The Genetics and Epidemiology of Myopia in the ALSPAC Cohort

George Mc Mahon

PhD 2010

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date
...08/03/2011...

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD

Signed ..... (candidate) Date
...08/03/2011...

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed ..... (candidate) Date
...08/03/2011...

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date
...08/03/2011...

To Shan Shan,

Contents

Chapter 5

Chapter 6

List of Figures and Tables

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Acknowledgements

Thanks to Dr. Cathy Williams and Dr. Jez Guggenheim, my supervisors throughout this thesis. I hope to have the opportunity to investigate myopia with you again in the future and to learn interesting things about the development of myopia. Thanks to my advisors Professors James Morgan and Maggie Woodhouse. Thanks also to Dr. Beate Glaser and John Kemp.

Many thanks to the School of Optometry and Vision Sciences, Cardiff University and the Department of Social Medicine, Bristol University, for providing a research friendly environment. Many thanks to the staff also, for their support throughout this thesis.

Thanks to other postgraduates in the School of Optometry and Vision Sciences and the Department of Social Medicine for helpful advice on research.

Finally, thanks to Cardiff University and Bristol University for providing a challenging and fun environment to carry out this PhD thesis, and thanks *very much* to the National Eye Research Centre (NERC) for providing funding.

Thesis Summary

The aim of this thesis is to investigate the genetics and epidemiology of myopia in the Avon Longitudinal Study of Parents and Children cohort (ALSPAC). To balance compliance and accuracy, the ALSPAC cohort uses non-cycloplegic autorefraction to measure refractive error. Non-cycloplegic autorefraction of young people is documented to display a negative offset which can be partly mediated by calibration with a more accurate technique. Subjective refractions of ALSPAC participants were collected from their optometrists and calibration of non-cycloplegic autorefraction measures was undertaken. It was observed that non-cycloplegic autorefraction has a high sensitivity (89%) and specificity (96%) to infer myopia.

Identification of a genetic factor underlying disease progression requires that the disease displays a genetic component, which can be identified by a heritability study. ALSPAC is a birth cohort and the majority of data collected is on the refractive error of unrelated people of a similar age. To conduct a heritability study, measures of subjective refraction of relatives were collected. A heritability study was then undertaken. Refractive error was observed to display a heritability of 50%.

An aim of this thesis is to map a genetic factor related to myopia progression. A genome-wide association study of myopia, refractive error and two ocular determinants of refractive error, axial length and corneal curvature was undertaken. A number of genetic locations were identified and extra genotyping and replication in an independent cohort is underway. A further aim of this thesis is epidemiological analyses of two myopia risk factors, birth order and season of birth. It was observed that an increased number of myopes were found in those with a higher birth order (odds ratio of 1.5, $P = 0.016$) and in those born in the summer and autumn (odds ratios and P-values of 1.17, $P = 0.006$, 1.16, $P = 0.007$ respectively).

Chapter 1

General Introduction

Myopia is a condition in which patients cannot see objects in the distance clearly. This thesis aims to understand the basis of myopia in terms of genetics and epidemiology. The results shown in Chapters 2-5 use data from the Avon Longitudinal Study of Parents and Children (ALSPAC) when children were aged fifteen. Chapter 6 uses data from ALSPAC when children were aged 11 and data from the International High Myopia Genetics Consortium. Chapter 7 uses data from an independent cohort, drawn from 19 optometric practises in the UK.

ALSPAC is a birth cohort that seeks to better understand how genetics and the environment play a role in health and development (Golding, Pembrey and Jones, 2001). It started in 1991, in Avon in South West England, where 85% of children born (approximately 14,000) at that time (from 1991 to 1992) and their mothers were recruited (Golding, 2004). It has collected a diverse set of measures on a subject's phenotype (Golding, 2004) and created a DNA bank (Jones et al., 2000) to allow researchers the opportunity to better understand health and development, research which is ongoing at present. The study is longitudinal in nature, following subjects from birth into adulthood. Efforts to better understand myopia are ongoing in ALSPAC and as such, this thesis forms part of a continuum of those investigations. The thesis seeks in particular to push forward investigations on myopia in the cohort to allow genetic and epidemiological components to be identified. In doing so five main aims were identified prior to the beginning of this PhD. Chapters 2 to 6 follow each of the below aims from numbers 1 to 5.

1) Additional measures of refraction were necessary to supplement measures taken at an ALSPAC clinic. These measures would be used for two purposes: a) to calibrate non-cycloplegic autorefraction measures of refractive error of young people in ALSPAC and b) to allow the estimation of heritability of refractive error. A data collection exercise would be necessary to gather highly accurate measures of refractive error.

2) Validation of existing measures of refraction taken by ALSPAC would be necessary. To balance cost, accuracy and compliance, non-cycloplegic autorefraction was used as the primary measure of refractive error in ALSPAC. There is evidence that such methods can lead to a small systematic error in the measurement of refractive error in young people. Quantification and correction of the error would be necessary for better use of the measures. In studies of the epidemiology and genetics of a disease the accuracy of a definition of a phenotype is important to the study (Aylsworth, 1998; Woodward, 2005).

3) The heritability of refractive error in the cohort needed to be established. As a birth cohort, ALSPAC's main participants are subjects born in Avon between 1991 and 1992. Refractive measurements of their parents and siblings were needed to carry out a heritability study. The heritability study would suggest whether refractive error was a good candidate for gene mapping studies in the cohort.

4) The genetic basis of myopia would need to be investigated directly. The most systematic method to achieve this in the ALSPAC cohort is by a genome-wide association study. The development of myopia depends in part on a number of ocular components (Rosenfield, 2006) (including axial length and corneal curvature). Measures of these components of the eye had been taken in the ALSPAC cohort. It was possible to examine the genetic basis of these traits (critical to myopia) in the cohort.

5) Birth order, a risk factor that contributes to myopia development would be investigated in the cohort. Environmental factors have been shown to play a role in myopia development in the cohort (Williams et al., 2008a). A relationship between myopia and this novel risk factor would be investigated.

Chapter 6 uses data from the International High Myopia Genetics Consortium (IHMGC) in conjunction with data from the ALSPAC cohort. The IHMGC is a collaborative of high myopia studies (Young, 2009) in which families are recruited if

they contain at least one high myope, over 250 of which were collected. In Chapter 6 a relationship between birth order and myopia is investigated in the ALSPAC cohort, with particular attention to possible confounding. Evidence of an association between birth order and myopia in the IHMGC is also examined.

Chapter 7 is different from previous chapters. Data from the ALSPAC cohort is absent, instead, data from a large (90,884), adult patient population is analysed. The data was collected in 2000 to 2001 from UK optometric practises (Farbrother et al., 2004a). Differences in the amount of natural light at the month of birth have been associated with myopia prevalence during adolescence (Mandel et al., 2008). The aim of Chapter 7 is to examine if an association between season of birth or/and variation in natural light (photoperiod) and myopia in adults is present in a UK cohort.

The following pages of this introduction review the literature of the genetics and epidemiology of myopia. Findings in the fields of genetics and epidemiology of myopia are summarized in Appendix A and B. Methods in the study of the genetics of myopia are briefly introduced, such as linkage analysis and genome-wide association. Considerations from epidemiology are introduced, such as myopia prevalences and myopia risk factors. Theories on the prevention of myopia are also discussed. Chapters 2 to 7 are concerned with experimental work during the course of the thesis. Chapter 8 provides a general discussion.

1.1 Genetics

1.1.1 Monogenic disorders

In the field of genetics, disease can be thought monogenic or multifactorial (although there is emerging evidence that some monogenic disease on closer inspection show evidence of multifactorial aetiology). The term monogenic is related to the term Mendelian in reference to a set of principles that accurately described the passage of some loci from one generation to the next (Klug, 2009). These principles were set out, around 1850, by Gregor Mendel, a German monk, academic (having studied and taught physics and natural science) and researcher. His work in the pea plant led to two hypotheses (among others) regarding inheritance that are still important today (King, Stansfield and Mulligan, 2006);

Segregation: The factors or pair of factors that determine a trait *segregate*, i.e. they separate and can combine independently of their original pair

Independent assortment: Members of different pairs of factors assort independently, i.e. such factors can combine independently.

It is interesting to note that the definition of segregation given above indicates that a pair of factors can determine a trait. In the study of monogenic diseases such a pair of factors is referred to as alleles. An allele (or allelomorph) is one possible form of a gene (with one possible function). In diploid organisms, two alleles (and possibly two functional elements) are present. Pea plants are diploid, i.e. they have two homologous chromosomes and therefore two genes at a locus, as are many other plants, animals and humans (many bacteria, viruses and some plants such as bryophytes, and animals are haploid, i.e. have only one copy of each chromosome (King et al., 2006)). The combination of alleles in a diploid organism is given two primary designations; heterozygosity (a pair of dissimilar alleles) or homozygosity (a pair of similar alleles). Their effect on a phenotype depends on the relationship between alleles in a pair. Two other terms are necessary to complete this brief overview. Dominance refers to the expression of an allele in a pair at the loss of expression of the other. This other allele is often termed recessive. It is important to note that not all alleles follow Mendel's laws. Mendel worked with pea plants and his hypothesis was based on observations of traits transmitted from one generation to the next. There are cases when alleles are not transmitted to the next generation due to complications associated with segregation. For example there is a Drosophila melanogaster gene (segregation distorter, SD) that results, after meiosis, in one daughter cell being transmitted preferentially (King et al., 2006). Some alleles do not assort independently. In this case two alleles from more than one gene are inherited together more often than would be expected if they combined independently (King et al., 2006). Linkage is a term used to describe genes that are co inherited together.

High myopia shows evidence of dominant and recessive forms. Both can be deduced from inheritance patterns in a pedigree in a large well defined family. Autosomal (genes located on one of the autosomes) dominant and recessive, and X-linked (the

presence of a gene on the X chromosome) inheritance of high myopia has been identified in a number of families (Drack, 1998; Young et al., 1998a; Young et al., 1998b).

1.1.2 Linkage analysis

Efforts have also been made to find the location of alleles which are responsible for high myopia. There have been many successes in the identification of genes underlying monogenic human diseases (diseases that are caused by one gene) but it is debatable whether myopia fits into this category. To identify a human disease gene *linkage analysis* has been previously widely used. As noted previously, linkage refers to alleles from different locations in the genome which are inherited together more often than would be expected due to independent assortment. Linkage analysis is a technique that allows a lack of independent assortment to be observed and a relationship between a disease and a region of the genome to be inferred.

Linkage analysis in humans depends upon the availability of genetic markers (segments of DNA whose sequence and location are known). These markers typically reside on chromosomal locations that confer relatively little change in a phenotype (this can be termed phenotypic silence (Terwilliger, 1998)). When mapping a disease gene, its location is unknown but of primary interest. However it is possible to infer the presence of a disease causing gene if the disease trait is observed. In linkage analysis a relationship between the unobserved disease gene and the observed genetic markers is examined (Terwilliger, 1998).

Linkage analysis is a separate entity from linkage. The former is a technique that relies upon the ability to observe and assess linkage. Linkage occurs due to a lack of recombination that normally drives independent assortment during meiosis. The amount of recombination (recombination fraction) that occurs between two alleles at different locations in the genome can be measured as the proportion of progeny that display recombination out of the total number of progeny (King et al., 2006). The recombination fraction of two unlinked alleles is 0.5, that is, for example, of four parental chromosomes, two recombinants and two parental chromosomes are present in the progeny. When alleles are in linkage there are more parental chromosomes

present than recombinants in the progeny, which leads to a recombination fraction of less than 0.5.

In linkage analysis, the recombination fraction between the observed genetic marker and trait (unobserved disease gene) can be measured, along with its likelihood. If the genetic marker and trait (unobserved disease gene) are linked, it is expected that the recombination fraction will be less than 0.5. In other words, loci that are linked will not show equal numbers of recombinants and parental chromosomes. A LOD score allows inferences to be made about a recombination fraction. The LOD score is the log to the base 10 of a likelihood ratio of two hypotheses or more formally the LOD score, Z(θ),

$Z\left(\theta\right) = \log_{10}\left[L\left(\theta\right)/L\left(\frac{1}{2}\right)\right]$, where θ is a recombination fraction, L(θ) is the likelihood of a recombination fraction given the observed data and L(1/2) is the likelihood of a 0.5 recombination fraction given the observed data. The maximum LOD score, $Z\left(\hat{\theta}\right)$, is the given by $Z\left(\hat{\theta}\right) = \log_{10}\left[L\left(\hat{\theta}\right)/L\left(\frac{1}{2}\right)\right]$, where $\hat{\theta}$ is the maximum likelihood value of the recombination fraction (Olson et al., 1999). When $L\left(\hat{\theta}\right)$ is much greater than $L\left(\frac{1}{2}\right)$ it is evidence that the maximum likelihood estimate of the recombination fraction is more likely than a recombination fraction of 0.5.

Although it is outside the scope of this thesis, it is worthy to note that determination of recombination between unobserved trait genotypes and observed genetic markers relies on a number of genetic factors. The probability of an individual displaying a particular unobserved genotype is a function of allele frequency. Allele frequency can be derived for unobserved genotypes via estimation of disease prevalence and knowledge of the genetic mode of transmission. For example if a disease prevalence of 0.1 and an autosomal recessive mode of transmission are assumed, the frequency of the recessive genotype is 0.1, while the allele frequency of the minor allele would be 0.1 square rooted.

The proportion of individuals with a particular genotype who display a certain phenotype is termed penetrance (King et al., 2006). If the unobserved genotype is present there is a secondary probability to be considered. The genotype may be present without the trait due to incomplete penetrance (i.e. some individuals do not develop the disease). Therefore when estimating a joint probability of individuals who display the trait and unobserved trait genotype, only a proportion of this probability will reflect the true number of times the unobserved trait genotype and trait are present. These three probabilities influence the estimation of a recombination fraction. They are taken into account by a genetic model (model-based linkage analysis (Olson et al., 1999)). In the case of Mendelian traits, estimation of a genetic model is possible. When a disease shows indeterminate mode of transmission (and therefore difficult to estimate transmission probabilities), variable penetrance due to environmental modulators (and therefore difficult to estimate true penetrance) and many possible causal loci, estimation of a genetic model is difficult. Instead model-free linkage analysis, which relies upon correlation between marker IBD (identical-by –descent) and trait similarity, can be used. If a marker and trait are linked, affected relatives (such as affected siblings) would be expected to share more marker alleles IBD than by chance alone (Olson et al., 1999).

A large number of Mendelian disorders and information on 12,000 genes is listed online at the Online Mendelian Inheritance in Man (OMIM). There are approximately 3,000 phenotypes listed with a known molecular basis (http://www.ncbi.nlm.nih.gov/Omim/mimstats.html). Linkage analysis has helped identify a significant proportion of Mendelian (monogenic) disorders. For example genes underlying cystic fibrosis, X-linked muscular dystrophy and other human disease genes have been mapped using linkage analysis (Botstein and Risch, 2003). These successes can speed up efforts to find better treatments for patients of these diseases.

1.1.2.1 Effect size

The success of linkage analysis partly depends upon a hypothetical gene displaying a certain size of effect. Effect size can be distinguished from penetrance by noting that an allele can confer a small change in a phenotype or a large change in a phenotype while penetrance indicates the proportion of individuals with a genotype who also

display the disease. Power (i.e. the chance of finding a true positive) of linkage analysis increases with increasing effect size. More frequent diseases do not display typical Mendelian inheritance and may result from genes of an effect size that are small and require very large number of observations to be mapped effectively. This is not true in all cases. Diabetes mellitus is a disorder resulting in an excess of glucose in the blood stream. It is relatively common (4-5% prevalence in adults (Jobling et al., 2004b)). Many genes of small effect are hypothesized to be involved in diabetes development, 10 of which have been identified in a genome-wide association study using single nucleotide polymorphism (SNP) data from the Wellcome Trust Case-control Consortium (Wellcome Trust Case Control Consortium, 2007). However linkage analysis using family data has been successful in mapping genes underlying diabetes susceptibility.

1.1.2.2 Heterogeneity

Effective linkage analysis depends on the effects of an allele on a disease phenotype and also the number of genes which have a role in disease development. In general the term heterogeneity refers to *not just one allele alone* being responsible for disease development. Allelic heterogeneity describes a disease for which one gene is causal but many alleles exist (Leal, 1997). Cystic fibrosis (CF) is a Mendelian disease that shows allelic heterogeneity. Over 1,000 mutations are known, most of which are very rare (Jobling et al., 2004b). One allele containing a 3 base pair (bp) deletion that leads to a misfolded protein accounts for approximately two thirds of cystic fibrosis chromosomes (Jobling et al., 2004b). CF has being mapped using linkage analysis. Linkage analysis of families is relatively unaffected by allelic heterogeneity compared to linkage disequilibrium association in a population analysis because within a family allelic heterogeneity will be reduced. A causal allele will have been passed on the same stretch of chromosome across generations within the family. When genotypes of unrelated individuals are examined for association as in linkage disequilibrium association mapping in the presence of allelic heterogeneity, many alleles may be present. The power to detect an association will depend upon how well each of these alleles is tagged (i.e. the degree of linkage disequilibrium between the marker allele and causal allele).

Locus heterogeneity (non-allelic heterogeneity) refers to different individual genes that are responsible for development of a single disease. Locus heterogeneity will lead to a recombination fraction between disease and genetic marker that is inconsistent across families and would make linkage analysis more difficult (Leal, 1997). The genes underlying myopia development have yet to be discovered but there is evidence that myopia displays locus heterogeneity. Close to 50% of patients with certain rare syndromes display myopia. Examples of these are; Stickler syndrome, Marfan syndrome, Down syndrome, spondyloepiphyseal dysplasia congenital, Fabry disease, Ehlers-Danlos syndrome, postaxial polydactyly and progressive myopia, homocystinuria (Drack, 1998), Kniest syndrome, Pierre Robin syndrome, Noonan syndrome, De Lange syndrome and albinism (Wildsoet, 1998). Stickler syndrome is caused by mutations in the type II collagen gene (COL2A1) on chromosome 12. The gene for Marfan's syndrome, the fibrillin gene (FBN1), is on a different chromosomal region (15q21). Fabry disease is X-linked with an aetiology due to defective activity of alpha-galactosidase. No mutations in any of the genes underlying syndromic myopia have been associated with myopia independent of these syndromes but it can be concluded that there is evidence that more than one gene is critical to myopia. In other research, LOD scores of above 3 have been observed when data from families with high myopia have been analysed at a number of genomic locations including chromosome 18 and 12 (Young et al., 1998a; Young et al., 1998b) and when data from families with common myopia have been analysed at chromosome 22 (Stambolian et al., 2004) and chromosome 1 (Wojciechowski et al., 2006). Appendix A contains a list of high and common myopia loci. Success in linkage analysis also depends on factors associated with mutation, a major source of genetic diversity.

1.1.2.3 Penetrance

The penetrance of an allele is important when estimating the recombination fraction in a linkage analysis, as discussed previously. In this section, how differing penetrance can occur is highlighted (due to the result of mutation) and an important point about the penetrance of genotypes underlying myopia is noted, that myopia genotypes show evidence of being less than 100% penetrant. The process whereby a gene undergoes a change in its structure is mutation (King et al., 2006). Mutations are of many different types. A point mutation is in general a small mutation usually resulting in a substitution of one nucleotide base for another. Other mutations are deletions (loss of

a segment of DNA) and insertions (addition of one nucleotide base or more into a DNA segment (Klug, 2009)). Mutations can be classified according to their phenotypic effects. Neutral mutations have no measurable phenotypic effect and make up the vast majority of mutations. Loss-of-function mutations reduce or eliminate the function of an allele. Gain-of-function mutations result in a new function for an allele. The degree to which a gene function is affected by one or more mutations can directly influence a phenotype.

The proportion of individuals with a particular genotype who display a certain phenotype is termed penetrance (King et al., 2006) (when all or most individuals with a particular mutation display a certain phenotype, penetrance is high) and mutations show varying level of penetrance. Penetrance is intrinsically linked to the type of functional change elicited by a genetic mutation. Certain forms of retinoblastoma show reduced penetrance. Retinoblastoma is a cancerous neoplasm of the retina that occurs in children before the age of three years (King et al., 2006). About 90% of children who carry a retinoblastoma chromosome develop the disease while 10% remain unaffected (Millodot and Laby, 2002). A loss of function (LOF) mutation that is inherited in the retinoblastoma (RB1) gene is located on the long arm of chromosome 13 (13q14). An inherited mutation in RB1 shows autosomal dominant inheritance. However, this mutation does not lead to generation of a tumour immediately. A second loss of function mutation is required in the sister allele for cancer to develop. As mentioned there is a 90% chance of a child who has already inherited one copy of a LOF RB1 mutation suffering a spontaneous second mutation.

The penetrance of an allele underlying myopia can be inferred to be less than 100%. Risk of myopia is associated with number of myopic parents. Children with two myopic parents often show an increased risk of developing myopia compared to children with one or no myopic parents (Zadnik et al., 1994; Low et al., 2010). This may indicate that a mutation responsible for myopia development is being inherited from one generation to the next. In the case of children with two myopic parents only a proportion go on to develop myopia and it can be inferred that a gene shows incomplete penetrance.

An allele which has lost complete functionality is likely to display a high penetrance. High penetrance increases power to map a disease gene in linkage analysis. Genes that have high penetrance can be quite rare. Mutations that display Mendelian inheritance and are detrimental to health often are very rare and highly penetrant. Not all traits that show Mendelian inheritance are rare, but these tend to be traits that confer little or no change in general health. Common baldness, chin fissure and mid-digital hair growth (hair on middle segment of fingers) are all Mendelian traits in humans that are relatively common (Hartl, 1983).

A process that results in loss of a highly penetrant deleterious allele is selection. Selection refers to different chances of reproduction of an organism given its genetic information. Negative selection describes a reduction in fitness (or ability to reproduce). Negative selection can act upon deleterious alleles, reducing the chance that the allele is passed on to the next generation via reproduction. A highly penetrant disease causing gene will (in most cases) be under negative selection which will serve to reduce the frequency of the allele and disease.

1.1.2.4 Conclusions: Linkage analysis and myopia

Successes in linkage analysis tend to be for rare diseases. There are exceptions to this trend. Breast cancer is one of the most common cancers in the world with a lifetime risk of about 12% (Klug, 2009). Only 5% to 10% of breast cancer cases are familial indicating either spontaneous mutation or environmental risk factors in its pathology. BRCA1 was the first gene to be identified as responsible for the development of breast cancer. It is located on chromosome 17q21 (King et al., 2006). 85% of subjects with a BRCA1 chromosome (it is autosomal dominant) go on to develop the disease. BRCA1 was mapped using linkage analysis in 1990. Researchers achieved this by examining families with high incidence of early onset breast cancer (Schildkraut, 1998). Therefore identifying genes underlying complex disease is possible via linkage analysis. However researchers focussed on a subset of patients with age specific prevalence of the disease, which in this case helped reduce heterogeneity. Also BRCA1 accounts for a large proportion of familial breast cancer (85%) indicating the gene displays a large effect size, an element that increases the chance of finding a true positive. A drawback of linkage analysis is that for genes of moderate effect size, linkage will be underpowered to detect a true association. As discussed above, genes

of large effect are likely to be rare (given they will be removed by negative selection) and consistent with this, success in linkage analysis has tended to be for rare disorders. High myopia is a relatively rare condition with a prevalence of 1-2%. It typically has an early age of onset being present in early childhood and also shows Mendelian inheritance. As such it is considered a good candidate for mapping by linkage analysis.

Linkage analysis of pedigrees displaying autosomal dominant and recessive inheritance, and X-linked inheritance has been successful in identifying a number of cytogenetic locations for high and common myopia. Other disorders similarly show multiple individual patterns of Mendelian inheritance. Retinitis pigmentosa (RP) shows a distribution of 60% autosomal recessive, 10%-25% autosomal dominant and 5%-18% X-linked (Bird, 1998) inheritance. The cytogenetic locations linked to myopia are listed in Online Mendelian Inheritance in Man (OMIM), a database devoted to cataloguing human genetic variation that leads to disease, with a prefix MYP ranging from MYP1-17. After a linkage signal is obtained the chromosomal region which harbours the causative gene can be narrowed by positional cloning (which involves identification of genetic markers closest to the gene). Young et al. (Young et al., 2004) undertook more detailed mapping of the MYP1 locus (Xq27-28) and hypothesized that genes involved in colour vision contained mutations associated with high myopia. Young et al. (Young et al., 2001) undertook more detailed mapping of the MYP2 locus on 18p11 and a number of plausible genes in the region have been investigated as candidates genes that may harbour a myopia mutation. More dense genetic maps of regions identified by linkage analysis and candidate gene analysis of these regions have been undertaken for the majority of MYP loci (see Appendix A). However positional cloning (narrowing the linkage signal to one causative gene) is still ongoing and a mutation that leads to myopia is yet to be identified. In an effort to increase the chances of finding a gene that underlies myopia development, myopia researchers have employed various strategies.

1.1.3 Strategies in gene mapping

The mapping of BRCA1 highlights an important part of genetic analysis, strategising to improve the chance of a true positive. One such strategy takes advantage of a

reduced genetic variation which can occur naturally via genetic drift and effective population size, which are now briefly introduced. Variation in allele frequencies between generations can be thought to occur through random (stochastic) sampling (Jobling et al., 2004b), formally termed genetic drift. Genetic drift leads to a reduction or increase in allele frequencies and is strongest in small populations (King et al., 2006). 'Small populations' is a loose term, its meaning more closely related to effective population size. A smaller effective population size leads to a stronger impact of genetic drift. Effective population size is related to inbreeding, the probability that two alleles in a population share a common ancestor (will be identical by descent) (Jobling et al., 2004b). When effective population size is small, the amount of genetic variation is decreased compared to larger populations. This characteristic is useful when mapping genes underlying complex diseases as there is a tendency in these populations for reduced genetic heterogeneity (Stephens and Bamshad, 2007). Migration also serves to reduce genetic heterogeneity. Migration of part of a community from the main to a new environment and genetic isolation, has the effect of reducing the amount of genetic variation in the isolated population (King et al., 2006). This is termed founder effect, evidence of such is found by clines in allele frequency from the founding population along the route of migration.

Jewish populations have shown close to 50 rare Mendelian disorders at significantly higher frequencies than in the general population (Jobling et al., 2004b). For example Tay-Sachs disease, an autosomal recessive disease that leads to destruction of the central nervous system is one hundred times more common in Ashkenazi Jews (Klug, 2009). This unusual increase of genetic susceptibility is partly due to genetic drift, founder effect also being a component. Myopia researchers have also examined genes underlying refractive error in Ashkenazi Jewish communities, mainly in America (Stambolian et al., 2004; Stambolian et al., 2006; Wojciechowski et al., 2006) and have demonstrated significant linkage between chromosome 22q12 and myopia. Reduced genetic heterogeneity in founder populations leads to another advantage in genetic mapping studies.

Linkage disequilibrium (LD) refers to the non-random distribution of alleles that reside at different locations in the genome in a population. The combination of alleles that are found on the same chromosomal segment in a population reflects the degree

of independent assortment that has occurred during meiosis. Alleles that reside on the same portion of chromosome are considered here to form a haplotype. For example, consider two bi-allelic SNPs at different genomic locations but on the same chromosome (i.e two syntenic SNPs). If they displayed independent assortment the proportion of haplotypes observed in the population would reflect the products of the allele frequencies at the SNPs. That is two SNPs which both have two alleles with a frequency of 0.25 and 0.75 for allele A1 and A2 at the first SNP and 0.3 and 0.7 for allele B1 and B2 at the second SNP would display the following haplotype frequencies due to independent assortment;

SNP1	Allele1	A1	0.25
	Allele2	A2	0.75
SNP2	Allele1	B1	0.3
	Allele2	B2	0.7

Haplotypes	A1 B1	=0.25*0.3	0.075
	A1 B2	=0.25*0.7	0.175
	A2 B1	=0.75*0.3	0.225
	A2 B2	=0.75*0.7	0.525

Deviations from these frequencies indicate that some alleles at different points in the genome are being inherited together. In a population this is manifest as linkage disequilibrium. A measure of linkage disequilibrium is given by D which is defined as

$$D = P_{AB} - P_A P_B$$

where P_{AB} is the proportion of observed haplotypes in a population for one allele at either locus of two loci and $P_A P_B$ is the product of their allele frequencies (Jobling et al., 2004b).

Founder populations show increased linkage disequilibrium between genetic markers. When a disease mutation arises on a haplotype, over generations it will gain many different haplotypic backgrounds (i.e. it will occur with different upstream and downstream alleles) due to crossing over. In founder populations the number of *different* alleles proximal to the causal mutation will be decreased. An association study or linkage analysis uses genetic markers and assumes if the markers are more

often in affected individuals, the disease mutation is close by. Affected individuals displaying high genetic diversity will display multiple haplotypes on which the disease mutation resides, while affected individuals displaying low genetic diversity will show a small number of haplotypes, increasing the strength of signal of a marker close by. Other founder populations have been studied in an effort to map genes underlying myopia.

The Old Order Amish are an isolated community in the USA which emigrated from south western Germany in the 18th century (King et al., 2006). They have high levels of endogamy (marriage within the community) and therefore also display reduced genetic variability. Rare diseases such as Ellis-van Creveld syndrome and cartilage-hair hypoplasia are at an increased prevalence in Amish populations (King et al., 2006). Myopia researchers mapped a quantitative trait locus for myopia to chromosome 8p23 in an Old Order Amish community (Stambolian et al., 2005) confirming an earlier finding in a UK twin cohort (Hammond et al., 2004). A genome-wide linkage scan for high myopia was undertaken in the Hutterite community (a genetically isolated community in the USA originating from Germany and Switzerland) and resulted in significant linkage on chromosome 10q21.1 (Nallasamy et al., 2007). Other genetic mapping efforts in myopia research in founder populations have been in Sardinia (Biino et al., 2005) and the Croatian Island of Korcula (Vatavuk et al., 2009).

