression and has already been used in assessment of clinical progression in patients with the m.3243A>G mtDNA mutation, although this was not a longitudinal study (Whittaker *et al.*, 2009). The Newcastle Paediatric Mitochondrial Disease Scale (NPMDS) is a similar scale used to monitor paediatric mitochondrial disease patients (Phoenix *et al.*, 2006).

Development of the NMDAS was prompted by success of other similar assessments for other mainly neurological conditions, such as those developed for Parkinson's Disease (Ebersbach *et al.*, 2006). It was developed to meet several key objectives; (1) to reflect the multi-dimensional nature of mitochondrial disease, (2) to monitor both the progress of the underlying dysfunction and the functional impact of the dysfunction (3) to allow

input from multiple sources, including the patient, carer, clinician, and clinical records (4) to be concise and straightforward to complete. To achieve this aim, the test comprises 4 sections, or domains; I, which assesses current function; II, which measures system specific involvement; III, which is a current clinical assessment; and IV, a quality of life survey.

Sections I to III are made up of several questions, each scoring from 0 (no involvement) to 5 (severe involvement). Section I comprises 10 questions assessing current function, and is in the main concerned with the impact of disease on the daily life and functions of the patient. Section II comprises 9 questions assessing system specific involvement, including neurological, respiratory and the cardiovascular system. Section III comprises 10 questions to summarise a general and neurological clinical examination, including three cognition tests to generate a combined cognition score. The final section of the assessment comprises the SF-12v2 quality of life survey (Ware *et al.*, 1996). The NMDAS questionnaire can be found in Appendix I.

1.6. Aims

From a clinical care perspective, it is vital to improve understanding of the progression of disease in patients with mtDNA disorders, in order that management of the disease can be tailored to patient needs, and care and monitoring be pre-emptive rather than reactive. Thus the overarching aim of this study is to improve understanding of the clinical progression of disease associated with mtDNA mutations.

To achieve this aim, I chose to study the mutations that are most prevalent in the cohort of adult patients regularly monitored in the Newcastle National Commissioning Group (NCG) Mitochondrial Disease Service. These are the patients carrying the m.3243A>G mutation and the patients with single large-scale mtDNA deletions.

There were two main focal points in the study of each patient group. The first was to improve understanding of the progression of total disease burden of patients, as measured by the score achieved on the NMDAS assessment. The second was to improve understanding of the development and progression of individual phenotypic features associated with each of these mutations, using the individual features examined in the NMDAS assessment. For both of these foci I aimed to investigate not only the predicted progression over time, using age as a predictor of disease burden, but also genetic factors which influenced the progression of disease in patients.

Chapter 2

Chapter 2. Statistical methodology

2.1. Introduction

The use of statistics in medical and biomedical fields is widely accepted as a powerful and necessary tool in the scientific research process. Statistical methods have been derived specifically to address the needs of researchers (Altman, 1981; Altman, 1982; Altman *et al.*, 1983), along with guidelines developed for reviewers of statistical publications to ensure high standards of peer review (Gore *et al.*, 1992; Altman, 1998; Goodman *et al.*, 1998; Altman *et al.*, 2002). However, there is a consensus that statistical reporting in medical journals is of low standard, with frequent erroneous use of statistical methodology and reporting (Gore *et al.*, 1977; Gardner *et al.*, 1983; Andersen and Forrest, 1987; Dar *et al.*, 1994; Porter, 1999; Gardenier and Resnik, 2002; Nagele, 2003; Marshall, 2004). Misuse of statistics is unethical and has the potential to lead to serious clinical consequences (Strasak *et al.*, 2007).

As this thesis is focussed on using statistical methodology to understand mitochondrial disease, this section will give a brief overview of all the statistical methods employed throughout this study. The less familiar statistical techniques will be discussed in more detail.

2.2. Basic statistical techniques

Basic statistical approaches can be classified into two broad categories of approach, parametric and non-parametric. The majority of parametric tests are suitable for use with data that is normally distributed. The t-test is ubiquitous and is the basic parametric test used throughout statistical analysis. Non-parametric approaches to analysis are subject to less stringent criteria regarding the shape or form of the data, but are also generally less powerful than the equivalent parametric test. The non-parametric equivalent of the t-test is the Mann-Whitney U test. Note that there are still assumptions associated with the Mann-Whitney U test; the data must be ordinal for instance. Additionally, all observations must be independent.

2.2.1. Linear regression

Linear regression is at the heart of many parametric statistical analysis techniques. The principles of linear regression are taught from an early age in school (the concept of the line of best fit, or regression line). Computational methods for deriving the regression line and quantifying the variability of data are more advanced. An understanding of

linear regression is a vital foundation for understanding the more complex statistical techniques used in this study.

Figure 2.1 illustrates an example of linear regression (or simple linear regression), and defines the residual, which is the distance of each data point from the regression line. Residuals allows us to calculate R^2, the coefficient of determination, a measure of how well the model (regression line) fits the data; it is defined by summing the squares of the residuals in the fitted model, summing the squares of the residuals without fitting the model (so the differences from the average y value), and taking the ratio of these quantities. It is subtracted from 1 to give a value that is 1 for perfect correlation and 0 for no correlation. Mathematically, R^2 is also equal to the square of the Pearson's correlation coefficient.

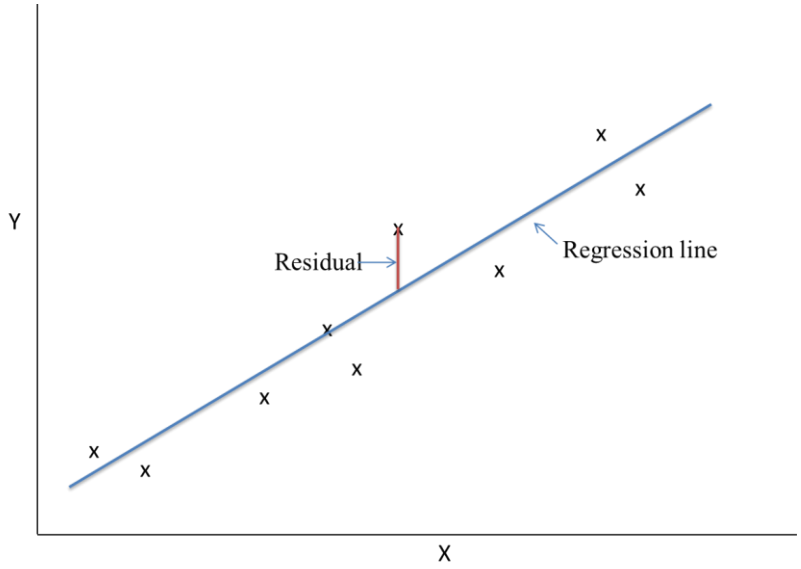


Figure 2.1 Example of linear regression.

The regression line (blue) is shown. The *residual* for each point, the vertical distance from the regression line, can be calculated. Computational methods for regression calculate the regression line by Ordinary Least Squares (OLS), which positions the regression line such that the sum of the squares of the residuals is minimized.

2.3. Multiple regression

Linear regression can be extended to encompass more than one predictor. This is difficult to visualise, as it requires a dimension for each predictor (or X). However, the principles of regression are the same as for simple linear regression; for two predictors the line of regression becomes a plane of regression in a total of three dimensions, with three predictors we would need to visualise a four dimensional space, and so on.

2.4. General linear model (GLM)

The t-test is the simplest test for comparison of the difference between the means of two groups. ANOVA is the extension of the t-test to multiple groups. The general linear model (GLM) subsumes the t-test, linear regression, ANOVA, and several other related statistical techniques such as ANCOVA, MANOVA, and MANCOVA into a single umbrella analysis; this is possible as they are all based on the same basic linear modelling principles. The general linear model can be written as shown in Equation 2.1. where y is the vector of observations of the dependent variables(s), X is the vector of observations of the predictors, β is the design matrix (that we are estimating with our model), and ε is the vector of random errors.

$$y = \boldsymbol{X}\beta + \varepsilon$$

Equation 2.1 General linear regression (GLM).
y is the vector of observations of the dependent variables(s), X is the vector of observations of the predictors, β is the design matrix (that we are estimating with our model), and ε is the vector of random errors.

GLM assumes that the errors in ε are independent and normally distributed, with a mean of zero. Note that β and X are one larger than the number of dependent variables in the model as they include the intercept term of the model.

For many analyses, there is only one dependent variable, in which case y, β, and ε are vectors. For multivariate analysis (so the extension of MANOVA and MANCOVA) these become matrices.

2.5. General linear mixed model (GLMM)

General linear mixed models, also known simply as mixed models, are statistical models that incorporate both fixed effects and random effects. A major development that separates mixed modelling from its predecessors is that it allows for modelling of observations that are not independent, by incorporating into the model the correlation structure of the errors. They are particularly useful in situations where repeated measurements are taken from the same observation subject (Macchiavelli and Moser, 1997).

The mathematical description of the mixed model is shown in Equation 2.2.

$$y = \boldsymbol{X}\beta + \boldsymbol{Z}\gamma + \varepsilon$$

Equation 2.2 General linear mixed model (GLMM).

β and γ are vectors of fixed effect and random effect regression parameters respectively which are to be determined, $\boldsymbol{X}$ and $\boldsymbol{Z}$ are matrices relating the observations y to β and γ, and ε is a vector of normally distributed errors with zero mean and equal variance.

This model has two critical qualities; the (potentially) correlated random effects described by $\boldsymbol{Z}\gamma$, and the fact that the rows of the vector ε are not required to be mutally independent as for traditional regression, but can be structured to allow covariance between errors. This allows repeated measures analysis to be performed.

In contrast to GLM, mixed models are fitted using a maximum likelihood approach, which unlike least squares methodology is robust in the case of missing data where multiple recordings are made from subjects (Moser, 2004). Principal assumptions of GLMM are the normality of residuals ε, which can be checked by residual diagnostics, and the homogeneity of variance across between-subject factors (homoscedacity) (Chiarotti, 2004).

Mixed models are extensively used in biostatistical studies and many other fields as they permit a flexibility of model that cannot be achieved by GLM or ordinary linear regression (Wolfinger, 1997). The approach has been widely used to understand disease progression in many neurodegenerative conditions including dementia (den Heijer *et al.*, 2010; Galvin Je and *et al.*, 2005; Hassing *et al.*, 2004; Johnson *et al.*, 2009; Knopman *et al.*, 2009; Tornatore and Grant, 2002), Parkinson's disease (Dobkin *et al.*, 2011; Johnson and Galvin, 2011; Nandhagopal *et al.*, 2009; Vu *et al.*, 2012) and Multiple Sclerosis (Meier *et al.*, 2007).

2.5.1. Repeated measures

A key characteristic of the cohort data is that it is longitudinal. *Repeated measures* is used to describe data composed of multiple observations of the same sampling unit. Usually it is the case that repeated observations are correlated, and this correlation must be incorporated into the linear model for appropriate inferences to be made (Littell *et al.*, 2000; Moser, 2004).

2.5.2. Fixed and random effects

Random effects allow modelling in which experimental data can be considered samples from a larger population. As an example, in a drug trial of two drugs A and B, the effect of the drug is a fixed effect, whereas variation attributable to individual patients is

considered a random effect. This distinction is important in order to correctly estimate variation in the model (Laird and Ware, 1982).

In a model incorporating random effects the observations are no longer independent, even if the errors are, since all observations are dependent on a shared set of random predictors.

For fixed effects we are generally concerned with estimating means, whereas for random effects we are not generally interested in specific differences in means between one factor and another, but how much variance in the dependent variable can be explained by the random factor.

2.5.3. Covariance structure

Traditional repeated measures analysis is designed to deal with experimental design that has an equal number of measurements taken for each subject at equivalent intervals. The data used in this study are considerably more complex in form; data for each subject have a generally unique temporal layout, depending on when patients are assessed in clinic, and the number of data points available varies considerably.

The most appropriate covariance structure for such data is a spatial power structure (Moser, 2004), the simplest of which, denoted in SAS as SP(POW), can be written as Equation 2.3.

$$Cov\left(\varepsilon_{ij}, \varepsilon_{i'j}\right) = \sigma^2 \rho^{\delta_{ii'}}$$

Equation 2.3 Spatial power covariance structure.

j indexes the subjects, i indexes the time points, σ^2 is the common variance of the error terms, $\delta_{ii'}$ is the temporal distance between two data points (measured in arbitrary units of years), and ρ is the correlation parameter to be determined.

2.5.4. Model selection

The prime imperative of model selection is that the model structure is representative of the data and appropriate for the objectives of the model (Diggle, 1988; Lindsey, 1993).

Secondly, modelling assumptions (normality of residuals, indicator variable independence) should be verified. Many graphical approaches used in other modelling approaches common to ordinary regression are applicable, including influence and residual diagnostics (Christensen *et al.*, 1992).

2.5.5. Assessing model fit

There is no simple objective measure of model fitness for GLMM. However, the likelihood based approach provides several useful measures for assessing model fitness, such as Akaike's Information Criterion (AIC) (Akaike, 1974) or Bayesian Information Criterion (BIC) (Schwarz, 1978). All such measures utilise the log likelihood and penalize in some way to promote model parsimony. AIC is generally used where model accuracy of prediction is considered more important than how well the model represents the true underlying data structure (Macchiavelli and Moser, 1997). In the model evaluation in this thesis a pragmatic approach was taken to use both AIC and BIC where they agreed on model preference, and AIC where they did not (Kuha, 2004). It should be noted that the lower the AIC or BIC the better the model describes the data. The AIC can be used for comparing non-nested models as long as the likelihood estimation procedure considers both fixed effects and random effects, i.e. it uses ML (maximum likelihood) not REML (restricted or residual maximum likelihood) (Kreft and de Leew, 1998).

Generation of a mixed effects model is a multi-step process involving iterative changes to fixed effects, random effects and covariance structure, with repeated re-testing of previously optimised model choices.

2.5.6. Restrictions

Multicollinearity occurs where two correlated fixed effects are included in a model together. This must be avoided in model specification, as it leads to imprecise estimations and inflated variance of parameter estimates (Silvey, 1969; Wißmann *et al.*, 2007). Independence of model effects must be verified.

2.6. Key concepts in statistical analysis

Many critical considerations of statistical modelling apply to a number of modelling approaches including linear regression, GLM, or mixed modelling. In all cases, the task of the modeller is to ensure that the model fits the data appropriately and that any modelling assumptions are not severely violated, so that reliable inferences can be made from the outputs of the model.

2.6.1. Residuals diagnostics

Analysis of the residuals from a fitted model is vital to assess whether a model is appropriate for the data, and they must meet certain requirements. At a basic level, residuals should be compared against the predicted value and each of the predictors, to

ensure that there are no trends in the data that show the residuals are not independent. If residuals are ill formed then the model needs to be revised. A common approach to address such issues would be transformation of the dependent or independent variables, discussed in 2.7, in order to improve the linearity of the data.

2.6.2. Influence and leverage

Leverage is a measure of how much a change in a data point's y value will change the outcome of a regression. Each point has an associated leverage. High leverage points are those that are far away from the average x value.

Influential data points are those that would result in a significantly changed outcome if they were omitted from the calculation. Influential data points will tend to have high leverage, but high leverage points are not necessarily influential; if the points happen to lie on or near the line of regression that fits all the other points in the data set then the influence will be small, despite high leverage. Influential points are those that have high leverage but also do not fit the pattern of the other data points very well.

Looking at residuals does not necessarily reveal influential points, since an outlier at a position of high leverage will tend to drag the regression line towards it strongly, which masks the effect. Thus specific influence and leverage analysis needs to be carried out. The Cook's distance (Cook's D) statistic (Cook and Weisberg, 1982) is useful in influence diagnostics. Cook's D for each point is calculated by measuring the effect of removal of the point on the errors of the other points in a regression.

Points with high influence need to be carefully considered, as removal from the model may be warranted.

2.7. Data transformation – Box-Cox analysis

The most powerful modelling techniques currently available generally rely on a linear relationship between independent and dependent variables. However, it is rare for actual data, biological or otherwise, to exhibit a direct linear relationship. Non-linear modelling techniques are possible and frequently employed, but often the simplest and most powerful approach is to attempt to transform the data so that a non-linear relationship can be analysed using linear techniques.

Box-Cox analysis is a specific approach of power transform that is commonly used in statistical analysis (Box and Cox, 1964). In essence, Box-Cox analysis is used to find the transformation of the dependent variable (either a power, e.g. y^2 or $y^{3.5}$, or the log

transformation) that minimizes the variance of the dependent variable when regressed against the independent variable(s). A result of this variance minimization is usually a reduction in skew or other distributional features that complicate analysis, and thus tends to produce data that is more normal in distribution.

Box-Cox is by no means a panacea, and will fail to find a suitable transformation where no simple transformation is optimal. Additionally, it is of limited use in situations where the variance in the model is truly heteroscedastic (Sakia, 1992).

2.8. Logistic regression

Linear regression, GLM, and mixed modelling are all restricted to dependent variables that are continuous; discrete data cannot be analysed with such techniques. Discrete data is commonly encountered however, whether in the form of yes/no binary data, or data on a discrete and limited scale, such as the responses to individual questions of the NMDAS which have a 6 point (0 to 5) scale.

Logistic regression is useful for analysis of such data. In contrast to linear regression, where coefficients are determined for the linear relationship between predictors and the outcome variable, in logistic regression parameters are determined that define the probability of an event occurring, e.g. for a given set of predictors the probability of the response variable being 'yes'. The link between the predictors and the outcome is no longer linear but defined by the logistic function, which can be written as seen in Equation 2.4 for a logistic regression with one predictor:

$$y \; or \; P(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

Equation 2.4 Logistic regression.

Though superficially mathematically more complex than the equation $y = \beta_0 + \beta_1 x$, which is solved for linear regression, the mathematical principles for identifying parameters are the same, and involve estimating the optimal parameters (β).

Similar to mixed modelling, maximum likelihood estimation is generally used to solve the equations to find the optimal parameters, which is an iterative procedure that begins with a tentative solution and iteratively tests solutions until convergence is achieved (any change in the estimated parameters results either in a poorer solution or one that is better by a negligible amount).

Logistic regression has an intuitive interpretation, in that the parameter for each predictor is the log of the odds ratio for a unit change in the predictor variable.

2.8.1. Assessing model fit

Similar to mixed modelling, the goodness of fit of a model can be analysed using the likelihood function, and the Akaike Information Criteria (AIC) provides a method for assessing the effectiveness of model fit whilst controlling for the number of parameters in the model (Akaike, 1974).

Dichotomous models can also be evaluated by using the ROC (receiver operating characteristic) curve (Hastie *et al.*, 2009). The area under the ROC curve (AUC) has a simple interpretation; the AUC is the probability that the regression result for a randomly selected affected individual will be greater than that of a randomly selected unaffected individual. Thus a poor regression has an AUC of 0.5, as this is no better than chance, and a perfect regression an AUC of 1.

2.8.2. Coefficient of determination (R^2)

There is general consensus about the use of the coefficient of determination (R^2) in ordinary least squares (OLS) multiple regression to describe the proportion of explained variance in a model. However, this is not the case in logistic regression. The problem occurs because in OLS there is only one reasonable measure of residual (unexplained) variance, the sum of squares of the deviations from predicted values. For logistic regression the situation is not so simple, as we have several ways to measure deviation (squared difference, entropy, qualitative difference) which are not mathematically (i.e. calculated the same way) or conceptually the same (Efron, 1978). There are also two very different ways of looking at the outcome of the model; either consideration of the (continuous) predicted probabilities that the model generates, or the accuracy of the (discrete) classification of the model. (Menard, 2000). Several pseudo-R^2 statistics have been derived for use in logistic regression, which are discussed by Menard extensively (Menard, 2000) and will not be discussed in detail here. However, in agreement with several published commentaries on the issue (Menard, 2000; Shtatland *et al.*, 2002) I use the pseudo-R^2 defined by McFadden to compare models (McFadden, 1974) where no other comparison was possible.

2.8.3. Ordered logit and multinomial logistic regression

Where discrete data has more than two levels there are two approaches to modelling the data with logistic regression. Multinomial logistic regression can be used, which

requires a set of parameters for each change in level of the dependent variable; for the NMDAS data this would require 5 sets of parameters to define the change from each level 0 to 4 to the next level. A simpler approach, requiring a much reduced number of parameters is to use ordered logit, which makes the assumption of proportional odds (McCullagh, 1980); for the NMDAS data, this would mean the odds of scoring 0 vs 1-5 is the same as scoring 1 vs 2-5, or 2 vs 3-5, etc. This assumption is restrictive and must be tested for data conformity, but if the data structure is appropriate it is a statistically efficient model for analysing discrete ordinal response data.

2.9. Generalised linear modelling

Generalised linear modelling is an extension of mixed modelling that allows non-identity link functions (linking the predictors to the outcomes) to be used; this opens up the possibility of creating models with both fixed and random effects but non-linear outcome variables (Nelder and Wedderburn, 1972). For instance, using the logit function mixed logistic models can be created, and other link functions extend the principles of linear mixed modelling to non-linear modelling.

2.10. Bootstrapping

Bootstrapping is a computational method for estimating the accuracy of a sampling statistic, first developed by Bradley Efron (Efron and Tibshirani, 1986). It is a simple and universally applicable technique for almost any estimation problem and facilitates calculation of the accuracy of estimates in situations where the sampling distributions are too complex for parametric statistical analysis (Johnson, 2001; Christie, 2004). It is one of a set of techniques known more generally as resampling.

Bootstrapping is particularly useful when we are interested in estimating properties of an estimator itself; for example, calculating the accuracy of a measure such as the standard error. For illustration, in a particular population we may take a random sample of people, measure their heights, and use this to calculate the mean height and the standard error. The standard error gives an estimate of the variation of heights in the population. However, there is no straightforward statistical way to estimate the accuracy of the standard error itself. In this situation, we can use bootstrapping to evaluate how accurate the estimate of the population variance is.

Bootstrapping estimates properties of an estimator by resampling, with replacement, a random sample from the original sample. A resample in a bootstrap contains the same number of samples as the original sample. For each resample, the relevant estimators

(for example sample mean or standard error) are calculated. The set of resampled values then gives us an empirical distribution from which we can determine confidence intervals for the estimator (or indeed, calculate the standard error of the standard error, if the distribution of the estimator is Gaussian).

Bootstrapping and resampling methodology is used in two contexts in this study. The first is in Chapter 3, where bootstrapping is used to estimate the variability in the real-time assays for measuring mtDNA heteroplasmy and total mtDNA copy number. Secondly, it is used several times in Chapter 4, in particular to evaluate which measurement of heteroplasmy (blood or urine) best predicts disease progression in the m.3243A>G patient cohort.

The bootstrapping was executed in matlab. The code used can be found in Appendix II.

2.11. General Methodology

2.11.1. Disease progression modelling

As several chapters of this thesis use similar statistical methodology, the generic aspects of the modelling will be detailed here and referred to in each chapter as appropriate.

Disease burden analysis utilises the NMDAS assessment scores, as described in section 1.5. The total NMDAS score is a summary of the disease burden for a patient at a given time (assessment date). Though the NMDAS score is actually a discrete scale, in effect we can consider it a quasi-continuous scale and thus it can be modelled using techniques such as linear regression. The total score of all NMDAS questions was used, excepting the score for respiratory function; this was excluded as the scoring system for this aspect of the NMDAS is currently under review and the existing scores are not considered to reflect the respiratory function of the patients.

For each analysis, Box-Cox analysis is used to stabilise the variance of the dependent variables and to identify optimal transformations of the variables to satisfy assumptions of normality,(Box and Cox, 1964) thus enabling the use of parametric testing. Where data is already satisfactorily Gaussian in distribution Box-Cox should identify the optimal transform as identity.

For all analyses conducted using the total NMDAS score, Box-Cox analysis has shown that the optimal transformation to achieve normality is to take the fourth root of the

NMDAS score ($NMDAS^{0.25}$) as the dependent variable. This transformation is used in all cases, unless otherwise stated.

2.11.1.1. Total Disease Burden Basic Statistical Analyses

For the majority of patients we have several NMDAS assessments recorded. Basic analyses assume independence of all data, and thus multiple data for each patient cannot be included in basic modelling. Hence, for basic analyses I use a single summary data point for each patient, determined by taking the mean NMDAS score and mean age at assessment. Basic analyses are conducted using SAS PROC GLM for simple linear regression and multiple linear regression.

2.11.1.2. Total Disease Burden Longitudinal Modelling

Though basic analyses are useful for a summary understanding of data, longitudinal modelling is required for a deeper understanding of disease progression over time.

PROC MIXED was used for longitudinal mixed modelling using a spatial power structure to model covariance of repeated data from the same patient.(Singer, 1998; Moser, 2004). Polynomial terms of time up to cubed ($time^3$) were included. The Akaike Information Criterion (AIC) was used to compare models (Akaike, 1974). Model validation included the use of residual, influence, and leverage diagnostics to check model assumptions and verify model stability. SAS version 9.2 (Cary, NC) was used throughout.

2.11.2. Statistical reporting conventions

In all statistical analyses reported in this thesis, the same general conventions are adhered to, unless otherwise stated.

For pre-determined hypothesis testing, significance was determined at $P < 0.05$, high significance at $P < 0.001$.

For multiple regression, I report the standardized coefficient (B) (standardized to have unit variance) and significance value (P value) for each parameter estimate, together with the number of subjects (N) and the adjusted coefficient of determination (R^2) for the overall regression.

For simple linear regression I report N, the Pearson's correlation coefficient (r) and the P value.

For multiple logistic regression I report N, and the standardized coefficient and P value for each parameter estimate.

Chapter 3

Chapter 3. **Methods development**

3.1. Large scale single-mtDNA deletion level measurement assay

3.1.1. Introduction

The *MT-ND4*/*MT-ND1* assay was developed for measuring the deletion level, or heteroplasmy level, of single large-scale mtDNA deletions (He *et al.*, 2002). For diagnostic or prognostic purposes an approximate estimation of the deletion level in skeletal muscle homogenate is often sufficient. However, for certain studies a more accurate measure of heteroplasmy is required. Resistance training, for example, has been shown to improve muscle oxidative capacity in patients with single large-scale mtDNA deletions (Murphy *et al.*, 2008). This is thought to be the result of activation, proliferation, and incorporation of satellite cells; though satellite cells have similar deletion levels to the mature muscle cells, the mutations have been shown to be lost during the transition from satellite cell to myoblast (Spendiff *et al.*, 2013). It is hypothesised that analysis of deletion levels in patients undergoing resistance training will reveal lowering of heteroplasmy levels in resistance trained muscles; however, sensitive measurement of heteroplasmy in muscle homogenate is a requirement for this analysis.

It has long been known that the *MT-ND4*/*MT-ND1* assay is more accurate at high heteroplasmy levels (over 70%) than at lower levels, with an empirically estimated cut-off of around 30% as a baseline for accurate quantification (Prof Doug Turnbull, personal communication). It has also recently been reported that the measurement error in real-time PCR (qPCR) is dependent on DNA concentration (Sochivko *et al.*, 2013), with empirical error increasing as DNA concentration decreases.

To date the variation in the *MT-ND4*/*MT-ND1* assay has not been quantified. Accurate quantification of this variation is necessary to plan experiments that require subtly changes in deletion level to be measured.

3.1.2. Aim

The aim of this study is to quantify the variation in the assay with respect to two variables, the deletion level (heteroplasmy) of the sample, and the concentration of the DNA.

3.1.3. Methods

3.1.3.1.1. Methodological approach

The methodological approach to the analysis was to repeat each sample the maximum number of times possible on a qPCR plate in order to be able to estimate with highest accuracy the variability of the assay. Bootstrapping was further used to quantify the precision of this variability estimate.

The accuracy of the assay was then verified with samples of similar deletion level to confirm that small changes in deletion level are measurable.

3.1.3.2. Real-time PCR methodology

All molecular analyses were performed using total skeletal muscle DNA extracted using standard protocols. mtDNA deletion level in muscle homogenate was quantified using a validated, multiplex qPCR *MT-ND4*/*MT-ND1* assay (He *et al.*, 2002; Krishnan *et al.*, 2007).

Wild-type skeletal muscle DNA from a control subject was used to produce the standard curve for each real-time analysis.

All mtDNA deletion level quantification was performed by Dr Julie Murphy.

3.1.3.3. Intra-plate variability measurement

3.1.3.3.1. Study design

Six patients with a range of previously quantified deletion levels were selected, with deletion levels ranging from around 6% up to 80%. All DNA samples were diluted to approximately the same concentration on a trial run of qPCR to achieve a target of around 20 CTs. These samples were then serially diluted to a range of concentrations to achieve final CT values of around 27, 31, and 35 CTs.

Each plate was made up of 84 replicates; the remaining wells on each 96 well plate were used up by no-template controls and serially diluted wild-type total DNA for standard curve estimation.

3.1.3.3.2. Statistical methodology

Severe outliers, defined as more than 3 times the interquartile range below the lower quartile or above the upper quartile, were removed from the data prior to further analysis.

The standard deviation of the replicates on each plate was calculated. Bootstrapping was then used to calculate 95% confidence intervals for the observed standard deviation.

The standard deviation at high DNA concentration (< 23 CTs) was used to calculate minimum detectable differences observable using a two tailed t-test for a range of N values at β=0.80 and α=0.05.

3.1.3.4. Empirical verification of deletion level

To confirm the reliability of the assay to measure deletion level changes, 3 different samples at a low deletion level (20%-25%) were run on three separate plates. Each sample was run with 24 replicates on each plate. This was used to calculate independent values for the differences between samples per. The difference between samples 1 and 2 and the difference between samples 2 and 3 are reported for each plate (samples 1 and 3 are not compared as this is not an independent comparison, it can be derived from the other two differences).

The same experiment was performed with three samples at high deletion levels (70%-80%).

For each pair of samples, the true difference was estimated by taking the mean of all replicates on each plate. The deviation for each replicate from this true value was then calculated.

3.1.4. Results

3.1.4.1. *MT-ND4*/*MT-ND1 a*ssay variability is dependent on DNA concentration and heteroplasmy level

The quantified variation in the assay is depicted in Figure 3.1. As shown in Figure 3.1A, for each DNA sample, dilution of the DNA to lower concentration increases the variance of the measurement, and heteroplasmy also decreases with increasing heteroplasmy. The relationship between heteroplasmy and the assay variation is linear at high DNA concentration (below 23 CTs), as seen in Figure 3.1B ($r = -0.996$, $P < 0.0001$). The parameters for the regression equation are found in Table 3.1.

Whilst Figure 3.1 depicts the data for a selection of the data produced for this study, Figure 3.2 depicts all of the data produced, including data at very low DNA concentration (> 31 CTs). This illustrates the extreme increase in variation as DNA concentration decreases.

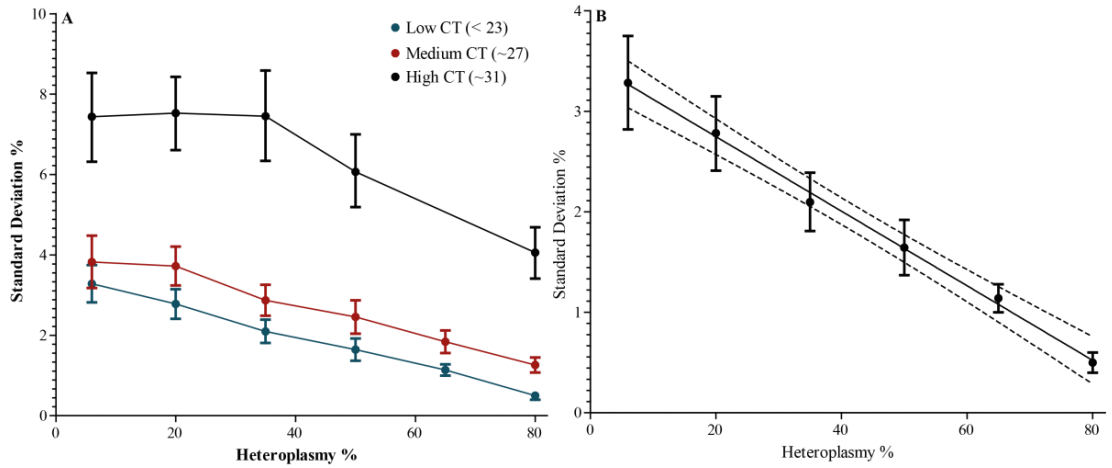


Figure 3.1 Variation in the *MT-ND4*/*MT-ND1* assay is dependent on heteroplasmy and DNA concentration.

The y axis depicts the standard deviation in absolute percentage points; for example, a 3% standard deviation (95% confidence interval approx. ±6%) for a measured 20% heteroplasmy would be 14%-26%. **(A)** Standard deviation of the *MT-ND4*/*MT-ND1* assay as measured for three concentrations of source DNA; the average CT of the assay is indicative of the DNA concentration, with the most concentrated DNA at lowest average CT. Variance increases with decreasing DNA concentration (increasing CT). At each DNA concentration, assay precision increases as heteroplasmy increases. **(B)** Relationship between assay variation and heteroplasmy at low (< 23) CT. Assay precision is linearly related to the heteroplasmy level. r = -0.996, P < 0.0001.

Parameter	Estimate	Std Dev	95% Lower	95% Upper
Slope	-0.03711	0.002468	-0.04234	-0.03188
Intercept	3.491	0.1225	3.231	3.751

Table 3.1 Regression coefficients for relationship between assay standard deviation and heteroplasmy.

The estimated value, standard deviation, and 95% confidence intervals are shown for each parameter.

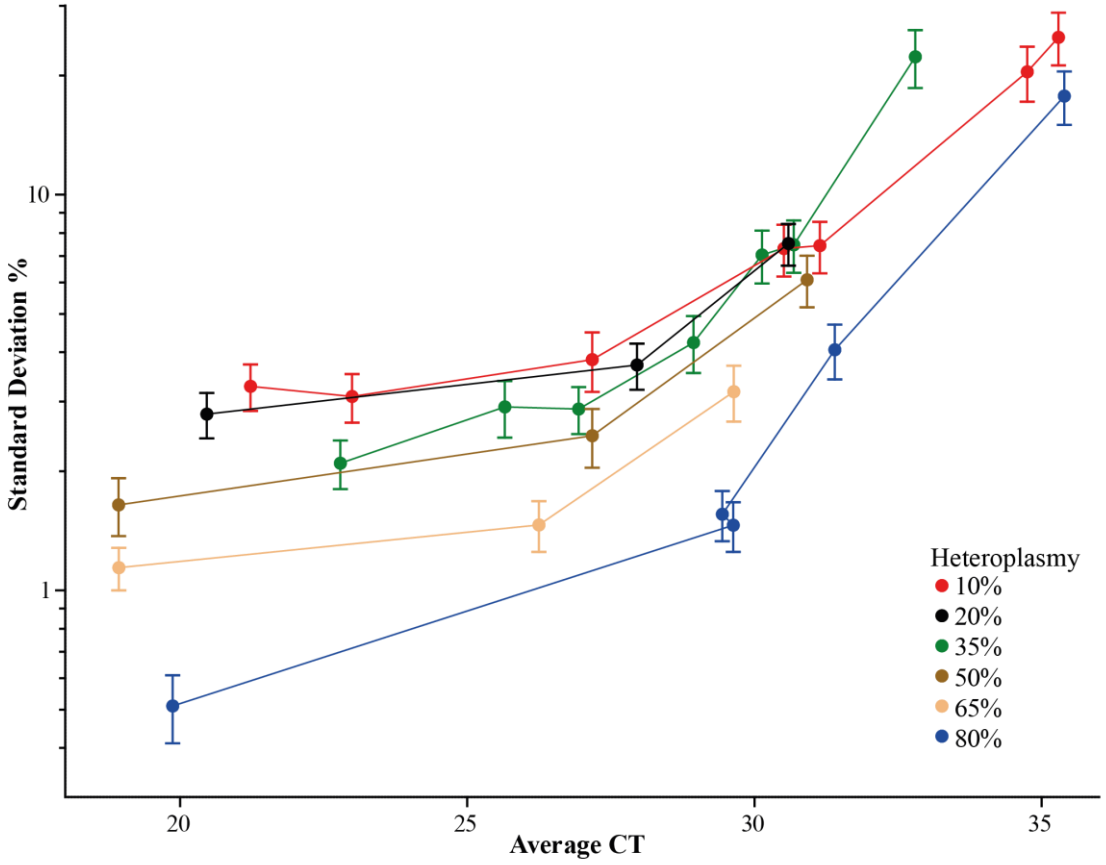


Figure 3.2 Relationship between DNA concentration and assay variation at various levels of heteroplasmy.

The y axis is log scaled. At any heteroplasmy level, assay variation increases with increasing CT (decreasing DNA concentration). Between 18 and around 28 CTs variation increases slowly at any heteroplasmy level. At very low DNA concentrations (above approximately 31 CTs) assay variation is extremely high and increases rapidly with average CT.

3.1.4.2. Verification of intra-plate and quantification of inter-plate variability

Results for the difference in samples are shown in Figure 3.3. The empirical standard deviation for the difference is 3.14% at 20%-25% deletion level and 1.28% at 70%-80% deletion level. This can be verified by scrutiny of the figure; the majority of values for the deviation at low deletion level are within ±6%, which is an approximate 95% confidence interval based on a standard deviation of 2.5%.

The expected standard deviation for the difference between two deletion levels can also be calculated. Using the linear regression coefficients in Table 3.1, at 25% deletion level the standard deviation for a single heteroplasmy measure is calculated as 2.56%. Thus the standard deviation of the difference of two deletion levels (using the fact that variance of the difference of two normally distributed variables is the sum of the variances) is calculated as $\sqrt{2 \times 2.56^2}$ =3.62%.

At 70% deletion level the standard deviation is calculated as 0.89%. The difference of two means at this level is, therefore, calculated as $\sqrt{2 \times 0.89^2}$ =1.26%.

The estimate for the inter-plate variation is 0.496% at low deletion level and 0.479% at high deletion level, with a pooled estimate of 0.485% using all samples.

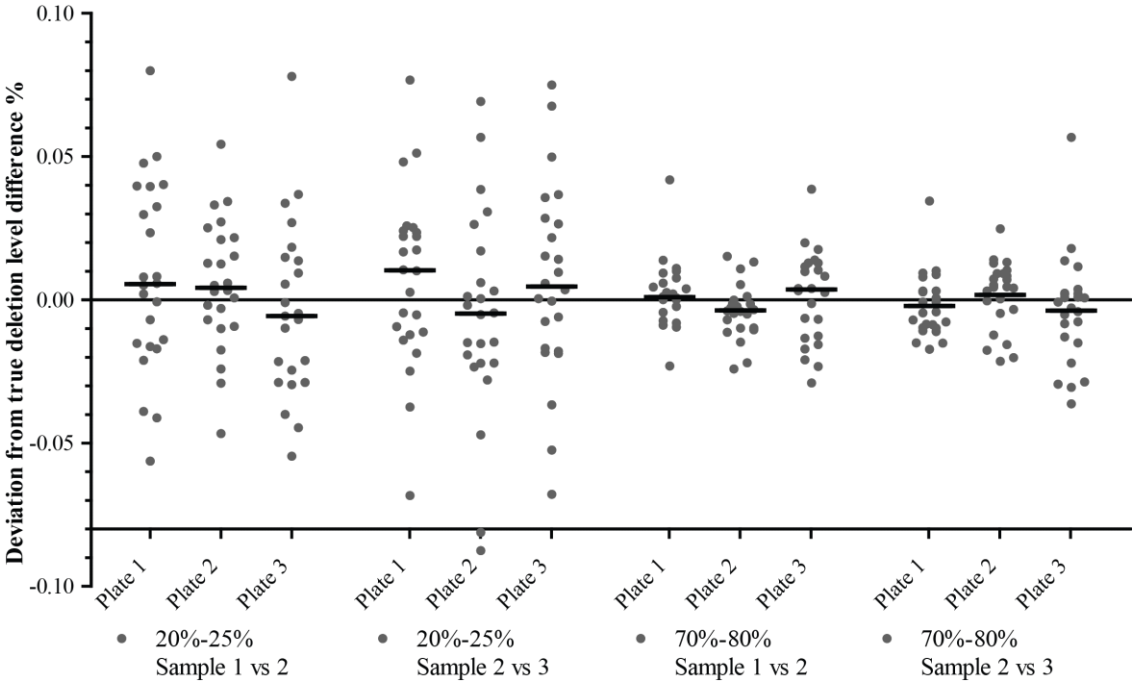


Figure 3.3 Inter- and intra-plate variability in the measurement of the difference in deletion level between samples.

To confirm the reliability of the assay to measure deletion level changes, 3 samples at a low deletion level (20%-25%) were run on three separate plates, 24 replicates of each sample per plate. This was used to calculate independent values for the difference between samples. The same experiment was performed with three samples at high deletion levels (70%-80%). The graph illustrates the variability in this difference; the data is centred by subtracting the best estimate of the actual deletion level difference. The mean value for each plate was within 1% of the actual deletion level difference. Variability at high deletion level was lower than at low deletion level.

Del Level	Standard Deviation	3	5	10	20
10%	3.12%	9.6%	6.3%	4.1%	2.8%
20%	2.75%	8.4%	5.6%	3.6%	2.5%
30%	2.38%	7.3%	4.8%	3.2%	2.2%
60%	1.26%	3.9%	2.6%	1.7%	1.1%
80%	0.52%	1.6%	1.1%	0.7%	0.5%

Table 3.2 Relationship between standard deviation, heteroplasmy levels and sample replicates.
For the given number of replicates, this table indicates the minimum detectable heteroplasmy difference.

3.1.5. Discussion

Severe outliers were removed from the data prior to further analysis. Though this may seem questionable considering the assay variability is under investigation, this was considered an appropriate pre-filter, in order to eliminate wells which were clearly erroneous. A basic interpretation of an outlier taught in many high school textbooks is based on points more than 1.5 times the interquartile range below or above the data quartiles; using a cut-off of 3 times the interquartile range makes the test much more liberal but avoids severely erroneous points being included in the analysis.

The results in section 3.1.3.4 show that the calculated standard deviation for the difference between two samples at high deletion levels (1.26%) is very close to the empirical level (1.28%). However, at low deletion level the empirical value (3.14%) is somewhat lower than the calculated value of 3.62%. The reason for this unexpected precision is unknown, but the value is not sufficiently different from the expected value to cause concern.

The estimate for the inter-plate standard deviation is low, under 0.5%. This implies that the advantage to be gained from running repeat plates for a given result is minimal.

The results from this analysis are important in two respects. Firstly, it has quantified for the first time the variability in this assay, and decisions can be made on the number of replicates required to precisely estimate deletion level. Many researchers use three replicates in any qPCR estimation as a matter of course; it is clear from this analysis that this approach is not valid at either low deletion level or low DNA concentration.

There are limitations to be acknowledged. Though we studied the inter-plate variation, this aspect of the study was not comprehensive as only three plates were run for each

investigation (six plates in total). It would be desirable to measure this more precisely to be confident that a single real-time PCR plate is sufficient to accurately quantify the heteroplasmy to a given required precision. Additionally, the linear regression provided guidelines for quantification of the variation only for concentrated DNA. The same was not done for lower DNA concentrations; however, rough estimates can be obtained by interpreting the data from the graphs.

3.1.5.1. Theoretical basis for the variation in the assay with heteroplasmy level.

The increasing variability of the assay with decreasing heteroplasmy can be explained by the mathematical process used to calculate heteroplasmy.

The real-time assay measures the relative difference in the expression of the two genes *MT-ND1* and *MT-ND6*. At low heteroplasmy levels, the difference between these two quantities is small. For instance, at 10% heteroplasmy and a perfectly efficient real-time reaction (doubling each time) the CT difference would be 0.152 ($\log_2(0.9)$, or the log of the heteroplasmy subtracted from 100%). The CT difference for a 90% heteroplasmy sample in the same circumstance would be 3.32, or $\log_2(0.1)$. If the errors are constant in the CT difference, for example ±0.1 CT, then the CT difference for the 10% sample will be measured within (0.052, 0.252), whilst the 90% heteroplasmy sample will be measured within (3.22, 3.42). These two ranges then relate to heteroplasmy measures of (4%, 19%) and (89.3%, 90.1%) respectively. Thus, the difference in accuracy arises most likely because the errors are independent of the CT differences, and thus will be proportionately larger for a small difference than a large difference.

3.2. Total mtDNA copy number measurement assay

3.2.1. Introduction

Low mtDNA copy number has been repeatedly shown to be associated with disease (Moraes *et al.*, 1991; Poulton *et al.*, 1995; Barthélémy *et al.*, 2001; DiMauro and Hirano, 2005). It has also been shown that exercise improves physical capacity of patients mtDNA disease, and that this is associated with an increased mtDNA copy number (Taivassalo *et al.*, 1998; Taivassalo *et al.*, 2006). Understanding the mitochondrial changes in muscle that lead to improved physical wellbeing is imperative, and mtDNA copy number is a potentially significant aspect of this. Copy number determination has not been a significant issue within exercise trial studies when single muscle fibres are the focus, as mitochondrial DNA levels are normalised per unit area. However, to understand overall changes in muscle it is important to look at homogenate muscle samples.

Gel electrophoresis and Southern blotting of nuclear DNA and mtDNA simultaneously has been used to measure the relative quantity of mtDNA to nuclear DNA, but this is labour intensive and only semi-quantitative (Shanske and Wong, 2004). Most published approaches to measuring mtDNA copy number have used real-time PCR (qPCR), using regions of the nuclear and mtDNA genomes where polymorphisms are rare (Venegas and Halberg, 2012). The *B2M* gene has been used in the literature as it is a convenient single copy gene in this respect (Malik *et al.*, 2011), and we already use the *MT-ND1* gene to measure heteroplasmy for mtDNA deletions as deletions covering this region are very rare, hence we chose to use this as the mtDNA target for quantification.

Multiplexing mtDNA and nDNA targets together is problematic, since several thousand copies of mtDNA are present in muscle cells but nuclear DNA is single copy. This presents problems with PCR kinetics and saturation effects in the multiplexed PCR reactions. Multi-copy genes can be used that ameliorate this problem to a large extent, such as 18S ribosomal DNA which is present around 600 times in the nuclear genome (Schmickel, 1973; Stults *et al.*, 2008); this is suitable for normalization for repeated samples from the same individual, but unsuitable otherwise due to inter-subject variation in the repeat numbers. Additionally, in a recent study it was found that the accuracy of quantitative real time PCR was severely affected by the concentration of DNA within the sample (Sochivko *et al.*, 2013) further highlighting the issue of comparing high copy number targets to lower copy number targets.

To address these issues, it was decided to develop a non-multiplexed assay with an extra dilution of the mtDNA target to ensure that both *B2M* and *MT-ND1* are within a reasonable range for quantification by real-time PCR (above approximately 17 CTs but below 30 CTs, to ensure quantification is in the exponential phase of growth).

3.2.2. Aims

As it was expected that the extra dilution step and the lack of multiplexing would introduce experimental error into the quantification process, it was decided to quantify the experimental variability in order to make recommendations on the number of replicates required to ascertain the relative copy number of two samples within a pre-determined tolerance. As the experimental error is multi-level (intra-plate variability nested within inter-plate variability) simulation will be used to determine the number of replicates required to determine a given change in copy number.

3.2.3. Methods

3.2.3.1. Real-time PCR methodology

All molecular analyses were performed using total skeletal muscle DNA extracted using standard protocols. mtDNA copy number in muscle homogenate was quantified by measuring the relative expression of the mtDNA *MT-ND1* gene and the nDNA *B2M* gene.

MT-ND1 and *B2M* reactions were run separately, sequentially on the same real time machine. Primers and probes used are detailed in Table 3.3, the mastermixes for each reaction are found in Table 3.4 and Table 3.5, and standard cycling conditions are described in Table 3.6. The same wild-type skeletal muscle DNA from a control subject was used to produce the standard curve for each real-time analysis. Sample were run in the same well on the paired *MT-ND1*and *B2M* plates to minimize well-to-well error. The target DNA concentration for *B2M* quantification was 100 fold higher than that for *MT-ND1* quantification, to ensure that the resulting CTs for both samples would be within useful range (18-25). Thus the samples for *MT-ND1* quantification were diluted 1 in 100 relative to the samples for *B2M* quantification, using two serial 1:10 dilutions to reduce dilution error.

All real-time PCR quantification was performed by Dr Helen Tuppen.

Primers:	**Sequence**
***MT-ND1* (NC_012920.1)**	
L3485-3504	5'- CCCTAAAACCCGCCACATCT -3'
H3553-3532	5'- GAGCGATGGTGAGAGCTAAGGT -3'
***B2M* (NG_012920.1)**	
B2F (8969-8990)	5'- CCAGCAGAGAATGGAAAGTCAA -3'
B2R (9064-9037)	5'- TCTCTCTCCATTCTTCAGTAAGTCAACT -3'
Probes:	
***MT-ND1* (L3506-3529)**	VIC-5'- CCATCACCCTCTACATCACCGCCC -3'-MGB
***B2M* (9006-9032)**	6-FAM-5'- CCGACATCATTACCGGGTTTTCCTCTTG -3'-MGB

Table 3.3 Primers and probes.

Taqman universal PCR mastermix	10µl
B2M F (10µM)	0.6µl (final concentration: 300nM)
B2M R (10µM)	0.6µl
B2M probe (5µM)	0.4µl (final concentration: 100nM)
$MgCl_2$ (50mM)	1.2µl (final concentration: 3mM)
DNA (optimal concentration ~ 10ng/µl)	5µl
dH_2O	2.2µl

Table 3.4 Mastermix for the *B2M* reaction.

Taqman universal PCR mastermix	10µl
MT-ND1 L3485 (10µM)	0.6µl (final concentration: 300nM)
MT-ND1 H3553 (10µM)	0.6µl
MT-ND1 probe (5µM)	0.4µl (final concentration: 100nM)
DNA (optimal concentration ~ 0.1ng/µl)	5µl
dH_2O	3.4µl

Table 3.5 Mastermix for the *MT-ND1* reaction.

2 minutes at 50°C
10 minutes at 95°C
40 cycles of 15 seconds at 95°C and 1 minute at 60°C

Table 3.6 Standard cycling conditions.

3.2.3.2. Intra-plate variability measurement

3.2.3.2.1. Study design

Test runs of several samples were executed to find two samples with high relative copy number difference. These two samples were diluted to approximately equal DNA concentration, and then mixed together in three different ratios. This resulted in 5 samples of different relative copy number, the original two samples plus three interpolated DNA copy numbers.

On each real-time PCR plate 12 replicates of each sample were run, totalling 60 wells. The remaining wells were used for standard curve generation and no template controls. Six replicates of five standard curve points (30 in total) were run, to ensure highly accurate standard curves for optimal quantification.

3.2.3.2.2. Statistical methodology

Severe outliers, defined as more than 3 times the interquartile range below the lower quartile or above the upper quartile, were removed from the data prior to further analysis.

The *MT-ND1/B2M* ratio was calculated for each sample well in each plate, yielding 12 replicate values per *MT-ND1/B2M* plate pair. These were used to calculate 12 independent values for the relative copy number of each of samples 2 to 5 compared to sample 1, which resulted in 12 replicates of 4 independent sample comparisons per *MT-ND1/B2M* plate pair. Four *MT-ND1/B2M* plate pairs were run, to examine inter-plate variability.

The standard deviation of the replicates on each plate was calculated. Bootstrapping was then used to calculate 95% confidence intervals for the observed standard deviation.

A linear model was used to verify whether the variance of the assay was independent of relative copy number.

Using the calculated inter-plate and intra-plate variability, sample size calculations were performed in SAS using PROC POWER, using the formula for the sample standard deviation as shown in Equation 3.1, as the intra-plate (replicate) variability is nested within the inter-plate (plate-to-plate) variability.

$$sd_{exp} = \sqrt{sd_{p2p}^2 + \frac{sd_{rep}^2}{n_{rep}}}$$

Equation 3.1 Calculation of experimental standard deviation.

The experimental standard deviation sd_{exp} is calculated from the inter-plate (p2p) standard deviation and the intra-plate (rep) standard deviation using the formula for the summation of variances. The sample variance for the replicates is divided by the number of replicates n_{rep}, which is then used to calculate the number of plates required.

3.2.4. Results

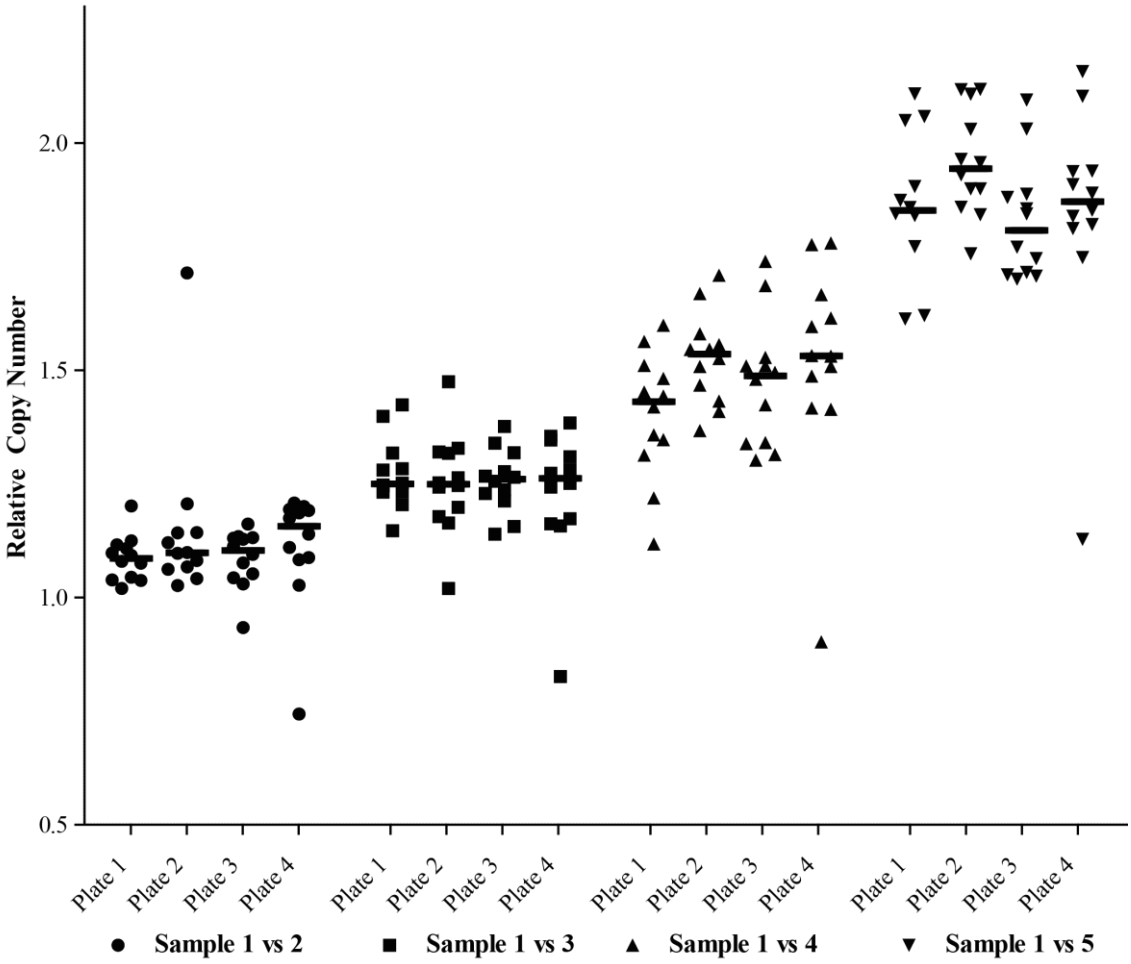


Figure 3.4 Intra- and inter-plate variability of the copy number assay.

5 samples of progressively decreasing copy number are analysed on each plate; each sample is replicated 12 times on each plate, and 4 plates are run in total. This was used to calculate 12 independent samples of the relative copy number of sample 1 compared to each of the other 4 samples. The mean relative copy numbers for each sample comparison are comparable from plate to plate.

The relative copy number of sample 1 compared to samples 2 to 5 are shown in Figure 3.4. With 12 replicates of each measure, these are 12 x 4 x 4 = 192 independent samples of copy number.

The intra-plate (replicate) and inter-plate (plate to plate) variability for the relative copy number of each pair of samples is shown in Figure 3.5. These are not all independent samples (the first 4 are independent), but are shown to allow comparison of variability across the samples. 95% confidence intervals are calculated using bootstrapping methodology.

To calculate a global estimate for the standard deviations a modified bootstrapping was employed, using the pool of all sample comparisons (10 comparisons, 12 replicates of each comparison on 4 plates, totalling 480 relative copy number measures in total). Each copy number estimate was scaled first by dividing by the best estimate of the relative copy number of the two samples. This yielded an intra-plate standard deviation of 1.066 (95% CI 1.038-1.101) and an inter-plate standard deviation of 1.027 (95% CI 1.0047-1.069).

The inter-plate and intra-plate standard deviations were also compared to the relative copy number. The results are shown in Figure 3.6. There is no evidence of a relationship between the relative copy number and the variability.

The number of plates/replicates per plate required to detect a given change in copy relative number based on the estimated inter- and intra- plate variation is shown in Table 3.7.

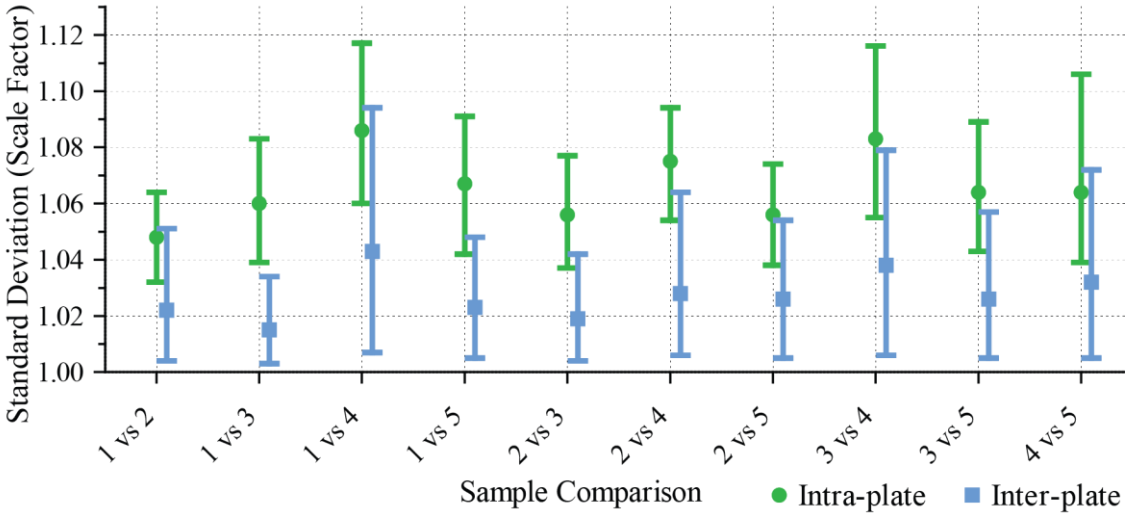


Figure 3.5 Inter-plate and intra-plate variability across samples.

The relative copy number (ratio) of each pair of samples was calculated, 12 replicates on 4 plates, a total of 48 values per ratio. All 48 values were used to calculate the true value, and each sample divided by this to normalize the ratios. The intra-plate (replicate) and inter-plate (plate to plate) standard deviation and 95% confidence intervals were then calculated for each sample comparison using bootstrapping. All variability was examined on the log values; the graph shows inverse log standard deviations, which are scale factors. The intra-plate variability is consistently higher than the inter-plate variability. The mean inter-plate variability is 1.027 and the mean intra-plate variability 1.066.

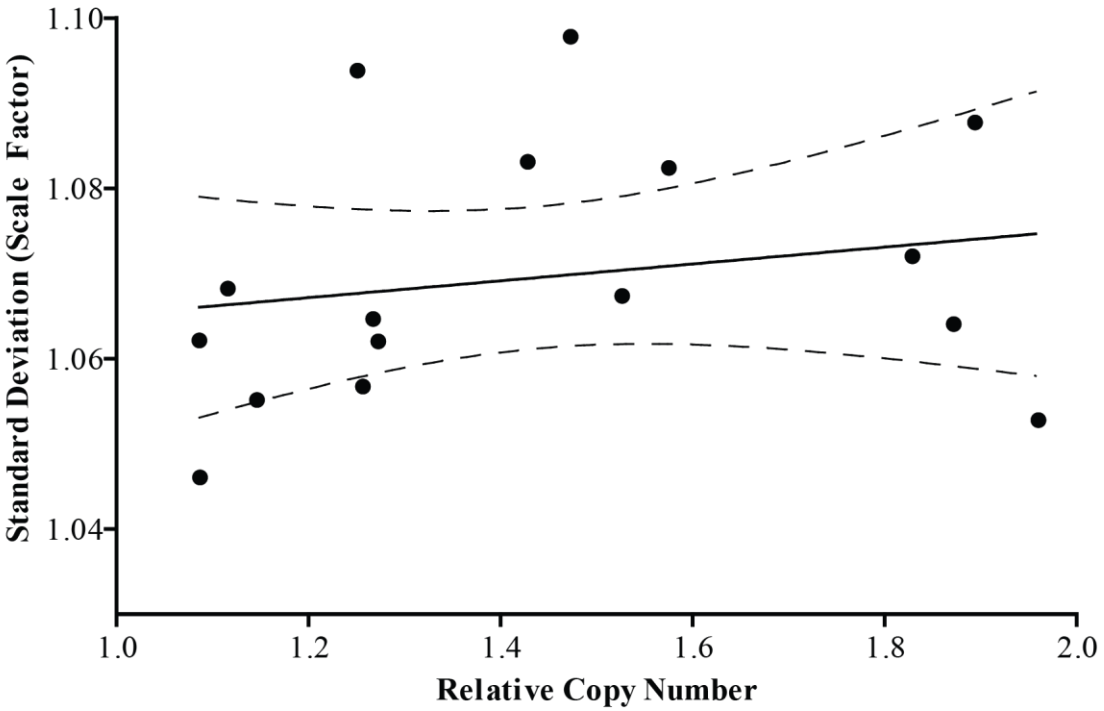


Figure 3.6 Linear regression of relative copy number against intra-plate assay variation.

The graph shows no evidence of a linear correlation between relative copy number and the assay intra-plate variability (P = 0.4614, r = 0.20). Each point is an independent estimate of the standard deviation from a single plate.

Relative Copy Number Change	5%	10%	15%	20%	50%
Log_{10} factor	0.021189	0.041393	0.060698	0.079181	0.176091
1 replicate	26 / 0.81 72 / 0.80	8 / 0.83 20 / 0.81	4 / 0.81 10 / 0.83	3 / 0.84 6 / 0.81	2 / 0.97 3 / 0.96
2 replicates	16 / 0.82 48 / 0.80	5 / 0.83 14 / 0.82	3 / 0.85 7 / 0.83	3 / 0.96 5 / 0.87	2 / 0.99 2 / 0.84
3 replicates	12 / 0.82 40 / 0.80	4 / 0.83 12 / 0.83	3 / 0.92 6 / 0.83	2 / 0.80 4 / 0.83	2 / >0.999 2 / 0.89
6 replicates	8 / 0.80 32 / 0.8	3 / 0.82 9 / 0/00	3 / 0.98 5 / 0.83	2 / 0.91 4 / 0.90	2 / >0.999 2 / 0.93
12 replicates	7 / 0.84 28 / 0.81	3 / 0.900 8 / 0.81	2 / 0.85 5 / 0.87	2 / 0.96 3 / 0.81	2 / >0.999 2 / 0.96

Table 3.7 Number of real-time PCR plates/replicates required to detect a given change in relative copy number.

Each cell shows the number of plates / actual power for a two sample t-test with the given number of replicates (per sample) per plate, to detect the given change in relative copy number at minimum power of 0.80. In each cell, the top line shows the calculation using the standard deviation estimates; the bottom line shows the calculation using the 95% upper bound of the standard deviation estimates.

3.2.5. Discussion

This investigation has quantified the variability in this novel approach to measurement of mtDNA copy number, and provides guidelines on the number of replicates and plates of each sample that should be run to achieve a required accuracy.

The variation in the assay is confirmed as independent of the relative copy number, within the range of relative copy number that was investigated (approximately 2 fold difference). This is a practically useful finding, as it implies that the number of replicates to determine copy number can be made independently of the copy number of the samples under investigation.

It was anticipated that the extra dilutions of the *MT-ND1* samples, necessary to ensure both targets were within conveniently measurable CT range, may have introduced unacceptable experimental error into the procedure. As each plate was run from re-diluted samples, the inter-plate variability incorporates any variation introduced by this extra dilution. The inter-plate variability (1.027 scale factor) is lower than the intra-plate variability (1.066), however it is still variability that needs to be accounted for when using the assay to accurately determine copy number.

Regarding recommendations for the number of replicates/plates to use to detect specific minimum changes in copy number, use of the t-test for requires a minimum of two plates in order that the sample variance is estimable. This restriction comes from the fact that the intra-plate variation is nested within the inter-plate variation. For large enough changes of copy number it seems intuitive that a single plate would be sufficient to estimate this change, but formal statistical analysis requires a minimum of two plates.

There are important limitations to be acknowledged.

Firstly, the number of plates run was small (four), and more plates would allow more accurate quantification of the inter-plate variation, a narrower 95% confidence interval for the estimates, and one would hope a lower estimate for the number of replicates required. The estimates using the 95% upper bound for the standard deviation estimates are impractically high for the smaller changes in copy number, and some reduction in the confidence interval would be very beneficial in order to have confidence in the results of the assay with the minimum experimental effort. The assay is limited by the number of samples that can be run on the same PCR plate, which is limited to 96 wells, several of which are used to calculate standard curves and for no-template controls.

Secondly, the inter-plate variability was confounded with the dilution error in this experiment; it may have been valuable to run separate repeated dilutions of samples on the same plate in order to separate the inter-plate and dilution variation in this assay, which may have identified the true inter-plate error to be small enough to ignore for practical purposes. However, since the two targets are intended to be run on separate plates (to avoid multiple standard curves on the same plate) some level of inter-plate variability is inevitable, thus is seems a prudent approach to combine the dilution and intra-plate variation and deal with them both concurrently. This is not necessarily a trade-off, as reducing the number of replicates per plate allows more samples to be run on each plate.

Though this assay optimisation and variation quantification has been focussed on relative copy number (intended for measuring change in copy number in repeated samples from the same patient), it can be extended to calculate absolute copy number. This would require an absolutely quantified reference sample to be run on each plate requiring absolute quantification.

3.2.5.1. Muscle specific issues

Normalisation of the mtDNA copy number by a single copy nuclear gene presents difficulties that are specific to muscle homogenate analysis, with two main issues to contend with.

Firstly, muscle fibres are multinucleated, each fibre formed from the fusion of multiple mononucleated myogenic cells (Okazaki and Holtzer, 1966). It is thought that each nucleus serves a certain volume of cytoplasm, called the myonuclear domain (Hall and Ralston, 1989). Historically, it was thought that the size of the domain was relatively constant, supported by general biological considerations that nuclear DNA content universally appears to be related to cell size (Gregory, 2001) and by specific examination of muscle fibres (Landing *et al.*, 1974). However, there are many findings from later studies that contradict this; domain size is related to body mass (Liu *et al.*, 2009), domain size isn't conserved during hypertrophic growth (Wada *et al.*, 2003), domain size varies along the length of muscle fibres and increases with age (Rosser *et al.*, 2002), and domain size decreases with muscle wasting and atrophy (Allen *et al.*, 1995; Ohira *et al.*, 1999; Bruusgaard and Gundersen, 2008). Furthermore, domain size has been shown to vary according to fibre type; Type I fibres (slow twitch, and heavily dependent on mitochondria for oxidative phosphorylation) have the smallest domain, Type IIa fibres (fast twitch, but also oxidative) are intermediate in size, and Type IIx (fast twitch glycolytic fibres with few mitochondria) have the largest myonuclear domain (Roy *et al.*, 1999; Van Der Meer *et al.*, 2011).

The second major issue is related to the cell types found in muscle homogenate. Non-muscle cells, such as adipocytes or cells of the vascular system are present and will contribute to both the nuclear and mtDNA assessment. The relative proportion of mitochondrial fibre types is also critical, particularly with reference to the proportion of Type IIx fibres which have very low mitochondrial content, though Type I fibres also have higher mitochondrial density than Type IIa fibres (Sjostrom *et al.*, 1982; Yu-Wai-Man *et al.*, 2010b) . Thus care must be taken to ensure that muscle samples for comparison are comparable. However, it is important to note that for analysis of samples to evaluate exercise trail intervention, the relative proportion of fibre types can be significantly altered by training; for example, endurance training has been shown to significantly increase the proportion of Type IIa fibres at the expense of Type IIx fibres (Ingjer, 1979). With this in mind, fibre type changes between biopsies may be genuine

changes induced by the intervention rather than sampling errors attributable to heterogeneity of muscle tissue.

Both of these issues must be considered when using normalisation of myDNA copy number by nuclear content with this assay. Prior to analysis, biopsies may need to be examined histochemically or otherwise to assess the relative proportions of fibre types and non-myogenic cells in order to be satisfied that the analysis is correctly determining any alterations in mitochondrial content of muscle cells.

Chapter 4

Chapter 4. Disease associated with the mt.3243A->G mutation

4.1. Introduction

The m.3243A>G mutation was one of the first identified causes of inherited mitochondrial disease (Goto *et al.*, 1990b). Of the over 250 known pathogenic mtDNA mutations (Tuppen *et al.*, 2010), m.3243A>G is the most commonly inherited, at an estimated prevalence in the adult population of 236/100,000 (Manwaring *et al.*, 2007), though the estimated minimum prevalence of clinically affected people is lower, at 3.65/100,000 (Schaefer *et al.*, 2008) or 16.5/100,000 (Majamaa *et al.*, 1998). In the cohort of 671 patients seen at the Newcastle mitochondrial disease clinic, 199 (29.7%) carry the m.3243A>G mutation.

The mutation is an A to G transition at position 3243 of the mitochondrial genome, in the dihydrouridine loop (D-loop) of the mitochondrial $tRNA^{Leu(UUR)}$ (Goto *et al.*, 1990b).

4.1.1. Phenotypic presentation

The earliest patients to be associated with the m.3243A>G mutation were almost exclusively MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes) (Goto *et al.*, 1990b), though a single patient with the mutation presented with CPEO, an early indication of the heterogeneous phenotypic variability shortly to be associated with the mutation (Hirano *et al.*, 1992). MELAS is a severe and progressive neurological disorder, and though it is not exclusively associated with the m.3243A>G mutation, around 80% of patients with MELAS syndrome are carriers of the m.3243A>G (Goto *et al.*, 1991). Clinical diagnostic criteria were published in 1992 and defined as (a) stroke-like episodes before the age of forty (b) encephalopathy (denoted by seizures and/or dementia), and (c) lactic acidosis and/or ragged-red fibres (RRF) (Hirano *et al.*, 1992).

Other common phenotypic presentations that were soon to be associated with the mutation were maternally inherited diabetes and deafness (MIDD) (van den Ouweland *et al.*, 1994), further CPEO cases (Goto *et al.*, 1990a; Moraes *et al.*, 1993; Koga *et al.*, 2000) and Leigh syndrome (Rahman *et al.*, 1996; Koga *et al.*, 2000). In common with other mtDNA mutations the clinical spectrum associated with the mutation is very broad. All of the constituent phenotypic features of both MELAS and MIDD are frequently seen in patients carrying the mutation, but there are a host of other common

features, both neurological and otherwise, including myopathy and exercise intolerance, cerebellar ataxia, visual impairment, migraine, cardiomyopathy, and gastro-intestinal disturbance (Ciafaloni *et al.*, 1992; Majamaa-Voltti *et al.*, 2006; Kaufmann *et al.*, 2011; de Laat *et al.*, 2012; Nesbitt *et al.*, 2013). Distinct phenotypic characterisation is problematic; a recent study has reported that although MIDD is the most common classical phenotype observed in patients, less than half of patients present with a classical recognised phenotype (Nesbitt *et al.*, 2013).