Another strategy to increase the chance of finding a true positive linkage signal is to increase the number of individuals examined. Linkage analysis can often rely on large multigenerational families. In the absence of such a 'family pooling' linkage strategy (Wright et al., 1997) may suffice to identify a gene underlying the disease. In this strategy data from small nuclear families are pooled together and analysed via linkage. Myopia researchers have used this to confirm linkage signals for a number of MYP loci (Farbrother et al., 2004b; Li et al., 2009b).

1.1.4 A new strategy

Risch and Merikangas, (Risch and Merikangas, 1996) estimated that linkage studies are sufficient to identify mutations that confer a 4 fold increase in genotypic relative

risk (GRR, an increased chance that an individual with a certain genotype has the disease), but smaller effects may be impractical to detect (a GRR of 2 may require 2,000 sibling pairs). Common diseases are hypothesized to be caused partly by genes of moderate effect. Thus a new experimental paradigm was developed to allow genes underlying common disease to be mapped. To detect genes of moderate effect many subjects need to be recruited. This is less feasible to achieve when families are the units of analysis. Unrelated individuals who are enrolled in a cohort allow for large numbers of subjects to be analysed in a genetic study and DNA to be stored for analysis. Such studies are becoming increasingly widespread with participants in the thousands (Gurwitz et al., 2009).

The new strategy relies on an idea that common variants underlie susceptibility to common diseases. The common disease/common variant hypothesis (CD/CV) is critical to success in genome-wide association studies of complex disease. It relies on theoretical considerations of population genetics (Reich and Lander, 2001), the main points of which are summarized below.

- Consider one rare (severe) and one common (mild) monogenic disorder.
- Each has the same mutation rate but has a different frequency of disease alleles (e.g. 0.001 and 0.3 respectively).
- The difference in frequencies is due to selection pressure.
- In an ancestral population, all loci had a simple allelic spectrum with a predominant disease causing allele that was responsible for more than 90% of cases.
- The ancestral population underwent an expansion to reach its current size (75,000 years ago). This should have lead to all disease loci having a complex spectrum (many alleles with varying frequencies).
- However, allelic diversity increases for rare disease rapidly as old disease alleles are lost to selection. Alleles of common, milder diseases are maintained in the new large population.

In terms of medical genetics, it is suggested that common genetic variants are responsible for common diseases. Therefore an association study of common variants may be able to identify variants underlying disease risk (Reich and Lander, 2001).

In linkage analysis 300 markers were sufficient to identify a region of 5-10 cM apart which would be then analysed in finer detail with more genetic markers. One of the reasons linkage analysis is not preferable to map common diseases when compared to genome-wide association studies is as follows; even when the region of interest has been identified and a huge number of markers have been genotyped, the relatively small number of individuals in a large multigenerational family will not provide the necessary number of recombination events to narrow down the interval which harbours the genetic mutation (Botstein and Risch, 2003). By sampling large populations of unrelated individuals more haplotypes will be observed (which can be inferred to have occurred from previous recombination). Comparing the same stretch of DNA in two unrelated individuals is analogous to comparing two related individuals from a massive multigenerational pedigree which stretches back to the last common ancestor. The ancestral haplotype will have been broken up by meiotic events during its descent which are now visible in both individuals. Thus the region that harbours the disease mutation can be narrowed.

Linkage analysis has identified about 1,200 disease genes. Linkage analysis has primarily used microsatellite markers (chromosomal segments that contain a sequence of nucleotides that repeats for a short distance (King et al., 2006)) to facilitate genetic mapping. Microsatellite markers are relatively evenly spaced across the genome (King et al., 2006), display high heterozygosity (Gulcher, 2007) and are amenable to amplification by polymerase chain reaction (Weber and May, 1989). By the mid 1990s comprehensive genetic maps (linear arrangement of polymorphic sites on a chromosome) of microsatellites were available (Gyapay et al., 1994; Dib et al., 1996) containing over 5,000 microsatellite markers covering approximately 3,500 centimorgans (cM, 1% rate of crossing over (King et al., 2006)) with an average interval of about 1.6 cM. In an initial linkage scan, when no prior indication of where in the genome a casual mutation is located, researchers set out to undertake a genome-wide scan. Typically 300 microsatellite markers were used in this approach with an average spacing of 10 cM (Antonarakis, 1994). Myopia research is no exception with

many genome-wide linkage analyses already undertaken (Andrew et al., 2008; Lam et al., 2008a; Paget et al., 2008a) (see Appendix A for more). There are exceptions to the use of microsatellites for genome-wide linkage analysis but they occur after the sequencing of the human genome. Replication of a number of linkage signals for myopia were undertaken using genome-wide single nucleotide polymorphism (SNP) linkage analysis in 2008 (Nurnberg et al., 2008) and 2009 (Li et al., 2009b).

Single nucleotide polymorphisms are the most frequent type of genetic variant and are amenable to high through-put genotyping (Wang et al., 1998). The human genome project provided the location of 2.1 million SNPs in the human genome (Venter et al., 2001). Gabriel et al. (Gabriel et al., 2002) showed that the human genome could be considered as consisting of stretches of DNA that are not separated by recombination. The HapMap project began in 2002 to determine the haplotype structure of the human genome (Jobling et al., 2004b). By 2007 over 3 million SNPs had been mapped to the human genome. It has been estimated that the average amount of linkage disequilibrium between two SNPs is close to 3 kilobases (kb) and that roughly 500,000 SNPs would be needed to conduct a genome-wide scan for a complex disease (Kruglyak, 1999). It was estimated that using commercially available genotyping platforms (500,000 SNPs) 80% of the 3 million SNPs identified by the HapMap project would be captured (Kruglyak, 1999). In other words due to linkage disequilibrium and haplotype blocks, using a subset of SNPs in the human genome, variation in a much larger set could be studied.

1.1.5 Multiple testing

A P-value measures the chance of observing an outcome (test statistic) at least as large as that observed when the null hypothesis is true. The P-value provides evidence in favour or against a null hypothesis. When the chance of observing an outcome given the null hypothesis is very small it provides evidence that the null hypothesis is not true. A threshold P-value can be used to decide whether to accept or reject a null hypothesis. If an outcome is observed with a P-value of α and α is less than a given threshold it can be declared as evidence that the null hypothesis is not true with an α chance of being incorrectly declared significant. α can be termed the false positive rate or the chance of falsely rejecting the null hypothesis also known as a type I error.

Multiple testing refers to considering many similar hypotheses on one set of data. Multiple testing can lead to extreme P-values occurring when the null hypothesis is true. The chance of observing a false positive can be measured by the family-wise error rate which is defined as the probability of at least one type I error. This probability can be controlled by multiple comparison procedures that modify the expected type I error rate and allow for a more conservative threshold for declaring significance to be defined.

Risch and Merikangas (Risch and Merikangas, 1996), consider the effect of using 500 markers in a linkage analysis on the chance of observing a false positive result. They note that a lod score of 3 (i.e. a logarithm of odds in favour of linkage of $\frac{1000}{1}$) gives an α of 1x10e-4 (or $\frac{1}{1000}$). In a linkage analysis of 500 markers the probability of no false positives is 95%. That is 500 tests each with an α of 1x10e-4 (500 x 10e-4 = 0.05) or in other words 5% of results are potentially false positives. They note that the equivalent false positive rate for 1,000,000 tests is 5 x 10e-8. More generally the relationship between false positive rate before and after multiple testing can be expressed by $\alpha_{after} \times N = \alpha_{before}$ where N is the number of tests. It can easily be seen that the false positive rate for 1,000,000 tests given by Risch and Merikangas can be derived by $\alpha_{after} \times 1{,}000{,}000 = 0.05$. It is noted in that paper that 1,000,000 tests (markers) may not be necessary to cover the entire human. Since markers that are located in close proximity may display linkage disequilibrium the number of independent tests may be less.

It is noted by Dudbridge and Gusnanto (Dudbridge and Gusnanto, 2008) that the effective number of independent tests was estimated in the International HapMap project to be approximately 150 per 500kb. They go on to note that to maintain a family-wise error rate of 5% for a 3 Gb genome a significant threshold of 5.5 x 10e-8 would be needed. This threshold can be calculated as follows; 150 independent tests per 500kb translates to 900,000 independent tests per 3Gb. Subbing into $\alpha_{after} \times N = \alpha_{before}$ gives $\alpha_{after} \times 900{,}000 = 0.05$ or 5.56 x 10e-8.

Another approach to multiple testing of genetic data involves a permutation test that allows a threshold P-value that maintains a 5% false positive rate to be estimated from the data (Churchill and Doerge, 1994). A permutation test can be explained by considering two group of observations (both of size N) with an observed difference in their means, m_1 and m_2, of d_1. Group membership is can be shuffled where there are still N observations per group but now each contains a dissimilar set of observations. For example if N = 4 and the two observed groups were (a,b) and (c,d) with $m_1 = 0.2$, $m_2 = 0.3$ and $d_1 = 0.1$, the permutation test involves calculating $d_1$for (a,c) versus (b,d) and (b,c) versus (a,d) and construction of a probability distribution. The P-value of the original observation can be then calculated empirically. A d_1 or P-value that occurs 5% or less times can be calculated and the threshold for determining significance obtained; however a more realistic N would be necessary. Dudbridge and Gusnanto (Dudbridge and Gusnanto, 2008) used a permutation test approach to estimate a genome-wide P-value for a 5% family-wise error rate using genotype data obtained from a commercial 500K array and approximately 3,000 samples of 7.2×10^{-8}.

1.2 The myopia phenotype

In this thesis two measures of refractive error are discussed; subjective refractions collected from optometric practises and non-cycloplegic autorefraction measures of myopia taken during a visit to an ALSPAC clinic. Myopia is defined as -0.5 D or less if the measure was a subjective refraction. If the measure was a non-cycloplegic autorefraction the point at which myopia is inferred to be present depends upon the age of the subject. Subjects were close to the age of fifteen for data used in Chapters 2 to 5 and myopia is defined as less than -1 D by non-cycloplegic autorefraction. However, in Chapter 6 subjects were age eleven and myopia is defined as less than -1.5 D by non-cycloplegic autorefraction. Both these definitions of myopia are supported by validation studies indicated in the text.

1.2.1 Classification

One possible difficulty in finding a cause of a disorder is due to classification of the phenotype. Classification of a disorder must be precise in order to identify affected and unaffected individuals (Woodward, 2005). Classification can be made on anatomic, physiological and pathological grounds (Aylsworth, 1998). Syndromic forms of a disease offer a clear classification from non-syndromic forms. For example retinitis pigmentosa, although showing clear Mendelian patterns of inheritance, also occurs with severe congenital sensory deafness and neurofibromatosis type 1 (Bird, 1998).

Myopia shows syndromic expression and (much more frequently) non-syndromic forms. For example the MYP1 locus is located on chromosome Xq28 (Schwartz, Haim and Skarsholm, 1990). MYP1 was mapped by linkage analysis in families with a syndromic form of myopia with associated hypoplasia of the optic nerve head and colour blindness and other characteristics collectively known as Bornholm eye disease named after the Danish Island where the families were resident. This syndrome has been identified in other families in Minnesota (Young et al., 2004) and in the UK (Michaelides et al., 2005). Myopia forms part of a distinct phenotype in Bornholm eye disease (BED) and it can be inferred that linkage analysis of families with BED represents analysis of a distinct phenotype. This distinction is more general however,

as in two of three families examined with BED, protanopia was present while in the third family, deuteranopia was observed.

Pathological and physiological characteristics of myopia are used to identify different forms of myopia. 'Pathological myopia' is first present in childhood, displays a quick progression and exhibits pathologies of the choroid and retina (Edwards, 1998b). In some cases those with pathological myopia find it difficult to attain normal vision after correction with spectacles or contact lenses (Edwards, 1998b). 'Physiological myopia' is a classification of myopia based on whether ocular components seem abnormal in comparison to those in individuals with good vision (incidentally this classification would require measures of axial length, corneal power and crystalline lens power (Edwards, 1998b)). Physiological myopia is used to describe eyes with no changes to the fundus which are typically seen in pathological myopia. It can be hypothesized that pathological and physiological forms of myopia have different aetiologies. In turn different genes may underlie their development. Myopia scientists have used such distinctions to investigate the genetics of myopia. The first published genome-wide association study of pathological myopia classified affected status by an axial length greater than 28 mm (Nakanishi et al., 2009b).

In high myopes, axial length is nearly always abnormally large. High myopia is often associated with pathological changes in the retina (when both are present the myopia can be termed pathological). However it is not infrequent that subjects with common myopia (less than 6 dioptres (D)) display myopic crescents (changes in the fundus) (Edwards, 1998b) a characteristic of high myopia. It has been observed that temporal crescents are much more common in Chinese than Caucasian eyes, with 84% of Chinese eyes with between 2 to 4 D myopia displaying such a pathology (Edwards, 1998b). Studies in samples from pathologically unique populations of myopes may improve the chance of successful mapping of genes responsible for disease pathology.

1.2.2 A continuous trait

Ideally (for the purposes of mapping a disease gene or epidemiological analysis) medical conditions are recognizably 'abnormal' or 'normal'. Furthermore an examination by a trained clinician would lead to identification of the presence or

absence of a disorder (Aylsworth, 1998). Diseases with full penetrance and a well defined phenotype are most amenable to linkage analysis. An example is achondroplasia or dwarfism, a rare autosomal dominant condition where patients are typically under 4ft 4in in height (Hartl, 1983). The causal gene, Fibroblast Growth Factor Receptor 3 (FGFR3), has been mapped to 4p16.3 (King et al., 2006).

Refractive error defines part of a continuous scale from hyperopia through emmetropia to myopia. The definition of myopia is based on a patient's difficulty in seeing objects in the distance. This coincides roughly with less than -0.5 D refractive power. However, the exact point when a person becomes myopic is somewhat arbitrary (Edwards, 1998b). A similar decision exists for high myopia which is generally diagnosed when a subject displays close to – 6 D refractive error. Traits that are continuous, such as height and blood pressure that are then dichotomised into what is considered a normal phenotype or a medical condition will require a threshold that is biologically meaningful to minimize genetic heterogeneity (Aylsworth, 1998) and increase the chance of successful mapping of a disease gene. Myopia falls into this category. The presence or absence of either common myopia or high myopia can be defined, on examination by a clinician, but within that definition there is a high degree of phenotypic variability (ranging from -0.5 D or less for common myopia or -6 D or less for high myopia).

High and common myopia display varying degrees of refractive error. In linkage analysis of high myopia, family members have displayed some myopia but the degree of which varies. It can be inferred that other factors modulate the development of myopia in these families and/or that many genes are responsible (due to the phenotypic diversity). Myopia is part of a continuous spectrum of refractive power of the eye. This will increase difficulty to identify a causative gene using linkage analysis.

1.2.3 Age of onset

Studies concerned with linkage analysis of high myopia often examined families in which the age of onset of myopia was young and myopia was progressive (Young et al., 1998a; Young et al., 1998b). While age of onset in some members of the families

examined tended to be early childhood others displayed more variation, from young to adult onset myopia. The degree to which myopia manifests has been associated with age. Some forms of myopia occur at birth and remain throughout life (congenital), some develop from the ages of 5 to 15 (youth onset) and from the ages of 20 to 40 (adult onset). It is unclear what factors determine age of onset of myopia (Attebo, Ivers and Mitchell, 1999; Giordano et al., 2009).

The observation that myopia varies depending on age complicates definition of affected and unaffected individuals in a pedigree and unfortunately is further complicated by whether the presenting myopia is derived from genetic or environmental factors. Huntington disease (HD, detailed in 1872 by an American physician George Huntington) is a degenerative brain disorder which is caused by an autosomal dominant mutation in the Huntington gene (King et al., 2006). Linkage analysis led to a region of chromosome 4 to be identified that harboured causative mutations of the disease in 1983. It wasn't for ten more years that the causative gene was to be identified (The Huntington's Disease Collaborative Research Group, 1993). HD patients display an expanded CAG repeat at the 5' end of the gene. Huntington disease displays complete penetrance and individuals carrying more than 35 copies of the CAG repeat develop HD. The age of onset of HD is inversely correlated with CAG repeat length. It has been noted that half the total amount of variation in age of onset of patients with HD is explained by the causative mutation with correlations ranging from -0.69 to -0.75 (Farrer and Cupples, 1998).

The age of onset of myopia may be under both genetic and environmental control. Unlike Huntington disease where presence of a genetic aberration defines in part the age of disease onset, the amount of influence of either genetics or the environment in myopia onset is unknown. There is evidence to suggest that exposure to certain amounts of nearwork is associated with youth onset (Saw et al., 2002a) and adult onset (Zadnik and Mutti, 1987; McBrien and Adams, 1997) myopia. Myopia in very young children may be less due to nearwork (there is little time for exposure) and a recent study found no association between the two (Low et al., 2010), furthermore genetic susceptibility to environmental influences of myopia cannot be ruled out.

1.2.4 Ocular components

It has been observed that axial length is negatively correlated with refractive error (Gonzalez Blanco, Sanz Fernandez and Munoz Sanz, 2008). Similarly it has been noted, myopes and emmetropes display differences in corneal curvature (Grosvenor and Goss, 1999) and corneal power (Rosenfield, 2006). In addition to the crystalline lens, the refractive power of these components collectively alter the convergence of light on the retina (Erickson, 1991).

The posterior focal length of the eye is defined by a number of components; corneal power, anterior chamber dept and crystalline lens power (Edwards, 1998b). If posterior focal length exceeds axial length (the distance from anterior to posterior poles (Millodot and Laby, 2002)) rays of light will focus ahead of the retina (creating myopic vision). Furthermore if axial length is shorter than posterior focal length, light rays will focus behind the retina (leading to hyperopic vision). It is thought that myopia is a result of imbalance between posterior focal length and axial length. Myopes often display longer axial lengths than non-myopes. There is also a strong correlation between refractive error and axial length (Rosenfield, 2006). However in the case of low amounts of myopia, axial length is not thought to be significantly different from those with normal vision (Wildsoet, 1998). Instead low myopia may be due to an aberration of the combination of the components of posterior focal length.

Corneal curvature also contributes to myopia development. The cornea is at the anterior portion of the eye with a slightly curved surface compared to the rest of the globe (Millodot and Laby, 2002). It is composed of five distinct layers and is transparent due to the regular arrangement of collagen fibres. It is a refractive surface having a power of about 42 D (Millodot and Laby, 2002). It has been noted that there is sometimes a correlation between axial length and corneal curvature (Wildsoet, 1998; Gonzalez Blanco et al., 2008). This implies that as axial length increases, the cornea becomes flatter and in turn may allow a balance between the focal length and axial length to achieve emmetropisation and normal vision. However in other findings only a weak relationship between corneal power and refractive power have been found (Rosenfield, 2006). This implies that for a subset of subjects the cornea plays a role in the development of emmetropia. However it has been noted that in some cases of myopia the corneal curvature tends to steepen (Grosvenor and Goss, 1999) which

has the effect of focusing light ahead of the retina and increasing myopia. Furthermore it has also been noted that there is a significant difference between myopic and emmetropic subjects in terms of corneal power (Rosenfield, 2006). This would suggest that corneal power plays a role in the development of refractive error and it is possible to suggest that the balance between the cornea and axial length is sometimes aberrant.

In a myopic eye, the size of the globe increases leading to stretching of the ocular tissues. The crystalline lens which is suspended between the iris and vitreous humour by the ciliary body (Millodot and Laby, 2002) may be stretched by enlargement of the orbit (Wildsoet, 1998). During progressive eye growth during childhood, the crystalline lens displays a process of thinning (Zadnik, 1997). This process leads to a hyperopic shift; a shortening of the posterior focal length (Wildsoet, 1998), and is hypothesized to allow for normal vision to develop during emmetropisation (Wildsoet, 1998). The lens of children aged 6 to 16 has been shown to decrease in power by approximately 3 D (Garner et al., 1998). It is also been noted that the power of the lens varies. The lens also shows small variations in its relative refractive index (the ratio of speed of light in air to the speed of light in another medium (Millodot and Laby, 2002)). It is the least optically dense at the surface and most at the centre. This variation is related to protein concentrations of the lens (Grosvenor and Goss, 1999). The difference between the centre and surface of the crystalline lens has been measured in terms of refractive index as approximately 0.2 (Garner et al., 1998).

It is not just the optical properties of each component individually that can change refractive power. Alignment of each component with each other is also critical to the resolving power of the eye. For example a displacement of 1 mm of the cornea can induce approximately 0.5 D of myopia. A similar amount of myopia is produced if the crystalline lens rotates forwards by 11 degrees (Erickson, 1991).

A balance between focal length and axial length defines refractive error. At least three distinct components are involved in regulating this balance; axial length, corneal curvature and the crystalline lens. These ocular components are morphologically and functionally different and it can be hypothesized that they are under control of at least some distinct genetic factors. In terms of mapping a genetic element or identifying a

novel risk factor, the cascade that normal vision requires is similar to the complexity of hearing. Disruption of one of the components that is required for normal vision may lead to myopia, each of which may have a distinct genetic architecture and a number of environmental modulators.

The physical dimensions of these components each effects the power of the eye. Each of the ocular components can be measured. The curvature of the anterior corneal surface can be measured with a keratometer (Millodot and Laby, 2002). It is typically close to 7.8 mm (Millodot and Laby, 2002). Ultrasonography is used to measure axial length and lens thickness (Millodot and Laby, 2002). Ultrasound waves are emitted at high frequency close to the eye. Reflections of ultrasound waves (echoes) allow biometric measurements to be made (Millodot and Laby, 2002). The above measurements require specialised equipment. In an effort to balance accuracy and cost some studies measure only some of the ocular components.

1.3 Epidemiology

1.3.1 Prevalence of myopia

Prevalence can be defined as the number of existing cases of a disease at a particular point in time (Woodward, 2005). Prevalence studies often indicate the degree of burden of a disease within a country and inform on allocation of resources at the governmental level (Woodward, 2005). Prevalence studies rely on random sampling (in which each individual in the population has an equal chance of inclusion in the study (Woodward, 2005)). The prevalence of myopia varies in four dimensions; a) across regions b) age c) gender and d) time.

a) The prevalence of myopia varies across regions (Zadnik and Mutii, 1998) with highs in Asian countries. The prevalence of myopia in children aged 15 in China was between 35% to 55% (Zhao et al., 2002), 76% for children of the same age in Taiwan (Lin et al., 1999), while in Poland the prevalence of myopia was 13% in school age children (Czepita, Zejmo and Mojsa, 2007). It has been hypothesized that differences in environmental exposures may underlie variation in refractive error prevalence across regions. An extension of this is regional differences in exposure and disease risk introduce new possibility of confounding (Woodward, 2005). It is also possible that variation between regions in myopia prevalence is due to genetics. Diseases with a specifically genetic cause show region specific prevalences. Huntington disease is frequent in populations of Western Europe but it is 10 to 100 times more prevalent in Finland (Jobling et al., 2004b). Similarly cystic fibrosis is found at a high frequency in Western Europe compared to other populations (Jobling et al., 2004b). It has been observed that mutations leading to human disease show region specific frequencies and high to low gradients across geographical areas. Galactosemia is a hereditary disease that leaves patients unable to digest galactose. Symptoms can be severe but can be avoided by employing a diet free of galactose (King et al., 2006). It is caused by mutations in the enzyme galactose-1-phosphate uridyl-transferase, the gene of which is located on chromosome 9p.13. The two most common galactosemia mutations (Q188R and K285N) show peaks in Ireland (93%) and Eastern Europe (34%) (Flanagan et al., 2010).

Investigation into whether genetics or environmental factors contribute to differences in myopia prevalence across geographical areas can be achieved by examining the prevalence of myopia in differing ethnic groups. In Singapore those of Malay ethnicity displayed lower prevalence of myopia than Singaporeans of Chinese ethnicity (Saw et al., 2008). A study of children of Asian ethnicity now living in Australia found a much lower prevalence (3.3% versus 29.1%) of myopia in children compared to their counterparts in Singapore (Rose et al., 2008b). Another study of Canadians of Chinese ethnicity found comparable prevalence (64%) to children in urban East Asia (Cheng, Schmid and Woo, 2007). In the Australia study exposures to risk factors and prevalence in both groups were measured directly. There were large differences between groups in terms of time spent reading and time engaged in outdoor activities. Adjustment for time outdoors or nearwork did not account for the difference in prevalence, although both were independently associated with myopia. Wu et al., found that Singaporean individuals of Chinese ethnicity displayed significantly more myopia (82%) than those of Malay (65%) or Indian (69%) ethnicity. After adjusting for education these differences persisted (Wu et al., 2001). Education was independently associated with risk of myopia.

b) Although prevalence studies employ random sampling it is often better to draw a sample from a separate subgroup of a population. This is termed stratification and can improve precision (Woodward, 2005). The prevalence of myopia in preschool children (age range 6-71 months) was less than one percent in Caucasian children (Giordano et al., 2009). Myopia shows age specific prevalence rates rising steadily from ages 7 to 15 years (Zadnik and Mutii, 1998). Rajan et al. estimated age stratified prevalences of myopia for Singaporean children at ages 7, 9 and 12 of 25%, 32% and 51% respectively (Rajan et al., 1998). In another large study in China, myopia was almost absent at age five but increased to close to 50% by age 15 (Zhao et al., 2004). The Refractive Error study in Children was designed to allow the prevalence of myopia to be estimated in a precise and accurate way in age and sex strata in different countries (Negrel et al., 2000). In Chile myopia prevalence increased from 3.4% at age five to close to 20% at age 15 (Maul et al., 2000). However an increase in myopia prevalence between the ages of 5 to 15 is not always observed. Pokharel et al. found that prevalence of myopia was 3% for children aged 15 in Nepal (Pokharel et al., 2000) and Dandona et al. (Dandona et al., 2002) found a prevalence of 4% in children

aged 15 in rural India. These findings may be exasperated by differing exposures. For example urban residence is known to be associated with higher prevalence of myopia and that may partly explain the lower prevalence found in rural India. In the Nepalese study the sample was drawn from a region that has agriculture as its main economy. This could explain the lower prevalence, however the area also is above the national average in terms of economic wealth, which itself is a pre-disposer to myopia. The Chilean study was drawn from the metropolitan area of Santiago, the country's capital. A proportion of cohort studies follow subjects through time to record exposure and risks at different ages and can be termed longitudinal cohorts. These types of cohorts are especially valuable when comparing changes in prevalence with age. A longitudinal cohort in Japan observed a myopia prevalence of 35% on entering school and 58% upon leaving for the same students over a six year period (Hirai, Saishin and Yamamoto, 1998). Similarly Edwards et al. found a myopia prevalence of 11% at age seven and 55% to 58% by the age of 12 for longitudinal data of children in China (Edwards, 1998a).

c) Differences in the prevalence of myopia between differing genders are also reported. Myopia was found to be higher in females in China (Zhao et al., 2002) and in Malaysia (Goh et al., 2005). A relationship between gender and myopia may be related to age (Zadnik and Mutii, 1998) as in some older cohorts females do not show higher myopia prevalences. A cohort sampled from the elderly population in Taiwan found the prevalence of hyperopia to be increased in females (Cheng et al., 2003) and in Australian adults, females were found to have a more hyperopic refraction than males (Attebo et al., 1999). It is hypothesized that puberty may play a role in gender differences observed in myopes in early teens (Zadnik and Mutii, 1998).

d) Furthermore comparison of present and past risk factor studies is complicated by changing exposures. Myopia prevalence was significantly higher in the USA in 2004 (41.6%) than in 1971 (25%) in both white and black populations (Vitale, Sperduto and Ferris, 2009). In a Chinese population, the prevalence in the elderly was 19% similar to Caucasian populations (Cheng et al., 2003), but the prevalence in school ages children is above 50% (Zhao et al., 2000). In both studies the effect of changing exposures are hypothesized to account for some of the increase in myopia prevalence.

1.3.2 Numerous risk factors

There are many risk factors that have been associated with myopia (see Appendix B). Myopia is similar to other complex diseases in having a range of risk factors (over 200 risk factors have been associated with coronary heart disease (Woodward, 2005)).

1.3.2.1 Correlation between myopia risk factors

A number of myopia risk factors are correlated. Myopia is associated with occupation. A study of Japanese subjects found that men in managerial roles and females in clerical roles displayed increased amounts of myopia (Shimizu et al., 2003) and clinical microscopists (McBrien and Adams, 1997) and law students (Zadnik and Mutti, 1987) have been shown to have increased myopia levels. Income is associated with myopia. Wong et al. found that individuals in Singapore with higher incomes had an increased rate of myopia (Wong et al., 2000), a relationship also found in the adult population of Sumatra and in the USA (Sperduto et al., 1983). Education is associated with myopia. Increased levels of education are associated with higher rates of myopia in Singapore military conscripts (Saw et al., 2001). Increasing levels of education have been associated with myopia in the USA (Sperduto et al., 1983). Higher levels of education, increased income and non-manual occupations are closely related. An individual with a college education is more likely to receive a higher income and to work in a non-manual occupation. Other factors that are correlated with levels of education, occupation and income are urban versus rural residence and nearwork. Residence in urban centres is associated with increased levels of myopia. Ip et al. found that after adjusting for a number of myopia risk factors (including nearwork, age and gender) Australian children living in urban areas were at an increased risk of myopia (Ip et al., 2008a). Zhang et al. found that after correcting for a number of risk factors urban residence was associated with levels of myopia in Chinese children (Zhang et al., 2010). Urban residence, non-manual occupation, increased levels of education and increased income may be related to nearwork. Visual activity is associated with myopia (discussed below) and it is hypothesized that environmental stimulation via mechanisms of normal vision can increase risk of myopia. The associations between myopia and occupation, income, education and urban residence could be due to increased stress on the visual system associated with these tasks.