Stroke-like episodes are a hallmark feature of the MELAS syndrome. Also known as sudden neurological deficits, these often present with hemiparesis, hemianopsia and/or cortical blindness (Iizuka and Sakai, 2005). Magnetic resonance imaging (MRI) of the brain after a stroke-like episode usually shows changes in both grey and white matter, mainly in the occipital and parietal lobes, which strongly resemble infarctions; however, the distribution is unrelated to vasculature and often shows a progressive spread (Ito *et al.*, 2011). The pathophysiology is still controversial, with several competing theories as to the cause of the neurological dysfunction; proposals include mitochondrial dysfunction in the vasculature, known as the mitochondrial angiopathy theory; a generalised cytopathic mechanism caused by oxidative phosphorylation deficits in neurons and glial cells; and a non-ischemic neurovascular cellular mechanism, in which the stroke-like episodes are caused by neuronal hyper-excitability and an ATP deficit due to oxidative phosphorylation deficits(Iizuka and Sakai, 2005). Diffusion weighted imaging (DWI) has been used to examine brains of patients after stroke-like episodes and many early reports showed an elevated apparent diffusion coefficient (ADC), which contrasts with the reduction seen in ischaemic stroke, and which was consistent with vasogenic oedema and the mitochondrial angiopathy theory (Yoneda *et al.*, 1999; Oppenheim *et al.*, 2000; Yonemura *et al.*, 2001; Kolb *et al.*, 2003). However, more recently patients have been reported to demonstrate a reduced ADC more consistent with cytotoxic oedema and thus one of the other two other competing theories (Wang *et al.*, 2003; Karkare *et al.*, 2009; Tzoulis and Bindoff, 2009). Whatever the precise pathophysiology, the consequences for patients suffering stroke-like episodes are severe.

Diabetes mellitus is a second prominent characteristic feature of m.3243A>G; while it has been found to be associated with a small number of other mtDNA mutations at higher penetrance than in m.3243A>G, on the whole mitochondrial diabetes is dominated by the m.3243A>G mutation (Whittaker *et al.*, 2007). Diabetes can be

divided into two broad classes, type I insulin dependent and type II non-insulin dependent. Early studies observed a connection between type II diabetes and maternal transmission, which put mitochondrial DNA under the spotlight (Alcolado and Alcolado, 1991), and it had also been noted in maternal relatives of MELAS sufferers (Obermaier-Kusser *et al.*, 1991). Indeed, in the diabetic population as a whole m.3243A>G is thought to account for 1%-3% of all cases (Gerbitz *et al.*, 1995), and an even higher proportion of familial diabetes.

Myopathy is a commonly reported phenotypic presentation (Karppa *et al.*, 2005), in common with many mitochondrial disorders. Cardiomyopathy, long recognised in advanced cases of MELAS (Hirano and Pavlakis, 1994) is also recognised as prevalent (Bates *et al.*, 2012a; Bates *et al.*, 2012b), and the pathology has been shown to correlate with skeletal muscle heteroplasmy (Hollingsworth *et al.*, 2012).

Renal disease is also increasingly recognised as a phenotypic presentation of m.3243A>G (Damian *et al.*, 1995; Cheong *et al.*, 1999; Guillausseau *et al.*, 2001; Hotta *et al.*, 2001; Iwasaki *et al.*, 2001; Suzuki *et al.*, 2003; Piccoli *et al.*, 2012; Seidowsky *et al.*, 2013). Proximal tubular cells have a high mitochondrial content and are frequently reported as affected in mitochondrial disease patients with Fanconi syndrome like features (Rötig *et al.*, 1997; Emma *et al.*, 2011), though focal segmental glomerular sclerosis (FSGS) is reported as specifically related to m.3243A>G (Hotta *et al.*, 2001; Emma *et al.*, 2011).

Regarding mortality, cardiac and neurological problems are reported as the most common cause of early death in m.3243A>G, and reports of sudden death are common (Majamaa-Voltti *et al.*, 2002; Uusimaa *et al.*, 2007; Vydt *et al.*, 2007; Bates *et al.*, 2012a).

4.1.2. Genotype-phenotype linkage

Understanding the connection between phenotype and genotype is complicated by heteroplasmy, as for all mtDNA disease (Wallace, 1992). Section 4.4 will consider in detail the issue of heteroplasmy in the m.3243A>G mutation.

Early studies on the m.3243A>G mutation suggested that patients with a high mutation load in muscle present at a young age with a MELAS-like phenotype, whilst those with lower mutation load present later in life with CPEO, myopathy, and deafness (Chinnery *et al.*, 1997). Though this is superficially paradoxical, it was suggested that this may be due to focal accumulation of mutant DNA in myopathic patients as compared to more

uniform levels of the mutation in MELAS patients (Petruzzella *et al.*, 1994), though this study was on a small number of patients from only two families. Other studies support this to an extent, showing that patients with myopathic phenotypes have a high prevalence of COX negative fibres in muscle but relatively low levels of the mutation compared to MELAS patients (Moraes *et al.*, 1993; Hammans *et al.*, 1995). However, it must be considered that in general genotype-phenotype correlation has long been considered very weak in patients with the m.3243A>G mutation (Kobayashi *et al.*, 1992; Martinuzzi *et al.*, 1992; Shiraiwa *et al.*, 1993; Liou *et al.*, 1994; Morgan-Hughes *et al.*, 1995; Chinnery *et al.*, 1997).

4.1.3. Pathogenesis

Early *in-vitro* studies demonstrated that cells harbouring the m.3243A>G mutation (as well as those with the m.8344A>G mutation) impair oxidative phosphorylation at a cellular threshold of around 85% heteroplasmy (Chomyn *et al.*, 1991; Kobayashi *et al.*, 1991; King *et al.*, 1992a). Cells harbouring even higher levels of mutation, around 95%, have been shown to be more acutely impaired (Dunbar *et al.*, 1996).

Processing of the polycistronic transcript is not affected by the mutation (King *et al.*, 1992b; Koga *et al.*, 1993; Kaufmann *et al.*, 1996). However the proportion of $tRNA^{Leu(UUR)}$ that is aminoacylated is reduced (Janssen *et al.*, 1999; Chomyn *et al.*, 2000; Park *et al.*, 2003), most likely on account of a dramatic 25 fold reduction in the efficiency of the aminoacylation of the mutant tRNA as compared to wild-type (Park *et al.*, 2003). Compounding this is the reduced steady state levels of the tRNA species that has been repeatedly reported (Chomyn *et al.*, 1992; Janssen *et al.*, 1999; Chomyn *et al.*, 2000; Park *et al.*, 2003). Though the precise nature of the pathogenesis is still under debate, certain post translational modifications of the tRNA are impaired (Helm *et al.*, 1999; Yasukawa *et al.*, 2000b), which results in impaired protein translation (Yasukawa *et al.*, 2000a; Yasukawa *et al.*, 2001).

However, somewhat contradictory is the report that mitochondrial protein translation is not seriously affected by the mutation even at levels which severely impair cellular respiration, despite highly reduced aminoacylation of $tRNA^{Leu(UUR)}$ in these cells; accelerated protein degradation instead was suggested as the pathogenic mechanism (Janssen *et al.*, 1999).

Complex I deficiency has been frequently reported in tissue analysis from patients with m.3243A>G (Goto *et al.*, 1992; Morgan-Hughes *et al.*, 1995) and also in cybrid cell

analysis (Dunbar *et al.*, 1996). Of all the respiratory chain complexes, complex I contains the highest proportion of mitochondrial encoded leucine (UUR) residues, which may explain the selective vulnerability of this complex. It has recently been demonstrated in induced pluripotent cells (iPSCs) generated from m.3423A>G fibroblasts that the respiratory chain dysfunction caused by the m.3243A>G mutation is very much tissue dependent (Hämäläinen *et al.*, 2013). For instance, a combined respiratory chain deficiency in the parent fibroblast cells became a specific complex I deficiency in differentiated neurons, which is consistent with the pathology commonly reported in m.3243A>G (Moraes *et al.*, 1992). Indeed, this study clearly demonstrated that Complex I is specifically degraded by sequestration into PTEN-induced putative kinase 1 (PINK1) and Parkin-positive autophagosomes, suggesting that the observed Complex I defect is potentially a cellular auto-protective response to mitochondrial dysfunction, perhaps to reduce potentially harmful ROS production.

4.1.4. Sex differences

The prospect of a difference between males and female in the burden of mtDNA disease (Frank and Hurst, 1996; Frank, 2012) has prompted the study of sex differences in the m.3243A>G mutation, which is covered in section 4.6.

4.1.5. Therapeutic strategies

There are few therapeutic treatments for patients with m.3243A>G. Studies on the efficacy of antioxidants and vitamins have failed to demonstrate clear benefit to patients (Marriage *et al.*, 2003). Treatment with Coenzyme Q_{10} (ubiquinone) and idebenone (Ihara *et al.*, 1989), an ATP production modulator and antioxidant, is routine and supported by research (Haefeli *et al.*, 2011). However, these approaches are focused on amelioration of ROS induced cell damage and do not address the underlying problems with oxidative phosphorylation.

There are several promising avenues of research into future therapeutics. tRNA import into mitochondria has been demonstrated to ameliorate respiratory defects *in vitro (Karicheva et al., 2011).* Similarly, overexpression of mitochondrial leucyl-tRNA synthetase (LARS2) has been shown to restore wild-type levels of respiration in cells harbouring the m.3243A>G mutation (Park *et al.*, 2008). This study showed dramatically increased steady-state levels of tRNA leucine and mitochondrial translation products. Interestingly, protein synthesis levels did not exceed those of the mutant cells without the overexpression, supporting the hypothesis that protein stabilization is the critical effect. More recently it has been shown that the C-terminus of

the tRNA synthetase molecule is necessary and sufficient for this effect (Perli *et al.*), and that the effect extends to defects from mutations in non-cognate tRNA species (Hornig-Do *et al.*). However, translating successful *in-vitro* therapies to patients is a currently insurmountable obstacle.

4.2. Methods

4.2.1. Heteroplasmy quantitation

Heteroplasmy levels are used in many of the analyses in this chapter. Heteroplasmy was quantitated using pyrosequencing (White *et al.*, 2005), though older samples were quantitated using last-cycle hot PCR (Moraes *et al.*, 1992). All quantitation was performed by the NHS Highly Specialised Service for Rare Mitochondrial Disorders in Newcastle upon Tyne.

4.2.2. Statistical methodology

The statistical methodology employed in each section will be detailed prior to the presentation of the results in each section.

4.3. Disease Progression

4.3.1. Introduction

Though many aspects of disease associated with the m.3243A>G mutation are considered progressive by clinicians, the nature and speed of this progression, and its relation to predictive factors such as age and heteroplasmy, are poorly understood. Understanding of likely progression of the disease burden in patients is critical for clinicians to be able to support and care for patients.

Longitudinal studies are a necessity in understanding disease progression. However, systematic recording of disease progression in m.3243A>G mutation carriers is poor. Though there are two recent longitudinal studies of patients with the m.3243A>G mutation (Majamaa-Voltti *et al.*, 2006; Kaufmann *et al.*, 2011), the follow-up period for each study (three and four years respectively) was limited. Additionally, neither study considered predictive factors such as heteroplasmy and age. Several other recent studies have either examined in detail the clinical phenotype of patients (Nesbitt *et al.*, 2013) or examined the correlation between disease burden and heteroplasmy (Whittaker *et al.*, 2009; de Laat *et al.*, 2012; Liu *et al.*, 2013) but without looking at progression or accounting for age.

It has long been thought that clinical variability is at least in part due to heteroplasmy (Ciafaloni *et al.*, 1991; Damian *et al.*, 1995), though association between heteroplasmy and clinical presentation in patients harbouring the m.3243A>G mutation have proved to be weak (Kobayashi *et al.*, 1992; Martinuzzi *et al.*, 1992; Shiraiwa *et al.*, 1993; Liou *et al.*, 1994; Chinnery *et al.*, 1997). However, recent studies have found that urine heteroplasmy correlates better with the severity of clinical features than other sources such as hair follicles, buccal mucosa, blood, and even muscle (Ma *et al.*, 2009; Whittaker *et al.*, 2009; de Laat *et al.*, 2012). Additionally, urine heteroplasmy is reported as stable over time (Blackwood *et al.*, 2010), and is currently widely accepted as the most suitable non-invasive measure of heteroplasmy (Whittaker *et al.*, 2009; de Laat *et al.*, 2012).

4.3.2. Aims

Improved understanding of the disease progression is critical both for clinicians caring for patients and for patients themselves, particularly regarding clinical management and planning of health and social care. To this end, I aim to utilise the MRC Mitochondrial Disease Patient Cohort, which contains a large sub-cohort of patients with the

m.3243A>G mutation, to ascertain the progressiveness of disease associated with this mtDNA mutation, and to quantify the correlation between disease burden and predictive factors of age, urine heteroplasmy, and familial lineage.

I will use longitudinal mixed modelling to understand the progression of overall disease burden (as measured by total NMDAS score), and multiple logistic regression to examine the relationship between specific phenotypic features (measured by individual NMDAS questions) and the predictive factors of age and heteroplasmy.

As intra-familial clustering of symptoms has been reported (Hammans *et al.*, 1995) I will also look at the significance of familial lineage on disease progression.

All modelling in this section uses heteroplasmy measured in the urinary sediment.

4.3.3. Cohort summary

4.3.3.1. Age and heteroplasmy

For modelling disease progression, all patients with both NMDAS data and urine heteroplasmy have been included. Of the 152 total patients with NMDAS data and urine heteroplasmy, 95 are female. One male patient was excluded due to renal transplant, and one further male and four female patients were excluded due to insufficiently complete NMDAS assessments (only those with at least 26 of 29 questions completed were used). Thus the final cohort consisted of 91 females and 55 males, 146 patients in total.

The age and heteroplasmy profile of the cohort are shown in Figure 4.1, along with the number of assessments per patient. The median age of the cohort is 42 years, the youngest patient is 16 years old and the oldest 73 years old. The median heteroplasmy level is 57%, and the range extends from only just detectable (0.1%) to 99%. The median number of assessments is 3 per patient, with a maximum of 15. The clinical and molecular characteristics for the cohort can be found in Appendix III.

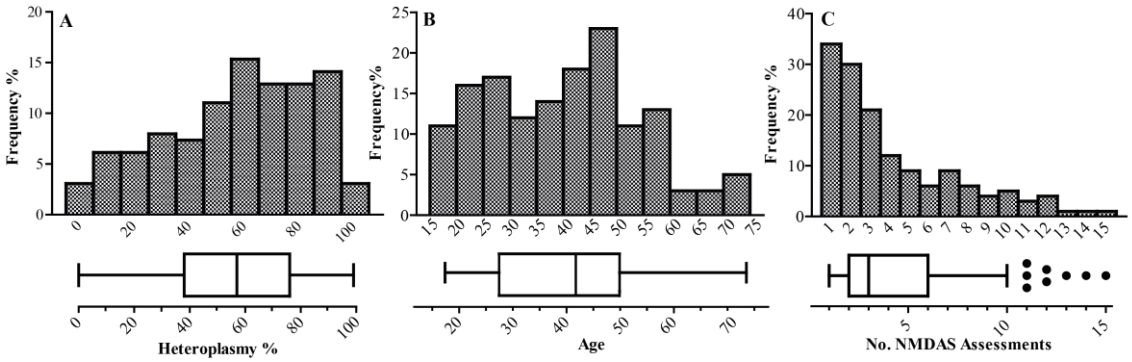


Figure 4.1 The distribution of urine heteroplasmy, age, and number of assessments in the cohort.

The cohort comprises 148 patients. **(A)** The distribution of urine heteroplasmy in our cohort. The median is 57%, with lower and upper quartiles of 38% and 76%. The distribution is skewed, with a long tail in the low heteroplasmy range. **(B)** The distribution of age in our cohort. The median age is 42, range 17-73. **(C)** Number of NMDAS assessments per patient. The median is 3 assessments, 90th percentile is 10 assessments.

4.3.3.2. Phenotypic spectrum

The phenotypic overview of the cohort is depicted in Figure 4.2. The most common features at any severity are deafness and exercise intolerance, followed by GI dysfunction, psychiatric disturbance and ataxia. For scores of 2 or above (panel B), deafness and GI dysfunction and are both prevalent at over 60%. Additionally, from the current function section of the NMDAS, exercise tolerance and gait stability stand out as prevalent at over 40%. From the symptoms (section II) psychiatric disturbance, migraine, and diabetes have a prevalence of at least 40%. Seizures affect around 22% of patients, stroke-like episodes 17% and encephalopathic episodes 18%. In section III of the NMDAS (clinical signs) myopathy and cerebellar ataxia most prevalent, with cognitive impairment also relatively common. Ptosis affects 20% of patients and CPEO 16%.

4.3.3.3. Common phenotypic features distribution with age

Figure 4.3 illustrates the proportion of patients in each age group that have the given phenotypic features (defined as scoring 2 or above in the NMDAS). Deafness has increasing incidence with age, as does CPEO, which only substantially affects the oldest patients in the cohort. Ataxia affects a proportion of the youngest patients, and a steady proportion of those aged over 45. Stroke and seizure are rare but do not appear linked with age, though both are absent in the patients over 60.

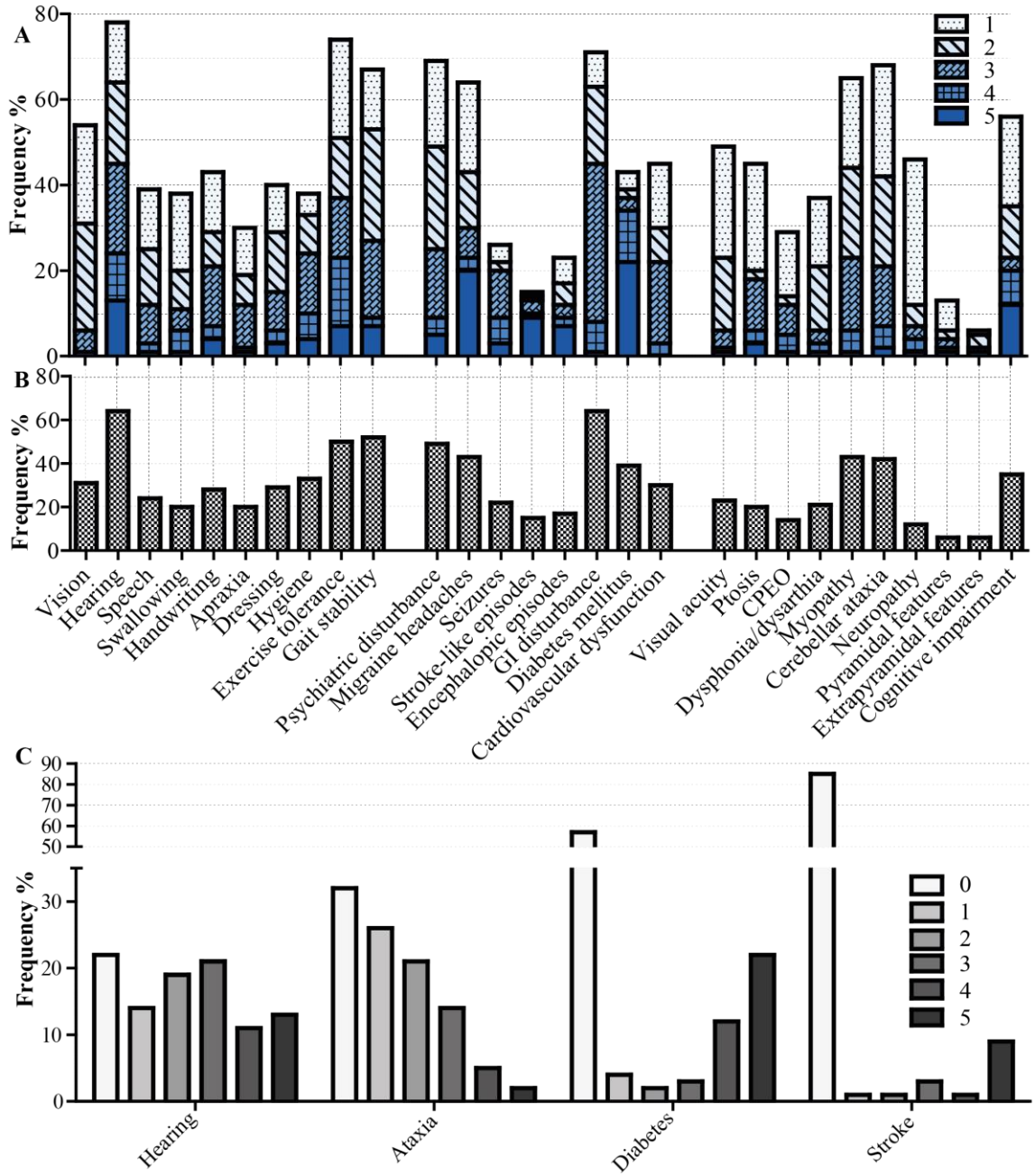


Figure 4.2 Phenotypic characterisation of the cohort.

(A-B). Profile of phenotypic symptoms from the cohort. For each patient, the maximum score in any NMDAS assessment for each NMDAS question is calculated. **A** is a stacked bar chart of the frequencies of each score (the proportion of the 146 patients that reach or exceed the given score); **B** summarises the frequencies for patients with moderate or severe symptoms (scoring 2 and above). The most common symptoms at moderate to severe level in our cohort are hearing loss and GI disturbance, followed by gait instability, exercise intolerance and psychiatric disturbance. Seizures and stroke affect 15%-20% of patients. **(C)** NMDAS score profile for four significant phenotypic characteristics of patients with the m.3243A>G mutation. Scores range in each case from no symptoms (0) to severe impairment (5).

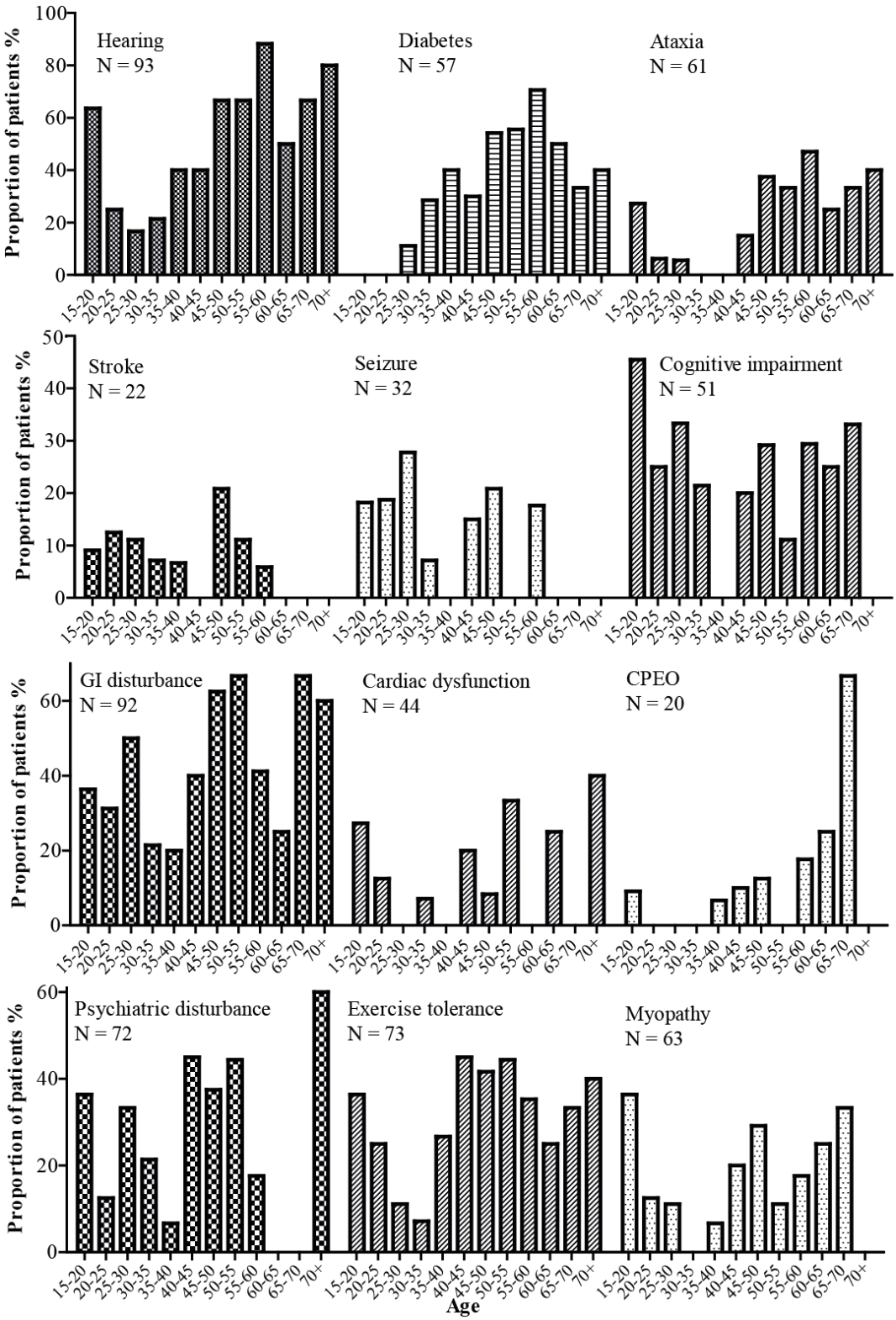


Figure 4.3 Proportion of patients in each age group with particular phenotypic characteristics.

For each feature, N indicates the total number of patients scoring 2 or above. Each bar shows the proportion of patients in that age group scoring 2 or above (the number of patients in each age group can be found in Figure 4.1B). The presentation of most phenotypic features is fairly constant across age groups, including stroke and seizures, which may have been expected to peak in frequency in younger patients; though stroke and seizure are notable absent in the 60+ age groups. Deafness is generally increasing in incidence with age; incidence of diabetes peaks around 55-60 and decline thereafter. CPEO is more common in older patients.

4.3.4. Total disease burden

4.3.4.1. Introduction

My examination of total disease burden is done in two parts. The first uses a single summary NMDAS score for each patient to get an understanding of how disease burden and predictors are related. I look at the relationship of urine heteroplasmy, age and disease burden.

In the second part I use repeated measures mixed modelling to model disease progression over time. I used urine heteroplasmy and age as predictors in this model. As inter-family clustering of disease phenotype and severity has been noted (Hammans *et al.*, 1995), I also introduce familial lineage as a predictor to evaluate the predictive effect.

4.3.4.2. Methods

The general methods for basic statistical analyses are described in section 2.11.1.1.

Longitudinal modelling is conducted as described in section 2.11.1.2. Appendix IV contains the SAS code to generate the model.

Each analysis uses the largest cohort of patients available with the appropriate predictors.

4.3.4.2.1. Familial lineage

To incorporate the family lineage into the longitudinal modelling, I divided the cohort into two groups; those with other family members in the cohort (familial) and those without other family members (non-familial). The repeated measures statement used this grouping to define the covariance structure of the data. To avoid over-parameterisation, familial lineage was introduced as a random effect.

4.3.4.2.2. Early-onset

To investigate whether patients with early-onset had significantly different disease progression from other patients, a flag was introduced to identify these patients. Early-onset was defined as any NMDAS assessment on record under the age of twenty-five that scored above a nominally low value of 3.

Longitudinal mixed modelling used this as a grouping for defining the covariance structure of the data. Models incorporating both family and early-onset used both of

these groups to define the covariance structure, i.e. in four groups (early-familial, early-non-familial, late-familial, late-non-familial).

4.3.4.3. Results

4.3.4.3.1. NMDAS score is predicted by age and heteroplasmy

Multiple linear regression was used to investigate the relationship between total disease burden and the predictors of age and urine heteroplasmy. Box-Cox analysis identified the fourth root of the NMDAS ($NMDAS^{0.25}$) as the optimal transformation, and this transformed variable was used as the dependent variable in all analyses.

Table 4.1 lists the P-values for each effect and overall R^2 for the model; standardized coefficients are also shown for analyses with multiple continuous predictors. Though heteroplasmy (P = 0.0063) and age (P = 0.0030) are both significant predictors of total disease burden, R^2 is extremely low for each predictor (5.1% and 6.0% respectively). Using multiple linear regression with both predictors, adjusted R^2 increases to 17.6%, and both predictors are highly significant (P < 0.0001). Diagnostic plots for the regression are shown in Figure 4.4.

Model Predictors	N	Heteroplasmy	Age	R^2	Adjusted R^2
Age	146	n/a	0.0030	6.0%	6.0%
Heteroplasmy	146	P = 0.0063	n/a	5.1%	5.1%
Age and Heteroplasmy	146	P < 0.0001 b = 0.37	P < 0.0001 b = 0.40	18.6%	17.6%

Table 4.1 Proportion of variability in total disease burden (as measured by NMDAS score) explained by predictive factors.

P values are shown for each predictive factor, and standardised parameters where all predictors are continuous. Age and heteroplasmy together explain 17.6% of the variance in NMDAS score.

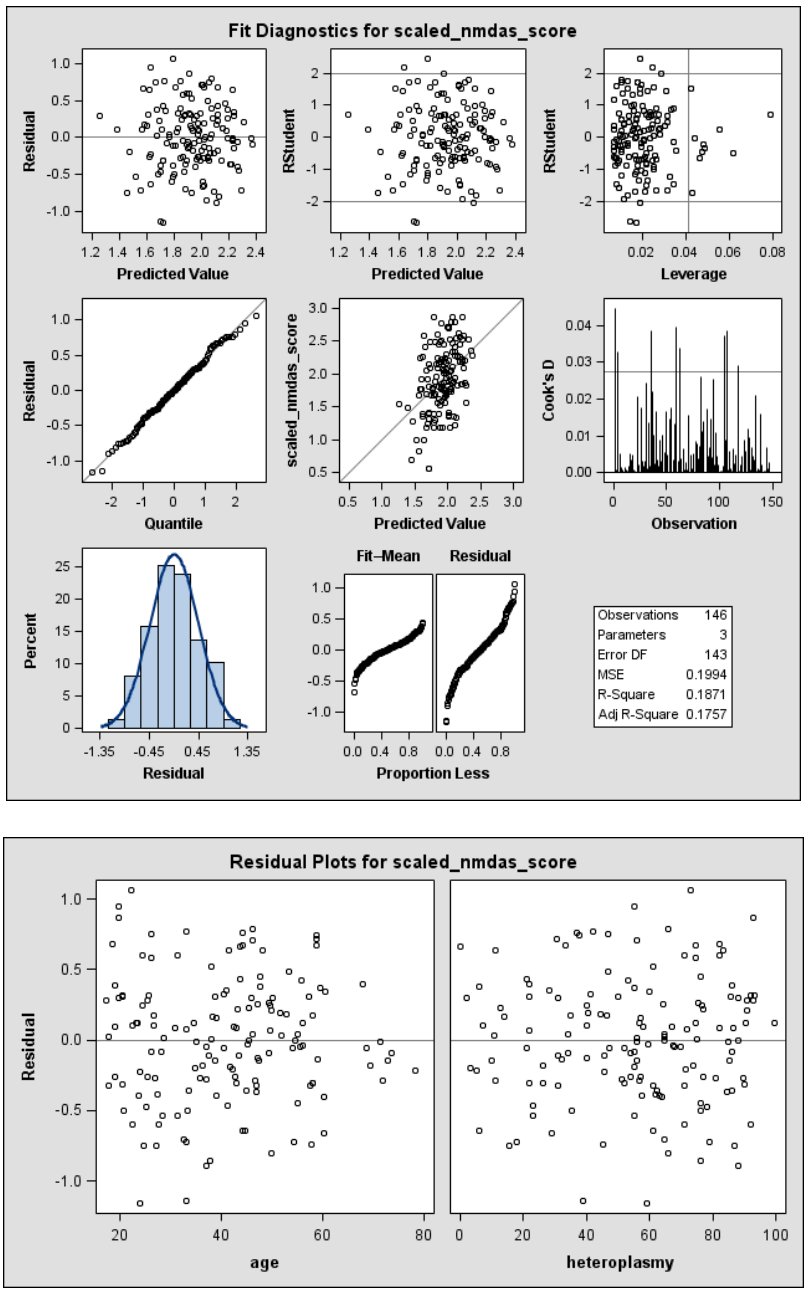


Figure 4.4 Model fit diagnostics for the multiple regression of age and hetereoplasmy on scaled NMDAS score.

The residuals are normally distributed and there are no highly influential points. There are no observable trends between the residuals and either the predicted value of the model or the two predictors of age and heteroplasmy. There is evidence of heteroscedacity in the model; the extreme residuals are only found in the youngest part of the cohort. The White test (White, 1980) identifies this as non-significant (P = 0.0554).

4.3.4.3.2. Longitudinal model fit and development

A preliminary longitudinal model using age and urine heteroplasmy as predictors was created and the fit of the model examined in detail. The residuals of the model were analysed in two ways; the distribution of the total set of studentised residuals (N=606), for a general understanding of model fit, and also the average studentised residuals for

each patient (N=146), to ensure independence of the data since there is intra-patient correlation of the residuals.

Figure 4.5 (left panel) illustrates that the model is somewhat ill fitting; there is a cluster of data points with high residuals that do not adequately fit the model. Though the average studentised residuals are independent of heteroplasmy (P = 0.7176, r = 0.02) and age (P = 0.3427, r = -.0.08), they residuals did show a significant relationship with the predicted value of the model (P = 0.0012, r = -0.26) as seen in Figure 4.11.

Scrutiny of the data indicated that the patients that were not well described by the model were young and with heavy disease burden. To better model these patients, a further predictor was added into the model to identify early-onset patients, defined as those under twenty five with non-trivial NMDAS scores. This flag dichotomises the cohort into two groups, early-onset (N = 28) and late-onset (N = 118). Figure 4.5 (right panel) illustrates the residuals of the model with this new parameter included; they are now well formed and normally distributed. Analysis of the average studentised residuals found they were now uncorrelated with the predicted value (P = 0.3636, r = -0.08), heteroplasmy (P = 0.5565, r = 0.05) and age (P = 0.4469, r = 0.07). Additionally, the AIC for the model dropped from 18.5 to -57.2 indicating a significant improvement in the model fit.

All three parameters (time, heteroplasmy, and early-onset) are highly significant (P < 0.0001).

Covariance parameters for the model are found in Table 4.2. The model includes separate residual variances for the early- and late-onset patients, i.e. residual variance is grouped according to this parameter. If the variance is not modelled separately, AIC of the model is -34.2, as compared to -57.2, indicating that allowing separate residual variances improves the model.

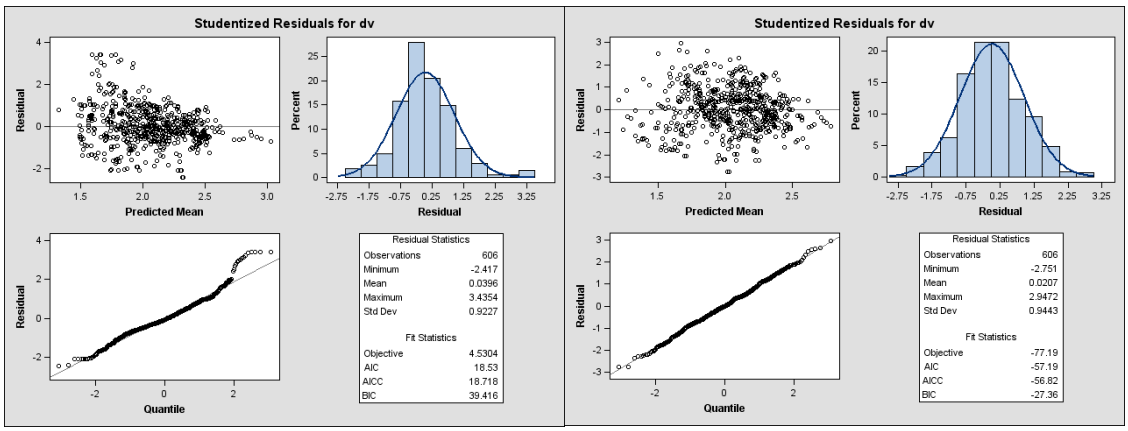


Figure 4.5 Residual diagnostics for models with and without early-onset indicator variable.

(Left) Residuals for the model without early-onset predictor. The residuals are non-Gaussian, indicating an underlying non-random error in the model specification. The cluster of points in the top left of the diagram have high residual error and low predicted mean; these correspond to early-onset patients with severe disease burden. **(Right)** The same model with early-onset predictor. The residuals are now balanced and normally distributed.

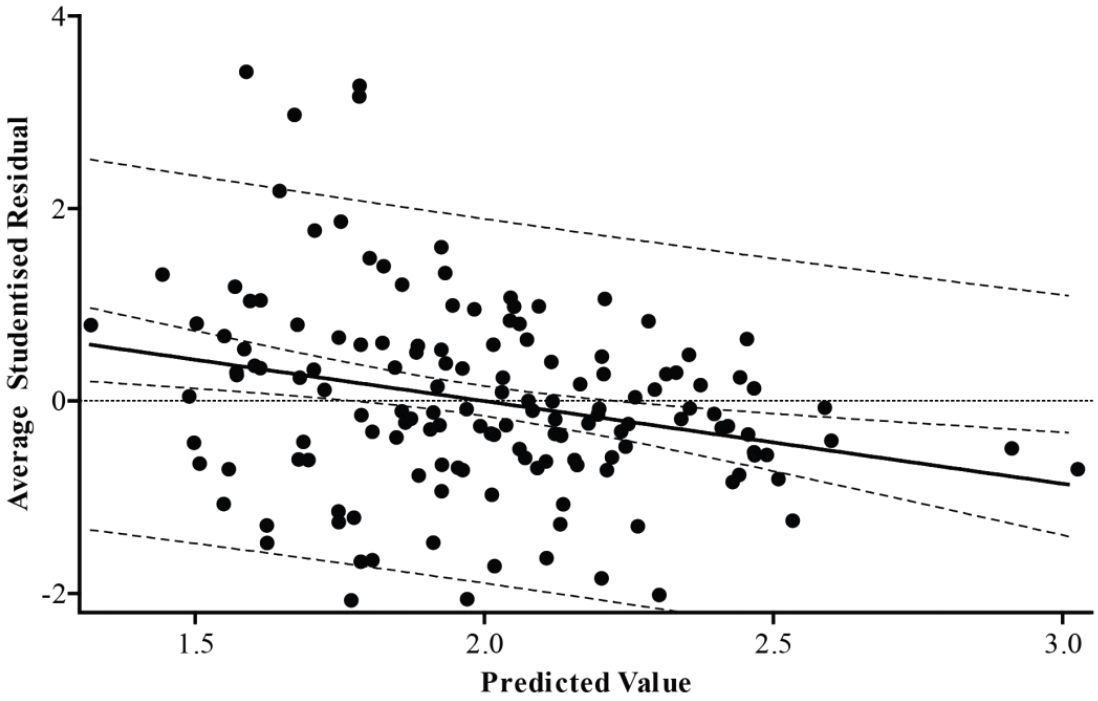


Figure 4.6 Average studentised residuals vs. predicted model value for a linear mixed model with age and urine heteroplasmy as predictors.

The graph shows that the residuals are correlated with the predicted value (P = 0.0012, r = -0.26). This indicates the model is not a good representation of the underlying data structure. The cluster of points in the top left of the graph is particularly problematic, as they are very poorly fitted by the model.

Group	Spatial Power	Residual Variance
Time		0.000083
Late-onset	0.9783	0.02352
Early-onset	0.9998	0.1915

Table 4.2 Covariance parameters for the longitudinal mixed model.

The residual variance of the early-onset group is far greater than that of the late-onset group; this is a reflection of the more erratic scores from patients in the early-onset group.

4.3.4.3.3. Longitudinal modelling of disease progression

Figure 4.7 illustrates the actual NMDAS scores for each patient in the cohort, grouped by the heteroplasmy quartiles, which demonstrates the variability in disease burden for patients with similar levels of heteroplasmy.

Figure 4.8 illustrates the longitudinal modelling of total disease burden in patients with the m.3243A>G mutation. Both heteroplasmy and age are highly significant predictors ($P < 0.0001$). The graph shows predicted progression for a nominal patient with selected heteroplasmy levels, with 95% confidence intervals.

Figure 4.9 depicts the actual and predicted progression of total disease burden for several patients with comparable heteroplasmy levels. The median (60%) and a high (90%) level of heteroplasmy have been utilised for illustration purposes.

Figure 4.7 Actual NMDAS scores for each patient in the cohort.

Patients are grouped into 4 quartiles by heteroplasmy. NMDAS scores for each patient are joined by solid lines, and labelled with urine heteroplasmy %. At low heteroplasmy (**A**, <35%), disease progression is slow and relatively homogenous. In the second quartile of patients (**B**) heterogeneity is seen to increase. The two upper quartiles (**C** & **D**) appear similar in the disease burden, though there is a shift to the left (earlier presentation) in the highest quartile (**D**). There is considerable overlap; patients with 2% and 7% heteroplasmy (**A**) have similar disease burden to patients with 87% and 92% heteroplasmy (**D**) at similar ages.

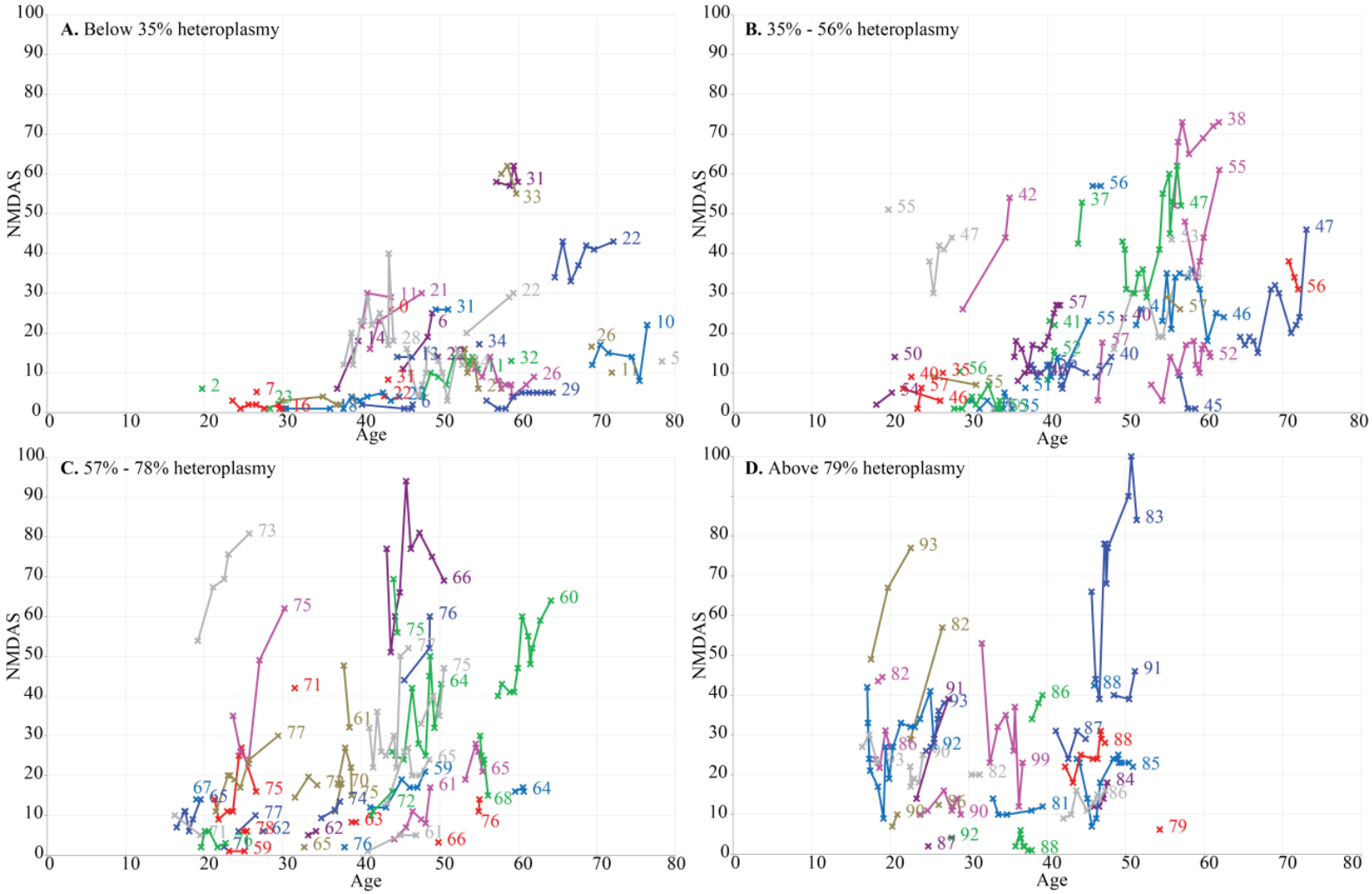
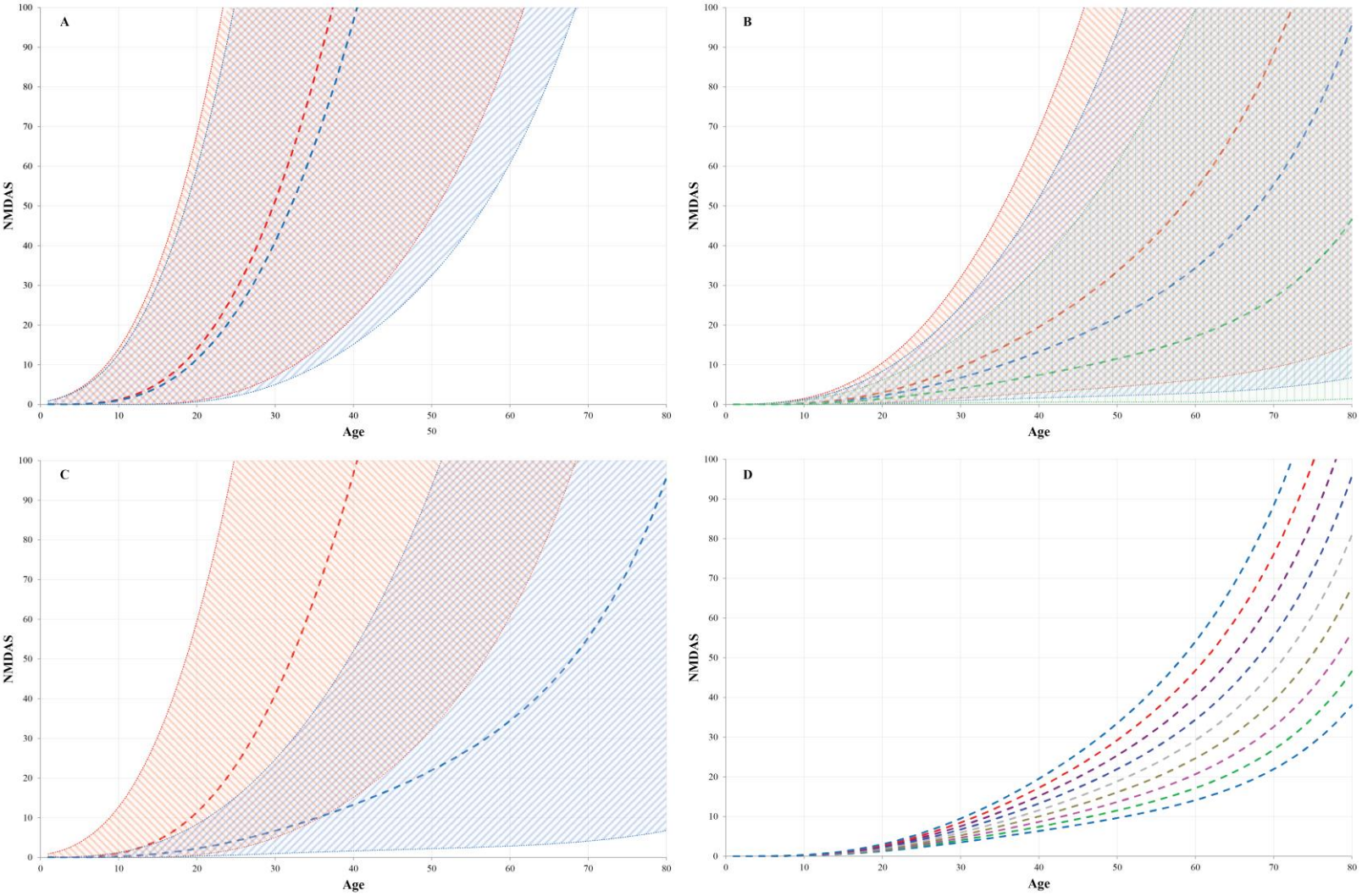


Figure 4.8 Longitudinal modelling of total disease burden.



Predicted progression depicted as dashed lines, 95% confidence intervals also shown except panel **D**. Both heteroplasmy and age are highly significant predictors in all cases ($P < 0.0001$). **(A)** Predicted progression of total disease burden for early-onset patients with 60% (blue) and 90% heteroplasmy (red). **(B)** Predicted progression for late-onset patients, with 20% (green), 60% (blue) and 90% (red) heteroplasmy. **(C)** Comparison of predicted progression for early (red) and late (blue) onset patients at 60% heteroplasmy. **(D)** Predicted progression for patients with 10% to 90% heteroplasmy in 10% intervals.

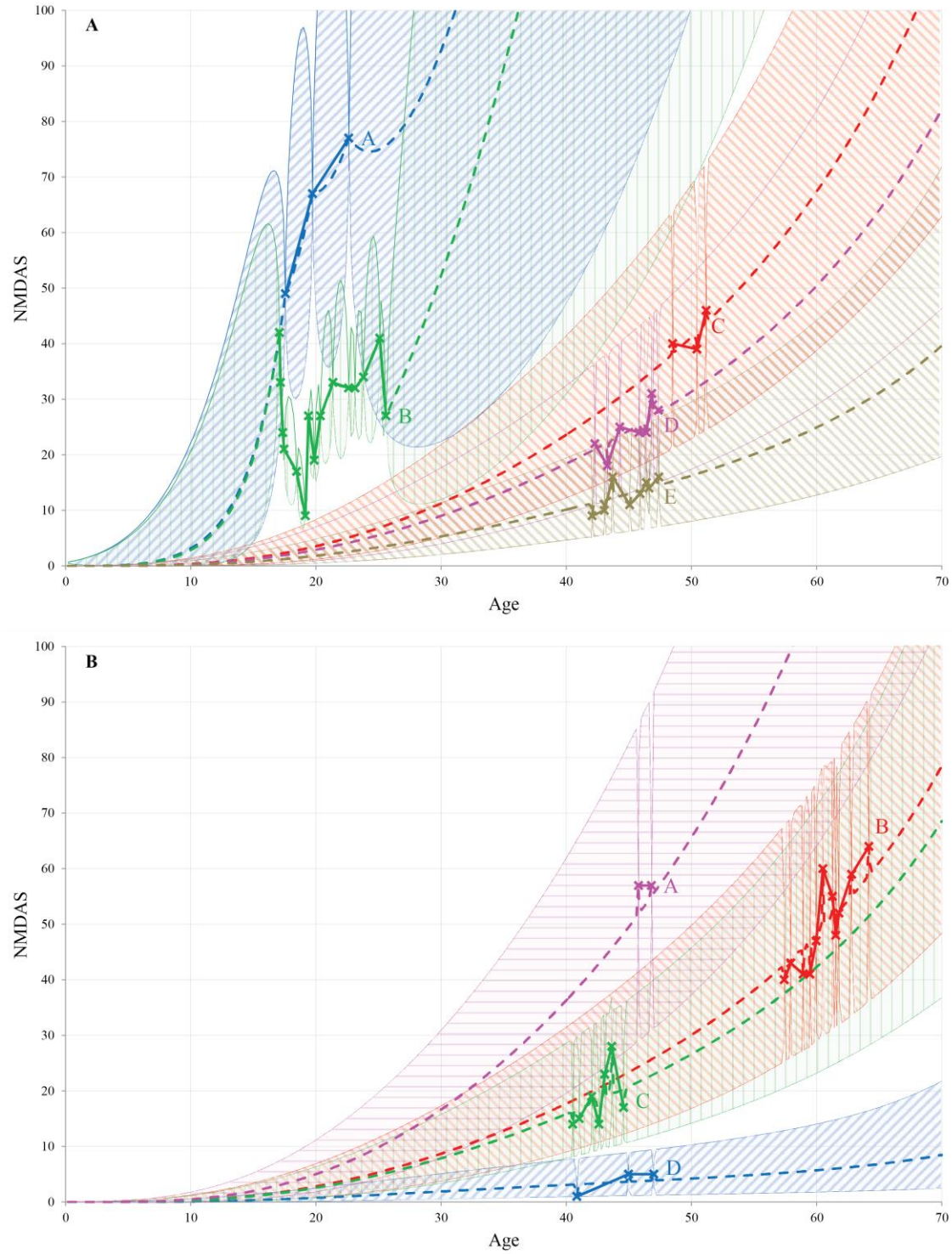


Figure 4.9 Examples of individual patient predicted progression.

Predicted and actual progression for a selection of patients with approximately 90% (A) and 60% (B) heteroplasmy. Actual assessment scores are depicted as crosses joined by solid lines. Dotted lines with shaded 95% confidence intervals denote predicted progression for each patient. These patients illustrate the huge variability of disease burden progression for patients with similar heteroplasmy levels of the m.3243A>G mutation. The only early-onset patients are patient A and B in panel A; the confidence intervals for their progression are much wider than the other patients, indicative of the higher variability for early-onset patients.

4.3.4.3.4. Familial lineage is predictive of disease progression

With mixed modelling familial lineage can be introduced into the model as a random effect, allowing the effect of family to be represented with a single parameter that represents the variability of the family lineage.

A number of models with family as a random effect were investigated. The models investigated all included early-onset as a fixed effect, and various forms of the residual covariance were investigated to find the optimal representation.

Table 4.3 illustrates the evaluation of model fit. The fall in the AIC with the inclusion of family as a random effect indicates the model is improved; this is shown to be significant by the likelihood ratio test (P = 0.0343).

Figure 4.10 shows an example the intra-familial clustering. The graph shows two familial lineages with sharply different total disease burden progression rates. The average progression rates for all families with multiple members in the cohort (N = 36) are show in Figure 4.11.

		Repeated measures groups		
		Early-onset * familial	Early-onset	Familial
Random effects	Patient	-80.5	-57.2	-52.8
	Family	NC	78.9	72.5
	Patient & family	-89.2	-65.6	-64

Table 4.3 Investigation of optimal model with family as a random effect.

AIC is shown for each model. NC indicates non-convergence of the model. The familial grouping for the repeated measures allows the variance of familial patients (those with other maternal relatives in the cohort) and non-familial (those with no maternal relatives in the cohort) to be modelled separately. The early-onset repeated measures group allows the variance of early-onset and late-onset patients to be separately estimated. An improvement in the AIC indicates that these extra parameters improve the model. The optimal model includes both patient and family as random effects, as this has the lowest AIC. The repeated measures grouping is the interaction of early-onset and family.

.

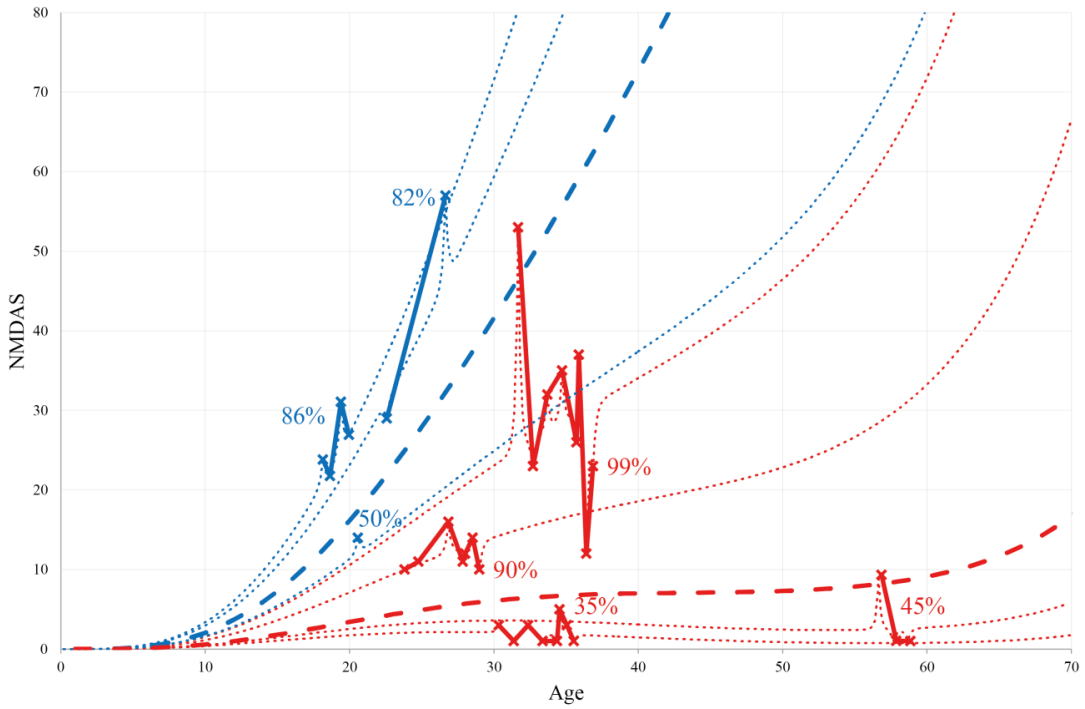


Figure 4.10 Intra-familial clustering of total disease burden progression.

The predicted progression for members of two different families are shown. Actual assessment scores for individual patients are shown as crosses joined by solid lines, patients are labelled with their heteroplasmy level. Predicted progression for individual patients are shown as dotted lines. The predicted progression for a nominal patient from each family with 60% heteroplasmy is shown as a heavy dashed line. The family depicted in red has slower disease progression than the family in blue.

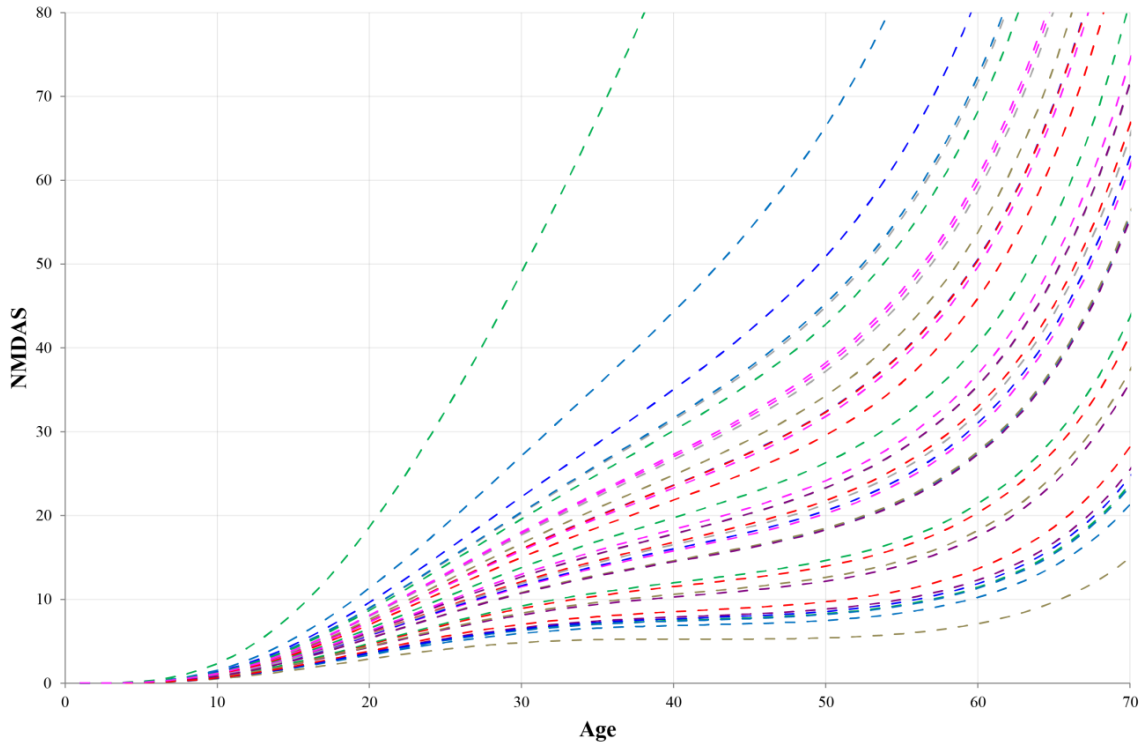


Figure 4.11 Predicted progression for a nominal patient with 60% heteroplasmy for each family in the cohort.

Heteroplasmy ($P < 0.0001$), age ($P < 0.0001$), and family ($P = 0.0003$) are all highly significant predictors.

4.3.4.4. Total disease burden discussion

The phenotypic variability associated with the m.3243A>G mutation has been frequently remarked upon (Hammans *et al.*, 1995; Chinnery *et al.*, 1997; Kaufmann *et al.*, 2011; de Laat *et al.*, 2012; Nesbitt *et al.*, 2013) and is well recognised by clinicians, however little attempt has been made to quantify this variability. The NMDAS scale allows such quantification to be made. I have shown that age and the urine heteroplasmy level together account for only around 18% of the variance in total disease burden, which is noticeably low, but resonates with anecdotal evidence from clinicians on the variability of disease in patients with the m.3243A>G mutation, as well as previous studies that have noted the weak association between heteroplasmy and phenotype (Kobayashi *et al.*, 1992; Martinuzzi *et al.*, 1992; Shiraiwa *et al.*, 1993; Liou *et al.*, 1994; Chinnery *et al.*, 1997). It is a clear that there is a vast amount of variation in disease burden that is unexplained by the predictors currently available in this analysis.

The detailed diagnostic examination of the model is cumbersome but gives reassurance about several key points. Firstly, there are no observable trends between the residuals and the predicted value, or the predictors. This is an important validation of the transformation used ($NMDAS^{0.25}$) and reassurance that no further transformation of the predictors are required. Secondly, the spread of the data is even; there are no highly influential points at either high or low heteroplasmy or age. This gives confidence in the interpretation of the R^2, as this can be strongly affected by influential data points. This reassurance about the shape and composition of the data set is important, not just for the basic modelling using single summary data points, but also for the longitudinal modelling.

The heteroscedacity of the residuals in the basic modelling with respect to age needs to be noted, however it is not unexpected. Patients in their early twenties or younger have a hugely varied disease burden, including severe burden associated with MELAS associated stroke-like episodes, but also much milder presentation, and there are a number of more or less asymptomatic children who are known to have inherited m.3243A>G from affected mothers who contribute to this variation. It is important to note that these basic statistical analyses are conducted to gain an overview of the relationship between the predictors and total disease burden; longitudinal mixed modelling allows modelling any heteroscedacity inherent in the system, and indeed, the

covariance parameters from the longitudinal models indicate that this heteroscedacity is incorporated.

The longitudinal modelling graphically illustrates two key points in this regard. Firstly, the unexplained variance is noticeably large; though both m.3243A>G heteroplasmy and age are highly significant predictors of total disease burden, the confidence intervals for progression of a nominal patient at a given heteroplasmy level are wide. Secondly, the predictive power of the family lineage is clearly illustrated, which concurs with previously reported observation (Hammans *et al.*, 1995), and suggests that a reasonable proportion of this unexplained variance can be explained by a combination of genetic factors of either nuclear or mtDNA origin, though environmental or other epi-genetic modulatory factors cannot be disregarded. Indeed, the nuclear genetic background has previously been shown *in vitro* to impact on m.3243A>G phenotypic expression (Dunbar *et al.*, 1995).

At an individual patient level, the power of longitudinal mixed modelling is well demonstrated by the graphical outputs shown in Figure 4.9; the introduction of random effects into the model demonstrates that individual patient progression can be modelled in spite of the wide confidence intervals observed using heteroplasmy and age as fixed effects. This modelling is critical in understanding the expected disease progression of individual patients.

The development of the longitudinal model demonstrates important points. Firstly, the model is ill-specified without accounting for a distinct group of early-onset patients. This division must be included in the model as a predictor (the early-onset flag); the residuals of the model clearly show that the model misrepresents the underlying data structure. There are some limitations in this approach. Firstly, the definition of early-onset relies on an arbitrary cut-off, which I set at the age of twenty-five. This cut-off point was derived on preliminary observation of the dataset that indicated poor model fit for a large number of patients under this boundary. However, I investigated models with lower and higher cut-offs and twenty five was more or less optimal for this data set. Secondly, it is a potential limitation of the methodology that inclusion within the early-onset group is contingent on having an NMDAS assessment under this age; it is feasible that there are patients within the cohort that were symptomatic before the age of twenty-five but not seen in the Newcastle clinic until over this age. However, despite both of these caveats, the inclusion of this predictor in the model comprehensively resolved

issues regarding the fit of the model, which is indicative of the success of the methodology. It might be argued that dividing the cohort by age is superfluous, since age is already included in the model as a predictor. However, it is critical to note that both age and the early-onset predictor are simultaneously predictive in the model, and thus the predictor has been shown to be empirically valuable.

An imperative to understanding m.3243A>G disease is to deepen understanding of the pathological mechanisms that cause variability in phenotypic expression, and early-onset of symptoms is a fundamental facet of this variability. It is clear from this modelling that heteroplasmy alone is not an adequate descriptor of the underlying pathology. There are several potential areas for further exploration of this, which will be discussed in the discussion on general disease progression in m.3243A>G in section 4.3.6.

4.3.5. Individual phenotypic features

4.3.5.1. Introduction

An understanding of the expected development of specific phenotypic features is of vital import to clinicians in care and management of patients.

As a semi-quantitative assessment scale, the NMDAS enables modelling of phenotypic features. As discussed in section 2.8, logistic regression is necessary to model discrete data such as individual NMDAS question scores. However, longitudinal modelling of discrete data is considerably less flexible as compared to continuous or pseudo-continuous data such as the overall NMDAS score, and thus the approach I take is constrained by the tools available to analyse the data.

There are prior studies that have considered the appearance of specific phenotypic features and the correlation with heteroplasmy (Chinnery *et al.*, 1997; Liu *et al.*, 2013). However, though both acknowledged age as a potential confounding factor, neither study incorporated age into their analysis. Since many features of mitochondrial disease are considered progressive this is a significant limitation.

4.3.5.2. Aims

My aim in this section is to use multiple logistic regression to study the relationship between individual phenotypic features and the predictive factors of age and urine heteroplasmy. I aim to distinguish which features are correlated with either one of these predictors or both.

I also will compare the multiple regression models with the simple logistic regression models with a single predictor of either age or heteroplasmy, to better understand the relationship between these predictors and each phenotypic feature.

4.3.5.3. Methods

The general methodology for logistic regression is explained in section 2.8.

I used two approaches to logistic regression. Both used the predictors of age and urine heteroplasmy. For each NMDAS question, a single summary score was determined for each patient by taking the maximum score achieved on each question and the age at which that score was first recorded.

The first approach used binary (dichotomous) logistic regression by identifying the optimal cut-off point for each NMDAS question to divide the cohort into two groups. To identify the optimal cut-off point the area under the ROC curve (AUC) was maximized. This was done separately for each individual predictor and for the model with both predictors, to allow a full understanding of the use of each predictor. Only cut-off points that partitioned the data into sets containing at least 10 patients in each group were considered to avoid small numbers of patients skewing the results. The code for this model can be found in Appendix V.

The second approach used a proportional odds multiple logistic regression. NMDAS scores were re-categorised as asymptomatic (NMDAS = 0), moderate (1-3), and severe (4-5). This re-categorisation was necessary for a majority of phenotypic features for the model to conform to the proportional odds assumption. Pseudo-R^2 values, as described in 2.8.2, were used to compare models. The code for this model can be found in Appendix VI.

In both approaches, standardised parameters for the predictors were calculated.

4.3.5.4. Results

4.3.5.4.1. Binary logistic regression

Using a model with age and urine heteroplasmy as predictors, the area under the ROC (AUC) curve was calculated for each phenotypic feature at each dichotomous cut-off point that had at least 10 patients in each group (below the cut-off point, and above or including the cut-off point). The AUCs for each cut-off point are shown in Figure 4.12. The higher the AUC the better the predictive power of the model. Steep gradients or inflections in the graphs indicate instability in the predictability of the features, i.e. prediction success is heavily dependent on where the cut-off is drawn. Hearing and cerebellar ataxia are the only two features to achieve an AUC over 0.8, however diabetes scores very close to 0.8. Hearing and diabetes are consistently good predictors at almost any cut-off level, and several other features are consistent across the board, albeit with less predictive power, including cognitive impairment, migraine, and encephalopathic episodes. GI disturbance is poorly predicted at any cut-off point.

Using the optimal cut-off for each NMDAS feature, models were generated using each predictor in isolation and both predictors together. The results are shown in Figure 4.14. As seen in panel C, the model with both predictors is the optimal model in almost all cases, other than GI disturbance, where the model with age alone was marginally better. Age is a better predictor than heteroplasmy for 17 of the 28 features. Migraine is the only features negatively associated with age, though not significantly so.

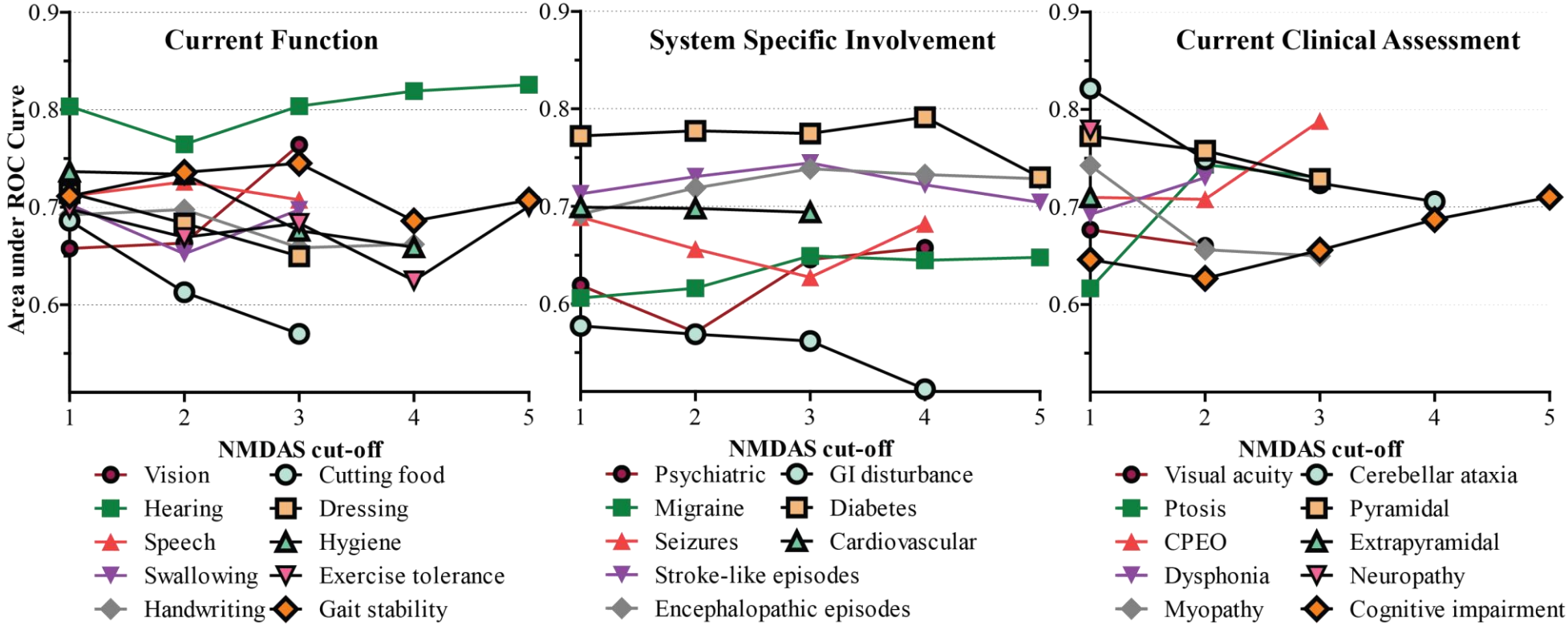


Figure 4.12 Area under the ROC curve (AUC) for each phenotypic feature and various score cut-off points with all three predictors in the model (age, heteroplasmy, and deletion size).