1.3.2.2 Measurement of risk factors and their quality

In an effort to balance cost and accuracy measures of a risk factor can vary in quality. Self reporting is a cheaper alternative to direct measurements but may generate some misleading information (Woodward, 2005). For example there may be differences in the reporting of a child's reading habits depending on whether parents are interviewed or children. Objective measures serve to reduce uncertainty in an estimate. Objective measures however can sometimes be distressing or difficult to carry out which in turn can lead to missing values. Missing values may be biased, being more likely in certain people than others (Woodward, 2005). Myopia researchers use measures of different levels of objectivity. Myopia is associated with outdoor activity in studies that use self reported (Dirani et al., 2009b) and parental reported (Rose et al., 2008a) measures as well as objective measures of activity (Deere et al., 2009). Myopia is associated with school achievement using self reported measures of number of books read and scores in objective measures such as standard IQ tests for reading and maths (Saw et al., 2004b; Williams et al., 2008a). Height, weight and birthweight can be measured objectively (although studies involving measures of weight after the teens may suffer from increased missing values of the obese). Small significant associations have been found between myopia and height (Saw et al., 2002b).

1.3.3 Causality

Causality is a primary concern of epidemiology. Association studies seek to identify whether an exposure to a certain factor is related to a disease. In a genome-wide association study, subjects with a disease (cases) are compared to subjects without the disease (controls). In some epidemiological studies published, subjects with a certain exposure (exposed) are compared to subjects without exposure (non-exposed) (Gordis, 2009). This type of study is termed a cohort study. In cohort studies the proportion of subjects with the exposure is compared in subjects with the disease and in subjects without the disease. Epidemiology research aims to identify whether there is a causal relationship (causality) between exposure and disease. An association study may be thought of as the first step to identifying causality (Gordis, 2009). In a large study of British doctors Doll and Hill (in 1964) identified an association between smoking and lung cancer (Woodward, 2005). They found that the chance of death due to lung cancer was lowest in individuals who never smoked and that the

chance of death increased as subjects who smoked some, often and a lot were considered. That subjects with the least exposure are at least risk and that increasing exposure is associated with an increased chance of morbidity are important signs of association that are still examined in cohort studies. Parental history of myopia (when one or more of a subject's parents display the disorder) is consistently associated with myopia (Mavracanas et al., 2000; Mutti et al., 2002; Ip et al., 2007; Jones et al., 2007; Konstantopoulos, Yadegarfar and Elgohary, 2007). A number of cohort studies examining whether number of myopic parents is associated with myopia have been undertaken (Mutti et al., 2002; Jones et al., 2007). A dose dependent effect of number of myopic parents and development of myopia has been demonstrated (Mutti et al., 2002; Ip et al., 2007; Jones et al., 2007).

If an association is to suggest causality there should be a plausible biological explanation (Woodward, 2005). An association between parental myopia and myopia in the next generation suggests a genetic factor that predisposed families to myopia development. Heritability studies of refractive error show that it is heritable with a portion of variability in the trait explained by additive genetic factors. Therefore it is probable that an association between myopia and parental myopia is due in part to genetics. However individuals from the same family tend to have the same lifestyle, in terms of reading habits for example, which is known to be a risk factor for myopia. Therefore amount of time spent reading may explain part of the relationship between myopia and parental myopia. In other words the familial relationship may be a confounding factor.

One definition of a confounder is a factor that is associated with both exposure and disease but a consequence of neither (Woodward, 2005). A good example is grey hair and age-related diseases, for example stroke (Woodward, 2005). Stroke patients are likely to have grey hair but this does not indicate that grey hair is a risk factor of stroke. Being in a family is not a risk factor for myopia, but an increased amount of nearwork is a risk factor. Strategies for dealing with confounding use a) a priori knowledge about the biological mechanism at work and b) analytical methods that examine a relationship of interest with and without a confounder present (Woodward, 2005).

a) It is known a priori that myopia is heritable and therefore parental myopia may be due to a genetic element. In a study of refractions of children aged less than six years, individuals with two myopic parents were at an increased risk of myopia (Low et al., 2010). Due to the young age of subjects, little exposure to nearwork can be inferred. A link between family history and myopia via genetic factors is therefore strengthened. Similarly in another study it was found that children with two myopic parents had longer eyes before the time of myopia progression (Zadnik et al., 1994).

b) Studies that find an association between parental myopia and myopia will often assess the relationship with and without a number of confounders present. Adjusted linear or logistic regression is used to estimate risk of myopia after the effect of a confounder has been identified. For example family history is associated with myopia independent of nearwork (Zadnik, 1997; Mutti et al., 2002).

An association study should also be repeated in other cohorts to protect against a chance finding that may occur due to sampling issues. In studies of smoking and lung cancer, subjects were followed over a period of 40 years, over which time smokers were consistently more at risk of death of lung cancer (Woodward, 2005). The practise of replicating associations between myopia and a risk factor is undertaken for the vast majority of myopia risk factors. Sometimes this can be critical to understanding the relationship suggested by the association findings.

Strong evidence for an association between the use of night lighting (Quinn et al., 1999) for children and myopia was published in the journal Nature with strong implications for the development of myopia; by discontinuing the use of night lights the chance of myopia would be greatly reduced. Replication was attempted in a number of studies, all of which failed to find a similar association (Gwiazda et al., 2000; Guggenheim, Hill and Yam, 2003; Konstantopoulos et al., 2007). See Appendix B for examples of risk factors associated with myopia.

Longitudinal data will also help to identify a possible causal link between exposure and disease. Exposure to a risk factor should precede onset of disease. Longitudinal data can help establish an order of events particularly when subjects in the initial cohort are without the disease at the outset (Woodward, 2005). Differences between

future myopes (third grade children who would eventually go on to develop myopia) and future non-myopes (children who remained emmetropic) were observed in the number of hours spent outside (Jones et al., 2007). In another study of parental myopia, children with two myopic parents tended to be less hyperopic before myopia progression (Lam et al., 2008b).

1.3.3.1 Intervention studies

An intervention study is an experiment that allows a clinician or researcher to evaluate the usefulness of a therapy which is designed to treat or prevent a disease (Woodward, 2005). Intervention studies rely on either a known biological pathway or evidence that such a pathway operates. There are examples of intervention studies for the treatment and prevention of myopia. These are largely based on theories of myopia development which centre on evidence that refractive error is modulated by the environment. There is strong evidence in animal studies that myopia can be induced by depriving an animal of patterned vision (Wallman, Turkel and Trachtman, 1978). Form deprivation (FD) is a term given to removal of patterned visual stimulation of the retina (Smith, 1991). It can be achieved by surgically suturing eyelids closed in an animal model of myopia. This technique has been demonstrated to produce large amounts of myopia (up to 15 D) in monkeys and tree shrews (Smith, 1991). Furthermore FD is also associated with increased vitreous chamber dept (Smith, 1998), a feature of myopia in humans. However lid fusion does not only result in loss of pattern vision and it is possible that the ensuing myopia results from loss of illumination (Smith, 1991). It is noted that illumination is necessary for form deprivation (eyelid closure in monkeys and subsequent transfer to a completely dark environment fails to produce myopia (Smith, 1998)). Lid fusion also reduces spatial frequencies and image contrast (Smith, 1991).

Form deprivation myopia is also studied in the chick. Covering the chick eye with a translucent material (occluder) can induce large amounts of myopia (10 D or more) and eye enlargement (Wallman et al., 1978). There is also evidence that form deprivation myopia is controlled in part by cells of the retina. The retina is a light-receptive tissue layer in the eye and responsible for visual activity (Millodot and Laby, 2002). Wallman et al. showed that by partially covering the chick eye with a

translucent occluder, the covered portion grew enlarged and became myopic (Wallman et al., 1987), indicating that the retina responds to FD locally. It has also been shown that FD occurs even when the optic nerve had been severed, indicating a level of control of eye growth that exists at the retina (Wallman, 1991).

Despite complexities in the mechanism of form deprivation myopia, it has led to the conclusion that myopia, in part, may be mediated environmentally via the visual experience. This has stimulated hypotheses about whether the same could be true in the case of human myopia. Form deprivation in animals increases the amount of blur in images presented to the eye. It has been hypothesized that retinal blur due to incomplete accommodation during nearwork, stimulates the eye to grow and leads to myopia. Associations between nearwork and myopia have been made for over a century (Edwards, 1998b). Furthermore myopia prevalences increase at the same time schoolwork begins to increase (Maul et al., 2000; Zhao et al., 2000). It has been noted that extended periods of nearwork can lead to a failure to relax accommodation (Grosvenor and Goss, 1999). This is termed as nearwork-induced transient myopia. It is defined as a short-term myopic shift, on average -0.25 D that occurs immediately after engagement in nearwork for at least 30 seconds (Gilmartin, 1998). It is hypothesized that such short term myopic shifts could act to produce myopia (Gilmartin, 1998). In support of this it has been noted that the accommodative response in myopes is relatively lower than emmetropes (Rosenfield, 1998).

Intraocular pressure (IOP) is a pressure occurring within the eye due to the constant increase and removal of aqueous humor (Millodot and Laby, 2002). The aqueous humour is a clear, colourless ocular fluid that is formed in the ciliary processes and fills the eye, in the anterior and posterior chambers and leaves via the trabecular meshwork (Millodot and Laby, 2002). It has a relative index of refraction of low power (Millodot and Laby, 2002). There is evidence to suggest that IOP is related to the development of myopia. Experimentally induced myopia can be induced by combining raised temperature and increasing IOP in rabbits (Edwards, 1998b). IOP is increased in myopic children (Lam et al., 1998) and in myopic adults (Rosenfield, 1998). Glaucoma results from elevated levels of IOP leading to optic atrophy. The prevalence of myopia is increased in adult glaucoma patients (Mitchell et al., 1999). The relationship between IOP and myopia underpins a number of candidate gene

analyses of myocilin as a myopia susceptibility gene (Tang et al., 2007; Vatavuk et al., 2009; Zayats et al., 2009).

It has been hypothesized that accommodation leads to small changes in intraocular pressure (Rosenfield, 1998) which in turn may lead to myopia. Although the link between accommodation and myopia (either via retinal blur or raised IOP) is tenuous (Rosenfield, 1998) it forms the theoretical basis of some intervention studies of myopia. Cycloplegia refers to paralysis of the ciliary muscles, loss of accommodation and often, dilation of the pupil (Millodot and Laby, 2002). A cycloplegic agent can induce loss of accommodation. A number of cycloplegic agents are antimuscarinic in action. They block acetylcholine from stimulating contraction of the ciliary muscle via muscarinic receptors at parasympathetic nerve endings (Millodot and Laby, 2002). Cycloplegics are used to treat myopia progression due to their action on accommodation and include tropicamide and atropine (Grosvenor, 1998). Treatment with atropine has shown retardation of myopia although with a small effect (approximately 1 D) (Chew et al., 1998). However the efficiency of atropine and other cycloplegics is questioned (Grosvenor, 1998; Grosvenor and Goss, 1999). The use of cycloplegics as a treatment also leads to reading problems and photophobia (high sensitivity to light) and possible adverse reactions (Grosvenor and Goss, 1999). The use of an add lens can help reduce the amount of retinal blur due to lack of accommodation response by decreasing the dioptric stimulus and simultaneously reduces accommodation levels that lead to increased IOP (Grosvenor and Goss, 1999). A number of studies have investigated the use of add lenses (bifocals and progressive addition lenses) on the control of myopia, with mixed results (Grosvenor, 1998; Grosvenor and Goss, 1999).

Adrenergic blocking agents are also used to try to prevent myopia progression. Adrenergic receptors are located on the ciliary epithelium which produces aqueous humour. Adrenergic receptors are stimulated by adrenaline or noradrenaline. Adrenergic blocking agents such as timolol and labetalol can inhibit secretion of aqueous humour from the ciliary epithelium and reduce IOP (Millodot and Laby, 2002). The efficiency of adrenergic blocking agents in control of myopia progression has yet to show efficiency (Grosvenor, 1998; Grosvenor and Goss, 1999).

Chapter 2
Collection of Subjective Refractions

2.1 Introduction

2.1.1 Aims

ALSPAC subjects attended a clinic at age fifteen where an objective measure of refraction (autorefraction) without cycloplegia (paralysis of accommodation) was undertaken. The need to collect extra data arose due to two reasons.

a) Non-cycloplegic autorefraction measures refractive error with a systematic error, this error needs to be quantified before the presence of myopia can be inferred with a high accuracy. The error associated with non-cycloplegic autorefraction is hypothesized to be related to an excess of accommodation that exists in children. Other measures that do not suffer from bias introduced by anomalous accommodation can be used to quantify the systematic error.

Yet other errors can be introduced that will make quantifying the error of interest difficult. Refractive error changes with age. A measure of a subject's refractive error at age 14 and then age at 15 with the same technique may have considerably differences due to the development of refractive error in the intervening year. Therefore it would be erroneous to compare one technique measured when the subject was aged 14 to another technique when the subject was age 15. This is an example of repeating a measure under *changing* conditions which is undesirable compared to *unchanging* conditions (Kirkup and Frenkel, 2006). To validate objective measures of refraction taken at an ALSPAC clinic at age 15 the extra data for validation with another technique was taken at age 15 also, leading to essentially paired data.

b) ALSPAC or The Children of the Nineties focuses primarily on the general health of children born in Avon between 1991 and 1992. Information on ocular health of siblings and parents of children participating in the cohort is limited. ALSPAC also has a definite genetic interest (Jones et al., 2000). It is hypothesized that myopia is

caused in part by genes and furthermore myopia is heritable. To investigate the genetics of myopia in the ALSPAC cohort it is critical to establish that the disorder has a genetic component. This can be achieved via a heritability study. To estimate the heritability of refractive error in the cohort refraction data of relatives of study participants would be necessary. Refractive errors of both parents and at least two siblings (if present) of each study child were targeted for data collection.

In summary, the aim of this study was collection of subjective refraction data for a sub-sample of the ALSPAC cohort. Subjective refraction is a measurement of refraction (the change in direction of light as it passes through the eye) which is based on patient judgement (Millodot and Laby, 2002). Directly opposed to subjective refraction is objective refraction which is a measure of refraction which is not based on a patient's judgement (Millodot and Laby, 2002). Both measures seek to estimate the refractive error of the eye. Refractive error can be termed as the dioptric power of the ametropia of the eye (Millodot and Laby, 2002). Ametropia refers to an aberration of the eye which leads to the image of objects at infinity not forming on the eye when accommodation is relaxed. There are three common ametropias; **myopia**, hyperopia and astigmatism.

2.1.2 Measures

A measure of subjective refraction will often begin with measures of visual acuity. Visual acuity is the capacity to see objects distinctly and in detail (Millodot and Laby, 2002). Visual acuity can be measured using a Snellen chart, which consists of a set of letters viewed at a distance. For example an emmetrope will be able to read all letters on the chart while a myope will have difficultly reading smaller letters. Visual acuity is not a direct measure of how well an image forms on the retina and relies on a patient's judgement as an indication of the refractive error of the eye. Therefore visual acuity is a subjective measure of refraction and as such was one of the measures of refractive error that was collected in this study.

An instrument that can measure the refractive state of the eye more precisely than a visual acuity test is an optometer. The principle of an optometer relies on placing a lens between the eye and a target and obtaining whether a clear image is formed on

the patient's retina. The positioning of the target when a clear image is observed indicates the refractive error of the eye. An optometer can be either subjective or objective. An objective optometer relies on vergence of light reflected on the subject's retina (light rays shone on a patient's retina will converge when a clear image has been obtained). Another instrument that may be used by an optometrist is a retinoscope. It also provides an objective measure of refraction. A retinoscope relies on the direction of movement of reflected light on a patient's fundus (the interior of the eye) after refraction. If the reflected light appears not to move, no ametropia exists but movement in either direction indicates that light is focused either too far ahead or behind the retina (Millodot and Laby, 2002). A subjective optometer relies on a patient's feedback to determine when a clear image has formed. When an optometrist is undertaking a subjective refraction a process similar to that of an optometer is used. Lenses of different power are placed in front of the eye until vision is achieved.

A subjective refraction can be measured via subjective methods such as visual acuity or the choice of lens that allows clear vision. Furthermore an objective technique, such as retinoscopy can be used by an optometrist to inform upon the nature of the refractive error that may be present in a subject's eye. An optometrist can then use professional judgement to discern the nature of the refractive error. If an accommodative anomaly is suspected (such as accommodative spasm or an excess of accommodation) cycloplegic agents can be administered to estimate the error in refractive measurement produced by abnormal accommodation (accommodation and accommodative anomalies are discussed in detail in Chapter 3). A measure of refraction with a simple objective measure relies totally on optical theory which allows anomalous properties of the refraction of the eye (which are not accounted for by the optical theory) to interfere with measurements. Subjective refractions are more precise than objective measures alone for this reason.

It would be hard to identify a measure of refraction that has been biased by either a systematic or random error based on measurements from one instrument. It is important to estimate the measure with a number of instruments (without a reduction in accuracy) to uncover a possible error due to inadequacies of the instrument of measurement (Kirkup and Frenkel, 2006). An optometrist uses different instruments to estimate refractive error in one examination.

Optometrists often will take more than one measure of refraction with the same instrument to obtain a mean reading. This has the effect of reducing errors due to random uncontrolled sources of environmental variation. An example of some sources of such errors are electrical interference, mechanical vibration or changes in temperature (Kirkup and Frenkel, 2006). Since these errors are hypothesized to be random they will occur equally in either direction (i.e. some add a little to the measure in the plus direction, others in the minus direction) and when summed will cancel out. The most accurate measure available is one based on many measures and averaged because when making an average the random errors will also tend to cancel out (as is hypothesized due to their random nature). Therefore optometrists when taking more than one reading with the same instrument will give an accurate measure of refraction.

The unit of refractive error is a dioptre (D). A dioptre is the reciprocal of an eye's focal length (the distance between an eye and point of focus). An eye that can focus an object 1 metre in the distance has a dioptric power of 1 D (Millodot and Laby, 2002). A highly myopic eye (-6 D) can focus an object 1/6 metres without the need for glasses. Refractive error measured as dioptric power of the eye is the main measure of refraction in this study.

2.1.3 Potential biases

This study collected subjective refractions for study children who attended an ALSPAC clinic at age 15. ALSPAC began in 1991 and involved 85% (approximately 14,000) of babies born in the district of Avon in that year. The number of study children remaining at the year 15 clinic was considerably reduced. The reduction in number of participants from 1991 to present represents withdrawals, the nature of which are unknown (it can be speculated for example, that a number of participants will have moved from the Avon region making attendance at a study clinic, which is located in the city of Bristol, difficult and that a number of individuals have left due to the development of a morbid disease or death). This study does not deal with the effect of withdrawal directly but it is pertinent to note that the subjective refractions were collected for a reduced sample. ALSPAC is a birth cohort and representative of the general population of Britain, although there is some bias towards having a father

in a non-manual occupation and living in owned accommodation (Golding et al., 2001). Subjects are not selected for any particular disease or exposure. It can be hypothesized that withdrawal due to reasons connected with a disease being studied are randomly occurring in the study sample (unlike for example a study of smoking and lung cancer where there are reduced numbers in the smoking group due to drinking related deaths as individuals who smoke tend to drink more heavily).

ALSPAC is concerned with general health and an ALSPAC clinic consists of many different measures of well being. Withdrawal due to discomfort of a particular measure during clinic attendance may have occurred (for example fear of a low score on a psychology measure) but since visits contained many different measures of general health, withdrawal of this nature could be hypothesized to be random.

Each participant that attended the visual examination of a study clinic at age 15 was asked to fill out a form that indicated willingness to participate in the present study. The number of individuals eligible to participate would be close to those attending the ALSPAC clinic. A selection process was not employed and therefore bias due to selection can be hypothesized to be negligible. After collection of the participation forms a number of individuals were excluded due to incomplete information. Incomplete information was due to poor hand writing on the form and loss of information during digitalisation (both relatively small number of instances). It can be hypothesized that ocular health is related to none of these and loss of information due to incomplete forms was random. The possibility that some individuals were more likely to attend the visual examination than others (for example individuals with good vision may have been more likely to skip the vision related part of the general clinic) or some individuals were more likely to return a completed form (for example subjects from families with poor vision) may have lead to a measure of bias. The amount of bias is investigated directly in results.

The aims of data collection were a) collection of age matched vision data for a number of ALSPAC children, b) collection of vision data for nuclear families recruited into ALSPAC. Any bias in the data is relevant only in terms of these two aims. No accurate inference can be made about the prevalence of refractive errors in the cohort from collected data. ALSPAC 'was specifically designed to determine

ways in which an individual's genotype combines with environmental pressures to influence health and development' (Golding et al., 2001). The data collection follows this principle in that its purpose is to support further genetic and environmental investigation of health (in this case myopia). Studies of refractive anomalies that estimate prevalences can be found elsewhere (Williams et al., 2008c) and are made using the larger ALSPAC cohort, not a subset as is used in this study.

2.2 Methods

2.2.1 Data preparation

Figure 2.1 shows a picture of a consent form. Each form had room for details of optometrists of the study young people and their parents or guardians and at least two siblings. On the back of each form, there was a section to obtain written parental consent to allow the measures to be collected from optometrists. A database was created prior to the beginning of this study consisting of copies of consent forms given to participants when they attended an ALSPAC clinic at age 15. These forms were digitalised before the beginning of this study and were recorded in a Microsoft Access database. A summary of the number of participants that had indicated willingness to participate is given in Figure 2.2.

Prior to the beginning of this study the method to transfer information on participants and their optometrists from paper (consent form) to electronic storage (access database) was via scanning forms automatically. The access database contained information on subjects in random order yet there was a need to formally group participants according to optometrist. Each optometrist would have to be contacted individually and only one visit would be made to each optometrist. At this visit, vision measures of ALSPAC participants who were willing to participate would be collected.

Forms were filled out by hand, but scanned in by a computer and it was observed that certain letters were occasionally read with poor accuracy. A common example of this is the entry of a '5' rather than an S. For example in the section of the form 'Name of Optometrist' entries that should read 'Specsavers Opticians' were replaced by '5pecsavers Opticians'. Another example is in a postcode when number '1' is replaced by the letter 'I'. For example in the section of the form 'Address of Optometrist' the Cardiff postcode 'CF10 4BT' for an optometrist located in South Wales may have been recorded as 'CFI0 4BT'.

The effect of these random errors would be to reduce the number of collectable subjects. In turn sample size would be decreased leading to less power when validating non-cycloplegic autorefraction measures and estimation of heritability. The

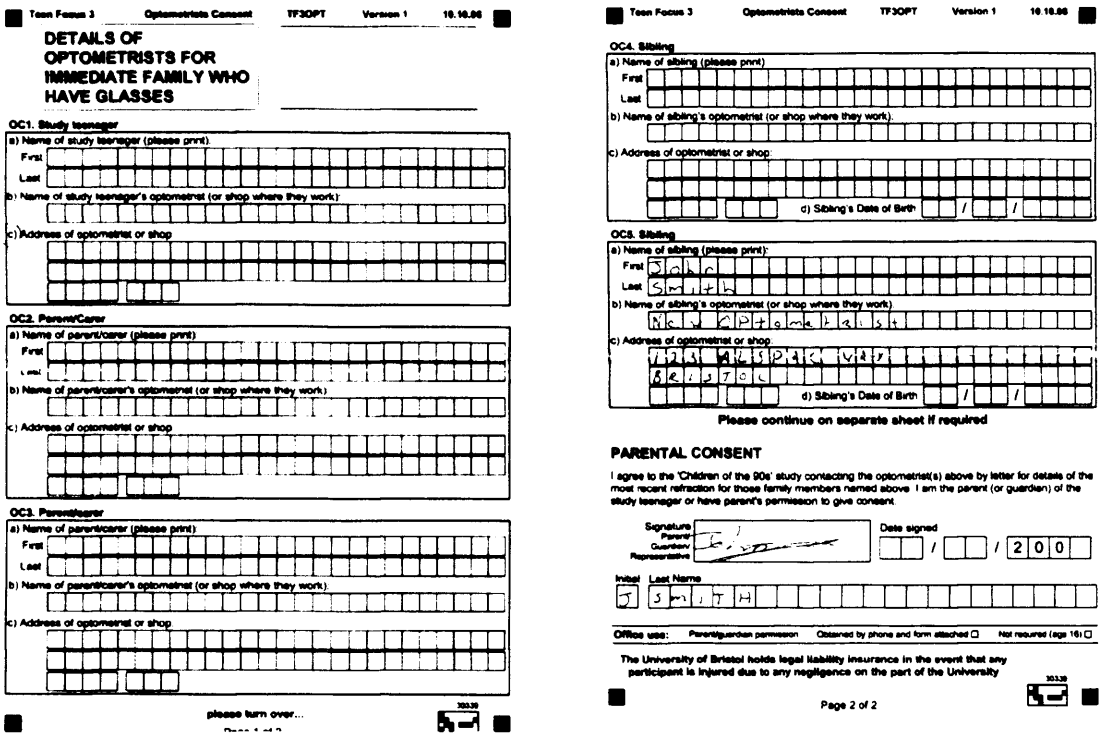


Figure 2.1) Consent form. Front and back of a form used to collect the name and address of the optometrist used by ALSPAC families. Information filled in on the form is handwritten. There is a space (right bottom) for parental consent.

impact of the correctable errors could be reduced. The effect of such errors is to increase difficulty to find all participants from any one optometrist. For example if these errors were ignored, finding those who attend Specsavers Opticians in postcode BS1 1DD would leave out those individuals who were listed under 5pecsavers Opticians in postcode B51 IDD. An advantage of these scanning errors is the mistake is reasonable obvious. '5pecsavers Opticians' indicates 'Specsavers Opticians' while an incorrect postcode such as B51 IDD can be rectified by reference to the address listed and vice versa.

To group individuals according to their optometrist without loss of data, due to correctable random errors accumulated during scanning, the following strategy was employed. The digitalised database was exported to Microsoft Excel (2003, Microsoft Corporation) and an algorithm was created in Visual Basic (version 6.5, Microsoft Corporation) to pick out subjects with the same optometrist. Subjects grouped by optometrist were then allocated one worksheet, giving a workbook of one optometrist per worksheet (approximately 500 worksheets were necessary). The algorithm searched for predefined partial matches of information that could identify groups of individuals by their optometrist. For example, to find individuals from Boots, 1 High Street, Weston Super Mare, the algorithm could search for rows where 'Boot' and 'High' and 'Weston' were present and group all instances together. Then this smaller list of individuals could be checked by eye to make sure each subject attended the correct optometrist. The whole database was treated in this fashion until a small number of subjects were left which the algorithm could not group according to optometrist. These were examined by eye and were in most cases disregarded due to lack of information. The advantages of this strategy were a) all individuals who were willing to participate and filled out a complete form were included b) if errors that occurred during digitalisation had been removed by hand the task would have taken many hours. Furthermore a new source of error may have been introduced if the task had been achieved manually; researcher derived errors due to the repetitive nature of the task. The power of a computer to achieve a laborious, repetitive task, a large number of times with great accuracy seemed more appropriate. After removal of correctable data errors, 7,311 subjects remained. These collectable subjects had also been organised, in the process, by optometrist (of which approximately 550 were listed).

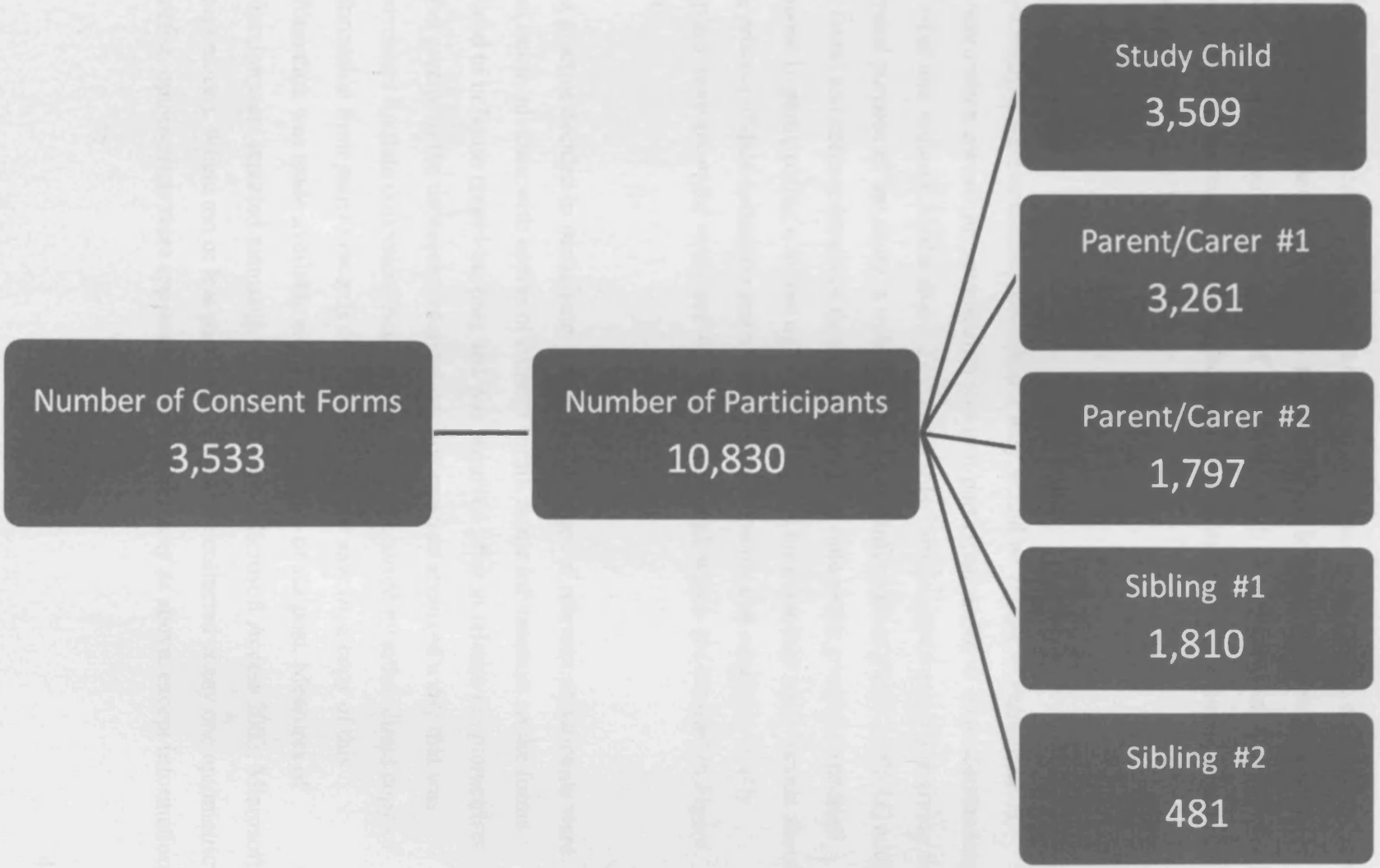


Figure 2.2) Number of eligible participants. Prior to the beginning of this study, approximately four thousand forms had been completed and returned, indicating willingness of subjects and their families to participate. The consent forms contained information for a study child plus their immediate family. Left, the number of each type of participant is displayed.