The cut-off score is used to divide the cohort into two groups; those scoring the cut-off or above versus those scoring below. For each feature, the cut-offs tested were those with a minimum of 10 patients in each group. Only hearing and cerebellar ataxia achieve an AUC of 0.8 at any cut-off, though diabetes scores close to 0.8 at almost all cut-off points. GI disturbance is notably poor at all cut-offs.

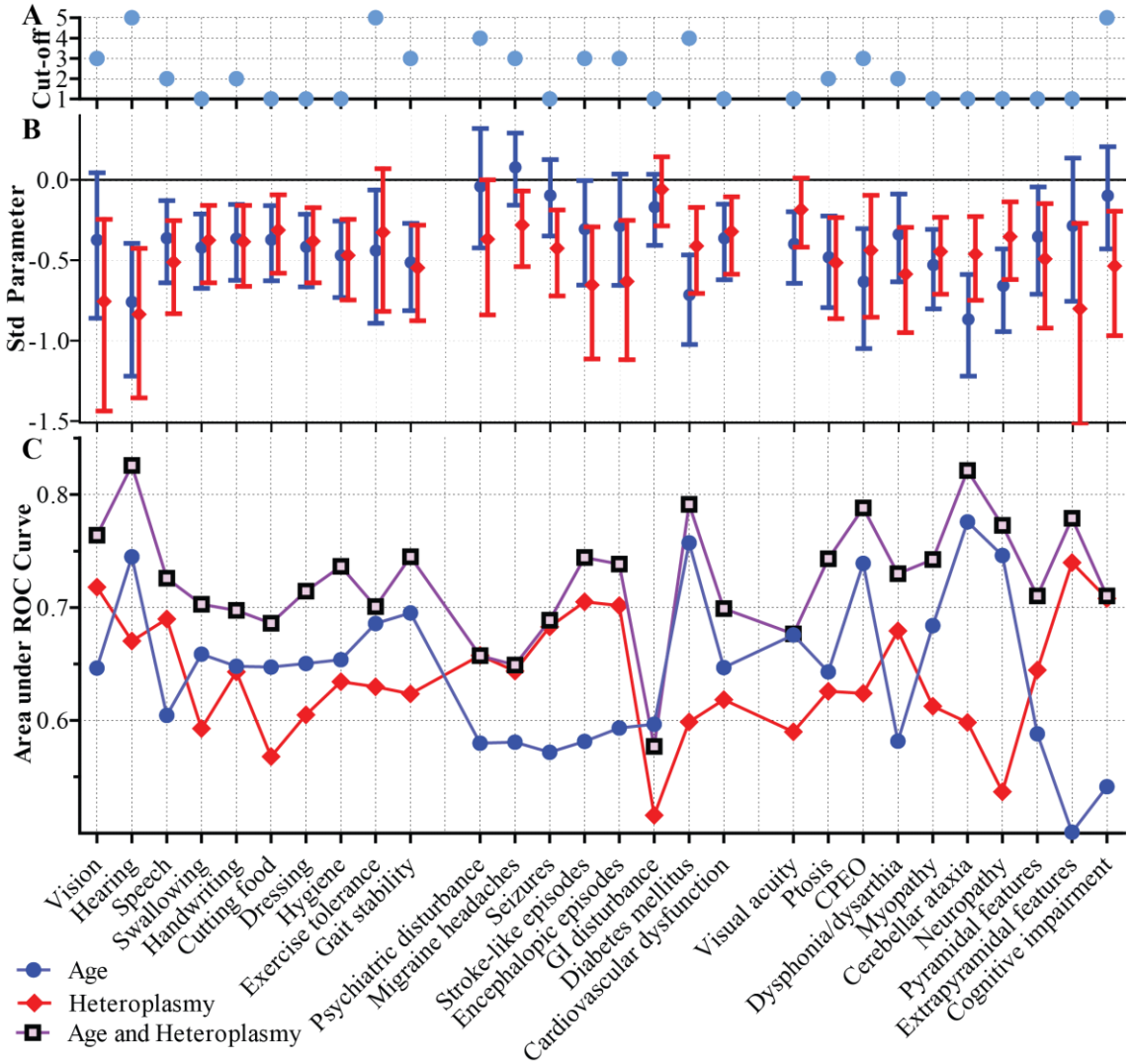


Figure 4.13 Standardised parameters and area under ROC curves (AUC) for each phenotypic feature from the NMDAS using age and urine heteroplasmy as predictors.

(A) Optimal cut-off point for each NMDAS feature, by maximizing AUC for the model with all three predictors. Cohort is dichotomised into those scoring the cut-off or above against those scoring below the cut-off. **(B)** Standardised parameters with 95% confidence intervals for the binary logistic model using the optimal cut-off with all three predictors. Parameters are statistically significant if the confidence interval does not cross the line at Y = 1. **(C)** The AUC using the optimal cut-off for each lone predictor and both predictors together. Hearing and ataxia score above 0.8, indicating strong predictive power. Diabetes achieves close to 0.8. In all cases, the regression with both predictors is optimal. Only migraine is (non-significantly) negatively associated with age. Twenty-five of 28 features are significantly associated with heteroplasmy, 20 of 28 with age. GI disturbance is very poorly predicted, but several other features have AUC values under 0.7, including cutting food, psychiatric disturbance, migraine, and seizures.

4.3.5.4.2. Phenotypic features predicted by age and heteroplasmy using multilevel ordered logistic regression

For multilevel-ordered logistic regression the odds ratios are presented; odds ratios relate to decades, i.e. a 10% change in heteroplasmy or 10 years for age. For example, the odds of having unaffected hearing decreases by a factor of 0.63 for each 10% change in heteroplasmy or by a factor of 0.42 for each decade of ageing. Calculation of the actual risk requires intercepts .The odds ratios for both age and heteroplasmy for the features in the NMDAS assessment are shown in Figure 4.14 as well as standardised parameters which allow comparison of the relative utility of age and heteroplasmy as predictors. Additionally, Appendix VII details all parameters required for risk calculation, including intercepts, and Appendix VIII is an explanation of how to calculate the risk for a given age and heteroplasmy with these parameters.

All features excepting gastro-intestinal disturbance, visual acuity, and ptosis were predictable by heteroplasmy. Several key features of classical m.3243A>G phenotypic presentation, including stroke-like episodes, seizures, and migraine, were not predictable by age.

Logistic models using a single predictor (age or urine heteroplasmy) were also generated, and compared with the multiple regression model containing both predictors, using the pseudo-R^2. To interpret the results of this comparison, an R^2 inflation factor was calculated, defined as the multiple regression R^2 divided by the sum of the individual R^2 values for the two simple regression models. An R^2 value over 1 indicates that the regression with the two predictors in the same model is more explanatory than the individual models. This is graphically illustrated in Figure 4.15, which shows both the R^2 inflation factor and the individual pseudo-R^2 values for each feature from the NMDAS. Almost all features are explained better by multiple regression. Most features from the current clinical assessment (section I) are substantially better predicted by both predictors together. There are no features that are significantly predicted by both age and heteroplasmy that have an inflation factor of 1 or below. The models for several features that are explained by only one of the predictors, for example migraine or seizures, are not improved by the inclusion of the second predictor. Figure 4.16 illustrates the change in P-value of each predictor when moving from simple logistic regression to multiple regression with both age and heteroplasmy as predictors. Several features are only significantly predicted by both age and heteroplasmy when both are included in the model together; cardiovascular dysfunction, ataxia, diabetes, exercise

intolerance, and neuropathy are not predicted by heteroplasmy unless age is included in the model; and cognition, dysphonia, pyramidal features, and speech are not predicted by age unless heteroplasmy is also in the model.

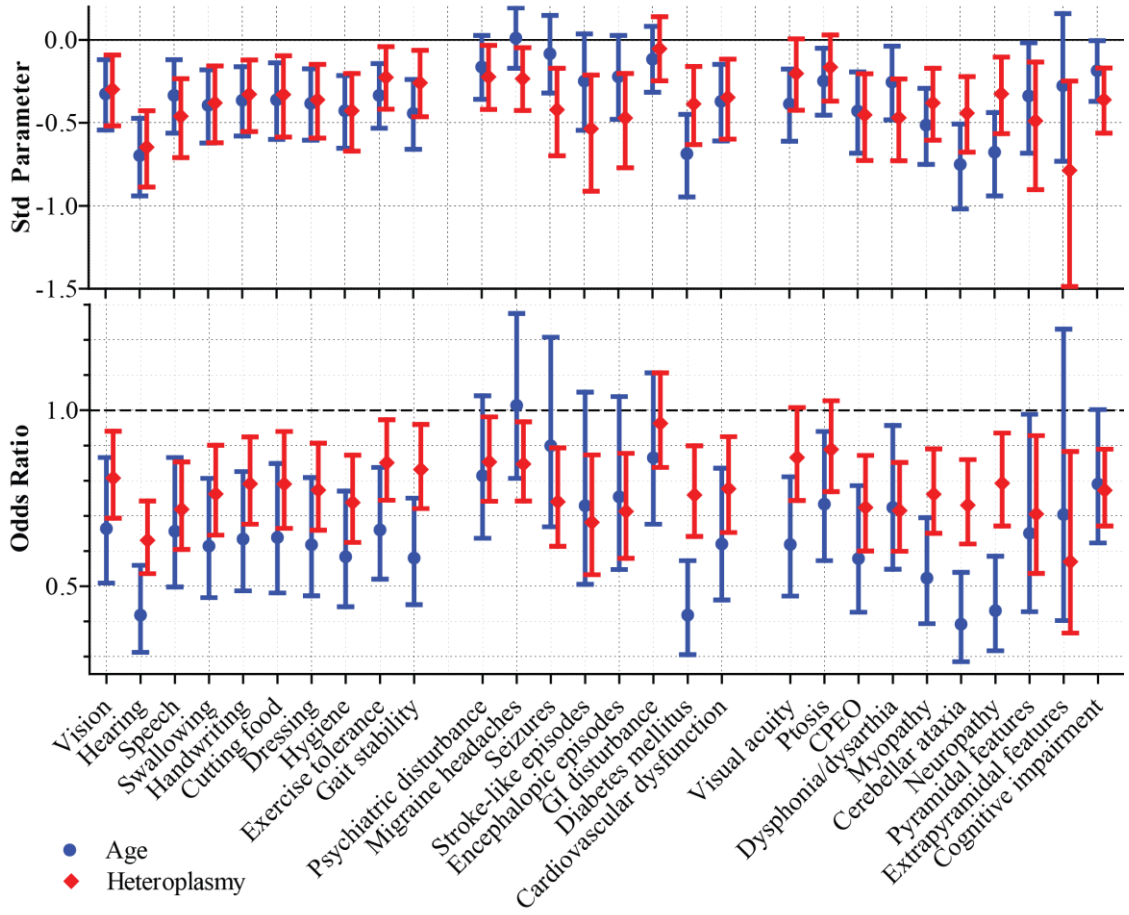


Figure 4.14 Risk of developing specific symptoms as predicted by age and heteroplasmy.
Standardised parameters (top) and odds ratios with 95% confidence intervals (bottom). Intervals crossing the line with odds ratio of 1 or standardised coefficient of 0 are not statistically significant. See the text for model development methodology. Phenotypic features are divided into the three sections of the NMDAS assessment. Ataxia, hearing, diabetes, and myopathy are the most predictive by heteroplasmy. Most features are predicted by heteroplasmy, other than GI disturbance, visual acuity, and ptosis, though both of the latter two are only just beyond statistical significance. Several features are not well predicted by age, all in the symptoms section of the NMDAS other than extrapyramidal features, though these are very rare in the cohort. GI disturbance is the only feature neither predicted by age nor heteroplasmy. Comparison of the standardised coefficients shows that age is a stronger predictor than heteroplasmy for 17 of the 28 features of the NMDAS.

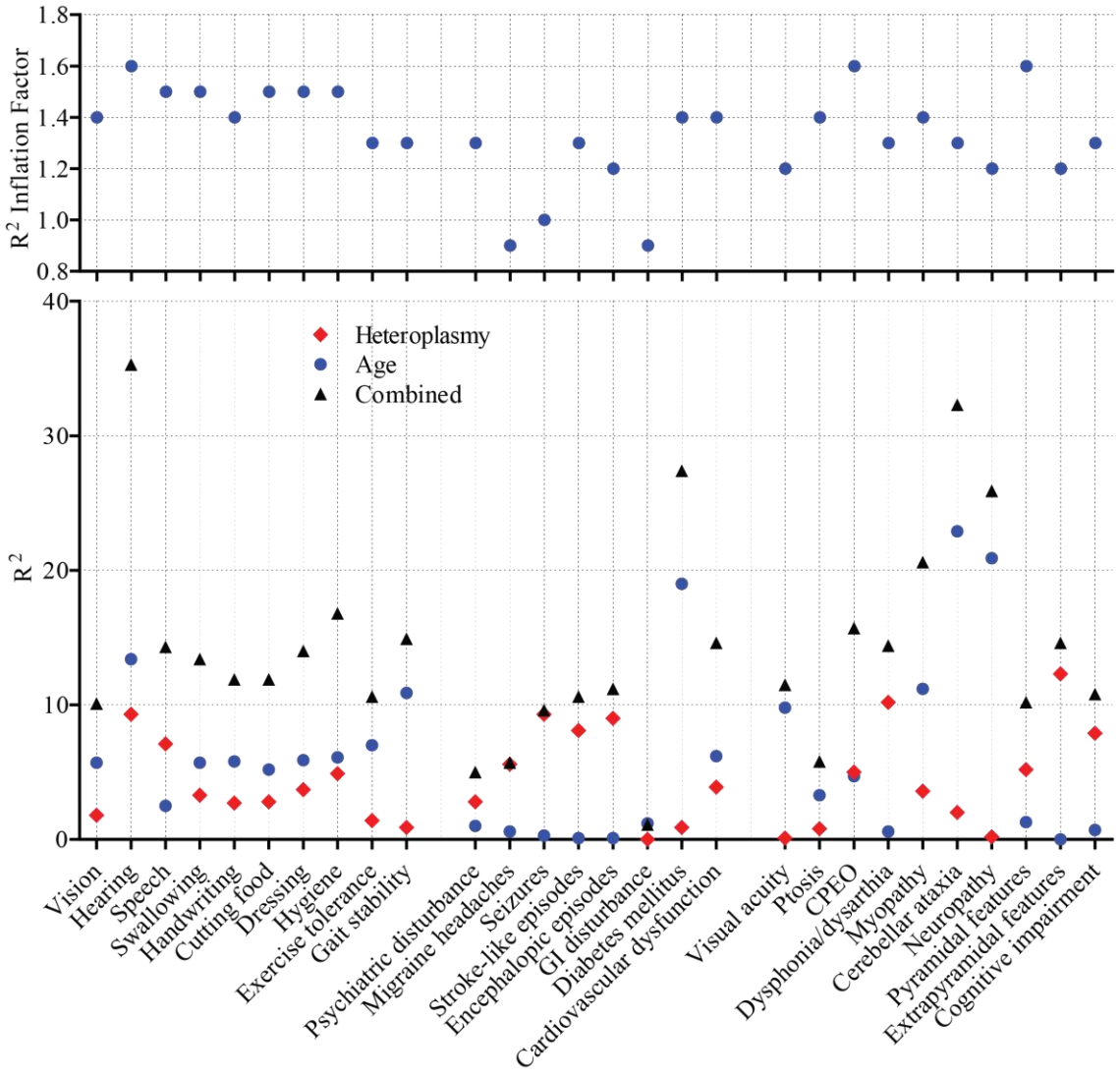


Figure 4.15 Pseudo-R^2 values for logistic regression with age, heteroplasmy, and both predictive factors, and R^2 inflation factors.

An R^2 inflation factor greater than 1 indicates that both predictive factors together are more predictive than the predictive factors in isolation. Only migraine, seizures, and GI disturbance are not predicted better by multiple regression. Hearing, CPEO, and pyramidal features are most improved by multiple regression.

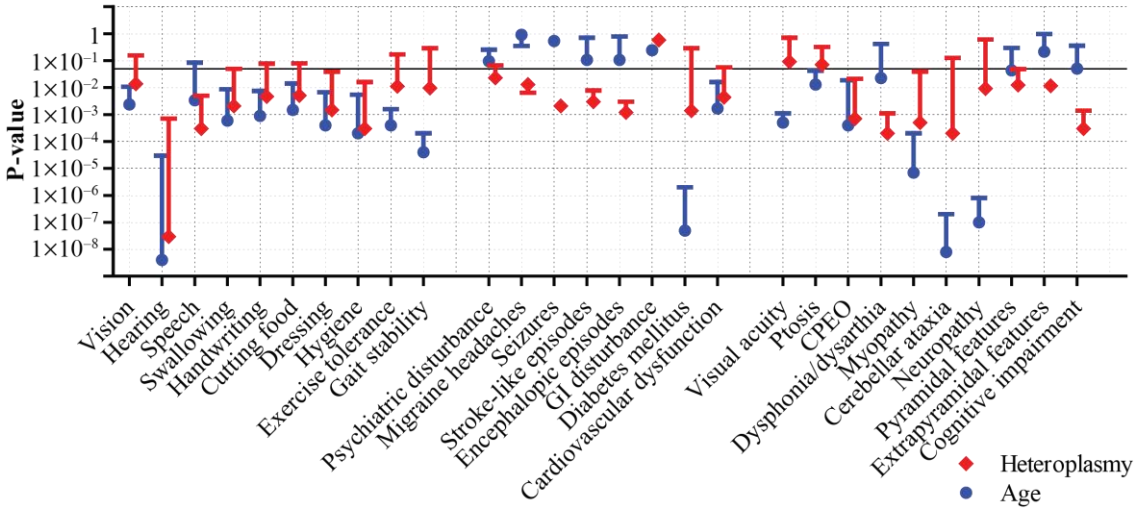


Figure 4.16 Improvement in statistical significance of predictors using multiple logistic regression rather than simple logistic regression.

Graph illustrates how the P-value for each predictor is affected when moving from simple logistic regression (horizontal bar) to multiple regression (symbol). The solid horizontal line is drawn at P = 0.05, indicating statistical significance. Almost all features show improvement in significance with multiple regression, most notably hearing, ataxia, diabetes, and myopathy. Heteroplasmy is not a significant predictor of several features including ataxia, neuropathy, and diabetes, unless multiple regression is used.

4.3.5.4.3. Risk profiles of deafness with age and heteroplasmy

I have chosen hearing to illustrate the outputs of the logistic regression model. The changing risk of deafness associated with age and heteroplasmy is depicted in Figure 4.17. Both age and heteroplasmy are highly significant predictors of deafness ($P < 0.0001$). At a given age, the risk of moderate or severe deafness increases as heteroplasmy increases; and conversely, at a given heteroplasmy the risk of more severe deafness increases with age.

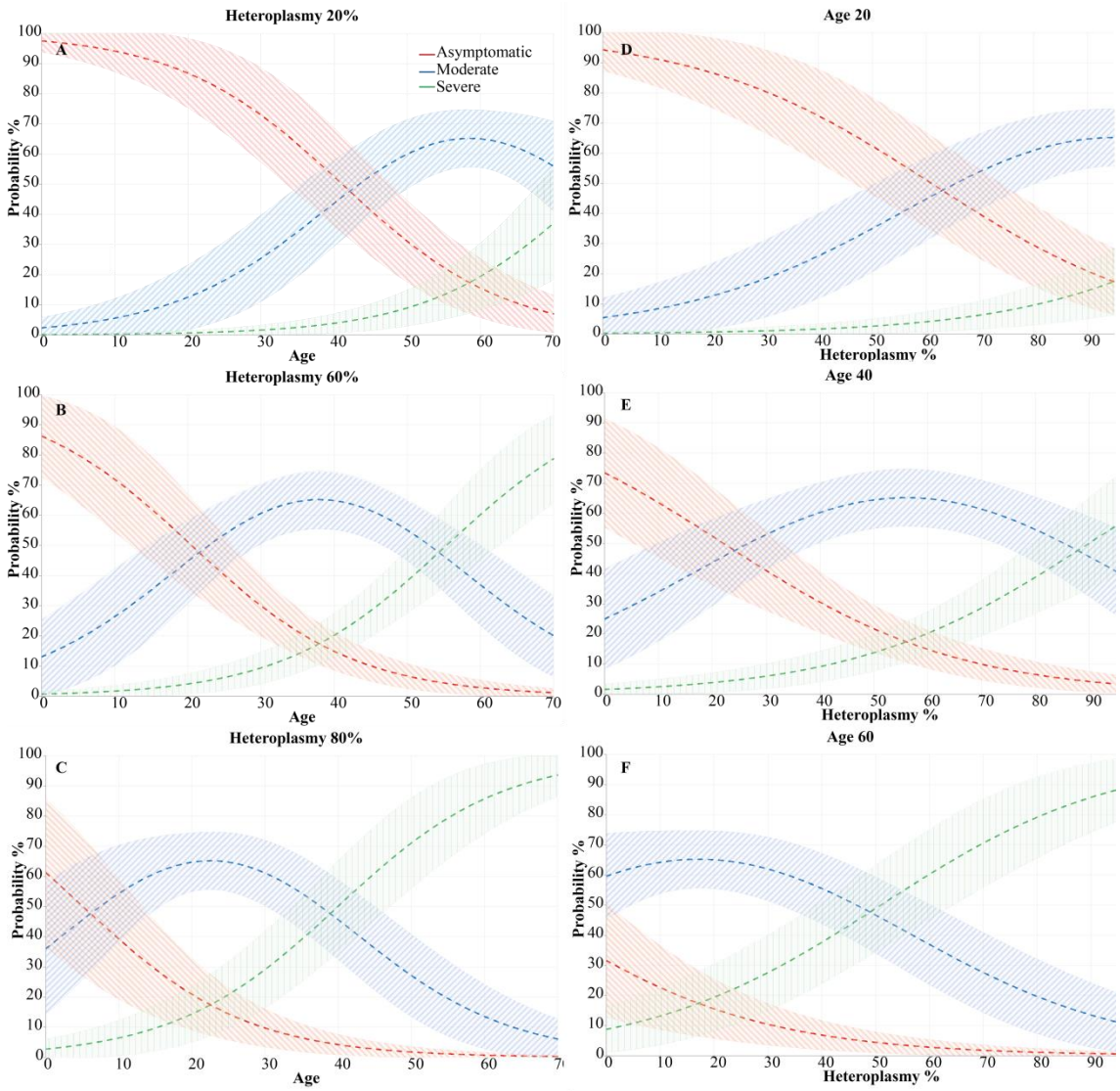


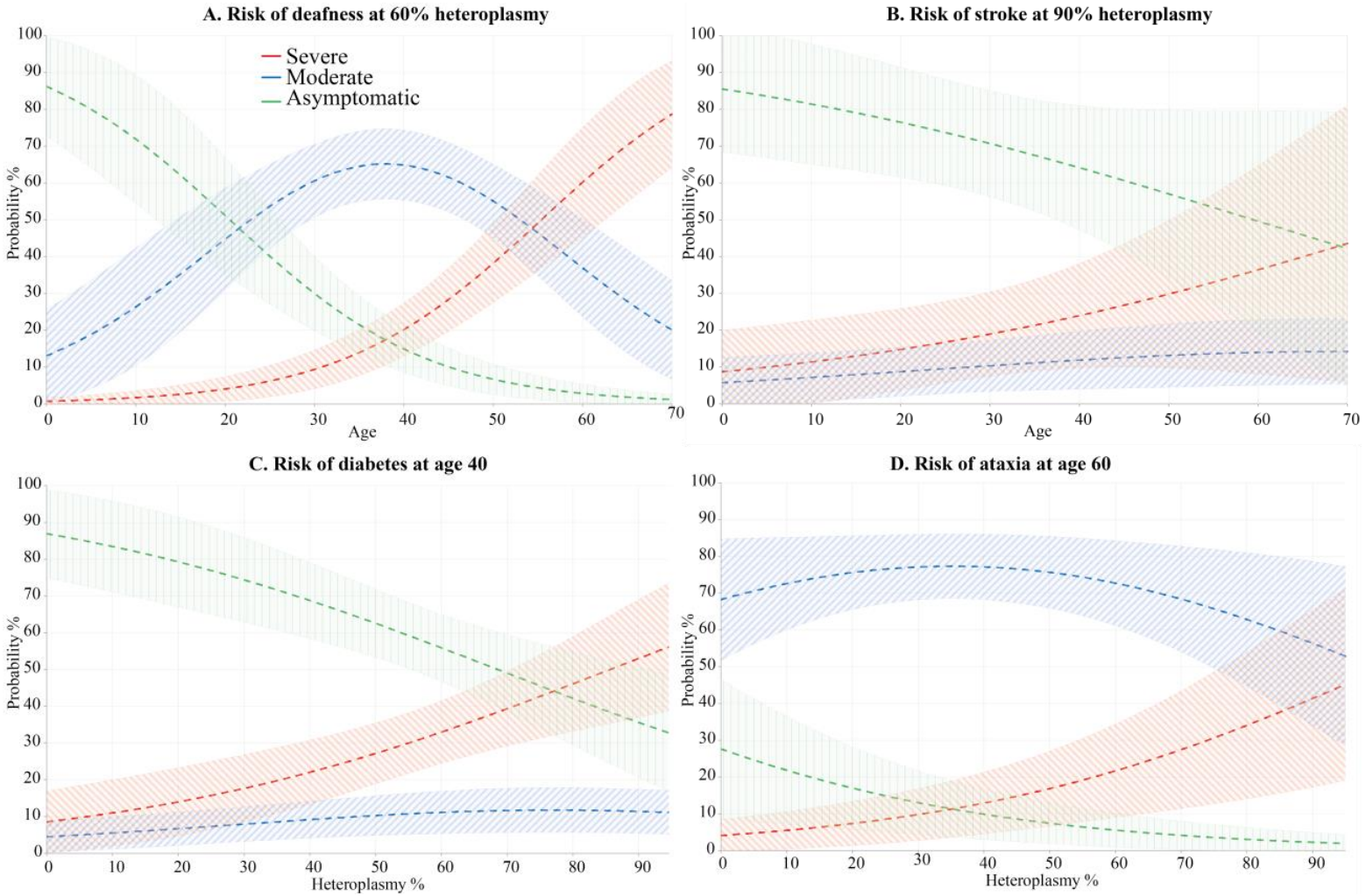
Figure 4.17 Risk of deafness and its relationship to age and heteroplasmy.

Each panel indicates the changing risk profile for one of the predictors (age or heteroplasmy) whilst the other is fixed at a determined value. **(A-C)** Changing risk of deafness with age for patients with 20%, 60%, or 80% heteroplasmy. **(D-F)** Impact of heteroplasmy on the risk of deafness for patients aged 20. At age 60, with 50% heteroplasmy, there is an approximately equal probability of moderate or severe deafness and a small probability of being asymptomatic.

4.3.5.4.4. Risk profile of principal phenotypic features

Risk profiles for a selection of other phenotypic features are shown in Figure 4.18. Each graph is shown with one predictive factor (heteroplasmy or age) fixed at an arbitrary value, for illustrative purposes. Both age and heteroplasmy are highly significant predictors for deafness ($P < 0.0001$). For stroke, heteroplasmy is predictive ($P=0.0044$) but age is not ($P = 0.1078$). Heteroplasmy is a significant predictor of cerebellar ataxia ($P = 0.0003$) and diabetes ($P = 0.0026$), and age is a highly significant predictor of both ($P < 0.0001$).

Figure 4.18 Risk profiles for four key phenotypic features of the m.3243A>G mutation.



Predicted probability of being moderately affected (score 1-3), severely affected (4-5), or asymptomatic (0). **(A)** Risk of deafness with increasing age at 60% heteroplasmy. **(B)** Risk of stroke with age at 90% heteroplasmy. The probabilities have wide confidence intervals and change little with age compared to features such as deafness, illustrating that risk of stroke is not well predicted by age ($P = 0.1078$). **(C)** Risk of diabetes at age 40 by heteroplasmy. The majority of patients are either asymptomatic or severely affected, illustrated by the uniformly low risk of moderate disease. Risk increases with increasing heteroplasmy ($P < 0.0001$). **(D)** Risk of ataxia with by heteroplasmy ($P < 0.0001$). By the age of 60, patients are likely to be moderately or severely ataxic.

4.3.5.4.5. Interaction of Heteroplasmy and Age

To examine whether there is an interaction between heteroplasmy and age in the model (i.e. the effect of heteroplasmy being different at different ages, or the effect of age being different depending on heteroplasmy) the interaction term was introduced into the model. The parameters estimates for the interaction term in the model are depicted in Figure 4.19. Out of twenty-eight features the interaction was only significant for diabetes (P = 0.0025) and cognitive impairment (P = 0.0357). The interaction for diabetes showed that the effect of either predictor is stronger at higher levels of the other predictor, i.e. a multiplicative effect. For cognition, the effect was reversed, i.e. heteroplasmy is a stronger predictor at a young age than at an older age.

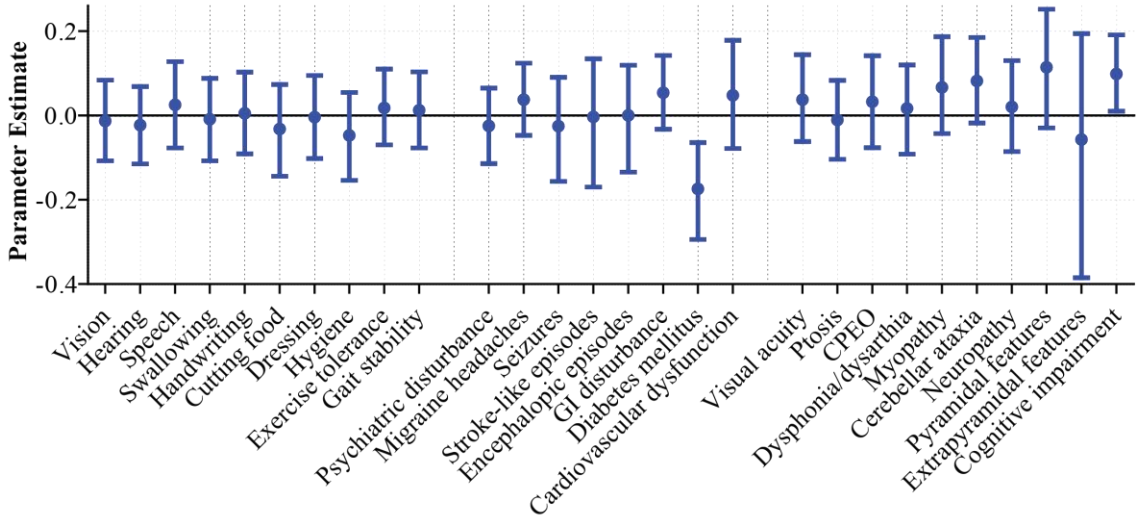


Figure 4.19 Age/heteroplasmy interaction in the logistic regression model.

Parameter estimate and 95% confidence interval are shown for each NMDAS feature. A confidence interval crossing zero is not statistically significant. Only diabetes (P = 0.0025) and cognitive impairment (P = 0.0357) have the interaction as a significant term. The negative estimate for diabetes indicates that the effect of heteroplasmy is increased as age increases, and similarly the effect of age is more profound at higher heteroplasmy. The positive estimate for cognitive impairment indicates that the effect of heteroplasmy is lessened as age increases.

4.3.5.5. Individual phenotypic features discussion

It has previously been suggested that heteroplasmy is a critical determinant of disease phenotype and not age (Chinnery *et al.*, 1997). However, comparing standardised coefficients of regression age is a better predictor for 17 out of the 28 measures reported by the NMDAS. This is an important characterisation of the disease progression in

m.3243A>G that has not been comprehensively considered in the literature of m.3243A>G disease to date.

There are several interesting observations to be made from the modelling of specific phenotypic features. Firstly, certain critical features associated with m.3243A>G, including stroke-like episodes, seizures, and migraine, are not significantly associated with age, but are predicted by the urine heteroplasmy level. For the former two symptoms the lack of association with age is apparent from the graphs of symptoms by age group in Figure 4.3; both phenotypic features are present throughout the age range of patients. Migraine, on the other hand, is negatively associated with age, though non-significantly; however, this concurs with reports in the literature of declining susceptibility to migraine with age (Dahlof *et al.*, 2009). Secondly, deafness and ataxia are the most highly correlated with age and heteroplasmy, suggesting that these features are particularly degenerative in nature. Thirdly, though gastro-intestinal dysfunction is prevalent, the severity is predicted neither by age nor heteroplasmy. The reasons for this are unclear.

Two previous studies have reported that some features of m.3243A>G disease are negatively correlated with heteroplasmy, namely myopathy, CPEO, and deafness in a study by our own group (Chinnery *et al.*, 1997), and deafness and diabetes in another (Liu *et al.*, 2013). In this study, I find to the contrary that these features are all directly correlated with heteroplasmy. However the prior studies did not incorporate age as a predictive factor, which was indeed acknowledged as a concern by both studies and I believe to be a critical omission, as demonstrated by the results from multilevel logistic regression models with only one predictor (age or urine heteroplasmy). For example, if I exclude age as a predictor in the modelling, heteroplasmy is not predictive of many common phenotypic features including diabetes, deafness, and cerebellar ataxia. The apparently paradoxical findings of these two previous studies clearly indicate the necessity of taking into account confounding factors such as age.

The motivation for examining the interaction between heteroplasmy and age is that it might be expected that some features would show a different relationship with heteroplasmy for different age groups, or that the progressiveness of the features in time would be modulated by heteroplasmy. The results from this section showed that the interaction is only significant in two features, diabetes mellitus and cognitive impairment. However, the most useful finding of this analysis is that no highly

significant interactions were found, which may have undermined the findings of the model without interaction terms.

Deafness is the most prominent symptom in this cohort; although only around 12% of patients suffer from complete deafness, 75% of patients have some form of hearing loss. It is also the symptom most predictable using heteroplasmy and age. It should be observed that prevalence does not go hand in hand with predictability; gastro-intestinal disturbance is almost as prevalent yet not predictable by either age or heteroplasmy. Deafness is so common in m.3243A>G patients that it has been suggested as a critical diagnostic criteria for identifying carriers of the mutation indeed, in a study of 1482 patients with post-lingual hearing loss, over 1% of patients were found to carry the m.3243A>G mutation (Iwanicka-Pronicka *et al.*, 2012). Another study found that the proportion was even higher (7.8%) in those with known sensorineural hearing loss (Majamaa *et al.*, 1998). The progression of deafness in m.3243A>G patients has been characterised as abrupt and stepwise usually occurring with encephalopathic or stroke-like episodes (Sue *et al.*, 1998a; Chinnery *et al.*, 2000a). However, the prevalence of deafness in this cohort far exceeds that of encephalopathic or stroke-like episodes, and thus most of the hearing loss, though it may be stepwise, cannot be so associated. The correlation between hearing loss and muscle heteroplasmy has previously been reported (Chinnery *et al.*, 2000a), and this study has augmented that finding by incorporating age as a co-predictor. Deafness was amongst the symptoms with high R^2 inflation factors and demonstrated a dramatic reduction in P-value with multiple regression, indicating that the population level inter-correlation of age and heteroplasmy masks the strong association of both of these predictors with symptom severity.

Sixty-three percent of patients are ataxic to some degree, and 39% are at least moderately affected. The prevalence of ataxia in mitochondrial disease has been noted in the literature (Lax *et al.*, 2012), and though it is generally acknowledged as a progressive symptom, this has not been fully characterised. The modelling documented here demonstrates that ataxia is both predictable by heteroplasmy and age. Indeed, the pseudo-R^2 value indicates that it is the most clearly progressive symptom of all in this cohort.

Though gastro-intestinal disturbance stands with deafness as one of the two most common phenotypic features seen in patients with m.3243A>G, it is the only feature not predicted by either age or heteroplasmy. The reasons for this are unclear. Whilst gastro-

intestinal disturbance is common within the general population, the incidence in our patients is striking. Previous studies have demonstrated respiratory chain deficiency of the bowel smooth muscle which could explain the symptoms (Betts *et al.*, 2008), and a recent study also found a previously unreported high incidence of coeliac disease in children with mtDNA mutations (Mazzaccara *et al.*, 2012). However, it remains unclear why the dysfunction would not correlate better with age and heteroplasmy. It is clear from the phenotypic characterisation of the cohort seen in Figure 4.2 that NMDAS scores of 5 for GI disturbance are very rare indeed, and scores of 4 are also rare. Thus either high levels of GI disturbance are rare in the cohort, or this analysis has highlighted an issue with the NMDAS data collection which needs to be addressed; patients with severe disease who have been hospitalised may not be being accurately assessed for criteria such as GI disturbance due to difficulty in performing the NMDAS assessment fully (Dr Andy Schaefer, personal communication).

The methodology employed in this section requires some discussion. To avoid the need to deal with repeated measures and the implicit correlation between repeated scores for individual patients, I have taken a single summary score for each patient as described in section 4.3.5.3. However, a necessary shortcoming of this approach is that the full potential of the dataset is not being exploited; nor is it longitudinal analysis. Longitudinal analysis using techniques analogous to repeated measures mixed modelling would support inferences about individual patients and their expected progression by the incorporation of random effects into the model; repeated measures multiple logistic regression with random effects would be a desirable approach to modelling this data in order to provide prognostic information for individual patients based on their phenotypic progression to date. However, current analytical tools for conducting such analysis are limited. Though SAS does support incorporation of random effects into logistic regression using PROC GLIMMIX, and modelling of the covariance structure in repeated measures from the same patient, this analysis is limited to dichotomous data. Thus progression from asymptomatic to moderate to severe disease cannot be conducted. An additional and more serious issue currently is that for most patients there is no noticeable progression in many of the clinical symptoms over the period of assessment. As more NMDAS data is collected this form of modelling will become more feasible.

4.3.6. Disease progression discussion

This section has used a variety of approaches to examine disease progression in a large cohort of patients with m.3243A>G, both examining the progression in total disease burden, and the risk of developing specific phenotypic features according to age and heteroplasmy.

Taken together, these analyses have shown that heteroplasmy has some utility as predictor of both overall disease burden and most individual phenotypic features, and that multiple regression with age is vitally important in order to fully recognise the prognostic potential. However, there is a strong imperative to research a more clinically relevant and easily measurable predictor of disease progression.

The basic challenge facing understanding of disease associated with the m.3243A>G mutation is identifying further predictors that will transform unexplained variance in phenotype and disease progression into explained variance. Understanding the links between genetics (both nuclear and mitochondrial), epigenetic modifiers and the resulting clinical phenotype is critical future work.

There are several interesting lines of enquiry in this regard. Firstly, a recent study has indicated that wild-type mtDNA copy number, rather than heteroplasmy, may be a more useful prognostic indicator than heteroplasmy (Liu *et al.*, 2013). Investigation of mtDNA copy number is therefore a priority in furthering this modelling work.

Secondly, it has been demonstrated in a homoplasmic point mutation in the mitochondrial tRNA isoleucine gene (*MT-TI*) that levels of the cognate tRNA synthetase modulate the penetrance of the mutation and the resulting clinical phenotype, in this case hypertrophic cardiomyopathy (Perli *et al.*, 2012). To my knowledge, there has been no study of patients carrying the m.3243A>G mutation to identify if tRNA synthetase levels correlate with clinical phenotype. Indeed, the leucyl tRNA synthetase is particularly interesting as it has been shown to act as an amino acid sensor modulating the mTORC1 pathway that regulates protein translation, cell size, and autophagy (Han *et al.*, 2012), and overexpression of the synthetase corrects the m.3243A>G phenotype in cybrid cells (Li and Guan, 2010). Thus there are several motivations for understanding how tRNA synthetases modulate disease burden and phenotypic expression in m.3243A>G patients.

Thirdly, it has been reported that a critical difference between MELAS and other phenotypic presentations, such as CPEO, is that MELAS patients have a relatively

homogeneous distribution of mutant mtDNA in muscle fibres, whereas patients with a CPEO-like phenotype appear to have focal accumulation of mutant mtDNA in selected fibres (Moraes *et al.*, 1993; Petruzzella *et al.*, 1994). It is speculated that this distinction may also be apparent in clinically affected tissues of the CNS though this is yet to be reported on. However, it should be noted that in these studies patient age is a fundamental confound, since the accumulation of defective mtDNA in fibres is likely to be a time-dependent process, and both studies had a significantly younger MELAS cohort than the comparator cohort.

The statistical modelling I have employed is novel in two regards; firstly, in comprehensive longitudinal modelling of total disease burden; and secondly, the predictive modelling of individual phenotypic features, using both age and heteroplasmy as predictors. Though previous studies have used heteroplasmy as a predictor of phenotype (Chinnery *et al.*, 1997; de Laat *et al.*, 2012; Liu *et al.*, 2013) this approach is limited and risks confounding any findings by not accounting for age. Indeed, multiple regression using both factors simultaneously is critical, since at the population level urine heteroplasmy and age are negatively correlated, as previously reported (de Laat *et al.*, 2012) and confirmed in this cohort. This significant Pearson's correlation is the cause of the increased explanatory power of heteroplasmy and age together as compared to either factor as singular predictors.

There are some limitations in this study which need to be acknowledged. Firstly, I have used urine heteroplasmy throughout; though it has been shown to correlate well with clinically affected tissues (Blackwood *et al.*, 2010), a more invasive measure of heteroplasmy, for example skeletal muscle, may reduce the unexplained variance and improve the modelling. This will be discussed further in the next section on m.3243A>G heteroplasmy.

Secondly, I have not considered in the model the baseline level of disease found in the normal ageing population, which would be necessary in order to ascertain the degree of disease progression attributable to the mtDNA mutation alone. However, from a clinical care perspective the overall disease burden for patients is the more relevant quantity.

Thirdly, as previously discussed in the section on individual phenotypic features, longitudinal modelling of specific features has not been conducted, and it is anticipated this would be valuable information for clinicians and patients alike.

In conclusion, I have produced two model systems for understanding disease burden and progression in patients with the m.3243A>G mutation, both utilising age and heteroplasmy as predictive factors; a longitudinal model of overall disease burden progression, and models to predict the severity of specific phenotypic features. Both these models provide critical information to clinicians for the care and management of patients with the m.3243A>G mutation.

4.4. Heteroplasmy in m.3243A>G

4.4.1. Introduction

Heteroplasmy levels in m.3243A>G remain a rich area for research, and though much has been studied and reported there are still large gaps in knowledge.

Since the earliest reports of blood heteroplasmy decline in the m.3243A>G mutation (Poulton and Morten, 1993; Hammans *et al.*, 1995), the decline has been repeatedly studied. An earlier study of 18 patients reported an average linear decline of 0.69 percentage points per year (t Hart *et al.*, 1996); Rahman *et al.*, in a study of six patients, compared blood from Guthrie cards taken at birth to blood taken at MELAS disease diagnosis and found a linear decline of an average 1.4 percentage points per year (Rahman *et al.*, 2001); Pyle *et al.* reported a linear decrease of 0.6 percentage points per year in a longitudinal study of 11 patients (Pyle *et al.*, 2007); Mehrazin *et al.* quantified the loss as 0.53 percentage points per year in MELAS patients and 0.22 percentage points per year in carrier relatives in a study of 34 patients (Mehrazin *et al.*, 2009). Rajasimha *et al.*, in a departure from the other studies, proposed that the decline in blood was exponential rather than linear, provided a coherent model based on progressive loss of haematopoietic stem cells which had accumulated high levels of mutant mtDNA by random genetic drift, and validated their model by comparison with experimental data (Rajasimha *et al.*, 2008). They quantified the decline as an approximate compound 2% loss per year. However, it should be noted that some studies that have examined longitudinal changes in blood heteroplasmy did not find the change significant (Kaufmann *et al.*, 2011), though this is generally attributed to small sample size or insufficient time between measures.

Heteroplasmy changes in other tissues are less well understood. It was shown early on that intra-patient heteroplasmy levels in many post-mitotic tissues are similar, for instance skeletal muscle heteroplasmy is thought to be broadly representative of the level in the neurons within the CNS (Ciafaloni *et al.*, 1991; Macmillan *et al.*, 1993). Several studies have reported a negative correlation between heteroplasmy and age in diverse samples including urine, skeletal muscle, and buccal mucosa (Frederiksen *et al.*, 2006; Kaufmann *et al.*, 2011; de Laat *et al.*, 2012), though as none of the studies were longitudinal it was not possible to draw firm conclusions about decline with age.

However, there are wide variations in heteroplasmy levels in different tissues. Shiraiwa *et al.*, for instance, found the lowest mutation level in the spleen (Shiraiwa *et al.*, 1993).

Interestingly, there are reported consistent differences in the average heteroplasmy levels between certain tissues. Chinnery *et al.* reported that the distribution of mutant mtDNA was non-random, and that the mutation load in tissues correlated with the cell turnover rate in each tissue; levels in muscle were highest, followed by hair follicles, buccal mucosa, and lowest in blood (Chinnery *et al.*, 1999). This hierarchy is supported by data from several other research groups, many of whom also added heteroplasmy measured in urine, which is generally found at a level similar to that found in muscle (Shanske *et al.*, 2004; Frederiksen *et al.*, 2006; de Laat *et al.*, 2012). De Laat, for instance, quantified that heteroplasmy levels in urine were 11 percentage points higher than in buccal mucosa and 23 percentage points higher than blood.

Non-random segregation of mutant mtDNA was consistent with the finding that foetal heteroplasmy levels are strikingly homogeneous (Matthews *et al.*, 1994), and strong evidence against the hypothesis that heterogeneous heteroplasmy levels in tissues were caused by random mitotic segregation in early embryogenesis (Huang *et al.*, 1996). It was also reported that though heteroplasmy levels vary from tissue to tissue there are no differences observable between the average levels of tissues originating from the three germ layers (Frederiksen *et al.*, 2006), further evidence against early embryonic random mitotic segregation, and evidence that variations in heteroplasmy are mainly due to selection pressures on either wild-type or mutant DNA during life.

The significant loss of the mutation in blood contrasts with other rapidly dividing tissues which do not exhibit such decline, such as hair follicles (Sue *et al.*, 1998b) and the urinary epithelium (Blackwood *et al.*, 2010), though has been speculated as due to the low energy requirement of such cells, and it is suggested that in rapidly dividing tissues random genetic drift accounts for alterations in heteroplasmy unless a severe respiratory deficiency exerts pressure and results in active selection of wild-type mtDNA (Rahman *et al.*, 2001).

Interestingly, variation between hair follicles from the same individual has been shown to be very large (Shanske *et al.*, 2004); this was attributed to the fact that individual hair follicles develop from one or a small number of stem cells in each follicle bulb (Ghazizadeh and Taichman, 2001) and thus random segregation will result in such mosaicism whilst maintaining a uniform average heteroplasmy level.

4.4.2. Aims

In this section, I examine the Newcastle cohort data to probe the inter-relationships between the various measures of heteroplasmy available for this cohort (urine, blood, and skeletal muscle), and also age.

I have two major aims. The first is to examine urine and muscle heteroplasmy for evidence of longitudinal change and to characterise any dynamic shift. The second is to examine the decline of blood heteroplasmy in the Newcastle cohort and to validate and characterise the proposed exponential decline model (Rajasimha *et al.*, 2008).

4.4.3. Methods

4.4.3.1. Exponential decline of blood heteroplasmy (longitudinal)

This analysis was restricted to those patients with multiple measurements of heteroplasmy more than two years apart.

Exponential decline, defined by a rate parameter here called *Rate*, can be expressed as shown in Equation 4.1.

$$Final\ Het\ (\%) = Initial\ Het\ (\%) \times Rate^{Years\ between\ measurements}$$

Equation 4.1 Exponential decline.

For convenience I define the *decline factor* to be *100% - rate*. Thus a decline factor of 5% indicates a compound decline of 5% (the result of multiplication by 0.95) each year. A negative decline factor indicates an increase in heteroplasmy.

The decline factor is hence calculated as shown in Equation 4.2; e.g. for a patient with a decline factor of 5%, if the starting heteroplasmy was 50%, after one year the heteroplasmy would decline to $50\% \times 0.95 = 47.5\%$, after two years $47.5\% \times 0.95 = 45.125\%$, and so on for each year of decline.

$$Decline\ factor\ (\%) = 100\% - e^{\frac{\ln(Initial\ Ht) - \ln(Final\ Het)}{Years\ between\ measurements}}$$

Equation 4.2 Decline factor calculation.

4.4.3.2. Exponential decline of blood heteroplasmy (using urine heteroplasmy)

A second method I chose to investigate the decline of blood heteroplasmy was to use muscle, urine heteroplasmy, or another putatively static heteroplasmy measure as a

nominative starting heteroplasmy level at birth. As I have only a limited number of muscle heteroplasmy levels recorded, I used urine heteroplasmy as the nominative level. Where multiple urine heteroplasmy records are available, I take the mean value.

Similarly to section 4.4.3.1, I define an exponential decline factor, but this time, I use age plus 9 months as the time period, as I am assuming that the urine heteroplasmy level is indicative of blood heteroplasmy at conception and that any selective pressure is exerted from that point.

4.4.4. Results

4.4.4.1. Cohort

There are 192 patients with m.3243A>G in the Newcastle cohort in total, excluding two patients who have had renal transplants and are excluded from heteroplasmy investigations for this reason. Figure 4.20 illustrates the number of patients available with the respective heteroplasmy levels. The full data can be found in Appendix III.

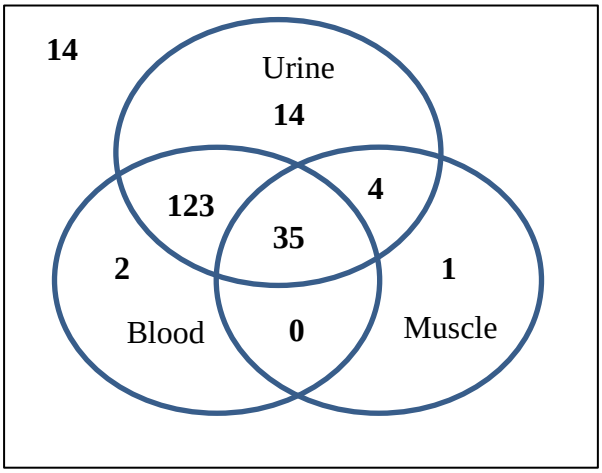


Figure 4.20 Number of patients with heteroplasmy data available.
Venn diagram illustrating the number of patients with urine, blood, and skeletal muscle heteroplasmy. Fourteen patients have no measure of heteroplasmy; 35 have all three measures. For each analysis. I use the maximum number of patients available with the required data.

4.4.4.2. Relationship between heteroplasmy and age

4.4.4.2.1. Urine

The regression of urine heteroplasmy and age is shown in Figure 4.21A. Although there is a clear correlation, and the relationship is reasonably linear, the distribution is asymmetric. However, Box-Cox identifies the identity function (no transformation) as optimal.

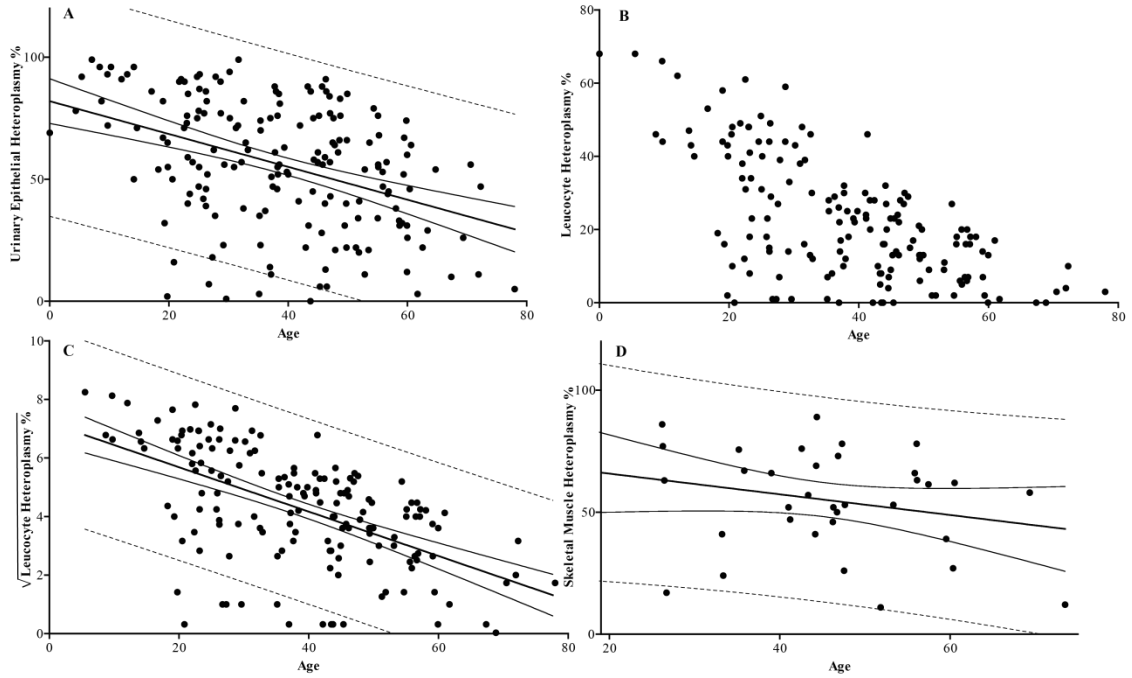


Figure 4.21 Heteroplasmy and age. Linear regression with 95% confidence and prediction intervals.

(A) Though there is a clear correlation between urine heteroplasmy and age, the relationship is not quite linear; the line is reasonably well fitting, but the residuals are asymmetric and do not follow a Gaussian distribution (P = 0.0004 Shapiro-Wilk). N = 175, R = -0.41, P < 0.0001. **(B)** Non-linear relationship between blood heteroplasmy and age. Box-Cox identifies the square root as the optimal transform. **(C)** Once transformed the relationship between blood and age is more clearly linear (R = -0.49, P < 0.0001). However the residuals are non-Gaussian and do not pass normality tests (P < 0.0001, Shapiro-Wilk). N = 159 **(D)** The relationship between SKM heteroplasmy and age is not significant though there is a negative trend (P = 0.0664, R = -0.30, N=39).

4.4.4.2.2. Blood

The relationship of blood heteroplasmy and age is shown in Figure 4.21B. Age and blood are clearly correlated (P < 0.0001); Box-Cox identifies the square root transform as optimal, and the linear regression of this is shown in Figure 4.21C. Though this improves the fit, the residuals still remain non-normal.

4.4.4.2.3. Muscle

Figure 4.21D shows the regression of SKM heteroplasmy against age. Though there is a declining trend with age, it is not statistically significant in this cohort (N = 39, P = 0.0664).

4.4.4.3. Urine and blood heteroplasmy levels are linearly related

On inspection, the relationship between urine and blood heteroplasmy levels is clearly non-linear (data not shown). Box-Cox identifies the appropriate transformation of the blood heteroplasmy as the square root. Figure 4.22 illustrates the relationship between urine heteroplasmy and the square root of the blood heteroplasmy level. The two quantities are highly significantly correlated ($P < 0.0001$, $R^2 = 66.4\%$, $N = 158$). Residual diagnostics confirm that the errors in this relationship are clearly Gaussian in distribution, indicating that the relationship is well formed.

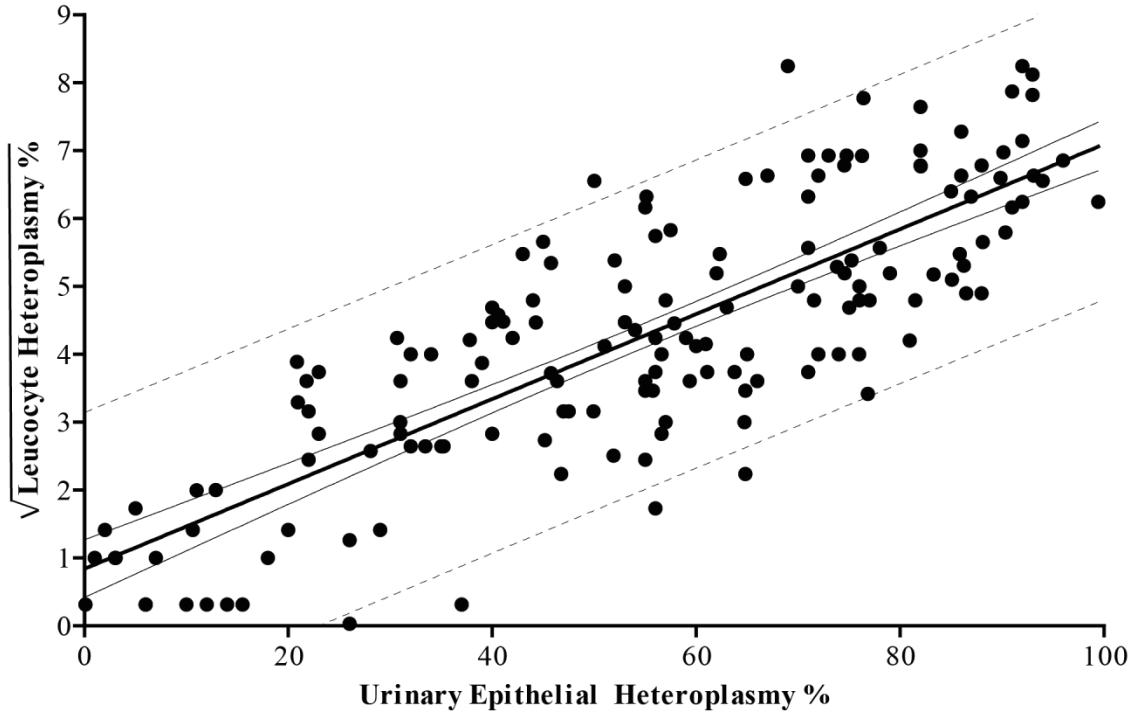


Figure 4.22 Urine heteroplasmy is linearly correlated with the square root of blood heteroplasmy.
The regression line and both 95% confidence and prediction intervals for the linear correlation are shown. The square root relationship is identified by Box-Cox as the appropriate transformation. Residuals are Gaussian in distribution, indicating that the relationship is well formed. $N = 158$, $R^2 = 66.1\%$, $P < 0.0001$.

4.4.4.4. Age at biopsy improves the linear relationship between urine and blood heteroplasmy

As blood heteroplasmy is thought to decline with age, I also examined the relationship between urine heteroplasmy, blood heteroplasmy, and age. This analysis also uses the square root transformation of blood heteroplasmy.

Figure 4.23 illustrates the outcome of the multiple linear regression. Both urine heteroplasmy ($B = 0.69$, $P < 0.0001$) and age ($B = -0.30$, $P < 0.0001$) are highly significant predictors of blood heteroplasmy ($R^2 = 74.0\%$). The interaction of age and urine heteroplasmy is also highly significant ($P < 0.0001$) yielding standardised

parameters for urine heteroplasmy (B = 0.69, P < 0.0001), age (B = -0.33, P < 0.0001) and the interaction (B = -0.15, P < 0.0001) (R^2 = 76.2%).

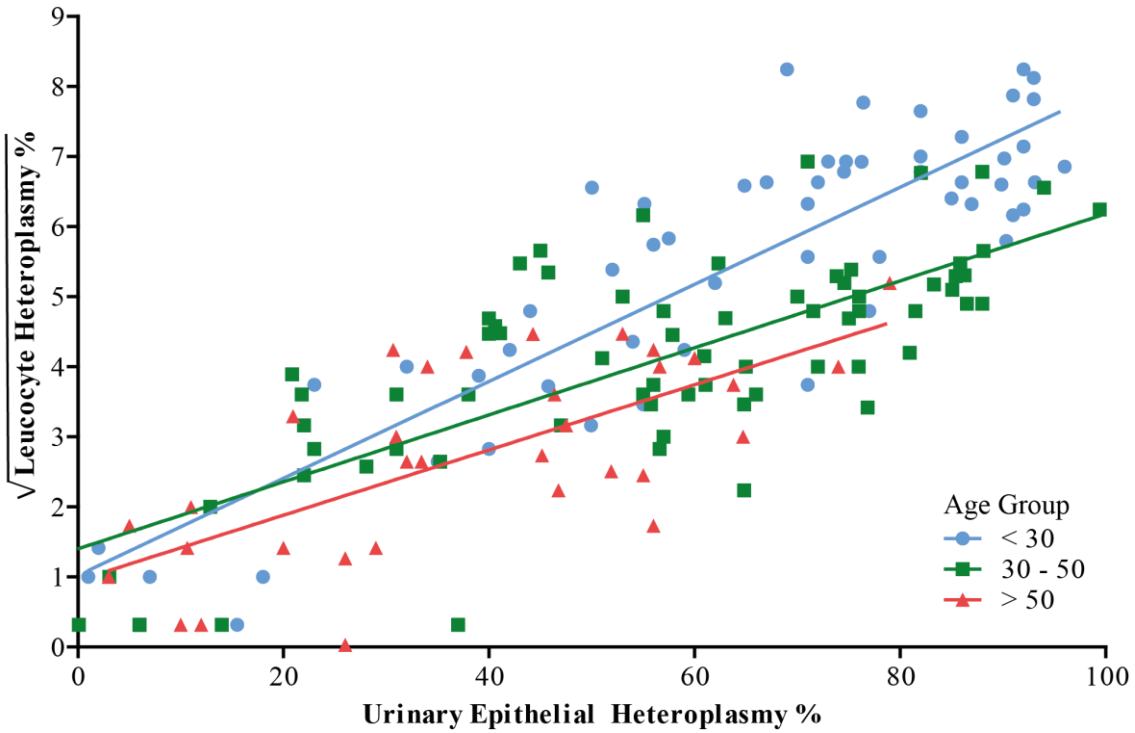


Figure 4.23 Multiple regression with age at biopsy improves the correlation between urine and blood heteroplasmy.

The linear regression lines for three age groups are depicted. Adjusted R^2 increases to 76.2%, from 66.4% for the model without age (Figure 4.22). Both age and urine heteroplasmy are highly significant predictors (P < 0.0001). The gradient of the line relating urine and blood heteroplasmy is age dependent (P < 0.0001). N = 158. P-values are for age as a continuous predictor; graphs show age as a categorical variable for illustrative purposes only.

4.4.4.5. Urine and SKM heteroplasmy levels are linearly related

Figure 4.24 illustrates the relationship between urine and muscle heteroplasmy. Where repeated measures of urine heteroplasmy are available the mean of the two values is used. The two measures are highly significantly linearly correlated (P = 0.0006, r = 0.53). The regression line defines the relationship between the two measures of heteroplasmy shown in Equation 4.3.

$$Muscle\ heteroplasmy\ (\%) = 28.07 + 0.53 \times Urine\ heteroplasmy\ (\%)$$

Equation 4.3 Relationship between muscle and urine heteroplasmy.

Age is not a significant factor in this relationship (P = 0.8535). The 95% confidence interval for the gradient in this equation is (0.24, 0.81).

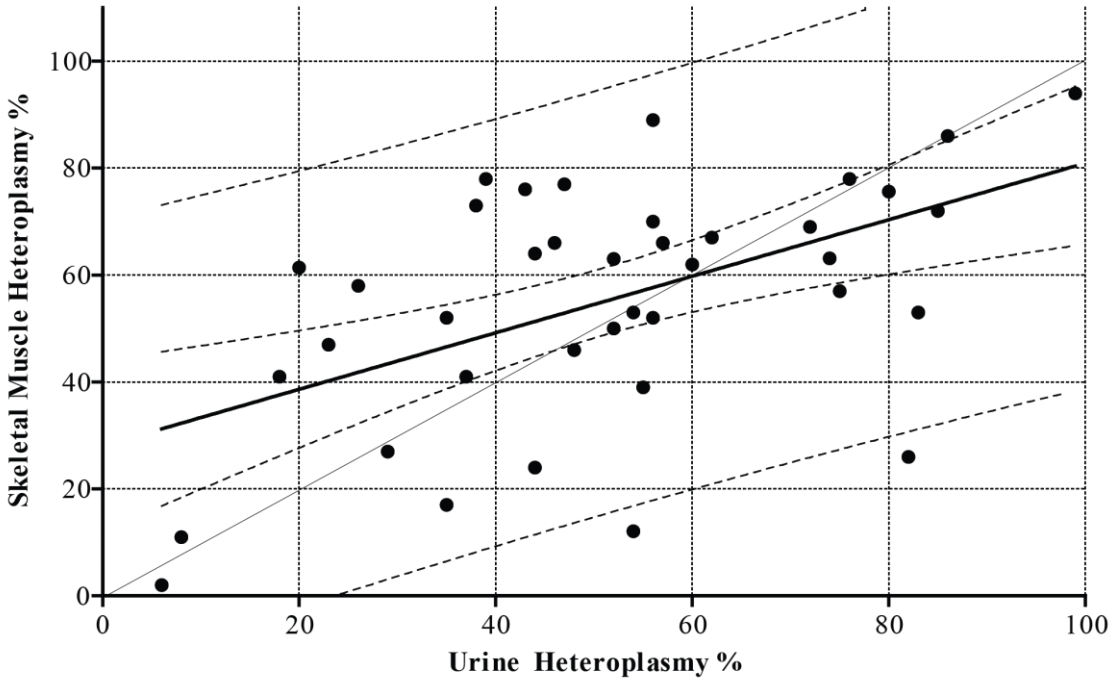


Figure 4.24 Urine and muscle heteroplasmy levels are linearly correlated.

Regression line with 95% confidence and prediction intervals are shown. r = 0.53, P = 0.0006, N = 39. The gradient (0.53) is significantly non-zero (95% CI 0.24, 0.81). The line y = x is also drawn on for comparison, as this would be the regression line describing equality of urine and muscle heteroplasmy values.

4.4.4.6. Longitudinal urine heteroplasmy analysis

For the patients with at least two sequential urine heteroplasmy measurements (N = 79) the percentage point change in heteroplasmy was calculated for the points furthest apart in time. As shown in Figure 4.25A, the changes in heteroplasmy were not significantly different from zero (P = 0.8487). The mean change is 0.26% (95% CI -2.4% - 2.9%). The sample standard deviation is quantified as 11.8%.

The changes in heteroplasmy were also regressed against several other quantities to investigate any trends, as depicted in Figure 4.25B-D. Heteroplasmy change was not significantly correlated with age at sampling (P = 0.1791, r = 0.15), though there was a non-significant trend towards increasing heteroplasmy with age. Nor was it correlated with the years between measurements (P = 0.6126, r = -0.06). Only the initial heteroplasmy level showed a significant correlation with the change over time with a trend towards falling heteroplasmy at high initial levels (P = 0.0035, r = -0.33).

The White test for heteroscedacity indicated that the variance in the heteroplasmy change was not correlated with years between measurement (P = 0.3652) or age (P = 0.6569).

Multiple regression with age and initial heteroplasmy finds that the significance of age as a predictor of heteroplasmy changes from P = 0.01791 to 0.9830, whilst initial heteroplasmy remains significant (P = 0.01) (R^2 = 0.11).

The standard deviation (11.8%) of the urine heteroplasmy was analysed using bootstrapping to establish a confidence interval for the estimate. This generated a 95% confidence interval for the standard deviation of (9.62%, 14.01%).

4.4.4.7. Longitudinal blood heteroplasmy decline is exponential

The decline factor is calculated as explained in section 4.4.3.1. Twenty-seven patients in the cohort had serial blood heteroplasmy records at least two years apart. The results are shown in Figure 4.26. Panel A shows that the decline factor is independent of age; panel B shows that the factor is not independent of the initial heteroplasmy level.

As the distribution of the decline factor is non-Gaussian I use non-parametric testing to assess whether the decline factor is significantly non-zero. The median decline factor is 1.30% (95% CI 0.00% to 2.53%). The result is not significantly non-zero (P = 0.0773).

4.4.4.8. Urine as a nominative initial blood heteroplasmy level provides a consistent blood heteroplasmy decline level

The concept of using urine as a nominative initial starting heteroplasmy to investigate blood heteroplasmy decline is considered in Figure 4.27. There are weak and insignificant trends with both the decline period and initial heteroplasmy level. Mean decline is 2.99% per year.

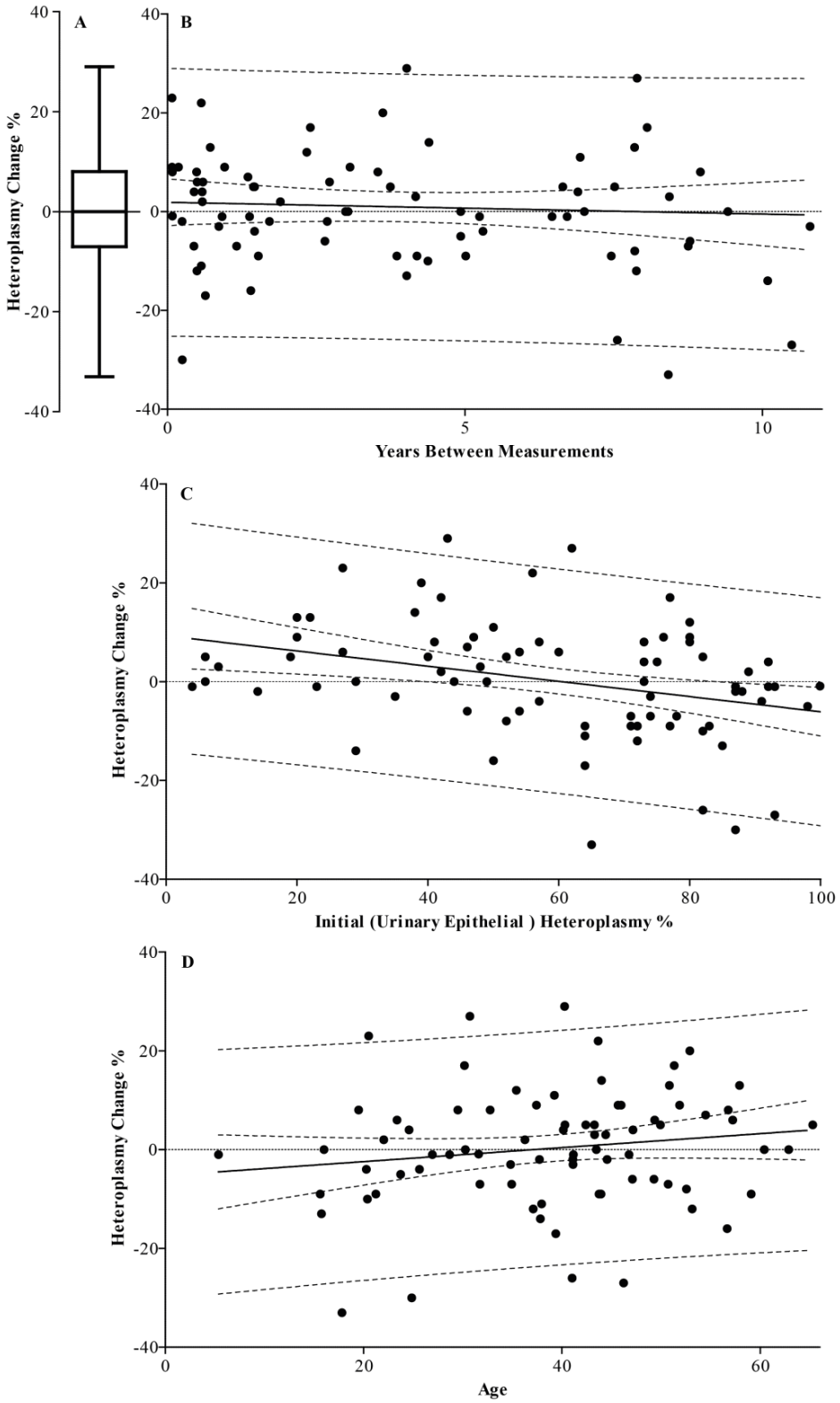


Figure 4.25 Longitudinal changes in urine heteroplasmy over time.