A second independent source of random error was present, introduced prior to digitalisation during the filling out of forms. Some forms were incomplete. Examples include the name of the participant but no optometrist name or address and in some cases full data was given but the signature for consent was left empty. Unlike errors introduced by digitalisation, where it was possible to infer what information was missing, incompletion of forms led to loss of data. Incomplete forms could be reduced by requesting the data from subjects a second time. However it was not necessary to do this given a reasonable number of individuals with collectable data. A summary of the number of instances of incomplete forms and forms with no consent is given in Figures 2.3a and 2.3b respectively.

2.2.2 Contact

Optometrists were contacted and asked if they would be willing to participate. A concern when contacting optometrists was non-participation due to misunderstanding of what was required by the study. To counter this, initial contact was made giving the general purpose of the study, a website where more information about ALSPAC could be found and contact details of those involved. If an optometric practice expressed interest in participating a follow up letter was sent. This contained more details about the process of data collection and a number of documents that sought to briefly explain more about the study and ALSPAC in general, which are detailed in Figure 2.4.

If a practice decided to participate, the names and ages of relevant participants were sent out in advance with copies of consent forms. Some information on the forms related to in house record keeping and was removed prior to release to optometrists. After receipt of the necessary information, optometrists arranged a day that was convenient for data collection. Acquisition of data occurred by either direct copy of information from patient records or if the optometrist saw fit, a copy of this information was made available either at the practise or via post. Measures of refraction were inputted manually into a database (Microsoft Access 2003, Microsoft Corporation). Where ten or less participants could be collected at any one optometric practise, optometrists were approached in the same way as above, except information

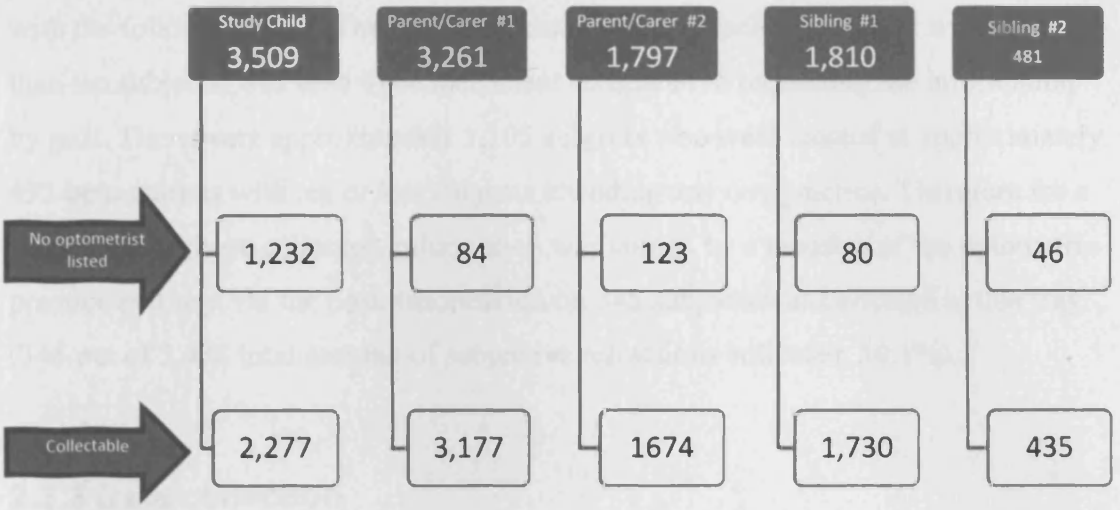


Figure 2.3a) Incomplete data I. Numbers of instances where no optometrist was listed by study participant.

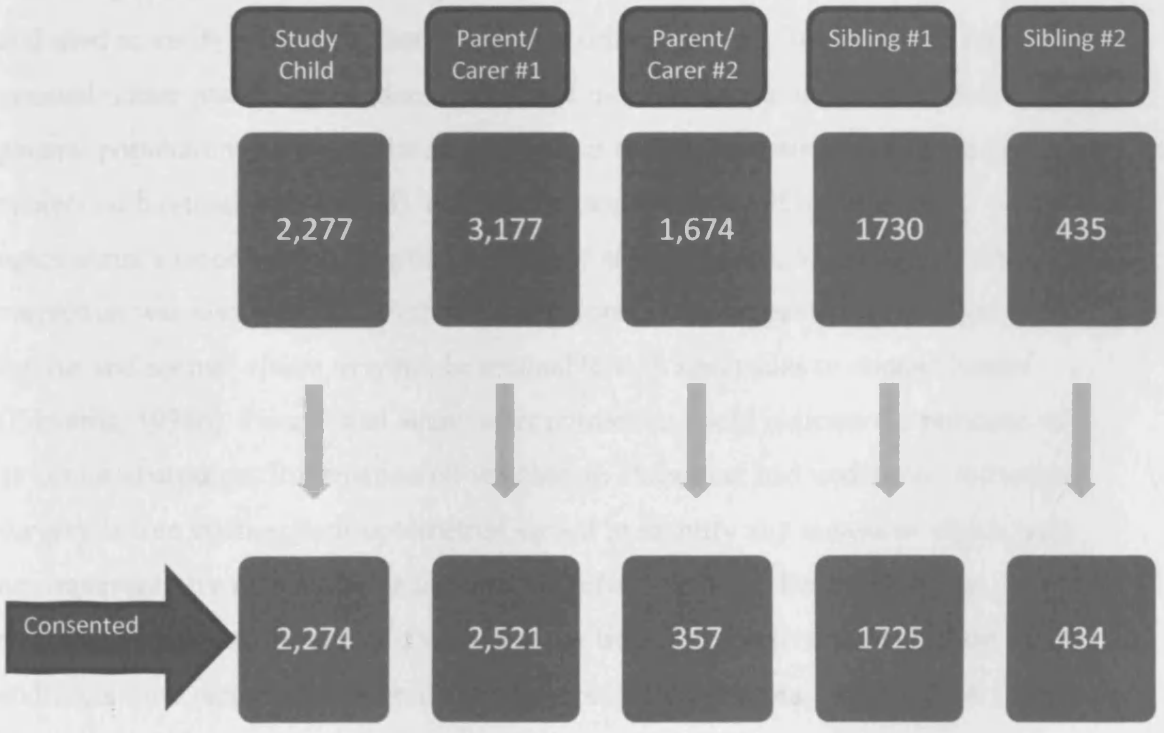


Figure 2.3b) Incomplete data II. Number of participants after those without consent have been identified.

was requested via return of a self addressed stamped envelope which was supplied with the follow up letter. The time and cost of visiting each optometrist with fewer than ten subjects, was seen to be inefficient compared to requesting the information by post. There were approximately 1,105 subjects who were located at approximately 452 optometrists with ten or less subjects attending any one practice. Therefore for a proportion of those collected, information was copied by a member of the optometric practice and sent via the post. Information on 345 subjects was collected in this way (345 out of 3,428 total number of subjective refractions collected, 10.1%).

2.2.3 Data collection

The main measure to be collected (which would be used to form the unit of analysis in the validation study and heritability analysis) was uncorrected refractive error in the form of average spherical equivalent. Therefore information on sphere, cylinder and axis of right and left eyes were of primary importance. Visual acuities were collected and used to verify refraction data. It was also critical that any subjects who had unusual ocular pathology be identified. These individuals are unrepresentative of the general population and would be removed from further analysis (an example is a subject with retinal detachment). A subject's ocular history, if listed in the optometrist's records, was recorded to identify atypical cases. Visual acuity after correction was also recorded. Pathological myopia is accompanied by changes to the fundus and normal vision may not be attainable with spectacles or contact lenses (Edwards, 1998b). Poor visual acuity after correction could indicate the presence of an ocular aberration. Information on whether an individual had undergone refractive surgery before visiting their optometrist served to identify any measures which were not representative of a subject's uncorrected refractive error. Date of test was recorded to estimate the age of a subject at the time of refractive measurement. Near additions were recorded, to identify the presence of presbyopia. Family history was also recorded.

Collected data was entered into an electronic database. A number of constraints were used to minimise random errors accumulating during transfer from the optometric records. These typing errors were hypothesized to be either a) incorrect transfer of many measures for an individual or b) incorrect transfer of values for a particular

a)

Paediatric and Perinatal Epidemiology 2001, 15, 74–87

ALSPAC–The Avon Longitudinal Study of Parents and Children
I. Study methodology

Jean Golding[a], Marcus Pembrey[b], Richard Jones[a,b] and the ALSPAC Study Team[a]
[a]*Institute of Child Health, University of Bristol, Bristol and* [b]*Institute of Child Health, University College, London, UK*

b)

BJO

Prevalence and risk factors for common vision problems in children: data from the ALSPAC study

C Williams, K Northstone, M Howard, I Harvey, R A Harrad and J M Sparrow

c)

Collection of refractive error information for the Children of the Nineties research study

We send you by way of registered post a list of your patients who are participating, along with copies of their consent forms.

"Mail call, from Children of the Nineties"

We ask that you choose a convenient day for a researcher to visit your practice to collect the required information.

On that day, we ask your receptionist to have the patients' record cards pulled out, in preparation for our visit. If you use an electronic patient record system, then we can simply collect a list of each patient's prescription, v.a.'s and date-of-test.

Collection only takes about one minute per patient. A small space is required for data entry.

That was easier than I thought

d)

This is an Example of the Database We are Collecting; to Help Establish the Causes of Refractive Error

The patients current prescription will be linked to previously collected data on the patients health. We will use this to identify the causes of refractive error.

We need this to establish the correct age of the participant when refractive state was measured.

Measures of Visual Acuity will help determine risk factors for ocular health in our broad based study population located in the environs of Bristol

Ocular History can help identify a persons refractive error in greater detail. For example, if a person has undergone refractive surgery, this will inform on their current prescription.

A participants medical history and that of their family's is important to general eye health. This information can identify those most at risk of myopia and other disorders of the eye.

Figure 2.4) Data pack. Information sent out after optometrists expressed an interest in participating in the study. Each document sought to briefly explain an element of the study; a) published papers about ALSPAC, b) a copy of a blank consent form, c) a cartoon depicting what was expected of the optometrist if they were to agree to full participation, d) a diagram explaining how the information would be useful.

measurement. Incorrect transfer of *many* measures was made minimal by allocating one electronic page per individual. Each record was typed into one page, necessitating creation of a new electronic record for every paper record. If a subject's data was entered in one row on one page, it would be easier to lose track of the current row number and data could be misappropriated. Incorrect transfer of *particular* measurements could occur during long sessions of data entry. For example a measure of spherical power in dioptres could be typed -12.5 when the correct value would be -1.25. It was possible to restrict values accepted by the database within a normal range and where possible unusual values were met with a warning message (Figure 2.5). This logic of this was extended to all data types by ensuring that only reasonable values were accepted without warning.

2.2.4 Quality control

A number of checks on the collected data were undertaken to improve data integrity. Duplicates were identified (n = 1). A number of cases (43) were collected without consent. This happened predominately in the first data collection visit (22 out of the 43). A number of entries were identified as likely typing errors and are detailed. Less than 10 astigmatism measures were listed with no sphere. This is most likely a case of the optometrist leaving the sphere blank to indicate zero dioptres or the sphere reading was entered in place of astigmatism. These subjects were removed. One data error was identified by the difference between right and left sphere. In this case the original optometrist record was available and the error (a missing minus) was corrected. One visual acuity was entered as 66/6. This was probably mistyped as 66/6 rather than 6/6. The data was removed. Two near additions were typed 0.125 and 0.175 instead of 1.25 and 1.75, these have been indicated. The range for date of clinic visit was within acceptable limits from 2006 to 2008.

Less than 10 instances of study participants indicated with a questionable date of birth (i.e. A study Parent/Carer born in 1992) were identified. Most likely these individuals had been indicated as Parent/Carer or study child erroneously and were removed. Date of eye exam showed two instances of likely data errors with both indicating dates of exam in either April or May 2010, after data collection had been completed. The optometrist records were available and checked. One was a typing error and was

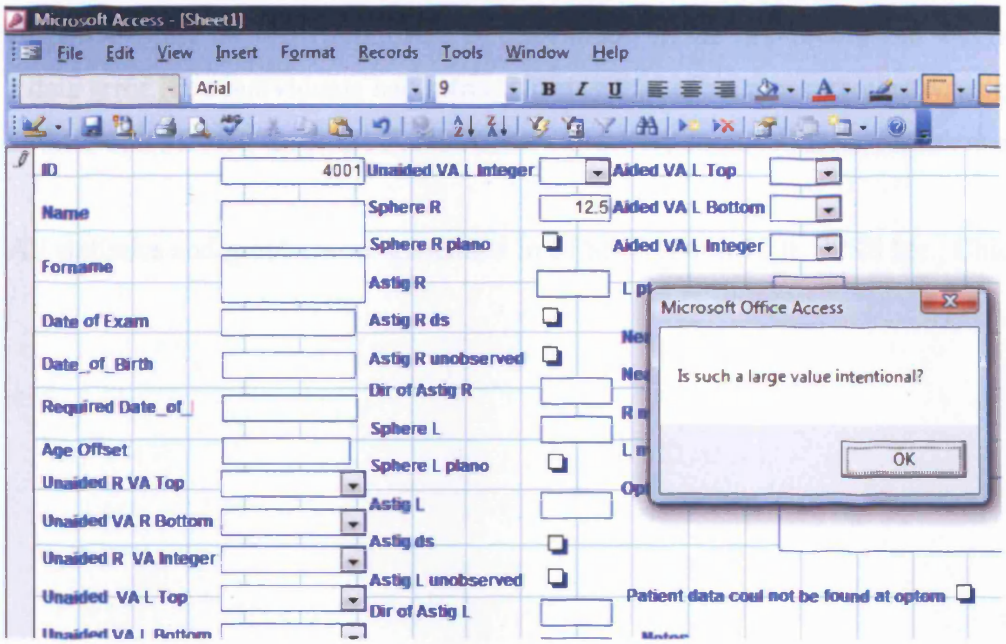


Figure 2.5) Data constraints. An example of a warning message issued in the electronic database (Microsoft Access) for unusual values of refraction.

corrected while the other replicated an error made on the first record and was noted as a data error. Ten individuals had refractive surgery before measurement was obtained at their optometrist; these are Parent/Carers and were removed from further analysis.

All statistics and graphs were generated in SPSS (version 16.0, SPSS Inc., Chicago).

2.3 Results

A summary of overall response rate is given in Table 2.1. All optometrists were approached but a 47% positive response indicates that a significant number declined participation. 3,091 average spherical equivalents were collected. This is slightly less than the number of collected vision data (3,330) due in part to some participants having had tests of visual acuity but no measure of refraction. Subjects with a valid measure of refraction were made up of three distinct study participants, the study child, born in 1991 or 1992, parents or guardians and siblings (Table 2.2). 375 out of 1,016 study children had a date of test that was within six months of their visit to the ALSPAC clinic at age 15. The average spherical equivalent of these subjects was matched to non-cycloplegic measures taken at the ALSPAC clinic (see Chapter 3).

The average spherical equivalent for children and their siblings were similar ($t = 0.2$, mean difference 0.01, $P = 0.88$) but parents showed a more negative value than siblings ($t = 6.3$, mean difference 0.68, $P < 0.001$) and study children ($t = 6.9$, mean difference 0.67, $P < 0.001$). The percentage myopia for parents, young people and siblings was 46%, 33% and 35% respectively. The number of myopes in the parental group was higher than either study young person ($\chi^2 = 39.1$, $df = 1$, $P < 0.001$) or siblings ($\chi^2 = 27.5$, $df = 1$, $P < 0.001$) but did not differ between study children and siblings ($\chi^2 = 0.64$, $df = 1$, $P = 0.42$). Table 2.3 gives the distribution of refractive states for each of the study participants. The majority of study children (58%) and siblings (50%) display emmetropia while the main refractive state in the parental group was myopia (46%). Across groups there were decreasing amounts of refractive error as severity increased. Figure 2.6 gives a histogram of the average spherical equivalent for each of the study participants, each shows reasonable symmetry and high kurtosis which is common for average spherical equivalent.

The group means of average spherical equivalent (as measured by non-cycloplegic autorefraction) for young people collected with a subjective measure within six months of non-cycloplegic autorefraction ($n = 375$) versus the larger sample from which they were drawn (those who attended an ALSPAC clinic at age 15, $n = 4,987$) was compared. The mean (standard deviation) of non-cycloplegic autorefraction for

Number of collectable participants	7311
Number collected	3428
Response rate	47%
Collected after data checks	3330
Number of refractions collected	3091

Table 2.1) A summary of data collection.

	Study Child	Parent/Carer	Sibling
Number of refractions	1016	1168	907
Mean (AveSph)	-0.41	-1.08	-0.40
Standard deviation	1.83	2.67	2.30
Refractions within six +/- 6 months of clinic	375	-	-

Table 2.2) Number of refractions by study participant. Mean (AveSph) refers to the group mean for average spherical equivalent.

	Emmetropia	Mild myopia	Moderate myopia	High myopia	Mild hyperopia	Moderate hyperopia	High hyperopia	Anisometropia	Total
Study Child	585 (0.58)	239 (0.24)	81 (0.08)	12 (0.01)	40 (0.04)	29 (0.03)	12 (0.01)	18 (0.02)	1016 (1)
Parent/Carer	448 (0.38)	311 (0.27)	163 (0.14)	61 (0.05)	89 (0.07)	46 (0.04)	11 (0.01)	39 (0.03)	1168 (1)
Sibling	473 (0.5)	216 (0.25)	76 (0.09)	20 (0.03)	38 (0.04)	42 (0.05)	19 (0.02)	23 (0.03)	907 (1)
Total	1506 (0.49)	766 (0.25)	320 (0.1)	93 (0.03)	167 (0.05)	117 (0.04)	42 (0.01)	80 (0.03)	3091 (1)

Table 2.3) Distribution of refractive errors. Categories were defined as follows; emmetropia: 0.5 to -0.5 D, mild myopia: -0.5 to -3 D, moderate myopia: -3 to -6 D, high myopia: less than -6 D, mild hyperopia: 1 to 2.25 D, moderate myopia: 2.25 to 5 D, high hyperopia: greater than 5, anisometropia: > 2 D absolute difference right and left spherical equivalent.

those collected was -0.58 (1.7) versus -0.39 (1.3) for those who attended the ALSPAC clinic (mean difference = 0.19, t =2.0, $P < 0.001$).

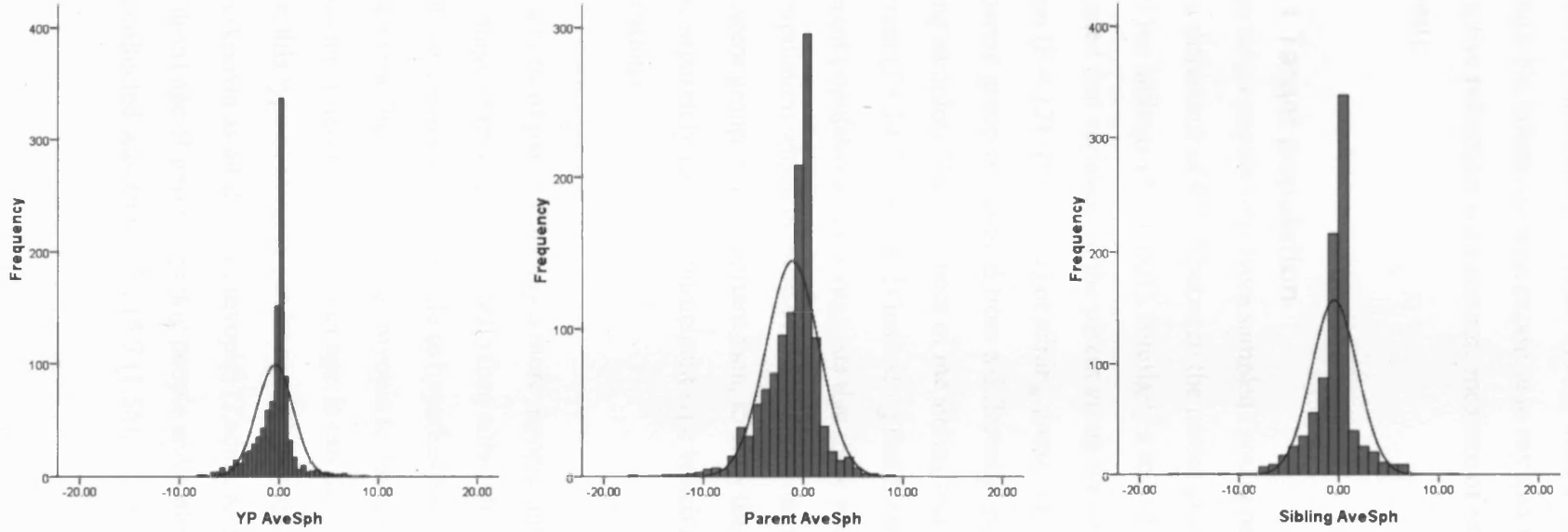


Figure 2.6) Histograms of average spherical equivalent. Subjective refractions are shown for each study participant. A black line denotes a normal curve. AveSph (average spherical equivalent), YP (young person/study child).

2.4 Discussion

The aims of this data collection exercise were largely met, a) collection of age matched subjective refractions for young people who visited an ALSPAC clinic at age fifteen b) collection of subjective refractions of young people and their parents and siblings. No information on exposure to myopia risk factors was collected; however subjective refraction is an accurate measure of refractive error (the phenotype of interest).

2.4.1 Target population

Three target populations were sampled (young person, parent/carer or sibling). There was a difference of 0.7 D between the parent group and either young people ($P < 0.001$) or siblings ($P < 0.001$). Similarly a test of equal variances (Levene's test) indicated that variation in the parent group was not at similar levels in the young person ($F = 121$, $P < 0.001$) or sibling groups ($F = 39$, $P < 0.001$). This suggests that the parent group is sampled from a different population than the young person or sibling samples. The variances of the sibling and young person groups were also different ($F= 14$, $P = 0.0002$) indicating that these two groups may be drawn from different populations. This suggests that there was a certain amount of stratification in the population which is more evident when comparing the subjective refractions for the parent group. It is important then, to view the means and percentages of refractive errors separately for each participant type which represent possibly distinct populations.

Refractions of parents display a more myopic mean ($P < 0.001$) and a higher percentage of myopia ($P < 0.001$) than either the subjective refractions of young people or siblings. It is possible to hypothesize that the increase in amount of myopia is due to the increased time for myopia to develop. Myopia can develop at any age but occurs most frequently at school age. It can develop before school also but this is rarer (often this type of myopia can be severe). Myopia can also develop after the teenage years (known as adult onset myopia) (Zadnik and Mutti, 2006). The mean (standard deviation) age of parents, young people and siblings for whom refractive data had been collected was 46 (4.73), 15.7 (1.56), 15.7 (3.87) respectively. Parents had on

average thrice as much time to allow for myopia to develop. Also myopia that develops in early adulthood tends to have a slower progression rate than myopia that develops earlier (Zadnik and Mutti, 2006). The percentage refractive errors in each study group seem similar to each other apart from moderate myopia where the parent/carer group displays a percentage of 14% compared to 8% and 9% for young people and siblings. This may be due to the presence of adult onset myopia in the parent/carer group which is more absent among young people and siblings. It is also possible that deviations due to sampling are responsible.

Increased amounts of time for myopia to develop would also be expected to lead to more variability in the subjective refractions of older subjects. This is borne out by a large standard deviation for subjective refractions of parents (2.67) compared to refractions of young people (1.83) and siblings (2.3) and by the low P values observed for the test of equal variance ($P < 0.001$ in both cases). There also is an increased amount of variation when comparing the refractions of siblings and young people, which is supported by a significant difference in variation by Levene's test of equal variances ($P < 0.01$). This difference may also be explainable by differences in age of the two groups, young people and siblings. Although each group displays a similar mean age (15 years) the standard deviation for the age of the sibling group is larger. This is reflected in the maximum age at test observed for each group (18 for young persons and 31 for siblings).

ALSPAC has a defined study area (Avon), which has a population of close to one million and a major urban area (Bristol). The majority (88%) of subjects in the current study attended an optometric practise with a postcode beginning with BS (greater Bristol including Bristol city centre). 86% of participants in the larger cohort indicated an optometrist with a postcode beginning with BS. Although a significant difference in the location (with a BS postcode versus outside) of practises visited compared to all eligible practises was observed (χ^2, $P = 0.0004$), similarity between percentage participants attending optometric practises in Bristol for the current study and the larger cohort was evident (86% of the larger cohort compared to 88% of the current study). Therefore it is suggested that data collected was clustered in a similar way to the larger cohort of those who attended the vision examination at age 15. A small

number of practises were visited outside Bristol (12%) and may represent subjects that have moved from the area (outward migration). It is concluded that the geographical dimensions of the study reflect the geographic area of the ALSPAC study.

2.4.2 Sampling

Cohort studies will often protect against bias, by sampling large numbers of observations randomly from the target population (Woodward, 2005) where most individuals will have an equal chance of selection. The ALSPAC cohort sampled 85% of individuals born in between 1991-2 in the Avon region (the eligible population) and is representative of the UK as a whole (apart from ALSPAC subjects being less likely to rent accommodation or have a father with a manual occupation). Of the approximately 14,000 pregnancies initially enrolled, close to 8,000 children attended clinics at age 7. Approximately 5,000 young people attended the clinic at age 15. Of these subjects approximately 3,500 filled out a form indicating willingness to participate in the current study. Approximately 2,200 subjects were eligible to take part after forms with incomplete data were identified. Therefore there is a sampling issue.

The ALSPAC cohort was designed with random sampling. The current study was not designed to involve random sampling specifically. However as mentioned previously, a visit to ALSPAC clinics entails other tasks that are non-visual which would tend to help reduce withdrawals that were related to vision examination specifically and randomise missing data. It is possible to hypothesize that there may be some bias in the decision to fill out and return a complete consent form. Subjects with visual problems may be more likely to find the time to return the necessary information. This would be reflected in the mean difference in measures taken at the ALSPAC clinic between those who filled out a completed consent form and those who did not. However it is more valuable to know if a difference exists for those collected compared to all subjects who attended the clinic at 15 years as those collected form part of a later analysis (the validation of non-cycloplegic autorefraction). A comparison of non-cycloplegic autorefraction of those collected (375) versus those not collected (4,987) with the same age, showed a mean difference of 0.19 D ($P <$

0.001). Since this difference is statistically significant it is evident that some non-random sampling has occurred during collection. However a mean difference of 0.19 D is clinically insignificant and poses no major obstacles for successful completion of the aims of the study. Validation of non-cycloplegic autorefraction would be biased by a large mean difference because it could be argued that validation would be achieved for only a subset of individuals of the larger sample. For example a mean difference of 1-2 D more hyperopia may suggest that a proportion of subjects with high myopia are in reduced numbers compared to other refractive errors in the collected data. Inference to the larger sample would then be more difficult. Since the mean difference is small it suggests that the distribution of refractive errors is similar between those with paired data (both subjective refractions and non-cycloplegic autorefraction) and the larger cohort. Therefore inferences from the validation study on the large cohort will be more accurate.

This study sought to collect data for a subset of subjects who attended an ALSPAC clinic at age 15. Therefore it is a data collection exercise. Similar to other epidemiological data collection exercises, forms (self reported) were administered to collect the data. However it differs from other data collection in that no exposure measures were being recorded. When an exposure to a possible risk factor is measured by self reporting, bias may be introduced by participants feeling embarrassed at repeating the true answer or due to a considerable amount of time having passed between exposure and administration of the questionnaire. In the current study very little of such bias could have accumulated because the forms sought objective information on the name and address of the participants' optometrist. The other information that was sought from participants was consent from a parent or guardian to participate in the study. It is possible that a small number of individuals did not want to participate because they did not want to share personal information held with their optometrist but it is unlikely that choosing not to participate for this reason was related to ocular health.

2.4.3 Summary of measures collected and their accuracy

There is evidence that non-cycloplegic autorefractive meausures of children are less precise when hyperopes are considered only and more precise when myopes are considered (Zhao et al., 2004) which may be related to a reduced accommodative response observed in some myopes (Rosenfield, 1998). It is therefore important to note that reasonable numbers of each refractive state were collected for study children (Table 2.3) a subset of which would be drawn for validation against non-cycloplegic autorefraction. It is conceivable that if a majority of samples collected during data collection were myopes the amount of bias observed in the validation study would be affected. However since reasonable numbers of each refractive state were collected any potential bias due to uneven numbers of one refractive state could be avoided.

Similar numbers of refractions were collected for each study participant (Table 2.2) with 45% of young people (1,016/2,274), 41% of parents (1,168/2,878) and 42% of eligible siblings (907/2,159), indicating no large excess of one group. Therefore comparison between pairs of study participants (such as mother-offspring pairs) in a heritability study would be amenable with close to equal numbers of participants collected in each group.

The ratio of refractions of young people collected to refractions of siblings collected (1016:907 or 1.12) indicated that for every refraction of a young person collected one refraction of a sibling was also collected. There was also an almost one to one ratio of collected refractions of young people to parent/carers (1016:1168 or 0.87). To carry out a heritability study at least one family member other than the study young person would be necessary. Classical heritability analysis requires at least one parent and offspring or a pair of siblings. In the sample collected, both designs are feasible given the number of each type of study participant collected. However the one to one ratios of parents to young people and siblings to young people do not necessarily mean that for each refraction of a young person collected, a refraction of *their* parent and a refraction of one of *their* siblings were collected. Forms allowed each participant to indicate their optometrist and it was the case that study participants attended *individual* optometrists as well as participants attending a *family* optometrist. It is possible that refractive data for a young person was held at one optometrist but refractive data for the parents or siblings were held at a different optometrist. If either

of these optometrists declined participation the young person or one of their immediate family members would be excluded from the analysis. Due to these considerations approximately 600 subjective refractions from parent young person pairs and 600 subjective refractions from young person sibling pairs were collected.

An implication of a lower average spherical equivalent observed in the parent/group (close to -1 D indicating on average 0.5 D more myopia than would be observed for a population that was on average emmetropic) is that the parents of young people sampled for heritability analysis display slightly more myopia. However the mean amount of myopia is small (close to 1 D) and allows the results of a heritability analysis in those with subjective refractions to give an estimate of the heritability in the larger ALSPAC cohort. Furthermore since the primary focus is on myopia, the tendency of slightly more myopia in parents used in a heritability analysis will increase confidence that such an analysis will provide an estimate of whether myopia is heritable in the larger ALSPAC cohort.

Chapter 3

Validation of Non-cycloplegic Autorefraction

3.1 Introduction

3.1.1 Validation

The use of non-cycloplegic autorefraction has been largely found to be biased towards overestimation of myopic refractions. Subjective refractions were chosen to validate non-cycloplegic autorefraction because they are generally thought to be more free from error or uncertainty in measuring refractive error. When measuring ametropia via subjective refraction an optometrist takes more than one measure of a person's refractive error using various instruments and uses patient feedback to estimate refractive measurement. Therefore a number of pieces of information will contribute to the final measure leading it to be more accurate. There would still be random errors in subjective refractions, from uncontrollable and small changes in the environment (Kirkup and Frenkel, 2006) but here these are not quantified directly and are considered negligible. Furthermore an optometrist takes repeated measures of refraction with one instrument and can use a mean value of measurements which tends to cancel out these small environmental changes.

The error of primary concern in non-cycloplegic autorefraction is a *systematic* bias. This is a non-random error that results in measurements being inaccurate by an amount that is constant. Unlike random errors, systematic errors are not improved by taking repeated measures (Kirkup and Frenkel, 2006). A systematic error can be *additive* in that when a measurement is made with a particular instrument, an amount is either added or subtracted from the true measure during use without the knowledge of the technician. For example refractive error measured by non-cycloplegic autorefraction is known to subtract an amount from a subject's true refractive error for most cases and this subtraction is relatively constant. This type of systematic error is termed an offset (i.e. non-cycloplegic autorefraction measures display a negative offset of some degree). Another type of systematic error is *multiplicative.* In this case there is a constant error only over a particular range of measurements (Kirkup and Frenkel, 2006). The amount of error varies depending on the size of measurement. For

example there is evidence that non-cycloplegic autorefraction displays a large offset when measuring above 0.5 – 1 D (hyperopic refractive error) (Krantz et al., 2010) and a lesser offset when measuring below -0.5 D (myopic refractive error) (Zhao et al., 2004). Calibration of an instrument against a standard (an instrument of higher accuracy) can reveal a systematic error. In this study non-cycloplegic autorefraction measures are calibrated against subjective refractions. Calibration will provide an estimate of the degree of systematic error and its variability. After calibration the systematic error can be removed by applying an accurate correction. However the variability associated with the systematic error may still be present (Kirkup and Frenkel, 2006).

A Bland Altman plot (Bland and Altman, 2010) indicates whether agreement between paired measures is constant or changing according to an accurate measure. The difference between measures for the same individual (when differences are large agreement is observed to be questionable) are plotted against the mean of the two measures (which can be considered an accurate measure). If the magnitude of the difference changes according to the mean it is suggested that one measure is a constant multiple of the other measure (Woodward, 2005). The error can be then thought to be multiplicative. The mean difference between measures and the 95% confidence intervals may then apply only over a certain range of measurements. If the magnitude of the difference shows no discernible trend as the mean measure changes, it suggests that one measure is different from another by a constant amount. The error can be thought to be additive. The mean difference between measures and 95% confidence intervals may be applied to the whole range of measurements investigated.

For the purpose of carrying out epidemiological investigation of myopia in analyses of the ALSPAC cohort, it is necessary to classify individuals into a binary disease status (myopia/not myopia). True disease status cannot always be obtained as the procedure for diagnosis may not be 100% reliable (Woodward, 2005). This is true when using non-cycloplegic autorefraction to infer the presence of myopia. Therefore it is important to quantify how reliable a diagnosis may be with a particular test. This can be achieved by calibration of the test by a standard (in this case, subjective refractions). The value of using a non-standard test is that it may be quicker and more convenient. To quantify the non-standard test two types of errors are important. The

test could wrongly decide that a subject with the outcome (a myope) does not have it or the test could wrongly decide that a subject without the outcome (a non-myope) does have it. This is more often expressed in a complementary sense, the probability that the right decision is made when the subject has the disease is termed the sensitivity of the test, the probability the right decision is made when a subject does not have the disease is called the specificity of the test.

	True disease status		
Test result	Positive	Negative	Total
Positive	*a*	*b*	*a + b*
Negative	*c*	*d*	*c + d*
Total	*a + c*	*b + d*	*n*

Sensitivity = a/ (a + c)

Specificity = d/ (b + d)

Table 3.1) A diagnostic test. Assessing the results of a diagnostic test (modified from (Woodward, 2005))

Sometimes a diagnostic test does not indicate disease status but measures a trait which is used to infer presence of the disease. It is possible to test the reliability of the test to make the correct diagnosis at different cut points given by the measure of the trait. It may be that sensitivity and specificity differ depending on the severity of the disease. The reliability of non-cycloplegic autorefraction to infer myopia was tested against a standard diagnosis of myopia that is displaying less that -0.5 D in a subjective refraction measure.

3.1.2 Autorefraction and accommodation

An autorefractor is an objective instrument that can measure the refractive power of the eye. Different autorefractors measure the refractive error of a patient by employing a number of optometric principles (for review see (Campbell, Benjamin and Howland, 2006)). Important to this study is that an objective autorefraction measure of refractive power does not use a clinician's professional judgement or a patient's feedback to obtain the measure. The process is automated through the use of

an optical technique and computer power. Autorefractors have a high speed of measurement (McBrien and Millodot, 1985) and allow measurement of refractive error by a non-optometrist, trained in the use of autorefraction. The use of autorefraction is common (Campbell et al., 2006).

The need for cycloplegia (paralysis of the ciliary muscles) during autorefraction arises due to the natural process of accommodation of the eye. Accommodation is a modification of the refractive power of the eye (Millodot and Laby, 2002). A neural signal in the innervations of the ciliary muscle causes contraction, which in turn allows the lens to become more convex, which will change the focus of parallel rays of light entering the eye (Millodot and Laby, 2002). Accommodation allows an image of an object of regard to be obtained and held in focus on the retina (Ciuffreda, 2006). Autorefractors measure the refractive power of the eye (the ability to focus a clear image on the retina) taking into account its refractive power *when accommodation is relaxed.* When an autorefractor is used to make a measure of refraction, the subject has placed their head on a rest and is observing a target, even when the target image is blurred (Campbell et al., 2006). If the accommodative response (which partly determines the refractive power of the eye when an image is either blurred or in focus) behaves inconsistently with the target of fixation then the autorefractor measures only part of the refractive state of the eye being complicated by the effects of accommodation.

An example of accommodation affecting the accuracy of a measurement of refractive error is given by accommodative spasm. Accommodative excess is a term describing a situation where a subject over-accommodates in response to a visual stimulus either exerting more accommodation than is necessary or by a failure to relax accommodation (Millodot and Laby, 2002). It can occur for a number of reasons including too much nearwork, latent hyperopia and emotional distress among other reasons (Millodot and Laby, 2002). Accommodative spasm is a type of accommodative excess, in this case due to involuntary stimulation of the ciliary muscle. Office workers, school goers and other subjects, who have recently spent a prolonged time reading before measurement, may display a small myopic shift due to an accommodative anomaly similar to accommodative spasm. This will subtract 0.25 to 1 dioptre from their refractive error as measured by an autorefractor (Campbell et

al., 2006). Other accommodative anomalies will similarly add uncertainty to a measure of refraction by an autorefractor. A clinician carrying out a subjective examination will be able to investigate such anomalies and identify pseudomyopia, latent hyperopia and so on by changing the conditions of measurement (using different instruments or methods that can reduce the influence of accommodation).

Another source of error when measuring refraction using an autorefractor is the amplitude of accommodation. The amplitude of accommodation is defined as the maximum amount the eye can accommodate (Millodot and Laby, 2002). The near point (a point in space that can be observed when accommodation is at a maximum) depends on a subject's amplitude of accommodation. The difference (in dioptres) between the near point and a point in space which is in focus when accommodation is relaxed (far point) indicates a subject's amplitude of accommodation. The amplitude of accommodation is larger for younger individuals (at age 10 it is approximately 14 D, at age 60 it is less than 2 D (Ciuffreda, 2006)). When refractive error is measured with an autorefractor the accommodative response should ideally be relatively stable. In subjects with a large amplitude of accommodation, stability of accommodation is harder to achieve. Also when an autorefractor is used it is important that the subject relaxes and attends to the target of fixation. Older patients would tend to display more motivation, attention and understanding of the task and the importance of looking at the target even when it is blurred.

3.1.3 Cycloplegia

Relaxation of accommodation can be achieved via paralysis of the ciliary muscles. This is known as cycloplegia (Millodot and Laby, 2002). Cycloplegia can be induced by antimuscarinic agents known as cycloplegics. Cycloplegics block the action of the neurotransmitter acetylcholine at the iris sphincter muscle and ciliary body and lead to paralysis of accommodation (Bartlett, Jaanus and Blaho, 2001). Cycloplegia is recommended if an accommodative abnormality is suspected to be present during refractive error measurement. Furthermore it has been noted that in younger age groups, the amount of myopia present in cycloplegic versus non-cycloplegic measures of refractive error can be significant, a trend that decreases with age (Grosvenor, 2002). In measuring the refractive error of young children it is recommended that

either cycloplegia is used to relax accommodation or a full subjective examination is undertaken. Similarly it is also recommended that cycloplegia is used to measure the refractive error of children from infancy to 48 months (Bartlett et al., 2001) as subjective refraction is not possible given the young age of the patient. The advantage of a subjective examination is that non-cycloplegic measures using an objective instrument can be compared to visual acuity measures and a subject's responses and if the clinician suspects an accommodative anomaly, the need for cycloplegia to obtain a more precise measure of refraction can be assessed.

In studies which estimate the prevalence of refractive error, cycloplegia is generally used when subjects are of a young age. A common method to ensure accuracy of prevalence measures of refractive errors in children across ethnic and geographic groups has been published (Negrel et al., 2000). The method includes cycloplegic retinoscopy, autorefraction and if uncorrected visual acuity is less than 0.625, subjective refraction. In this case estimation of refraction can be based on two to three methods, with any incongruent readings due to ocular abnormalities being identified via comparison between techniques. A number of large studies on the prevalence of refractive error in children utilise the method (Maul et al., 2000; Pokharel et al., 2000; Zhao et al., 2004).

As age increases a subject without any accommodative anomalies, may become more easily tested and cycloplegia may be unnecessary. However the age of a subject when cycloplegia is no longer needed is debated (Bartlett et al., 2001). A number of comparisons between autorefraction without cycloplegia and other measures of refraction have been published. In an early study it was found that autorefraction readings in a group of young adults (18-25) were more negative than subjective refractions. The statistical significance of the results was not approached but the authors conclude that the difference was clinically significant (McBrien and Millodot, 1985). In another study it was found that autorefraction readings were more negative but only in younger subjects. The offset was less pronounced or absent in subjects older than 40 years of age (total age range for the study was 6-75 years). Furthermore subjects with higher refractive errors showed less negative offset. The authors concluded that accommodation was the critical factor leading to the bias in non-cycloplegic autorefraction measurements (Ghose, Nayak and Singh, 1986) due to the

ability of those displaying a large offset to accommodate more readily than subjects who showed lesser offset. In related work, the authors examined the effect of cycloplegia on the accuracy of autorefraction observed in young people (8 to 25 years of age) with low or absent refractive errors. They observed that the offset was neutralised with addition of a cycloplegic agent (Nayak, Ghose and Singh, 1987).

The need for cycloplegic autorefraction in adults has been addressed formally in two modern studies. Non-cycloplegic autorefraction was found to give more negative measures of refraction than autorefraction with cycloplegia (Jorge et al., 2005), (sample size 199 and age range 18-34). In a large study (approximately 3,000 individuals) with a wider age range (22-84) the effect of cycloplegia on autorefraction measures was also examined. Autorefraction without cycloplegia was found to display a small negative offset (0.29 D) that decreased with age (to 0.15 D over 50 years). The authors concluded that the overall difference between autorefraction measures with and without cycloplegia was clinically insignificant (Krantz et al., 2010). A number of more recent studies have investigated the effect of cycloplegia on autorefraction measurements of refractive error in those in younger age groups. A negative offset has been found for non-cycloplegic autorefraction versus cycloplegic autorefraction up to the age of 18 (Zhao et al., 2004), at the ages of 6 and 12 (Fotedar et al., 2007), in subjects age 3 to 15 (Rotsos et al., 2009) and from ages 6 to 13 (Funarunart et al., 2009). From these studies, two points may be of interest. Firstly it is possible to conclude that the studies support the general agreement among practitioners that autorefraction with cycloplegia is necessary for those in younger age groups but a need for cycloplegia in adult subjects diminishes (Bartlett et al., 2001). Secondly when non-cycloplegic autorefraction leads to a negative offset the addition of cycloplegia is found to diminish the effect.

A systematic error observed for non-cycloplegic autorefraction is present when using a range of autorefractors. Non-cycloplegic autorefractions have been found to be more negative compared to cycloplegic autorefraction for the RMA-3000 autorefractometer (Topcon, Tokyo, Japan) (Rotsos et al., 2009), Retinomax K-Plus autorefractor (Nikon, Tokyo, Japan) (Zhao et al., 2004), Nikon autorefractor (NRK-8000) (Funarunart et al., 2009) and Canon-RKF1 autorefractor (Tokyo, Japan) (Fotedar et al., 2007).

Studies designed to make inferences about refractive errors try to balance compliance, accuracy and cost when choosing the type of measure of refraction to be used. Cycloplegic drugs have a number of disadvantages that can limit their usefulness. Modern cycloplegics (cyclopentolate, tropicamide) have a duration of effect of 4 hours or more (Bartlett et al., 2001). When other measures apart from refractions are to be collected on one day, a duration of large magnitude may be inconvenient. In the ALSPAC cohort, a visit to a clinic involved various measures of general health and cycloplegia may have resulted in the inability of subjects to participate in other tests. Paralysis of accommodation may leave a subject with blurry vision even after leaving the clinic or test room. This can inhibit normal routine. Cycloplegics can also increase sensitivity to bright light (Bartlett et al., 2001). Both of these side effects may convince subjects that the benefit of participating is not enough compared to uncomfortable side effects which may lead to withdrawal from the study. Rarely cycloplegics agents can elicit a severe adverse response. This places an increased "duty of care'' on clinicians. Despite these drawbacks, cycloplegia remains the norm in cohort studies of the refractive error in children (Mutti et al., 2002; Saw et al., 2007; Giordano et al., 2009). There are exceptions (Deng, Gwiazda and Thorn, 2010).

3.2 Methods

A description of data collection methods and results is given in Chapter 2. The sample consisted of 375 individuals aged 15. Non-cycloplegic autorefraction measures were taken during a visit to an ALSPAC clinic. Subjective refraction measures were obtained from optometrists. This resulted in measures of refraction within 6 months of the clinic visit. For both non-cycloplegic autorefraction and subjective refraction, the refractive power of the sphere and cylinder were obtained. Spherical equivalents were estimated by sphere + ½ cylinder for the right and left eyes. Average spherical equivalent was used in further analysis, given by the mean of the right and left spherical equivalents. Myopia was defined as <=- 0.5 D, emmetropia as > -0.5 to < 1 D, and hyperopia as >-1 D by subjective refraction. Statistics and graphs were generated in SPSS (version 16.0, SPSS Inc., Chicago). A Wilcoxon and Sign test were used to measure the difference between non-cycloplegic autorefraction and subjective refractions. A Bland Altman plot was employed to examine the nature of any offset between the two measures. Sensitivity and specificity was generated via logistic regression and Receiver Operating Curve (ROC) analysis.

3.2.1 Error checking

Before comparing non-cycloplegic autorefraction and subjective refractive measures, both types of measures were checked to identify errors accrued during accumulation of the data (termed data errors). Errors due to mislabelling or mistyping during data entry may add uncertainty to calibration. This type of error can sometimes be readily identified. One indication of such an error is a large difference between right and left spherical equivalent (Table 3.2). Three individuals showed a large difference (> 8 dioptres) between right and left spherical equivalents as measured by non-cycloplegic autorefraction (no subjective refractions showed large differences). Each case was examined separately with reference to other measures of vision if available. Regarding case 15,376, the subjective measure, visual acuity and family history suggest myopia; the case was excluded from further analysis. Case 14,284 seems to have a particular problem with the right eye. Checking the original optometric record indicated 'balance' entered under right sphere, however, unaided vision was 6/120-1. Due the uniqueness of this case the measurements are excluded. In this instance the

type of refractive error is not suitable for a validation study which seeks to identify bias in non-cycloplegic autorefraction measures of refractive error in the general population. Removal of a case with atypical pathology will reduce power slightly (due to loss of an observation) but it will guard against the chance of such an abnormal case being a data error which would introduce a large amount of noise into the analysis. Case 1,180 is excluded due to the large difference between right and left autorefraction measures. No other data was available for this case. The non-cycloplegic autorefraction measures suggest an atypical pathology or a data error.

Extremely large differences between non-cycloplegic autorefraction and subjective refraction (on average more than 4 D difference, nine instances, Table 3.3) were considered another indication of whether a data error was present. These differences were not representative of the study sample (mean difference of 0.25 D, standard deviation 0.5 D). It is observed that in some cases the subjective refractions were considerably more negative than non-cycloplegic autorefraction. This is also unlikely due an offset in non-cycloplegic autorefraction which results typically in small negative readings compared to subjective refractions. In six out of the nine instances, subjective refractions are markedly more negative (Cases: 21,218, 20,368, 3,507, 5,309, 20,130 and 20,693). It is possible that in these cases autorefraction measures were taken while the subject was wearing corrective lenses (over refraction). In support of this conclusion three of the cases (3,507, 20,130 and 15,946) had an optometric record that strongly disagreed with the non-cycloplegic autorefraction measures. Each of the six cases was excluded on the basis that the autorefraction measures did not represent their refractive measures. Three more cases were excluded. Cases 15,946 and 20,693, display values located at eight standard deviations from the mean difference (over 4 D difference). They were excluded on the basis that such large differences are atypical and likely to overly influence the estimation of the offset. Case 8,266 showed a difference of 4 standard deviations from the mean. This case was also excluded.

	Autorefraction				Subjective						
Case number	SphEL	SphER	AveSph	Difference RL	SphEL	SphER	AveSph	Difference auto-sub	Query	Checked optometric record	Vision
*15,376	-3.50	5.63	1.06	9.13	-3.63	-3.38	-3.50	4.56	R eye	Yes	R:6/9 L:6/12
$14,284	-3.88	-14.38	-9.13	10.50	-3.00	0.00	-1.50	7.63	R eye	Yes	R:6/120-1
1,180	3.63	-9.75	-3.06	13.38	4.38	3.50	3.94	7.00	R eye	NA	NA

*Father and brother have myopia.

$ Case had strabismus. Optometric records indicate right eye has balance prescription, but vision is 6/120-1. This suggests the right eye is problematic as confirmed by autorefraction.

Table 3.2) Data check I. Individuals with a large difference between right and left spherical equivalent measured by autorefraction. R (right), L (left), SphE (spherical equivalent), AveSph (average spherical equivalent), Difference auto-sub (difference between subjective refraction and autorefraction for a case measured in dioptres), Difference RL (difference between autorefraction measures between right and left spherical equivalents), Checked optometric record (when a copy of the optometrist record was still available it was consulted), NA (none available).

	Autorefraction				Subjective						
Case number	SphEL	SphER	AveSph	Difference RL	SphEL	SphER	AveSph	Difference auto-sub	Query	Checked optometric record	Vision
21,218	-0.25	-0.63	-0.44	0.38	-2.63	-2.25	-2.44	2.00	Both eyes	No	NA
20,693	0.50	0.75	0.63	0.25	-6.50	-6.63	-6.56	7.19	Both eyes	No	NA
20,368	-0.63	-0.50	-0.56	0.13	-3.38	-3.50	-3.44	2.88	Both eyes	No	NA
3,507	-0.63	-0.50	-0.56	0.13	-3.13	-3.88	-3.50	2.94	Both eyes	Yes	R:6/60-1 L:6/60-1
5,309	-0.75	-0.88	-0.81	0.13	-4.25	-4.25	-4.25	3.44	Both eyes	No	NA
20,130	-0.50	-0.13	-0.31	0.38	-4.25	-4.50	-4.38	4.06	Both eyes	Yes	NA
15,946	-4.50	-4.50	-4.50	0.00	0.00	0.00	0.00	4.50	Both eyes	Yes	R:6/6 L 6/6
8,266	-3.13	-1.38	-2.25	1.75	0.00	0.25	0.13	2.38	Both eyes	No	NA
13,478	0.38	-0.13	0.13	0.50	6.00	6.00	6.00	5.88	Both eyes	No	NA

21,218: Sister mild myope.

8,266: NA.

20,368: Father myope.

3,507: Mother myope. Three optometric records indicated case has moderate myopia.

5,309: Mother and brother are moderate myopes.

20,130: Father is high myope, brother has moderate myopia.

15,946: Mother moderate myope. Two records, from 2006 and 2008 indicated the case is an emmetrope.

13,478: Mother emmetrope.

20,693: NA.

Table 3.3) Data check II. Individuals with a large difference between autorefraction and subjective refractions. The values noting differences between autorefraction and subjective refractions (column 'Difference auto-sub') are presented without sign (i.e. negative or positive difference). For abbreviations see previous.

3.3 Results

A scatter plot (Figure 3.1) was employed to give a general indication of the correlated nature of the measures from two different sources. Since each point represents two readings from the same individual (paired data) it would be expected that the plot depicts a reasonable relationship. Correlation measures association, not agreement (Woodward, 2005; Bland and Altman, 2010), and since the data is paired a strong association is likely, which is reflected in a high R square (R Sq) of 0.9. The plot makes clear that there is a relationship between the measures.

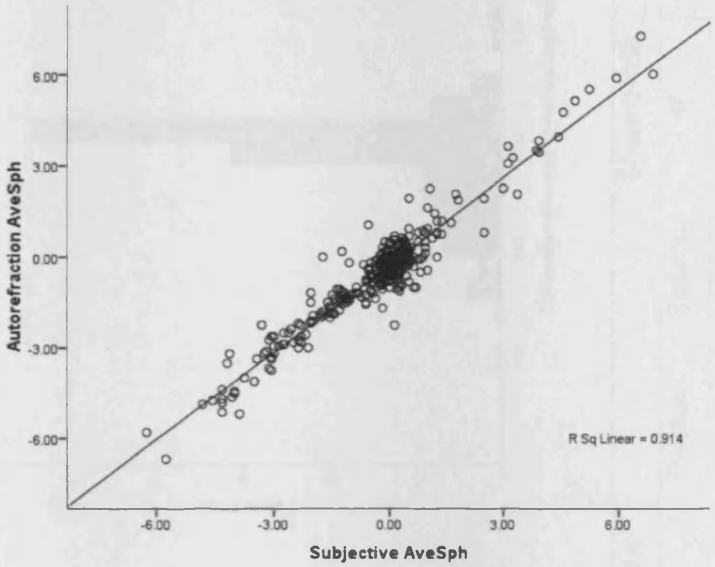


Figure 3.1) Measure of association. A scatter plot of non-cycloplegic autorefraction versus subjective refraction. AveSph (average spherical equivalent).

A negative trend in non-cycloplegic autorefraction is present when mean values for each group are compared (-0.58 vs. -0.33 for non-cycloplegic autorefraction and subjective refractions respectively, Table 3.4). This is supported by the median value for each group (-0.38 and 0.00 for non-cycloplegic autorefraction and subjective refractions respectively). Both these statistics indicate more negative measures of refraction in the non-cycloplegic autorefraction group. The distributions of both measures deviated significantly from normality (Figure 3.2) and therefore

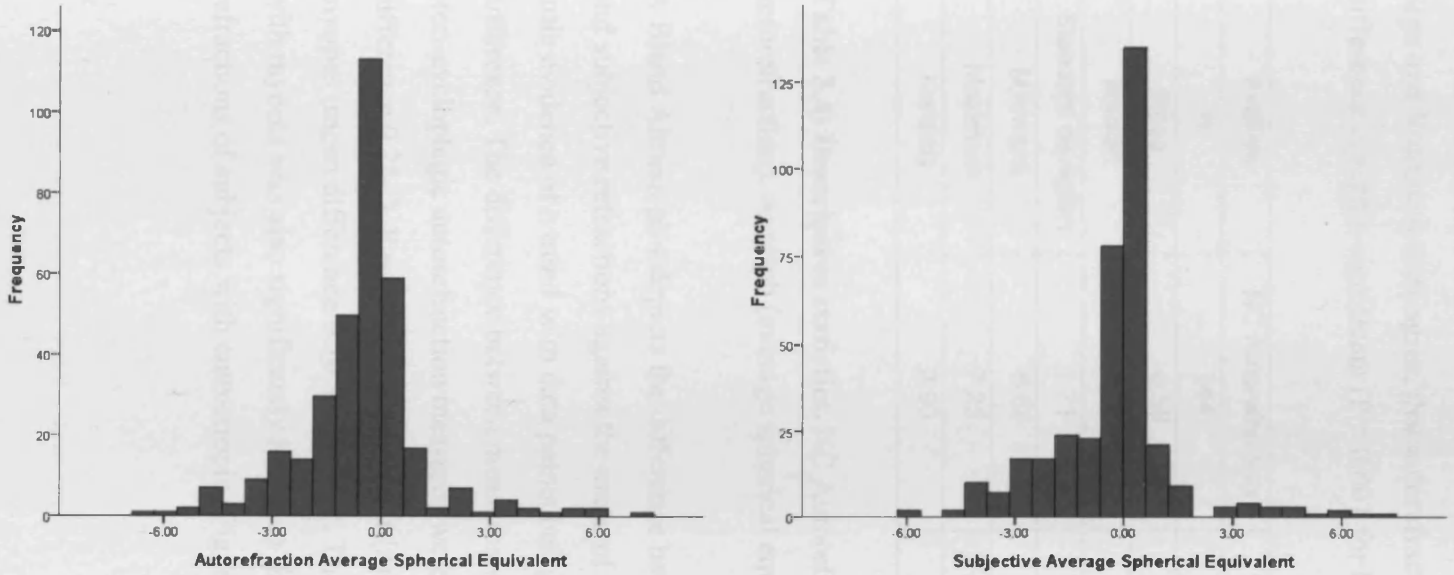


Figure 3.2) Tests of normality. Left, histograms of average spherical equivalent and below, normality tests for the two datasets, df (degrees of freedom). Low P values indicate significant deviations from a normal distribution. NC Autorefraction (non-cycloplegic autorefraction).

Tests of Normality	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	P-value	Statistic	df	P-value
Subjective refractions	0.176	363	1.E-30	0.888	363	1.E-15
NC Autorefraction	0.143	344	9.E-19	0.899	344	2.E-14

NC Autorefraction – Subjective refraction		N	Mean Rank	Wilcoxon Test	Sign Test
a. NC Autorefraction < Subjective refraction	Negative Ranks	251	178.07	7.E-22	3.E-20
b. NC Autorefraction > Subjective refraction	Positive Ranks	82	133.13		
c. NC Autorefraction = Subjective refraction	Ties	11			
	Total	344			

Table 3.5) Sign and Wilcoxon tests. Sign and Wilcoxon tests for differences between non-cycloplegic autorefraction and subjective refractions.

non-parametric statistics were used. The Sign and Wilcoxon test investigate if there is a statistically significant difference between groups in terms of their average spherical equivalent (Table 3.5). Both tests do not assume data is normally distributed. The Sign test investigates whether there are equal numbers of negative and positive differences between the two groups. The Wilcoxon test is similar but more powerful to detect differences because it takes into account the magnitude of the differences and not just their sign. However as a prerequisite for the Wilcoxon test the distributions should be symmetric. Observing the histograms of subjective and autorefraction measures, the distributions look reasonably similar. In both cases, the Sign and Wilcoxon tests agree, that autorefraction measures are more negative. This difference is highly significant ($P < 0.001$ for both tests).