For each patient with sequential urine measurements (N = 79), the change in heteroplasmy was calculated and compared to various potential correlates. Regression graphs depict regression lines with 95% confidence and prediction intervals**. (A)** Change in heteroplasmy is not significantly different from zero (P = 0.8487). **(B)** Change in heteroplasmy over time is not correlated with time between measurements (P = 0.6126, r = - 0.06). **(C)** Change in heteroplasmy is correlated with the initial level of heteroplasmy (P = 0.0035, r = -0.33) **(D)** Change in heteroplasmy over time in not significantly correlated with age (P = 0.1791, r = 0.15).

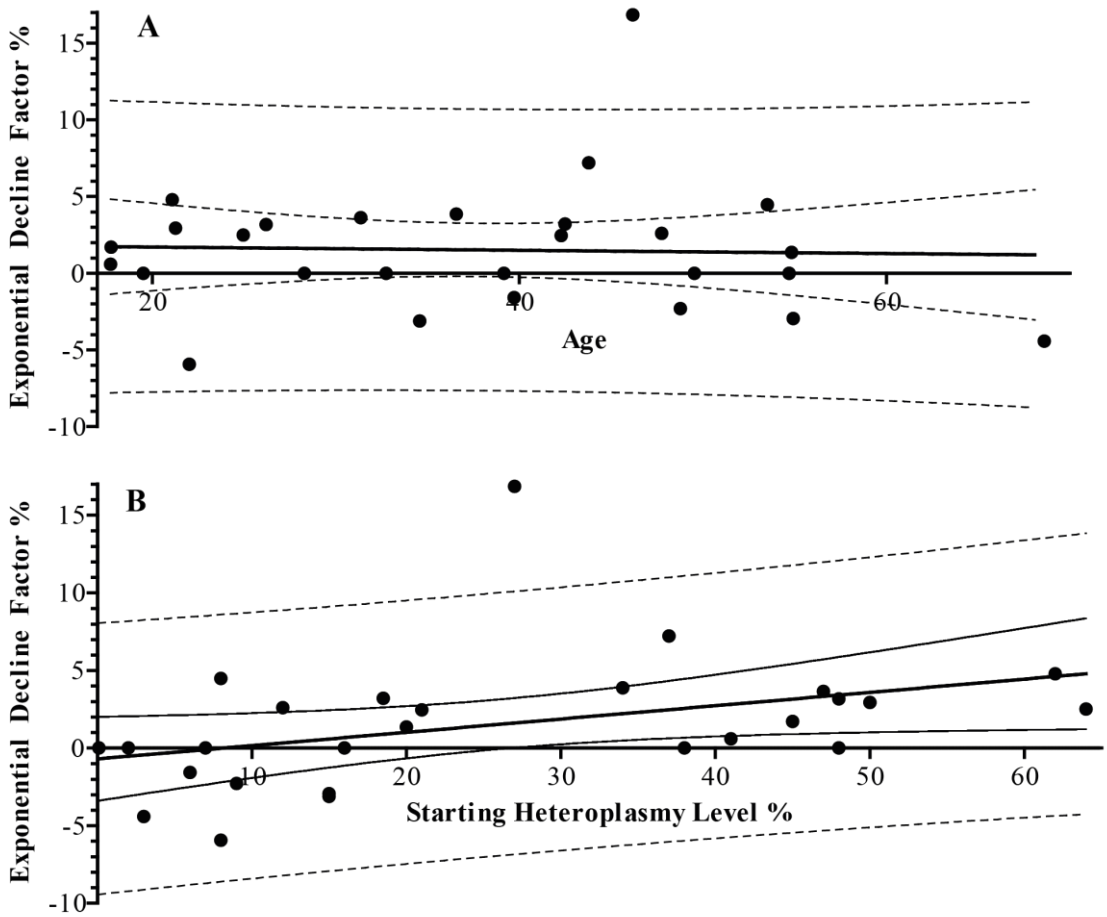


Figure 4.26 Examining the exponential decline of blood heteroplasmy.

The annual rate of exponential decline of blood heteroplasmy is computed for those patients with serial blood measurements more than two years apart (N=27). Linear regression with 95% confidence and prediction intervals. A positive factor indicates decline in heteroplasmy over time, a negative factor indicates increasing heteroplasmy **(A)** The decline factor is independent of age (P = 0.8661, r = - 0.034). **(B)** The decline factor is directly correlated with the initial heteroplasmy level (P = 0.0460, r = 0.387).

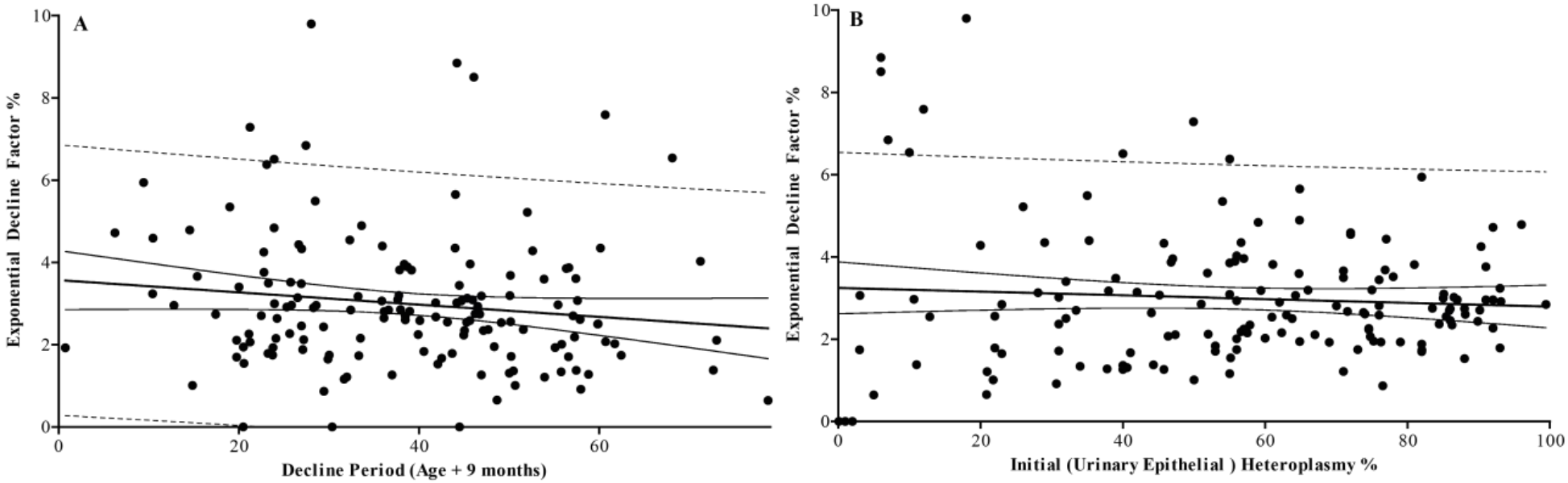


Figure 4.27 Exponential decline of blood heteroplasmy using urine heteroplasmy as nominative initial level.

The urine heteroplasmy level is taken as a nominative initial heteroplasmy level at birth, and is used to calculate an exponential decline factor for blood heteroplasmy (N = 155). Regression line with 95% confidence and prediction intervals is shown. **(A)** The relationship between the period of decline and the decline factor. There is a downward trend but it is not statistically significant (P = 0.0923, r = -0.14). **(B)** The relationship between the initial heteroplasmy level and decline factor. There is a non-significant downward trend (P = 0.3848, r = -0.07), the mean decline is 2.99% per year.

4.4.5. Discussion

4.4.5.1. Heteroplasmy and age

Several studies have reported a negative correlation of heteroplasmy with age, using samples from a variety of sources including buccal mucosa, urine, and skeletal muscle (Frederiksen *et al.*, 2006; Kaufmann *et al.*, 2011; de Laat *et al.*, 2012), and these observations are confirmed in my analysis of urine and skeletal muscle, though the correlation with skeletal muscle in the cohort is not significant. It needs to be considered that the sample size of the muscle heteroplasmy data is much lower than that of urine and blood. Though some researchers extended these observations to speculation about longitudinal changes in patients (de Laat *et al.*, 2012), this has not been studied in a large cohort to date in any sample source other than blood. However, it should be noted from the outset that a negative correlation in heteroplasmy in all tissues with age is potentially explained by the fact that older patients in the cohort are likely to have presented with disease later in life and thus be expected to have lower heteroplasmy. Thus a negative correlation may be simply a result of sampling bias.

4.4.5.2. Is urine heteroplasmy static or dynamic?

The data in section 4.4.4.6 shows that there is no trend in urine heteroplasmy levels either upwards or downwards in general. Nor were there any significant trends with the age of the patients or the period of measurement. These findings in a large sample are strongly supportive of a conclusion that urine heteroplasmy levels do not systematically increase or decrease over time. Although this has previously been speculated to be the case based on a small cohort of 11 patients (Blackwood *et al.*, 2010), this is the first comprehensive report of longitudinal change in heteroplasmy recorded from any source other than blood.

There is, however, a significant correlation with the initial heteroplasmy level; those with high heteroplasmy levels are more likely to show a decrease over time, and those with low heteroplasmy levels are more likely to show an increase. The association, though significant ($P = 0.0035$), is weak ($r = 0.33$), and drawing conclusions from this association is difficult, as such trend could easily be accounted for by floor and ceiling effects when interpreting changes in extreme (low or high) heteroplasmy levels. Multiple serial measurements in a large cohort would be needed to better understand if there are true longitudinal changes.

This observation should be considered in tandem with the relationship observed between urine and muscle heteroplasmy levels. It is interesting to note that the linear regression implies that for low heteroplasmy levels muscle heteroplasmy exceeds urine, but at high levels, muscle is lower than urine. This is a curious observation; though urine heteroplasmy levels are not systematically moving either up or down observing the cohort as a whole, there is systematic divergence of urine and muscle heteroplasmy levels at low and high heteroplasmy levels. These two observations are superficially in direct conflict; for example, a urine heteroplasmy level of 80% would tend to show decrease over time, but relative to muscle heteroplasmy it appears to be inflated. How can these observations be reconciled?

One explanation is that one or both of these observations is a statistical anomaly, perhaps derived from ascribing a linear relationship where there is non-linear relationship, which in this case is likely on account of the finite limits (0% to 100%) of heteroplasmy measurements. This can only be correctly ascertained by collecting further data, as the sample size is not large enough to draw firm conclusions. However, it should be considered that divergence of heteroplasmy levels between tissues suggests that at some point during development or later life heteroplasmy levels have dynamically shifted, though it is unclear here whether the shift has occurred in muscle, the cells found in the urinary sediment, or both. Heteroplasmy levels in muscle, as post-mitotic tissue, are expected to be relatively static, and thus the cells that are found in urinary sediment are put into the spotlight. Indeed, in section 4.6, I show that urine heteroplasmy in males is significantly higher than females despite similar blood heteroplasmy levels, which is further evidence of a potential dynamic shift in heteroplasmy in the tissues from which cells in urinary sediment originate.

A major question is the origin of mtDNA in urinary sediment. Though the cells are generally expected to be urothelial cells from the bladder lining and urinary tract (Blackwood *et al.*, 2010), this has not been specifically investigated in mitochondrial disease and m.3243A>G in particular. Renal dysfunction is common in patients with m.3243A>G (Cheong *et al.*, 1999; Guillausseau *et al.*, 2001; Hotta *et al.*, 2001; Iwasaki *et al.*, 2001; Suzuki *et al.*, 2003; Piccoli *et al.*, 2012; Seidowsky *et al.*, 2013) and renal dysfunction may lead to the presence of renal cells in urine (Simerville *et al.*, 2005). Thus it may be that the elevated levels of heteroplasmy in urinary sediment in comparison to other measures of heteroplasmy may be indicative of renal dysfunction and denote a genuine accumulation of the mutation in such tissue. It may therefore be

useful to investigate the relationship between measures of kidney function, such as the urine/creatinine ratio, with heteroplasmy or disease burden in this cohort.

In conclusion, this section does satisfy the question of whether urine heteroplasmy levels in general move upwards or downwards; they appear to remain static. The question of whether individual patients can or do show systematically changing heteroplasmy over time cannot be answered by this dataset however; for this we require patients with three or more longitudinal measurements of heteroplasmy.

4.4.5.3. Variability in urine heteroplasmy measurement

I have quantified the variation in urine heteroplasmy measurement in this cohort as 11.8% (the standard deviation of the sequential measurements) and used bootstrapping to estimate the confidence interval for this estimate (9.62%, 14.01%). The White test for heteroscedacity illustrates that this variance is not dependent on the time between measurements. This supports the conclusion that this variation originates either in random fluctuations within the biological systems that contribute to the cells found in urinary sediment or the heteroplasmy measurement assay itself; however, the same assay is used for measuring heteroplasmy from other sources without such extreme variation and thus the heteroplasmy of the cells in urine themselves are the most likely source.

Though previous studies have commented on the variation in urine heteroplasmy (Blackwood *et al.*, 2010), to my knowledge, this is the first formal quantification of the variation. The standard deviation of urine heteroplasmy is troublingly large. Urine heteroplasmy has been recommended as the best correlate for disease burden (Whittaker *et al.*, 2009), but a confidence interval of ±23% (calculated from the mean standard deviation of 11.8%) seriously impacts on the ability of clinicians to understand the expected level of heteroplasmy in other tissues when using urine heteroplasmy as a guide. The source of this variation is currently unknown. A prior study investigated urine heteroplasmy levels, and found that the time of day that urine was collected was not a significant factor in the variation (Blackwood *et al.*, 2010). Interestingly, they found that even measuring heteroplasmy on the same day sometimes recorded differences of up to 20% from the median measurement, which is consistent with the standard deviation reported here.

If urine levels are to be used, it would appear that a single value alone is not sufficient. Multiple repeated testing will reduce the variation, for instance 4 samples will halve the

standard error. Such an approach is laborious however; though it may be preferable to an invasive technique such as a muscle biopsy.

4.4.5.4. Blood heteroplasmy decline

The results in section 4.4.4.6 expand current understanding of blood heteroplasmy decline whilst raising new questions.

The exponential decline rate is shown to be independent of age in this cohort. This is supportive of the hypothesis that the blood decline is indeed exponential. Additionally, though the exponential rate is not statistically significantly non-zero in this cohort, this may be due to insufficient power, as I had a limited number of serial measurements in blood, and the data is suggestive of blood heteroplasmy decline over time, with a median decline factor of 1.3%.

The significant relationship between the decline factor and the initial heteroplasmy level is of concern, as it would indicate that the rate of decline is dependent on the heteroplasmy level (note that since we have calculated an exponential decline factor, the relationship we are looking at is multiplicative, and should be independent of the actual heteroplasmy levels). Scrutiny of Figure 4.26B reveals that this trend is perhaps due to the existence of patients with apparently increasing heteroplasmy in the low (0-20%) heteroplasmy range. A potential contributor to this is that small errors in measurement at low heteroplasmy levels translate into large errors in the multiplicative factor; a change from 4% to 5% heteroplasmy in a year translates as a decline factor of -25%, but a change from 30 to 31% is a mere -3.3%. To minimise such errors, N must be increased or heteroplasmy measured over as long a time period as possible, so that the effect of measurement error and random fluctuations are diluted. However, even if we restrict the data to measurements taken 5 or more years apart, this relationship still exists, albeit not statistically significant ($P = 0.1233$), though the sample is much reduced ($N = 14$). The cluster of points that show negative decline (and thus increasing heteroplasmy) suggest that at heteroplasmy levels below around 20%, decline is no longer occurring in general. Furthermore, several patients above this level also show no decline over prolonged periods; for instance, a patient with 48% blood heteroplasmy demonstrated no decline over a period of 7 years, and several others have similar periods of static and relatively high blood heteroplasmy (the full dataset can be found in Appendix III). This suggests that some patients may have an asymptotic level of blood heteroplasmy far above zero. To my knowledge, there is no discussion in the literature

of the asymptotic level to which heteroplasmy is declining, nor a consideration of the intra-patient variability in this decline. Indeed, all the discussion appears to implicitly or explicitly assume that heteroplasmy levels decline to zero or negligible levels, most apparent in the published exponential decline model (Rajasimha *et al.*, 2008). This belief is strongly reinforced by the fact that older patients tend to have very low levels of heteroplasmy. However, interpreting these very low levels at older age in this way misses the point that patients presenting at older age generally have milder disease phenotype and are thus likely to have had lower heteroplasmy levels initially. The only way to correctly analyse intra-patient variation in heteroplasmy decline is a thorough longitudinal analysis, but further investigation of this will require a larger cohort of patients studied for a longer period. However, we can conclude that non-zero asymptotic levels of blood heteroplasmy is a possible explanation of the shape of the data seen here, and this is a potentially significant observation. It will be considered further later in this discussion, and in the discussion of disease progression in m.3243A>G in section 4.5.5.

A second approach to studying blood heteroplasmy decline is seen in section 4.4.4.8, which uses urine heteroplasmy as a nominative initial level at conception. This approach is justifiable on account of the stability over time of urine heteroplasmy, and the close correlation of urine heteroplasmy and levels in other tissues, both of which have been previously discussed. In section 4.4.4.8, I have demonstrated that there is a noticeable, albeit non-significant, relationship between the decline rate and the period of decline. It is also important to note that the mean decline factor of around 3% is considerably higher than that calculated using the repeated blood heteroplasmy data. There are several potential explanations of this.

The first is that urine heteroplasmy could systematically overestimating the embryonic haematopoietic heteroplasmy level and thereby inflate the decline factor estimate. However, the relationship with skeletal muscle heteroplasmy suggests this is not the case, though it may be overstated at high heteroplasmy levels.

Secondly, though the mechanism of haematopoietic heteroplasmy decline is not fully understood, it has been proposed to occur due to loss of stem cells through a mitosis-dependent mechanism (Rajasimha *et al.*, 2008). With this in mind, comparative stem cell cycling rates in embryogenesis and later life may be critically important. It has been shown in mice that during embryogenesis almost the entire population of

haematopoietic stem cells (HSCs) cycle every twenty four hours, from around embryonic day 7 (E7) until around 3 weeks post natal; after this point 95% become dormant and cycle every 145 days, whilst the remaining 5% cycle every 36 days (Pietras *et al.*, 2011). If the cell cycling rates are similarly rapid during human embryonic development it would be expected that blood heteroplasmy decline would be greatly accelerated during this period. This would be expected to significantly inflate the decline factor.

However, it must also be considered that embryonic stem cells are not heavily reliant on oxidative phosphorylation for their energy requirements and are more reliant on glycolysis in the hypoxic embryonic environment (Shyh-Chang *et al.*, 2013); thus oxidative phosphorylation defects are not expected to express a strong phenotype in stem cells. However, it is also true that adult stem cells are also glycolytic in general, and in particular HSCs (Suda *et al.*, 2011) (Simsek *et al.*, 2010), speculated as a necessity in the hypoxic environment of the bone marrow niche. Thus any argument against a stem cell driven mechanism applies equally to adult and embryonic HSCs. Importantly, however, any long term changes in heteroplasmy in the haematopoietic system are almost certain to originate in the HSCs. Though the problem of lack of an oxidative phosphorylation driver for selection has been long problematic, recently it has been shown that mitochondrial transcription is heavily regulated in HSCs, despite exhibiting relative low mitochondrial membrane potential and thus low oxidative phosphorylation activity (Norddahl *et al.*, 2011). It has also been shown that even early embryonic changes in HSCs are observable in mutagenic mtDNA mice, which are hypothesised to be caused by subtle ROS or redox environment changes (Ahlqvist *et al.*, 2012). Thus even in the absence of active oxidative phosphorylation, mtDNA dysfunction has been shown to affect HSC activity and differentiation. Rajasimha *et al.* proposed a mechanism whereby stem cells with high levels of mutation are progressively lost (Rajasimha *et al.*, 2008); a feasible hypothesis is that oxidative phosphorylation defects may alter the differentiation products of stem cells, as suggested by the aforementioned studies (Norddahl *et al.*, 2011; Ahlqvist *et al.*, 2012). Stem cells can divide in three fundamental ways; symmetrically into two further stem cells, symmetrically into two progenitors, or asymmetrically into one of each kind. A subtle shift towards either of the latter two mechanisms would result in a declining heteroplasmy level in the stem cell population. Alternatively, the shift in heteroplasmy may result from the ROS-induced senescence of HSCs with high levels of mutant

mtDNA (Shao *et al.*, 2011). Whatever the mechanism, a study on the relationship between blood heteroplasmy at birth (measured from Guthrie cards) and sources of relatively stable heteroplasmy measured in adults (such as urine or muscle) may give useful insight into embryonic blood heteroplasmy decline.

4.4.5.5. Relationship between urine and blood heteroplasmy

The linear relationship between urine heteroplasmy and the square root of blood heteroplasmy has not been previously reported and is intriguing. Firstly, it is further evidence against blood heteroplasmy decline to a nominally low level and in support of an asymptotic level of blood heteroplasmy related to the initial level of heteroplasmy in the embryo. This does need to be verified, by long-term longitudinal observation of blood heteroplasmy in m.3243A>G patients. However, this study does suggest that in the majority of older patients blood heteroplasmy levels have neared a plateau level. This offers an insight into the mechanism of blood heteroplasmy decline, and contradicts the mechanism suggested by Rajasimha *et al.* that assumed an asymptotic level of zero irrespective of the starting point (Rajasimha *et al.*, 2008). The significance of age in this relationship suggests that heteroplasmy does continue to decline throughout life, though as said the asymptotic level is related to the initial heteroplasmy level.

4.4.5.6. Conclusion

The variability in urine heteroplasmy has been quantified (s.d. 11.9%), which demonstrates that a single measurement of urine heteroplasmy is imprecise and undermines confidence in this measurement for prognostic purposes. It has also been confirmed that urine heteroplasmy does not appear to change systematically over time in patients at a population level.

Blood heteroplasmy has been shown to decline as previously reported, however evidence against the previously proposed model of exponential decline to a negligible level has been presented; the asymptotic level of blood heteroplasmy appears to be linearly related to levels in other tissues, as demonstrated by the relationship to urine heteroplasmy.

4.5. Comparative value of blood, urine, and SKM heteroplasmy in prediction of disease progression

4.5.1. Introduction

Though skeletal muscle was long been considered the gold standard for measuring heteroplasmy levels, urine, as discussed, is generally accepted as the most suitable non-invasive measure of heteroplasmy, and indeed has been shown to be more correlated to disease phenotype (Whittaker *et al.*, 2009; de Laat *et al.*, 2012).

Blood heteroplasmy has repeatedly been shown to be unrelated to disease phenotype (Mehrazin *et al.*, 2009), in contrast to studies that report the utility of skeletal muscle (Macmillan *et al.*, 1993; Chinnery *et al.*, 1997; Jeppesen *et al.*, 2006). When blood heteroplasmy has been found useful, it is not in comparison with other measures of heteroplasmy (Laloi-Michelin *et al.*, 2009). However, though blood heteroplasmy has historically been disregarded as suitable for analysis of disease phenotype, the results from section 4.4 suggest that it may in fact be of utility, since, when suitably transformed, blood heteroplasmy is highly linearly correlated with urine heteroplasmy and thus should be expected to exhibit reasonably similar predictive properties, and indeed improved, if the variability in blood heteroplasmy measurement is lower.

4.5.2. Aims

I aim in this section to compare the comparative value of the three measures of heteroplasmy that are available in the cohort (blood, urine, and skeletal muscle) to assess which is the best predictor for disease burden and progression, particularly in the light of observations made in the previous section regarding the variability of urine heteroplasmy. Though I also consider muscle heteroplasmy, this is more limited in scope for drawing conclusions as the number of patients with samples available is considerably smaller.

4.5.3. Methods

I look at a variety of the analyses from the previous section on disease progression and compare the merit of blood and urine heteroplasmy as predictive factors.

For the section comparing blood and urine, I use only the cohort of patients that have both urine and blood heteroplasmy measured, hence N values are smaller than for previous sections. For other comparisons, e.g. with muscle heteroplasmy, I use the largest cohort available with all appropriate measures of heteroplasmy.

To facilitate the ability to draw firm conclusions regarding superiority of one predictor over another I implemented a bootstrapping (resampling) methodology, as described in section 2.10. This extends the modelling technique and allows me to calculate a confidence interval for the comparative success of predictors, rather than a simple point estimate using the entire cohort.

4.5.3.1. Bootstrapping

For each comparison, bootstrapping was used to resample the dataset. Bootstrapping for the basic statistical comparisons using a single estimate is a straightforward resampling of the data. For longitudinal modelling, bootstrapping was done at the patient level, not the individual observations, i.e. patients were chosen with replacement from the pool of patients, and all observations for each patient were included in the analysis each time. This was to ensure that the resampling was concerned with statistically independent units. The longitudinal modelling utilised for comparison did not contain random effects, as the intention was to evaluate the effectiveness of the fixed effects.

Each bootstrapping used 1000 resamples to evaluate the distribution of the statistic under investigation and 95% confidence intervals were calculated.

A variety of statistical measures were suitable for bootstrapping. For multiple regression I used the difference in R^2 as the statistic for comparison. For longitudinal modelling the difference in log likelihood is utilised. The methodology is in accordance with published guidelines (Lewis *et al.*, 2011).

For each bootstrapping the distribution of the chosen statistic was examined to ensure the distribution was reasonable and the statistical approach appropriate.

4.5.4. Results

4.5.4.1. Cohort

134 patients with NMDAS data have both urine and blood heteroplasmy recorded. 34 patients have both urine and SKM heteroplasmy data, 31 have both blood and SKM heteroplasmy.

4.5.4.2. Blood heteroplasmy is a better predictor of NMDAS score than urine using multiple regression

I use multiple logistic regression with age and heteroplasmy as independent variables, and the summary NMDAS score as the dependent variable. Parameters and data from

the model are found in Table 4.4. Blood heteroplasmy increases R^2 to 27.8% from 19.7% with urine heteroplasmy.

Bootstrapping was used to compare blood and urine heteroplasmy as predictors. Figure 4.28A depicts the distribution of the percentage point ratio increase observed when using blood rather than urine heteroplasmy. The mean increase in R^2 is 7.7%, and the change is significantly above zero (P = 0.042).

Model	N	Age	Heteroplasmy	R^2
urine	134	P < 0.0001 B = 0.416	P < 0.0001 B = 0.401	19.7%
blood	134	P < 0.0001 B = 0.608	P < 0.0001 B = 0.585	27.8%
urine	134	n/a	P = 0.0070	5.3%
blood	134	n/a	P = 0.0125	4.6%

Table 4.4 Comparison of the correlation between NMDAS score and blood or urine heteroplasmy.
Using the same cohort of patients, multiple regression with age and either urine and blood heteroplasmy is used to compare model fit. P values and standardised coefficients for each predictor are shown, with the R2 for the overall model. Blood heteroplasmy increases R2 to 27.8% from 19.7%, and standardised coefficients for age and heteroplasmy both improve. Simple linear regression without age was also conducted with each predictor; urine is shown to be marginally more predictive.

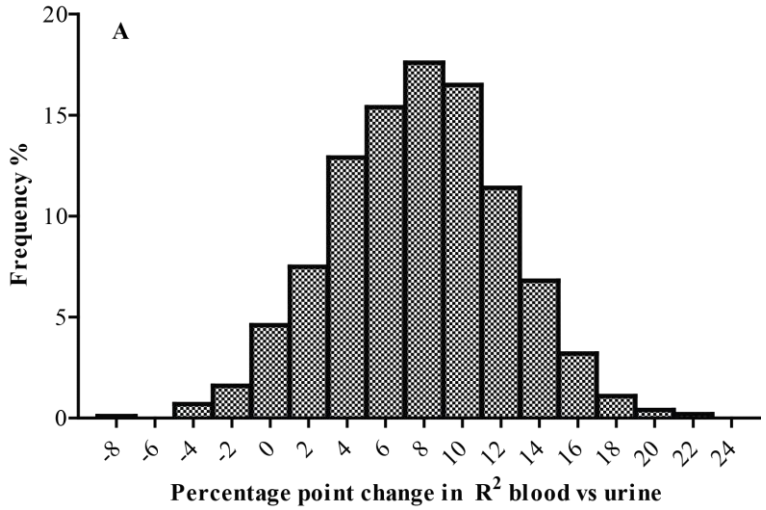


Figure 4.28 Resampling comparison of blood and urine heteroplasmy for the total disease burden multiple regression model.
N = 134 1000 resamples with replacement of the patient pool are used to generate the above distributions. The percentage point increase of R^2 when using blood instead of urine heteroplasmy is shown, using multiple regression with age and heteroplasmy as predictors. Mean increase is 7.7%, and the change is significantly greater than zero (P=0.042).

4.5.4.3. Blood heteroplasmy is better than urine heteroplasmy in longitudinal modelling of total disease burden; SKM heteroplasmy is not significantly better than either urine or blood heteroplasmy.

The longitudinal model is the same as that described in section 4.3.4, however, the cohort is slightly smaller as it is restricted to those patients with both urine and blood heteroplasmy (N = 134). The model includes the early-onset flag as a predictor, as this is necessary for correctly formed residuals (N = 27 early-onset, 107 late-onset).

Both urine and blood heteroplasmy are highly significant predictors of total disease burden in this slightly smaller cohort ($P < 0.0001$), as are age and early-onset (both $P < 0.0001$ in each model).

The models were initially compared using the AIC, as shown in Table 4.5. The blood model is superior to urine by comparing the AIC values in each case. Note that the AIC value for urine differs from that previously described as the cohort is smaller. In addition, bootstrapping was performed to investigate the stability of the improvement in model fit by using blood heteroplasmy; the results are shown in Figure 4.29.

To further validate the model using blood heteroplasmy, the average (per patient) studentised residuals were examined against the predictors of heteroplasmy and age. The results are shown in Figure 4.30. The residuals in the model are not significantly correlated with either predictor, demonstrating that the model is well formed.

Muscle heteroplasmy was also compared to both urine and blood. The results are also found in Table 4.6. The cohort in both cases is small, and no statistical significance is reported in either comparison, however, both blood and urine explain more variance in these smaller cohorts than muscle heteroplasmy.

Model	AIC	P value	R^2
Urine	-45.0	0.1285	1.74%
Blood	-57.3	0.3636	0.62%

Table 4.5 Comparison of blood and urine heteroplasmy models using AIC.

The table shows the AIC value for the overall model. The P value and R^2 relate to the regression of the predicted value of the model versus the average studentised residuals; significance in this relationship would indicate poor model fit. Both urine and blood have residuals that are uncorrelated with the predicted values of the model.

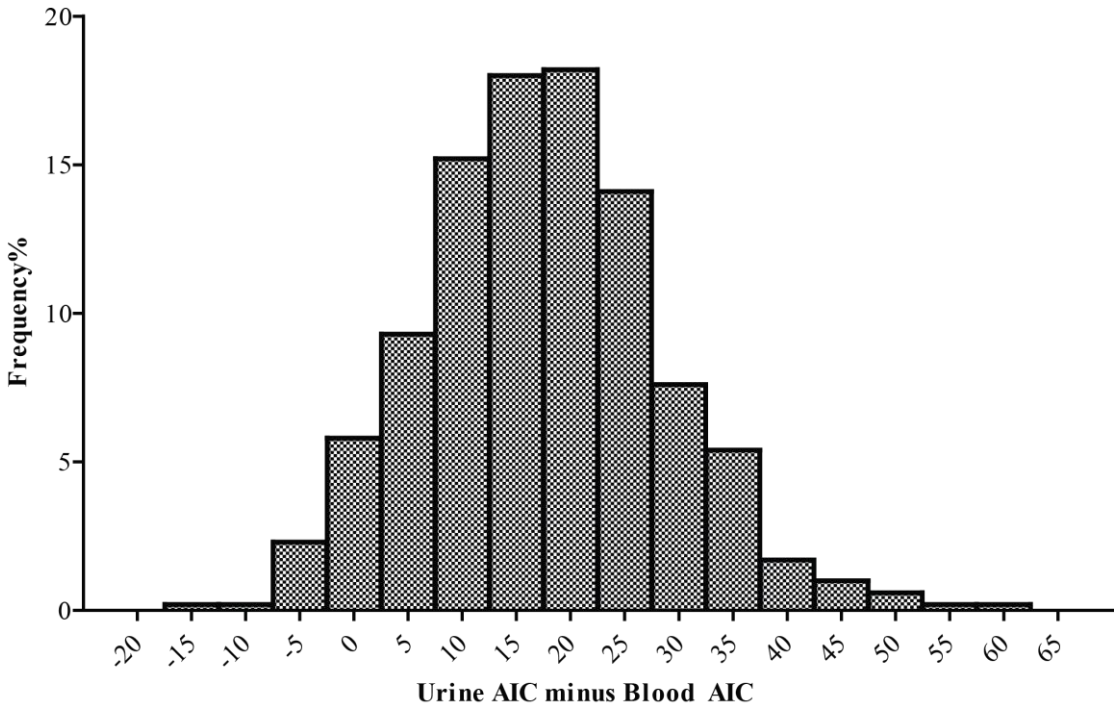


Figure 4.29 Resampling comparison of urine and blood heteroplasmy using the longitudinal mixed effects model.

The AIC difference for 1000 bootstrapped resamples is shown; a positive AIC difference indicates blood is a better predictor than urine. The AIC difference is significantly above 0 (P = 0.046) indicating that blood is a consistently better predictor than urine.

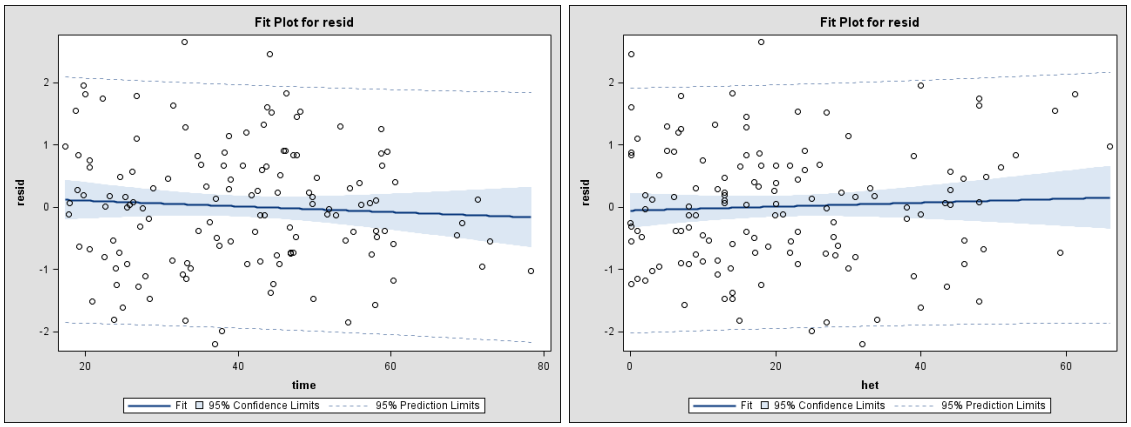


Figure 4.30 Average studentised residuals of the longitudinal mixed model compared to the model predictors of age and blood heteroplasmy.

(Left) Residuals are not significantly correlated with age (P = 0.4302, r = -0.05). **(Right)** Residuals are not significantly correlated with blood heteroplasmy (P = 0.5565, r = 0.03).

	AIC difference				
Comparison	N	Mean	95%L	95% U	P(diff>0)	**Conclusion**
Urine minus Blood	134	17.5	-2.8	41.9	0.0460	blood significantly better than urine
Muscle minus Blood	31	5.7	-12.9	23.2	0.1840	blood better than muscle, not significant
Muscle minus Urine	34	12.9	-3.7	32.6	0.0600	urine better than muscle, not significant

Table 4.6 Comparison of urine, blood, and muscle heteroplasmy for longitudinal modelling of total disease burden.

The mean AIC difference and 95% confidence limits calculated using bootstrapping are shown. The significance value (P) is the proportion of resampled values with a difference greater than 0. Blood is significantly better than urine heteroplasmy. The cohort of patients with muscle is small and no significant difference can be found between muscle and either blood or urine, though both blood and urine are in general more predictive than muscle heteroplasmy.

4.5.4.4. Comparison of blood and urine heteroplasmy in predicting individual features from the NMDAS.

Bootstrapping was used to probe the logistic regression model incorporating age and heteroplasmy as predictors. For each simulation the blood and urinary heteroplasmy models were compared to identify the best fitting model. Figure 4.31 illustrates the proportion of resamples in which urine heteroplasmy is the better predictor; the line at 50% separates those features better predicted by urinary heteroplasmy (upper half) and those better predicted by blood heteroplasmy (lower half). Cut-off lines at 20% and 80% are also drawn to identify those with a strong bias in favour of either heteroplasmy measure. The features that are better predicted by urinary heteroplasmy are ptosis, pyramidal and extra-pyramidal features, and migraine. The first three of these are rare in the cohort; migraine is common but not strongly correlated with either age or urinary heteroplasmy. Several common features are predicted better by blood heteroplasmy, including cerebellar ataxia, diabetes mellitus, cognition, exercise intolerance, and visual impairment.

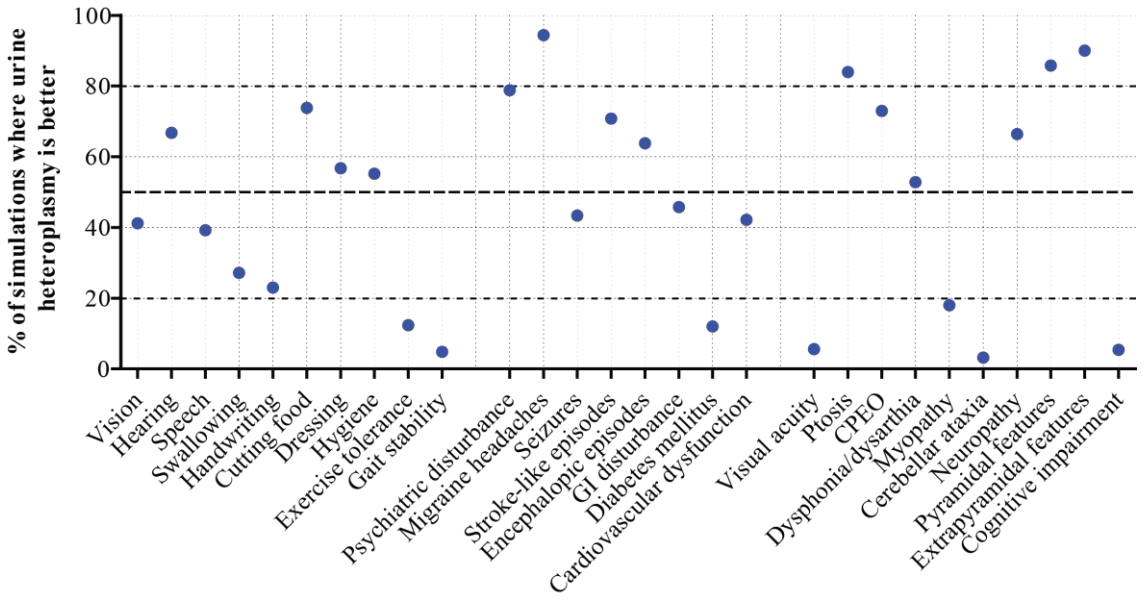


Figure 4.31 Blood and urine heteroplasmy model comparison for individual phenotypic features.
For each feature, the patient pool was resampled 1000 times and the models compared to see which of the two models (blood or urine heteroplasmy) was the better predictor. Cerebellar ataxia, cognitive impairment, diabetes mellitus, exercise tolerance, gait stability, myopathy, and visual acuity are better predicted by blood heteroplasmy; extra-pyramidal, migraine, ptosis and pyramidal by urinary heteroplasmy.

4.5.5. Discussion

I have used urine heteroplasmy throughout the modelling of disease burden and progression. This is on account of various studies that have identified urine heteroplasmy as the preferred non-invasive heteroplasmy measure in m.3243A>G (Whittaker *et al.*, 2009; de Laat *et al.*, 2012). However, the results in this section show that blood heteroplasmy is more predictive of both total disease burden and also several individual phenotypic features of m.3243A>G disease.

Blood heteroplasmy has been repeatedly discredited as being of prognostic value (Mehrazin *et al.*, 2009; Whittaker *et al.*, 2009); this was intuitively consistent with the finding that blood heteroplasmy levels are not static and decline over time. However, the results presented in section 4.4 provide substantial evidence in support of the merit of blood heteroplasmy as a more useful predictor than urine heteroplasmy. It is significantly better in predicting total disease burden, both using basic statistical analysis of the average NMDAS score for each patient, and also using longitudinal modelling of disease progression. With regard to individual phenotypic features the picture is not so clear; there are several features that are better predicted in our cohort by urine rather than blood. However, the features with a strong bias towards urine heteroplasmy are generally rare in the m.3243A>G cohort, such as extra-pyramidal and

pyramidal features. Migraine is the only feature in this group that is common, however it has already been shown in section 4.3.5 that migraine is poorly predicted by either age or urine heteroplasmy. Conversely, the features that are strongly biased towards blood heteroplasmy are both common and also strongly predictable by urine heteroplasmy, including cerebellar ataxia, diabetes mellitus, exercise tolerance, myopathy, and cognitive impairment. The notable exception to this is deafness; 67% of resamples found urine a better predictor than blood.

Muscle is not significantly better than urine or blood heteroplasmy in this cohort, and indeed the trend is in the opposite direction. However, the sample size (31/34 patients) is relatively small, and statistical significance would be difficult to achieve with this size of cohort, particularly considering that I have shown heteroplasmy predicts little of the overall variation in total disease burden. However, if urine is indeed more predictive than muscle, despite being highly variable as reported in section 4.4, the reason for this is unknown. I will return to this discussion in section 4.6.4.4 with regard to sex differences in the m.3243A>G mutation.

I speculate that blood heteroplasmy may be a useful prognostic indicator for two reasons. The first is that the level of heteroplasmy in blood, though in decline, are more consistent than measures of urine or muscle heteroplasmy. I have already been shown that urine heteroplasmy measurements vary considerably for single patients; if blood is also superior to muscle heteroplasmy it may indicate that a single muscle biopsy heteroplasmy level is also a poor estimate of average heteroplasmy.

The discussion on the utility of heteroplasmy must be put into the context of how useful heteroplasmy as a measure of the underlying pathology. It is clear from the basic modelling using GLM that heteroplasmy is a poor predictor of disease burden. This begs the question of whether another easily measurable quantity may be more useful, for example wild-type mtDNA copy number. It has been shown in a recent study that wild-type mtDNA copy number is a more useful predictor of disease phenotype than heteroplasmy (Liu *et al.*, 2013). As for previous studies (Whittaker *et al.*, 2009) urine was noted as superior to blood, however both of these studies fail to incorporate age into their modelling.

Liu *et al.* convincingly shows that wild-type mtDNA copy number is a better indicator than heteroplasmy; however, their results are potentially confounded by the lack of age as a predictor. They found that younger subjects had lower mtDNA copy number. Total

mtDNA copy number was significantly higher in the m.3243A>G patients than in controls. The data shown for total copy number by age group was interesting. For both blood and urine total mtDNA copy number, the highest copy number was found in the 10-20 age group; the lowest was found in the < 10 age group, and above 20 years of age the copy number was in between the two. The control data is perhaps the more informative of the two data sets here, since it is not biased by more severe disease phenotype found in younger patients in the m.3243A>G patients. It would appear that copy number is significantly higher in the 10-20 group than the other groups, in both blood and urine. This would suggest a non-linear trend of total mtDNA copy number with age, which would potentially confound any use of this measure for prediction, unless age adjusted correctly.

There are some caveats to add to this analysis, and all analyses using heteroplasmy measures. The methodology for assessment of heteroplasmy has changed and improved over the years, and heteroplasmy levels used in this analysis span over 13 years. Though this is not expected to materially affect the analysis in general, it may affect the variability in the heteroplasmy measurements.

4.6. Sex differences in m.3243A>G mutation

4.6.1. Introduction

There are diseases associated with mtDNA mutations that have a clear difference in expression between sexes, for instance Leber's hereditary optic neuropathy (LHON), to which males are four to five times more susceptible (Yu-Wai-Man *et al.*, 2009). However, it is less clear whether there are sex differences in other mitochondrial mutations. The prospect of more severe male disease was raised in 1996 by Frank and Hurst, who hypothesised that maternal inheritance would open up the prospect of more severe phenotypic disease for males, particularly relating to infertility, but also other systems exhibiting male-female dimorphism as mutations deleterious only to males would not be selectively filtered from the germline (Frank and Hurst, 1996).

From the phenotypic perspective, there is recent work that highlights the potentially critical importance of male-female dimorphism, particularly as regards ageing. Camus and colleagues showed that ageing is accelerated in *Drosophila melanogaster* by deleterious mitochondrial mutations, but only in males, not females (Camus *et al.*, 2012).

4.6.2. Aims

The aim of this chapter is to explore sex differences in the m.3243A>G cohort. I chose to focus on m.3243A>G, as it is the largest set of patients in the MRC cohort.

The principal question for exploration in this chapter is whether there any differences in the phenotypic expression of the m.3243A>G mutation between sexes. This will encompass examining differences in heteroplasmy levels, disease burden, and specific phenotypic features.

Several large studies in m.3243A>G have been conducted, and in each case females outnumber males in a ratio of approximately 2:1. To explore the cause of this imbalance, I chose to investigate maternal family inheritance to establish whether there is any difference in the ratio of males and female children born to mothers carrying the m.3243A>G mutation.

4.6.3. Methods

4.6.3.1. Cohort

I utilise the same cohort of patients as described in section 4.3. All families with recorded pedigrees were used in the pedigree analysis. For other analyses all patients in

the cohort with the appropriate factors (blood, urine, and muscle heteroplasmy) were included.

4.6.3.2. Meta-analysis

For analysis of the sex ratio at birth, I also collated data on family trees from published literature to compare with the data from our own cohort.

A search was done in PubMed for any of 3243, MELAS or MIDD, together with PEDIGREE, and any publications with defined pedigrees were included. In addition, any pedigrees from publications that were reviewed in the wider investigation into m.3243A>G were included.

4.6.3.3. Sex ratio methodology

Each maternal family lineage was considered separately, and a female/male birth ratio was calculated for each family.

All generations of descendants through the maternal line below the earliest known m.3243A>G mutation carrier were included; the total number of females and total number of males were summed and the ratio calculated. The top level mother was not included in the ratio. It was not always necessary for the top-level mother to be a confirmed carrier by genetic testing. For instance, if two or more siblings (non-identical twins excepted) at any level were confirmed by genetic diagnosis to be carriers of the mutation then their mother is assumed to be a carrier, since the chance of separate sporadic mutation events in two siblings is negligible.

4.6.3.4. Statistical analyses

For sex differences in disease burden I used multiple regression to look at heteroplasmy, age, and NMDAS score.

All tests on sex ratio differences were non-parametric.

4.6.4. Results

4.6.4.1. Cohort

The cohort is the same as used in section 4.3, consisting of 146 patients in total.

4.6.4.2. Gender balance in large cohort studies of m.3243A>G

Table 4.7 shows the balance of male and female participants in several recent large cohort studies on m.3243A>G, including our own. Each study comprises an independent and non-overlapping cohort.

Source	Male	Female	Female Proportion
Newcastle cohort	72	124	63.3%
(Majamaa-Voltti *et al.*, 2006)	33	67	67.0%
(Kaufmann *et al.*, 2011)	30	56	65.1%
(de Laat *et al.*, 2012)	42	85	66.9%
(Suzuki *et al.*, 2003)	36	77	68.1%

Table 4.7 Male female balance in large cohort studies on m.3243A>G.

All studies show a consistent proportion of around two thirds female to one third male participants.

4.6.4.3. Sex ratio at birth

4.6.4.3.1. General population

Table 4.8 shows the females proportion in live-births in the in the general population.

Population	Proportion at birth (Estimated)	Proportion in population over 65
World	48.3%	56.2%
UK	48.8%	55.6%

Table 4.8 Proportion of females in the population.

At birth the proportion of females in the population is slightly less than half, at similar proportions both worldwide and in the UK. Above the age of 65 the balance shifts to a female dominated population. (Central Intelligence Agency, 2013).

4.6.4.3.2. Analysis of m.3243A>G carriers

Nineteen separate publications were found for the meta-analysis, detailing 39 separate pedigrees. Twenty-two separate pedigrees were identified from the patients in the Newcastle cohort. Source data for the analyses can be found in Appendix IX and Appendix X.

The proportion of female live-births was calculated for each pedigree. The distribution for each data set is depicted in Figure 4.32. In the Newcastle cohort the median female proportion (58.8%) is significantly higher than that of the general population (P = 0.045, N = 44). The median proportion in the meta-analysis (61.5%) is also significantly higher than the general population (P = 0.0005, N = 39).

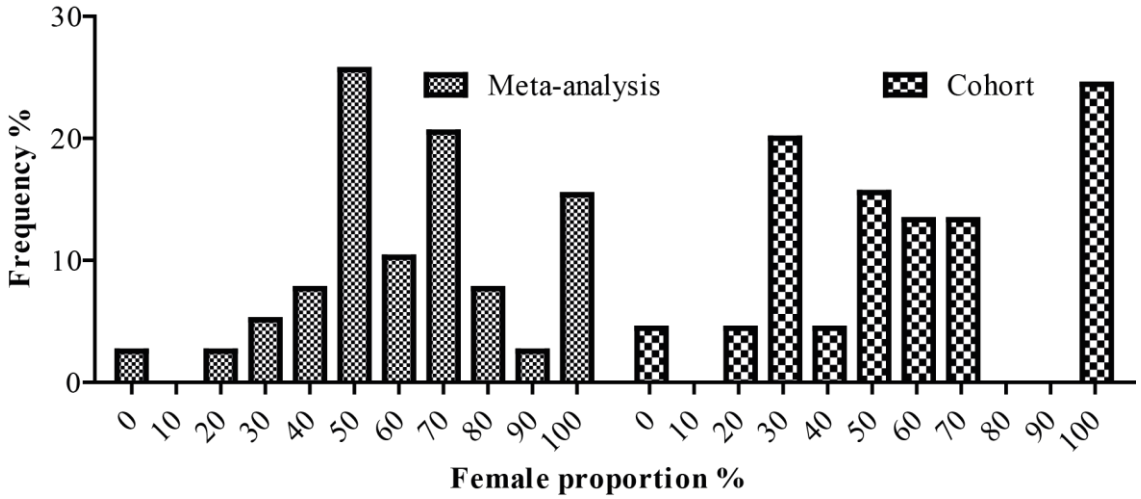


Figure 4.32 Proportion of live-births that are female in the meta-analysis and this cohort.

The proportion of female live births is significantly different from the proportion of the general population. In the meta-analysis the estimated median is 61.5% (P = 0.0005, N = 39), in the Newcastle cohort the estimated median is 58.5% (P = 0.045, N = 44). The proportion for the general population is taken as 48.8%, as seen in Table 4.8.

4.6.4.4. Sex differences in heteroplasmy

For each of the three heteroplasmy measures (urine, blood, and muscle) I used multiple regression to examine the effect of age (at biopsy) and sex as predictors.

As shown in Figure 4.33, urine heteroplasmy is significantly higher in men than in women (P < 0.0001, R^2 = 26.5%, N=145). It is not for significantly different in blood (P = 0.8770, R^2 = 37.0%, N=134) nor in muscle (P = 0.1784, R^2 = 10.6%, N = 36).

The relationship between blood and urine heteroplasmy is also modulated by sex. In multiple regression with urine heteroplasmy, age at biopsy, and sex as predictors, and the square root of blood heteroplasmy as the dependent variable, sex is a highly significant predictor (P < 0.0001) as well the other two predictors. (N = 159, R^2 =78.0%).

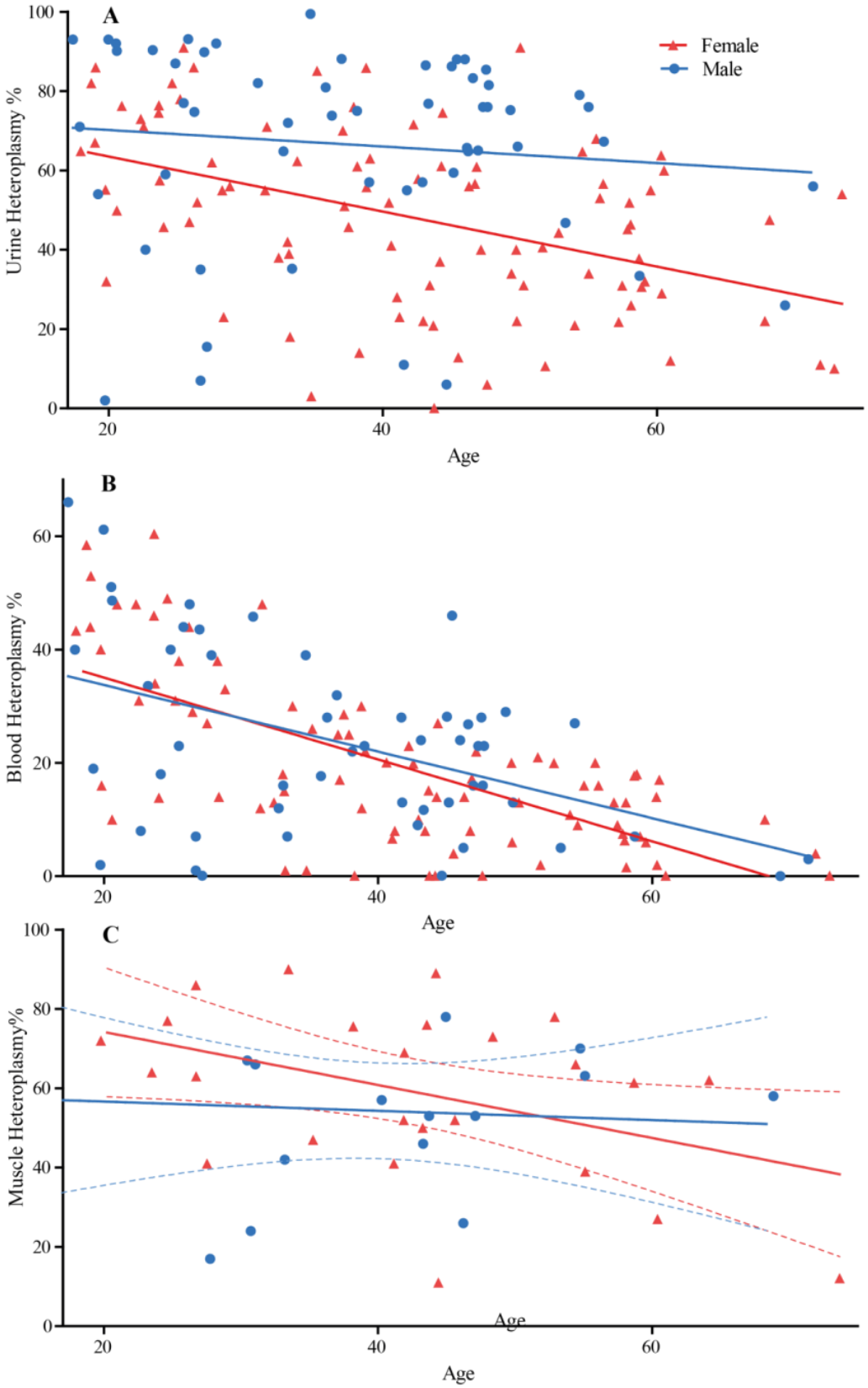


Figure 4.33 Sex differences in heteroplasmy levels.

(A) Urine heteroplasmy levels are significantly lower in females than in males ($P < 0.0001$, $R^2 = 26.5$). **(B)** Blood heteroplasmy levels are not significantly different between males and females ($P = 0.8770$, $R^2 = 37.0\%$). **(C)** Muscle heteroplasmy is not significantly different between males and females ($P = 0.1784$, $R^2 = 10.6\%$).

4.6.4.5. Sex differences and NMDAS progression

To investigate whether sex has a significant impact on NMDAS score and progression, I first conducted multiple regression with age and sex as predictors, using the scaled NMDAS score ($NMDAS^{0.25}$) as the dependent variable. I first examined two models; one without an interaction term (allowing different intercepts but not gradients for sex), and the second with the interaction term (allowing both intercept and gradient to differ).

The first model finds that the gradients are not significantly different (P = 0.6552, N = 146, R^2 = 6.1%). The second model is illustrated in Figure 4.34. The gradients are not significantly different (P = 0.2403, R^2 = 7.0%).

I also investigated whether sex differences in NMDAS progression are observable when urine heteroplasmy is included in the model. In this model, sex is also not found to have a significant effect on NMDAS score (P = 0.3392, R^2 = 18.3%, N = 146).

With blood heteroplasmy, sex is also not a significant effect (P = 0.6005, R^2 = 27.3%, N = 135).

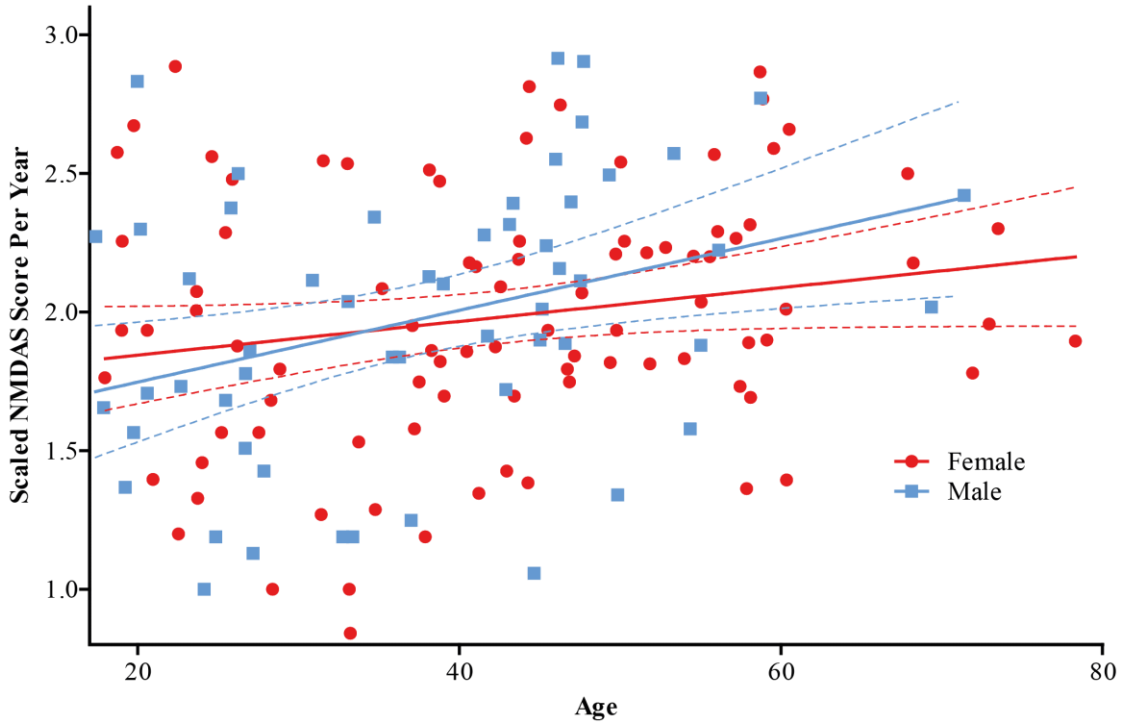


Figure 4.34 Scaled NMDAS score, age and sex.

Linear regression with 95% confidence intervals are shown for males and females. Sex is not a significant effect in NMDAS score (P = 0.6552) and the gradients are not significantly different (P = 0.2403). N = 146. There are no significant sex differences observable in NMDAS score.

4.6.4.6. Specific phenotypic features

As in section 4.3.5, I used logistic regression to model the effect of age and heteroplasmy (both urine and blood) on the severity of phenotypic features as identified by the NMDAS. I use the same modelling approach here, but introduce sex as a third predictor. The results are shown in Figure 4.35. Using urine heteroplasmy, only the severity of ptosis is significantly more severe in males than females; however, exercise intolerance and myopathy are more severe in females. Using blood heteroplasmy, males are significantly more severely affected in several features, namely ptosis, vision, cutting food, hearing, cardiovascular dysfunction, and encephalopathic episodes. There are no features where females are significantly more severely affected than males. The odds ratios for almost all other features show that the (non-statistically significant) trend is for females to be less affected than males; the exceptions to this are diabetes, GI disturbance, exercise tolerance, and myopathy.

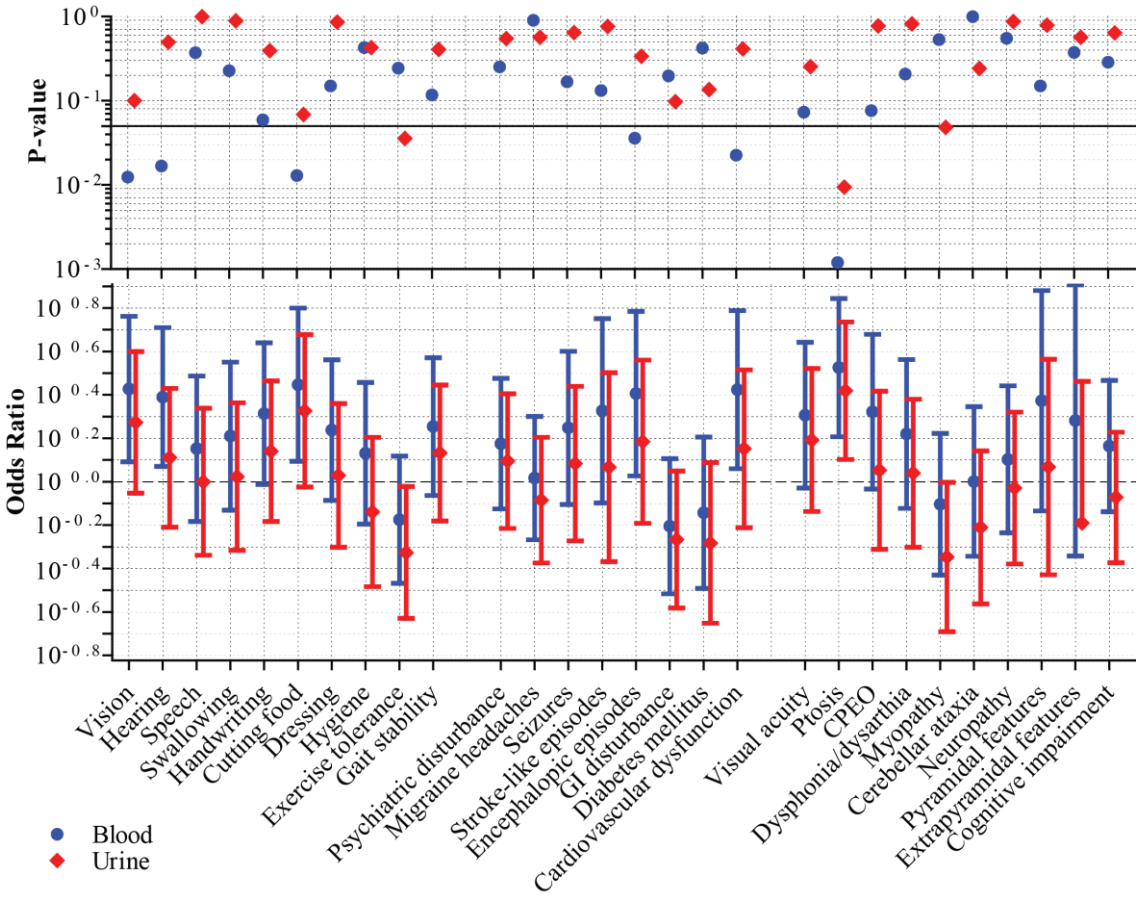


Figure 4.35 P-values and odds ratios for the predictive effect of sex on the risk of developing specific phenotypic features.

For each feature, I look at the effect of sex on the logistic model incorporating also age and heteroplasmy (either urine or blood). Both graphs use log scales. Solid line on the P-value graph indicates P = 0.05; blow this line indicates statistical significance. An odds ratio significantly above 1 implies that males are more affected than females at the same age and heteroplasmy; an odds ratio below 1 that males are less affected than males. Using urine heteroplasmy, only ptosis is significantly more severe in males, whereas exercise intolerance and myopathy are more severe in females. Using blood heteroplasmy, males are more severely affected in several features, namely ptosis, vision, cutting food, hearing, and encephalopathic episodes. There are no features where females are significantly more severely affected than males.

4.6.5. Sex differences discussion

As shown in Table 4.7, cohort studies on m.3243A>G are almost invariably female heavy in their makeup, with females outnumbering males in a ratio of approximately 2:1. This is not necessarily an indication that there are more female carriers of the m.3243A>G mutation than male, for several reasons. The first is that females are more frequent users of health services than men in general (ONS, 2010), for which reason one might expect more women to be identified as potential carriers of the mutation than men. The second is related to a peculiarity of maternal inheritance of mitochondrial mutations; when a patient is identified as a carrier of a potentially inherited mtDNA

mutation such as m.3243A>G, it would be normal clinical practice to assess the mother of the patient to identify whether the mutation is sporadic or inherited, which would swell the number of females assessed (Dr Andrew Schaefer, personal communication). The third is a confound that derives from the asymmetrical ageing profile of the sexes; due to various factors, and in spite of a slightly higher male birth rate, the population above the age of 65 is female dominated (Central Intelligence Agency, 2013).

To avoid population dynamics confounding my investigation, I chose to look at the female/male ratio in live births to females known to carry the m.3243A>G mutation. Conventionally the ratio is expressed as the ratio of male births to female births (1.07 for the worldwide ratio in 2013 (Central Intelligence Agency, 2013)), but when looking at small populations such as individual maternal lineages proportions are preferred to ratios, to avoid division by zero when a family has only female children. Hence I consider the proportion of female births, with extreme values of 0 (all male offspring) and 1 (all female offspring).

The data shows that both in the Newcastle cohort and the meta-analysis there is a significant deviation from the normal birth ratio in favour of female offspring. This is a novel observation, and has not been reported previously in literature to my knowledge. There are two potential causes of this imbalance; an influence on the sex determination of the embryo pre-fertilization, or a pathogenic effect specific to males.

The concept of an adaptive sex ratio has long been of interest to researchers, since the seminal theory of Trivers and Willard that first proposed a link between maternal characteristics and the sex ratio of offspring (Trivers and Willard, 1973). A second theory by Myers agreed with the outcome, but not the reasoning; he theorised that mothers under nutritional stress are more likely to give birth to daughters as daughters are nutritionally less expensive than sons (Myers, 1978). Factors which modulate the sex ratio in mammals have been studied in a variety of species. Diet, for example, has been shown to affect the sex ratio (Rosenfeld and Roberts, 2004). Indeed, it has been shown in a large number of studies that mothers in good physical condition produce more male offspring (Rosenfeld and Roberts, 2004). In humans, a study of the changing sex ratio during and after the 1959-1961 famine in China concluded that mothers in poor condition are more likely to give birth to daughters (Song, 2012). Pregravid diabetes has been shown to bias the sex ratio in favour of female offspring (Ehrlich *et al.*, 2012); this is particularly relevant to the discussion regarding m.3243A>G

considering the prevalence of diabetes in the cohort. The physiological mechanism underlying this adaptive response remains unknown (Helle *et al.*, 2008), but in line with the prior research it seems reasonable to infer that mitochondrial dysfunction could lead to energy stress in mothers and thus a trend towards female progeny.

The second possible explanation of the altered sex ratio is that the mutations are more pathogenic for male offspring. Sex differences in mitochondrial disease have been discussed recently in the scientific literature, prompted by the prospect of mitochondrial replacement (MR) therapy (Reinhardt *et al.*, 2013). The argument made by Reinhardt *et al.* applies to any of the proposed mechanisms of MR, both pro-nuclear transfer (Craven *et al.*, 2010) and mitotic spindle transfer (Tachibana *et al.*, 2009). The argument rests on the co-ordination between nuclear and mitochondrial gene expression, which is highly integrated and tightly co-regulated (Woodson and Chory, 2008). From the perspective of mitochondrial evolution males are an evolutionary dead end, since the effects of mitochondrial genetic mutations on males are irrelevant as they are not passed on down the male line (Parsch, 2011). This leads to the curious situation that mutations that are deleterious to males but not females can potentially be successfully passed on without fear of genetic selection (Frank and Hurst, 1996). Indeed, mitochondrial replacement in fruit flies has been shown to lead to a major change in nuclear gene expression in males but not females (Innocenti *et al.*, 2011), which is suggestive of potential major differences between mitochondrial gene expression between males and females.

If the m.3243A>G mutation were indeed more pathogenic for males than females, it could be expected that males would exhibit faster disease progression and more severe disease phenotype. However, the most appropriate way to assess this is not straightforward. I have already shown in 4.6.4.4 that urine heteroplasmy is higher in males, but that blood heteroplasmy is not significantly different between sexes. I chose to retain heteroplasmy (either blood or urine) and age in the model when looking at both overall disease burden progression and also the severity of specific phenotypic features, so as to control for the correlations observed between sex and heteroplasmy. However, I also investigated the relationship between sex and disease burden and severity with heteroplasmy absent from the model altogether.

Regarding overall disease burden progression (section 4.6.4.5), sex is not a significant predictor. This is true whether heteroplasmy of any source is included as a predictor or not. This lack of an effect with or without heteroplasmy is slightly puzzling, since I

have also shown that urine heteroplasmy is significantly higher in the male patients of the cohort (section 4.6.4.4). It was therefore expected that one of the two models, either blood heteroplasmy or urine heteroplasmy, would exhibit sex differences.

The modelling of individual phenotypic features is more illustrative, particularly comparing the effect of sex as a predictor when using the different heteroplasmy measures as predictors. Here there are several phenotypic features in which males are more severely affected when using blood heteroplasmy. Several of these features are not observably different between the sexes using urine heteroplasmy as a predictor. This suggests that the sex difference in the heteroplasmy level observed in urine is potentially explanatory of some of the variation in phenotypic severity; however, it should be noted that blood heteroplasmy is in general a better predictor for almost all phenotypic features. Of the features that have more severe disease burden for females (albeit non-significantly for blood heteroplasmy), it is notable that both myopathy and exercise intolerance are features that are more severe in females than in males; both of these are related to muscle, in which there are known sex differences (Staron *et al.*, 2000).

The discussion of sex differences in m.3243A>G must be examined in the context of fundamental physiological differences between males and females. It is not disputed that females live longer than males (Austad, 2006; Frank, 2012). There are numerous reported differences in males and females reported that are related to ageing however, and no consensus on the root causes of the difference in longevity and whether a slower ageing process is the root cause of this. However, it has been shown that females demonstrate slower immune system ageing (Hirokawa *et al.*, 2013), suffer less oxidative DNA damage (Proteggente *et al.*, 2002), and, most relevant to a discussion about mitochondrial disease, it has been shown that females have higher mitochondrial antioxidant levels and lower mitochondrial oxidative damage than males, directly attributable to oestrogen levels (Borras *et al.*, 2003),which is highly supportive of the oxidative stress hypothesis of ageing. Thus there may be a natural asymmetry in disease progression that is not attributable to mitochondrial disease *per se* and more related to general differences also observed in the general non-diseased population.

To truly understand whether the mechanism underlying the altered sex ratio is a pathogenic mechanism post-conception or selection pre-conception would require a study of pregnancies that fail to carry to term and an examination of the sex ratio in

unviable foetuses. However, on balance it seems more reasonable to assume that the changes in the live birth sex ratio in m.3243A>G are caused by the same mechanism as other changes discussed previously, such as seen in nutritional deprivation or diabetes.

It is important to acknowledge potential limitations of this study. Firstly, any potential source of bias in the reporting of family pedigrees must be considered. Secondly, caution must be expressed with regard to conclusions of statistical significance; the analysis we are carrying out does involve multiple testing, and caution should be observed in interpreting the p-values of the effects. However, whilst acknowledging the limitation of determining statistical significance in analyses such as this, I believe that the analysis has merit regardless and is informative as to the effect of sex on disease progression and severity.

4.7. m.3423A>G discussion

There are several novel aspects to the work presented here on disease progression associated with the m.3243A>G mutation.

Regarding disease progression, to my knowledge this is the first study to investigate disease burden and progression using multiple regression techniques, which has been shown as invaluable in understanding the relationship between disease burden, heteroplasmy, and age. Without multiple regression techniques the correlation between heteroplasmy and age at a population level severely hampers any attempt using a single predictor at a time.

Secondly, longitudinal mixed modelling has been used to provide predictive models of total disease burden progression. Again, this is novel, and is important for clinicians in understanding the likely course of progression of disease burden.

Thirdly, the risk profiles for individual phenotypic features using age and heteroplasmy as predictive factors are novel and are fundamentally important for understanding the link between age, heteroplasmy, and specific features of disease in patients with m.3243A>G.

The investigation into heteroplasmy also demonstrates several novel findings. Firstly, I have quantified the variability of urine heteroplasmy. This is vitally important, as this is routinely used by clinicians to understand the likely disease burden of patients and this high variability is a serious limitation of urine heteroplasmy as a prognostic measure. The characterisation of the relationship between urine and blood heteroplasmy is also

novel, as is the observation that blood heteroplasmy levels are likely to be asymptotically non-zero in many patients.

The report of the altered sex ratio in favour of female progeny is novel, and opens up further questions about the mechanism behind this alteration. The investigation into sex differences in disease associated with m.3243A>G has demonstrated clear differences between males and females as regards urine heteroplasmy, which is of concern. However, the examination of disease burden and specific phenotypic features indicate that the higher urine heteroplasmy level in males may potentially reflect a genuinely higher disease burden.

In conclusion, though much has progress has been made in understanding disease progression in patients with the m.3243A>G mutation, the limitations of heteroplasmy as a prognostic measure have been clearly demonstrated, and the case for improved metrics more useful than heteroplasmy in predicting disease progression has been clearly made.

Chapter 5

Chapter 5. Disease progression in single large-scale mtDNA deletions

5.1. Introduction

Single large-scale mtDNA deletions are amongst the more commonly reported mutations in mitochondrial disease. Population prevalence in adults is estimated at between 1.2/100,000 (Chinnery *et al.*, 2000b) and 2.9/100,000 (Remes *et al.*, 2005). In the cohort of 671 patients seen at the Newcastle mitochondrial disease clinic, 86 (12.8%) have single large-scale mtDNA deletions.

Single large-scale mtDNA deletions were the first identified genetic cause of mtDNA disease (Holt *et al.*, 1988). The clinical phenotypic associated with single large-scale mtDNA deletions is traditionally divided into three main presentations (Schon *et al.*, 2012). The mildest phenotypic presentation is chronic progressive external ophthalmoplegia (CPEO) (Moraes *et al.*, 1989), which principally involves weakness of the extra ocular muscles and ptosis, but is often associated with more widespread muscular weakness. A second major presentation is Kearns-Sayre syndrome (KSS) (Kearns and Sayre, 1958), a multi-system childhood or teenage onset syndrome. KSS was first characterised with ophthalmoplegia, retinal pigmentary degradation, and cardiomyopathy as principal features; the criteria for diagnosis have since been adjusted to onset under twenty years of age, PEO, and pigmentary retinopathy. However, other features such as cerebellar ataxia, broader myopathy, and other organ involvement are also common. Pearson Syndrome (Pearson *et al.*, 1979) is an infantile onset syndrome that is often fatal, characterised by sideroblastic anaemia and exocrine pancreatic dysfunction. In those who survive infancy it often develops into KSS (Simonsz *et al.*, 1992).

As many studies attempting to analyse phenotypic presentation have noted, these divisions are somewhat imprecise and misrepresent the true spectrum of single mtDNA deletion-related disease; patients denoted CPEO are often sub-divided into groups such as ‘classic CPEO’ at the mild end of the spectrum, ‘severe CPEO’ or ‘CPEO+’, and ‘partial KSS’ or ‘CPEO + Multisystem’ (Auré *et al.*, 2007). Discrete phenotypic classification is severely problematic.

The earliest investigations of disease associated with single large-scale mtDNA found little connection between mitochondrial genetics, biochemical defects and clinical

phenotype (Zeviani *et al.*, 1988; Holt *et al.*, 1989a; Moraes *et al.*, 1989; Rotig *et al.*, 1995). More recent studies are rather contradictory. For example, SKM heteroplasmy is has been reported as non-predictive of either phenotype or age at onset (Yamashita *et al.*, 2008), predictive of onset but not phenotype (López-Gallardo *et al.*, 2009), or predictive of both to an extent (Sadikovic *et al.*, 2010). Similarly, mtDNA deletion size is variously reported as either predictive of phenotype and disease severity (Yamashita *et al.*, 2008) or not predictive (López-Gallardo *et al.*, 2009). Two recent studies found that the location on the genome of the mtDNA deletion was predictive of phenotype or age at onset, though with contrary findings. One study found that mtDNA deletions including one of the three mtDNA-encoded structural components of cytochrome *c* oxidase (COX) (*MT-CO1, MT-CO2* or *MT-CO3* genes) or complex V (*MTATP6* or *MTATP8* genes) had significantly earlier-onset (Yamashita *et al.*, 2008). However, another study found that deletion of the mitochondrial cytochrome *b* (*MT-CYB*) gene was significantly associated with the more severe KSS phenotype as compared to CPEO (López-Gallardo *et al.*, 2009).

To improve analysis of disease burden and phenotype, researchers have attempted to better segregate patients into groups that are more appropriate for analysis. Auré et al segregated patients by the presence of a neurological phenotype, distinguished by the presence of cerebellar involvement (Auré *et al.*, 2007); a second group looked at whether the phenotype was purely myopathic in nature (López-Gallardo *et al.*, 2009). However, as is true for many mitochondrial genetic disorders, the disease spectrum is multidimensional, with a wide range of potential systems involved, each to varying degrees of severity. This severely hampers any attempt to discretely classify the disease. In this context, the NMDAS is a significant aid in understanding disease severity, progression, and the correlation with predictive factors, and offers a route to understanding disease progression that sidesteps the issue of discrete phenotypic classification.

5.2. Aims

In this study, I chose to use repeated measures mixed modelling with the collated NMDAS data to test the hypothesis that mtDNA heteroplasmy levels and mtDNA deletion size and location are predictive of single, large-scale mtDNA deletion disease progression.

I also aimed to apply multiple linear regression analyses to examine the relationship between these predictors and outcome measures such as age at onset, clinical phenotype, the level of biochemical defect (as determined by COX-deficient fibre density) in both our cohort and a meta-analysis of previously published data.

Finally, I aimed to use logistic regression to examine the relationship between individual phenotypic features and predictive factors.

5.3. Methods

5.3.1. Newcastle patient cohort

All the patients in this cohort were investigated by the NHS Highly Specialised Service for Rare Mitochondrial Disorders in Newcastle upon Tyne. The majority (55 out of 87) were recruited to the MRC Mitochondrial Disease Patient Cohort UK and followed in our clinic at regular 6 or 12 monthly intervals, and assessed using the NMDAS at each visit (Schaefer *et al.*, 2006). The remaining 32 patients comprise individuals not seen in our clinic and for whom we have limited clinical information, and those for whom we do not have NMDAS data available. For some of these patients we have been able to ascertain disease onset from medical records. A full listing of the clinical and molecular characteristics of the patient cohort can be found in Appendix XI.

For phenotypic analysis the patients were divided into two groups, those classified either as CPEO or as CPEO with myopathy (combined together, N = 54) and those classified as KSS (N = 9).

5.3.2. Meta-analysis

The meta-analysis is based on a previous meta-analysis by López-Gallardo *et al.* (López-Gallardo *et al.*, 2009) which includes data on patient age at disease onset, age at biopsy, clinical phenotype, SKM heteroplasmy, mtDNA deletion size, and in many cases specific mtDNA deletion breakpoints. I used the data as published by López-Gallardo *et al.* except where review of the literature identified differences between the published data and that used by this group (three cases) or where there was inconsistency between the reported mtDNA deletion size and location of the breakpoints (one case). I also eliminated those cases where the reported mtDNA deletion was characterised by restriction endonuclease digests and there was uncertainty as to whether specific genes under scrutiny (*MT-CO* or *MT-CYB* genes) were deleted (seven cases) or other mutations were reported concurrently for the same patient (nine cases). One patient with a highly unusual mtDNA deletion in the minor arc of the

mitochondrial genome was also excluded. Also excluded are any patients that have been reported in studies from this centre.

The data was augmented with other cases from subsequent studies or those not included by López-Gallardo *et al.* that were identified through PubMed (Johns *et al.*, 1989; Mita *et al.*, 1989; Larsson and Holme, 1992; Oldfors *et al.*, 1992; Ishikawa *et al.*, 2000; Gellerich *et al.*, 2002; Jacobs *et al.*, 2004; Sadikovic *et al.*, 2010).

All data from the meta-analysis can be found in Appendix XII. All amendments and exclusions are listed in Appendix XIII.

For phenotypic analysis, patients classified by López-Gallardo *et al.* as "CPEO without non-muscular signs" were combined with patients classified "CPEO" or "CPEO + myopathy" from the other sources (101 cases in total). This patient group was compared to the patients classified as "KSS" from all sources (104 cases in total). For other analyses, we used the patients in the meta-analysis that had all the appropriate information available.

5.3.3. Muscle biopsy histochemistry data

Sequential cytochrome *c* oxidase (COX)/succinate dehydrogenase (SDH) histochemical reactions were performed on SKM sections following standardised protocols (Old and Johnson, 1989). The percentage of COX-deficient fibres (measuring the extent of COX deficiency) was determined by counting all the fibres in a section (minimum 200). Counting was performed by a single researcher to ensure consistency.

All muscle biopsy histochemistry was conducted by Gavin Falkous.

5.3.4. Determination of level of mtDNA deletion

All molecular analyses were performed using total SKM DNA extracted using standard protocols. mtDNA deletion levels in muscle homogenates was quantified using a validated, multiplex real-time PCR (qPCR) *MT-ND1*/*MT-ND4* assay (He *et al.*, 2002; Krishnan *et al.*, 2007).

All mtDNA deletion level quantification was performed by Georgia Campbell and Thiloka Ratnaike.

5.3.5. Determination of mtDNA deletion size and location

Long-range PCR was used to amplify ~ 9.5 kb region of the mitochondrial genome across the major arc using a single primer set corresponding to nucleotides 6378-15896 (GenBank Accession number: NC_012920.1). PCR reactions used ~100ng of DNA

which was added to PCR mastermix (dH2O, LA Taq buffer (TaKaRa), 10mM dNTPs, 20mM forward and reverse primers and XX units of LA Taq enzyme (TaKaRa)) to a total volume of 50µl and subjected to the following cycling conditions: 94°C for 1 minute; 35 cycles of 94°C for 30 seconds, 58°C for 30 seconds and 68°C for 11 minutes; final extension of 72°C for 10 minutes. Amplified products were separated through a 0.7% agarose gel, using a 1kb DNA ladder to estimate product size and determine mtDNA deletion sizes.

Long-range PCR products were further assessed by restriction digests to map the precise location of the mtDNA deletion breakpoints (Khrapko *et al.*, 1999). PCR amplimers (5µl) were digested to a series of DNA fragments of known length and position within the mitochondrial genome using restriction enzymes *Xho* I, *BamH* I, *Xcm* I and *Dra* I (New England Biolabs) prior to separation through a 0.7% agarose gel. The size of restriction products allows the location of the mtDNA deletion within the genome to be estimated, guiding the choice of appropriate sequencing primers (Supplementary Table 4) to characterise mtDNA deletion breakpoints by Sanger sequencing. (BigDye v3.1 terminator cycle sequencing chemistries using an ABI 3130xl Genetic Analyser; Applied Biosystems).

All mtDNA deletion breakpoint determination was carried out by Georgia Campbell.

5.3.6. Statistical analyses

5.3.6.1. Basic statistical modelling

The general methods for basic statistical analyses are described in section 2.11.1.1.

5.3.6.2. Longitudinal modelling of disease progression

Longitudinal modelling is conducted as described in section 2.11.1.2. Appendix XIV lists the SAS code to generate the models used in this section.

5.3.6.3. Data transformation

As described in section 2.11.1.2, Box-Cox analysis was used to investigate the relationship between the NMDAS score and age at assessment. In addition to this, I wished to investigate the relationship between the other independent variables (heteroplasmy and deletion size) with the scaled NMDAS score.

This was done in two ways; the first used PROC TRANSREG and Box-Cox analysis, but with the QPOINT option to allow examination of quadratic surfaces, i.e. the

introduction of simple powers of the dependent variables and cross interactions to see the effect on variance stability.

QPOINT identifies simple powers of independent variables that stabilize the variance of the data, but does not explore other transformations such as the log function. It was hypothesised that the effect of heteroplasmy may have increased exponentially as heteroplasmy approaches 100%. Hence the transformation het_{log}=log(1-het) was examined to see the effect of this transformation on the fit of the model and variance stability.

5.3.6.4. Individual phenotypic features

The general methodology for logistic regression is explained in section 2.8.

I used two approaches to logistic regression. Both used the three predictors of age, muscle heteroplasmy, and deletion size. For each NMDAS question, a single summary score was determined for each patient by taking the maximum score achieved on each question and the age at which that score was first recorded.

The first approach used binary (dichotomous) logistic regression by identifying the optimal cut-off point for each NMDAS question to divide the cohort into two groups. To identify the optimal cut-off point the area under the ROC curve (AUC) was maximized. This was done separately for each individual predictor, each pair of predictors, and the three predictors together in the model, to allow a full understanding of the use of each predictor. Only cut-off points that partitioned the data into sets containing at least 6 patients (10% of the cohort) in each group were considered to avoid small numbers of patients skewing the results.

The second approach used a proportional odds multiple logistic regression. NMDAS scores were re-categorised as asymptomatic (NMDAS = 0), moderate (1-3), and severe (4-5); this re-categorisation was necessary for a majority of phenotypic features for model to conform to the proportional odds assumption. Pseudo-R^2 values, as described in 2.8.2, were used to compare models.

In both approaches, standardised parameters for the predictors were calculated.

5.4. Results

5.4.1. Patient cohort

The number of patients used in the various analyses are detailed in Table 5.1.

The distributions of age, heteroplasmy, deletion size, and number of NMDAS assessments in the Newcastle cohort are shown in Figure 5.1.

5.4.1.1. Phenotypic spectrum

The phenotypic profile of patients in the cohort is shown in Figure 5.2. There are several features that are absent or rare in the cohort, including seizures, stroke-like episodes, diabetes, and both pyramidal and extra-pyramidal features. The majority of the other features are present in the cohort at level of at least 50%. CPEO and ptosis are 100% penetrant, vision is also affected in 80% of people according to the current function (section I), whilst visual acuity is impaired in 70% of patients. Exercise tolerance is affected in 77% of patients, though 93% of patients are myopathic. Cerebellar ataxia is prevalent at 80%, with gait stability affected in 70% of patients.

	With NMDAS	Without NMDAS	Cohort (total)	Meta-analysis
Number of patients with mtDNA deletion size and muscle mtDNA heteroplasmy data	55	32	**87**	256
Number of patients with mtDNA breakpoints identified	52	31	**83**	184
Number of patient with age at onset data	52	8	**60**	117
Number of patients with age at onset data and mtDNA breakpoints identified	49	7	**56**	83
Number of patients with COX-deficient fibre density data and mtDNA breakpoints identified	49	23	**72**	40
Number of patients presenting with a CPEO, CPEO + Myopathy or KSS phenotype	36	27	**63**	205
Number of patients presenting with a CPEO, CPEO + Myopathy or KSS phenotype and mtDNA breakpoints identified	35	26	**61**	149

Table 5.1 Summary of Newcastle patient cohort and available data in literature.

SKM heteroplasmy and deletion size data was available for all patients. The sub-cohort of patients with identified mtDNA breakpoints is used where gene location is under investigation; the larger cohort with mtDNA deletion size is used where location is not considered. The small number of patients with phenotype data reflects the fact that patients with a multisystem phenotype are excluded from the analysis; this number is restricted to patients with CPEO, CPEO + Myopathy, or KSS.

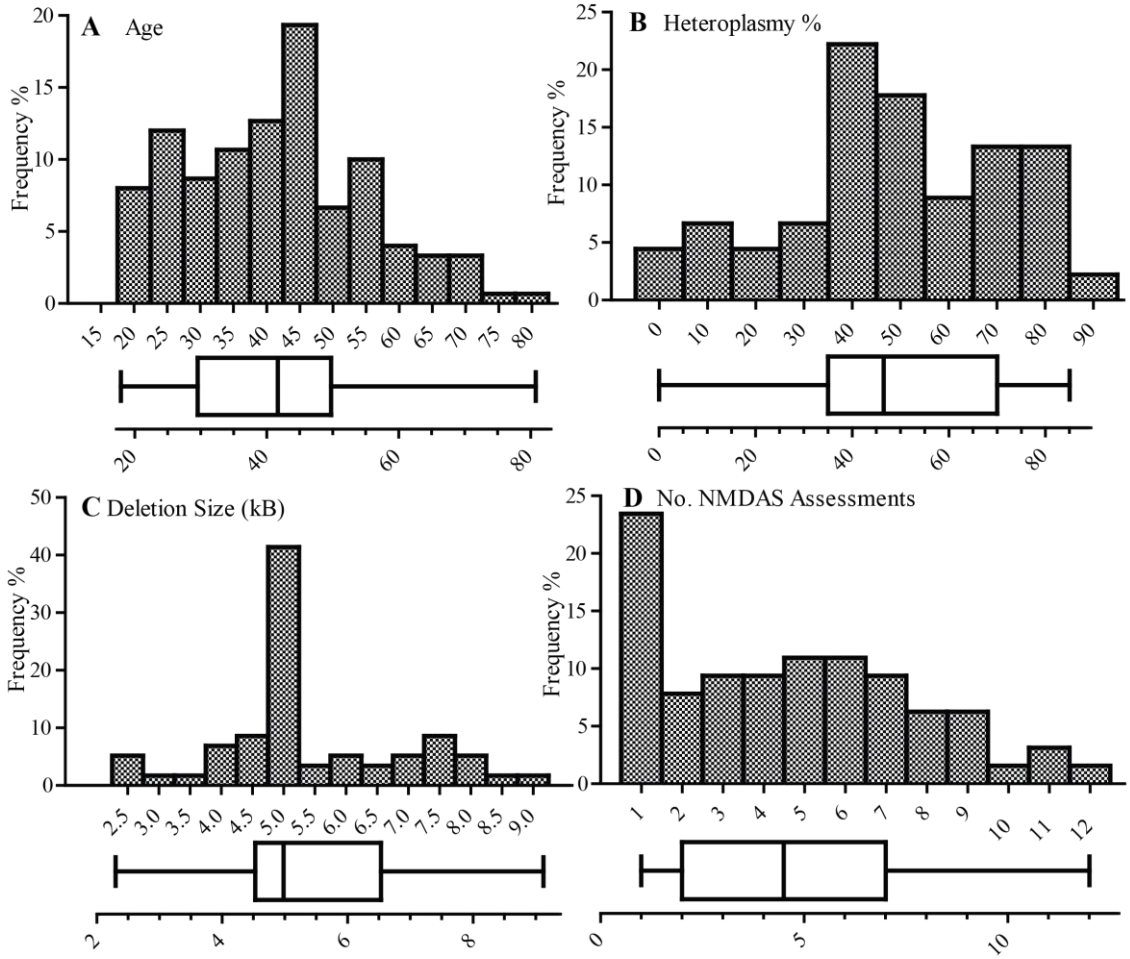


Figure 5.1 Age, heteroplasmy, deletion size, and number of assessments in the cohort.

The patients with heteroplasmy, deletion size, and NMDAS data are considered in this graph (N = 55). **(A)** Age is positively skewed with a median age of 42 year, minimum 18 years and maximum 82 years. **(B)** Muscle heteroplasmy is negatively skewed, with a median of 46.5%. Heteroplasmy ranges from undetectable (0%) to 85%, with an interquartile range of 35%. **(C)** Deletion sizes range from 2.3kB to 9.1kB, and the distribution is fairly uniform other than the great peak at around 5kB, which are in the main common deletion (4,977kB) patients. **(D)** The median number of assessments is 4.5, with a maximum of 12, and an interquartile range of 5.

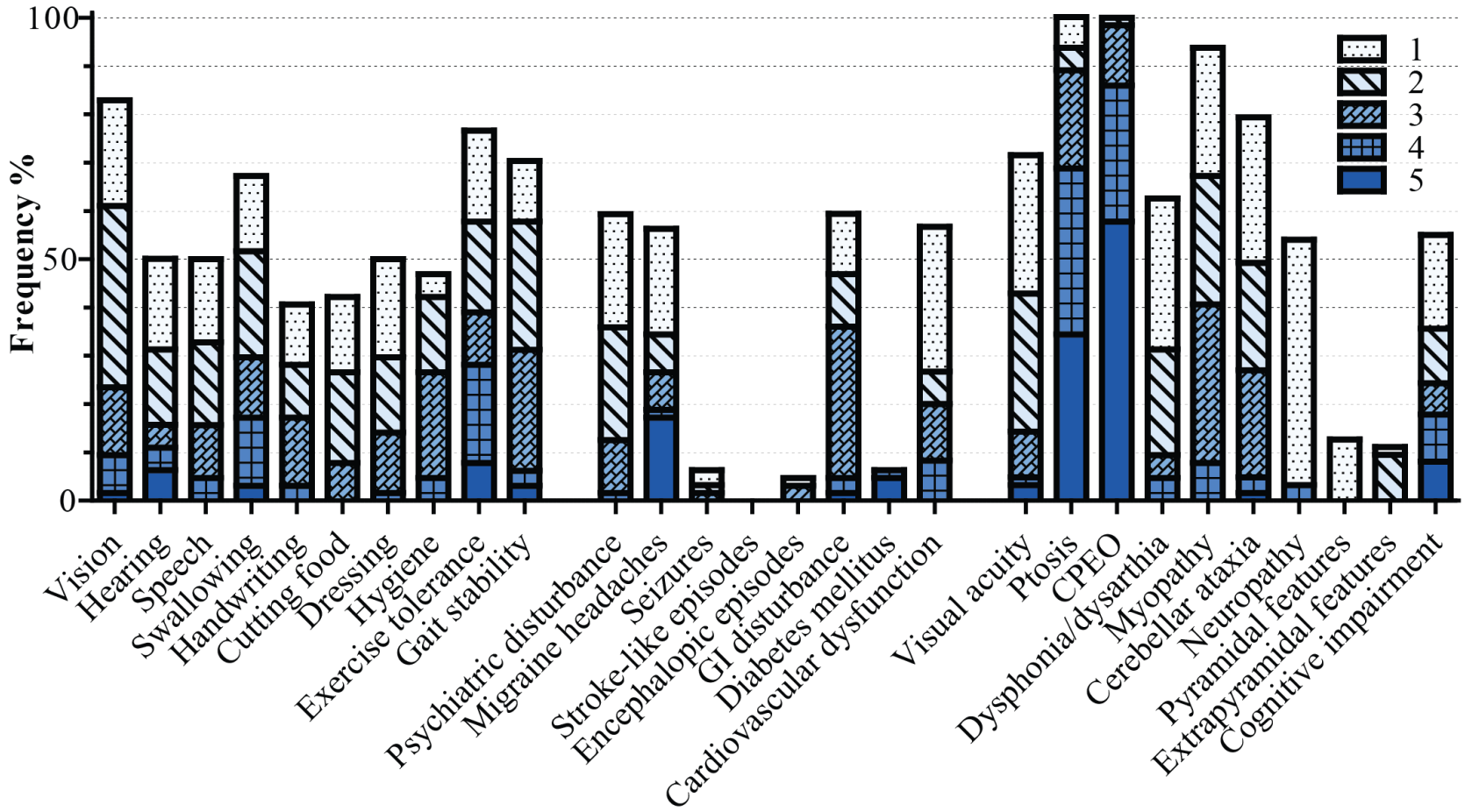


Figure 5.2 Phenotypic profile of the single large-scale deletion cohort.

All patients with NMDAS data are included (N = 55). Bars show the stacked NMDAS score. All patients suffer from ptosis and CPEO to some extent, and 94% from more widespread myopathy. Visual impairment and cerebellar ataxia are the next most common symptoms. General encephalopathy, including stroke and seizure, are almost entirely absent from the cohort, as is diabetes mellitus. Pyramidal and extra-pyramidal features are rare. Mild neuropathy is common but more severe neuropathy is very rare.

5.4.2. Association of NMDAS score and traditional phenotypic classification

NMDAS score was shown to correlate with the traditional phenotype classification, as shown in Figure 5.3 NMDAS score and classical phenotype are highly significantly correlated.

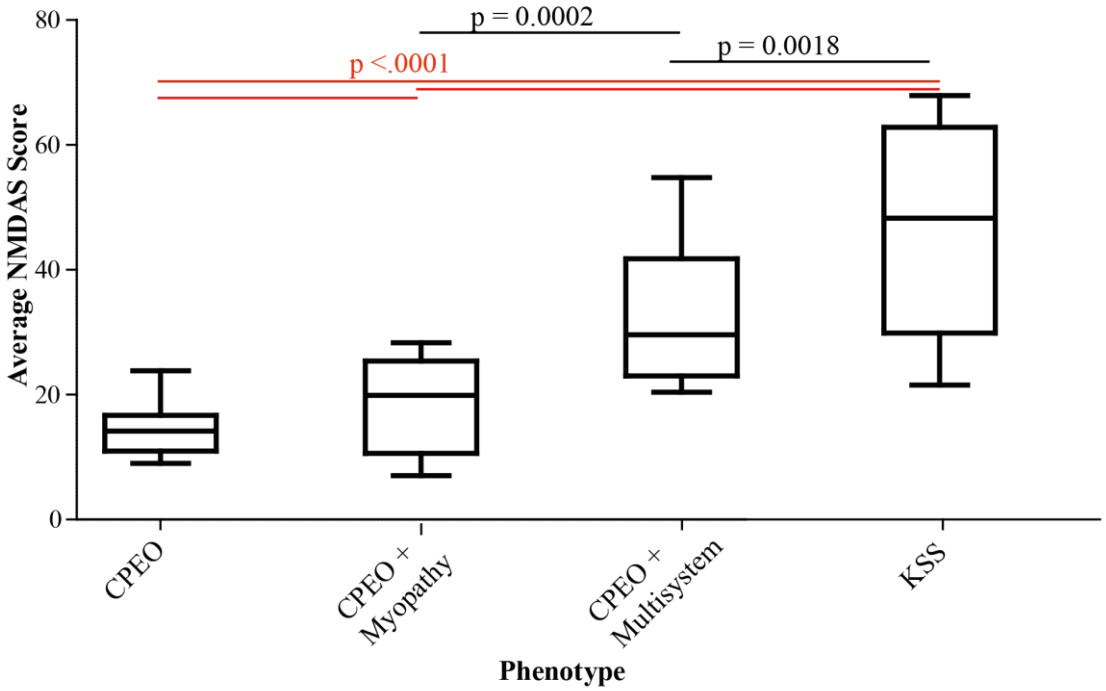


Figure 5.3 NMDAS score and classical phenotype are highly significantly correlated.

P-values for significant individual comparisons are shown; other comparisons are not significant ($P \geq 0.05$). The three comparisons significant at $P < 0.0001$ are highlighted in red. Pure CPEO phenotype patients have a narrower range of NMDAS scores than the other phenotypes; KSS patients have the largest range. KSS patients have the highest NMDAS scores, followed by multisystem, CPEO + myopathy, then CPEO patients. NMDAS scores for patients with a purely CPEO phenotype patients are highly significantly lower than the other three phenotype groups.

5.4.3. Data transformation

Box-Cox identified the fourth root of the NMDAS score ($NMDAS^{0.25}$) as the optimal transformation for linear regression with the single independent variable of age, hereafter called *scaled NMDAS score*.

The QPOINT option of PROC TRANSREG in SAS with heteroplasmy, deletion size and age as independent variables found the square of the deletion size ($size^2$) to be optimal. The transformation het_{log}=log(1-het) was also found to stabilise the variance and improve the model fit. These transformations were used throughout the disease progression modelling. Only the basic investigations into inter-correlations between the predictors used untransformed heteroplasmy and deletion size.

5.4.4. Putative predictors of disease burden and progression are inter-correlated

Before considering the relationship between the putative predictors of disease progression and disease progression itself, I chose to look for inter-correlations between the various predictors. In the Newcastle patient cohort I observed a strong negative linear correlation between the mtDNA heteroplasmy and mtDNA deletion size (N=87, r=-0.49, p <0.0001) (Figure 5.4A). This observation was confirmed in the meta-analysis (N=256, r =-0.18, p=0.0032).

mtDNA deletion size and location were also highly significantly correlated in the Newcastle cohort (N=83, r=-0.48, p<0.0001) (Figure 5.4B), and again this was confirmed by the meta-analysis (N=184, r=-0.29, p<0.0001).

Highly significant correlations were also found between mtDNA heteroplasmy, mtDNA deletion size, and the two proposed genetic loci (*MT-CO* and *MT-CYB* genes) that were identified as significant in previous literature (Yamashita *et al.*, 2008; López-Gallardo *et al.*, 2009) (Table 5.2)

5.4.5. Clinical phenotype, age at disease onset, and NMDAS progression, are correlated with muscle heteroplasmy and mtDNA deletion size

The square root of age at onset was used in these analyses, which was identified by Box-Cox as the optimal transform. For the subjects in the Newcastle cohort with known age at onset (N=60), age at onset was significantly correlated with both mtDNA deletion size (b=-0.41, p=0.0039) and muscle mtDNA heteroplasmy (b=-0.42, p=0.0027) using multiple linear regression (R^2=0.18) (Figure 5.5A). Similarly, in the meta-analysis (N=117), both mtDNA deletion size (b=-0.30, p=0.0008) and muscle mtDNA heteroplasmy (b=-0.30, p=0.0010) were significantly correlated with age at onset (R^2=0.15).

In the phenotypic analysis comparing KSS patients to CPEO patients (excluding intermediate phenotypes, as described in the methods), in the Newcastle cohort (N=64) I found that both mtDNA heteroplasmy (b=-1.4, p=0.0020) and mtDNA deletion size (b=-0.74, p=0.0318) were significantly correlated with phenotype using multiple regression. Similarly, in the meta-analysis (N=192), both mtDNA heteroplasmy (b=-0.56, p<0.0001) and mtDNA deletion size (b=-0.18, p=0.0453) were significantly correlated with phenotype.

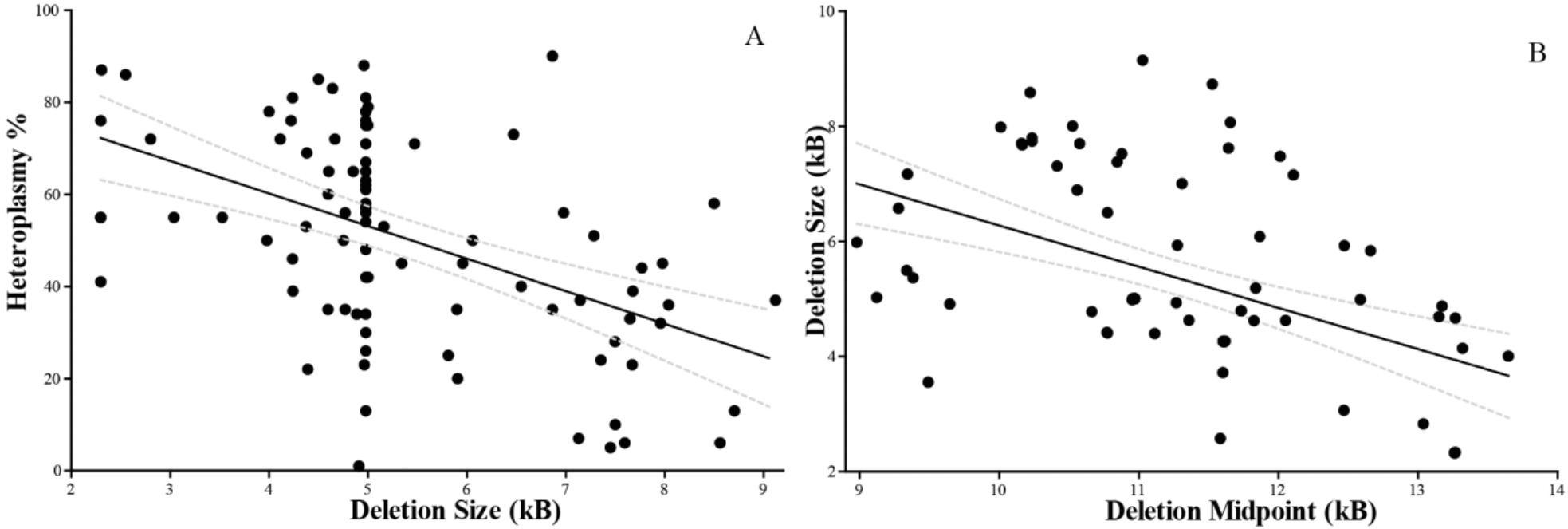


Figure 5.4 Putative predictors of disease progression are inter-correlated.

Regression lines with 95% confidence intervals are shown in each case. **(A)** SKM heteroplasmy is negatively correlated with mtDNA deletion size in the Newcastle cohort. N=87, r=-0.49, p<0.0001. The dense cluster of points just below 5.0kB represents the cohort of patients with the 4,977bp common mtDNA deletion. **(B)** mtDNA deletion size is negatively correlated with the location of the mtDNA deletion midpoint in the Newcastle cohort. N=83, r=-0.48, p<0.0001.

The scaled NMDAS score was divided by the age at assessment to measure NMDAS progression. In the Newcastle cohort (N=55), I found that both mtDNA deletion size (b=0.49, p<0.0001) and mtDNA heteroplasmy (b=0.70, p<0.0001) were significant predictors of NMDAS progression using multiple regression (R^2=0.49) (Figure 5.5B).

	mtDNA heteroplasmy (%)	**mtDNA deletion Size (kB)**	**mtDNA deletion midpoint kB)**	***MT-CO* gene deletion**	***MT-CYB* gene deletion**
mtDNA heteroplasmy (%)		<0.0001 -0.49 (-9.8, -4.3) %/kB	0.3918 +0.10 (-2.6, 6.5) %/kB	0.0423 -0.22 (-27,-0.50) %	0.0199 -0.26 (-27, -2.4) %
mtDNA deletion Size (kB)	0.0032 -0.18 (-4.1, -0.8) %/kB		<0.0001 -0.48 (-0.99,-0.42)	<0.0001 +0.54 (1.6, 3.2) kB	0.0069 +0.29 (0.34, 2.0) kB
mtDNA deletion midpoint (kB)	0.1260 +0.11 (-0.67, 5.4) %/kB	<0.0001 -0.29 (-0.62, -0.21)		<0.0001 -0.73 (-2.7,-1.8) kB	<0.0001 +0.46 (0.72, 1.8) kB
***MT-CO* gene deletion**	0.1262 -0.11 (-15, 1.9)%	<0.0001 +0.47 (1.42, 2.48) kB	<0.0001 -0.72 (-2.3, -1.7) kB		0.0180 -0.28 N/A
***MT-CYB* gene deletion**	0.1625 +0.10 (-2.1, 13)%	<0.0001 +0.39 (0.93, 1.89) kB	<0.0001 +0.55 (1.1, 1.6) kB	<0.0001 -0.28 N/A	

Table 5.2 Inter-correlations between putative predictors of disease burden and progression.
Shaded cells (upper right triangle) show for our cohort p values, correlation coefficients, and 95% confidence intervals for linear regression gradient or estimated difference due to specific mtDNA gene deletion; unshaded cells (bottom left triangle) show the same for the meta-analysis. Units for 95% confidence intervals identify y and x for linear regression, except in the case of mtDNA deletion size (kB) vs. mtDNA deletion midpoint (kB). The strongest correlations in each dataset are between mtDNA deletion midpoint or size and *MT-CO* gene deletion; larger mtDNA deletions tend to include *MT-CO* genes. *MT-CYB* deletion is also associated with larger mtDNA deletion size and mtDNA deletions. The same trends are seen in all cases in our cohort and the meta-analysis, excepting that in our cohort *MT-CYB* gene deletion is significantly associated with lower mtDNA heteroplasmy, whereas in the meta-analysis there is a non-significant trend relating *MT-CYB* gene deletion with higher mtDNA heteroplasmy levels.

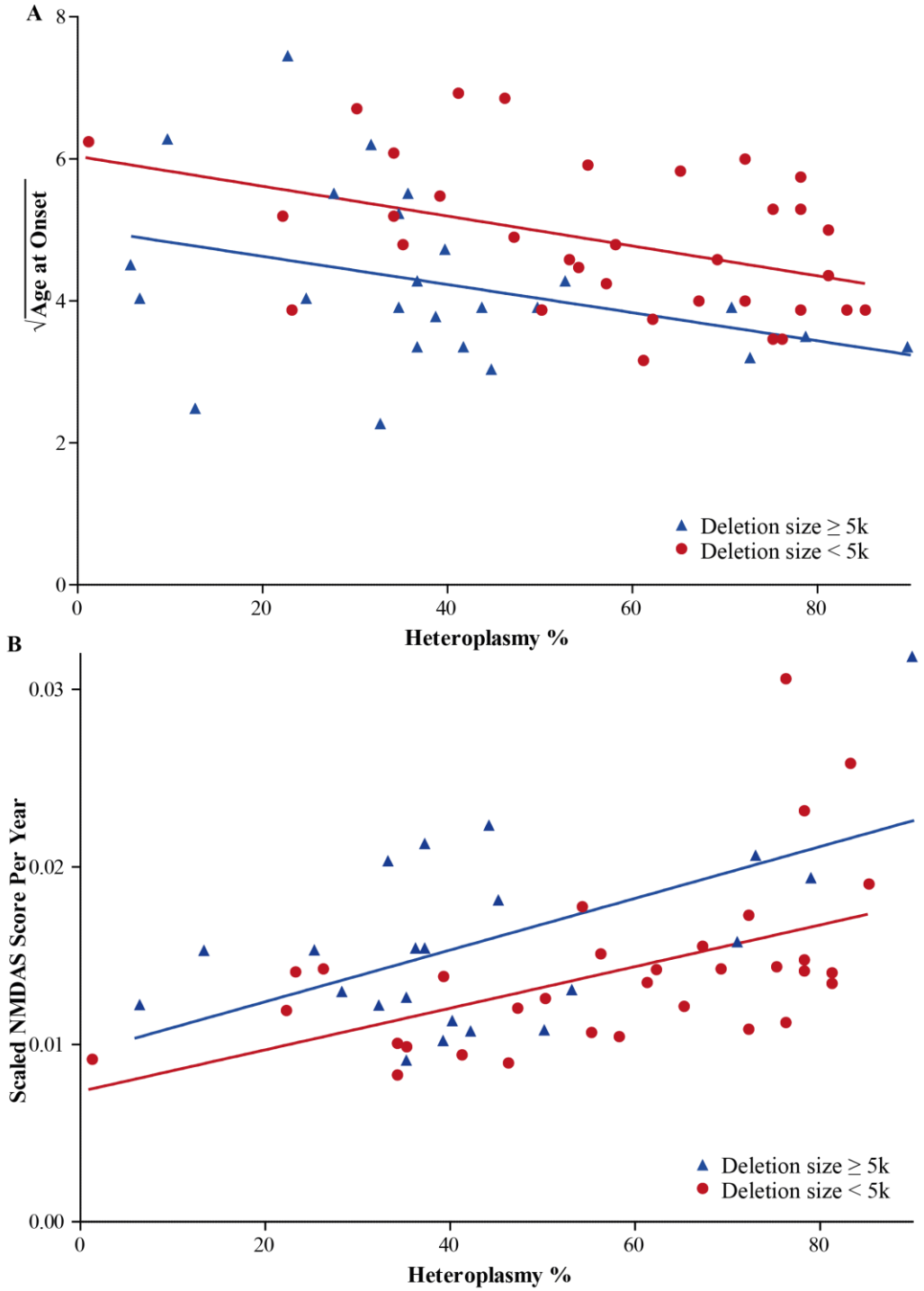


Figure 5.5 Heteroplasmy and deletion size are linearly correlated with age at onset and NMDAS score progression.

For both graphs, mtDNA deletion size is dichotomous for visualization only; regression is performed with deletion size as a continuous predictor. **(A)** Age at onset is predicted by both mtDNA heteroplasmy and deletion size. Y axis shows the square root of age at onset. Data is from the Newcastle cohort. N=60, R^2=0.18. Both mtDNA heteroplasmy (p=0.0027) and deletion size (p=0.0039) are significantly correlated with age at onset using multiple regression. **(B)** NMDAS progression (scaled NMDAS points per year) is highly significantly correlated with both mtDNA deletion size ($p < 0.0001$) and heteroplasmy ($p < 0.0001$). N=55, R^2=0.49. Y axis shows scaled NMDAS score per year.

5.4.6. Disease burden and progression of patients with the 4,977bp common mtDNA deletion are correlated with heteroplasmy

I also examined the cohort of patients whose mtDNA disease is associated with a common 4,977 bp single, large-scale mtDNA deletion (Schon *et al.*, 1989). For this analysis I combined both the Newcastle cohort and the meta-analysis, as neither cohort was large enough in isolation for analysis of the common deletion patients. I observed that SKM mtDNA heteroplasmy was a significant predictor of both clinical phenotype (N=85, OR for 10% change in heteroplasmy 1.43 95%CI (1.13, 1.81), p=0.0030) and also age of disease onset (N=37, r=-0.44, p=0.0063).

5.4.7. COX deficient fibre density is dependent on muscle heteroplasmy and deletion location but not deletion size

I next studied the relationship between the proportion of COX-deficient muscle fibres in patient biopsies, mtDNA heteroplasmy, and *MT-CO* gene deletion (namely, deletion of part or all of at least one of the *MT-CO1*, *MT-CO2* or *MT-CO3* genes). The square root of COX-deficient fibre density was used in all analyses, which was identified by Box-Cox as the optimal transform.

In the Newcastle cohort (N=72) I observed that both SKM mtDNA heteroplasmy (b=0.68, p<0.0001) and *MT-CO* gene deletion (b=0.31, p=0.0018) were significantly correlated with COX-deficient fibre density using multiple linear regression (R^2=0.43) (Figure 5.6). Similarly, in the meta-analysis (N=39), both mtDNA heteroplasmy (b=0.49, p=0.0012) and *MT-CO* gene deletion (b=0.34, p=0.0192) were significantly correlated with COX-deficient fibre density (R^2=0.31).

In both the Newcastle cohort and the meta-analysis the inclusion of *MT-CO* genes within the deleted mtDNA region was significantly correlated with a larger mtDNA deletion size, as seen in section 5.4.4. Hence I chose to examine whether there was a correlation between mtDNA deletion size and COX-deficient fibre density. For this I used multiple regression with all three predictors. In the Newcastle patient cohort (N=72), mtDNA deletion size (b=-0.081, p=0.5104) was not a significant predictor, although both mtDNA heteroplasmy (b=0.65, p<0.0001) and *MT-CO* gene deletion (b=0.35, p=0.0028) remained significant even with deletion size in the model (R^2=0.44). Similarly, in the meta-analysis (N=39), mtDNA deletion size (b=-0.00010, p=0.5586) was not a significant predictor, however both mtDNA heteroplasmy (b=0.48, p=0.0016) and *MT-CO* gene deletion (b=0.43, p=0.0445) remained significant (R^2=0.32).

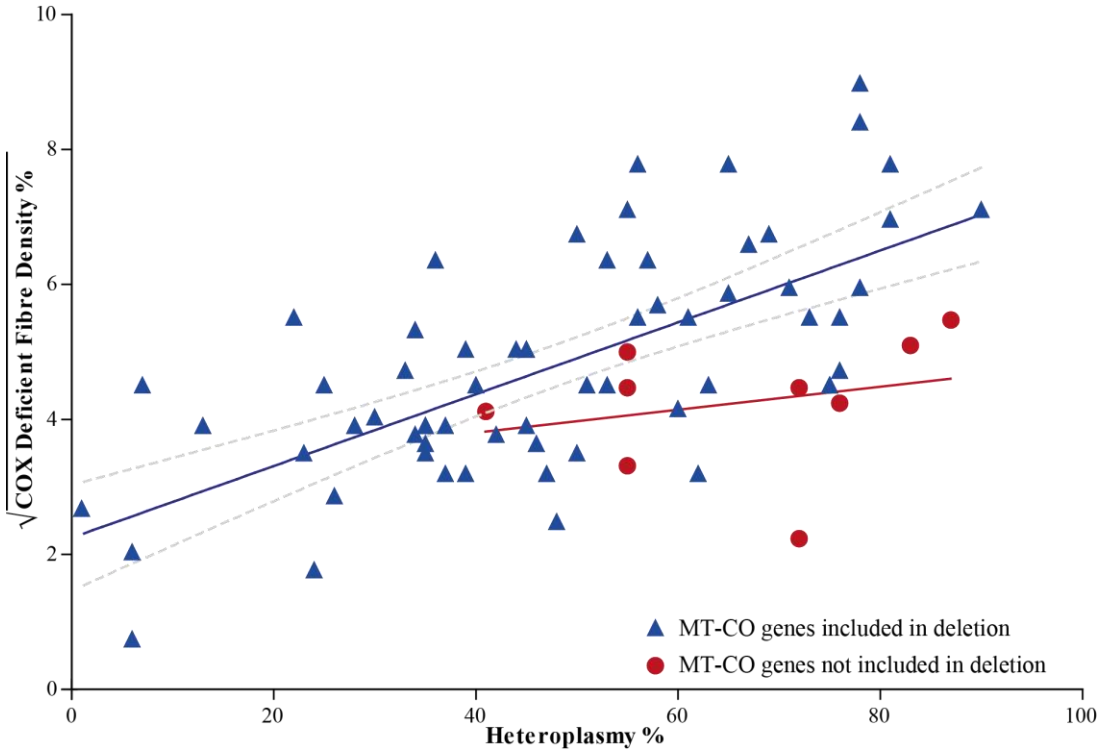


Figure 5.6 COX-deficient fibre density is dependent on SKM mtDNA heteroplasmy and deletion of *MT-CO* genes.

Y axis shows the square root of the COX-deficient fibre density %. Data are from the Newcastle cohort, N=72, R^2=0.43. Heteroplasmy (p<0.0001) and deletion of *MT-CO* genes (p=0.0018) are both significant predictors. Separate regression lines are shown for those that delete or partially delete one or more *MT-CO* genes (N=63, 95% confidence interval for regression line shown) and those that do not (N=9, gradient of regression line is not significantly non-zero, confidence interval not shown).

5.4.8. Longitudinal mixed modelling shows that mtDNA heteroplasmy, mtDNA deletion size and location are predictors of disease progression

Using longitudinal mixed modelling I found that muscle mtDNA heteroplasmy (p<0.0001) and mtDNA deletion size (p<0.0001) were both highly significantly correlated with NMDAS progression in our patient cohort (N=55) (Figure 5.7). The interaction between mtDNA deletion size and mtDNA heteroplasmy was also significant (p=0.0046), which is exemplified by comparing panels A to B, where mtDNA deletion size has a stronger effect at high mtDNA heteroplasmy levels than low mtDNA heteroplasmy levels, and C to E, where mtDNA heteroplasmy has a strong effect with large mtDNA deletions but negligible effect with small mtDNA deletions.

The location of the mtDNA deletion within the mitochondrial genome was also shown to affect disease progression. With mtDNA heteroplasmy and mtDNA deletion size in the model, deletion of the *MT-CYB* gene is significantly predictive of faster progression (p=0.0085, N=52) (Figure 5.74F).

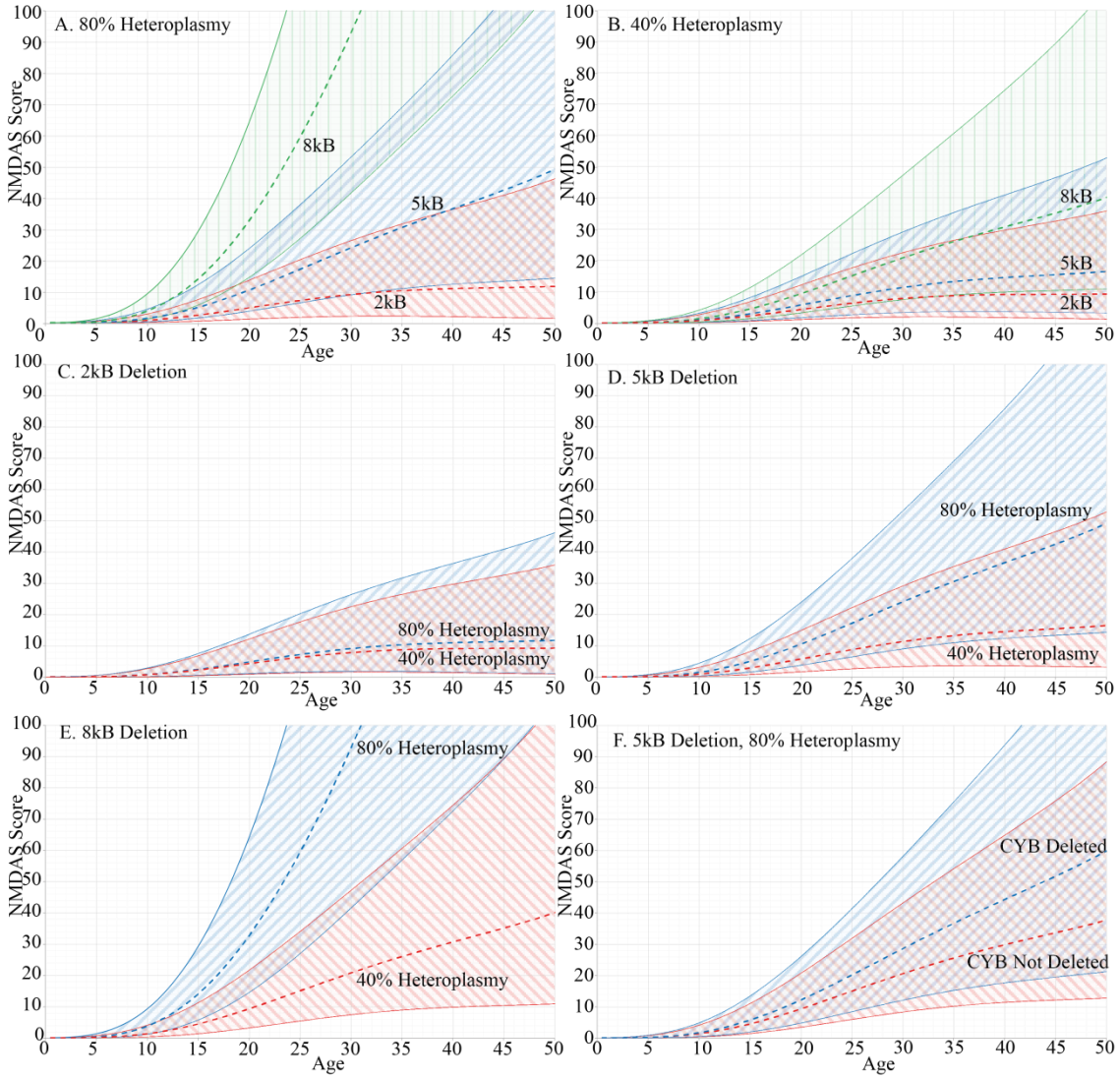


Figure 5.7 The effect of mtDNA deletion size and heteroplasmy on NMDAS progression.

All panels show 95% confidence intervals. **(A-B)** The effect of mtDNA deletion size at 80% and 40% heteroplasmy respectively. Deletion size is shown to have a greater impact at high heteroplasmy than at low heteroplasmy. **(C-E)** The effect of mtDNA heteroplasmy for a 2.0kB, 5.0kB and 8.0kB mtDNA deletion respectively. For small deletions the effect of heteroplasmy on NMDAS progression is negligible, but for larger deletions the effect is substantial. **(F)** The effect of deletion location for a 5.0kB deletion present at 80% heteroplasmy; progression is faster when *MT-CYB* is included in the deleted region. Panels A-E are generated from a model using time, deletion size and heteroplasmy as predictors. The model used for Panel F has an additional deletion location predictor (*MT-CYB* gene inclusion).

5.4.9. Individual patient progression can be modelled longitudinally

Mixed modelling allows incorporation of random effects to model unknown variance in a given system. Hence I was also able to use the clinical and molecular genetic data to model the progress of individual patients and predict the expected course of progression for their overall disease burden. Progression graphs were produced for all patients; in Figure 5.8 I present the predicted progression for five individual patients. The model used to produce these graphs incorporated muscle mtDNA heteroplasmy, mtDNA deletion size, and deletion of the *MT-CYB* gene as predictive factors. Expected progression is shown with 95% prediction intervals. Patients 4 and 5 have similar heteroplasmy but different deletion sizes, and thus exemplify the effect of deletion size on disease progression. Patients 1 and 3, conversely, have similar sizes of deletion (6.9kb and 6.5kB respectively) but different heteroplasmy levels. Patient 2 demonstrates rapid progression, despite low heteroplasmy, on account of the exceptionally large deletion size (9.1kB mtDNA deletion). Patient 1 demonstrates that even a single NMDAS score can represent a useful prognostic input.

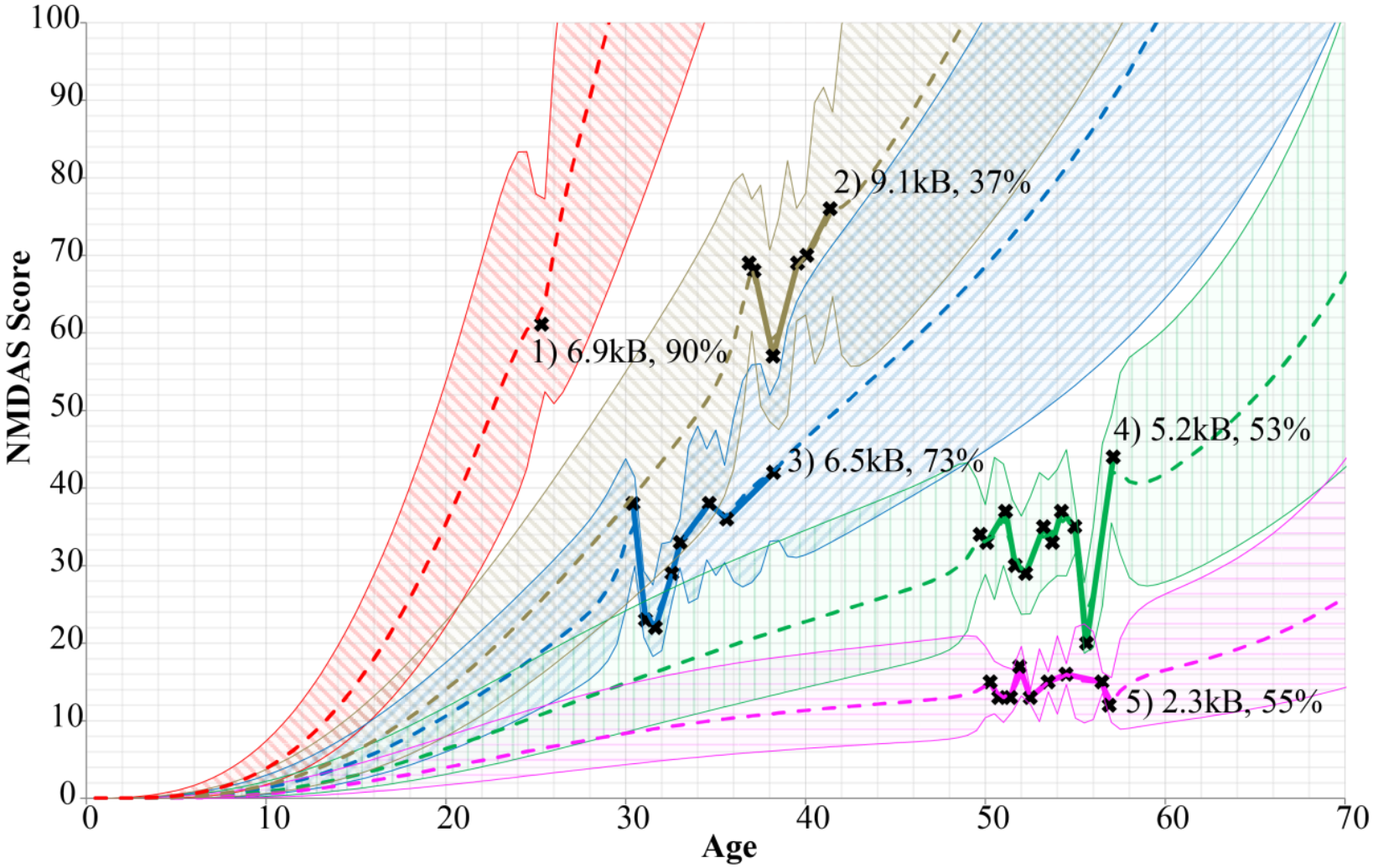


Figure 5.8 Longitudinal modelling of five individual patients with single, large-scale mtDNA deletion disease. The chosen patients are representative of the range of rates of disease progression found in our cohort. Actual NMDAS assessment scores are depicted as crosses joined by solid lines. Each patient is shown with their predicted progression trendline with 95% prediction intervals, and is labelled with deletion size and heteroplasmy. Only patient 2 includes part of the *MT-CYB* gene in their deletion.

5.4.10. Individual phenotypic features

Using a model with age, muscle heteroplasmy, and deletion size as predictors, the area under the ROC (AUC) curve was calculated for each phenotypic feature at each dichotomous cut-off point that had at least 6 patients (10% of the cohort) in each group (below the cut-off point, and above or including the cut-off point). The AUCs for each cut-off point are shown in Figure 5.9. The higher the AUC the better the predictive power of the model. Steep gradients or inflections in the graphs indicate instability in the predictability of the features, i.e. prediction success is heavily dependent on where the cut-off is drawn. Cognitive impairment is notably poor at a cut-off of 2. Hearing is highly predictive at high cut-off points but poor at low cut-offs.

Using the optimal cut-off for each NMDAS feature, models were generated with each combination of the three predictors of age, heteroplasmy, and deletion size. The results are shown in Figure 5.10. As seen in panel C, the combination of all three predictors was the optimal model in almost all cases, other than psychiatric and CPEO, where the addition of deletion size (psychiatric) or age (CPEO) into the model reduces the AUC slightly. Of the lone predictors, heteroplasmy is generally the most predictive, deletion size the least predictive. Age is highly predictive of several features, most noticeably neuropathy and migraine, as neither of the other two predictors contribute much to the AUC. The most highly predicted features were hearing, exercise tolerance, and ptosis, though speech, dressing, and visual acuity also achieved AUCs over 0.8. The lowest AUCs were observed in myopathy and psychiatric disturbance, though gait stability, cutting food, and handwriting are also noticeably poorly predicted. Migraine and psychiatric disturbance are the only features negatively associated with age.

A second approach used the proportional-odds logistic regression with three levels (asymptomatic, mild/moderate, and severe), results are shown in Figure 5.11. Again, heteroplasmy appears to be generally the most valuable predictor, out of age, heteroplasmy, and deletion size, and is significantly associated with half of the 22 features. Exercise tolerance is the most predictable feature, and all three predictors are statistically significant; the only other features significantly associated with all three predictors are ptosis and handwriting, though handwriting has a low pseudo-R^2. Deletion size is a statistically significant predictor for 7 of 22 examined features, most notable exercise tolerance and cerebellar ataxia. Hearing is noticeably poorly predicted in this model as compared to the dichotomous model.

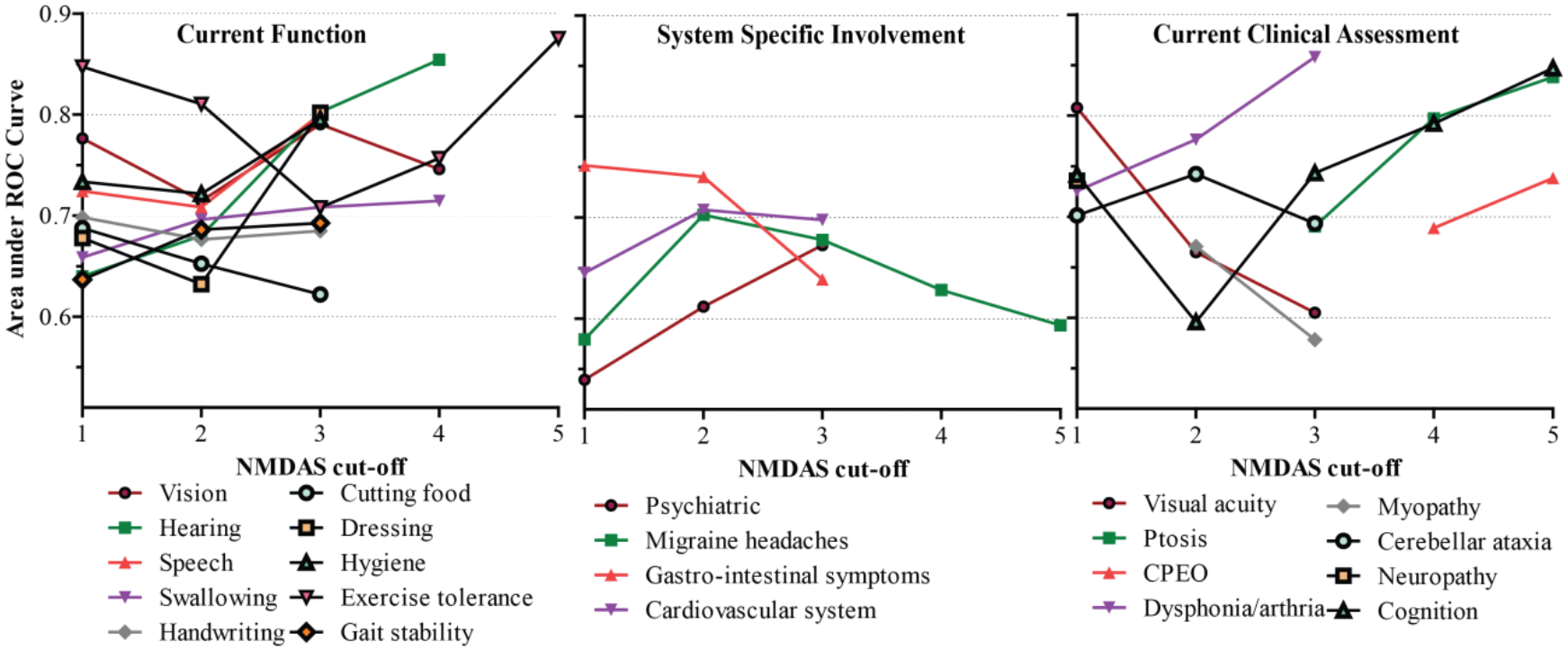


Figure 5.9 Area under the ROC curve (AUC) for each phenotypic feature and various score cut-off points with all three predictors in the model (age, heteroplasmy, and deletion size).

The cut-off score is used to divide the cohort into two groups; those scoring the cut-off or above versus those scoring below. For each feature, the cut-offs tested were those with a minimum of 6 patients (approximately 10% of the cohort) in each group. Exercise tolerance, dressing, hearing, visual acuity, dysphonia/dysarthria, cognition, and ptosis achieve an accuracy above 80% using the ROC curve, however most features achieve 70% accuracy for at least one cut-off point. Some features are poorly predicted at any cut-off, for instance psychiatric involvement, migraine, and myopathy, and several features from the current function assessment. Hearing is much more accurately predicted at higher levels of dysfunction, whilst exercise tolerance is fairly well predicted across the spectrum.

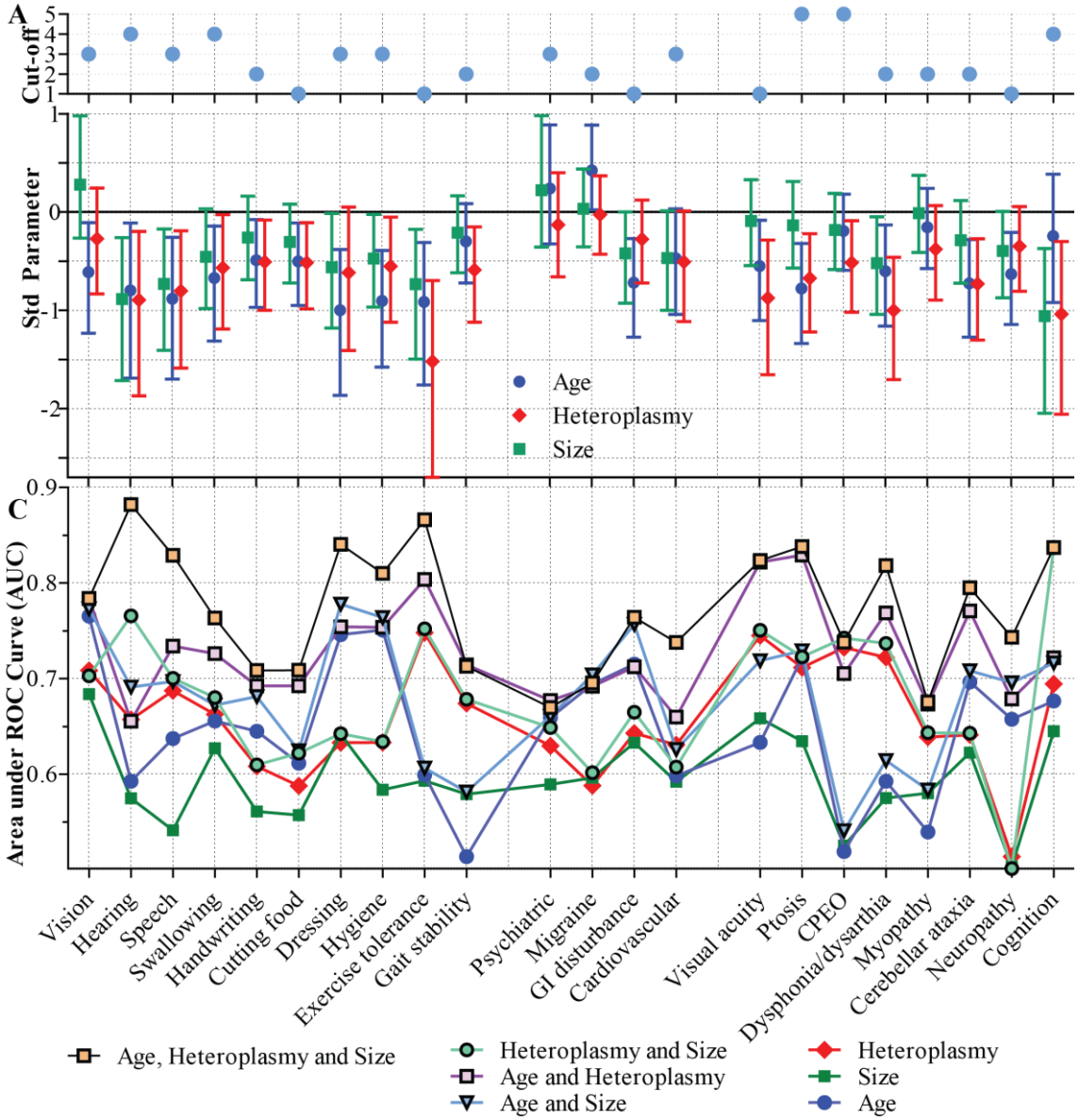


Figure 5.10 Standardised parameters and area under ROC curves (AUC) for each phenotypic feature from the NMDAS using age, urine heteroplasmy, and deletion size as predictors.

(A) Optimal cut-off point for each NMDAS feature, by maximizing AUC for the model with all three predictors. Cohort is dichotomised into those scoring the cut-off or above against those scoring below the cut-off. **(B)** Standardised parameters with 95% confidence intervals for the binary logistic model using the optimal cut-off with all three predictors. Parameters are statistically significant if the confidence interval does not cross the line at Y = 1. **(C)** The AUC using the optimal cut-off for each lone predictor and combination of predictors. Nine features achieve an AUC over 0.8, indicating strong predictive power. In almost all cases the regression with all three predictors is optimal, though in CPEO and cognition age is not predictive. Deletion size is generally the weakest predictor, though age is noticeably poorer in predicting gait stability and myopathy. Age is the principal predictor of migraine and neuropathy, and noticeable predictive of vision, dressing, hygiene, GI disturbance, ptosis, and ataxia. Heteroplasmy is better than the other lone predictors for many features, and significantly predictive for 15 out of 22 features. Migraine is the only feature that decreases in severity with age.

Appendix XV details the parameters required for risk calculation in the proportional-odds, including intercepts, and Appendix VIII is an explanation of how to calculate the risk for a given age and heteroplasmy with these parameters.

Models were also generated with a further predictor included in addition to age, heteroplasmy, and deletion size, indicating the presence or absence of specific genes in the deletion. Both *MT-CYB* and *MT-CO* genes were tested. Cerebellar ataxia was the only feature significantly predicted by *MT-CYB* gene inclusion (P = 0.0151). No features were significantly predicted by *MT-CO* gene inclusion.