AveSph	NC Autorefraction	Subjective refractions
N	344	363
Mean	-0.58	-0.33
Median	-0.38	0.00
Standard deviation	1.71	1.71
Minimum	-6.69	-6.25
Maximum	7.25	6.88
Kurtosis	3.90	3.34

Table 3.4) Descriptive statistics. NC Autorefraction (non-cycloplegic autorefraction). AveSph (average spherical equivalent)

A Bland Altman plot depicts the difference between non-cycloplegic autorefraction and subjective refractions against the mean of the two measures (Figure 3.4). There is little evidence of a trend with data points being evenly scattered around the mean difference. The difference between measures was examined within refractive status. Non-cycloplegic autorefraction measures were more negative for hyperopes (mean difference 0.25 D, $P = 0.033$), emmetropes (mean difference 0.36, $P < 0.001$) and myopes (mean difference 0.10, $P = 0.035$). The mean difference observed in subjects with myopia was also significantly less than the mean difference observed in the refractions of subjects with emmetropia (Figure 3.5).

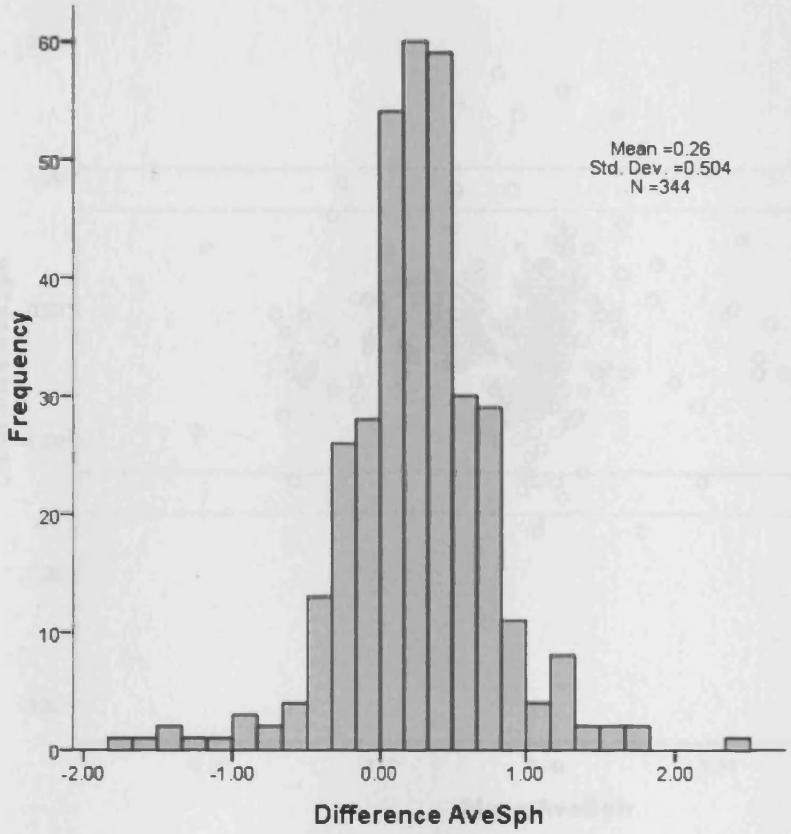


Figure 3.3) Histogram of differences. A histogram of differences between non-cycloplegic autorefraction and subjective refractions (Difference AveSph).

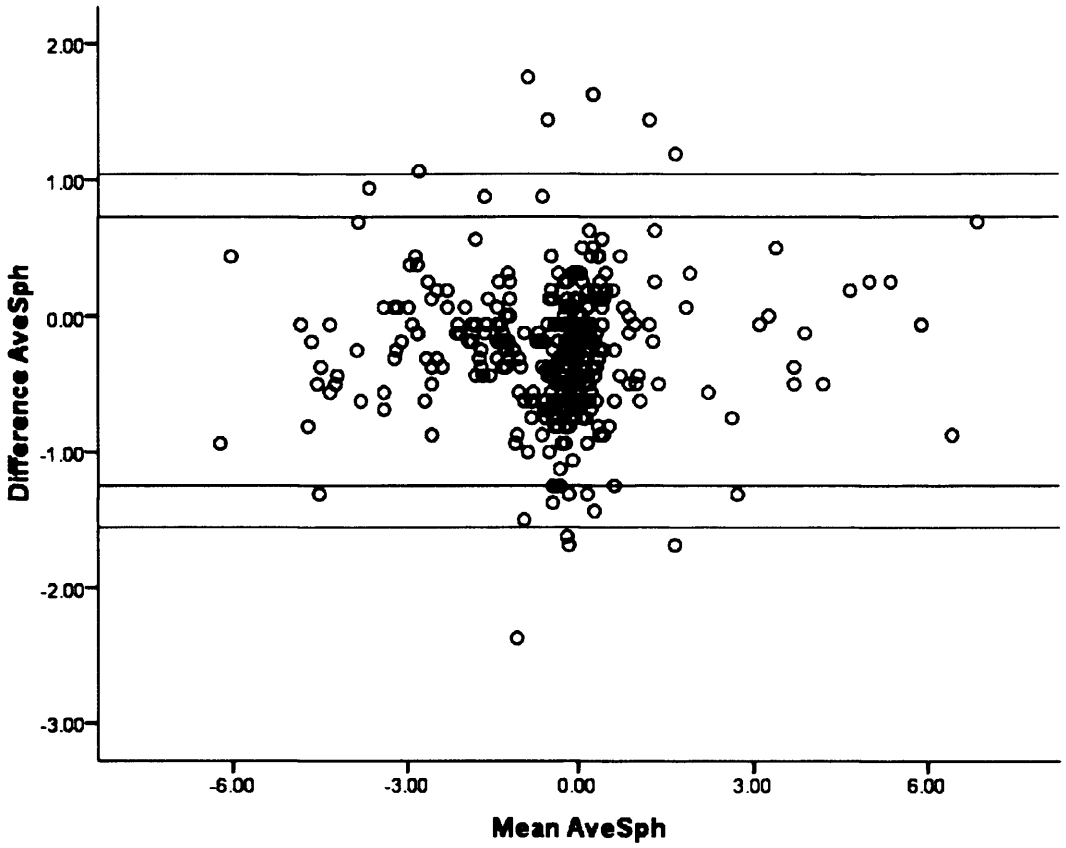


Figure 3.4) Bland Altman plot. Mean AveSph (average mean spherical equivalent), Difference AveSph (subjective refraction minus autorefraction). Red and green lines indicate 95% and 99% limits of agreement respectively for Difference AveSph. Below, mean and standard deviation (SD) for the difference between the two measures, used to identify 95% (by mean +/- 1.96*SD) and 99% (by mean +/- 2.58*SD) limits of agreement.

		Limits of agreement			
		95%		99%	
	Difference AveSph	lower	upper	lower	upper
Valid	344				
Missing	19				
Mean	-0.26	-1.25	0.73	-1.56	1.04
Standard deviation	0.50				

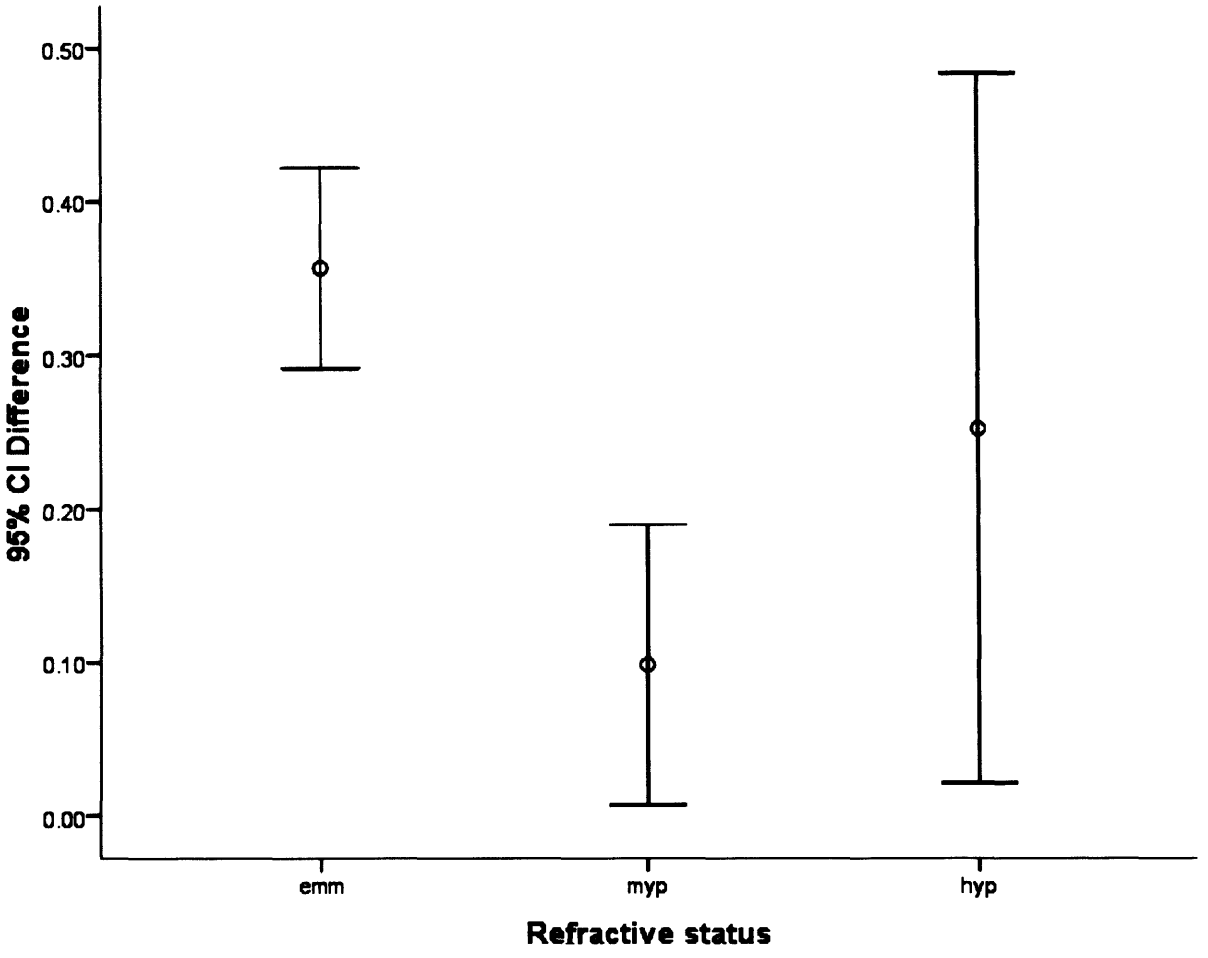


Figure 3.5) Offset by refractive state. Mean difference between measures plus 95% confidence intervals (95% CI difference) taken for (defined using subjective refractions) emmetropes (emm, > -0.5 to < 1 D), myopes (myp, <= -0.5 D) and hypermetropes (hyp >= 1 D).

The sensitivity and specificity of non-cycloplegic autorefraction to detect myopes (<-0.5 D from subjective measures) was optimal at -1 D at 89% and 96% respectively (Table 3.6b). Similarly the area under the ROC curve for this point is the highest of points tested at 0.92. This was significantly better than classifying myopia at random ($P << 0.001$).

	Myopia (Subjective refraction)	
	Frequency	Percent
No myopia	249	68.6
Myopia	114	31.4
Total	363	100

Table 3.6a) Distribution of myopia. Frequency distribution of myopia (<= -0.5 D) or no myopia (> -0.5 D) in subjective refractions.

Cut off (Autorefraction)	Sensitivity	Specificity	False positive rate	False negative rate	Area under curve (ROC C statistic)	P-value (C statistic)
<= - 0.5	92%	77%	23%	8%	0.85	6.E-25
<= - 0.75	90%	90%	10%	10%	0.90	3.E-32
<= - 1	89%	96%	4%	11%	0.92	1.E-35
<= - 1.25	76%	99%	1%	24%	0.88	2.E-28
<= - 1.5	60%	99%	1%	40%	0.80	2.E-18
<= - 1.75	55%	100%	0%	45%	0.77	6.E-16

Table 3.6b) Sensitivity and specificity. Sensitivity and specificity of autorefraction to detect myopia (<= -0.5 D) or no myopia (> -0.5 D) in the subjective refraction dataset at different cut offs.

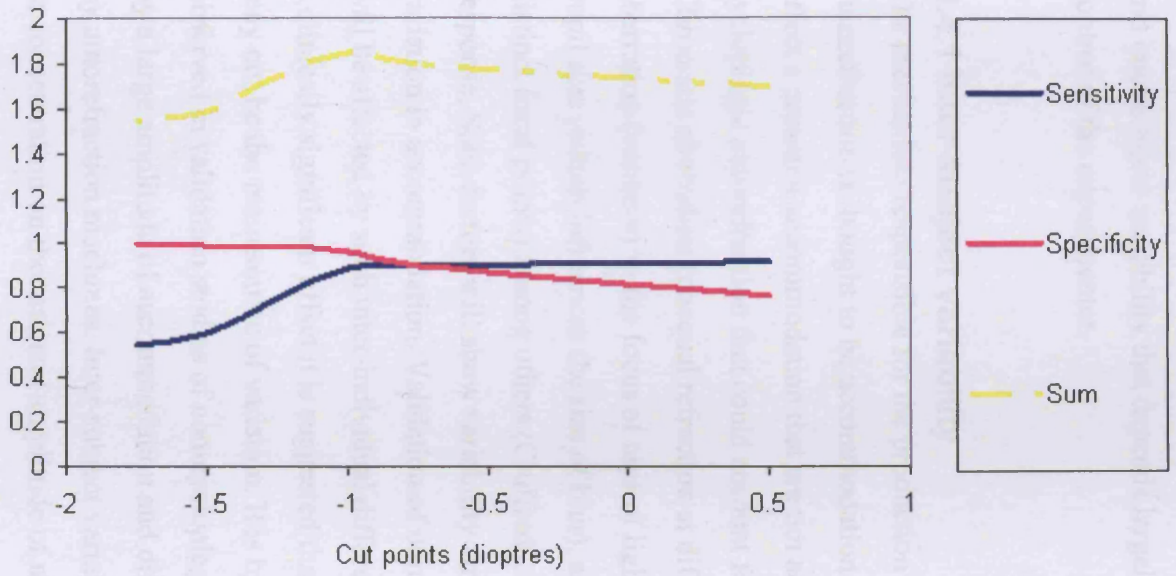


Figure 3.6) Optimal cut point. Sensitivity, specificity and their sum (y axis) against cut points (x axis, in dioptres) used to distinguish myopes from non-myopes.

3.4 Discussion

Here it is shown that there is an offset in the measure of refraction by non-cycloplegic autorefraction in the ALSPAC cohort. It is a clinically small difference of 0.26 D (95% limits of agreement -1.25 to 0.73). Other studies that have investigated bias in non-cycloplegic autorefraction in children and young adults have found a similar offset with a varying magnitude; -0.26 D in a sample of 120 individuals (Funarunart et al., 2009), -1.23 D in a sample of 5,000 individuals (Zhao et al., 2004), -0.84 D in a sample of 2,000 (Fotedar et al., 2007). The size of the offset will remain fairly constant in each sample as long as the sample size remains large enough to maintain random sampling. Differences in the size of the offset between studies are attributable to differences in study design (age of participants, the choice of a standard measure) and inter-subject variability that depends largely on environmental factors outside the control of the experimenter.

3.4.1 Inter-subject variability

The mechanism responsible for the production of error when using non-cycloplegic autorefraction is thought to be accommodation. There are a range of factors that can affect a person's accommodation that are not accounted for when calibrating non-cycloplegic autorefraction that could account for inter-individual differences. Chromatic aberration (unequal refraction at different wavelengths), spherical aberration (variation in the focus of rays of light entering the eye at different points), pupil size (which influences the size of blur), astigmatism (the presence of two distinct focal points) among others (Ciuffreda, 2006) affect the accommodative response. Such factors will show variability among individuals and therefore add to variation in accommodation. Validation of non-cycloplegic autorefraction in adults will be affected by such inter-individual differences and since such studies do not find a clinically significant offset it is suggested that the ocular factors suggested above may not be the main source of variation. It is hypothesized that a negative offset observed in validation studies of non-cycloplegic autorefraction in children is driven by a large amplitude of accommodation and difficulty attending to the task required by autorefraction machines. Inter-subject variability could be generated by variation in concentration on the task and amplitude of accommodation. In this study a mean

difference of -0.26 D is observed when comparing non-cycloplegic autorefraction to subjective refractions. The 95% limits of agreement are -1.25 D to 0.73 D. In other words, 95% of individuals display less than -1.25 D to 0.73 D difference. Figure 3.3 shows a histogram of the difference between the two measures. The precision of the 95% limits of agreement depend partly on the differences being normally distributed (Bland and Altman, 2010). In this case the differences show significant deviations from normality ($P < 0.001$) but good symmetry around the mean and by observation alone a reasonable approximation of normality. More importantly the mean difference is located at -0.26 D and there is a negative skew in the 95% limits of agreement with more negative points. This is to be expected given that non-cycloplegic autorefraction generates a negative bias when measuring refractive error in children. However as can be observed from Figure 3.3 some individuals show a positive difference, that is the subjective refraction reading was more negative than a non-cycloplegic autorefraction. Similarly, some individuals show little or no difference between measures while others show 1 D or more myopic refraction when measured by non-cycloplegic autorefraction. It is suggested that differences in the mechanism of accommodation between individuals and differences in attention to the task of autorefraction generate this variability.

3.4.2 Offset

In a calibration study it is possible to discern whether an offset is additive or multiplicative (i.e. whether a constant amount of error is present independent of the level of measurement or the error varies according to the magnitude of measurement). It has been shown in other calibration studies of non-cycloplegic autorefraction of children and young adults that a negative offset is observed to decrease when examining hyperopes, emmetropes and myopes (Zhao et al., 2004). Furthermore the largest error observed was for hyperopes, in a calibration study with adult subjects (Krantz et al., 2010). It has been observed that some myopes have a decreased accommodative response (Rosenfield, 1998). This suggests that refractive status influences accommodative response which would explain differences in error observed across ametropias. In this study it was found that myopes displayed a significantly lower offset when compared to emmetropes (Figure 3.5) with a mean difference of 0.1 D. This difference is lower than the average for the entire sample

(0.26 D) and is still significantly different from zero (P = 0.035). Considering differences in refraction of the hyperopic group separately, a significant difference was found between non-cycloplegic autorefraction and subjective refractions (mean difference 0.25 D, P = 0.033). The magnitude of this difference is somewhere in between myopes (0.1 D) and emmetropes (0.36 D) and displays large 95% confidence intervals (Figure 3.5). The size of the confidence intervals is confounded by sample size and reflects the decreased number of hyperopes in the study (n = 33) compared to myopes (n = 114) and emmetropes (n = 216).

No trend was discernible when observing the Bland Altman plot (Figure 3.4). If there was a linear increase in error that changed with the magnitude of the refractive error it would be expected to be visible using the Bland Altman method. However the method is confounded by differences in sample sizes across refractive states. It is possible to conclude that an additive offset is apparent in non-cycloplegic autorefraction. The error also shows evidence that it changes with the magnitude of measurement as non-cycloplegic measurements show less error when subjects are myopic. However it is unclear if there is a linear increase in the size of error as refractive measurement increases. It is likely given the results from other studies (Zhao et al., 2004; Krantz et al., 2010) that non-cycloplegic autorefraction measurement can display more error when hyperopes are examined separately and that sample size constraints limit the ability of this study to observe this trend.

3.4.3 Classification

Epidemiological investigations sometime require classification of individuals according to disease status. In Chapter 5 a case-control analysis is undertaken based on the presence or absence of myopia. Therefore it was important to accurately identify individuals as either myopes or not. The reliability of non-cycloplegic autorefraction to infer the presence of myopia has shown to vary (Choong, Chen and Goh, 2006; Fotedar et al., 2007). Its reliability was investigated in this study and found to be optimal when inferring myopia was present at -1 D or lower using non-cycloplegic autorefraction. In other words the number of times a myope was incorrectly classified as a non-myope or conversely the number of times a non-myope was incorrectly classified as a myope is minimized when -1 D by non-cycloplegic

autorefraction indicates the presence of myopia. Sensitivity and specificity were found to vary depending on what point myopia was inferred from the non-cycloplegic autorefraction measures (Figure 3.6). Sensitivity (the probability the test correctly identified myopia) and specificity (the probability the test correctly identified no myopia) were optimal at -1 D as can be seen from the sum of both measures in Figure 3.6.

Another indication of the performance of different cut points to infer myopia is the Receiver Operating Characteristic (ROC) plot (Woodward, 2005). This is a two dimensional plot of sensitivity against one minus specificity and is useful to compare tests or differing choices of cut points. A test that is perfect (i.e. compared to a standard the test always classifies individuals correctly and therefore has a sensitivity of one and a specificity of one) will display a straight line along the vertical axis at zero (one minus a specificity of one) and another straight line along the horizontal axis at one (sensitivity of one). The area under this graph is also one (and indicates a diagnosis that is as good as the standard diagnosis). When the test produces an equal chance of classifying a subject correctly when the subject has the disease or incorrectly when the subject does not have the disease, a diagonal line is produced on the ROC plot (Woodward, 2005). Tests with a large area under the ROC curve (and therefore a lesser number of instances where misdiagnosis occurred) are considered better at classifying individuals. The highest area under the ROC curve for classifying myopia in this study was observed at a -1 D cut point (0.92). This point was marginally better at identifying myopes based on non-cycloplegic autorefraction than -0.75 D (area under ROC 0.90) and -1.25 D (area under ROC 0.88).

3.4.4 Age

Refractive error is known to change with age (Rosenfield, 2006). Myopia increases rapidly from the ages of 5 to 15 (Maul et al., 2000; Zhao et al., 2000). Subjects in the present study were aged 15 when measures of refraction were taken. At this age it is expected that some myopia is still developing. Non-cycloplegic autorefraction was measured during a visit to an ALSPAC clinic. All measurements were not taken on the same day or by the same technician. All measurements were also taken over a period of a year. Therefore a certain amount of error due to small changes in the

environment would be expected in the non-cycloplegic autorefraction measures. However these errors are considered small and random and therefore may have no appreciable effect on calibration. The time at which non-cycloplegic autorefraction was taken is considered the first measure of refraction. Subjective refractions were collected from optometrists of young people at a later date. Refractions that had a date of test within six months (and therefore represent a subjective refraction taken close to the first measure) were collected to supplement the non-cycloplegic autorefraction measures. This differs from other studies where subjects are measured by non-cycloplegic autorefraction and another more accurate measure (subjective refraction, cycloplegic autorefraction, and cycloplegic retinoscopy) on the same day (Zhao et al., 2004; Choong, Chen and Goh, 2006). A plot of differences in measures by time between the first measure and second measure (subjective refraction) showed no trend towards larger differences in measures due to intervening time between measurements (data not shown). A correlation between intervening time and difference between measures was also not significantly different from zero (Pearson correlation 0.04, $P = 0.9$, Spearman rank correlation -0.014, $P = 0.8$). Although the choice of six months was slightly arbitrary it is shown that this time period is not large enough to bias the mean difference between measures appreciably. Even still it is suggested that some variability in measures is attributable to variation in the time between first and second readings. However these changes are observed to be small and random and may cancel out leading to little change in the estimate of offset between non-cycloplegic autorefraction and subjective refraction measures.

It has been observed that non-cycloplegic autorefraction has a negligible offset for older individuals (Krantz et al., 2010). In general, studies of refractive error in adults do not use cycloplegia when undertaking autorefraction, although many studies measure refraction subjectively as well as by autorefraction (Attebo et al., 1999; Wong et al., 2000; Xu et al., 2005; Saw et al., 2008). There is evidence that the accuracy of non-cycloplegic autorefraction in studies of refractive error of school children varies with age, being more accurate in older children. Fotedar et al. (Fotedar et al., 2007) found a mean difference between autorefraction measures pre- and post-cycloplegia of 0.84 D (95% CI 0.81 to 0.87 D) for children aged 12 and 1.18 D (95% CI 1.05 to 1.30 D) for children aged 6 years. ALSPAC is a longitudinal study and as such has recorded refractive error measures of children over time when they were

aged 7, 11 and as in the present study at age 15. Non-cycloplegic autorefraction has been used to measure refractive error in these age groups. A validation study of non-cycloplegic autorefraction was undertaken on refractive measures when participants visited a clinic at age seven. Calibration was undertaken in a subset of individuals at the clinic visit using cycloplegic retinoscopy (Williams et al., 2008b) by an experienced optometrist. The sensitivity and specificity to detect myopia was lower than the present study and may reflect the effect of age of subjects on the accuracy of non-cycloplegic autorefraction.

In summary, non-cycloplegic autorefraction is shown to generate more negative readings than subjective refractions for subjects in the ALSPAC cohort. The difference is clinically small (-0.26 D). The reliability of non-cycloplegic autorefraction to identify myopia was found to be good with a sensitivity and specificity of 0.86 and 0.96 respectively. Classification of subjects into myopes and non-myopes is the least bias at -1 D in non-cycloplegic autorefraction measures.

Chapter 4

Heritability of Refractive Error

4.1 Introduction

4.1.1 Heritability

Some conditions tend to cluster within families. This observation about a trait sometimes leads to the trait being termed *hereditary.* Heredity can be defined as the observation of a trait being transmitted from one generation to the next within a family (King et al., 2006). It is known that genes are the unit of transmission, an idea that can be traced to the work of Mendel (Hartl, 1983). That a trait appears more often in some families than others is a starting point to examine whether the trait is caused by a gene (Haines and Pericak-Vance, 1998). It had been noted as early as 1889 that myopia tends to cluster in some families (No authors listed, 1889).

If a gene is responsible for a disease, the mapping of the gene is important to discover its function and the aberration that causes disease pathology. However both genetics and the environment can play a role in the development of disease. If the environment is responsible for disease pathology then the importance of mapping a gene for the disease is absent. For example if lung cancer in patients who smoked was found to cluster in families it may be suggested that a genetic cause was responsible. Efforts to map the gene would be largely wasteful. It is known that smoking causes lung cancer (Woodward, 2005) and the development of lung cancer in those families is likely to be due to environmental exposure.

It is therefore important to evaluate the relative importance of genetic and environmental factors in determining the disease. This is the function of a heritability study. Heritability refers to the amount of variation in a trait determined by genetic factors. Heritability stems from research on quantitative traits, unlike the example above where the trait is an absence or presence of cancer, heritability is usually measured on a quantitative trait (also referred to as a metric character (Falconer and Mackay, 1996)). For a character to be metric is must be measureable, for example height and weight. A metric character is often continuous in distribution (King et al., 2006). Refractive error is measurable and displays a continuous distribution and many

studies on the heritability of refractive error use either spherical equivalent (refractive power of the sphere + ½ astigmatism) (Chen et al., 2007a; Klein et al., 2009) or average spherical equivalent (mean of right and left spherical equivalents) (Wojciechowski et al., 2005).

4.1.2 Polygenetic and environmental roles

Quantitative traits often do not show distinct patterns of Mendelian inheritance. High myopia has been found to segregate in some families in clear autosomal dominant (Young et al., 1998b) and recessive (Drack, 1998; Yang et al., 2009), and X-linked forms (Haim, Fledelius and Skarsholm, 1988). However other evidence suggests that myopia is not inherited as a Mendelian trait. In Alaskan Eskimos it was observed in 1969 that younger members of families displayed a prevalence of myopia of close to 45% while myopia was found in only 14% of adults. Similarly in present day China, the prevalence of myopia in school children is approximately 50% (Zhao et al., 2000), while in the elderly adult population it displays a prevalence of 20% (Cheng et al., 2003). A rapid change in prevalence (1 to 2 generations) suggests that other factors other than a single gene are important in the development of the disease. In studies of myopia and parental myopia, it is found that all children of parents who both display myopia do not go on to develop the disorder, instead only a low proportion develop the disease (28% (Mutti et al., 2002), 12.2% (Drack, 1998)). A gene inherited with a Mendelian pattern would not show low frequencies in the offspring of two affected parents.

Mendelian inheritance of a trait is evidence that a single gene is responsible for the development of the trait. When Mendelian inheritance is not evident, but there is evidence that the trait is hereditary, it can be hypothesized that the trait is under the influence of a number of genes (polygenic). There is evidence that myopia is influenced by a number of genes. Linkage analysis of high and common myopia has identified a number of cytogenetic locations throughout the genome (Young et al., 1998a; Young et al., 1998b; Hammond et al., 2004; Stambolian et al., 2005). Furthermore the role of genetics in myopia is supported by studies of family history of myopia that show an increased amount of myopia in children with two myopic

parents compared to one myopic or no myopic parents (Jones et al., 2007; Low et al., 2010).

Instead of Mendelian inheritance, a trait under the influence of polygenetic factors can exhibit a quantitative inheritance. An example of quantitative inheritance of ear length in maize demonstrates that when two parents from extremes of the quantitative trait are mated the offspring show intermediate values of the trait (King et al., 2006). This is found in families where high myopia is thought to be hereditary. In some families the degree of myopia is similar; in others it tends to vary widely (Drack, 1998; Young et al., 1998a; Young et al., 1998b).

In the above discussion, there is evidence that myopia occurs in families and that it does not always show Mendelian inheritance. Therefore it may be under the influence of a number of genes. However there is evidence that myopia develops due to environmental influences. It is found that myopia develops more often during the time when children first attend school (Maul et al., 2000; Zhao et al., 2000). There is an increased prevalence of myopia in jobs with more nearwork (Zadnik and Mutti, 1987; McBrien and Adams, 1997). Furthermore nearwork and myopia are found to be associated in many cohort studies (Zadnik and Mutii, 1998; Mutti et al., 2002; Saw et al., 2002a). In studies of family history of myopia when a parent has myopia the children do not always display the disease (Drack, 1998; Mutti et al., 2002). Myopia therefore can be considered to be under the influence of genetic and environmental influences.