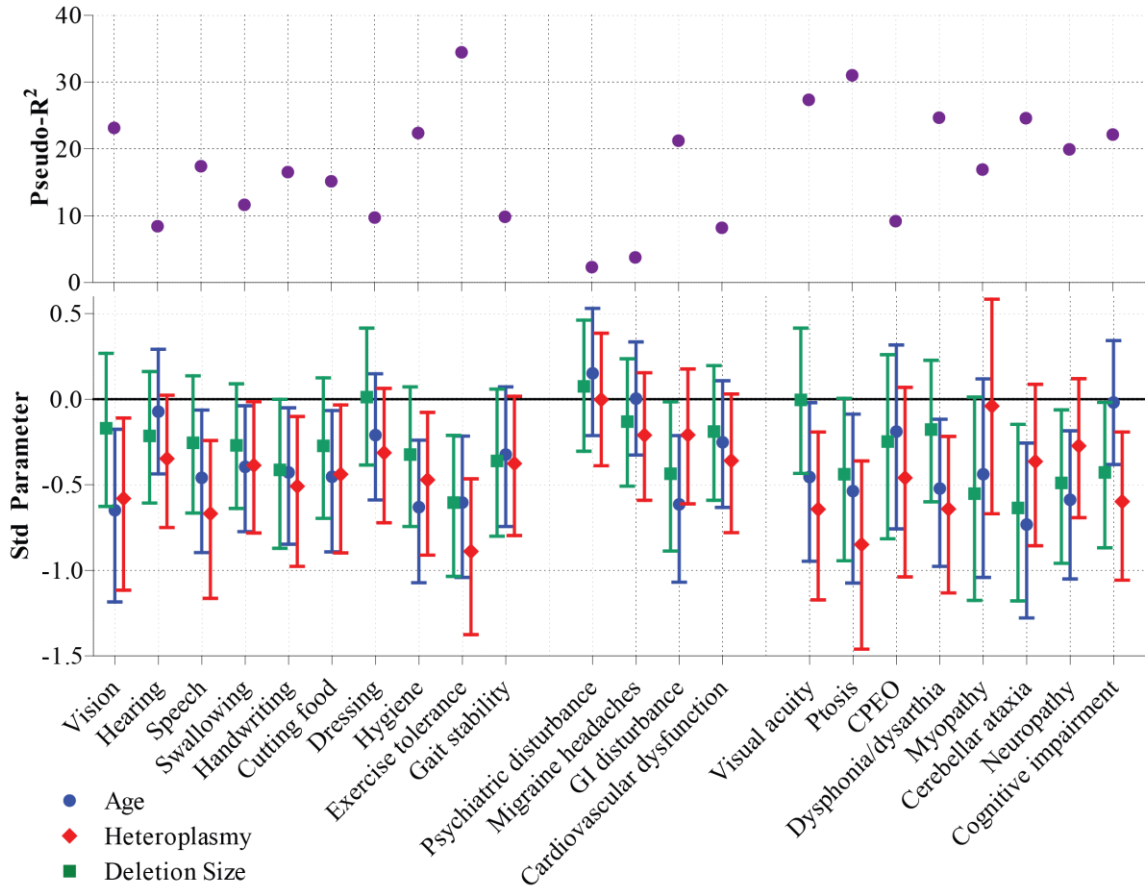


Figure 5.11 Pseudo-R^2 and standardised parameters for the logistic modelling of individual phenotypic features.

Logistic models include age, heteroplasmy, and deletion size as predictors. Pseudo-R^2 indicates that exercise tolerance and ptosis are the most predictable features using this model. Standardised parameters are statistically significant if the 95% confidence interval does not cross zero. Deletion size is only statistically significant for exercise tolerance, ataxia, neuropathy, GI disturbance, and cognitive impairment. Heteroplasmy is in general the strongest predictor of the three and is significant for 11 of the 22 features, including almost all of the current function assessment (section I). Cerebellar ataxia is the most predictable by age, indicating this is particularly progressive.

5.5. Discussion

Single large-scale mtDNA deletions are a common cause of mitochondrial disease, and the challenges in understanding the progression of disease for these patients highlights many of the challenges observed with other mtDNA deletions.

The significant finding of the work presented in this chapter is that SKM mtDNA heteroplasmy levels, mtDNA deletion size and mtDNA deletion location are all important in understanding the expression and progression of clinical disease in patients with single large-scale mtDNA deletions.

Two major factors have helped to reach this understanding. The first is a validated rating scale that covers the major clinical features of mitochondrial disease, the NMDAS. This, as a quantitative measure of mitochondrial disease burden, permits the study of disease progression in a manner which was previously unfeasible. Though I have presented the correlation between traditional phenotypic classification and NMDAS scores in the Newcastle cohort, the difficultly in using a discrete phenotype classification is clearly apparent. Longitudinal modelling of successive NMDAS assessments for patients demonstrates the benefits of a quantitative approach. It also highlights the dynamic nature of the patient phenotype.

The second major factor is the use of statistical techniques, including multiple regression analyses that I have applied to both the Newcastle cohort data and those previously published in the literature. This includes use of the Box-Cox transformation to identify optimal transformations to normality, which allows the use of more powerful parametric statistical tests where appropriate. Multiple regression, however, is of high valuable in this particular study; where there is inter-correlation between the predictive factors, such techniques are required to correctly identify significant findings.

5.5.1. Predictability of disease progression

The longitudinal modelling shows that disease burden and progression is predicted by muscle mtDNA heteroplasmy level, mtDNA deletion size and the location of the mtDNA deletion within the mitochondrial genome. These findings are supported by other findings from both the Newcastle cohort and the meta-analysis; both phenotype and age at onset are predicted by mtDNA deletion size and mtDNA heteroplasmy, and also heteroplasmy is a significant predictor of phenotype in the cohort of patients with the common 4,977bp mtDNA deletion. Previous studies have been contradictory regarding the utility of these factors as predictors. It has been reported that both

phenotype and age at onset were dependent on mtDNA deletion size, but not significantly related to mtDNA heteroplasmy (Yamashita *et al.*, 2008), however it has also been shown that the clinical phenotype was related to mtDNA heteroplasmy, but unrelated to mtDNA deletion size (although age at onset was related to both factors) (López-Gallardo *et al.*, 2009). In a study specifically addressing disease progression it was reported that SKM mtDNA heteroplasmy and mtDNA deletion size were not of utility in disease progression prediction (Auré *et al.*, 2007). More seminal studies also reported inconsistently on the utility of these predictors in predicting phenotype or age at onset (Holt *et al.*, 1989b; Moraes *et al.*, 1989; Goto *et al.*, 1990a). However, we found that multiple regression identifies consistent correlation between disease phenotype or progression and both mtDNA heteroplasmy and deletion size. Furthermore, multiple regression of previously published data is revealing; for example, whilst López-Gallardo and colleagues found that mtDNA deletion size is not predictive of phenotype ($P = 0.3953$), multiple regression with both deletion size and heteroplasmy shows that deletion size is indeed a significant predictor of phenotype ($P = 0.0330$); in this case, the negative correlation between heteroplasmy and deletion size leads to the masking of the effect of deletion size when simple linear regression is used.

5.5.2. Inter-correlation of heteroplasmy, deletion size, and location

The correlations observed between the predictive factors are themselves of interest. López-Gallardo and colleagues noted the negative correlation between muscle heteroplasmy levels and mtDNA deletion size (López-Gallardo *et al.*, 2009). From this they surmised that shorter mtDNA deletions do not have a replicative advantage, since one would expect in this case large deletions to be found at higher heteroplasmy levels, not lower. However, I would speculate that the observed correlation between muscle mtDNA heteroplasmy and mtDNA deletion size is more likely a reflection of the spectrum of disease that presents with a clinically-recognisable phenotype within the population, and is not an outcome of an intrinsic biochemically driven relationship. This speculation is supported by the results presented here on disease progression; small mtDNA deletions at low heteroplasmy levels would not be expected to be pathogenic and therefore patients with such deletions would not present clinically with symptoms. This is also consistent with the observation that potentially pathogenic mtDNA mutations are widespread at sub-threshold levels in the non-diseased population (Elliott *et al.*, 2008).

5.5.3. Deletion location and pathogenicity

I have shown that deletions including all or part of the *MT-CYB* gene are associated with faster disease progression. This is consistent with the report by López-Gallardo *et al.* that deletion of the *MT-CYB* gene was linked to a more severe phenotype (López-Gallardo *et al.*, 2009). The reason for this remains uncertain, but could potentially indicate that complex III deficiency has a particularly important role in the disease mechanism.

I also showed that *MT-CO* gene deletion (that is, deletion of part or all of any of *MT-CO1*, *MT-CO2* or *MT-CO3)* is predictive of COX-deficient fibre density. Though previous studies reported no such link (Goto *et al.*, 1990a; Oldfors *et al.*, 1992), using multiple regression to reanalyse these studies reveals the same trend in all cases, indeed at statistical significance when the study is large enough (Goto *et al.*, 1990a).

It has been widely accepted for some time that the root of the pathogenicity of mtDNA deletions is the deletion of mitochondrial tRNA (mt-tRNA) genes (Schon *et al.*, 2012). This hypothesis is somewhat challenged by the results presented in this study, consequently I will discuss the studies which have contributed to this hypothesis and re-evaluate them in the light of the current study.

Mita *et al.* showed that COX deficiency in SKM fibres was associated with low level wild-type mtDNA and high levels deleted mtDNA, which is a non-contentious finding (Mita *et al.*, 1989). They also found that the distribution of mRNA demonstrated that the deleted mtDNA species was undergoing transcription. However, although COX subunit IV (nuclear encoded) was present using immunocytochemistry, COX subunit II (mitochondrially encoded) was absent; this was despite the fact that COX subunit II was not part of the deleted region. This showed that a deletion could affect mitochondrially encoded translation products that were outside the deleted region. However, it should be noted that this study was on a single KSS patient.

A study published the following year showed that deletion of a single tRNA gene could compromise mitochondrial protein translation (Nakase *et al.*, 1990). Using cloned fibroblasts from two KSS patients with different deletions they showed that the mutant DNA was transcriptionally active and produced RNA species, though the fusion mRNA spanning the deleted region in each case was not translated. Their conclusion was that the biochemical defects were due to lack of translation of the mtDNA encoded proteins in mitochondria harbouring mutant mtDNA, most likely due to tRNA insufficiency. They concluded that the mutant DNA species must be segregated from the wild-type for

lack of complementation to occur (i.e. tRNA from the wild-type mtDNA supplementing the lack of tRNA from the deleted mtDNA species).

Hayashi *et al.* (Hayashi *et al.*, 1991) introduced deletions into HeLa cells, and found that below 60% heteroplasmy there was normal translation of proteins in mitochondria, but above this level mRNA levels mismatched protein levels and thus showed that translation was impaired. An interesting conclusion that can be drawn from this study is that there is functional complementation of mutant mtDNA by wild-type in the same mitochondrion, which contrasted with the conclusion of Nakase *et al.*; however, this complementation breaks down above approximately 60% heteroplasmy and overall translation is impaired. They suggested that the lack of complementation cited by Nakase *et al.* was most likely due to exceeding the threshold for normal translation, and competition for tRNA interfering with translation throughout the mitochondrion. This work was consistent with an earlier study by Shoubridge *et al.* that showed that deleted mtDNA species are functionally dominant over wild-type, since they accumulate in ragged red fibres and impair biochemical activity despite normal wild-type levels of mtDNA in these cells (Shoubridge *et al.*, 1990).

More recently this functional complementation has been confirmed by a study in which two cell lines, each with distinct and non-overlapping mtDNA deletions, were fused, and restoration of mitochondrial protein translation was demonstrated, despite the two mtDNA species existing in discrete nucleoids within the same mitochondrion (Gilkerson *et al.*, 2008).

In summary, the conclusion that pathogenicity is based on tRNA insufficiency is predicated on a lack of correlation between the quality (size and location) of the deletion and the resulting biochemical defect and clinical phenotype. However, in this study I have demonstrated that both the biochemical defect (COX deficiency) and the clinical phenotype (disease burden) are correlated with both deletion size and location, which is evidence that tRNA insufficiency is not entirely responsible for the pathogenicity of deletions. However, a more comprehensive data set would be required to elucidate the relative importance of mt-tRNA genes, specific oxidative phosphorylation protein genes, and other potential pathogenic mechanisms as regards biochemical defects and the resulting clinical phenotype.

5.5.4. Pathogenic threshold level

An important implication of this study regards is that threshold level for pathogenicity of a mtDNA deletion (Rossignol *et al.*, 2003). Since the biochemical defect and the resulting clinical phenotype are dependent on the deletion size and location, the threshold for phenotypic expression at the cellular level would be expected to be dependent on the location and size of the deletion.

5.5.5. Specific phenotypic features

It has been shown that many individual phenotypic features are predictable by either age, heteroplasmy, or deletion size, or some combination of these factors. Exercise tolerance and ptosis are notably predictable, and this is not unexpected since neuromuscular features are prominent in disease associated with single large-scale mtDNA deletions. There are several other features that are strongly predictable, for instance dysphonia/dysarthria, which is a noted feature of single deletion disease (Auré *et al.*, 2007), cerebellar ataxia, and cognitive impairment.

Though most features demonstrated similar outcomes in the two approaches to logistic modelling (binary and multi-level proportional odds), hearing is noticeably different in the two, predicted by the dichotomous model but not by the multi-level model. However, as was noted in Figure 5.9, hearing is not predicted well when looking at low cut-off points in the NMDAS score, and significant effects of the predictors may only be apparent at higher levels of impairment.

Hearing is also interesting from another perspective, in that the improvement seen in moving from a single- or two-predictor model to a model with all three predictors is very large; in particular, adding age into the model with heteroplasmy and deletion size improves the AUC from 0.7 to over 0.85. This is a strong reinforcement of the necessity for multiple regression techniques.

CPEO is a second feature that was not well predicted by the multi-level model but that is reasonably predicted by the dichotomous model; the reason for this is found in the optimal cut-off score for CPEO of 5, which is more or less the only level between which prediction can be made for CPEO. However, both 4 and 5 scores are grouped together in the multi-level model. This emphasises the importance of thorough analysis of data from a number of perspectives, and also highlights the importance of optimising the analysis on question by question basis. CPEO is also curious in that heteroplasmy alone is almost entirely responsible for the prediction of this feature; age and deletion size are

not contributory. In this CPEO stands out from the other features. This is surprising, since the very nature of CPEO is progressive, and thus association with age should be expected. However, this finding may be related to the rating scale itself; CPEO is a ubiquitous feature of single large-scale mtDNA deletions, and over 50% of the cohort score maximally on this feature, suggesting that if CPEO continues to be progressive for these patients the severity range is inadequate for this group.

Age appears to be almost the sole contributor to vision decline (from the current function part of the NMDAS), psychiatric disturbance, and migraine, suggesting that these features may be more related to general ageing than factors due to mtDNA deletion. Indeed all of these features are indeed common in the general population and associated with ageing. Interestingly migraine is the only NDMAS feature to significantly decrease with age, and this reflects the reported pattern in the general population (Dahlof *et al.*, 2009). Psychiatric disturbance is the only other feature to be negatively associated with age; again, this is the reported trend in the population (Jorm, 2000).

Heteroplasmy and age are shown to be reasonably consistent predictors of most phenotypic features, however deletion size is generally less effective and less consistent. This has already been seen in the longitudinal modelling, but it is assuring to see the same pattern at the level of the individual phenotypic features. That deletion size is predictive but secondary to heteroplasmy and age for many phenotypic features does suggest that the pathogenic nature of the deletions may be to a large extent deletion-size independent, and may be due the effect of tRNA gene deletion. However, the growing understanding of the role of supercomplexes (Vartak *et al.*, 2013) suggests a second pathway by which a deletion of any of the OXPHOS genes may affect the structural integrity and function of other complexes further up or down the respiratory chain.

The location of the deletion did not prove to be a significant predictor of any phenotypic features excepting cerebellar ataxia. Interpretation of this finding as statistically significant would be naïve, and no firm conclusions can be drawn from this. However, the lack of significant association is not of great concern; the dataset used for these analyses is summary, containing only a single predictive point per patient; with heteroplasmy, age, and deletion size in the model already it is no great surprise that the model lacks statistical power to expose any significant relationships between deleted genes and phenotypic features. This is an important area for future research with a

larger dataset however; though the study on total disease burden is informative regarding the pathology of the mtDNA deletions, characterising and understanding association between specific phenotypic features and the properties of the deletion would be fascinating indeed.

Though I have not displayed any graphical risk profiles of the individual feature modelling, the purpose of this section has not been to present the model outputs but to discuss the utility of the various predictors for understanding phenotypic feature development. However, the modelling of m.3243A>G did present models for several features, and the model outputs for this section would be similar in form. All models from this section of work will be available at https://research.ncl.ac.uk/mitoresearch as an aid to understanding the outcomes of this work.

5.5.6. Limitations and future work

Although this study provide new insights into disease progression in patients with single, large-scale mtDNA deletions, much further work is still required.

Firstly, both mtDNA deletion size and the location parameter used in the longitudinal modelling (*MT-CYB* gene deletion) are perhaps merely proxies for the underlying pathogenic nature of the mtDNA deletion. A more nuanced characterization of these mtDNA deletions may better predict pathogenesis and mitochondrial disease progression. It is unclear, for instance, why *MT-CO* gene deletion is associated with more severe COX deficiency in cells but it does not have a clear modulatory effect on overall disease progression. However, expanding the model to characterise the deletions more fully would potentially involve a large number of parameters, in which case a larger dataset would be required to achieve reasonable statistical power.

Secondly, it is notable that we do not have any predictor that encapsulates the sometimes multi-system nature of disease. In this regard, Auré *et al.* found that the presence of the mutation in blood was predictive of a neurological phenotype (Auré *et al.*, 2007), and a similar trend in urine, a result confirmed by Blackwood *et al.* with regard to severe early-onset disease as compared to milder phenotypic presentation (Blackwood *et al.*, 2010). These reports suggest that blood or urine mtDNA heteroplasmy levels, in tandem with SKM mtDNA heteroplasmy, may lead to an improved prediction of disease prognosis. For this to be included in the model blood or urine samples will need to be collected for the patients included in the model, but these are not available for the Newcastle patient cohort at the current time.

Thirdly, the impact of other mitochondrial DNA rearrangements has not been considered in this study, for instance duplications (Poulton *et al.*, 1989). However, it is not generally thought that duplications have a pathogenic effect.

Fourthly, though most of the findings from our cohort are corroborated by the meta-analysis, there are differences in the datasets, and in particular the correlations between predictors are on the whole stronger in the Newcastle cohort than in the meta-analysis. These differences may arise in part from the fact that the data from the Newcastle cohort are more homogeneous, but also because the makeup of this cohort is somewhat different; there are relatively few patients with the KSS phenotype as compared to the meta-analysis.

Finally, though I do not use Bonferroni adjustment in the analyses for reasons well documented in the literature (particularly where *a priori* hypotheses are under test) (Thomas, 1998), the large number of statistical tests employed do open up the potential for more frequent type I statistical errors. However, consistent findings in both the Newcastle cohort and the meta-analysis provide support for firm conclusions to be drawn. Multiple testing is particularly relevant in the section on the prediction of specific phenotypic features. However, though caution must be taken in interpretation of statistical significance, the analytical techniques employed remain valid.

5.5.7. Web tool

The finding that deletion size, heteroplasmy, and deletion location are predictors of disease progression is important for all clinicians looking after patients with single, large-scale mtDNA deletions. A web based tool is therefore under development (https://research.ncl.ac.uk/mitoresearch) to support clinicians in their management of these patients. This tool currently uses the predictive longitudinal model detailed in this chapter to provide progression graphs, but is intended to also provide individual phenotypic feature risk profiles.

5.5.8. Conclusion

I have demonstrated that muscle heteroplasmy, deletion size and deletion location are predictive of disease severity and progression in single large-scale mtDNA deletions. Using these predictors together with the NMDAS, a quantitative measure of total disease burden, longitudinal modelling of disease progression can be performed both at a population level and for individual patients, as well as calculation of risk of developing specific phenotypic features. The end result is that advice and care plans for

patients can be given on an individual basis, and discussions about the expected course of their disease burden can be provided.

Chapter 6

Chapter 6. **Discussion**

The principal aim of these studies has been to further understanding of the nature and progression of mtDNA disease. Despite the fact that the first mtDNA mutations were described over twenty five years ago (Holt *et al.*, 1988), this remains challenging.

There are several areas in which these studies have presented novel data and significantly progressed knowledge of disease progression in mtDNA disease. The study on m.3243A>G has expanded the current understanding regarding the predictive nature of disease associated with the mutation, both in terms of overall disease burden and also specific phenotypic features; it has provided new understanding of the dynamic (or not so dynamic) nature of heteroplasmy in various tissues of patients with the m.3243A>G mutation; it has generated several observations on sex differences in disease phenotype and heteroplasmy; it has presented novel data on the altered live-birth sex ratio; and it has shown that blood may be a more appropriate source of heteroplasmy than urine for use as a prognostic measure, in contradiction to long standing belief. Regarding single large-scale mtDNA deletions, there are several novel observations on stronger than previously understood connections between the genetic, biochemical, and clinical phenotype that shed new light on the pathology of large scale mtDNA deletions; and, as for m.3243A>G, the disease modelling provide useful prognostic data for clinicians in the care and management of patients.

These studies have convincingly demonstrated the importance of simultaneous consideration of multiple factors when trying to understand the link between genotype and phenotype in mitochondrial disease, due to population level inter-correlation of all the predictive factors including age, heteroplasmy, and in the single large-scale deletions, deletion size and location. Without proper consideration of these relationships either true relationships between putative predictive factors and disease burden and progression are obscured, or false causal relationships are suggested.

They have also clearly demonstrated the utility of the NMDAS as a tool for recording and understanding disease burden and progression in mitochondrial disease. This clinical tool has enabled the development of the prognostic models for overall disease burden and individual phenotypic feature progression, which are valuable tools for clinicians in the care and management of patients.

There are numerous avenues to explore in progressing the work begun in this thesis. Both in m.3243A>G and single large-scale mtDNA deletions factors to improve the prediction models have been identified, whether mtDNA copy number, alternative sources of heteroplasmy data, or investigations into the wider genetics and biochemical activity of the cells in mtDNA disease patients, for instance the role of tRNA synthetases. In particular, the limitations of heteroplasmy as a predictor of disease burden and progression in m.3243A>G have been clearly demonstrated, and improving on this prognostic ability is a critically important area for future research. The limitations of urine heteroplasmy in particular as a prognostic device have been clearly demonstrated, and the analysis strongly suggests that though urine has good properties for diagnostic identification of the presence of the m.3243A>G mutation, it performs poorly as an indicator of average heteroplasmy for a patient, and the existing recommendations for its use in this manner should be reviewed. The dynamism of heteroplasmy, particularly in patients with the m.3243A>G mutation, is also a rich area for exploration.

The scope for further research is much wider however. This study has considered only two of the most common of the large number of mtDNA mutations, and the scope of mitochondrial disease is far wider indeed than even these, encompassing a huge number of nDNA mutations. Extending the techniques and understanding gained in these mutations to a wider array of genetic defects will be important future work.

The models developed here are open for a great deal of further development. For instance, the logistic regression models presented are only analyses of the prevalence of each feature within the population of patients with the respective mutations; though this is an important characterisation of the phenotypic presentation of m.3243A>G and single large-scale mtDNA deletion disease, from the perspective of clinical care it is imperative to develop models to predict their expected progression of specific phenotypic features for individual patients, in a similar manner to the models provided for overall disease burden, by incorporating random effects and repeated measures analysis. This is important future work to be conducted on these cohorts of patients.

A final important consequence of this modelling is that it has a very practical use as a baseline from which to measure the effectiveness of any future treatments for mitochondrial disease. Without a thorough understanding of the expected progression of disease then interventions cannot be confidently assesses for efficacy.

In conclusion, much progress has been made in understanding disease progression in patients with the m.3243A>G mutation and single large-scale mtDNA deletions, but much progress remains to be made, not only in these particular mutations, but in mitochondrial disease in general.

Appendices

Appendix I The NMDAS Assessment.

THE NEWCASTLE MITOCHONDRIAL DISEASE ADULT SCALE (NMDAS)

Name: _____
Date of birth: _____
Date of assessment: _____
Height: _____
FVC - 1st attempt _____
FVC - 2nd attempt _____
FVC - 3rd attempt _____

Section I- Current Function

Rate function over the preceding **4 week period, according to patient and/or caregiver** interview only. The clinician's subjective judgement of functional ability should **not** be taken into account.

1. **Vision** with usual glasses or contact lenses
0. Normal.
1. No functional impairment but aware of worsened acuities.
2. Mild - difficulty with small print or text on television.
3. Moderate - difficulty outside the home (eg bus numbers, road signs or shopping).
4. Severe - difficulty recognising faces.
5. Unable to navigate without help (eg carer, dog, cane).

2. **Hearing** with or without hearing aid
0. Normal.
1. No communication problems but aware of tinnitus **or** deterioration from prior 'normal' hearing.
2. Mild deafness (eg missing words in presence of background noise). **Fully** corrected with hearing aid.
3. Moderate deafness (eg regularly requiring repetition). **Not fully** corrected with hearing aid.
4. Severe deafness - poor hearing even with aid (see 3 above).
5. End stage - virtually no hearing despite aid. Relies heavily on non-verbal communication (eg lip reading) **or** has cochlear implant.

3. **Speech**
0. Normal.
1. Communication unaffected but patient or others aware of changes in speech patterns or quality.
2. Mild difficulties - usually understood and **rarely** asked to repeat things.
3. Moderate difficulties - poorly understood by strangers and **frequently** asked to repeat things.
4. Severe difficulties - poorly understood by family or friends.
5. Not understood by family or friends. Requires communication aid.

4. **Swallowing**
0. Normal.
1. Mild - sensation of solids 'sticking' (occasional).
2. Sensation of solids 'sticking' (most meals) **or** need to modify diet (eg avoidance of steak/salad).
3. Difficulty swallowing solids - affecting meal size or duration. Coughing, choking **or** nasal regurgitation infrequent (1 to 4 times per month) but more than peers.
4. Requires adapted diet - regular coughing, choking, **or** nasal regurgitation (more than once per week).
5. Requiring enteral feeding (eg PEG).

5. **Handwriting**
0. Normal.
1. Writing speed unaffected but aware of increasing untidiness.
2. Mild – Has to write slower to maintain tidiness/legibility.
3. Moderate – Handwriting takes at least twice as long **or** resorts to printing (must previously have used joined writing).
4. Severe – Handwriting mostly illegible. Printing very slow and untidy (eg 'THE BLACK CAT' takes in excess of 30 seconds).
5. Unable to write. No legible words.

6. **Cutting food and handling utensils** (irrespective of contributory factors – eg weakness, coordination, cognitive function etc. This is also true for questions 7-10)
0. Normal.
1. Slightly slow and/or clumsy but **minimal** effect on meal duration.
2. Slow and/or clumsy with extended meal duration, but no help required.
3. Difficulty cutting up food and inaccuracy of transfer pronounced. Can manage alone but avoids problem foods (eg peas) or carer typically offers minor assistance (eg cutting up steak).
4. Unable to cut up food. Can pass food to mouth with great effort or inaccuracy. Resultant intake minimal. Requires major assistance.
5. Needs to be fed.

7. **Dressing**
0. Normal.
1. Occasional difficulties (eg shoe laces, buttons etc) but no real impact on time or effort taken to dress.
2. Mild – Dressing takes longer and requires more effort than expected at the patient's age. No help required.
3. Moderate - Can dress unaided but takes at least twice as long and is a major effort. Carer typically helps with difficult tasks such as shoe laces or buttons.
4. Severe – Unable to dress without help but some tasks completed unaided.
5. Needs to be dressed.

8. **Hygiene**
0. Normal.
1. Occasional difficulties only but no real impact on time or effort required.
2. Mild – hygienic care takes longer but quality unaffected.
3. Moderate - bathes and showers alone with difficulty **or** needs bath chair / modifications. Dextrous tasks (eg brushing teeth, combing hair) performed poorly.
4. Severe - unable to bathe or shower without help. Major difficulty using toilet alone. Dextrous tasks require help.
5. Dependent upon carers to wash, bathe, and toilet.

9. **Exercise Tolerance**
0. Normal.
1. Unlimited on flat - symptomatic on inclines or stairs.
2. Able to walk < 1000m on the flat. Restricted on inclines or stairs - rest needed after 1 flight (12 steps).
3. Able to walk < 500m on the flat. Rest needed after 8 steps on stairs.
4. Able to walk < 100m on the flat. Rest needed after 4 steps on stairs.
5. Able to walk < 25m on the flat. Unable to do stairs alone.

10. **Gait stability**
0. Normal.
1. Normal gait - occasional difficulties on turns, uneven ground, or if required to balance on narrow base.
2. Gait reasonably steady. Aware of impaired balance. **Occasionally** off balance when walking.
3. Unsteady gait. **Always** off balance when walking. **Occasional** falls. Gait steady with support of stick or person.
4. Gait grossly unsteady without support. **High likelihood** of falls. Can only walk short distances (< 10m) without support.
5. Unable to walk without support. Falls on standing.

Section II – System Specific Involvement

Rate function according to patient and/or caregiver interview and consultation with the medical notes. Each inquiry should take into account the situation for the preceding **12 month period** only, unless otherwise stated in the question.

1. **Psychiatric**
0. None.
1. Mild & transient (eg reactive depression) - lasting **less** than 3 months.
2. Mild & persistent (lasting **more** than 3 months) **or** recurrent. Patient has consulted GP.
3. Moderate & warranting specialist treatment (e.g. from a psychiatrist) - eg. bipolar disorder or depression with vegetative symptoms (insomnia, anorexia, abulia etc).
4. Severe (eg self harm - psychosis etc).
5. Institutionalised or suicide attempt.

2. **Migraine Headaches** During **the last 3 months**, how many days have headaches prevented the patient from functioning normally at school, work, or in the home?
0. No past history.
1. Asymptomatic but past history of migraines.
2. One day per month.
3. Two days per month.
4. Three days per month.
5. Four days per month or more.

3. **Seizures**
0. No past history.
1. Asymptomatic but past history of epilepsy.
2. Myoclonic or simple partial seizures only.
3. Multiple absence, complex partial, or myoclonic seizures affecting function **or** single generalised seizure.
4. Multiple generalised seizures.
5. Status epilepticus.

4. **Stroke-like episodes** (exclude focal deficits felt to be of vascular aetiology)
0. None.
1. Transient focal sensory symptoms only (**less** than 24 hours).
2. Transient focal motor symptoms only (**less** than 24 hours).
3. Single stroke-like episode affecting one hemisphere (**more** than 24 hours).
4. Single stroke-like episode affecting both hemispheres (**more** than 24 hours).
5. Multiple stroke-like episodes (**more** than 24 hours each).

5. **Encephalopathic Episodes**
0. No past history.
1. Asymptomatic **but** past history of encephalopathy.
2. Mild - single episode of personality or behavioural change but retaining orientation in time/place/person.
3. Moderate - single episode of confusion or disorientation in time, place or person.
4. Severe – multiple moderate episodes (as above) **or** emergency hospital admission due to encephalopathy **without** associated seizures or stroke-like episodes.
5. Very severe - in association with seizures, strokes or gross lactic acidaemia.

6. **Gastro-intestinal symptoms**
0. None.
1. Mild constipation only **or** past history of bowel resection for dysmotility.
2. Occasional symptoms of 'irritable bowel' (pain, bloating or diarrhoea) with long spells of normality.
3. Frequent symptoms (as above) most weeks **or** severe constipation with bowels open less than once/week **or** need for daily medications.
4. Dysmotility requiring admission **or** persistent and/or recurrent anorexia/vomiting/weight loss.
5. Surgical procedures **or** resections for gastrointestinal dysmotility.

7. **Diabetes mellitus**
0. None.
1. Past history of gestational diabetes or transient glucose intolerance related to intercurrent illness.
2. Impaired glucose tolerance (in absence of intercurrent illness).
3. NIDDM (diet).
4. NIDDM (tablets).
5. DM requiring insulin (irrespective of treatment at onset).

8. **Respiratory muscle weakness**
0. FVC normal (≥ 85% predicted).
1. FVC < 85% predicted.
2. FVC < 75% predicted.
3. FVC < 65% predicted.
4. FVC < 55% predicted.
5. FVC < 45% predicted **or** ventilatory support for over 6 hours per 24 hr period (**not** for OSA alone).

9. **Cardiovascular system**
0. None.
1. Asymptomatic ECG change.
2. Asymptomatic LVH on echo **or** non-sustained brady/tachyarrhythmia on ECG.
3. Sustained or **symptomatic** arrhythmia, LVH **or** cardiomyopathy. Dilated chambers **or** reduced function on echo. Mobitz II AV block or greater.
4. Requires pacemaker, defibrillator, arrhythmia ablation, **or** LVEF < 35% on echocardiogram.
5. Symptoms of left ventricular failure **with** clinical and/or x-ray evidence of pulmonary oedema **or** LVEF < 30% on echocardiogram.

Section III – Current Clinical Assessment

Rate current status according to examination performed at **the time of** assessment

1. **Visual acuity** with usual glasses, contact lenses or pinhole.
0. CSD ≤ **12** (ie normal vision - 6/6, 6/6 or better).
1. CSD **≤ 18** (eg 6/9, 6/9).
2. CSD **≤ 36** (eg 6/12, 6/24).
3. CSD **≤ 60** (eg 6/24, 6/36).
4. CSD **≤ 96** (eg 6/60, 6/36).
5. CSD **≥ 120** (eg 6/60, 6/60 **or worse**).

2. **Ptosis**
0. None.
1. Mild ptosis - not obscuring **either** pupil.
2. Unilateral ptosis obscuring < 1/3 of pupil.
3. Bilateral ptosis obscuring < 1/3 **or** unilateral ptosis obscuring > 1/3 of pupil **or** prior unilateral surgery.
4. Bilateral ptosis obscuring > 1/3 of pupils **or** prior bilateral surgery.
5. Bilateral ptosis obscuring >2/3 of pupils **or** >1/3 of pupils **despite** prior bilateral surgery.

3. **Chronic Progressive External Ophthalmoplegia**
0. None.
1. Some restriction of eye movement (any direction). Abduction complete.
2. Abduction of worst eye incomplete.
3. Abduction of worst eye below 60% of normal.
4. Abduction of worst eye below 30% of normal.
5. Abduction of worst eye minimal (flicker).

4. **Dysphonia/Dysarthria**
0. None.
1. Minimal - noted on examination only.
2. Mild – clear impairment but easily understood.
3. Moderate – some words poorly understood and infrequent repetition needed.
4. Severe – many words poorly understood and frequent repetition needed.
5. Not understood. Requires communication aid.

5**. Myopathy**
0. Normal.
1. Minimal reduction in hip flexion and/or shoulder abduction **only** (eg MRC 4+/5).
2. Mild but clear proximal weakness in hip flexion and shoulder abduction (MRC 4/5). Minimal weakness in elbow flexion and knee extension (MRC 4+/5 - both examined with joint at 90 degrees).
3. Moderate proximal weakness including elbow flexion & knee extension (MRC 4/5 or 4 -/5) **or difficulty** rising from a 90 degree squat.
4. Waddling gait. **Unable** to rise from a 90 degree squat (=a chair) unaided.
5. Wheelchair dependent **primarily** due to proximal weakness.

6. **Cerebellar ataxia**
0. None.
1. Normal gait but hesitant heel-toe.
2. Gait reasonably steady. Unable to maintain heel-toe walking **or** mild UL dysmetria.
3. Ataxic gait (but walks unaided) **or** UL intention tremor & past-pointing. Unable to walk heel-toe – falls immediately.
4. Severe - gait grossly unsteady without support **or** UL ataxia sufficient to affect feeding.
5. Wheelchair dependent **primarily** due to ataxia **or** UL ataxia **prevents** feeding.

7**. Neuropathy**
0. None.
1. Subtle sensory symptoms **or** areflexia.
2. Sensory impairment only (eg glove & stocking sensory loss).
3. Motor impairment (distal weakness) **or** sensory ataxia.
4. Sensory ataxia **or** motor effects severely limit ambulation.
5. Wheelchair bound **primarily** due to sensory ataxia or neurogenic weakness.

8. **Pyramidal Involvement**
0. None.
1. Focal or generalised increase in tone or reflexes only.
2. Mild **focal** weakness, sensory loss or fine motor impairment (eg cortical hand).
3. Moderate hemiplegia allowing unaided ambulation **or** dense UL monoplegia.
4. Severe hemiplegia allowing ambulation with aids **or** moderate tetraplegia (ambulant).
5. Wheelchair dependant **primarily** due to hemiplegia or tetraplegia.

9. **Extrapyramidal**
0. Normal.
1. Mild and unilateral. Not disabling (H&Y stage 1).
2. Mild and bilateral. Minimal disability. Gait affected (H&Y stage 2).
3. Moderate. Significant slowing of body movements (H&Y stage 3)
4. Severe. Rigidity and bradykinesia. Unable to live alone. Can walk to limited extent (H&Y stage 4).
5. Cannot walk or stand unaided. Requires constant nursing care (H&Y stage 5).

10. **Cognition**
Patients undergo testing using WTAR, Symbol Search and Speed of Comprehension Test.
0. Combined centiles **100 or more.**
1. Combined centiles **60 - 99**
2. Combined centiles **30 - 59**
3. Combined centiles **15 - 29**
4. Combined centiles **5 - 14**
5. Combined centiles **4 or below.**

Appendix II Matlab resampling code.

```
function [ stats answers] = Do_Resample( no_resamples, data,
nColsToUse, nRepsToUse, dopower, drawfigs, remove_outliers,
ratio_check)
%work out dimensions. each column is a single GROUP containing
replicates
nCols = size(data,2);
nReps = size(data,1);
nSets = size(data,3);
%Don't use remove outliers anymore! Remove them before calling.
removed = 0;
ratio = 0;
n = 0;
if remove_outliers == 1,
      %Work out the acceptable limits for the data. The output of this
      is the limits variable

      std_dev_before = 0;
      std_dev_after = 0;
      [data removed ratio] = remove_outliers_using_iqr_1(data);
end

number_of_vals = sum(isnan(data(:)));
fprintf('%d of %d reps, %d of %d groups, %d points total, %d outliers
removed\n', nRepsToUse, nReps, nColsToUse, nCols,
nColsToUse*nRepsToUse, removed);
stats = []; %3 rows, for each of the above estimate
samples = zeros(2,no_resamples);%we just record the differences

if drawfigs==1,
       figure;
       hist(data(:,:,1));
end

%Note - this actual answer is calculated from the values we supply,
but really we should be using the answer calculated from ALL the
values. So be careful about requesting ratio_check.
actual_answer =  nanmean(nanmean(data(:,:,1)));
answers = zeros(1,no_resamples);
for rs = 1:no_resamples;
        actualSet = randi(nSets, 1, 1);
        %Do it col by col, so that we can remove
        actualCols = randi(nCols,1,nColsToUse); %Multiple columns for
the group dist
        actualReps = randi(nReps,1,nRepsToUse); %Some of the summaries
use the actual reps
        vals = data(actualReps,actualCols,actualSet); %Both cols
        if ratio_check==1,
           vals = vals - actual_answer; %subtract the true ratio, so
that we are comparing relatively
        end

        switch dopower,
            case 1,
            vals = 10.^vals;
            case 2,
            vals = 1-10.^vals;
            case 3,
            vals = 10.^(vals+2); %This is for the B2M MT-ND1 assay
that is 100 fold different.
            case 4,
```

```
            vals = vals+2; %This is for the B2M MT-ND1 assay that is
100 fold different, looking at logs.
        end

%The groups are arranged in cols, so we need to take the mean of each
col and then look at std.
        samples(1,rs) = nanstd(nanmean(vals,1)); %rawgroup
        samples(2,rs) = nanmean(nanstd(vals)); %rep
        samples(4,rs) = nanmean(nanmean(vals,2)); %actual value.
        answers(rs) = samples(4,rs);
    end

%For the overall variance, this is tricky... Samples4 contains the
actual values we come out with each time.
%We need to look at the variance within this. So we actually need to
resample the resamples.... to get a variance for these
    x = samples(4,:);
    no_x = size(x,2);
    %just do 1000 for this
    overall_std = zeros(1,1000);
    for i = 1:1000,
        overall_std(1,i) = nanstd(x(randi(no_x,1,no_x)));
    end

    if no_resamples == 1,
%If we are just doing one resample, we pass back all the values as the
answer
        answers = vals';
    else
samples(3,:) = abs(sqrt(samples(1,:).^2 -
samples(2,:).^2/nRepsToUse)); %this is the adjusted VARIANCE, which
measures the GROUP. We need to take the square root to get the final
answer.
        stats(1,1:4) = calculate_rep_group_stats(samples(2,:),
nRepsToUse); %rep variance
        stats(1,5:8) = calculate_rep_group_stats(samples(3,:),
nColsToUse); %group variance
        stats(1,9:12) = calculate_rep_group_stats(samples(4,:), 0);
%actual values
        stats(1,13:16) = calculate_rep_group_stats(overall_std, 0);
%overall variance
        stats(1,17) = removed;
        stats(1,18) = ratio(1,:); %Note - this takes the ratio of the
front page of data. The second should be zeros anyway...
        stats(1,19) = nReps - number_of_vals; %This will give is
number of rows minus any NaNs that we have removed, so how many good
vals we have.
    end
function output = remove_outliers_using_iqr(data)
%Removes outliers using a threshold level that detects changes to the
%standard deviation above and beyond acceptable limits
sz = size(data,1);
mul = 2;
prcs = prctile(data, [25 75]);
iqr = prcs(2) - prcs(1);
lims = [(prcs(1) - iqr*mul) (prcs(2) + iqr*mul)];
sum(data < lims(1) | data > lims(2))
data(data < lims(1) | data > lims(2)) = NaN;
output = data;
```

Appendix III m.3243A>G Heteroplasmy data, with sequential measurements.

For each patient, the table lists the heteroplasmy data available and the age at sample. The age column records the age of the first sample, and age diff records the time in years between two samples(where two samples are available). A blank age diff implies only one sample is available; 0 implies repeated samples taken very close together in time. The number of NMDAS assessments for each patient is also shown.

	Urine Heteroplasmy				Blood Heteroplasmy				Muscle Heteroplasmy				
ID	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	No. NMDAS
1	34.8	10.8	35	32	34.8	9.8	6	7	45.6		52		11
2	46.2	10.5	93	66	52.9	3.8	20	19					6
3	37.8	10.1	29	15	37.1	10.8	18.5	13					3
4	62.9	9.4	49	49	72.3		10						12
5	29.5	9	80	88	36.9		26		34.1	4.1	67	91	7
6	50.7	8.8	74	67					54.9	0.2	62	64	14
7	47.1	8.8	54	48	55.9		5		47.1		53		9
8	44.4	8.4	8	11	44.4	10.3	2	2	44.4		11		10
9	17.8	8.4	65	32	17.8	8.4	8	13					2
10	51.3	8.1	42	59	50.4	6.3	8	6					12
11	30.7	7.9	62	89	30.7	7.7	15	19	30.5		67		5
12	37.1	7.9	72	60	37.1		14		41.9		69		3
13	52.5	7.9	52	44	60		13						11
14	50.8	7.8	20	33	50.8		9		54.7	4	60	62	3
15	41	7.6	82	56									9
16	21.2	7.5	77	68	21.2	7.5	64	53					13
17	42.4	7.5	82	87	42.4		28		46.2		26		7
18	16	7	73	73	16	7	48	48					5
19	40.1	6.9	75	79	46.1	3.3	12	11	40.3		57		6
20	39.2	6.9	50	61	46.2		13						9
21	46.8	6.7	23	22	44.3	8.8	9	11					4
22	65.3	6.6	6	11	65.3	6.6	3	4					1
23	28.7	6.5	4	3	35.1		1						4
24	25.6	5.3	57	53	25.6	5.3	38	38					2
25	41.2	5.2	87	86	41.2	5.2	37	25					8
26	43.7	5	83	74	43.7		16		43.7		53		3
27	23.7	4.9	98	93	23.8	4.9	48	41					8
28	30.3	4.9	44	44	30.3	4.9	7	7	30.7		24		7
29	44	4.4	38	52	44	4.4	27	12	48.4		73		6
30	20.4	4.4	82	72	20.4		46						7
31	15.6	4.2	64	55	15.6	4.2	41	40					1
32	43.3	4.2	48	51	43.3		5		43.3		46		6
33	40.3	4	43	72	40.3	4	21	19	41	2.5	73	78	8
34	15.7	4	85	72	15.7	4	45	42	19.8		72		6
35	43.9	3.9	72	63	43.9		16						10
36	43.3	3.7	52	57	43.3		8		43.3		50		2
37	52.9	3.6	39	59	52.9	3.6	16	16	52.9		78		2

	Urine Heteroplasmy				Blood Heteroplasmy				Muscle Heteroplasmy				
ID	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	No. NMDAS
38	19.5	3.5	73	81	19.5	3.5	50	45					5
39	45.7	3.1	80	89	45.5		23						10
40	43.5	3	6	6	43.5		0.1		-60.5	0	2	2	4
41	60.4	3	29	29	59.4		2		60.4		27		9
42	41.1	2.7	88	86	41.1		24						4
43	57.2	2.7	27	33	57.2		18						4
44	49.3	2.6	46	40	49.3		21						2
45	30.2	2.4	77	94	30.2	2.4	47	43					2
46	35.4	2.3	80	92	35.4	2.3	34	31					7
47	36.3	1.9	42	44	35.8	0.4	32	26					3
48	44.5	1.7	14	12	44.5		4						2
49	20.3	1.5	91	87	20.3	1.5	48	51					3
50	40.3	1.5	40	45	40.3	1.5	19	24					2
51	59.1	1.5	71	62	59.1		14						2
52	56.7	1.4	50	34	51.8	6.3	15	18					3
53	54.5	1.4	46	53	54.3	1.5	22	20	54.4		66		7
54	5.4	1.4	92	91	5.5		68						
55	49.9	1.4	19	24	49.9		13						1
56	31.7	1.2	71	64	32.7	0.1	12	11					1
57	46	1	76	85	46	1	25	27					4
58	26.9	0.9	93	92	26.9	0.9	39	39					1
59	41.1	0.9	74	71	41.1		23						3
60	57.9	0.7	22	35	56.3		7						3
61	43.7	0.6	56	78	43.7	0.6	27	27	44.2		89		2
62	22	0.6	89	91	21.5	0.5	33	34					2
63	24.6	0.6	92	96	24.6		44						5
64	23.4	0.6	54	60	23.4		34						7
65	38	0.6	64	53	38		12						2
66	39.4	0.6	64	47									4
67	49.3	0.5	60	66	49.3		13						1
68	32.8	0.5	57	65	32.8		30						4
69	53.1	0.5	72	60	53.1		9						2
70	34.9	0.4	78	71	34.9	0.4	28	28					4
71	47.2	0.4	73	77	47.6		29						3
72	24.9	0.2	87	57	24.9	0.2	40	29					1
73	37.7	0.2	87	85	37.7		30						3
74	37.4	0.2	47	56	37.6		10						
75	31.6	0.1	99.9	99	31.6	0	38	40					3
76	56.8	0.1	41	49	56.8	0	8	7					8
77	20.5	0.1	27	50	20.5		10						1
78	51.9	0.1	20	29	51.9		2						

	Urine Heteroplasmy				Blood Heteroplasmy				Muscle Heteroplasmy				
ID	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	No. NMDAS
79	36.2		37		36.2	6	0.1	0.1	41.2		41		2
80	13		93		19.7	2.8	62	54					3
81	22.5		71		21.8	0.7	35	31					1
82	39.3		63		38.8	0.6	23	21					2
83	45.7		88		45.5	0.5	24	24					2
84	19		82		18.5	0.5	57	60					1
85	24.7		92		24.7	0.2	54	48					15
86	64.7		54						73.5	0.2	12	66	11
87	7.1		99						6.9		94		7
88	14.1		96		13.8		47						12
89	8.4		96										2
90	10.3		96										5
91	9.7		93		9.7		66						
92	12		91		12		62						1
93	22		91		22		38						
94	46.3		91										1
95	43.3		88		41.3		46						2
96	17.1		86		16.7		53						1
97	26.2		86		26.2		44		26.7		86		1
98	23.2		85		23.2		41						1
99	26.3		82		26.3		49						5
100	8.7		82		8.7		46						1
101	54.3		79		54.3		27						6
102	25		78		25		31						10
103	4.4		78										5
104	25.9		77		25.9		23						10
105	38.2		76		38.2		25						3
106	45.2		76		44.9		23		44.9		78		1
107	55.1		76										3
108	37		75		37		22						3
109	59.8		74		57.1		16						8
110	9.7		72		9.7		44						8
111	31.6		72		31.6		16						4
112	31.2		71		31.2		48						4
113	14.6		71		14.6		40						
114	35.4		70		35.4		25						2
115	0		69		0		68						1
116	55.2		68										
117	19		67		19		44						
118	27.5		62		27.5		27						3
119	37.8		61										4
120	60		60		61		17		64.2		62		3

	Urine Heteroplasmy				Blood Heteroplasmy				Muscle Heteroplasmy				
ID	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	No. NMDAS
121	23.2		59		23.2		18						2
122	32.2		57		39.2		23		31		66		2
123	44.9		57		44.9		9						2
124	29.3		56		29.3		33						1
125	55.1		56		55.1		18		54.8		70		1
126	45.8		56		45.8		14		41.9		52		2
127	70.6		56		70.5		3						2
128	38.2		55		45.2		13						1
129	25.1		55		22.3		12						1
130	55.1		55		55.6		6		55.1		39		
131	18.2		54		18.2		19						
132	52.8		54										
133	39.8		53		39.8		25						
134	26.5		52		26.5		29		26.7		63		2
135	37.2		51		37.2		17						3
136	14.1		50		14.1		43						
137	25		47						24.6		77		1
138	23.5		44		23.5		23		23.5		64		
139	25.8		42		25.8		18						
140	46.2		40		46.2		22						1
141	49.7		40		49.7		20						1
142	23.2		40		23.2		8						1
143	26.2		39		26.2		15						2
144	32.5		38		32.5		13						1
145	27.7		35		27.7		7		27.7		17		1
146	20.5		34		20.5		48						1
147	55		34		55		16						1
148	51.7		34										1
149	19.3		32		19.3		16						1
150	59.1		32		59.1		7						2
151	49.5		31		49.5		13						2
152	43.4		31		43.4		8						
153	60		26		51.3		1.6						1
154	69.4		26		68.8		0.001		68.8		58		1
155	29.1		23		29.1		14						1
156	35.4		23		35.9		8		35.3		47		1
157	42.9		22		42.9		10						1
158	49.8		22		49.4		6						1
159	62.5		22										1
160	27.3		18		27.3		1		27.5		41		5
161	20.8		15.5		20.8		0.1						
162	36.9		14		36.9		0.1						12

	Urine Heteroplasmy				Blood Heteroplasmy				Muscle Heteroplasmy				
ID	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	Age	Age Diff	Het First	Het Last	No. NMDAS
163	59.9		12		59.9		0.1						3
164	37.1		11										7
165	67.4		10		67.4		0.1						3
166	26.7		7		26.7		1						2
167	45.3		6		45.3		0.1						
168	78		5		78		3						5
169	61.7		3		61.7		1						2
170	19.7		2		19.7		2						
171	29.6		1		29.6		1						
172	43.7		0.1		43.7		0.1						
173					41		30						
174					41.7		28						1

Appendix IV m.3243A>G Longitudinal Model SAS code.

```
Create table sasuser.n3243 as
Select patient_id, family_id, sex, aao, avg_age, no_ass,
het_wt_muscle*100 as hm, het_wt_urine*100 as hu, het_wt_blood*100 as
hb, (dt_nm - dt_birth)/365.0 as time, scaled_score, log_score,
max_response from connection to odbc
(Select dd.*, ddd.avg_age, stroke_seizure from data_sum_with_pd_ct dd
inner join patient_sum_ct ddd on dd.patient_id = ddd.patient_id
left join patient_melas_phenotype mel on dd.patient_id =
mel.patient_id) where genetics = 14 and exclude_flags is null  and
institution = 1 and max_response >= 130
 order by patient_id asc, time asc  ;

data n1;
set sasuser.n3243;
dv = scaled_score**0.25;
time_scaled = time*1000;
if family_id = 99999 then family_group = 'other'; else family_group =
'individual';
het = hu;
if not missing(hu); if not missing(hb);
run;quit;

proc sql;
create table p_fam_temp as select family_id, patient_id, count(*) from
n1 group by family_id, patient_id;

create table p_fam as select family_id, count(*) from p_fam_temp group
by family_id having count(*) < 2;
update n1 set family_id = 99999 where family_id in (select family_id
from p_fam);
create table p_fam_uniq as  select family_id, count(*) from n1 where
family_id < 99999 group by family_id ;
run;

proc sql;
create table young  as select patient_id, count(*) from n1 where time
< 25 and scaled_score > 3 group by patient_id;
create table n1 as select n1.*, case when young.patient_id is not null
then 1 else 0 end as young  from n1 left join young on n1.patient_id =
young.patient_id ;
create table n1 as select n1.*, case when melas.patient_id is not null
then 1 else 0 end as melas from n1 left join melas on n1.patient_id =
melas.patient_id ;
select young, count(distinct patient_id) from n1 group by young;
run;

proc mixed data=n1 method=ml plots=all;
class patient_id young;
    model dv =  time time*time time*time*time time*hu time*young /s
noint outp=PredR outpm = PredF residual;
      repeated /type=sp(pow)(time_scaled ) subject=patient_id
group=young ;
random time /subject=patient_id;
run;
```

Appendix V Binomial logistic regression SAS code.

```
data q;
set sasuser.allq;
t = age_at_ass/10;
h  = hu;
run;

data stuff; do max=1 to 5 by 1; output; end;
data q_id;  do q_id=1 to 29 by 1; output; end;
data q_id_filtered; set q_id;
if  q_id ne 18 ;
run;

proc sql;
create table q1 as select max, case when q_max >= max then 1 else 0
end as qval,q.* from stuff cross join q order by q.q_id, q.q_text,
max;
create table qs as select qif.q_id, q1.max, sum(case when qval = 1
then 1 else 0 end) as yes,  sum(case when qval = 1 then 0 else 1 end)
as no from q_id_filtered qif inner join q1 on qif.q_id = q1.q_id group
by qif.q_id, q1.max;
create table qs1 as select *, yes+no as total from qs where yes >= 10
and no >= 10;
create table qdata as select q1.* from q1 inner join qs1 on q1.q_id =
qs1.q_id and q1.max = qs1.max;
drop table qs; drop table q1; drop table qs1;
run;

proc sort data=q; by q_id; run;

*Find the optimal cutoff point;
ods html close;
proc logistic data = qdata plots = none; by q_id q_text max;
model qval =  h  /stb rsq CLPARM=PL;
ods output Association=Assoc1;
run;
ods html;

*Create the data sets using the optimal cutoff point;
proc sql;
create table out as select q_id, q_text, max,
round(sum(nValue2),.0001) as c from Assoc1 where label2 = 'c' group by
q_id, q_text, max;
create table best_cutoff as select t.q_id, t.max as max, t.c from out
t inner join (select q_id, max(c) as c from out group by q_id) a2 on
t.q_id = a2.q_id and t.c = a2.c;
create table qbestcutoff as select q.*, max as cutoff, case when
q.q_max >= qbest.max then 1 else 0 end as score from q inner join
best_cutoff qbest on q.q_id = qbest.q_id order by q_id, q_text;
drop table out; drop table Assoc1;
run;

proc logistic data = qbestcutoff plots = none;
by q_id q_text ;
model score =  h  /stb rsq CLPARM=PL;
ods output ParameterEstimates=ParsBest ClparmPL=ParsCLBest
Association=AssocBEST  FitStatistics = Fit2;;
output out = PredR p=Pred stdxbeta=StdErrPred lower=Lower upper=Upper
PREDPROBS=(I);
run;
```

Appendix VI Multilevel ordered logistic regression SAS code.

```
data q;
set sasuser.allq;
decade = age_at_ass/10;
het_10 = hu/10;

if q_max > 3 then q_hml = 2; else if q_max > 1 then q_hml =1 ; else
q_hml = 0;

het_10  = hu/10;
het_10  = hb/10;
if not missing(hu);
*if not missing(hb);
run;


proc sql;
create table young as select patient_id, count(*) from n1 where time <
25 and scaled_score > 3 group by patient_id;

create table q1 as select q.*, case when young.patient_id is not null
then 1 else 0 end as young  from q left join young on q.patient_id =
young.patient_id ;
run;

*this is to do analysis of frequencies;
proc sql;
create table q_summary as select q_text,
sum(case when q_max = 0 then 1 else 0 end)/count(*) as q_0,
sum(case when q_max = 1 then 1 else 0 end)/count(*) as q_1,
sum(case when q_max = 2 then 1 else 0 end)/count(*) as q_2,
sum(case when q_max = 3 then 1 else 0 end)/count(*) as q_3,
sum(case when q_max = 4 then 1 else 0 end)/count(*) as q_4,
sum(case when q_max = 5 then 1 else 0 end)/count(*) as q_5,
count(*) as n from q group by q_text;
run;

proc sort data = q; by q_text; run;

proc logistic data = q plots=EFFECT(X=decade at(het_10 = 0.5 4 6.5))
plots=EFFECT(X=het_10  at(decade = 2 4 6));
by q_text;
model q_hml =  decade het_10   /stb rsq CLPARM=PL;
ods output OddsRatios=OddsDia ParameterEstimates=ParsDia
CumulativeModelTest=TestDia FitStatistics=FitDia RSquare=R2Dia
ClparmPL=ParsCL;
output out = PredR p=Pred stdxbeta=StdErrPred lower=Lower upper=Upper
PREDPROBS=(I);
run;
```

Appendix VII. Parameters for calculation of risk profiles for m.3243A>G multilevel ordinal logistic model.

Each parameter is shown with its standard error. The het parameter is for urine heteroplasmy decades, age is for age decades. Standardised parameters are also shown in the final two columns.

Item	Int 0	S.E	P	Int 1	S.E	P	Het	S.E	P	Age	S.E	P	Het STD	Age STD
Cardiovascular	3.612	0.976	0.0002	7.251	1.195	1E-09	-0.253	0.089	0.0044	-0.478	0.152	0.0017	-0.347	-0.371
Cerebellar ataxia	4.660	0.939	7E-07	8.767	1.184	1E-13	-0.313	0.083	0.0002	-0.928	0.161	8E-09	-0.436	-0.742
CPEO	5.011	1.094	5E-06	7.134	1.185	2E-09	-0.324	0.095	0.0007	-0.547	0.156	0.0004	-0.452	-0.428
Cognition	2.002	0.748	0.0074	3.785	0.795	2E-06	-0.259	0.072	0.0003	-0.236	0.121	0.0511	-0.361	-0.185
Cutting food	4.106	0.991	3E-05	7.023	1.143	8E-10	-0.247	0.088	0.0051	-0.460	0.145	0.0015	-0.345	-0.369
Diabetes mellitus	5.384	1.025	2E-07	5.857	1.044	2E-08	-0.276	0.086	0.0014	-0.873	0.161	5E-08	-0.386	-0.685
Dressing	3.829	0.912	3E-05	6.353	1.015	4E-10	-0.258	0.081	0.0015	-0.482	0.137	0.0004	-0.361	-0.383
Dysphonia/dysarthria	3.820	0.975	9E-05	6.984	1.131	7E-10	-0.336	0.090	0.0002	-0.324	0.142	0.0228	-0.469	-0.254
Encephalopathy	4.252	1.137	0.0002	5.451	1.176	4E-06	-0.335	0.104	0.0012	-0.260	0.162	0.1078	-0.469	-0.203
Exercise Tolerance	1.699	0.730	0.0199	3.984	0.797	6E-07	-0.173	0.068	0.0111	-0.428	0.122	0.0004	-0.242	-0.341
Extrapyramidal	7.713	2.411	0.0014	10.151	2.616	0.0001	-0.564	0.224	0.0118	-0.352	0.285	0.2171	-0.787	-0.276
Gait stability	2.529	0.785	0.0013	5.719	0.913	4E-10	-0.188	0.073	0.0097	-0.539	0.131	4E-05	-0.263	-0.436
Gastro-intestinal	-0.174	0.757	0.8178	3.083	0.815	0.0002	-0.039	0.071	0.5865	-0.145	0.126	0.2472	-0.054	-0.116
Handwriting	3.429	0.876	9E-05	5.841	0.970	2E-09	-0.222	0.079	0.0047	-0.442	0.134	0.0009	-0.311	-0.352
Hearing	4.438	0.874	4E-07	7.664	1.050	3E-13	-0.460	0.083	3E-08	-0.873	0.148	4E-09	-0.644	-0.695
Hygiene	4.442	0.970	5E-06	6.340	1.043	1E-09	-0.308	0.086	0.0003	-0.538	0.142	0.0002	-0.431	-0.425
Migraine Headaches	0.308	0.717	0.6673	2.134	0.739	0.0039	-0.167	0.067	0.013	0.011	0.117	0.9229	-0.234	0.009
Myopathy	3.583	0.881	5E-05	7.376	1.065	4E-12	-0.275	0.080	0.0005	-0.652	0.145	7E-06	-0.383	-0.514
Neuropathy	4.706	0.971	1E-06	8.345	1.170	1E-12	-0.216	0.083	0.0093	-0.821	0.155	1E-07	-0.301	-0.657
Psychiatric	0.991	0.755	0.1895	4.149	0.841	8E-07	-0.161	0.071	0.0229	-0.209	0.125	0.0961	-0.226	-0.164
Ptosis	2.191	0.796	0.0059	4.792	0.886	6E-08	-0.132	0.073	0.0711	-0.313	0.126	0.0131	-0.184	-0.250
Pyramidal	5.855	1.594	0.0002	8.221	1.751	3E-06	-0.350	0.140	0.0125	-0.431	0.214	0.0438	-0.488	-0.337
Seizures	3.108	1.005	0.002	4.489	1.045	2E-05	-0.289	0.094	0.0021	-0.093	0.150	0.5361	-0.404	-0.073
Speech	3.853	0.942	4E-05	7.171	1.111	1E-10	-0.306	0.085	0.0003	-0.409	0.140	0.0034	-0.429	-0.324
Stroke-like episodes	5.085	1.361	0.0002	5.658	1.380	4E-05	-0.368	0.124	0.003	-0.299	0.186	0.107	-0.514	-0.234
Swallowing	3.972	0.938	2E-05	6.307	1.033	1E-09	-0.258	0.084	0.0021	-0.473	0.138	0.0006	-0.360	-0.381
Vision with glasses	2.536	0.827	0.0022	7.816	1.336	5E-09	-0.186	0.075	0.0139	-0.406	0.134	0.0024	-0.260	-0.321
Visual acuity	2.704	0.840	0.0013	7.053	1.148	8E-10	-0.128	0.076	0.0929	-0.472	0.136	0.0005	-0.179	-0.378

Appendix VIII Calculation of odds ratio and probability.

The tables in the following appendices contains the parameters for calculating individual features of m.3243A>G and single-large scale mtDNA deletions in model with age, heteroplasmy, and deletion size (for single deletionos) as predictors.

Int 0 is the intercept for calculating the probability of being asymptomatic.

Int 1 is used to calculate the probability of being asymptomatic or moderate (it is cumulative).

To calculate the probability of being moderately affected, subtract the probability of being asymptomatic (calculated using Int 0) from the probability of being asymptomatic or moderate (calculated using Int 1). To calculate the probability of being severely affected subtract the probability of being asymptomatic or moderate from 100%.

The formula is

$$\log(odds) = Intercept + Age\ decade\ \times Age\ par + Het\ decade \times het\ par$$

e.g. For m.3243A>G and cerebellar ataxia, the log odds of being asymptomatic at age 20 and heteroplasmy level 80% is $4.660 - 0.313 \times 8 - 0.928 \times 2 = 0.3$.

To convert a log odds to a probability,

$$Probability = \frac{e^{Log\ odds}}{1 + e^{Log\ odds}}$$

Thus in this case Probability = 3/(1 + 0.3) = 23.1%.

Appendix IX Proportion of female offspring in pedigrees from the Newcastle cohort.

The proportions are calculated from the total of live-born offspring for each generation below that of the first mother known or suspected to carry the m.3243A>G mutation.

No	Female	Male	Female Proportion
1	1	4	20%
2	1	4	20%
3	1	2	33%
4	2	0	100%
5	2	0	100%
6	1	2	33%
7	1	3	25%
8	0	2	0%
9	13	6	68%
10	1	1	50%
11	1	2	33%
12	1	2	33%
13	1	3	25%
14	3	0	100%
15	4	3	57%
16	4	3	57%
17	12	14	46%
18	1	1	50%
19	2	0	100%
20	7	0	100%
21	8	5	62%
22	2	2	50%
23	12	8	60%
24	1	1	50%
25	8	10	44%
26	10	7	59%
27	13	5	72%
28	1	2	33%
29	1	2	33%
30	2	1	67%
31	2	1	67%
32	1	1	50%
33	3	5	38%
34	2	0	100%
35	2	0	100%
36	3	3	50%
37	5	2	71%
38	6	4	60%
39	3	0	100%
40	0	1	0%
41	2	0	100%
42	4	2	67%
43	1	0	100%
44	2	0	100%

Appendix X Proportion of female offspring in pedigrees from the meta-analysis.

The proportions are calculated from the total of live-born offspring for each generation below that of the first mother known or suspected to carry the m.3243A>G mutation. Some studies contain multiple pedigrees; multiple values are included for each study in this case.

Source	Female	Male	Female Proportion
(Chen *et al.*, 2012)	2	2	50%
(Conway *et al.*, 2011)	5	5	50%
(de Wit *et al.*, 2012)	3	2	60%
(Doleris *et al.*, 2000)	5	3	63%
(Doleris *et al.*, 2000)	3	0	100%
(Dougherty *et al.*, 1994)	8	7	53%
(Gilchrist *et al.*, 1996)	1	4	20%
(Hammans *et al.*, 1995)	4	2	67%
(Hammans *et al.*, 1995)	7	2	78%
(Hammans *et al.*, 1995)	4	2	67%
(Hammans *et al.*, 1995))	4	2	67%
(Hammans *et al.*, 1995)	8	3	73%
(Hammans *et al.*, 1995)	13	9	59%
(Hammans *et al.*, 1995)	2	3	40%
(Hotta *et al.*, 2001)	1	3	25%
(Hotta *et al.*, 2001)	5	0	100%
(Hotta *et al.*, 2001)	2	1	67%
(Iwasaki *et al.*, 2001)	2	3	40%
(Iwasaki *et al.*, 2001)	5	2	71%
(Iwasaki *et al.*, 2001)	1	1	50%
(Iwasaki *et al.*, 2001)	1	0	100%
(Iwasaki *et al.*, 2001)	0	3	0%
(Iwasaki *et al.*, 2001)	1	1	50%
(Kobayashi *et al.*, 1992)	1	1	50%
(Kobayashi *et al.*, 1992)	1	1	50%
(Koga *et al.*, 2000)	5	1	83%
(Koga *et al.*, 2000)	2	0	100%
(Koga *et al.*, 2000)	1	1	50%
(Koga *et al.*, 2000)	4	2	67%
(Li *et al.*, 2008)	3	2	60%
(Massin *et al.*, 1999)	4	1	80%
(Salsano *et al.*, 2011)	5	0	100%
(Shanske *et al.*, 2004)	1	2	33%
(Sue *et al.*, 1998a)	8	7	53%
(Sue *et al.*, 1998a)	15	22	41%
(Sue *et al.*, 1998a)	1	1	50%
(Tay *et al.*, 2008)	7	1	88%
(Walter, 2009)	16	6	73%
(Zhang *et al.*, 2009)	4	0	100%

Appendix XI Single large-scale mtDNA deletion Newcastle cohort data.

No	Phenotype	Age at Onset	Age at Biopsy	COX %	Deletion Size	Break point 5'	Break Point 3'	NMDAS Assessments	Het %
1	KSS	11	27	10	9120	6468	15588	7	37%
2	Multisystem	6	46	15	8704	7175	15879	4	13%
3	CPEO		57	4	8560	5942	14502		6%
4	CPEO	30	34	40	8039	7637	15676	3	36%
5	CPEO		65	25	7977	6537	14514		45%
6	CPEO	38			7958	6033	13991	1	32%
7	CPEO+MM	15	25	25	7768	6352	14120	7	44%
8	Multisystem	14	59	10	7676	6323	13999	5	39%
9	Multisystem	55	85	12	7671	6741	14412		23%
10	Multisystem	5	28	22	7648	6341	13989	4	33%
11	CPEO	20	38	0.5	7595	7845	15440	4	6%
12	CPEO	39	72	5	7500				10%
13	CPEO	30		15	7498	7130	14628	2	28%
14			113		7451	8287	15738		5%
15	CPEO+MM		60	3	7355	7168	14523		24%
16	CPEO+MM		59	20	7284	6774	14058		51%
17	CPEO	18	30	15	7144	5772	12916	4	37%
18	CPEO	16	16	20	7129	8543	15672		7%
19	CPEO+MM		68		6978	7821	14799		56%
20	CPEO	27	63	13	6864	7128	13992	3	35%
21	KSS	11	23	50	6864	7128	13992	1	90%
22	CPEO+MM	22	36	20	6549	6006	12555	6	40%
23	KSS	10	24	30	6472	7540	14012	8	73%
24	Multisystem	15	41	45	6058	8838	14896	2	50%
25	CPEO+MM			15	5958	6002	11960		45%
26	CPEO		41		5906	8324	14230		20%
27	CPEO	15	40	12	5899	9523	15422	7	35%
28	CPEO	16	33	20	5813	9754	15567	1	25%
29	CPEO+MM	15	34	35	5470	6603	12073	7	71%
30	Multisystem	9	31	25	5340	6714	12054	5	45%
31	Multisystem	18	49	20	5160	9258	14418	11	53%
32	Multisystem	12	32	49	5000			4	79%
33	Multisystem	11	56	16	5000			1	42%
34	Multisystem	12	20	20	4999	6625	11624		75%
35	CPEO	23	51	32	4977	8470	13447	8	58%
36	Multisystem	28	40	20	4977	8470	13447	7	75%
37	CPEO+MM	14	31	10	4977	8470	13447	6	62%
38	Multisystem	16	37	43	4977	8470	13447	6	67%
39	Multisystem	10	34	30	4977	8470	13447	5	61%
40	CPEO+MM	27	55	14	4977	8470	13447	5	34%
41	Multisystem	28	41	80	4977	8470	13447	5	78%
42	CPEO	20	28		4977	8470	13447	3	54%
43	KSS	12	15	30	4977	8470	13447	2	76%
44	KSS	15	33	35	4977	8470	13447	1	78%
45	CPEO+MM	19	43	60	4977	8470	13447	1	81%

No	Phenotype	Age at Onset	Age at Biopsy	COX %	Deletion Size	Break point 5’	Break Point 3'	NMDAS Assessments	Het %
46	CPEO+MM	33	41	70	4977	8470	13447	1	78%
47	CPEO		35	8	4977	8470	13447	1	26%
48	CPEO+MM		43	30	4977	8470	13447	1	56%
49	CPEO+MM	18	57	40	4977	8470	13447		57%
50	CPEO	45		16	4977	8470	13447		30%
51	CPEO		18	6	4977	8470	13447		48%
52	CPEO+MM		25		4977	8470	13447		71%
53			37	14	4977	8470	13447		42%
54	CPEO		39	60	4977	8470	13447		65%
55	CPEO+MM		40	20	4977	8470	13447		63%
56	CPEO		46		4977	8470	13447		13%
57	CPEO+MM	15			4963	10105	15068	1	23%
58	KSS		16		4959	8474	13433		88%
59	CPEO	39	61	7	4909	8814	13723	4	1%
60	CPEO	37		28	4885	7205	12090	1	34%
61	CPEO		27		4851	10747	15598		65%
62	CPEO	23	63	15	4770	9349	14119	1	35%
63	CPEO+MM			60	4770	9349	14119		56%
64	CPEO	15	40	12	4752	8289	13041	9	50%
65	Multisystem	15	26	26	4641	10946	15587	3	83%
66	CPEO+MM	34	38	34	4604	9057	13661	5	65%
67	KSS		17	17	4599	9752	14351		60%
68	CPEO		41	15	4596	9528	14124		35%
69	KSS	15	29	50	4500			9	85%
70	Multisystem	27	58	30	4392	8576	12968	4	22%
71	CPEO+MM	21	32	45	4382	8586	12968	6	69%
72	CPEO	21	25	40	4372	8929	13301		53%
73	Multisystem	30	45	25	4241	9498	13739	8	39%
74	Multisystem	25	40	48	4237	9486	13723	6	81%
75	CPEO	47	60	13	4237	9486	13723	6	46%
76	CPEO		70	22	4223	9500	13723		76%
77	CPEO+MM	16	32	20	4113	11262	15375	9	72%
78	KSS	15			3979	11657	15636		50%
79	Multisystem	24	44	10	3693	9756	13449	7	47%
80	CPEO		63	50	3527	7729	11256		55%
81			105	11	3039	10950	13989		55%
82	Multisystem	36	50	5	2803	11637	14440	11	72%
83	CPEO+MM		40	30	2308	12113	14421		87%
84	CPEO	35	50	20	2300	12112	14412	9	55%
85	Multisystem	48	72	17	2300	12112	14412	9	41%
86	CPEO			18	2300	12112	14412	1	76%
87	CPEO		56	25	2297	12115	14412		55%

Appendix XII Single large-scale mtDNA deletion meta-analysis data.