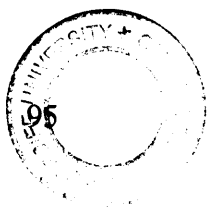
4.1.3 Previous estimates

A heritability study can measure the amount of variation in a trait that is attributable to genetic factors. The relative importance of genes in the development of the trait can be established via a heritability study. The heritability of refractive error has been demonstrated in a number of cohorts. The estimates of heritability vary considerably, being as high as 90% (Lyhne et al., 2001; Dirani et al., 2006) in twin studies to 50% to 60% in sibling and family studies (Chen et al., 2007a; Klein et al., 2009). A reason why estimates of the percentage of variation due to genes are higher in twin studies may be due to assumptions made in estimation. In such studies it is assumed that

genetic and environmental influences are not correlated (Falconer and Mackay, 1996). Identical twins (monozygotic, MZ) share the same genetic information, while fraternal twins (dizygotic, DZ) share 50%. It is possible that identical twins respond to environmental influences more similarly because of their genetic similarity. In twin studies, heritability is estimated from the difference in resemblance between MZ and DZ twins. It may be that this difference is larger due to genotype-environment interaction. This is not to be confused with a strength of twin studies, in that the amount of common environment MZ and DZ twins share is thought to be equal and therefore not a source of bias on the heritability estimate, unlike estimates of heritability in siblings which is inflated by a common environment (Falconer and Mackay, 1996).

Some studies have found lower estimates of heritability. In a study of a genetically isolated population in Sardinia, the heritability of refractive error was between 18% to 27% (Biino et al., 2005). The population is descended from a founder population, arriving on the island approximately 400 years ago and has a high level of endogamy (marriage within the population) (Biino et al., 2005). It is possible to hypothesize that the amount of genetic diversity in the population is reduced, as founder effect leads to such a reduction (Jobling et al., 2004b). A reduction in genetic diversity would mean changes in allele frequencies. Founder effect and long term isolation is found in Finland. Approximately 30 diseases show elevated rates in Finland, typically 75% to 80% of cases of each disease being caused by one mutation (Jobling et al., 2004b). It can be hypothesized that in populations with founder effect and relative isolation some diseases are caused by comparatively lower numbers of mutations. The heritability study of refractive error in Sardinia may have resulted in a different estimate (in this case a reduced estimate) due to the alleles responsible for refractive error being present at different frequencies when compared to a similar population with a larger effective population size.

Similarly another study estimated the heritability of refractive error at 20% (Paget et al., 2008b). This study examined extended families with at least two members that displayed non-syndromic high myopia. High myopia shows some notable similarities to rare Mendelian disorders. It has a low prevalence (1-2%) and it is a severe condition. This suggests that one rare gene with a large effect is responsible for



development of the disorder. Such a gene would be at a high frequency in the families selected for analysis. The frequencies of alleles that cause refractive error in the study population could be different to those in a similar population that was not selected for the presence of high myopia.

4.1.4 Summary

It is clear from the results of heritability studies of refractive error that the estimate of heritability varies. It may be that this variability depends upon genotype-environment interaction and the genetic heritage of the population. However in the majority of studies, the heritability of refractive error is demonstrated. Therefore evidence that genetics plays a role in the development of refractive error is strengthened. Determining whether a trait displays a genetic component is an important step in the process of understanding the biological pathway of disease development. The chance of mapping a gene underlying a disease is increased by first identifying that the disease is heritable (Haines and Pericak-Vance, 1998). It was important to demonstrate this in the ALSPAC cohort because a) family history and nearwork are independently associated with myopia in the ALSPAC cohort (Williams et al., 2008a). Therefore it may be that genetics and the environment influence myopia development in the study population. b) ALSPAC is designed to better understand the role of genetics and the environment in health and development. Gene mapping studies are ongoing in the cohort and it was important to establish a genetic component for refractive error before undertaking a mapping experiment.

Although other studies have demonstrated the heritability of refractive error (Chen et al., 2007a; Klein et al., 2009 ; Lyhne et al., 2001; Dirani et al., 2006) in this study it is considered more robust to estimate the heritability of refractive error in the study population (Visscher et al., 2008). The two main reasons for this are

a) The amount of phenotypic variability may change between populations due to non-genetic factors such as the environment. A change in phenotypic variability (V_p) will affect the denominator used to estimate narrow sense heritability ($h^2 = V_a/V_p$).

b) Differences in gene frequencies between populations may mean that heritability estimates vary from one group to another.

Other studies that seek to find a gene underlying refractive error, estimate the heritability of the trait in the specific cohort (Hammond et al., 2004; Biino et al., 2005).

It may have been noticed that a number of studies of the heritability of refractive error have been discussed but relatively little has been noted about the heritability of myopia *per say*. Heritability analysis is undertaken in traits with a continuous distribution, for example refractive error, while myopia is truncated, typically at -0.5 D. The trait is usually expressed in terms of two phenotypes, affected and unaffected (even though there is phenotypic variation within groups). As such myopia can be considered as a threshold trait (Hartl and Clark, 1997). It is possible to estimate the heritability of threshold traits. The estimation is based largely on a theoretical risk (or liability) to the trait, with affected individuals passing a threshold in risk of developing the disease. Comparison of this threshold in related populations leads to an estimate of heritability (Hartl and Clark, 1997).

Although the primary concern of this thesis is myopia, the heritability of refractive error was investigated. Myopia is defined as displaying a refractive error of less than -0.5 D. A subject can display mild (between -0.5 to -3 D), moderate (between -3 to -6 D) or high (less than -6 D) myopia. It is hypothesized that an allele that predisposes to myopia results in a shift towards less refractive power. The lack of such an allele would shift the strength of an individual's refractive power towards the more positive. As such, a gene that effects refractive error is either protective or increases risk of myopia.

There is evidence that the ametropias form a continuous spectrum of variation. Hyperopic eyes tend to be too short to allow light to focus clearly on the retina, while myopic eyes tend to be too long. The corneas of myopic eyes tend to have a short radius of curvature which results in light being focussed ahead of the retina. Similarly hyperopia can result from the flattening of the cornea (Rosenfield, 2006). Non-pathological ametropias are thought to occur due to a failure of correlation in the

refractive indices of the components of the eye (for example, axial length and corneal curvature) (Drack, 1998). This theory on the development of refractive error suggests genes underlying refractive error will influence myopia development.

4.2 Methods

Data collection of the study population is detailed in Chapter 2. Heritability was estimated using mother-offspring pairs (637), sibling-sibling pairs (527) and all available data (1898 individuals, full data) where subjective refraction data was available. Full data refers to a pedigree with information on at least one pair of individuals supplemented by information on other family members if present (the majority of which were mothers, young people and siblings). The frequencies of families of size 2, 3, and 4 in the analysis were 464, 200 and 55. In all estimates of heritability average spherical equivalent was used, defined as the mean spherical equivalent of right and left eyes (spherical equivalent, sphere + ½ astigmatism) by subjective refraction.

Heritability is defined as the ratio of additive genetic variance to the phenotypic variance (Falconer and Mackay, 1996)

$$h^2 = V_a/V_p$$

where h^2 is the heritability, V_a the additive genetic variance and V_p the phenotypic variance.

A number of methods were used to estimate h^2.

The intraclass coefficient (t) was used to estimate the heritability of refractive error in sibling pairs

$$t = \sigma_B^2 / \sigma_B^2 + \sigma_W^2$$

where σ_B^2 is the between group variances, and σ_w^2 is the within group variance.

$$2t = (V_a + ½ V_d + 2 V_{ec}) / V_p$$

where V_a is the variance of the trait due additive genetic effects, V_d due to dominance, V_{ec} due to a common environment and V_p is the phenotypic variance.

The regression coefficient (b) was used to estimate the heritability of refractive error in mother-young person pairs

$b = cov(x,y) / var(x)$

where x and y are the trait distributions for average spherical equivalent in the mother and young person groups respectively, cov is the covariance and var is the variance

$2b = V_a/V_p$.

Intraclass coefficients and regression based estimates were generated in SPSS (version 16.0, SPSS Inc., Chicago) using linear regression and a linear mixed model. Covariates (described below) were included in regression and linear mixed models (as fixed effects). Estimates of heritability adjusted for covariates, using full data were also generated in Sequential Oligogenic Linkage Analysis Routines (SOLAR, http://solar.sfbrgenetics.org) (Almasy and Blangero, 1998). The proportion of variance due to a common environment was estimated in SOLAR. Subjects in a pedigree were indicated as being a member of a sibship (all sibling-sibling pairs) or not (mother-offspring pairs, father-offspring pairs) (via the household option). Age and gender were included as covariates. Age was generated from the difference between date of birth and date of test. Gender was inferred from the gender of a subject's first name.

The normality of average spherical equivalent was assessed via the Shapiro-Wilk and Kolmogorov-Smirnova test of normality. Significant P values indicate large deviations from normality. Average spherical equivalent was transformed via an inverse normal rank distribution. The estimate of heritability was generated again with a normally distributed trait. SOLAR allows for traits to be automatically converted to a normal distribution via this transformation and the transformation was also undertaken manually in SPSS. To demonstrate the transformations were equivalent,

heritability estimates were compared for the mother-offspring analysis (no difference was observed, data not shown).

4.3 Results

Mother-offspring pairs were used to estimate the heritability of refractive error. Summary statistics of average spherical equivalent are given in Table 4.1. Note the adult population has more than twice the variance of the young person group. Since the square of the variance for average spherical equivalent of mothers would be a poor approximation to the product of both standard deviations ($8.62^2 \neq \sqrt{3.19} * \sqrt{8.62}$), the approximation of heritability using correlation may be biased. Regression, which does not rely on equal variances, was used to estimate the heritability. A linear regression of young person average spherical equivalent (y axis) on mother average spherical equivalent (x axis) is listed in Figure 4.1. The slope is 0.192 giving an estimate of heritability of 0.38 ($P << 0.001$). The sampling distributions of the two groups are not normally distributed ($P < 0.001$), therefore the heritability of refractive error was re-estimated after transformation to a normal distribution and found to be significant ($P < 0.001$).

Both gender and age show varying prevalence of myopia in the literature and are considered important here when estimating heritability. There is little evidence to suggest age of mother plays a role in refractive error development. Concurrently, age of mother was not associated with refractive error in young person (linear regression, $P = 0.152$). Therefore age of mother was not considered as a covariate.
To take into account variance due to differences in refractive error due to age and gender a multiple linear regression was undertake (Table 4.3). Concurrent with age of young person being a covariate with refractive error, a one year increase in age is negatively associated with a myopic shift of -0.14 dioptres ($P = 0.001$). The estimate of heritability remains largely unchanged after adjustment (0.37, $P < 0.001$). Multiple linear regression of the transformed trait of young person and mother was also undertaken. The estimate of heritability remained significant ($P < 0.001$).

To investigate the effect of a common environment the heritability of refractive error was examined in sibling-sibling pairs. Variance components methods with restricted maximum likelihood were used to estimate the heritability for average spherical

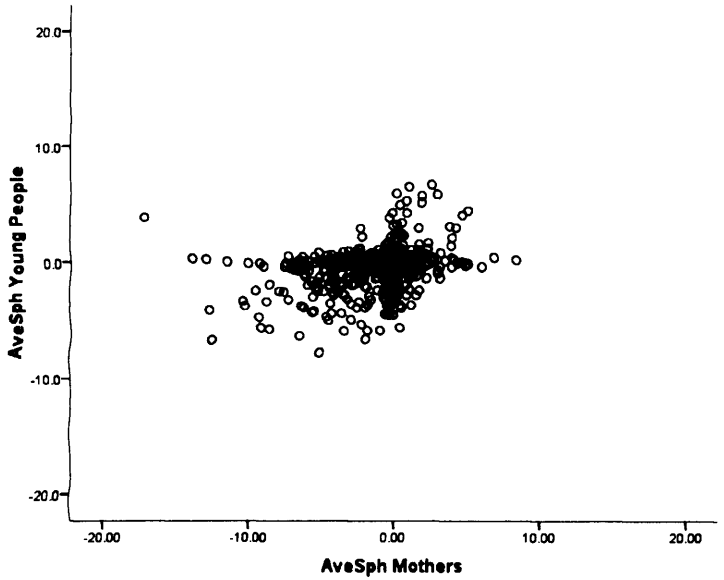
Statistics	AveSph	
	Young Person	Study Mother
N	640	607
Missing	0	33
Mean	-0.36	-1.15
Mode	0.00	0.25
Variance	3.19	8.62

Table 4.1) Descriptive statistics of mother-offspring pairs. AveSph (average spherical equivalent).

	Age (Young Person)	Age (Mother)
N	875	875
Missing	14	318
Mean	15.75	46.57
Median	16	47
Standard Deviation	1.57	4.65
Minimum	7	32
Maximum	19	59

Young People	Frequency	Valid Percent
Female	463	53.7
Male	399	46.3
Total	862	100
Missing	13	
Total	875	

Table 4.2) Descriptive statistics of covariates (mother-offspring pairs). Left, descriptive statistics of ages of young person and mothers and right, frequencies of males and females for young people.



	Unstandardized Coefficients			95% Confidence Interval for B	
	B	Standard Error	P value	Lower Bound	Upper Bound
(Constant)	-0.095	0.07	0.187	-0.24	0.05
AveSph Mothers	0.192	0.02	2.47E-16	0.15	0.24

Figure 4.1) Univariable analysis (mother-offspring pairs). Left a scatter plot of average spherical equivalent (young person versus mothers). Right, a simple linear regression of young person AveSph on mothers AveSph (giving a h^2 of 0.38, $P < 0.001$). AveSph (average spherical equivalent).

	Unstandardized Coefficients			95% Confidence Interval for B	
	B	Standard Error	P value	Lower Bound	Upper Bound
(Constant)	2.104	0.73	0.004	0.67	3.53
AveSph Mothers	0.187	0.02	1.64E-15	0.14	0.23
Age	-0.139	0.04	0.001	-0.22	-0.06
Gender	-0.003	0.14	0.983	-0.27	0.26

Table 4.3) Multivariate analysis (mother-offspring pairs). Multiple linear regression of AveSph of young person and mother.

equivalent among siblings. The intraclass correlation coefficient (t) was 0.35 giving an estimate of heritability of 0.70 ($P < 0.001$) (Table 4.5). A one year increase in age was associated with a -0.14 D myopic shift ($P <0.001$) in siblings and therefore included as a covariate. Gender is not associated with average spherical equivalent ($P = 0.328$) in this cohort but is known to be associated with refractive development and therefore was included as a covariate. The adjusted heritability estimate was 0.54 ($P < 0.001$) (Table 4.6), indicating that some similarity between sibling pairs was due to age or gender. This analysis was repeated and verified in SOLAR (heritability 0.541, $P <0.001$). Both analyses were repeated using a transformed distribution giving significant estimates of heritability ($P < 0.001$).

To estimate the heritability of refractive error using all available information, data from the full sample was entered into SOLAR as 719 pedigrees of sizes 2 to 4. This also allowed the proportion of variability due to a common environment shared between siblings to be estimated. The heritability of average spherical equivalent was 0.57 ($P < 0.001$) after adjustment for age ($P < 0.001$) and gender ($P = 0.26$). Analysis with a transformed trait showed the trait was still heritable ($P < 0.001$). The proportion of variance due to a common environment was 0.18 ($P < 0.001$) after adjustment for covariates. After taking into account the effect of a common environment the estimate of heritability was 0.50 ($P < 0.001$). Both the heritability and common environment remained significant after transformation ($P < 0.001$).

Statistics	AveSph	
	Young Person	Study Sibling
N	570	527
Missing	0	43
Mean	-0.40	-0.41
Median	0.00	-0.06
Mode	0.00	0.00
Variance	3.22	4.42
Minimum	-7.81	-9.38
Maximum	8.50	6.69

Table 4.4) Descriptive statistics of sibling-sibling pairs. Above, descriptive statistics of average spherical equivalent of sibling-sibling pairs. Below left, frequencies of males and females for young people and siblings. Below right, descriptive statistics of age of all siblings. AveSph (average spherical equivalent).

Siblings	Frequency	Valid Percent
Female	615	55.2
Male	500	44.8
Total	1115	100
Missing	21	
Total	1136	

Statistic	Age
N	1136
Missing	180
Mean	15.79
Median	16
Standard Deviation	2.58
Minimum	7
Maximum	31

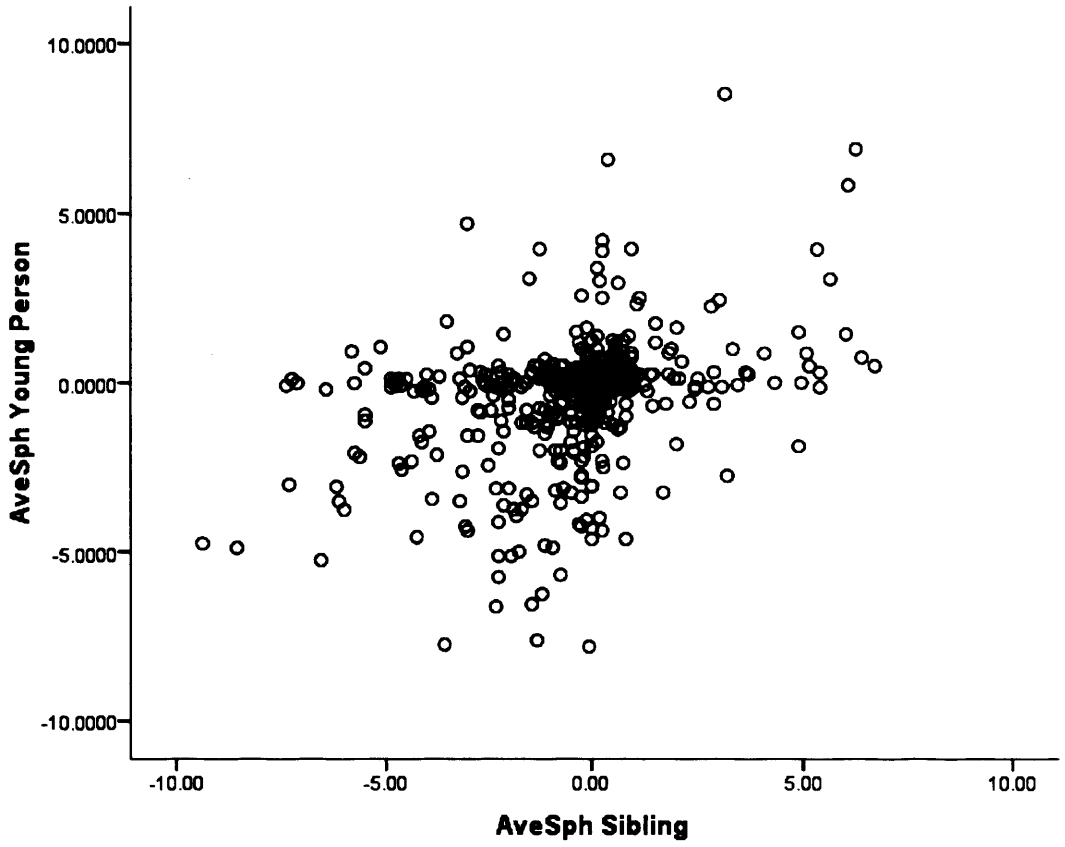


Figure 4.2) Scatter plot of sibling-sibling pairs. A scatter plot of average spherical equivalent of young persons on sibling. AveSph (average spherical equivalent).

					95% Confidence Interval	
Parameter	Estimate	Standard Error	Wald Z	P value	Lower Bound	Upper Bound
Residual	2.45	0.15	16.27	1.69E-59	2.17	2.76
SibShip	1.33	0.17	7.65	1.95E-14	1.03	1.71

Table 4.5) Variance components analysis of sibling-sibling pairs. Residual (within sibship variance), SibShip (between sibling variance).

						95% Confidence Interval	
Parameter	Estimate	Standard Error	df	t	P-value	Lower Bound	Upper Bound
Intercept	1.59	0.43	887.70	3.72	2.E-04	0.75	2.43
Age	-0.14	0.02	876.31	-5.78	1.E-08	-0.18	-0.09
Gender	0.12	0.12	905.34	0.98	0.328	-0.12	0.36

					95% Confidence Interval	
Parameter	Estimate	Standard Error	Wald Z	P value	Lower Bound	Upper Bound
Residual	2.60	0.18	14.16	1.74E-45	2.26	2.98
SibShip	0.96	0.18	5.29	1.22E-07	0.67	1.40

Table 4.6) Adjusted variance component analysis. Variance component analysis after inclusion of age and gender as fixed effects. Residual (within sibship variance), SibShip (between sibling variance).

4.4 Discussion

The heritability of refractive error in families from the ALSPAC cohort was estimated. After adjustment for variance attributable to a common environment in siblings and age and gender effects, the estimate of refractive error was 0.5 ($P < 0.001$). Other estimates of the heritability of refractive error generated by analyzing pedigrees from a population based cohort (0.62 (Klein et al., 2009)), a cohort with slightly increased amounts of myopia (0.5 (Chen et al., 2007a)) and a cohort of siblings (0.61 (Wojciechowski et al., 2005)) are similar.

4.4.1 Interpretation

Myopia shows evidence of being a multifactorial disease. Heritability studies indicate that only a proportion of the phenotypic variance is attributable to genetics and linkage studies indicate that some forms of myopia may be caused by genes that show Mendelian inheritance and display a severe phenotype (Young et al., 1998a; Young et al., 1998b). Furthermore evidence from animal studies indicates that myopia may be environmentally induced (Wallman et al., 1978; Smith, 1991). For these reasons heritability estimates need not be constant across populations. Factors that influence heritability estimates include changes in gene frequencies and changes in exposures to environmental conditions. The idea that heritability estimates need not be similar among different populations is supported by the observation that prevalences of refractive error vary across regions (Zadnik and Mutti, 2006). For example a study of myopia presenting in Jewish males found a higher prevalence in orthodox subjects (Zadnik and Mutii, 1998). It was hypothesized that the difference in amount of myopia present was due to the large amounts of nearwork orthodox subjects had undertaken as part of their studies. If the decision to practise orthodox customs clustered in orthodox families, heritability estimates in this group would be different, from estimates in other groups.

Estimates of heritability may vary between groups with little loss of value in interpretation because heritability estimates are often sought with a specific purpose. For example a breeder may want to know the heritability of a particular trait in similar populations before embarking on a breeding program. In the current study, the

heritability of refractive error was investigated to identify whether gene mapping studies of the trait would be amenable in the ALSPAC cohort. Heritability estimates were generated from a sample population with a refractive error distribution that was reasonably similar to the larger ALSPAC cohort (see discussion, Chapter 2) and suggests that some of the variation in refractive error in the cohort is due to genetics. The size of the heritability can be broadly indicative of the chance of success in a gene mapping experiment. However the interpretation of an estimate of heritability is more complicated.

4.4.2 A common environment

Heritability is often termed as being either narrow or broad sense. Narrow sense heritability is the amount of phenotypic variation that is attributable to genes passed on from one generation to the next. It takes no account of the necessary combination of these genes into genotypes. A change in a phenotype due to a person's genotype can be termed a dominance deviation. Estimates of heritability of siblings take into account that on average the probability that full sibs will share the same genotype is one quarter. Furthermore the relationship between alleles across loci (epistatis) may increase resemblance among individuals with similar genotypes. This can be termed interaction (Falconer and Mackay, 1996). In practise dominance deviation and interaction are not measured. In this study they are treated as negligible. Interaction between the environment and a person's genotype can also affect resemblance between individuals. Estimates of broad sense heritability contain dominance and interaction deviations.

In the current study, it was possible to estimate the effect of one type of environmental variance, the common environment between siblings. Covariance between refractions of relatives due to a common environment has been investigated in studies of refractive error with estimates varying from 2% (Lopes et al., 2008) to approximately 30% (Chen et al., 2007a). Due to differences between study design and study population, comparison between studies is not readily amenable and interpretation of previous findings is restricted to the observation that a proportion of variability in refractive error is attributable to the environment. In the current study the amount of variation in refractive error due to a shared environment was 18%. This

shared environment was specific to siblings. Other shared environments are hypothetically important when trying to estimate environmental and genetic components of refractive error. It has been noted that refractive errors were more similar between twins that engaged in similar vision activities than twins with dissimilar vision activities (Bear, 1991). This could be extended to parent-child relationships. It is possible that a parent's nearwork preferences have influenced the refractive error of the mother and the young person. Parents may set levels of nearwork (reading, studying, watching television, playing video games etc.) for their children and it is possible that parents, who are already engaged in certain nearwork activities, expose their children to a similar environment. Some authors now adjust heritability estimates by environmental risk factors (Chen et al., 2007a; Klein et al., 2009). Furthermore in a recent study two types of shared environment were examined; a nuclear family and sibling-sibling environments. It was found that the sibling-sibling environment was more accurate (Chen et al., 2007a).

The estimate of variation in refractive error due to a shared sibling-sibling environment in this study was estimated from pedigrees of families with various sizes. It was found to be a significant proportion of the phenotypic variance. Other evidence that a common environment leads to increased resemblance among siblings is gained when comparing the estimates of heritability generated from mother-offspring pairs and sibling-sibling pairs. The sibling-sibling estimate is inflated (0.54) compared to the mother-offspring pairs (0.37) which can be attributed in part to the difference between their phenotypic covariance. A mother-offspring regression estimates half the additive genetic variance while a sibling-sibling analysis estimates this component plus resemblance due to a common environment.

The mother-offspring regression is hypothesized to generate an estimate of heritability that is free from bias due to a common environment. However it is noted that the mother-offspring regression is not free from other sources of environmental variance. It is noted that the measure of heritability from a mother-offspring regression can be biased by maternal effects (Falconer and Mackay, 1996). A maternal effect refers to increased resemblance among individuals due to a maternal environment when offspring are young. An example of a trait that is influenced by maternal effects is birthweight. Consequently the heritability of birthweight has been shown to be

composed of at least three different maternal effects, two of which are the effect of the maternal genotype and non-genetic variation between mothers which account for 38% of variation in birthweight of children (Falconer and Mackay, 1996).

There is evidence that events at birth influence the development of myopia later in life. Photoperiod (number of daylight hours) (Mandel et al., 2008), season of birth (McMahon et al., 2009), birth order (Peckham, Gardiner and Goldstein, 1977) and gestation (Larsson, Rydberg and Holmstrom, 2003) have been associated with myopia. However these effects are small and there is little evidence to suggest that they are correlated across generations. If a maternal effect is to bias a heritability estimate using mother-offspring pairs it would have to increase the environmental covariance between mother and offspring. Furthermore birthweight is a trait that is present at a time when maternal effects can be thought to be at their most potent. Refractive error, most of which develops years later, may be less influenced by maternal factors.

Another factor that may be a source of covariance between the refractive error of relatives is age. Refractive errors tend to vary with age (Goss, 2006). It has been observed that the risk of myopia is higher for subjects with a myopic sibling and that risk increases when the age difference between siblings is small (The Framingham Offspring Eye Study Group, 1996). In the current study it was found that age is a significant covariate of refractive error (slope of -0.14 in regression and linear mixed models, $P < 0.001$). This indicates that age is responsible for some of the variability in the refractive error phenotype. After adjustment for age and gender the estimate of heritability generally decreased. For example the estimate based on sibling-sibling pairs was 0.7 before and 0.54 after adjustment. This further suggests that age can account for some of the covariance observed within sibships. Other evidence suggests that age is important in estimating the proportion of variability in refractive error due to a common environment. Lopes et al. (Lopes et al., 2008) found that the proportion of variability due to a shared environment fell from 7% to 2% after taking into account age.

4.4.3 Summary

Estimates of heritability using mother-offspring, sibling-sibling and pedigree data were generated in this study. This strategy is analogous to generating an estimate under *changing conditions*. The estimates of heritability varied from 0.37 to 0.54 but in each case a significant proportion of the phenotypic variance was attributable to genetics. Thus the evidence that refractive error is under some genetic control in the ALSPAC cohort is strengthened. To estimate the heritability of a trait using mother-offspring regression and a sibling-sibling intraclass correlation coefficient, each observation needs to be matched, creating a *balanced* design. Information was also available for additional siblings within each family. It is efficient to take into account information from many siblings in the estimate of heritability because the precision of the estimate will be increased. This can be achieved by use of restricted maximum likelihood estimation. The estimate of heritability using all available data was 0.5, after adjustment for age, gender and a common environment between siblings.

Two features differentiate this analysis from other pedigree based estimates. Firstly the study sample largely consisted of mothers and siblings. This is due in part to the overall design of the ALSPAC cohort which primarily focuses on mothers and children (Golding, 2004). In total 89% of parents in the pedigree based analysis were mothers. It is hypothesized that an effect on a heritability estimate of refractive error would be relatively small due to the refraction of mothers being analysed instead of both parents. There is little evidence that maternal effects play a large role in the development of refractive error.

Secondly the age of subjects in this study varied from a mean of 46.6 years for the parent group to 15.8 for the siblings group. In other studies of refractive error the study participants are typically over the age of eighteen (Lopes et al., 2008; Klein et al., 2009). A large proportion of myopia develops from ages 5-15 (Maul et al., 2000; Zhao et al., 2000). However a proportion of individuals develop myopia during early and late adulthood (Goss, 2006). It is noted that most subjects in the sibling group will yet to have developed refractive errors that have an adult onset. There is evidence that adult onset myopia is under genetic control. In a study of twins with adult onset myopia, Dirani et al. (Dirani, Shekar and Baird, 2008) found a significantly higher correlation between identical twins compared to fraternal twins. It is therefore

possible to suggest that genes underlying adult onset myopia account for a proportion of the heritability of refractive error. The influence of such genes on the heritability of refractive error in the current study must be largely absent given the mean age of the siblings group. Therefore the estimate of the heritability of refractive error in this study may be more accurately defined as the heritability of juvenile onset and early onset refractive errors.