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
1	López-Gallardo *et al.*, 2009	KSS				11041	4166	15207	84%	Yes
2	Marin-Garcia *et al.*, 2002	KSS	3	15		11000	4500	15500	20%	Yes
3	Kornblum *et al.*, 2005b	CPEO	55	59		10900			23%	Yes
4	Kornblum *et al.*, 2005b	CPEO	56	60		10900			49%	Yes
5	De Coo *et al.*, 1997					10000			1%	No
6	Solano *et al.*, 2003	KSS	8	17		9438	6003	15441	87%	Yes
7	Gellerich 2002			0	27	9000	6000	15000	69%	No
8	Fromenty *et al.*, 1996	CPEO	3	18		8800			32%	Yes
9	Ishikawa 2000					8731	6903	15634	55%	No
10	Emma *et al.*, 2006	KSS	1	14		8661	7836	16497	60%	Yes
11	López-Gallardo *et al.*, 2009	KSS	6			8477	6123	14600	82%	Yes
12	Solano *et al.*, 2003	KSS	9	20		8431	7515	15946	86%	Yes
13	Degoul *et al.*, 1991	CPEO	28	45		8137	5786	13923	25%	Yes
14	Simaan *et al.*, 1999	KSS	3	13		8000			18%	Yes
15	De Coo *et al.*, 1997	KSS				8000			60%	Yes
16	Blok *et al.*, 1995	KSS	10	14		7865	6238	14103	25%	Yes
17	Heddi *et al.*, 1994	KSS				7768	7669	15437	74%	Yes
18	Goto *et al.*, 1990	CPEO	7	15	14.5	7700	6000	13700	20%	Yes
19	Matsuoka *et al.*, 1992	CPEO Multi		15		7700			20%	Yes
20	Matsuoka *et al.*, 1992	KSS		21		7700			20%	Yes
21	Goto *et al.*, 1990	CPEO Multi	7	7	15.3	7700	6000	13700	30%	Yes
22	Goto *et al.*, 1990	CPEO	9	16	27.2	7700	6000	13700	40%	Yes
23	Matsuoka *et al.*, 1992	CPEO		16		7700			40%	Yes
24	Sadikovic 2010	Multisystem		40		7673	6331	14004	24%	No

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
25	Solano *et al.*, 2003	KSS				7663	6331	13994	64%	Yes
26	Johns *et al.*, 1989	CPEO				7650	6250	13900	25%	Yes
27	Montiel-Sosa, 2013			10		7628	7437	15065	85%	No
28	Sadikovic 2010	CPEO		59		7603	8469	16072	22%	No
29	Kiyomoto *et al.*, 1997	CPEO Multi	6	12	15	7565	7827	15392	66%	Yes
30	Sadikovic 2010	CPEO + MM		48		7544	7865	15409	14%	No
31	Oldfors 1992			56	9	7534	8366	15900	52%	No
32	Oldfors 1992			5	30	7534	8366	15900	73%	No
33	Odoardi *et al.*, 2003	KSS	7	21		7521	7983	15504	38%	No
34	Schroder *et al.*, 2000	CPEO		30		7500			15%	Yes
35	Kornblum *et al.*, 2005a	CPEO Multi	10	34		7500			16%	Yes
36	Barrientos *et al.*, 1995	CPEO Multi	10	47		7500			23%	Yes
37	Kornblum *et al.*, 2005a	CPEO Multi	14	42		7500			26%	Yes
38	Schroder *et al.*, 2000	CPEO		34		7500			26%	Yes
39	Schroder *et al.*, 2000	CPEO		26		7500			27%	Yes
40	Barrientos *et al.*, 1995	KSS	12	20		7500			40%	Yes
41	Odoardi *et al.*, 2003	CPEO + MM	8	18		7500			42%	No
42	Barrientos *et al.*, 1995	CPEO Multi	3	12		7500			50%	Yes
43	Kornblum *et al.*, 2005a	CPEO Multi	10	38		7500			58%	Yes
44	Sadikovic 2010	Multisystem		18		7436	8637	16073	45%	No
45	Goto *et al.*, 1990	CPEO	13	31	17.2	7300	6000	13300	40%	Yes
46	Oldfors 1992			13	5	7300	8600	15900	43%	No
47	Sadikovic 2010	Multisystem		46		7213	8427	15640	47%	No
48	Sadikovic 2010	KSS		31		7213	8427	15640	82%	No
49	Matsuoka *et al.*, 1992	CPEO		16		7200			5%	Yes
50	Moraes *et al.*, 1992	KSS				7100			70%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
51	Sadikovic 2010	Multisystem		6		7039	8623	15662	70%	No
52	Zeviani *et al.*, 1988	KSS	8	15		7025	5275	12300	45%	Yes
53	Schroder *et al.*, 2000	KSS		19		7000			20%	Yes
54	Kornblum *et al.*, 2005a	KSS	12	31		7000			66%	Yes
55	Goto *et al.*, 1990	CPEO	22	49	11.6	6800	8600	15400	30%	Yes
56	López-Gallardo *et al.*, 2009	CPEO	31			6798	6024	12822	30%	Yes
57	Barrientos *et al.*, 1995	KSS	14	24		6500			56%	Yes
58	Kornblum *et al.*, 2005b	CPEO	25	51		6500			58%	Yes
59	Schroder *et al.*, 2000	CPEO		50		6500			60%	Yes
60	Schroder *et al.*, 2000	KSS		21		6500			66%	Yes
61	Fromenty *et al.*, 1996	KSS	8	22		6495	7836	14331	31%	Yes
62	Solano *et al.*, 2003	KSS	12	13		6366	7949	14315	69%	Yes
63	López-Gallardo *et al.*, 2009	CPEO				6279	7409	13688	21%	Yes
64	Solano *et al.*, 2003	CPEO				6213	7407	13620	7%	Yes
65	Sadikovic 2010	CPEO + MM		26		6119	9516	15635	32%	No
66	Larsson and Holme, 1992	KSS	7	16	5	6100	8800	14900	59%	No
67	Kornblum *et al.*, 2005a	CPEO Multi	8	61		6000			46%	Yes
68	Schroder *et al.*, 2000	CPEO		49		6000			46%	Yes
69	De Coo *et al.*, 1997	CPEO				6000	9500	15500	70%	Yes
70	De Coo *et al.*, 1997					6000			82%	No
71	Zeviani *et al.*, 1988	KSS	4	7		5980	9020	15000	66%	Yes
72	Solano *et al.*, 2003	CPEO				5928	9816	15744	29%	Yes
73	Sadikovic 2010	KSS		42		5905	8467	14372	70%	No
74	Marie *et al.*, 1999	KSS	1	5		5900			30%	Yes
75	Sadikovic 2010	KSS		11		5867	8558	14425	65%	No
76	Okulla *et al.*, 2005	CPEO	12	13		5800			32%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
77	Sadikovic 2010	CPEO + MM		38		5630	8429	14059	33%	No
78	Kornblum *et al.*, 2005a	CPEO Multi	28	52		5600			85%	Yes
79	Sudoyo *et al.*, 1993	CPEO	12	34		5500	8000	13500	39%	Yes
80	Kornblum *et al.*, 2005b	KSS	16	47		5500			58%	Yes
81	Schroder *et al.*, 2000	KSS		40		5500			58%	Yes
82	Goto *et al.*, 1990	KSS	11	12		5500	7500	13000	80%	Yes
83	Zeviani *et al.*, 1988	KSS	3	12		5448	10600	16048	57%	Yes
84	Sadikovic 2010	MM		26		5438	8140	13578	11%	No
85	Goto *et al.*, 1990	KSS	2	13	23.9	5400	7000	12400	60%	Yes
86	Shanske *et al.*, 2002	CPEO				5355	10004	15359	40%	No
87	López-Gallardo *et al.*, 2009	KSS				5311	7450	12761	73%	Yes
88	Sadikovic 2010	CPEO + MM		57		5225	6076	11301	13%	No
89	Goto *et al.*, 1990	KSS	4	14	40.6	5200	8500	13700	90%	Yes
90	López-Gallardo *et al.*, 2009	CPEO				5113	8477	13590	80%	Yes
91	Sadikovic 2010	Multisystem		34		5112	8468	13580	50%	No
92	Goto *et al.*, 1990	KSS	12	30	14	5100	10059	15159	70%	Yes
93	Carta *et al.*, 2000	CPEO		28		5049	9570	14619	55%	Yes
94	Vazquez-Acevedo *et al.*, 1995	KSS	4	17		5026	10050	15076	86%	Yes
95	Johns and Hurko, 1989	KSS	13	43		5014	8708	13722	65%	Yes
96	Kornblum *et al.*, 2005a	CPEO Multi	55	61		5000			5%	Yes
97	Schroder *et al.*, 2000	CPEO		39		5000			26%	Yes
98	Moraes *et al.*, 1992	KSS				5000			36%	Yes
99	Barrientos *et al.*, 1995	CPEO Multi	8	16		5000			42%	Yes
100	Kornblum *et al.*, 2005a	CPEO Multi	12	50		5000			53%	Yes
101	Tanaka *et al.*, 1989	CPEO	14	32		5000	8600	13600	53%	Yes
102	Schroder *et al.*, 2000	CPEO		38		5000			53%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
103	Schroder *et al.*, 2000	KSS		18		5000			54%	Yes
104	Moraes *et al.*, 1992	CPEO				5000			55%	Yes
105	Schroder *et al.*, 2000	CPEO		30		5000			55%	Yes
106	Schroder *et al.*, 2000	KSS		19		5000			58%	Yes
107	Vielhaber *et al.*, 2002	CPEO		44		5000			58%	Yes
108	Vielhaber *et al.*, 2002	CPEO		29		5000			68%	Yes
109	Matsuoka *et al.*, 1992	CPEO		14		5000			70%	Yes
110	Kornblum *et al.*, 2005b	CPEO	30	48		5000			72%	Yes
111	Schroder *et al.*, 2000	CPEO		42		5000			72%	Yes
112	Vielhaber *et al.*, 2002	CPEO		47		5000			72%	Yes
113	Kornblum *et al.*, 2005a	CPEO Multi	16	57		5000			74%	Yes
114	Vielhaber *et al.*, 2002	KSS		13		5000			74%	Yes
115	Moraes *et al.*, 1992	KSS				5000			76%	Yes
116	Kornblum *et al.*, 2005a	KSS	9	39		5000			78%	Yes
117	Schroder *et al.*, 2000	KSS		29		5000			78%	Yes
118	Sadikovic 2010	PEO		24		4995	5835	10830	58%	No
119	López-Gallardo *et al.*, 2009	CPEO				4978	8482	13460	6%	Yes
120	Wong, 2001	KSS		26		4978	8482	13460	10%	Yes
121	Odoardi *et al.*, 2003	CPEO	39	43		4978	8482	13460	12%	Yes
122	Bernes *et al.*, 1993	CPEO	16	24		4978	8482	13460	15%	Yes
123	Pineda *et al.*, 2004	CPEO				4978	8482	13460	15%	Yes
124	López-Gallardo *et al.*, 2009	CPEO				4978	8482	13460	16%	Yes
125	Goto *et al.*, 1990	CPEO	20	53	5	4978	8482	13460	20%	Yes
126	Odoardi *et al.*, 2003	CPEO	12	14		4978	8482	13460	26%	Yes
127	Odoardi *et al.*, 2003	CPEO	35	45		4978	8482	13460	31%	Yes
128	Odoardi *et al.*, 2003	CPEO	15	36		4978	8482	13460	33%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
129	Wong, 2001	KSS		28		4978	8482	13460	33%	Yes
130	Degoul *et al.*, 1991	CPEO Multi	4	31		4978	8482	13460	34%	Yes
131	Gellerich 2002					4978	8482	13460	39%	No
132	Poulton *et al.*, 1991	KSS	14	14	3	4978	8482	13460	40%	Yes
133	Goto *et al.*, 1990	CPEO	14	32	12.8	4978	8482	13460	40%	Yes
134	López-Gallardo *et al.*, 2009	CPEO	20			4978	8482	13460	40%	Yes
135	Chen *et al.*, 1998	CPEO				4978	8482	13460	44%	Yes
136	López-Gallardo *et al.*, 2009	CPEO				4978	8482	13460	45%	Yes
137	Wong, 2001	KSS		20		4978	8482	13460	45%	Yes
138	Wong, 2001	CPEO		60		4978	8482	13460	45%	Yes
139	Goto *et al.*, 1990	CPEO	12	15	18.5	4978	8482	13460	50%	Yes
140	Shoffner *et al.*, 1989	CPEO Multi	20	61		4978	8482	13460	50%	Yes
141	López-Gallardo *et al.*, 2009	CPEO	48			4978	8482	13460	50%	Yes
142	Wong, 2001	KSS		34		4978	8482	13460	51%	Yes
143	López-Gallardo *et al.*, 2009	CPEO				4978	8482	13460	54%	Yes
144	Gellerich 2002					4978	8482	13460	61%	No
145	López-Gallardo *et al.*, 2009	KSS	4			4978	8482	13460	62%	Yes
146	López-Gallardo *et al.*, 2009	KSS				4978	8482	13460	62%	Yes
147	Obermaier-Kusser *et al.*, 1990	KSS	11	26		4978	8482	13460	63%	Yes
148	Wong, 2001	KSS		16		4978	8482	13460	63%	Yes
149	Wong, 2001	KSS		36		4978	8482	13460	64%	Yes
150	Boles *et al.*, 1998	KSS	1	5		4978	8482	13460	65%	Yes
151	Wong, 2001	KSS		6		4978	8482	13460	65%	Yes
152	Gellerich 2002				48	4978	8482	13460	66%	No
153	Degoul *et al.*, 1991	KSS	7	27		4978	8482	13460	68%	Yes
154	López-Gallardo *et al.*, 2009	CPEO				4978	8482	13460	68%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
155	Odoardi *et al.*, 2003	KSS	6	17		4978	8482	13460	69%	Yes
156	Goto *et al.*, 1990	KSS	10	14	41.4	4978	8482	13460	70%	Yes
157	Wong, 2001	KSS		44		4978	8482	13460	70%	Yes
158	López-Gallardo *et al.*, 2009	KSS	12			4978	8482	13460	72%	Yes
159	Ponzetto *et al.*, 1990	KSS				4978	8482	13460	72%	Yes
160	López-Gallardo *et al.*, 2009	KSS				4978	8482	13460	72%	Yes
161	Consalvo *et al.*, 1997	KSS	9	19		4978	8482	13460	80%	Yes
162	Degoul *et al.*, 1991	KSS	10	31		4978	8482	13460	80%	Yes
163	Goto *et al.*, 1990	CPEO	22	26	40.2	4978	8482	13460	80%	Yes
164	López-Gallardo *et al.*, 2009	KSS				4978	8482	13460	80%	Yes
165	Johns *et al.*, 1989	KSS				4978	8482	13460	83%	Yes
166	Odoardi *et al.*, 2003	KSS	6	10		4978	8482	13460	85%	Yes
167	Johns *et al.*, 1989					4978	8482	13460	86%	No
168	Sadikovic 2010	CPEO + MM		60		4977	8470	13447	6%	No
169	Sadikovic 2010	KSS		26		4977	8470	13447	10%	No
170	Kiyomoto *et al.*, 1997	CPEO	30	44	1	4977	8483	13460	14%	Yes
171	Sadikovic 2010	Multisystem		24		4977	8470	13447	24%	No
172	Sadikovic 2010	Multisystem		14		4977	8470	13447	27%	No
173	Sadikovic 2010	CPEO + MM		78		4977	8470	13447	27%	No
174	Sadikovic 2010	KSS		28		4977	8470	13447	33%	No
175	Sciacco *et al.*, 1994	KSS		35		4977	8470	13447	36%	Yes
176	Sadikovic 2010	CPEO + MM		39		4977	8470	13447	42%	No
177	Sadikovic 2010	KSS		20		4977	8470	13447	45%	No
178	Sadikovic 2010	CPEO + MM		60		4977	8470	13447	45%	No
179	Sadikovic 2010	Multisystem		45		4977	8470	13447	49%	No
180	Sciacco *et al.*, 1994	CPEO	61	66		4977	8470	13447	50%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
181	Shanske *et al.*, 1990	KSS				4977	8470	13447	50%	Yes
182	Kiyomoto *et al.*, 1997	CPEO	15	33	8.5	4977	8483	13460	54%	Yes
183	Sciacco *et al.*, 1994	KSS	11	13		4977	8470	13447	55%	Yes
184	Sadikovic 2010	CPEO + MM		56		4977	8470	13447	55%	No
185	Sadikovic 2010	CPEO		39		4977	8470	13447	60%	No
186	Sadikovic 2010	KSS		13		4977	8470	13447	61%	No
187	Sadikovic 2010	KSS		16		4977	8470	13447	62%	No
188	Sudoyo *et al.*, 1993	CPEO	4	41		4977	8470	13447	64%	Yes
189	Sadikovic 2010	MM		36		4977	8470	13447	64%	No
190	Sadikovic 2010	Multisystem		6		4977	8470	13447	65%	No
191	Sadikovic 2010	Renal tubular acidosis		8		4977	8470	13447	67%	No
192	Sadikovic 2010	CPEO + MM		9		4977	8470	13447	67%	No
193	Sadikovic 2010	CPEO + MM		36		4977	8470	13447	70%	No
194	Sadikovic 2010	KSS		44		4977	8470	13447	70%	No
195	Sciacco *et al.*, 1994	KSS		28		4977	8470	13447	75%	Yes
196	Kiyomoto *et al.*, 1997	CPEO Multi	12	14	18	4977	8483	13460	77%	Yes
197	Mita 1989	KSS		30	51	4977	8483	13460	80%	No
198	Sadikovic 2010	CPEO + MM		27		4977	8470	13447	85%	No
199	Sadikovic 2010	PEO		20		4977	8470	13447	88%	No
200	Solano *et al.*, 2003	CPEO				4958	8380	13338	30%	Yes
201	Pistilli *et al.*, 2003	KSS	10	36		4949	8631	13580	60%	Yes
202	Oldfors 1992			27	20	4914	7586	12500	79%	No
203	Barrientos *et al.*, 1995	KSS	3	28		4800			38%	Yes
204	Larsson and Holme, 1992	KSS	5	12		4800	10800	15600	87%	No
205	López-Gallardo *et al.*, 2009	KSS				4754	11292	16046	45%	Yes

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
206	Oldfors 1992			11	1	4700	10500	15200	87%	No
207	Degoul *et al.*, 1991	KSS	18	25		4500	11300	15800	30%	Yes
208	Johns *et al.*, 1989	CPEO				4500	9300	13800	55%	Yes
209	Zeviani *et al.*, 1988	KSS	14	26		4500	9000	13500	62%	Yes
210	Zeviani *et al.*, 1988	KSS	14	28		4500	9000	13500	75%	Yes
211	Solano *et al.*, 2003	KSS	7	18		4421	10951	15372	80%	Yes
212	Kiyomoto *et al.*, 1997	KSS	4	15	3.5	4420	10952	15372	20%	Yes
213	Sadikovic 2010	Multisystem		8		4420	10560	14980	55%	No
214	Wong, 2001	KSS		8		4420	10560	14980	55%	Yes
215	Sadikovic 2010	Multisystem		48		4369	9256	13625	38%	No
216	Carod-Artal *et al.*, 2003	CPEO	19	26		4238	9500	13738	55%	Yes
217	Solano *et al.*, 2003	CPEO				4238	9485	13723	55%	No
218	Goto *et al.*, 1990	CPEO	40	52	19	4200	5850	10050	30%	Yes
219	Zeviani *et al.*, 1988	KSS	17	27		4200	9000	13200	66%	Yes
220	Pineda *et al.*, 2006	KSS	7	8		4124	11033	15157	72%	Yes
221	Blakely *et al.*, 2004	CPEO			20	4115	11262	15375	66%	Yes
222	Gellerich 2002					4093	10057	14150	62%	No
223	De Coo *et al.*, 1997	KSS				4000			60%	Yes
224	Kiyomoto *et al.*, 2006	CPEO	60	62	14	3800			4%	Yes
225	Vielhaber *et al.*, 2002	CPEO		48		3800			10%	Yes
226	Kornblum *et al.*, 2004	CPEO	28	35		3800			45%	Yes
227	Solano *et al.*, 2003	KSS	14	31		3720	11727	15447	50%	Yes
228	Kiyomoto *et al.*, 1997	CPEO	27	43	8	3716	10845	14561	44%	Yes
229	Kiyomoto *et al.*, 2006	CPEO	18	28	12	3700			65%	Yes
230	Degoul *et al.*, 1991	CPEO	31	51		3513	7483	10996	40%	Yes
231	Wong, 2001					3500			16%	No

No	Original Study	Phenotype	Age at Onset	Age at Biopsy	COX %	Del Size	Break point 5'	Break point 3'	Het%	In López-Gallardo *et al.*
232	Gellerich 2002					3500	11300	14800	58%	No
233	Kornblum *et al.*, 2005a	KSS	20	37		3500			62%	Yes
234	Schroder *et al.*, 2000	KSS		30		3500			62%	Yes
235	Schroder *et al.*, 2000	CPEO		29		3500			65%	Yes
236	Lertrit *et al.*, 1999	KSS	35	37		3485	10280	13765	37%	Yes
237	Barbiroli *et al.*, 1995	CPEO	22	30		3300			20%	Yes
238	Wong *et al.*, 2003	KSS	29	36		3079	8419	11498	92%	Yes
239	Sadikovic 2010	Multisystem		25		3030	10958	13988	47%	No
240	Sadikovic 2010	KSS		41		2976	8388	11364	92%	No
241	Goto *et al.*, 1990	CPEO	35	45	16.6	2800	11500	14300	50%	Yes
242	Kornblum *et al.*, 2005a	CPEO Multi	44	49		2700			32%	Yes
243	Gellerich 2002				22	2600	11000	13600	62%	No
244	Kornblum *et al.*, 2005b	CPEO	48	48		2500			39%	Yes
245	Ohno *et al.*, 1996	KSS	26	27		2500			88%	Yes
246	Solano *et al.*, 2003	KSS				2434	10620	13054	77%	Yes
247	Solano *et al.*, 2003	KSS				2310	12112	14422	70%	Yes
248	Kiyomoto *et al.*, 1997	CPEO	20	47	20	2309	12113	14422	59%	Yes
249	Kiyomoto *et al.*, 1997	CPEO	21	31	4	2309	12113	14422	89%	Yes
250	Goto *et al.*, 1990	CPEO	39	70	10.7	2300	11000	13300	50%	Yes
251	Moraes *et al.*, 1992	CPEO				2300			80%	Yes
252	Goto *et al.*, 1990	CPEO	34	37	7.2	2200	12000	14200	50%	Yes
253	Goto *et al.*, 1990	CPEO	52	55	19.3	2200	12000	14200	60%	Yes
254	Zeviani *et al.*, 1988	KSS	24	33		2060	7440	9500	60%	Yes
255	Schroder *et al.*, 2000	CPEO		45		2000			32%	Yes
256	Goto *et al.*, 1990	KSS	35	36	12.6	1800	13000	14800	50%	Yes

Appendix XIII Single large-scale mtDNA deletion meta-analysis exclusions.

Details of the excluded case and reasons for exclusion are noted.

Study	Pheno-type	AAO	Age at Biopsy	COX %	Deletion Size	Break point 3'	Break point 5'	Het%	Notes	In López-Gallardo
Mori *et al.*, 1991	KSS	2	6		6596	6383	12979	60%	Deletion size does not match original publication	yes
Larsson *et al.*, 1992	KSS	16	26		4667	7697	12364	70%	Deletion size does not original publication	yes
Goto *et al.*, 1990	CPEO	13	31	50%	5300	9206	14506	90%	Deletion size is inconsistent with reported breakpoints	yes
Schaefer *et al.*, 2005	CPEO	15	24		7400			8%	In our cohort	yes
Zoccolella *et al.*, 2006		6	32		1813	3505	5318	70%	Location is unlike any other deletion in the study	yes
Reynier *et al.*, 1994	CPEO		41					50%	Other mutations; also contains a point mutation	yes
Odoardi *et al.*, 2003	KSS	20	27		4978	8482	13460	51%	Other mutations; Dimers	yes
Odoardi *et al.*, 2003	CPEO	30	40		7000			51%	Other mutations; Dimers	yes
Brockington *et al.*, 1995	KSS	12	19		8562	7354	15916	27%	Other mutations; Dimers and duplications	yes
Brockington *et al.*, 1995	KSS	4	22		4978	8482	13460	89%	Other mutations; Dimers and duplications	yes
Jacobs *et al.*, 2004	PS/KSS	0	3		8034	7934	15968		Other mutations; Dimers and duplications	no
Jacobs *et al.*, 2004	PS	0	1.5		3444	6097	9541	64%	Other mutations; Dimers and duplications	no
Tengan *et al.*, 1998	KSS	5	12		9660	5784	15444	65%	Other mutations; Duplications	yes
Vazquez-Acevedo *et al.*, 2002	KSS	8	23		4978	8482	13460	75%	Other mutations; multiple deletions	yes
Tanaka *et al.*, 1989	CPEO	39	70		2100			36%	Uncertain breakpoints	yes
Oldfors 1992			52	10%	3940	10060	14000	76%	Uncertain breakpoints	no
Kunz *et al.*, 1997	CPEO		34		3200			53%	Uncertain breakpoints	yes
Kunz *et al.*, 1997	CPEO		55		4700			67%	Uncertain breakpoints	yes
Kunz *et al.*, 1997	CPEO		32		2600			84%	Uncertain breakpoints	yes
Marzuki *et al.*, 1997	CPEO	20	59		5020	9780	14800	16%	Uncertain breakpoints and multiple deletions	yes
Sudoyo *et al.*, 1993	CPEO	25	60		2664	12336	15000	55%	Uncertain breakpoints and multiple deletions	yes

Appendix XIV Single large-scale mtDNA deletion longitudinal model SAS code.

```
data n1;
set sasuser.n3243;
if scaled_score = 0 then scaled_score = 0.1;
dv = scaled_score**0.25;
time_scaled = time*1000;
if missing(family_id) then family_id = 99999;
if family_id = 99999 then family_group = 'other';
else family_group = 'individual';
het = hu;
if not missing(hu);
if not missing(hb);
run;quit;


proc sql;
create table p_fam_temp as select family_id, patient_id, count(*) from
n1 group by family_id, patient_id;
create table p_fam as select family_id, count(*) from p_fam_temp group
by family_id having count(*) < 2;
update n1 set family_id = 99999 where family_id in (select family_id
from p_fam);
create table p_fam_uniq as  select family_id, count(*) from n1 where
family_id < 99999 group by family_id ;
run;

proc sql;
create table young  as select patient_id, count(*) from n1 where time
< 25 and scaled_score > 3 group by patient_id;
create table n1 as select n1.*, case when young.patient_id is not null
then 1 else 0 end as young, from n1 left join young on n1.patient_id =
young.patient_id ;
run;

proc mixed data=n1 method=ml plots=none RATIO COVTEST ;
      class patient_id young family_group family_id;
       model dv =  time time*time time*time*time time*hu time*young
/s noint outp=PredR outpm = PredF residual;
            repeated /type=sp(pow)(time_scaled )
subject=patient_id(family_id) group=young*family_group ;
      random time /subject=patient_id(family_id) ;
            random time /subject=family_id ;
      ods output FitStatistics=fm SolutionF=SFfm;
run;
```

Appendix XV. Parameters for calculation of risk profiles for single large-scale mtDNA deletion multilevel ordinal logistic model.

Each parameter is shown with its standard error. The parameters are age in decades (Decade), the square of the deletion size in kB ($Size^2$), and the natural log of (100%-heteroplasmy) (ln(1-h)).

Item	Int 0	Std	P	Int 1	Std	P	Decade	Std	P	ln(1-h)	Std	P	$Size^2$	Std	P
Vision	3.451	1.989	0.083	6.614	2.166	0.002	-0.559	0.263	0.033	1.523	0.689	0.027	-0.008	0.019	0.657
Hearing	5.241	2.297	0.023	6.616	2.371	0.005	-0.324	0.290	0.265	1.819	0.755	0.016	-0.041	0.021	0.054
Speech	7.904	2.583	0.002	10.941	2.846	0.000	-0.811	0.325	0.013	2.689	0.842	0.001	-0.037	0.022	0.096
Swallowing	5.748	2.096	0.006	7.705	2.219	0.001	-0.641	0.273	0.019	1.980	0.695	0.004	-0.044	0.019	0.021
Handwriting	6.313	2.388	0.008	9.606	2.674	0.000	-0.641	0.300	0.032	1.661	0.767	0.030	-0.031	0.021	0.148
Cutting food	5.819	2.454	0.018				-0.428	0.300	0.154	1.575	0.806	0.051	-0.043	0.023	0.058
Dressing	5.134	2.253	0.023	8.345	2.520	0.001	-0.488	0.285	0.088	1.006	0.728	0.167	-0.035	0.021	0.089
Hygiene	7.247	2.301	0.002	10.130	2.528	<.0001	-0.900	0.298	0.003	1.953	0.733	0.008	-0.034	0.020	0.086
Exercise tolerance	6.710	2.163	0.002	8.718	2.410	0.000	-0.733	0.271	0.007	2.858	0.789	0.000	-0.048	0.020	0.015
Gait stability	5.047	2.133	0.018				-0.584	0.282	0.039	1.821	0.726	0.012	-0.045	0.021	0.030
Psychiatric disturbance	0.123	1.959	0.950				0.176	0.270	0.514	0.225	0.654	0.730	-0.010	0.019	0.590
Migraine headaches	-1.430	1.986	0.472	-0.457	1.981	0.818	0.469	0.283	0.097	0.232	0.642	0.717	0.002	0.019	0.933
Seizures	13.613	8.089	0.092				-0.778	0.955	0.415	2.729	1.999	0.172	-0.106	0.057	0.064
Encephalopic episodes	18.914	17.916	0.291	8.421	2.283	0.000	0.507	1.811	0.780	8.162	7.682	0.288	-0.192	0.170	0.259
GI disturbance	4.653	2.097	0.027	8.028	2.379	0.001	-0.642	0.278	0.021	0.029	0.672	0.966	-0.055	0.021	0.010
Diabetes mellitus	9.718	4.806	0.043				-1.093	0.589	0.064	0.031	1.860	0.987	-0.046	0.035	0.180
Cardiovascular	7.145	2.620	0.006	8.958	2.742	0.001	-0.631	0.325	0.053	1.930	0.817	0.018	-0.051	0.022	0.021
Visual acuity	1.708	1.916	0.373	4.930	2.071	0.017	-0.297	0.255	0.243	0.649	0.645	0.314	0.011	0.019	0.562
Ptosis	3.190	2.284	0.162	5.473	2.345	0.020	-0.690	0.312	0.027	2.736	0.960	0.004	-0.036	0.022	0.108
CPEO				1.373	2.702	0.611	-0.254	0.369	0.491	1.758	1.119	0.116	-0.022	0.027	0.419
Dysphonia/dysarthia	9.760	2.984	0.001	12.907	3.293	<.0001	-0.942	0.356	0.008	3.492	0.986	0.000	-0.046	0.024	0.055
Myopathy	1.165	1.997	0.560	4.737	2.131	0.026	-0.266	0.274	0.332	0.965	0.693	0.164	-0.002	0.019	0.912
Cerebellar ataxia	8.338	2.484	0.001	11.819	2.814	<.0001	-1.048	0.321	0.001	2.327	0.799	0.004	-0.052	0.022	0.017
Extrapyramidal	18.161	7.621	0.017				-1.584	0.797	0.047	5.118	2.278	0.025	-0.088	0.052	0.092
Neuropathy	47.364	102.800	0.645				-0.838	10.196	0.935	-7.010	59.590	0.906	-0.647	0.996	0.516
Cognitive impairment	2.698	2.119	0.203	3.816	2.154	0.076	-0.038	0.278	0.890	0.959	0.765	0.210	-0.034	0.020	0.093

References

Ahlqvist, K.J., Hamalainen, R.H., Yatsuga, S., Uutela, M., Terzioglu, M., Gotz, A., Forsstrom, S., Salven, P., Angers-Loustau, A., Kopra, O.H., Tyynismaa, H., Larsson, N.G., Wartiovaara, K., Prolla, T., Trifunovic, A. and Suomalainen, A. (2012) 'Somatic progenitor cell vulnerability to mitochondrial DNA mutagenesis underlies progeroid phenotypes in Polg mutator mice', *Cell Metab*, 15(1), pp. 100-9.

Akaike, H. (1974) 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, AC-19(6), pp. 716-723.

Al Rawi, S., Louvet-Vallee, S., Djeddi, A., Sachse, M., Culetto, E., Hajjar, C., Boyd, L., Legouis, R. and Galy, V. (2011) 'Postfertilization autophagy of sperm organelles prevents paternal mitochondrial DNA transmission', *Science*, 334(6059), pp. 1144-7.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002) *Molecular Biology of the Cell*. 4th Edition edn.

Alcolado, J.C. and Alcolado, R. (1991) 'Importance of maternal history of non-insulin dependent diabetic patients', *BMJ*, 302(6786), pp. 1178-80.

Allen, D.L., Monke, S.R., Talmadge, R.J., Roy, R.R. and Edgerton, V.R. (1995) 'Plasticity of myonuclear number in hypertrophied and atrophied mammalian skeletal muscle fibers', *Journal of Applied Physiology*, 78(5), pp. 1969-1976.

Altman, D.G. (1981) 'Statistics and ethics in medical research. VIII. Improving the quality of statistics in medical journals', *British Medical Journal*, 282(6257), pp. 44-47.

Altman, D.G. (1982) 'Statistics in medical', *Journals. Stat Med*, 1, pp. 59-71.

Altman, D.G. (1998) 'Statistical reviewing for medical journals', *Stat Med*, 17(23), pp. 2661-2674.

Altman, D.G., Goodman, S.N. and Schroter, S. (2002) 'How statistical expertise is used in medical research', *Journal of the American Medical Association*, 287(21), pp. 2817-2820.

Altman, D.G., Gore, S.M., Gardner, M.J. and Pocock, S.J. (1983) 'Statistical guidelines for contributors to medical journals', *British medical journal (Clinical research ed.)*, 286(6376), pp. 1489-1493.

Andersen, B. and Forrest, M. (1987) 'Misuse of statistics. If neither SD nor SE--what then?', *Om fejlagtig brug af statistik. Hvis hverken SD eller SE--hvad så?*, 102(4), pp. 141-142.

Andreux, P.A., Houtkooper, R.H. and Auwerx, J. (2013) 'Pharmacological approaches to restore mitochondrial function', *Nat Rev Drug Discov*, 12(6), pp. 465-83.

Apabhai, S., Gorman, G.S., Sutton, L., Elson, J.L., Plötz, T., Turnbull, D.M. and Trenell, M.I. (2011) 'Habitual Physical Activity in Mitochondrial Disease', *PLoS One*, 6(7), p. e22294.

Arpa, J., Cruz-Martinez, A., Campos, Y., Gutierrez-Molina, M., Garcia-Rio, F., Perez-Conde, C., Martin, M.A., Rubio, J.C., Del Hoyo, P., Arpa-Fernandez, A. and Arenas, J. (2003) 'Prevalence and progression of mitochondrial diseases: a study of 50 patients', *Muscle Nerve*, 28(6), pp. 690-5.

Auré, K., Ogier De Baulny, H., Laforêt, P., Jardel, C., Eymard, B. and Lombès, A. (2007) 'Chronic progressive ophthalmoplegia with large-scale mtDNA rearrangement: Can we predict progression?', *Brain*, 130(6), pp. 1516-1524.

Austad, S.N. (2006) 'Why women live longer than men: sex differences in longevity', *Gend Med*, 3(2), pp. 79-92.

Bacman, S.R., Williams, S.L. and Moraes, C.T. (2009) 'Intra- and inter-molecular recombination of mitochondrial DNA after in vivo induction of multiple double-strand breaks', *Nucleic Acids Res*, 37(13), pp. 4218-26.

Barrell, B.G., Anderson, S., Bankier, A.T., de Bruijn, M.H., Chen, E., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J., Staden, R. and Young, I.G. (1980) 'Different pattern of codon recognition by mammalian mitochondrial tRNAs', *Proc Natl Acad Sci U S A*, 77(6), pp. 3164-6.

Barthélémy, C., Ogier De Baulny, H., Diaz, J., Armelle Cheval, M., Frachon, P., Romero, N., Goutieres, F., Fardeau, M. and Lombès, A. (2001) 'Late-onset

mitochondrial DNA depletion: DNA copy number, multiple deletions, and compensation', *Annals of Neurology*, 49(5), pp. 607-617.

Bates, M.G., Bourke, J.P., Giordano, C., d'Amati, G., Turnbull, D.M. and Taylor, R.W. (2012a) 'Cardiac involvement in mitochondrial DNA disease: clinical spectrum, diagnosis, and management', *Eur Heart J*, 33(24), pp. 3023-33.

Bates, M.G.D., Hollingsworth, K.G., Newman, J.H., Jakovljevic, D.G., Blamire, A.M., MacGowan, G.A., Keavney, B.D., Chinnery, P.F., Turnbull, D.M., Taylor, R.W., Trenell, M.I. and Gorman, G.S. (2012b) 'Concentric hypertrophic remodelling and subendocardial dysfunction in mitochondrial DNA point mutation carriers', *European Heart Journal – Cardiovascular Imaging*.

Benit, P., Lebon, S. and Rustin, P. (2009) 'Respiratory-chain diseases related to complex III deficiency', *Biochim Biophys Acta*, 1793(1), pp. 181-5.

Betts, J., Barron, M.J., Needham, S.J., Schaefer, A.M., Taylor, R.W. and Turnbull, D.M. (2008) 'Gastrointestinal tract involvement associated with the 3243A>G mitochondrial DNA mutation', *Neurology*, 70(15), pp. 1290-2.

Blackwood, J.K., Whittaker, R.G., Blakely, E.L., Alston, C.L., Turnbull, D.M. and Taylor, R.W. (2010) 'The investigation and diagnosis of pathogenic mitochondrial DNA mutations in human urothelial cells', *Biochem Biophys Res Commun*, 393(4), pp. 740-5.

Bogenhagen, D. and Clayton, D.A. (1977) 'Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle', *Cell*, 11(4), pp. 719-27.

Borras, C., Sastre, J., Garcia-Sala, D., Lloret, A., Pallardo, F.V. and Vina, J. (2003) 'Mitochondria from females exhibit higher antioxidant gene expression and lower oxidative damage than males', *Free Radic Biol Med*, 34(5), pp. 546-52.

Bowmaker, M., Yang, M.Y., Yasukawa, T., Reyes, A., Jacobs, H.T., Huberman, J.A. and Holt, I.J. (2003) 'Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone', *J Biol Chem*, 278(51), pp. 50961-9.

Box, G.E.P. and Cox, D.R. (1964) 'An Analysis of Transformations', *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), pp. 211-252.

Brockington, M., Alsanjari, N., Sweeney, M.G., Morgan-Hughes, J.A., Scaravilli, F. and Harding, A.E. (1995) 'Kearns-Sayre syndrome associated with mitochondrial DNA deletion or duplication: a molecular genetic and pathological study', *J Neurol Sci*, 131(1), pp. 78-87.

Brown, G.C. and Borutaite, V. (2012) 'There is no evidence that mitochondria are the main source of reactive oxygen species in mammalian cells', *Mitochondrion*, 12(1), pp. 1-4.

Brown, M.D., Voljavec, A.S., Lott, M.T., Torroni, A., Yang, C.C. and Wallace, D.C. (1992) 'Mitochondrial DNA complex I and III mutations associated with Leber's hereditary optic neuropathy', *Genetics*, 130(1), pp. 163-73.

Brown, T.A., Tkachuk, A.N., Shtengel, G., Kopek, B.G., Bogenhagen, D.F., Hess, H.F. and Clayton, D.A. (2011) 'Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction', *Mol Cell Biol*, 31(24), pp. 4994-5010.

Brown, W.M., George, M., Jr. and Wilson, A.C. (1979) 'Rapid evolution of animal mitochondrial DNA', *Proc Natl Acad Sci U S A*, 76(4), pp. 1967-71.

Bruusgaard, J.C. and Gundersen, K. (2008) 'In vivo time-lapse microscopy reveals no loss of murine myonuclei during weeks of muscle atrophy', *Journal of Clinical Investigation*, 118(4), pp. 1450-1457.

Camus, M.F., Clancy, D.J. and Dowling, D.K. (2012) 'Mitochondria, maternal inheritance, and male aging', *Curr Biol*, 22(18), pp. 1717-21.

Carafoli, E. (2003) 'Historical review: mitochondria and calcium: ups and downs of an unusual relationship', *Trends Biochem Sci*, 28(4), pp. 175-81.

Carrozzo, R., Tessa, A., Vazquez-Memije, M.E., Piemonte, F., Patrono, C., Malandrini, A., Dionisi-Vici, C., Vilarinho, L., Villanova, M., Schagger, H., Federico, A., Bertini, E. and Santorelli, F.M. (2001) 'The T9176G mtDNA mutation severely affects ATP production and results in Leigh syndrome', *Neurology*, 56(5), pp. 687-90.

Central Intelligence Agency, U.S. (2013) *The world factbook*. Available at: http://purl.access.gpo.gov/GPO/LPS35389 (Accessed: 18/10/2013).

Chang, D.D. and Clayton, D.A. (1984) 'Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA', *Cell*, 36(3), pp. 635-43.

Chang, D.T., Honick, A.S. and Reynolds, I.J. (2006) 'Mitochondrial trafficking to synapses in cultured primary cortical neurons', *J Neurosci*, 26(26), pp. 7035-45.

Chatre, L. and Ricchetti, M. (2013) 'Prevalent coordination of mitochondrial DNA transcription and initiation of replication with the cell cycle', *Nucleic Acids Research*.

Chen, C., Xiong, N., Wang, Y., Xiong, J., Huang, J., Zhang, Z. and Wang, T. (2012) 'A study of familial MELAS: evaluation of A3243G mutation, clinical phenotype, and magnetic resonance spectroscopy-monitored progression', *Neurol India*, 60(1), pp. 86-9.

Chen, H. and Chan, D.C. (2009) 'Mitochondrial dynamics--fusion, fission, movement, and mitophagy--in neurodegenerative diseases', *Hum Mol Genet*, 18(R2), pp. R169-76.

Cheong, H.I., Chae, J.H., Kim, J.S., Park, H.W., Ha, I.S., Hwang, Y.S., Lee, H.S. and Choi, Y. (1999) 'Hereditary glomerulopathy associated with a mitochondrial tRNA(Leu) gene mutation', *Pediatr Nephrol*, 13(6), pp. 477-80.

Chiarotti, F. (2004) 'Detecting assumption violations in mixed-model analysis of variance', *Ann Ist Super Sanita*, 40(2), pp. 165-71.

Chinnery, P.F., Elliott, C., Green, G.R., Rees, A., Coulthard, A., Turnbull, D.M. and Griffiths, T.D. (2000a) 'The spectrum of hearing loss due to mitochondrial DNA defects', *Brain*, 123(1), pp. 82-92.

Chinnery, P.F., Howell, N., Lightowlers, R.N. and Turnbull, D.M. (1997) 'Molecular pathology of MELAS and MERRF. The relationship between mutation load and clinical phenotypes', *Brain*, 120 (Pt 10), pp. 1713-21.

Chinnery, P.F., Johnson, M.A., Wardell, T.M., Singh-Kler, R., Hayes, C., Brown, D.T., Taylor, R.W., Bindoff, L.A. and Turnbull, D.M. (2000b) 'The epidemiology of pathogenic mitochondrial DNA mutations', *Annals of Neurology*, 48(2), pp. 188-193.

Chinnery, P.F., Zwijnenburg, P.J., Walker, M., Howell, N., Taylor, R.W., Lightowlers, R.N., Bindoff, L. and Turnbull, D.M. (1999) 'Nonrandom tissue distribution of mutant mtDNA', *Am J Med Genet*, 85(5), pp. 498-501.

Chomyn, A., Enriquez, J.A., Micol, V., Fernandez-Silva, P. and Attardi, G. (2000) 'The mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episode syndrome-associated human mitochondrial tRNALeu(UUR) mutation causes aminoacylation deficiency and concomitant reduced association of mRNA with ribosomes', *J Biol Chem*, 275(25), pp. 19198-209.

Chomyn, A., Martinuzzi, A., Yoneda, M., Daga, A., Hurko, O., Johns, D., Lai, S.T., Nonaka, I., Angelini, C. and Attardi, G. (1992) 'MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts', *Proc Natl Acad Sci U S A*, 89(10), pp. 4221-5.

Chomyn, A., Meola, G., Bresolin, N., Lai, S.T., Scarlato, G. and Attardi, G. (1991) 'In vitro genetic transfer of protein synthesis and respiration defects to mitochondrial DNA-less cells with myopathy-patient mitochondria', *Mol Cell Biol*, 11(4), pp. 2236-44.

Christensen, R., Pearson, L.M. and Johnson, W. (1992) 'Case-Deletion Diagnostics for Mixed Models', *Technometrics*, 34(1), pp. 38-45.

Christie, D. (2004) 'Resampling with Excel', *Teaching Statistics*, 26(1), pp. 9-14.

Ciafaloni, E., Ricci, E., Servidei, S., Shanske, S., Silvestri, G., Manfredi, G., Schon, E.A. and DiMauro, S. (1991) 'Widespread tissue distribution of a tRNALeu(UUR) mutation in the mitochondrial DNA of a patient with MELAS syndrome', *Neurology*, 41(10), pp. 1663-4.

Ciafaloni, E., Ricci, E., Shanske, S., Moraes, C.T., Silvestri, G., Hirano, M., Simonetti, S., Angelini, C., Donati, M.A., Garcia, C. and et al. (1992) 'MELAS: clinical features, biochemistry, and molecular genetics', *Ann Neurol*, 31(4), pp. 391-8.

Clapham, D.E. (2007) 'Calcium Signaling', *Cell*, 131(6), pp. 1047-1058.

Clayton, D.A. (1982) 'Replication of animal mitochondrial DNA', *Cell*, 28(4), pp. 693-705.

Cohen, J.E. and Fields, R.D. (2004) 'Extracellular calcium depletion in synaptic transmission', *Neuroscientist*, 10(1), pp. 12-7.

Coku, J., Shanske, S., Mehrazin, M., Tanji, K., Naini, A., Emmanuele, V., Patterson, M., Hirano, M. and DiMauro, S. (2010) 'Slowly progressive encephalopathy with hearing loss due to a mutation in the mtDNA tRNA(Leu(CUN)) gene', *J Neurol Sci*, 290(1-2), pp. 166-8.

Conway, L.J., Robertson, T.E., McGill, J.J. and Hanson, J.P. (2011) 'MELAS syndrome in an Indigenous Australian woman', *Med J Aust*, 195(10), pp. 581-2.

Cook, R.D. and Weisberg, S. (1982) *Residuals and Influence in Regression*. New York: Chapman and Hall.

Cortopassi, G.A., Shibata, D., Soong, N.W. and Arnheim, N. (1992) 'A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues', *Proceedings of the National Academy of Sciences of the United States of America*, 89(16), pp. 7370-7374.

Craven, L., Tuppen, H.A., Greggains, G.D., Harbottle, S.J., Murphy, J.L., Cree, L.M., Murdoch, A.P., Chinnery, P.F., Taylor, R.W., Lightowlers, R.N., Herbert, M. and Turnbull, D.M. (2010) 'Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease', *Nature*, 465(7294), pp. 82-85.

Cree, L.M., Samuels, D.C., De Sousa Lopes, S.C., Rajasimha, H.K., Wonnapinij, P., Mann, J.R., Dahl, H.H.M. and Chinnery, P.F. (2008) 'A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes', *Nature Genetics*, 40(2), pp. 249-254.

Dahlof, C.G., Johansson, M., Casserstedt, S. and Motallebzadeh, T. (2009) 'The course of frequent episodic migraine in a large headache clinic population: a 12-year retrospective follow-up study', *Headache*, 49(8), pp. 1144-52.

Damian, M.S., Seibel, P., Reichmann, H., Schachenmayr, W., Laube, H., Bachmann, G., Wassill, K.H. and Dorndorf, W. (1995) 'Clinical spectrum of the MELAS mutation in a large pedigree', *Acta Neurol Scand*, 92(5), pp. 409-15.

Danial, N.N. and Korsmeyer, S.J. (2004) 'Cell death: critical control points', *Cell*, 116(2), pp. 205-19.

Dar, R., Serlin, R.C. and Omer, H. (1994) 'Misuse of statistical tests in three decades of psychotherapy research', *Journal of Consulting and Clinical Psychology*, 62(1), pp. 75-82.

Davidzon, G., Mancuso, M., Ferraris, S., Quinzii, C., Hirano, M., Peters, H.L., Kirby, D., Thorburn, D.R. and DiMauro, S. (2005) 'POLG mutations and Alpers syndrome', *Ann Neurol*, 57(6), pp. 921-3.

Davis, A.F. and Clayton, D.A. (1996) 'In situ localization of mitochondrial DNA replication in intact mammalian cells', *J Cell Biol*, 135(4), pp. 883-93.

de Grey, A.D. (1997) 'A proposed refinement of the mitochondrial free radical theory of aging', *Bioessays*, 19(2), pp. 161-6.

de Laat, P., Koene, S., van den Heuvel, L.P., Rodenburg, R.J., Janssen, M.C. and Smeitink, J.A. (2012) 'Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A > G mutation', *J Inherit Metab Dis*, 35(6), pp. 1059-69.

de Wit, H.M., Westeneng, H.J., van Engelen, B.G. and Mudde, A.H. (2012) 'MIDD or MELAS : that's not the question MIDD evolving into MELAS : a severe phenotype of the m.3243A>G mutation due to paternal co-inheritance of type 2 diabetes and a high heteroplasmy level', *Neth J Med*, 70(10), pp. 460-2.

Delettre, C., Lenaers, G., Griffoin, J.M., Gigarel, N., Lorenzo, C., Belenguer, P., Pelloquin, L., Grosgeorge, J., Turc-Carel, C., Perret, E., Astarie-Dequeker, C., Lasquellec, L., Arnaud, B., Ducommun, B., Kaplan, J. and Hamel, C.P. (2000) 'Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy', *Nat Genet*, 26(2), pp. 207-10.

Diaz, F., Bayona-Bafaluy, M.P., Rana, M., Mora, M., Hao, H. and Moraes, C.T. (2002) 'Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control', *Nucleic Acids Res*, 30(21), pp. 4626-33.

Diggle, P.J. (1988) 'An Approach to the Analysis of Repeated Measurements', *Biometrics*, 44(4), pp. 959-971.

DiMauro, S. and Hirano, M. (2005) 'Mitochondrial encephalomyopathies: an update', *Neuromuscul Disord*, 15(4), pp. 276-86.

Doleris, L.M., Hill, G.S., Chedin, P., Nochy, D., Bellanne-Chantelot, C., Hanslik, T., Bedrossian, J., Caillat-Zucman, S., Cahen-Varsaux, J. and Bariety, J. (2000) 'Focal segmental glomerulosclerosis associated with mitochondrial cytopathy', *Kidney Int*, 58(5), pp. 1851-1858.

Dougherty, F.E., Ernst, S.G. and Aprille, J.R. (1994) 'Familial recurrence of atypical symptoms in an extended pedigree with the syndrome of mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes (MELAS)', *J Pediatr*, 125(5, Part 1), pp. 758-761.

Dunbar, D.R., Moonie, P.A., Jacobs, H.T. and Holt, I.J. (1995) 'Different cellular backgrounds confer a marked advantage to either mutant or wild-type mitochondrial genomes', *Proc Natl Acad Sci U S A*, 92(14), pp. 6562-6.

Dunbar, D.R., Moonie, P.A., Zeviani, M. and Holt, I.J. (1996) 'Complex I deficiency is associated with 3243G:C mitochondrial DNA in osteosarcoma cell cybrids', *Hum Mol Genet*, 5(1), pp. 123-29.

Ebersbach, G., Baas, H., Csoti, I., Mungersdorf, M. and Deuschl, G. (2006) 'Scales in Parkinson's disease', *J Neurol*, 253 Suppl 4, pp. IV32-5.

Efron, B. (1978) 'Regression and ANOVA with Zero-One Data: Measures of Residual Variation', *Journal of the American Statistical Association*, 73(361), pp. 113-121.

Efron, B. and Tibshirani, R. (1986) 'Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy', *Statistical Science*, 1(1), pp. 54-75.

Ehrlich, S.F., Eskenazi, B., Hedderson, M.M. and Ferrara, A. (2012) 'Sex ratio variations among the offspring of women with diabetes in pregnancy', *Diabet Med*, 29(9), pp. e273-8.

Elliott, H.R., Samuels, D.C., Eden, J.A., Relton, C.L. and Chinnery, P.F. (2008) 'Pathogenic Mitochondrial DNA Mutations Are Common in the General Population', *American Journal of Human Genetics*, 83(2), pp. 254-260.

Elson, J.L., Samuels, D.C., Turnbull, D.M. and Chinnery, P.F. (2001) 'Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age', *Am J Hum Genet*, 68(3), pp. 802-6.

Elson, J.L., Swalwell, H., Blakely, E.L., McFarland, R., Taylor, R.W. and Turnbull, D.M. (2009) 'Pathogenic mitochondrial tRNA mutations--which mutations are inherited and why?', *Hum Mutat*, 30(11), pp. E984-92.

Emma, F., Montini, G., Salviati, L. and Dionisi-Vici, C. (2011) 'Renal mitochondrial cytopathies', *Int J Nephrol*, 2011, p. 609213.

Enns, G.M., Kinsman, S.L., Perlman, S.L., Spicer, K.M., Abdenur, J.E., Cohen, B.H., Amagata, A., Barnes, A., Kheifets, V., Shrader, W.D., Thoolen, M., Blankenberg, F. and Miller, G. (2012) 'Initial experience in the treatment of inherited mitochondrial disease with EPI-743', *Mol Genet Metab*, 105(1), pp. 91-102.

Fan, W., Waymire, K.G., Narula, N., Li, P., Rocher, C., Coskun, P.E., Vannan, M.A., Narula, J., Macgregor, G.R. and Wallace, D.C. (2008) 'A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations', *Science*, 319(5865), pp. 958-62.

Fayet, G., Jansson, M., Sternberg, D., Moslemi, A.R., Blondy, P., Lombes, A., Fardeau, M. and Oldfors, A. (2002) 'Ageing muscle: clonal expansions of mitochondrial DNA point mutations and deletions cause focal impairment of mitochondrial function', *Neuromuscul Disord*, 12(5), pp. 484-93.

Ferguson, S.J. (2010) 'ATP synthase: From sequence to ring size to the P/O ratio', *Proceedings of the National Academy of Sciences*, 107(39), pp. 16755-16756.

Finkel, T. (1998) 'Oxygen radicals and signaling', *Curr Opin Cell Biol*, 10(2), pp. 248-53.

Frank, S.A. (2012) 'Evolution: mitochondrial burden on male health', *Curr Biol*, 22(18), pp. R797-9.

Frank, S.A. and Hurst, L.D. (1996) 'Mitochondria and male disease', *Nature*, 383(6597), pp. 224-224.

Fratter, C., Gorman, G.S., Stewart, J.D., Buddles, M., Smith, C., Evans, J., Seller, A., Poulton, J., Roberts, M., Hanna, M.G., Rahman, S., Omer, S.E., Klopstock, T., Schoser, B., Kornblum, C., Czermin, B., Lecky, B., Blakely, E.L., Craig, K., Chinnery, P.F., Turnbull, D.M., Horvath, R. and Taylor, R.W. (2010) 'The clinical, histochemical, and molecular spectrum of PEO1 (Twinkle)-linked adPEO', *Neurology*, 74(20), pp. 1619-26.

Frederiksen, A.L., Andersen, P.H., Kyvik, K.O., Jeppesen, T.D., Vissing, J. and Schwartz, M. (2006) 'Tissue specific distribution of the 3243A->G mtDNA mutation', *J Med Genet*, 43(8), pp. 671-7.

Frey, T.G. and Mannella, C.A. (2000) 'The internal structure of mitochondria', *Trends Biochem Sci*, 25(7), pp. 319-24.

Fukui, H. and Moraes, C.T. (2009) 'Mechanisms of formation and accumulation of mitochondrial DNA deletions in aging neurons', *Hum Mol Genet*, 18(6), pp. 1028-36.

Gardenier, J.S. and Resnik, D.B. (2002) 'The misuse of statistics: Concepts, tools, and a research agenda', *Accountability in Research*, 9(2), pp. 65-74.

Gardner, M.J., Altman, D.G., Jones, D.R. and Machin, D. (1983) 'Is the statistical assessment of papers submitted to the "British Medical Journal" effective?', *British medical journal (Clinical research ed.)*, 286(6376), pp. 1485-1488.

Gellerich, F.N., Deschauer, M., Chen, Y., Muller, T., Neudecker, S. and Zierz, S. (2002) 'Mitochondrial respiratory rates and activities of respiratory chain complexes correlate linearly with heteroplasmy of deleted mtDNA without threshold and independently of deletion size', *Biochim Biophys Acta*, 1556(1), pp. 41-52.

Gerbitz, K.D., van den Ouweland, J.M., Maassen, J.A. and Jaksch, M. (1995) 'Mitochondrial diabetes mellitus: a review', *Biochim Biophys Acta*, 1271(1), pp. 253-60.

Gerencser, A.A., Chinopoulos, C., Birket, M.J., Jastroch, M., Vitelli, C., Nicholls, D.G. and Brand, M.D. (2012) 'Quantitative measurement of mitochondrial membrane potential in cultured cells: calcium-induced de- and hyperpolarization of neuronal mitochondria', *J Physiol*, 590(Pt 12), pp. 2845-71.

Ghazizadeh, S. and Taichman, L.B. (2001) 'Multiple classes of stem cells in cutaneous epithelium: a lineage analysis of adult mouse skin', *EMBO J*, 20(6), pp. 1215-22.

Gilchrist, J.M., Sikirica, M., Stopa, E. and Shanske, S. (1996) 'Adult-Onset MELAS: Evidence for Involvement of Neurons as Well as Cerebral Vasculature in Strokelike Episodes', *Stroke*, 27(8), pp. 1420-1423.

Gilkerson, R.W., Schon, E.A., Hernandez, E. and Davidson, M.M. (2008) 'Mitochondrial nucleoids maintain genetic autonomy but allow for functional complementation', *Journal of Cell Biology*, 181(7), pp. 1117-1128.

Goodman, S.N., Altman, D.G. and George, S.L. (1998) 'Statistical reviewing policies of medical journals caveat lector?', *Journal of General Internal Medicine*, 13(11), pp. 753-756.

Gore, S.M., Jones, G. and Thompson, S.G. (1992) 'The Lancet's statistical review process: Areas for improvement by authors', *Lancet*, 340(8811), pp. 100-102.

Gore, S.M., Jones, I.G. and Rytter, E.C. (1977) 'Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976', *British Medical Journal*, 1(6053), pp. 85-87.

Goto, Y., Horai, S., Matsuoka, T., Koga, Y., Nihei, K., Kobayashi, M. and Nonaka, I. (1992) 'Mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS): a correlative study of the clinical features and mitochondrial DNA mutation', *Neurology*, 42(3 Pt 1), pp. 545-50.

Goto, Y., Koga, Y., Horai, S. and Nonaka, I. (1990a) 'Chronic progressive external ophthalmoplegia: a correlative study of mitochondrial DNA deletions and their phenotypic expression in muscle biopsies', *J Neurol Sci*, 100(1-2), pp. 63-9.

Goto, Y., Nonaka, I. and Horai, S. (1990b) 'A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies', *Nature*, 348(6302), pp. 651-653.

Goto, Y., Nonaka, I. and Horai, S. (1991) 'A new mtDNA mutation associated with mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS)', *Biochim Biophys Acta*, 1097(3), pp. 238-40.

Greaves, L.C., Yu-Wai-Man, P., Blakely, E.L., Krishnan, K.J., Beadle, N.E., Kerin, J., Barron, M.J., Griffiths, P.G., Dickinson, A.J., Turnbull, D.M. and Taylor, R.W. (2010) 'Mitochondrial DNA defects and selective extraocular muscle involvement in CPEO', *Invest Ophthalmol Vis Sci*, 51(7), pp. 3340-6.

Gregory, T.R. (2001) 'Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma', *Biological Reviews of the Cambridge Philosophical Society*, 76(1), pp. 65-101.

Griffiths, E.J. and Rutter, G.A. (2009) 'Mitochondrial calcium as a key regulator of mitochondrial ATP production in mammalian cells', *Biochim Biophys Acta*, 1787(11), pp. 1324-33.

Guillausseau, P.J., Massin, P., Dubois-LaForgue, D., Timsit, J., Virally, M., Gin, H., Bertin, E., Blickle, J.F., Bouhanick, B., Cahen, J., Caillat-Zucman, S., Charpentier, G., Chedin, P., Derrien, C., Ducluzeau, P.H., Grimaldi, A., Guerci, B., Kaloustian, E., Murat, A., Olivier, F., Paques, M., Paquis-Flucklinger, V., Porokhov, B., Samuel-Lajeunesse, J. and Vialettes, B. (2001) 'Maternally inherited diabetes and deafness: a multicenter study', *Ann Intern Med*, 134(9 Pt 1), pp. 721-8.

Haefeli, R.H., Erb, M., Gemperli, A.C., Robay, D., Courdier Fruh, I., Anklin, C., Dallmann, R. and Gueven, N. (2011) 'NQO1-dependent redox cycling of idebenone: effects on cellular redox potential and energy levels', *PLoS One*, 6(3), p. e17963.

Hagerhall, C. (1997) 'Succinate: quinone oxidoreductases. Variations on a conserved theme', *Biochim Biophys Acta*, 1320(2), pp. 107-41.

Hall, Z.W. and Ralston, E. (1989) 'Nuclear domains in muscle cells', *Cell*, 59(5), pp. 771-772.

Hämäläinen, R.H., Manninen, T., Koivumäki, H., Kislin, M., Otonkoski, T. and Suomalainen, A. (2013) 'Tissue- and cell-type–specific manifestations of heteroplasmic mtDNA 3243A>G mutation in human induced pluripotent stem cell-derived disease model', *Proceedings of the National Academy of Sciences*, 110(38), pp. E3622-E3630.

Hammans, S.R., Sweeney, M.G., Hanna, M.G., Brockington, M., Morgan-Hughes, J.A. and Harding, A.E. (1995) 'The mitochondrial DNA transfer RNALeu(UUR) A-->G(3243) mutation. A clinical and genetic study', *Brain*, 118 (Pt 3), pp. 721-34.

Han, J.M., Jeong, S.J., Park, M.C., Kim, G., Kwon, N.H., Kim, H.K., Ha, S.H., Ryu, S.H. and Kim, S. (2012) 'Leucyl-tRNA synthetase is an intracellular leucine sensor for the mTORC1-signaling pathway', *Cell*, 149(2), pp. 410-24.

Hardie, D.G., Ross, F.A. and Hawley, S.A. (2012) 'AMPK: a nutrient and energy sensor that maintains energy homeostasis', *Nat Rev Mol Cell Biol*, 13(4), pp. 251-262.

Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning*

Data Mining, Inference, and Prediction 2nd edn. New York: Springer.

Hatefi, Y. (1985) 'The mitochondrial electron transport and oxidative phosphorylation system', *Annu Rev Biochem*, 54, pp. 1015-69.

Hayashi, J., Ohta, S., Kikuchi, A., Takemitsu, M., Goto, Y. and Nonaka, I. (1991) 'Introduction of disease-related mitochondrial DNA deletions into HeLa cells lacking mitochondrial DNA results in mitochondrial dysfunction', *Proc Natl Acad Sci U S A*, 88(23), pp. 10614-8.

He, L., Chinnery, P.F., Durham, S.E., Blakely, E.L., Wardell, T.M., Borthwick, G.M., Taylor, R.W. and Turnbull, D.M. (2002) 'Detection and quantification of mitochondrial DNA deletions in individual cells by real-time PCR', *Nucleic Acids Res*, 30(14), p. e68.

Helle, S., Laaksonen, T., Adamsson, A., Paranko, J. and Huitu, O. (2008) 'Female field voles with high testosterone and glucose levels produce male-biased litters', *Animal Behaviour*, 75(3), pp. 1031-1039.

Helm, M., Florentz, C., Chomyn, A. and Attardi, G. (1999) 'Search for differences in post-transcriptional modification patterns of mitochondrial DNA-encoded wild-type and mutant human tRNALys and tRNALeu(UUR)', *Nucleic Acids Res*, 27(3), pp. 756-63.

Hirano, M. and Pavlakis, S.G. (1994) 'Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes (MELAS): current concepts', *J Child Neurol*, 9(1), pp. 4-13.

Hirano, M., Ricci, E., Koenigsberger, M.R., Defendini, R., Pavlakis, S.G., DeVivo, D.C., DiMauro, S. and Rowland, L.P. (1992) 'Melas: an original case and clinical criteria for diagnosis', *Neuromuscul Disord*, 2(2), pp. 125-35.

Hirokawa, K., Utsuyama, M., Hayashi, Y., Kitagawa, M., Makinodan, T. and Fulop, T. (2013) 'Slower immune system aging in women versus men in the Japanese population', *Immun Ageing*, 10(1), p. 19.

Hoekstra, A.S. and Bayley, J.P. (2013) 'The role of complex II in disease', *Biochim Biophys Acta*, 1827(5), pp. 543-51.

Hofer, A.M., Curci, S., Doble, M.A., Brown, E.M. and Soybel, D.I. (2000) 'Intercellular communication mediated by the extracellular calcium-sensing receptor', *Nat Cell Biol*, 2(7), pp. 392-8.

Hollingsworth, K.G., Gorman, G.S., Trenell, M.I., McFarland, R., Taylor, R.W., Turnbull, D.M., MacGowan, G.A., Blamire, A.M. and Chinnery, P.F. (2012) 'Cardiomyopathy is common in patients with the mitochondrial DNA m.3243A>G mutation and correlates with mutation load', *Neuromuscul Disord*, 22(7), pp. 592-6.

Holt, I.J., Harding, A.E., Cooper, J.M., Schapira, A.H., Toscano, A., Clark, J.B. and Morgan-Hughes, J.A. (1989a) 'Mitochondrial myopathies: clinical and biochemical features of 30 patients with major deletions of muscle mitochondrial DNA', *Annals of Neurology*, 26(6), pp. 699-708.

Holt, I.J., Harding, A.E., Cooper, J.M., Schapira, A.H., Toscano, A., Clark, J.B. and Morgan-Hughes, J.A. (1989b) 'Mitochondrial myopathies: clinical and biochemical features of 30 patients with major deletions of muscle mitochondrial DNA', *Ann Neurol*, 26(6), pp. 699-708.

Holt, I.J., Harding, A.E. and Morgan-Hughes, J.A. (1988) 'Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies', *Nature*, 331(6158), pp. 717-719.

Holt, I.J., Lorimer, H.E. and Jacobs, H.T. (2000) 'Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA', *Cell*, 100(5), pp. 515-24.

Hornig-Do, H.T., Montanari, A., Rozanska, A., Tuppen, H.A., Almalki, A.A., Abg-Kamaludin, D.P., Frontali, L., Francisci, S., Lightowlers, R.N. and Chrzanowska-Lightowlers, Z.M. 'Human mitochondrial leucyl tRNA synthetase can suppress non cognate pathogenic mt-tRNA

mutations', *EMBO Molecular Medicine*, Accepted for Publication. In Press.

Horvath, R., Gorman, G. and Chinnery, P.F. (2008) 'How can we treat mitochondrial encephalomyopathies? Approaches to therapy', *Neurotherapeutics*, 5(4), pp. 558-68.

Hotta, O., Inoue, C.N., Miyabayashi, S., Furuta, T., Takeuchi, A. and Taguma, Y. (2001) 'Clinical and pathologic features of focal segmental glomerulosclerosis with mitochondrial tRNALeu(UUR) gene mutation', *Kidney Int*, 59(4), pp. 1236-1243.

Huang, C.C., Chen, R.S., Chu, N.S., Pang, C.Y. and Wei, Y.H. (1996) 'Random mitotic segregation of mitochondrial DNA in MELAS syndrome', *Acta Neurol Scand*, 93(2-3), pp. 198-202.

Hudson, G. and Chinnery, P.F. (2006) 'Mitochondrial DNA polymerase-gamma and human disease', *Hum Mol Genet*, 15 Spec No 2, pp. R244-52.

Huttemann, M., Pecina, P., Rainbolt, M., Sanderson, T.H., Kagan, V.E., Samavati, L., Doan, J.W. and Lee, I. (2011) 'The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis', *Mitochondrion*, 11(3), pp. 369-81.

Ihara, Y., Namba, R., Kuroda, S., Sato, T. and Shirabe, T. (1989) 'Mitochondrial encephalomyopathy (MELAS): pathological study and successful therapy with coenzyme Q10 and idebenone', *J Neurol Sci*, 90(3), pp. 263-71.

Iizuka, T. and Sakai, F. (2005) 'Pathogenesis of stroke-like episodes in MELAS: analysis of neurovascular cellular mechanisms', *Curr Neurovasc Res*, 2(1), pp. 29-45.

Ikeda, K., Shiba, S., Horie-Inoue, K., Shimokata, K. and Inoue, S. (2013) 'A stabilizing factor for mitochondrial respiratory supercomplex assembly regulates energy metabolism in muscle', *Nat Commun*, 4, p. 2147.

Ingjer, F. (1979) 'Effects of endurance training on muscle fibre ATP-ase activity, capillary supply and mitochondrial content in man', *J Physiol*, 294, pp. 419-32.

Innocenti, P., Morrow, E.H. and Dowling, D.K. (2011) 'Experimental Evidence Supports a Sex-Specific Selective Sieve in Mitochondrial Genome Evolution', *Science*, 332(6031), pp. 845-848.

Ishikawa, Y., Goto, Y.I. and Minami, R. (2000) 'Progression in a case of Kearns-Sayre syndrome', *Journal of Child Neurology*, 15(11), pp. 750-755.

Ito, H., Mori, K. and Kagami, S. (2011) 'Neuroimaging of stroke-like episodes in MELAS', *Brain Dev*, 33(4), pp. 283-8.

Iwanicka-Pronicka, K., Pollak, A., Skorka, A., Lechowicz, U., Pajdowska, M., Furmanek, M., Rzeski, M., Korniszewski, L., Skarzynski, H. and Ploski, R. (2012) 'Postlingual hearing loss as a mitochondrial 3243A>G mutation phenotype', *PLoS One*, 7(10), p. e44054.

Iwasaki, N., Babazono, T., Tsuchiya, K., Tomonaga, O., Suzuki, A., Togashi, M., Ujihara, N., Sakka, Y., Yokokawa, H., Ogata, M., Nihei, H. and Iwamoto, Y. (2001) 'Prevalence of A-to-G mutation at nucleotide 3243 of the mitochondrial tRNA(Leu(UUR)) gene in Japanese patients with diabetes mellitus and end stage renal disease', *J Hum Genet*, 46(6), pp. 330-4.

Jacobs, L.J., Jongbloed, R.J., Wijburg, F.A., de Klerk, J.B., Geraedts, J.P., Nijland, J.G., Scholte, H.R., de Coo, I.F. and Smeets, H.J. (2004) 'Pearson syndrome and the role of deletion dimers and duplications in the mtDNA', *J Inherit Metab Dis*, 27(1), pp. 47-55.

Janssen, G.M., Maassen, J.A. and van Den Ouweland, J.M. (1999) 'The diabetes-associated 3243 mutation in the mitochondrial tRNA(Leu(UUR)) gene causes severe mitochondrial dysfunction without a strong decrease in protein synthesis rate', *J Biol Chem*, 274(42), pp. 29744-8.

Jenuth, J.P., Peterson, A.C., Fu, K. and Shoubridge, E.A. (1996) 'Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA', *Nat Genet*, 14(2), pp. 146-51.

Jeppesen, T.D., Schwartz, M., Olsen, D.B., Wibrand, F., Krag, T., Duno, M., Hauerslev, S. and Vissing, J. (2006) 'Aerobic training is safe and improves exercise capacity in patients with mitochondrial myopathy', *Brain*, 129(Pt 12), pp. 3402-12.

Johns, D.R., Rutledge, S.L., Stine, O.C. and Hurko, O. (1989) 'Directly repeated sequences associated with pathogenic mitochondrial DNA deletions', *Proc Natl Acad Sci U S A*, 86(20), pp. 8059-62.

Johnson, R.W. (2001) 'An Introduction to the Bootstrap', *Teaching Statistics*, 23(2), pp. 49-54.

Jokinen, R., Marttinen, P., Sandell, H.K., Manninen, T., Teerenhovi, H., Wai, T., Teoli, D., Loredo-Osti, J.C., Shoubridge, E.A. and Battersby, B.J. (2010) 'Gimap3 regulates tissue-specific mitochondrial DNA segregation', *PLoS Genet*, 6(10), p. e1001161.

Jorm, A.F. (2000) 'Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span', *Psychol Med*, 30(1), pp. 11-22.

Karicheva, O.Z., Kolesnikova, O.A., Schirtz, T., Vysokikh, M.Y., Mager-Heckel, A.M., Lombes, A., Boucheham, A., Krasheninnikov, I.A., Martin, R.P., Entelis, N. and Tarassov, I. (2011) 'Correction of the consequences of mitochondrial 3243A>G mutation in the MT-TL1 gene causing the MELAS syndrome by tRNA import into mitochondria', *Nucleic Acids Res*, 39(18), pp. 8173-86.

Karkare, S., Merchant, S., Solomon, G., Engel, M. and Kosofsky, B. (2009) 'MELAS with A3243G mutation presenting with occipital status epilepticus', *Journal of Child Neurology*, 24(12), pp. 1564-1567.

Karppa, M., Herva, R., Moslemi, A.R., Oldfors, A., Kakko, S. and Majamaa, K. (2005) 'Spectrum of myopathic findings in 50 patients with the 3243A>G mutation in mitochondrial DNA', *Brain*, 128(Pt 8), pp. 1861-9.

Kaufmann, P., Engelstad, K., Wei, Y., Kulikova, R., Oskoui, M., Sproule, D.M., Battista, V., Koenigsberger, D.Y., Pascual, J.M., Shanske, S., Sano, M., Mao, X., Hirano, M., Shungu, D.C., DiMauro, S. and De Vivo, D.C. (2011) 'Natural history of MELAS associated with mitochondrial DNA m.3243A>G genotype', *Neurology*, 77(22), pp. 1965-1971.

Kaufmann, P., Koga, Y., Shanske, S., Hirano, M., DiMauro, S., King, M.P. and Schon, E.A. (1996) 'Mitochondrial DNA and RNA processing in MELAS', *Ann Neurol*, 40(2), pp. 172-80.

Kazak, L., Reyes, A. and Holt, I.J. (2012) 'Minimizing the damage: repair pathways keep mitochondrial DNA intact', *Nat Rev Mol Cell Biol*, 13(10), pp. 659-71.

Khrapko, K., Bodyak, N., Thilly, W.G., van Orsouw, N.J., Zhang, X., Coller, H.A., Perls, T.T., Upton, M., Vijg, J. and Wei, J.Y. (1999) 'Cell-by-cell scanning of whole mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions', *Nucleic Acids Res*, 27(11), pp. 2434-41.

Kim, I., Rodriguez-Enriquez, S. and Lemasters, J.J. (2007) 'Selective degradation of mitochondria by mitophagy', *Arch Biochem Biophys*, 462(2), pp. 245-53.

King, M.P., Koga, Y., Davidson, M. and Schon, E.A. (1992a) 'Defects in mitochondrial protein synthesis and respiratory chain activity segregate with the tRNA(Leu(UUR)) mutation associated with mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes', *Mol Cell Biol*, 12(2), pp. 480-90.

King, M.P., Koga, Y., Davidson, M. and Schon, E.A. (1992b) 'Defects in mitochondrial protein synthesis and respiratory chain activity segregate with the tRNA(Leu(UUR)) mutation associated with mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes', *Molecular and Cellular Biology*, 12(2), pp. 480-490.

Kirichok, Y., Krapivinsky, G. and Clapham, D.E. (2004) 'The mitochondrial calcium uniporter is a highly selective ion channel', *Nature*, 427(6972), pp. 360-364.

Kobayashi, Y., Ichihashi, K., Ohta, S., Nihei, K., Kagawa, Y., Yanagisawa, M. and Momoi, M.Y. (1992) 'The mutant mitochondrial genes in mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS) were selectively amplified through generations', *J Inherit Metab Dis*, 15(5), pp. 803-8.

Kobayashi, Y., Momoi, M.Y., Tominaga, K., Momoi, T., Nihei, K., Yanagisawa, M., Kagawa, Y. and Ohta, S. (1990) 'A point mutation in the mitochondrial tRNA(Leu)(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes)', *Biochem Biophys Res Commun*, 173(3), pp. 816-22.

Kobayashi, Y., Momoi, M.Y., Tominaga, K., Shimoizumi, H., Nihei, K., Yanagisawa, M., Kagawa, Y. and Ohta, S. (1991) 'Respiration-deficient cells are caused by a single point mutation in the mitochondrial tRNA-Leu (UUR) gene in mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes (MELAS)', *Am J Hum Genet*, 49(3), pp. 590-9.

Koga, Y., Akita, Y., Takane, N., Sato, Y. and Kato, H. (2000) 'Heterogeneous presentation in A3243G mutation in the mitochondrial tRNA(Leu(UUR)) gene', *Arch Dis Child*, 82(5), pp. 407-11.

Koga, Y., Davidson, M., Schon, E.A. and King, M.P. (1993) 'Fine mapping of mitochondrial RNAs derived from the mtDNA region containing a point mutation associated with MELAS', *Nucleic Acids Res*, 21(3), pp. 657-62.

Kolb, S.J., Costello, F., Lee, A.G., White, M., Wong, S., Schwartz, E.D., Messé, S.R., Ellenbogen, J., Kasner, S.E. and Galetta, S.L. (2003) 'Distinguishing ischemic stroke from the stroke-like lesions of MELAS using apparent diffusion coefficient mapping', *Journal of the Neurological Sciences*, 216(1), pp. 11-15.

Kornblum, C., Nicholls, T.J., Haack, T.B., Scholer, S., Peeva, V., Danhauser, K., Hallmann, K., Zsurka, G., Rorbach, J., Iuso, A., Wieland, T., Sciacco, M., Ronchi, D., Comi, G.P., Moggio, M., Quinzii, C.M., DiMauro, S., Calvo, S.E., Mootha, V.K., Klopstock, T., Strom, T.M., Meitinger, T., Minczuk, M., Kunz, W.S. and Prokisch, H. (2013) 'Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease', *Nat Genet*, 45(2), pp. 214-219.

Korr, H., Kurz, C., Seidler, T.O., Sommer, D. and Schmitz, C. (1998) 'Mitochondrial DNA synthesis studied autoradiographically in various cell types in vivo', *Braz J Med Biol Res*, 31(2), pp. 289-98.

Kostyuk, P.G. (2007) 'Key role of calcium signaling in synaptic transmission', *Neurophysiology*, 39(4-5), pp. 248-250.

Kraytsberg, Y., Kudryavtseva, E., McKee, A.C., Geula, C., Kowall, N.W. and Khrapko, K. (2006) 'Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons', *Nature Genetics*, 38(5), pp. 518-520.

Kreft, I.G.G. and de Leew, J. (1998) *Introducing multilevel modeling*. London: Sage.

Krishnan, K.J., Bender, A., Taylor, R.W. and Turnbull, D.M. (2007) 'A multiplex real-time PCR method to detect and quantify mitochondrial DNA deletions in individual cells', *Anal Biochem*, 370(1), pp. 127-9.

Krishnan, K.J., Reeve, A.K., Samuels, D.C., Chinnery, P.F., Blackwood, J.K., Taylor, R.W., Wanrooij, S., Spelbrink, J.N., Lightowlers, R.N. and Turnbull, D.M. (2008) 'What causes mitochondrial DNA deletions in human cells?', *Nature Genetics*, 40(3), pp. 275-279.

Kuha, J. (2004) 'AIC and BIC', *Sociological Methods & Research*, 33(2), pp. 188-229.

Kukat, C., Wurm, C.A., Spahr, H., Falkenberg, M., Larsson, N.G. and Jakobs, S. (2011) 'Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA', *Proc Natl Acad Sci U S A*, 108(33), pp. 13534-9.

Laird, N.M. and Ware, J.H. (1982) 'Random-effects models for longitudinal data', *Biometrics*, 38(4), pp. 963-974.

Laloi-Michelin, M., Meas, T., Ambonville, C., Bellanne-Chantelot, C., Beaufils, S., Massin, P., Vialettes, B., Gin, H., Timsit, J., Bauduceau, B., Bernard, L., Bertin, E., Blickle, J.F., Cahen-Varsaux, J., Cailleba, A., Casanova, S., Cathebras, P., Charpentier, G., Chedin, P., Crea, T., Delemer, B., Dubois-Laforgue, D., Duchemin, F., Ducluzeau, P.H., Bouhanick, B., Dusselier, L., Gabreau, T., Grimaldi, A., Guerci, B., Jacquin, V., Kaloustian, E., Larger, E., Lecleire-Collet, A., Lorenzini, F., Louis, J., Mausset, J., Murat, A., Nadler-Fluteau, S., Olivier, F., Paquis-Flucklinger, V., Paris-Bockel, D., Raynaud, I., Reznik, Y., Riveline, J.P., Schneebeli, S., Sonnet, E., Sola-Gazagnes, A., Thomas, J.L., Trabulsi, B., Virally, M. and Guillausseau, P.J. (2009) 'The clinical variability of maternally inherited diabetes and deafness is associated with the degree of heteroplasmy in blood leukocytes', *J Clin Endocrinol Metab*, 94(8), pp. 3025-30.

Landing, B.H., Dixon, L.G. and Wells, T.R. (1974) 'Studies on isolated human skeletal muscle fibers: Including a proposed pattern of nuclear distribution and a concept of nuclear territories', *Hum Pathol*, 5(4), pp. 441-461.

Larsson, N.G. and Holme, E. (1992) 'Multiple short direct repeats associated with single mtDNA deletions', *Biochim Biophys Acta*, 1139(4), pp. 311-4.

Lax, N.Z., Hepplewhite, P.D., Reeve, A.K., Nesbitt, V., McFarland, R., Jaros, E., Taylor, R.W. and Turnbull, D.M. (2012) 'Cerebellar ataxia in patients with mitochondrial DNA disease: a molecular clinicopathological study', *J Neuropathol Exp Neurol*, 71(2), pp. 148-61.

Legros, F., Lombes, A., Frachon, P. and Rojo, M. (2002) 'Mitochondrial fusion in human cells is efficient, requires the inner membrane potential, and is mediated by mitofusins', *Mol Biol Cell*, 13(12), pp. 4343-54.

Leigh, T.F. and Thompson, E.A. (1951) 'Pulmonary metastatic sarcoma with associated pneumothorax', *Am J Roentgenol Radium Ther*, 66(6), pp. 900-2.

Lewis, F., Butler, A. and Gilbert, L. (2011) 'A unified approach to model selection using the likelihood ratio test', *Methods in Ecology and Evolution*, 2(2), pp. 155-162.

Li, M.Z., Yu, D.M., Yu, P., Liu, D.M., Wang, K. and Tang, X.Z. (2008) 'Mitochondrial gene mutations and type 2 diabetes in Chinese families', *Chin Med J (Engl)*, 121(8), pp. 682-6.

Li, R. and Guan, M.X. (2010) 'Human mitochondrial leucyl-tRNA synthetase corrects mitochondrial dysfunctions due to the tRNALeu(UUR) A3243G mutation, associated with mitochondrial encephalomyopathy, lactic acidosis, and stroke-like symptoms and diabetes', *Mol Cell Biol*, 30(9), pp. 2147-54.

Li, Z., Okamoto, K., Hayashi, Y. and Sheng, M. (2004) 'The importance of dendritic mitochondria in the morphogenesis and plasticity of spines and synapses', *Cell*, 119(6), pp. 873-87.

Liesa, M., Palacin, M. and Zorzano, A. (2009) 'Mitochondrial dynamics in mammalian health and disease', *Physiol Rev*, 89(3), pp. 799-845.

Lightowlers, R.N., Chinnery, P.F., Turnbull, D.M. and Howell, N. (1997) 'Mammalian mitochondrial genetics: heredity, heteroplasmy and disease', *Trends Genet*, 13(11), pp. 450-5.

Lindsey, J.K. (1993) *Models for repeated measurements*. Clarendon Press, UK.

Liou, C.W., Huang, C.C., Chee, E.C., Jong, Y.J., Tsai, J.L., Pang, C.Y., Lee, H.C. and Wei, Y.H. (1994) 'MELAS syndrome: correlation between clinical features and molecular genetic analysis', *Acta Neurol Scand*, 90(5), pp. 354-9.

Littell, R.C., Pendergast, J. and Natarajan, R. (2000) 'Modelling covariance structure in the analysis of repeated measures data', *Stat Med*, 19(13), pp. 1793-1819.

Liu, H., Ma, Y., Fang, F., Zhang, Y., Zou, L., Yang, Y., Zhu, S., Wang, S., Zheng, X., Pei, P., Li, L., Wu, H., Xiao, Y., Xu, Y., Wang, L., Cao, Y., Pan, H. and Qi, Y. (2013) 'Wild-Type Mitochondrial DNA Copy Number in Urinary Cells as a Useful Marker for Diagnosing Severity of the Mitochondrial Diseases', *PLoS One*, 8(6), p. e67146.

Liu, J.X., Höglund, A.S., Karlsson, P., Lindblad, J., Qaisar, R., Aare, S., Bengtsson, E. and Larsson, L. (2009) 'Myonuclear domain size and myosin isoform expression in muscle fibres from mammals representing a 100 000-fold difference in body size', *Experimental Physiology*, 94(1), pp. 117-129.

López-Gallardo, E., López-Pérez, M.J., Montoya, J. and Ruiz-Pesini, E. (2009) 'CPEO and KSS differ in the percentage and location of the mtDNA deletion', *Mitochondrion*, 9(5), pp. 314-317.

Luo, S.M., Ge, Z.J., Wang, Z.W., Jiang, Z.Z., Wang, Z.B., Ouyang, Y.C., Hou, Y., Schatten, H. and Sun, Q.Y. (2013) 'Unique insights into maternal mitochondrial inheritance in mice', *Proc Natl Acad Sci U S A*, 110(32), pp. 13038-43.

Ma, Y., Fang, F., Yang, Y., Zou, L., Zhang, Y., Wang, S., Xu, Y., Pei, P. and Qi, Y. (2009) 'The study of mitochondrial A3243G mutation in different samples', *Mitochondrion*, 9(2), pp. 139-143.

Macchiavelli, R.E. and Moser, E.B. (1997) 'Analysis of Repeated Measurements with Ante-Dependence Covariance Models', *Biometrical Journal*, 39(3), pp. 339-350.

Macmillan, C., Lach, B. and Shoubridge, E.A. (1993) 'Variable distribution of mutant mitochondrial DNAs (tRNA(Leu[3243])) in tissues of symptomatic relatives with MELAS: the role of mitotic segregation', *Neurology*, 43(8), pp. 1586-90.

Magnusson, J., Orth, M., Lestienne, P. and Taanman, J.W. (2003) 'Replication of mitochondrial DNA occurs throughout the mitochondria of cultured human cells', *Exp Cell Res*, 289(1), pp. 133-42.

Majamaa-Voltti, K., Peuhkurinen, K., Kortelainen, M.L., Hassinen, I.E. and Majamaa, K. (2002) 'Cardiac abnormalities in patients with mitochondrial DNA mutation 3243A>G', *BMC Cardiovasc Disord*, 2, p. 12.

Majamaa-Voltti, K.A., Winqvist, S., Remes, A.M., Tolonen, U., Pyhtinen, J., Uimonen, S., Karppa, M., Sorri, M., Peuhkurinen, K. and Majamaa, K. (2006) 'A 3-year clinical follow-up of adult patients with 3243A>G in mitochondrial DNA', *Neurology*, 66(10), pp. 1470-5.

Majamaa, K., Moilanen, J.S., Uimonen, S., Remes, A.M., Salmela, P.I., Karppa, M., Majamaa-Voltti, K.A., Rusanen, H., Sorri, M., Peuhkurinen, K.J. and Hassinen, I.E. (1998) 'Epidemiology of A3243G, the mutation for mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes: prevalence of the mutation in an adult population', *Am J Hum Genet*, 63(2), pp. 447-54.

Malik, A.N., Shahni, R., Rodriguez-de-Ledesma, A., Laftah, A. and Cunningham, P. (2011) 'Mitochondrial DNA as a non-invasive biomarker: Accurate quantification using real time quantitative PCR without co-amplification of pseudogenes and dilution bias', *Biochemical and Biophysical Research Communications*, 412(1), pp. 1-7.

Mancuso, M., Orsucci, D., Angelini, C., Bertini, E., Carelli, V., Comi, G.P., Minetti, C., Moggio, M., Mongini, T., Servidei, S., Tonin, P., Toscano, A., Uziel, G., Bruno, C., Caldarazzo Ienco, E., Filosto, M., Lamperti, C., Martinelli, D., Moroni, I., Musumeci, O., Pegoraro, E., Ronchi, D., Santorelli, F.M., Sauchelli, D., Scarpelli, M., Sciacco, M., Spinazzi, M., Valentino, M.L., Vercelli, L., Zeviani, M. and Siciliano, G. (2013a) 'Phenotypic heterogeneity of the 8344A>G mtDNA "MERRF" mutation', *Neurology*, 80(22), pp. 2049-54.

Mancuso, M., Orsucci, D., Ienco, E.C., Pini, E., Choub, A. and Siciliano, G. (2013b) 'Psychiatric involvement in adult patients with mitochondrial disease', *Neurol Sci*, 34(1), pp. 71-4.

Manfredi, G., Vu, T., Bonilla, E., Schon, E.A., DiMauro, S., Arnaudo, E., Zhang, L., Rowland, L.P. and Hirano, M. (1997) 'Association of myopathy with large-scale mitochondrial DNA duplications and deletions: which is pathogenic?', *Ann Neurol*, 42(2), pp. 180-8.

Manwaring, N., Jones, M.M., Wang, J.J., Rochtchina, E., Howard, C., Mitchell, P. and Sue, C.M. (2007) 'Population prevalence of the MELAS A3243G mutation', *Mitochondrion*, 7(3), pp. 230-3.

Margulis, L. (1971) 'Symbiosis and evolution', *Sci Am*, 225(2), pp. 48-57.

Marriage, B., Clandinin, M.T. and Glerum, D.M. (2003) 'Nutritional cofactor treatment in mitochondrial disorders', *Journal of the American Dietetic Association*, 103(8), pp. 1029-1038.

Marshall, S.W. (2004) 'Testing with confidence: The use (and misuse) of confidence intervals in biomedical research', *Journal of Science and Medicine in Sport*, 7(2), pp. 135-137.

Martin Negrier, M.L., Coquet, M., Moretto, B.T., Lacut, J.Y., Dupon, M., Bloch, B., Lestienne, P. and Vital, C. (1998) 'Partial triplication of mtDNA in maternally transmitted diabetes mellitus and deafness', *Am J Hum Genet*, 63(4), pp. 1227-32.

Martin, W. and Muller, M. (1998) 'The hydrogen hypothesis for the first eukaryote', *Nature*, 392(6671), pp. 37-41.

Martinuzzi, A., Bartolomei, L., Carrozzo, R., Mostacciuolo, M., Carbonin, C., Toso, V., Ciafaloni, E., Shanske, S., DiMauro, S. and Angelini, C. (1992) 'Correlation between clinical and molecular features in two MELAS families', *J Neurol Sci*, 113(2), pp. 222-9.

Massin, P., Virally-Monod, M., Vialettes, B., Paques, M., Gin, H., Porokhov, B., Caillat-Zucman, S., Froguel, P., Paquis-Fluckinger, V., Gaudric, A. and Guillausseau, P.-J. (1999) 'Prevalence of macular pattern dystrophy in maternally inherited diabetes and deafness', *Ophthalmology*, 106(9), pp. 1821-1827.

Matthews, P.M., Hopkin, J., Brown, R.M., Stephenson, J.B., Hilton-Jones, D. and Brown, G.K. (1994) 'Comparison of the relative levels of the 3243 (A-->G) mtDNA mutation in heteroplasmic adult and fetal tissues', *J Med Genet*, 31(1), pp. 41-4.

Mazzaccara, C., Iafusco, D., Liguori, R., Ferrigno, M., Galderisi, A., Vitale, D., Simonelli, F., Landolfo, P., Prisco, F., Masullo, M. and Sacchetti, L. (2012) 'Mitochondrial diabetes in children: seek and you will find it', *PLoS One*, 7(4), p. e34956.

McCullagh, P. (1980) 'Regression models for ordinal data (with Discussion)', *Journal of the Royal Statistical Society*, Series B(42), pp. 109-142.

McFadden, D. (1974) 'The measurement of urban travel demand', *Journal of Public Economics*, 3(4), pp. 303-328.

McFarland, R. and Turnbull, D.M. (2009) 'Batteries not included: diagnosis and management of mitochondrial disease', *J Intern Med*, 265(2), pp. 210-28.

McKinney, E.A. and Oliveira, M.T. (2013) 'Replicating animal mitochondrial DNA', *Genet Mol Biol*, 36(3), pp. 308-315.

Meeusen, S., McCaffery, J.M. and Nunnari, J. (2004) 'Mitochondrial fusion intermediates revealed in vitro', *Science*, 305(5691), pp. 1747-52.

Mehrazin, M., Shanske, S., Kaufmann, P., Wei, Y., Coku, J., Engelstad, K., Naini, A., De Vivo, D.C. and DiMauro, S. (2009) 'Longitudinal changes of mtDNA A3243G mutation load and level of functioning in MELAS', *Am J Med Genet A*, 149A(4), pp. 584-7.

Melov, S., Shoffner, J.M., Kaufman, A. and Wallace, D.C. (1995) 'Marked increase in the number and variety of mitochondrial DNA rearrangements in aging human skeletal muscle', *Nucleic Acids Res*, 23(20), pp. 4122-6.

Menard, S. (2000) 'Coefficients of Determination for Multiple Logistic Regression Analysis', *The American Statistician*, 54(1), pp. 17-24.

Miller, K.E. and Sheetz, M.P. (2004) 'Axonal mitochondrial transport and potential are correlated', *J Cell Sci*, 117(Pt 13), pp. 2791-804.

Mimaki, M., Wang, X., McKenzie, M., Thorburn, D.R. and Ryan, M.T. (2012) 'Understanding mitochondrial complex I assembly in health and disease', *Biochim Biophys Acta*, 1817(6), pp. 851-62.

Mita, S., Schmidt, B., Schon, E.A., DiMauro, S. and Bonilla, E. (1989) 'Detection of 'deleted' mitochondrial genomes in cytochrome-c oxidase-deficient muscle fibers of a patient with Kearns-Sayre syndrome', *Proceedings of the National Academy of Sciences of the United States of America*, 86(23), pp. 9509-9513.

Miyabayashi, S., Hanamizu, H., Nakamura, R., Endo, H. and Tada, K. (1992) 'Defects of mitochondrial respiratory enzymes in cloned cells from MELAS fibroblasts', *Journal of Inherited Metabolic Disease*, 15(5), pp. 797-802.

Montoya, J., Christianson, T., Levens, D., Rabinowitz, M. and Attardi, G. (1982) 'Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA', *Proc Natl Acad Sci U S A*, 79(23), pp. 7195-9.

Moraes, C.T., Ciacci, F., Silvestri, G., Shanske, S., Sciacco, M., Hirano, M., Schon, E.A., Bonilla, E. and DiMauro, S. (1993) 'Atypical clinical presentations associated with the MELAS mutation at position 3243 of human mitochondrial DNA', *Neuromuscul Disord*, 3(1), pp. 43-50.

Moraes, C.T., DiMauro, S., Zeviani, M., Lombes, A., Shanske, S., Miranda, A.F., Nakase, H., Bonilla, E., Werneck, L.C., Servidei, S., Nonaka, I., Koga, Y., Spiro, A.J., Brownell, A.K.W., Schmidt, B., Schotland, D.L., Zupanc, M. and DeVivo, D.C. (1989) 'Mitochondrial DNA deletions in progressive external ophthalmoplegia and Kearne-Sayre syndrome', *New England Journal of Medicine*, 320(20), pp. 1293-1299.

Moraes, C.T., Ricci, E., Bonilla, E., DiMauro, S. and Schon, E.A. (1992) 'The mitochondrial tRNA(Leu(UUR)) mutation in mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes (MELAS): genetic, biochemical, and morphological correlations in skeletal muscle', *Am J Hum Genet*, 50(5), pp. 934-49.

Moraes, C.T., Shanske, S., Tritschler, H.J., Aprille, J.R., Andreetta, F., Bonilla, E., Schon, E.A. and DiMauro, S. (1991) 'mtDNA depletion with variable tissue expression: A novel genetic abnormality in mitochondrial diseases', *American Journal of Human Genetics*, 48(3), pp. 492-501.

Moreno-Lastres, D., Fontanesi, F., Garcia-Consuegra, I., Martin, M.A., Arenas, J., Barrientos, A. and Ugalde, C. (2012) 'Mitochondrial complex I plays an essential role in human respirasome assembly', *Cell Metab*, 15(3), pp. 324-35.

Morgan-Hughes, J.A., Sweeney, M.G., Cooper, J.M., Hammans, S.R., Brockington, M., Schapira, A.H., Harding, A.E. and Clark, J.B. (1995) 'Mitochondrial DNA (mtDNA) diseases: correlation of genotype to phenotype', *Biochim Biophys Acta*, 1271(1), pp. 135-40.

Morris, R.L. and Hollenbeck, P.J. (1993) 'The regulation of bidirectional mitochondrial transport is coordinated with axonal outgrowth', *J Cell Sci*, 104 (Pt 3), pp. 917-27.

Moser, E.B. (2004) 'Repeated measures modeling with PROC MIXED', *Proceedings of the 29th SAS Users Group International Conference*.

Murphy, J.L., Blakely, E.L., Schaefer, A.M., He, L., Wyrick, P., Haller, R.G., Taylor, R.W., Turnbull, D.M. and Taivassalo, T. (2008) 'Resistance training in patients with single, large-scale deletions of mitochondrial DNA', *Brain*, 131(Pt 11), pp. 2832-40.

Murphy, M.P. (2009) 'How mitochondria produce reactive oxygen species', *Biochem J*, 417(1), pp. 1-13.

Myers, J.H. (1978) 'Sex Ratio Adjustment Under Food Stress: Maximization of Quality or Numbers of Offspring?', *The American Naturalist*, 112(984), pp. 381-388.

Nagele, P. (2003) 'Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals', *British Journal of Anaesthesia*, 90(4), pp. 514-516.

Nakamura, M., Yabe, I., Sudo, A., Hosoki, K., Yaguchi, H., Saitoh, S. and Sasaki, H. (2010) 'MERRF/MELAS overlap syndrome: a double pathogenic mutation in mitochondrial tRNA genes', *J Med Genet*, 47(10), pp. 659-64.

Nakase, H., Moraes, C.T., Rizzuto, R., Lombes, A., DiMauro, S. and Schon, E.A. (1990) 'Transcription and translation of deleted mitochondrial genomes in Kearns-Sayre syndrome: Implications for pathogenesis', *American Journal of Human Genetics*, 46(3), pp. 418-427.

Nelder, J. and Wedderburn, R. (1972) 'Generalized Linear Models', *Journal of the Royal Statistical Society*, Series B(135), pp. 370-384.

Nesbitt, V., Pitceathly, R.D., Turnbull, D.M., Taylor, R.W., Sweeney, M.G., Mudanohwo, E.E., Rahman, S., Hanna, M.G. and McFarland, R. (2013) 'The UK MRC Mitochondrial Disease Patient Cohort Study: clinical phenotypes associated with the m.3243A>G mutation--implications for diagnosis and management', *J Neurol Neurosurg Psychiatry*, 84(8), pp. 936-8.

Nishino, I., Spinazzola, A. and Hirano, M. (2001) 'MNGIE: from nuclear DNA to mitochondrial DNA', *Neuromuscul Disord*, 11(1), pp. 7-10.

Norddahl, G.L., Pronk, C.J., Wahlestedt, M., Sten, G., Nygren, J.M., Ugale, A., Sigvardsson, M. and Bryder, D. (2011) 'Accumulating mitochondrial DNA mutations

drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging', *Cell Stem Cell*, 8(5), pp. 499-510.

O'Brien, T.W. (2003) 'Properties of human mitochondrial ribosomes', *IUBMB Life*, 55(9), pp. 505-13.

Obermaier-Kusser, B., Paetzke-Brunner, I., Enter, C., Muller-Hocker, J., Zierz, S., Ruitenbeek, W. and Gerbitz, K.D. (1991) 'Respiratory chain activity in tissues from patients (MELAS) with a point mutation of the mitochondrial genome [tRNA(Leu(UUR))]', *FEBS Lett*, 286(1-2), pp. 67-70.

Ohira, Y., Yoshinaga, T., Ohara, M., Nonaka, I., Yoshioka, T., Yamashita-Goto, K., Shenkman, B.S., Kozlovskaya, I.B., Roy, R.R. and Edgerton, V.R. (1999) 'Myonuclear domain and myosin phenotype in human soleus after bed rest with or without loading', *Journal of Applied Physiology*, 87(5), pp. 1776-1785.

Ojala, D., Montoya, J. and Attardi, G. (1981) 'tRNA punctuation model of RNA processing in human mitochondria', *Nature*, 290(5806), pp. 470-4.

Okazaki, K. and Holtzer, H. (1966) 'Myogenesis: fusion, myosin synthesis, and the mitotic cycle', *Proc Natl Acad Sci U S A*, 56(5), pp. 1484-90.

Old, S.L. and Johnson, M.A. (1989) 'Methods of microphotometric assay of succinate dehydrogenase and cytochrome c oxidase activities for use on human skeletal muscle', *Histochem J*, 21(9-10), pp. 545-55.

Oldfors, A., Larsson, N.G., Holme, E., Tulinius, M., Kadenbach, B. and Droste, M. (1992) 'Mitochondrial DNA deletions and cytochrome c oxidase deficiency in muscle fibers', *Journal of the Neurological Sciences*, 110(1-2), pp. 169-177.

ONS (2010) *United Kingdom Health Statistics 2010*. London, UK: Office for National Statistics.

Oppenheim, C., Galanaud, D., Samson, Y., Sahel, M., Dormont, D., Wechsler, B. and Marsault, C. (2000) 'Can diffusion weighted magnetic resonance imaging help differentiate stroke from stroke-like events in MELAS?', *Journal of Neurology Neurosurgery and Psychiatry*, 69(2), pp. 248-250.

Orrenius, S., Zhivotovsky, B. and Nicotera, P. (2003) 'Regulation of cell death: the calcium-apoptosis link', *Nat Rev Mol Cell Biol*, 4(7), pp. 552-65.

Orsucci, D., Calsolaro, V., Siciliano, G. and Mancuso, M. (2012) 'Quality of life in adult patients with mitochondrial myopathy', *Neuroepidemiology*, 38(3), pp. 194-5.

Palade, G.E. (1953) 'An electron microscope study of the mitochondrial structure', *J Histochem Cytochem*, 1(4), pp. 188-211.

Paradkar, P.N., Zumbrennen, K.B., Paw, B.H., Ward, D.M. and Kaplan, J. (2009) 'Regulation of mitochondrial iron import through differential turnover of mitoferrin 1 and mitoferrin 2', *Mol Cell Biol*, 29(4), pp. 1007-16.

Park, H., Davidson, E. and King, M.P. (2003) 'The pathogenic A3243G mutation in human mitochondrial tRNALeu(UUR) decreases the efficiency of aminoacylation', *Biochemistry*, 42(4), pp. 958-64.

Park, H., Davidson, E. and King, M.P. (2008) 'Overexpressed mitochondrial leucyl-tRNA synthetase suppresses the A3243G mutation in the mitochondrial tRNA(Leu(UUR)) gene', *RNA*, 14(11), pp. 2407-16.

Parsch, J. (2011) 'The Cost of Being Male', *Science*, 332(6031), pp. 798-799.

Patergnani, S., Suski, J.M., Agnoletto, C., Bononi, A., Bonora, M., De Marchi, E., Giorgi, C., Marchi, S., Missiroli, S., Poletti, F., Rimessi, A., Duszynski, J., Wieckowski, M.R. and Pinton, P. (2011) 'Calcium signaling around Mitochondria Associated Membranes (MAMs)', *Cell Commun Signal*, 9, p. 19.

Perli, E., Giordano, C., Pisano, A., Montanari, A., Campese, A.F., Reyes, A., Ghezzi, D., Nasca, A., Tuppen, H.A., Orlandi, M., Di Micco, P., Poser, E., Taylor, R.W., Colotti, G., Francisci, S., Morea, V., Frontali, L., Zeviani, M. and d'Amati, G. 'The isolated carboxy-terminal domain of human mitochondrial leucyl-tRNA synthetase

rescues the pathological phenotype of mitochondrial tRNA mutations in human cells', *EMBO Molecular Medicine*, Accepted for Publication. In Press.

Perli, E., Giordano, C., Tuppen, H.A., Montopoli, M., Montanari, A., Orlandi, M., Pisano, A., Catanzaro, D., Caparrotta, L., Musumeci, B., Autore, C., Morea, V., Di Micco, P., Campese, A.F., Leopizzi, M., Gallo, P., Francisci, S., Frontali, L., Taylor,

R.W. and d'Amati, G. (2012) 'Isoleucyl-tRNA synthetase levels modulate the penetrance of a homoplasmic m.4277T>C mitochondrial tRNA(Ile) mutation causing hypertrophic cardiomyopathy', *Hum Mol Genet*, 21(1), pp. 85-100.

Petruzzella, V., Moraes, C.T., Sano, M.C., Bonilla, E., DiMauro, S. and Schon, E.A. (1994) 'Extremely high levels of mutant mtDNAs co-localize with cytochrome c oxidase-negative ragged-red fibers in patients harboring a point mutation at nt 3243', *Human Molecular Genetics*, 3(3), pp. 449-454.

Pfeffer, G., Majamaa, K., Turnbull, D.M., Thorburn, D. and Chinnery, P.F. (2012) 'Treatment for mitochondrial disorders', *Cochrane Database Syst Rev*, 4, p. CD004426.

Phoenix, C., Schaefer, A.M., Elson, J.L., Morava, E., Bugiani, M., Uziel, G., Smeitink, J.A., Turnbull, D.M. and McFarland, R. (2006) 'A scale to monitor progression and treatment of mitochondrial disease in children', *Neuromuscul Disord*, 16(12), pp. 814-20.

Piccoli, G.B., Bonino, L.D., Campisi, P., Vigotti, F.N., Ferraresi, M., Fassio, F., Brocheriou, I., Porpiglia, F. and Restagno, G. (2012) 'Chronic kidney disease, severe arterial and arteriolar sclerosis and kidney neoplasia: on the spectrum of kidney involvement in MELAS syndrome', *BMC Nephrol*, 13, p. 9.

Pietras, E.M., Warr, M.R. and Passegue, E. (2011) 'Cell cycle regulation in hematopoietic stem cells', *J Cell Biol*, 195(5), pp. 709-20.

Porter, A.M.W. (1999) 'Misuse of correlation and regression in three medical journals', *Journal of the Royal Society of Medicine*, 92(3), pp. 123-128.

Poulton, J., Deadman, M.E. and Gardiner, R.M. (1989) 'Duplications of mitochondrial DNA in mitochondrial myopathy', *Lancet*, 1(8632), pp. 236-40.

Poulton, J. and Holt, I.J. (1994) 'Mitochondrial DNA: does more lead to less?', *Nat Genet*, 8(4), pp. 313-5.

Poulton, J. and Morten, K. (1993) 'Noninvasive diagnosis of the MELAS syndrome from blood DNA', *Annals of Neurology*, 34(1), pp. 116-116.

Poulton, J., O'Rahilly, S., Morten, K.J. and Clark, A. (1995) 'Mitochondrial DNA, diabetes and pancreatic pathology in Kearns-Sayre syndrome', *Diabetologia*, 38(7), pp. 868-71.

Proteggente, A.R., England, T.G., Rehman, A., Rice-Evans, C.A. and Halliwell, B. (2002) 'Gender differences in steady-state levels of oxidative damage to DNA in healthy individuals', *Free Radic Res*, 36(2), pp. 157-62.

Pyle, A., Taylor, R.W., Durham, S.E., Deschauer, M., Schaefer, A.M., Samuels, D.C. and Chinnery, P.F. (2007) 'Depletion of mitochondrial DNA in leucocytes harbouring the 3243A->G mtDNA mutation', *J Med Genet*, 44(1), pp. 69-74.

Raap, A.K., Jahangir Tafrechi, R.S., van de Rijke, F.M., Pyle, A., Wählby, C., Szuhai, K., Ravelli, R.B.G., de Coo, R.F.M., Rajasimha, H.K., Nilsson, M., Chinnery, P.F., Samuels, D.C. and Janssen, G.M.C. (2012) 'Non-Random mtDNA Segregation Patterns Indicate a Metastable Heteroplasmic Segregation Unit in m.3243A>G Cybrid Cells', *PLoS One*, 7(12), p. e52080.

Rahman, S., Blok, R.B., Dahl, H.H., Danks, D.M., Kirby, D.M., Chow, C.W., Christodoulou, J. and Thorburn, D.R. (1996) 'Leigh syndrome: clinical features and biochemical and DNA abnormalities', *Ann Neurol*, 39(3), pp. 343-51.

Rahman, S., Poulton, J., Marchington, D. and Suomalainen, A. (2001) 'Decrease of 3243 A-->G mtDNA mutation from blood in MELAS syndrome: a longitudinal study', *Am J Hum Genet*, 68(1), pp. 238-40.

Rajasimha, H.K., Chinnery, P.F. and Samuels, D.C. (2008) 'Selection against pathogenic mtDNA mutations in a stem cell population leads to the loss of the 3243A-->G mutation in blood', *Am J Hum Genet*, 82(2), pp. 333-43.

Rebelo, A.P., Dillon, L.M. and Moraes, C.T. (2011) 'Mitochondrial DNA transcription regulation and nucleoid organization', *J Inherit Metab Dis*, 34(4), pp. 941-51.

Reeve, A.K., Krishnan, K.J., Taylor, G., Elson, J.L., Bender, A., Taylor, R.W., Morris, C.M. and Turnbull, D.M. (2009) 'The low abundance of clonally expanded mitochondrial DNA point mutations in aged substantia nigra neurons', *Aging Cell*, 8(4), pp. 496-8.

Reinhardt, K., Dowling, D.K. and Morrow, E.H. (2013) 'Mitochondrial Replacement, Evolution, and the Clinic', *Science*, 341(6152), pp. 1345-1346.

Remes, A.M., Majamaa-Voltti, K., Kärppä, M., Moilanen, J.S., Uimonen, S., Helander, H., Rusanen, H., Salmela, P.I., Sorri, M., Hassinen, I.E. and Majamaa, K. (2005) 'Prevalence of large-scale mitochondrial DNA deletions in an adult Finnish population', *Neurology*, 64(6), pp. 976-981.

Rhee, S.G., Bae, Y.S., Lee, S.R. and Kwon, J. (2000) 'Hydrogen peroxide: a key messenger that modulates protein phosphorylation through cysteine oxidation', *Sci STKE*, 2000(53), p. pe1.

Richardson, D.R., Lane, D.J.R., Becker, E.M., Huang, M.L.H., Whitnall, M., Rahmanto, Y.S., Sheftel, A.D. and Ponka, P. (2010) 'Mitochondrial iron trafficking and the integration of iron metabolism between the mitochondrion and cytosol', *Proceedings of the National Academy of Sciences*, 107(24), pp. 10775-10782.

Rorbach, J. and Minczuk, M. (2012) 'The post-transcriptional life of mammalian mitochondrial RNA', *Biochem J*, 444(3), pp. 357-73.

Rosenfeld, C.S. and Roberts, R.M. (2004) 'Maternal Diet and Other Factors Affecting Offspring Sex Ratio: A Review', *Biology of Reproduction*, 71(4), pp. 1063-1070.

Rosser, B.W.C., Dean, M.S. and Bandman, E. (2002) 'Myonuclear domain size varies along the lengths of maturing skeletal muscle fibers', *International Journal of Developmental Biology*, 46(5), pp. 747-754.

Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J.P. and Letellier, T. (2003) 'Mitochondrial threshold effects', *Biochem J*, 370(Pt 3), pp. 751-62.

Rotig, A., Bessis, J.L., Romero, N., Cormier, V., Saudubray, J.M., Narcy, P., Lenoir, G., Rustin, P. and Munnich, A. (1992) 'Maternally inherited duplication of the mitochondrial genome in a syndrome of proximal tubulopathy, diabetes mellitus, and cerebellar ataxia', *Am J Hum Genet*, 50(2), pp. 364-70.

Rotig, A., Bourgeron, T., Chretien, D., Rustin, P. and Munnich, A. (1995) 'Spectrum of mitochondrial DNA rearrangements in the Pearson marrow-pancreas syndrome', *Hum Mol Genet*, 4(8), pp. 1327-30.

Rötig, A., Lehnert, A., Chrétien, D., Rustin, P., Gubler, M.C., Munnich, A. and Niaudet, P. (1997) 'Kidney involvement in mitochondrial disorders', *Atteinte renale au cours des cytopathies mitochondriales*, 13(1), pp. 18-27.

Rouault, T.A. and Tong, W.H. (2005) 'Iron-sulphur cluster biogenesis and mitochondrial iron homeostasis', *Nat Rev Mol Cell Biol*, 6(4), pp. 345-51.

Roy, R.R., Monke, S.R., Allen, D.L. and Edgerton, V.R. (1999) 'Modulation of myonuclear number in functionally overloaded and exercised rat plantaris fibers', *Journal of Applied Physiology*, 87(2), pp. 634-642.

Sadikovic, B., Wang, J., El-Hattab, A., Landsverk, M., Douglas, G., Brundage, E.K., Craigen, W.J., Schmitt, E.S. and Wong, L.-J.C. (2010) 'Sequence Homology at the Breakpoint and Clinical Phenotype of Mitochondrial DNA Deletion Syndromes', *PLoS One*, 5(12), p. e15687.

Sakia, R.M. (1992) 'The Box-Cox Transformation Technique: A Review', *Journal of the Royal Statistical Society. Series D The Statistician*, 41(2), pp. pp. 169-178.

Salsano, E., Giovagnoli, A.R., Morandi, L., Maccagnano, C., Lamantea, E., Marchesi, C., Zeviani, M. and Pareyson, D. (2011) 'Mitochondrial dementia: a sporadic case of progressive cognitive and behavioral decline with hearing loss due to the rare m.3291T>C MELAS mutation', *J Neurol Sci*, 300(1-2), pp. 165-8.

Samuels, D.C., Schon, E.A. and Chinnery, P.F. (2004) 'Two direct repeats cause most human mtDNA deletions', *Trends in Genetics*, 20(9), pp. 393-398.

Sarzi, E., Goffart, S., Serre, V., Chretien, D., Slama, A., Munnich, A., Spelbrink, J.N. and Rotig, A. (2007) 'Twinkle helicase (PEO1) gene mutation causes mitochondrial DNA depletion', *Ann Neurol*, 62(6), pp. 579-87.

Schaefer, A.M., McFarland, R., Blakely, E.L., He, L., Whittaker, R.G., Taylor, R.W., Chinnery, P.F. and Turnbull, D.M. (2008) 'Prevalence of mitochondrial DNA disease in adults', *Annals of Neurology*, 63(1), pp. 35-39.

Schaefer, A.M., Phoenix, C., Elson, J.L., McFarland, R., Chinnery, P.F. and Turnbull, D.M. (2006) 'Mitochondrial disease in adults: a scale to monitor progression and treatment', *Neurology*, 66(12), pp. 1932-4.

Schagger, H. and Pfeiffer, K. (2000) 'Supercomplexes in the respiratory chains of yeast and mammalian mitochondria', *EMBO J*, 19(8), pp. 1777-83.

Schmickel, R.D. (1973) 'Quantitation of Human Ribosomal DNA: Hybridization of Human DNA with Ribosomal RNA for Quantitation and Fractionation', *Pediatr Res*, 7(1), pp. 5-12.

Schon, E.A., Dimauro, S. and Hirano, M. (2012) 'Human mitochondrial DNA: Roles of inherited and somatic mutations', *Nature Reviews Genetics*, 13(12), pp. 878-890.

Schon, E.A., Rizzuto, R., Moraes, C.T., Nakase, H., Zeviani, M. and DiMauro, S. (1989) 'A direct repeat is a hotspot for large-scale deletion of human mitochondrial DNA', *Science*, 244(4902), pp. 346-9.

Schon, E.A., Santra, S., Pallotti, F. and Girvin, M.E. (2001) 'Pathogenesis of primary defects in mitochondrial ATP synthesis', *Semin Cell Dev Biol*, 12(6), pp. 441-8.

Schultz, B.E. and Chan, S.I. (2001) 'Structures and proton-pumping strategies of mitochondrial respiratory enzymes', *Annu Rev Biophys Biomol Struct*, 30, pp. 23-65.

Schwartz, M. and Vissing, J. (2002) 'Paternal Inheritance of Mitochondrial DNA', *New England Journal of Medicine*, 347(8), pp. 576-580.

Schwarz, G. (1978) 'Estimating the dimension of a model', *Annals of Statistics*, 6(2), pp. 461-464.

Seidowsky, A., Hoffmann, M., Glowacki, F., Dhaenens, C.M., Devaux, J.P., de Sainte Foy, C.L., Provot, F., Gheerbrant, J.D., Hummel, A., Hazzan, M., Dracon, M., Dieux-Coeslier, A., Copin, M.C., Noel, C. and Buob, D. (2013) 'Renal involvement in MELAS syndrome - a series of 5 cases and review of the literature', *Clin Nephrol*, 80(12), pp. 456-463.

Sena, L.A. and Chandel, N.S. (2012) 'Physiological roles of mitochondrial reactive oxygen species', *Mol Cell*, 48(2), pp. 158-67.

Shadel, G.S. (2008) 'Expression and maintenance of mitochondrial DNA: new insights into human disease pathology', *Am J Pathol*, 172(6), pp. 1445-56.

Shanske, S., Pancrudo, J., Kaufmann, P., Engelstad, K., Jhung, S., Lu, J., Naini, A., DiMauro, S. and De Vivo, D.C. (2004) 'Varying loads of the mitochondrial DNA A3243G mutation in different tissues: implications for diagnosis', *Am J Med Genet A*, 130A(2), pp. 134-7.

Shanske, S., Tang, Y., Hirano, M., Nishigaki, Y., Tanji, K., Bonilla, E., Sue, C., Krishna, S., Carlo, J.R., Willner, J., Schon, E.A. and DiMauro, S. (2002) 'Identical mitochondrial DNA deletion in a woman with ocular myopathy and in her son with pearson syndrome', *Am J Hum Genet*, 71(3), pp. 679-83.

Shanske, S. and Wong, L.J. (2004) 'Molecular analysis for mitochondrial DNA disorders', *Mitochondrion*, 4(5-6), pp. 403-15.

Shao, L., Li, H., Pazhanisamy, S.K., Meng, A., Wang, Y. and Zhou, D. (2011) 'Reactive oxygen species and hematopoietic stem cell senescence', *Int J Hematol*, 94(1), pp. 24-32.

Shiraiwa, N., Ishii, A., Iwamoto, H., Mizusawa, H., Kagawa, Y. and Ohta, S. (1993) 'Content of mutant mitochondrial DNA and organ dysfunction in a patient with a MELAS subgroup of mitochondrial encephalomyopathies', *J Neurol Sci*, 120(2), pp. 174-9.

Shoffner, J.M., Lott, M.T., Lezza, A.M., Seibel, P., Ballinger, S.W. and Wallace, D.C. (1990) 'Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation', *Cell*, 61(6), pp. 931-7.

Shoffner, J.M., Lott, M.T., Voljavec, A.S., Soueidan, S.A., Costigan, D.A. and Wallace, D.C. (1989) 'Spontaneous Kearns-Sayre/chronic external ophthalmoplegia plus syndrome associated with a mitochondrial DNA deletion: A slip-replication model and metabolic therapy', *Proceedings of the National Academy of Sciences of the United States of America*, 86(20), pp. 7952-7956.

Shokolenko, I., Venediktova, N., Bochkareva, A., Wilson, G.L. and Alexeyev, M.F. (2009) 'Oxidative stress induces degradation of mitochondrial DNA', *Nucleic Acids Research*, 37(8), pp. 2539-2548.

Shoubridge, E.A., Karpati, G. and Hastings, K.E. (1990) 'Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease', *Cell*, 62(1), pp. 43-9.

Shtatland, E.S., Kleinman, K. and Cain, E.M. (2002) *One More Time About R^2 Measures of Fit in Logistic Regression*. Available at: http://nesug.org/proceedings/nesug02/st/st004.pdf (Accessed: August 12, 2012).

Shyh-Chang, N., Daley, G.Q. and Cantley, L.C. (2013) 'Stem cell metabolism in tissue development and aging', *Development*, 140(12), pp. 2535-47.

Silvey, S.D. (1969) 'Multicollinearity and Imprecise Estimation', *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(3), pp. 539-552.

Simerville, J.A., Maxted, W.C. and Pahira, J.J. (2005) 'Urinalysis: a comprehensive review', *Am Fam Physician*, 71(6), pp. 1153-62.

Simonsz, H.J., Barlocher, K. and Rotig, A. (1992) 'Kearns-Sayre's syndrome developing in a boy who survived pearson's syndrome caused by mitochondrial DNA deletion', *Doc Ophthalmol*, 82(1-2), pp. 73-9.

Simsek, T., Kocabas, F., Zheng, J., Deberardinis, R.J., Mahmoud, A.I., Olson, E.N., Schneider, J.W., Zhang, C.C. and Sadek, H.A. (2010) 'The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche', *Cell Stem Cell*, 7(3), pp. 380-90.

Singer, J.D. (1998) 'Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models', *Journal of Educational and Behavioral Statistics*, 23(4), pp. 323-355.

Sjostrom, M., Angquist, K.A., Bylund, A.C., Friden, J., Gustavsson, L. and Schersten, T. (1982) 'Morphometric analyses of human muscle fiber types', *Muscle Nerve*, 5(7), pp. 538-53.

Smits, P., Smeitink, J. and van den Heuvel, L. (2010) 'Mitochondrial translation and beyond: processes implicated in combined oxidative phosphorylation deficiencies', *J Biomed Biotechnol*, 2010, p. 737385.

Sochivko, D.G., Fedorov, A.A., Varlamov, D.A., Kurochkin, V.E. and Petrov, R.V. (2013) 'Accuracy of quantitative real-time PCR analysis', *Doklady Biochemistry and Biophysics*, 449(1), pp. 105-108.

Song, S. (2012) 'Does famine influence sex ratio at birth? Evidence from the 1959–1961 Great Leap Forward Famine in China', *Proceedings of the Royal Society B: Biological Sciences*, 279(1739), pp. 2883-2890.

Spendiff, S., Reza, M., Murphy, J.L., Gorman, G., Blakely, E.L., Taylor, R.W., Horvath, R., Campbell, G., Newman, J., Lochmuller, H. and Turnbull, D.M. (2013) 'Mitochondrial DNA deletions in muscle satellite cells: implications for therapies', *Hum Mol Genet*, 22(23), pp. 4739-47.

Staron, R.S., Hagerman, F.C., Hikida, R.S., Murray, T.F., Hostler, D.P., Crill, M.T., Ragg, K.E. and Toma, K. (2000) 'Fiber type composition of the vastus lateralis muscle of young men and women', *J Histochem Cytochem*, 48(5), pp. 623-9.

Stewart, J.B., Freyer, C., Elson, J.L. and Larsson, N.G. (2008a) 'Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease', *Nat Rev Genet*, 9(9), pp. 657-62.

Stewart, J.B., Freyer, C., Elson, J.L., Wredenberg, A., Cansu, Z., Trifunovic, A. and Larsson, N.G. (2008b) 'Strong purifying selection in transmission of mitochondrial DNA', *PLoS Biol.*, 6.

Strasak, A.M., Zaman, Q., Pfeiffer, K.P., Gobel, G. and Ulmer, H. (2007) 'Statistical errors in medical research--a review of common pitfalls', *Swiss Med Wkly*, 137(3-4), pp. 44-9.

Stults, D.M., Killen, M.W., Pierce, H.H. and Pierce, A.J. (2008) 'Genomic architecture and inheritance of human ribosomal RNA gene clusters', *Genome Res*, 18(1), pp. 13-8.

Suda, T., Takubo, K. and Semenza, G.L. (2011) 'Metabolic regulation of hematopoietic stem cells in the hypoxic niche', *Cell Stem Cell*, 9(4), pp. 298-310.

Sue, C.M., Crimmins, D.S., Soo, Y.S., Pamphlett, R., Presgrave, C.M., Kotsimbos, N., Jean-Francois, M.J., Byrne, E. and Morris, J.G. (1998a) 'Neuroradiological features of

six kindreds with MELAS tRNA(Leu) A2343G point mutation: implications for pathogenesis', *J Neurol Neurosurg Psychiatry*, 65(2), pp. 233-40.

Sue, C.M., Quigley, A., Katsabanis, S., Kapsa, R., Crimmins, D.S., Byrne, E. and Morris, J.G. (1998b) 'Detection of MELAS A3243G point mutation in muscle, blood and hair follicles', *J Neurol Sci*, 161(1), pp. 36-9.

Sugioka, K., Nakano, M., Totsune-Nakano, H., Minakami, H., Tero-Kubota, S. and Ikegami, Y. (1988) 'Mechanism of O2- generation in reduction and oxidation cycle of ubiquinones in a model of mitochondrial electron transport systems', *Biochim Biophys Acta*, 936(3), pp. 377-85.

Sutovsky, P., Moreno, R.D., Ramalho-Santos, J., Dominko, T., Simerly, C. and Schatten, G. (2000) 'Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos', *Biol Reprod*, 63(2), pp. 582-90.

Suzuki, S., Oka, Y., Kadowaki, T., Kanatsuka, A., Kuzuya, T., Kobayashi, M., Sanke, T., Seino, Y. and Nanjo, K. (2003) 'Clinical features of diabetes mellitus with the mitochondrial DNA 3243 (A-G) mutation in Japanese: maternal inheritance and mitochondria-related complications', *Diabetes Res Clin Pract*, 59(3), pp. 207-17.

t Hart, L.M., Jansen, J.J., Lemkes, H.H., de Knijff, P. and Maassen, J.A. (1996) 'Heteroplasmy levels of a mitochondrial gene mutation associated with diabetes mellitus decrease in leucocyte DNA upon aging', *Hum Mutat*, 7(3), pp. 193-7.

Tachibana, M., Sparman, M., Sritanaudomchai, H., Ma, H., Clepper, L., Woodward, J., Li, Y., Ramsey, C., Kolotushkina, O. and Mitalipov, S. (2009) 'Mitochondrial gene replacement in primate offspring and embryonic stem cells', *Nature*, 461(7262), pp. 367-372.

Taivassalo, T., De Stefano, N., Argov, Z., Matthews, P.M., Chen, J., Genge, A., Karpati, G. and Arnold, D.L. (1998) 'Effects of aerobic training in patients with mitochondrial myopathies', *Neurology*, 50(4), pp. 1055-60.

Taivassalo, T., Gardner, J.L., Taylor, R.W., Schaefer, A.M., Newman, J., Barron, M.J., Haller, R.G. and Turnbull, D.M. (2006) 'Endurance training and detraining in

mitochondrial myopathies due to single large-scale mtDNA deletions', *Brain*, 129(Pt 12), pp. 3391-401.

Tang, Y., Manfredi, G., Hirano, M. and Schon, E.A. (2000) 'Maintenance of human rearranged mitochondrial DNAs in long-term cultured transmitochondrial cell lines', *Mol Biol Cell*, 11(7), pp. 2349-58.

Tay, S.K.H., DiMauro, S., Pang, A.Y.W., Lai, P.-S. and Yap, H.-K. (2008) 'Myotoxicity of Lipid-Lowering Agents in a Teenager With MELAS Mutation', *Pediatr Neurol*, 39(6), pp. 426-428.

Taylor, R.W., Barron, M.J., Borthwick, G.M., Gospel, A., Chinnery, P.F., Samuels, D.C., Taylor, G.A., Plusa, S.M., Needham, S.J., Greaves, L.C., Kirkwood, T.B. and Turnbull, D.M. (2003) 'Mitochondrial DNA mutations in human colonic crypt stem cells', *J Clin Invest*, 112(9), pp. 1351-60.

Taylor, R.W. and Turnbull, D.M. (2005) 'Mitochondrial DNA mutations in human disease', *Nat Rev Genet*, 6(5), pp. 389-402.

Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R.N. and Chrzanowska-Lightowlers, Z.M. (2010a) 'Hungry codons promote frameshifting in human mitochondrial ribosomes', *Science*, 327(5963), p. 301.

Temperley, R.J., Wydro, M., Lightowlers, R.N. and Chrzanowska-Lightowlers, Z.M. (2010b) 'Human mitochondrial mRNAs--like members of all families, similar but different', *Biochim Biophys Acta*, 1797(6-7), pp. 1081-5.

Terzioglu, M., Ruzzenente, B., Harmel, J., Mourier, A., Jemt, E., López, Marcela D., Kukat, C., Stewart, James B., Wibom, R., Meharg, C., Habermann, B., Falkenberg, M., Gustafsson, Claes M., Park, Chan B. and Larsson, N.-G. (2013) 'MTERF1 Binds mtDNA to Prevent Transcriptional Interference at the Light-Strand Promoter but Is Dispensable for rRNA Gene Transcription Regulation', *Cell Metab*, 17(4), pp. 618-626.

Thomas, V.P. (1998) 'What's wrong with Bonferroni adjustments', *BMJ*, 316(7139), pp. 1236-1238.

Thyagarajan, D., Shanske, S., Vazquez-Memije, M., De Vivo, D. and DiMauro, S. (1995) 'A novel mitochondrial ATPase 6 point mutation in familial bilateral striatal necrosis', *Ann Neurol*, 38(3), pp. 468-72.

Trivers, R.L. and Willard, D.E. (1973) 'Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring', *Science*, 179(4068), pp. 90-92.

Tuppen, H.A.L., Blakely, E.L., Turnbull, D.M. and Taylor, R.W. (2010) 'Mitochondrial DNA mutations and human disease', *Biochimica et Biophysica Acta - Bioenergetics*, 1797(2), pp. 113-128.

Turrens, J.F. and Boveris, A. (1980) 'Generation of superoxide anion by the NADH dehydrogenase of bovine heart mitochondria', *Biochem J*, 191(2), pp. 421-7.

Twig, G., Elorza, A., Molina, A.J., Mohamed, H., Wikstrom, J.D., Walzer, G., Stiles, L., Haigh, S.E., Katz, S., Las, G., Alroy, J., Wu, M., Py, B.F., Yuan, J., Deeney, J.T., Corkey, B.E. and Shirihai, O.S. (2008) 'Fission and selective fusion govern mitochondrial segregation and elimination by autophagy', *EMBO J*, 27(2), pp. 433-46.

Tzoulis, C. and Bindoff, L.A. (2009) 'Serial Diffusion Imaging in a Case of Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-Like Episodes', *Stroke*, 40(2), pp. e15-e17.

Ugalde, C., Vogel, R., Huijbens, R., van den Heuvel, B., Smeitink, J. and Nijtmans, L. (2004) 'Human mitochondrial complex I assembles through the combination of evolutionary conserved modules: a framework to interpret complex I deficiencies', *Human Molecular Genetics*, 13(20), pp. 2461-2472.

Uusimaa, J., Moilanen, J.S., Vainionpaa, L., Tapanainen, P., Lindholm, P., Nuutinen, M., Lopponen, T., Maki-Torkko, E., Rantala, H. and Majamaa, K. (2007) 'Prevalence, segregation, and phenotype of the mitochondrial DNA 3243A>G mutation in children', *Ann Neurol*, 62(3), pp. 278-87.

van den Ouweland, J.M., Lemkes, H.H., Trembath, R.C., Ross, R., Velho, G., Cohen, D., Froguel, P. and Maassen, J.A. (1994) 'Maternally inherited diabetes and deafness is a distinct subtype of diabetes and associates with a single point mutation in the mitochondrial tRNA(Leu(UUR)) gene', *Diabetes*, 43(6), pp. 746-51.

Van Der Meer, S.F.T., Jaspers, R.T., Jones, D.A. and Degens, H. (2011) 'Time-course of changes in the myonuclear domain during denervation in young-adult and old rat gastrocnemius muscle', *Muscle and Nerve*, 43(2), pp. 212-222.

Vartak, R., Porras, C.A. and Bai, Y. (2013) 'Respiratory supercomplexes: structure, function and assembly', *Protein Cell*, 4(8), pp. 582-90.

Venegas, V. and Halberg, M.C. (2012) 'Measurement of mitochondrial DNA copy number', *Methods Mol Biol*, 837, pp. 327-35.

Vogel, F., Bornhovd, C., Neupert, W. and Reichert, A.S. (2006) 'Dynamic subcompartmentalization of the mitochondrial inner membrane', *J Cell Biol*, 175(2), pp. 237-47.

Vydt, T.C., de Coo, R.F., Soliman, O.I., Ten Cate, F.J., van Geuns, R.J., Vletter, W.B., Schoonderwoerd, K., van den Bosch, B.J., Smeets, H.J. and Geleijnse, M.L. (2007) 'Cardiac involvement in adults with m.3243A>G MELAS gene mutation', *Am J Cardiol*, 99(2), pp. 264-9.

Wada, K.I., Katsuta, S. and Soya, H. (2003) 'Natural occurrence of myofiber cytoplasmic enlargement accompanied by decrease in myonuclear number', *Japanese Journal of Physiology*, 53(2), pp. 145-150.

Wallace, D.C. (1992) 'Diseases of the mitochondrial DNA', *Annu Rev Biochem*, 61, pp. 1175-212.

Wallace, D.C. (2010) 'Mitochondrial DNA mutations in disease and aging', *Environ Mol Mutagen*, 51(5), pp. 440-50.

Wallace, D.C., Singh, G., Lott, M.T., Hodge, J.A., Schurr, T.G., Lezza, A.M., Elsas, L.J., 2nd and Nikoskelainen, E.K. (1988) 'Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy', *Science*, 242(4884), pp. 1427-30.

Walter, J.H. (2009) 'Genes, patients, families, doctors—Mutation analysis in clinical practice', *Journal of Inherited Metabolic Disease*, 32(3), pp. 441-446.

Wang, C. and Youle, R.J. (2009) 'The role of mitochondria in apoptosis*', *Annu Rev Genet*, 43, pp. 95-118.

Wang, G.J., Nutter, L.M. and Thayer, S.A. (1997) 'Insensitivity of cultured rat cortical neurons to mitochondrial DNA synthesis inhibitors: evidence for a slow turnover of mitochondrial DNA', *Biochem Pharmacol*, 54(1), pp. 181-7.

Wang, X.Y., Noguchi, K., Takashima, S., Hayashi, N., Ogawa, S. and Seto, H. (2003) 'Serial diffusion-weighted imaging in a patient with MELAS and presumed cytotoxic oedema', *Neuroradiology*, 45(9), pp. 640-3.

Ware, J., Jr., Kosinski, M. and Keller, S.D. (1996) 'A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity', *Med Care*, 34(3), pp. 220-33.

Westermann, B. (2010) 'Mitochondrial fusion and fission in cell life and death', *Nat Rev Mol Cell Biol*, 11(12), pp. 872-884.

White, H. (1980) 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity', *Econometrica*, 48(4), pp. 817-38.

White, H.E., Durston, V.J., Seller, A., Fratter, C., Harvey, J.F. and Cross, N.C. (2005) 'Accurate detection and quantitation of heteroplasmic mitochondrial point mutations by pyrosequencing', *Genet Test*, 9(3), pp. 190-9.

Whittaker, R.G., Blackwood, J.K., Alston, C.L., Blakely, E.L., Elson, J.L., McFarland, R., Chinnery, P.F., Turnbull, D.M. and Taylor, R.W. (2009) 'Urine heteroplasmy is the best predictor of clinical outcome in the m.3243A>G mtDNA mutation', *Neurology*, 72(6), pp. 568-9.

Whittaker, R.G., Schaefer, A.M., McFarland, R., Taylor, R.W., Walker, M. and Turnbull, D.M. (2007) 'Prevalence and progression of diabetes in mitochondrial disease', *Diabetologia*, 50(10), pp. 2085-9.

Wiesner, R.J., Ruegg, J.C. and Morano, I. (1992) 'Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues', *Biochem Biophys Res Commun*, 183(2), pp. 553-9.

Winge, D.R. (2012) 'Sealing the mitochondrial respirasome', *Mol Cell Biol*, 32(14), pp. 2647-52.

Wißmann, M., Toutenburg, H. and Shalabh 8 (2007) 'Role of Categorical Variables in Multicollinearity in the Linear Regression Model'. Available at: http://epub.ub.uni-muenchen.de/2081/.

Wolfinger, R.D. (1997) 'An example of using mixed models and PROC MIXED for longitudinal data', *J Biopharm Stat*, 7(4), pp. 481-500.

Woodson, J.D. and Chory, J. (2008) 'Coordination of gene expression between organellar and nuclear genomes', *Nat Rev Genet*, 9(5), pp. 383-95.

Yamashita, S., Nishino, I., Nonaka, I. and Goto, Y.I. (2008) 'Genotype and phenotype analyses in 136 patients with single large-scale mitochondrial DNA deletions', *Journal of Human Genetics*, 53(7), pp. 598-606.

Yasukawa, T., Reyes, A., Cluett, T.J., Yang, M.Y., Bowmaker, M., Jacobs, H.T. and Holt, I.J. (2006) 'Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand', *EMBO J*, 25(22), pp. 5358-71.

Yasukawa, T., Suzuki, T., Ishii, N., Ohta, S. and Watanabe, K. (2001) 'Wobble modification defect in tRNA disturbs codon-anticodon interaction in a mitochondrial disease', *EMBO J*, 20(17), pp. 4794-802.

Yasukawa, T., Suzuki, T., Ishii, N., Ueda, T., Ohta, S. and Watanabe, K. (2000a) 'Defect in modification at the anticodon wobble nucleotide of mitochondrial tRNA(Lys) with the MERRF encephalomyopathy pathogenic mutation', *FEBS Lett*, 467(2-3), pp. 175-8.

Yasukawa, T., Suzuki, T., Ueda, T., Ohta, S. and Watanabe, K. (2000b) 'Modification defect at anticodon wobble nucleotide of mitochondrial tRNAs(Leu)(UUR) with pathogenic mutations of mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes', *J Biol Chem*, 275(6), pp. 4251-7.

Yatsuga, S., Povalko, N., Nishioka, J., Katayama, K., Kakimoto, N., Matsuishi, T., Kakuma, T. and Koga, Y. (2012) 'MELAS: a nationwide prospective cohort study of 96 patients in Japan', *Biochim Biophys Acta*, 1820(5), pp. 619-24.

Yoneda, M., Chomyn, A., Martinuzzi, A., Hurko, O. and Attardi, G. (1992) 'Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy', *Proc Natl Acad Sci U S A*, 89(23), pp. 11164-8.

Yoneda, M., Maeda, M., Kimura, H., Fujii, A., Katayama, K. and Kuriyama, M. (1999) 'Vasogenic edema on MELAS: A serial study with diffusion-weighted MR imaging', *Neurology*, 53(9), pp. 2182-2184.

Yoneda, M., Miyatake, T. and Attardi, G. (1994) 'Complementation of mutant and wild-type human mitochondrial DNAs coexisting since the mutation event and lack of complementation of DNAs introduced separately into a cell within distinct organelles', *Mol Cell Biol*, 14(4), pp. 2699-712.

Yoneda, M., Tanno, Y., Horai, S., Ozawa, T., Miyatake, T. and Tsuji, S. (1990) 'A common mitochondrial DNA mutation in the t-RNA(Lys) of patients with myoclonus epilepsy associated with ragged-red fibers', *Biochem Int*, 21(5), pp. 789-96.

Yonemura, K., Hasegawa, Y., Kimura, K., Minematsu, K. and Yamaguchi, T. (2001) 'Diffusion-weighted MR imaging in a case of mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes', *American Journal of Neuroradiology*, 22(2), pp. 269-272.

Yu-Wai-Man, P., Griffiths, P.G., Hudson, G. and Chinnery, P.F. (2009) 'Inherited mitochondrial optic neuropathies', *J Med Genet*, 46(3), pp. 145-58.

Yu-Wai-Man, P., Lai-Cheong, J., Borthwick, G.M., He, L., Taylor, G.A., Greaves, L.C., Taylor, R.W., Griffiths, P.G. and Turnbull, D.M. (2010a) 'Somatic mitochondrial DNA deletions accumulate to high levels in aging human extraocular muscles', *Invest Ophthalmol Vis Sci*, 51(7), pp. 3347-53.

Yu-Wai-Man, P., Sitarz, K.S., Samuels, D.C., Griffiths, P.G., Reeve, A.K., Bindoff, L.A., Horvath, R. and Chinnery, P.F. (2010b) 'OPA1 mutations cause cytochrome c oxidase deficiency due to loss of wild-type mtDNA molecules', *Hum Mol Genet*, 19(15), pp. 3043-52.

Zeviani, M., Moraes, C.T., DiMauro, S., Nakase, H., Bonilla, E., Schon, E.A. and Rowland, L.P. (1988) 'Deletions of mitochondrial DNA in Kearns-Sayre syndrome', *Neurology*, 38(9), pp. 1339-46.

Zhang, S., Tong, A.L., Zhang, Y., Nie, M., Li, Y.X. and Wang, H. (2009) 'Heteroplasmy level of the mitochondrial tRNaLeu(UUR) A3243G mutation in a Chinese family is positively associated with earlier age-of-onset and increasing severity of diabetes', *Chin Med Sci J*, 24(1), pp. 20-5.

Zhu, X., Peng, X., Guan, M.-X. and Yan, Q. (2009) 'Pathogenic mutations of nuclear genes associated with mitochondrial disorders', *Acta Biochim Biophys Sin (Shanghai)*, 41(3), pp. 179-187.

Zollo, O., Tiranti, V. and Sondheimer, N. (2012) 'Transcriptional requirements of the distal heavy-strand promoter of mtDNA', *Proc Natl Acad Sci U S A*, 109(17), pp. 6508-12.

Zwirner, P. and Wilichowski, E. (2001) 'Progressive sensorineural hearing loss in children with mitochondrial encephalomyopathies', *Laryngoscope*, 111(3), pp. 515-21.