To summarize it is shown here that a) refractive error is heritable in the ALSPAC cohort and therefore is a candidate for gene mapping studies. The heritability is demonstrated in two classic heritability designs and using measures from multiple family members b) a shared common environment contributes to the variation in refractive error, which indicates that the environment plays a role in the development of refractive error in the cohort. Evidence of a shared environment was found by estimation of the amount of variation due to being a member of a sibship in an unbalanced design and by comparing one parent-offspring regression to sibling-sibling intraclass correlation.

Chapter 5

Genome-wide Association Study

5.1 Introduction

The aim of this study is the identification of a genetic element that may be causally related to refractive error or one of the ocular components involved in refractive development. The study uses a genome-wide scan of hundreds of thousands of genetic locations in the form of single nucleotide polymorphisms (SNPs) that are available largely due to the sequencing of the human genome (Venter et al., 2001), the HapMap project (International HapMap Consortium et al., 2005) and commercially available genotyping platforms. SNPs are the most abundant genetic markers in the human genome (Wang et al., 1998) and allow for the genome to be densely mapped which will decrease the distance between a marker and the causal genetic mutation. Linkage analysis has revealed a number of cytogenetic locations where a mutation resides that may lead to myopia (termed MYP1-17). In a genome-wide linkage analysis a large segment of the genome is identified (1-10 centimorgans, cM) which contains hundreds of genes and then a positional cloning strategy is undertaken, where the region under the linkage signal is mapped in more detail to find markers that enclose the causative gene. Although genome-wide linkage analysis has been successful in finding strong linkage for a myopia gene (Young et al., 1998a; Young et al., 1998b; Hammond et al., 2004) and efforts to narrow down promising linkage signals have been undertaken (Young et al., 2001; Lam et al., 2003b; Nurnberg et al., 2008), a causative mutation for myopia has yet to be identified.

5.1.1 A genome-wide approach

An advantage of using current SNP data in a genome-wide association analysis is that the genetic location identified is smaller leading to lesser area to search to find a causal mutation. Using a set of 300,000 to 500,000 SNPs up to 80% of common variation in the genome (variation that leads to common diseases) may be examined (Frazer et al., 2009). On top of this regions typically identified by genome-wide association analysis are from 10 to 100 kb (as opposed to 2 to 10 Mb from family based linkage analysis) (Altshuler, Daly and Lander, 2008). Although genome-wide

association studies are underway for myopia and ocular traits very few have been published. It can be hypothesized that these studies will lead to the identification of regions of the genome that are smaller than those identified in linkage analysis and there will be an increased chance of finding a causal mutation.

A number of candidate genes have been proposed to be involved in myopia development due to evidence from experimental work implicating their biological pathways in the development of myopia. A candidate gene can be described as a gene with a biological function that is similar to the physiological or biochemical basis of the phenotype. Variation in physiological mechanisms of complex diseases such as heart disease and obesity (Hirschhorn and Daly, 2005) have been found using a candidate gene approach. An increase in axial length often will accompany myopia development in humans and animals and for this to occur the sclera must accommodate an increase in size. Evidence from animal and human studies indicates that the sclera changes during myopia development (Curtin, Iwamoto and Renaldo, 1979; Rosenfield, 1998; McBrien, Jobling and Gentle, 2009). This has led to genes involved in re-modelling of the sclera to be examined in genetic association studies. Similarly genes that code for molecules whose normal physiological function is disrupted in animal and human studies of myopia have been investigated in genetic association studies of myopia. While some association signals have been observed, no causative mutation has been identified (see Appendix A for summary of candidate genes investigated in myopia research).

There may be a disconnect between animal studies of a disease and the human physiology due to a significant time for genetic divergence to have occurred in between the present day and the existence of a last common ancestor (Jobling et al., 2004b). In myopia research, a study has found that mice lacking genes coding for lumican and fibromodulin display larger axial lengths than wild type mice. Association studies of these genes in humans have failed to find a causative mutation (Paluru et al., 2004; Wang et al., 2009b). It may be that differences in function exist between the human and mouse homologues of these genes. In other studies of myopia in animals, a number of biological molecules have been observed to be disrupted during myopia development. For example the expression of matrix metalloproteinases and their activity have been shown to be dysregulated in animal models of myopia

(Rada and Brenza, 1995; Guggenheim and McBrien, 1996; Siegwart and Norton, 2002). Association studies between the genes encoding matrix metalloproteinases in humans and myopia have been undertaken (Nakanishi et al., 2009a; Hall et al., 2009; Wojciechowski, Bailey-Wilson and Stambolian, 2010) with association found but no causative mutation identified. It may be that evidence of disruption of endogenous levels of a molecule may indicate involvement in development of the disorder downstream from the cause so that the biological *effects* of a causative mutation are being observed. Therefore investigation of the gene of a molecule disrupted during myopia development may not yield a causative mutation.

Another disadvantage of the candidate gene approach is that there can be a difference between genes that when disrupted lead to a severe phenotype and those that lead to milder variation in the same phenotype. For example a number of genes involved in pigmentation when disrupted are known to cause severe phenotypes like the PAX3 gene and Waardenburg syndrome but the involvement of such genes in variation of normal pigmentation is not clear (Jobling et al., 2004b). In myopia research a number of genes that have been implicated in severe phenotypes which list myopia as one of the symptoms have been investigated. Mutations in the COL2A1 (chromosome 12q13) gene cause Stickler syndrome, a progressive connective tissue disorder that leads to deafness, progressive arthritis, cleft palate and myopia among other symptoms (Millodot and Laby, 2002). Researchers have investigated a relationship between the COL2A1 locus and high myopia (Metlapally et al., 2009a) and low myopia (Mutti et al., 2007a). Significant association signals between SNPs and the disorders have been demonstrated but no causative mutation has been identified. A relationship between both the glutamate receptor metabotropic 6 gene (mutations in this gene cause congenital night blindness) on chromosome 5q35 and myopia (Dryja et al., 2005; O'Connor et al., 2006; Xu et al., 2009) and myocilin (mutations in this gene cause glaucoma which is sometimes accompanied by myopia) on chromosome 1q24 and myopia have been investigated with similar results (Tang et al., 2007; Vatavuk et al., 2009; Zayats et al., 2009).

An advantage of a genome-wide analysis to investigate a gene underlying the progression of a disorder is that no prior knowledge is required about the functioning of such a gene. In other words the investigation can be initially hypothesis free and

each genomic location can be considered to have an equal chance of being the location of a causative mutation. This advantage of genome-wide association studies (that it does not need an underlying biological hypothesis to drive successful mapping) has be cited to have been helpful in understanding complex diseases. For example genes implicated in multiple sclerosis by genome-wide association studies derive from different biological pathways involving immune function but also axonal function (Frazer et al., 2009).

A genome-wide association study may not identify a causal variant but rather may identify a genomic region where such a variant can be found. One of the benefits of a genome-wide association study is that common variation across the genome can be surveyed with a reduced number of genotyped SNPs due to patterns of linkage disequilibrium (LD) which allow some SNPs to effectively tag others that are in high LD. However there are properties of the SNPs examined in a genome-wide study which can affect the chance of finding an association.

Wang et al. indicate that as minor allele frequency (MAF) decreases, to maintain a power of 80% at a $P < 10\text{-}6$ significance threshold, the number of cases and controls rises dramatically (Wang et al., 2005). Apart from minor allele frequency the degree to which a causal SNP is in LD with a genotyped marker will affect an ability to find an association. Causal SNPs that are in high LD with a genotyped marker may give a larger association signal compared to less well tagged causal SNPs. As LD between a genotyped marker and causal SNP decreases the power to detect an association follows at a rate proportional to r^2. The HapMap consortium estimates that a causal SNP with an r^2 of 0.5 with a genotyped marker requires twice the sample size to detect a similar strength of association (International HapMap Consortium et al., 2005). Furthermore a relationship between minor allele frequency and patterns of LD are not necessarily independent as it is estimated that in the HapMap data SNPs, with a MAF of less than 10% are less well tagged (International HapMap Consortium et al., 2005)

Power to detect an association also depends on samples size and effect size. These two properties are not independent; as an effect size decreases, more samples are necessary to detect a significant association. It has been noted that associations found in well powered genome-wide association studies for height, Crohn's disease and

breast, prostate and colorectal cancers may be biased towards identification of loci with a large effect size for a common disease and that many more causal loci with a smaller effect size exist (Park et al., 2010). In another study of the power of genome-wide association to detect causal loci it was observed that current genome-wide association studies may be underpowered to detect loci of small effect size (an increase in risk of 1.1 to 1.2) but that large scale meta-analysis and follow up in larger studies will increase power (Spencer et al., 2009).

5.1.2 Ocular determinants

Myopia is thought to develop due to an imbalance between the ocular components, leading to light rays being focussed ahead of the retina. Axial length is defined as the distance between the anterior and posterior poles of the eye (Millodot and Laby, 2002). Myopic eyes are found to have longer axial lengths (Gonzalez Blanco et al., 2008) in young adulthood, before the onset of myopia (Mutti et al., 2007b) and the progression of myopia in childhood is mediated by axial elongation of the vitreous chamber (Goss, 2006). Furthermore a strong negative correlation is observed between axial length and refractive error (i.e. as axial length increases subjects tend to become less able to focus objects in the distance) (Wildsoet, 1998; Goss, 2006). It is suggested that axial length plays a role in the development of myopia. Furthermore it can be hypothesized that mutations that lead to changes in axial length may be responsible in part, for the development of myopia.

However it is noted that in high myopia, where subjects display less than -6 D in refractive power and are at an increased risk of pathological complications including retinal detachment, cataract and glaucoma (Saw et al., 2005), axial elongation is more distinct (Drack, 1998; Marsh-Tootle and Frazier, 2006). In emmetropic subjects axial length is typically 24 mm (Millodot and Laby, 2002) and can vary. Furthermore in more moderate myopia, the disorder is thought to develop due to a mismatch between the refractive power of the ocular components and axial length rather than simply an axial length that is abnormal in length (Drack, 1998; Marsh-Tootle and Frazier, 2006). The majority of subjects in the current study that have myopia display a moderate amount of the disorder. Therefore it is unlikely that a mutation in axial length that may be found in a genome-wide association study causes myopia development. It

seems more reasonably to suggest that the current study is powered to find mutations that are responsible for normal variation in axial length (i.e. found in both emmetropes and myopes). However it is possible that if such mutations were present along with mutations leading to changes in other ocular components, myopia may develop.

Another ocular component that is important in maintaining the refractive power of the eye is corneal curvature. Corneal curvature refers to radius of curvature of the cornea, the transparent anterior portion of the eye (Millodot and Laby, 2002). The cornea is a refractive surface that has a refractive power of approximately 42 D that varies with age (Grosvenor, 1991). Corneal curvature is typically 7.8 mm (Millodot and Laby, 2002) and also can vary. Corneal power has been shown to have a normal distribution (Rosenfield, 2006). Corneal power is positively correlated with corneal curvature and it has been observed that variations in corneal power are related to the development of refractive errors. In the corneal curvature of subjects with juvenile onset myopia and in those with moderate myopia a trend towards shorter radii has been observed (Gonzalez Blanco et al., 2008). Similarly, the mean corneal power between emmetropic and myopic subjects has been found to be significantly different (Rosenfield, 2006). Since corneal curvature is important in focussing light on the retina, it is hypothesized that genes underlying variation in corneal curvature can influence refractive error development.

An increased corneal curvature observed to occur in some cases of myopia increases the distance between the retina and a point where an image is brought into focus. During infancy and up to early adulthood the eye continues to grow. An increase in axial length occurs which increases the distance between the retina and light rays entering the eye. The need for more hyperopic refraction is met by a thinning of the crystalline lens (Wildsoet, 1998). Conversely it is thought that the cornea does undergo a significant change in refractive power during eye growth. It is noted that the refractive power of the cornea changes by approximately 1 D from the ages of 4 to 19 (Grosvenor, 1991) and that changes in refraction are not well correlated with corneal radius (Wildsoet, 1998).

It has been demonstrated in Chapter 4 that refractive error is heritable in the ALSPAC cohort. A number of other studies have similarly demonstrated that refractive error

has a genetic component (Chen et al., 2007a; Klein et al., 2009, Wojciechowski et al., 2005). Axial length has been shown to be heritable in a number of studies with estimates of 92% and 94% (Lyhne et al., 2001; Dirani et al., 2006) in twin studies and 67% and 73% in family studies (Chen et al., 2007a; Klein et al., 2009). Similarly corneal curvature has been shown to be heritable with varying estimates of 95% (Klein et al., 2009) and 16% (Chen et al., 2007a) in family studies and 90% to 92% (Lyhne et al., 2001) in a twin study.

A number of loci hypothesized to play a role in the development of low to high amounts of myopia have been identified by linkage analysis (Paluru et al., 2005; Chen et al., 2007b; Wojciechowski et al., 2009b) (see Appendix A). A smaller number of loci have been implicated in variation of axial length and corneal curvature. Suggestive evidence of linkage has been found for axial length and corneal curvature (Biino et al., 2005). However no causative mutation has been identified for common refractive error or axial length and corneal curvature. Given that axial length and corneal curvature are determinants of refractive error and are heritable, these traits were investigated in the current study via genome-wide association. Genes underlying myopia and refractive error were also investigated.

5.2 Methods

5.2.1 Genetic data

Genotyping was undertaken on two Illumina 317K and two 610K single nucleotide polymorphism (SNP) platforms. The 317K platforms contained genotype data on 1,760 and 2,030 individuals with approximately 90% overlap (i.e. 90% of subjects were represented on both platforms). The 610K platforms contained genotype data on 1,244 and 772 subjects. After error checking (detailed below) SNPs which were common to all platforms were retained (n = 285,537) for a total of 3,222 subjects. On average 165 individuals and approximately 8,000 SNPs were excluded per platform due to quality control.

Cleaned genetic data from each platform was merged into a single pedigree file using the genome-wide association program PLINK (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007). Imputation was undertaking in the program MACH 1.0 (http://www.sph.umich.edu/csg/yli/mach) (Li et al., 2009a) using data from the HapMap project (build 36, release 22) giving a total of 2,543,888 genotyped and imputed SNPs. Principal components of genetic variation were tested for association with the traits of interest. If an association was found it was taken as evidence of possible stratification of the phenotype by genetic background and the relevant principal components were included as covariates to mediate the generation of spurious signals. Association analysis was undertaken in PLINK (genotyped only SNPs) and Mach2qtl and Mach2dat (imputed and genotyped SNPs).

5.2.1.1 Missingness

Data from each platform was assessed individually. Missingness was used as an indicator of genotyping quality. Missingness for number of genotypes per individual and number of individuals per SNP were assessed. A threshold of 3% missing genotypes per individual was set (subjects that displayed more than 3% of genotypes missing that is not successfully genotyped, were excluded, on average five individuals per platform). A threshold of 5% or 3% was set for missing individuals per SNP

(SNPs that displayed more than 5% of individuals where genotyping failed were excluded, approximately 6,000 SNPs per platform).

5.2.1.2 Hardy-Weinberg equilibrium

A second indicator of genotyping quality was estimated with reference to Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (HWE) (Hartl and Clark, 1997) describes a mathematical relationship between allele frequencies and genotype frequencies such that

AA: p^2 Aa: 2pq aa: q^2

where A and a are major and minor alleles, p and q are their respective frequencies and the frequency of the major homozygotes, heterozygote and minor homozygotes are AA, Aa and aa. The principle relies on a number of assumptions to hold true, some of which are described below.

- two alleles at the locus
- an infinite population
- low migration
- low mutation
- no or negligible selection

If one of these assumptions is broken, deviations from equilibrium are expected. The allele frequencies of each SNP are used to estimate the number of expected genotypes at each locus. Deviations from such indicate an increase or decrease in the number of major or minor homozygotes or heterozygotes. This in turn is evidence that one of the assumptions of Hardy-Weinberg equilibrium has not been met. For example the locus that gives rise to sickle cell anemia (Hb^S) increases morbidity of homozygotes due to a red blood cell disorder (King et al., 2006). However heterozygotes are at a decreased risk of infection by a malaria causing parasite. This protective effect is thought to increase the prevalence of red blood cell disorders in areas of malaria endemicity (Jobling et al., 2004b). The decreased frequency of homozygotes for the sickle cell gene may be hypothesized to lead to a lack of Hardy-

Weinberg equilibrium at the locus. In this way, deviations from Hardy-Weinberg equilibrium can indicate alleles that are under a phenotypic pressure and by inference, are *causally* related to a phenotype.

Deviations from Hardy-Weinberg can indicate other phenomena. Systematic genotyping errors resulting from failure to amplify one or more of the homozygote and heterozygote classes would lead to deviations from equilibrium. Therefore HWE can be used to investigate genotyping quality. It is hypothesized that large systematic genotyping errors would lead to very large deviations. The effect size of alleles of potential interest is small in a genome-wide association study and the effect of selection on such an allele would also be small. Therefore it is further hypothesized that small deviations from HWE may indicate alleles of interest which are under selective pressure. For these reasons, a low threshold was applied for SNPs to be excluded on the basis of failure to meet Hardy-Weinberg equilibrium (P <10e-7, on average 1,200 SNPs per platform).

5.2.1.3 Minor allele frequency

The power to detect an effect decreases with minor allele frequency (MAF). Low minor allele frequency can lead to low sample numbers to estimate effect sizes and this in turn can lead to bias in a test statistic. Alleles displaying a minor allele frequency of below 0.5% were excluded (on average 700 SNPs per platform).

5.2.1.4 Other indicators of genotyping quality

X chromosome inbreeding (F) can be used as a genetic measure of gender. Identical by descent (IBD) refers to alleles which are identical because they were inherited from a common ancestor (Jobling et al., 2004b). The inbreeding coefficient F is the probability that two alleles at a locus are IBD (Hartl and Clark, 1997).

For the sex chromosomes the coefficient is close to one or zero for males and females respectively. Values that were between 0.2 and 0.8 were considered as indicators of poor quality samples or contamination. A small number of samples displaying such values were excluded. A small number of samples showed contradiction between X

chromosome inbreeding values and previously assigned gender. These instances were most likely labelling errors and were excluded.

If samples are contaminated during genotyping with DNA from another sample (an unrelated individual) the number of heterozygous alleles would be expected to increase substantially. Heterozygosity (the percentage of alleles that are heterozygote) was examined and samples that showed large excesses of heterozygosity were removed.

Individuals in a sample who have no known relation may upon inspection of their genetic data, display a degree of relatedness. If this relatedness is associated with the trait of interest then an association signal may be confounded by the relationship. It is possible to estimate the percentage of alleles shared identical by descent between pairs of individuals (i.e. the allele was derived from a common ancestor) by reference to the amount of alleles that are identical by state (i.e. alleles are similar but not due to a common ancestor). Pi Hat is the term given to the proportion of alleles that share IBD estimated as above (Purcell et al., 2007). Relatedness of individuals can be indicated by IBD. A value of 1 indicates monozygotic twins, 0.5 fraternal twins or full siblings, 0.25 half-siblings, 0.125 first cousins and so on. One of each pair of individuals that showed relatedness at the first cousin level (a Pi Hat above 0.1) was excluded.

5.2.2 Model determination

Ten principal components (PCs 1-10) were available, which summarize the extent of shared genetic background in the sample. All ten components were tested for each trait. For the three quantitative traits, Spearman rank correlation test was used to identify association. This uses ranks instead of actual values, it is non-parametric and does not require data to be normally distributed. For the categorical trait myopia status, a Mann Whitney U non-parametric test was used. PC5, PC6 and PC10 showed association to the traits of interest (Table 5.1). All other principal components and traits displayed no relationship ($P > 0.1$). It is noted that an association with PC6 recurs for average spherical equivalent and myopia status (which is derived from the former) and an association between axial length and PC5 and myopia status and PC5

Myopia status	PC5	1113/1176 (0.082)
Myopia status	PC6	1112/1180 (0.069)
AveSphTF3	PC6	-0.036 (0.085)
Axial length	PC5	0.051 (0.08)
Corneal curvature	PC10	0.069 (0.02)

Table 5.1) Model determination I (principal components). Summary of association between principal components and ocular traits. Either mean ranks of a Mann Whitney U test (myopia status; non-myope/myope) or Spearman rank correlations (AveSphTF3, axial length and corneal curvature) are listed with P values in brackets. Average spherical equivalent at age 15 (AveSphTF3).

Myopia status	1.21 (0.1)
AveSphTF3	1149/1099 (0.066)
Axial length	711/492 (5.5E-28)
Corneal curvature	656/509 (7.38E-14)

Table 5.2) Model determination II (gender). A relationship between gender and the traits of interest is examined. For myopia status the odds ratio of being a myope and female versus a myope and male is listed with a P value generated by a Chi square test in brackets. The three quantitative traits were examined using a Mann Whitney U test. Listed are the mean ranks for male/female with P values beside in brackets.

Myopia status	1124/1121 (0.943)
AveSphTF3	-0.014 (0.493)
Axial length	0.03 (0.297)
Corneal curvature	0.042 (0.155)

Table 5.3) Model determination III (age). A relationship between age at clinic and each trait is examined. For myopia status a Mann Whitney U test is used. Listed are the mean ranks for non-myopes/myopes with P values beside in brackets. For the three quantitative traits Spearman rank correlation test was used. Spearman rho is listed with P values in brackets.

is present (axial length is highly correlated with myopia status). Gender and age were also tested for association with the ocular traits. Gender displayed evidence of association (Table 5.2). Age does not show a relationship with any trait of interest (Table 5.3), which is supported by the study sample being drawn from a birth cohort. For a genome-wide association analysis it is important to correct for shared genetic ancestry and population stratification (Price et al., 2006). Therefore any principal components that showed mild association with the traits of interest were included as covariates (principal components are derived from genetic properties of the study sample). Also gender showed a mild (myopia status and average spherical equivalent) and strong (axial length and corneal curvature) relationship with the traits. Since it is possible some SNPs also show gender specific distributions (Payami et al., 2005), gender is included as a covariate. Age shows no association, largely due to a lack of variability in a birth cohort. The final models for each trait are listed Table 5.4.

5.2.3 Data checks: phenotype

The genome-wide association study sought to identify genes underlying common variation in four ocular traits; axial length, corneal curvature, refractive error (in this case average spherical equivalent) and presence or absence of myopia. Average spherical equivalent measured by non-cycloplegic autorefraction was used to analyse refractive error. Myopia was defined as -1 D or less on this scale. Thresholds were set to identify individuals with evidence of atypical pathology or extremely influential data points (Table 5.5). Individuals showing extreme values were removed prior to analysis for three reasons. Firstly such observations are by definition uncommon and therefore are not representative of the sample population. Also larger outliers have the potential to overly influence an effect size observed in the results of an association study. Thirdly some extreme outliers represent data errors accrued during data entry or during phenotype measurements.

Extreme outliers were identified as being four standard deviations (4SD) from a mean value and the trait value at 4SD is also listed. It is noted that the trait values at 4SD were considered large from a clinical point of view supporting the use of units of standard deviation to identify clinically unsuitable cases. An extremely small number

of observations were outside given thresholds indicating little loss of total power to detect a genetic association.

Trait	Covariates
Myopia status	PC5, PC6, Gender
AveSphTF3	PC6, Gender
Axial length	PC5, Gender
Corneal curvature	PC10, Gender

Table 5.4) Final model. List of covariates included in analysis of each ocular trait.

	Lower (4SD)	Upper (4SD)	N
R-L S	-3.08	3.11	34
R-L C	-2.12	2.01	35
Chg	-0.93	0.73	35
MeanAL	19.83	26.98	3
MeanCC	6.77	8.88	1
Astig	NA	4	37

Table 5.5) Outliers. Thresholds to exclude potentially overly influential data points (outliers) and clinically unusual observations. Lower/upper 4SD (the trait value four standard deviations from the mean in either direction, except for Astig values which are listed in dioptres), N (number of observations outside the threshold), R-L S (difference between right and left spheres), R-L C (difference between right and left cylinder), Chg (change in average spherical equivalent measured at two time points, the 15 and 11 year clinics, divided by the time between these clinics for each individual), MeanAL (mean of axial length in right and left eyes), MeanCC (mean of corneal curvature between right and left eyes), Astig (astigmatism in either eye with four dioptres being the upper limit for inclusions. Most individual have less than 1 dioptre astigmatism in this sample).

5.2.4 Data checks: analysis

3,222 individuals had genotypes and at least one phenotype available for analysis. Initially a set of 285,537 genotyped single nucleotide polymorphic (SNP) markers were analysed in PLINK. Then an imputed set for the same individuals was analysed in Mach2qtl for quantitative traits and Mach2dat for myopia status. This second set contained the original genotyped set plus SNPs in HapMap phase II, release 22, build 36. In total this gave 2,543,888 genotyped and imputed markers. In both sets the covariates analysed in each set were the same (see model determination above).

Moving from the first to second analysis required reformatting of both phenotype and genotype inputs (achieved in Perl and SPSS (version 16.0, SPSS Inc., Chicago)). To guard against any corruption or loss of data during transfer from one analysis to another, the number of individuals was checked in each phenotype file (3,222). Furthermore for genotyped markers, effect sizes and P values from both analyses with a P value under 10e-5 were compared (Tables 5.6a-d). If no systematic error was introduced during reformatting then these measures should be similar in each analysis. Between the first and second analysis, effect sizes and P values were similar with trivial differences (most likely due to slight differences between algorithms or choice of test statistic; Mach2qtl/2dat uses Chi square, PLINK a multivariate t distribution).

Since no large differences seem to exist between both types of analysis and the larger dataset incorporates all information contained in the smaller set, all further results refer to the larger set containing imputed and genotyped SNPs for 3,222 individuals. To further guard against systematic errors, output from the results file of each genome-wide association analysis which contains information of the mean and variance of the trait were compared to the mean and variance of each trait as generated by loading an earlier phenotype file (which was not formatted for entry into analysis) in SPSS. These two statistics (mean and variance) for each trait remained unchanged in both types of file (Figure 5.1).

Finally, the quantitative traits may resemble a normal distribution to make the assumptions of the test statistic used to generate P values. For each of the quantitative

Mach Output				PLINK	
Trait	Marker	OR	LRPvalue	OR	P value
MyopiaStatus	rs11745248	0.685	2.80E-05	0.685	4.09E-05
MyopiaStatus	rs12534172	1.701	1.62E-04	1.710	8.13E-05
MyopiaStatus	rs12744084	1.412	2.14E-04	1.425	8.57E-05
MyopiaStatus	rs1436093	1.430	7.94E-05	1.435	5.39E-05
MyopiaStatus	rs1843587	0.692	3.79E-05	0.696	7.51E-05
MyopiaStatus	rs1860094	0.725	6.80E-05	0.726	9.29E-05
MyopiaStatus	rs1934345	0.602	3.67E-05	0.606	9.83E-05
MyopiaStatus	rs2546968	1.355	1.66E-04	1.382	6.35E-05
MyopiaStatus	rs34583	0.547	6.02E-06	0.565	2.09E-05
MyopiaStatus	rs4145072	1.437	1.18E-05	1.423	1.46E-05
MyopiaStatus	rs4724206	1.612	3.05E-05	0.615	4.84E-05
MyopiaStatus	rs4745123	1.490	3.03E-05	1.484	2.70E-05
MyopiaStatus	rs4851079	0.711	2.67E-05	0.711	3.87E-05
MyopiaStatus	rs4946880	1.408	4.70E-05	1.401	5.42E-05
MyopiaStatus	rs7101596	0.705	2.80E-05	0.705	3.30E-05
MyopiaStatus	rs804134	1.391	9.92E-05	1.388	9.04E-05
MyopiaStatus	rs9521666	1.555	1.66E-06	1.562	1.16E-06

Table 5.6a) Data check I (myopia status). Effect sizes and P values for both datasets. The similarity highlights that no systematic error was introduced when moving between the smaller dataset of genotyped SNPs (PLINK) to the larger imputed set (Mach2dat). In some cases (for example marker rs1843587) a different reference allele was used by either program i.e. an A/C SNP could be analysed with either A or C as the reference allele. The effect then would indicate either an increase or decrease in myopia risk for those with the non-reference allele. In these cases the odds ratio was divided into one to facilitate comparison (SNPs rs11745248, rs1843587, rs1860094, rs1934345, rs34583, rs4724206, rs4946880, rs7101596, rs9521666). Odds ratio (OR), likelihood ratio P value (LRPvalue).

Mach Output				PLINK	
Trait	Marker	Effect	P value	Beta	P value
Mean Corneal Curvature	rs10500740	0.05	1.05E-04	-0.050	9.97E-05
Mean Corneal Curvature	rs10821278	-0.076	9.21E-06	0.076	7.97E-06
Mean Corneal Curvature	rs10861467	-0.046	1.48E-05	-0.045	1.77E-05
Mean Corneal Curvature	rs11112661	-0.043	2.58E-05	0.043	2.38E-05
Mean Corneal Curvature	rs12763439	-0.051	9.72E-05	-0.053	5.09E-05
Mean Corneal Curvature	rs1342761	-0.042	7.49E-05	-0.042	6.85E-05
Mean Corneal Curvature	rs1393350	0.045	2.31E-04	-0.047	9.75E-05
Mean Corneal Curvature	rs1402675	-0.048	3.49E-05	0.048	3.96E-05
Mean Corneal Curvature	rs153516	-0.049	1.31E-05	0.049	1.24E-05
Mean Corneal Curvature	rs1954343	0.047	4.45E-05	-0.047	3.71E-05
Mean Corneal Curvature	rs2156422	0.068	3.84E-05	0.068	3.75E-05
Mean Corneal Curvature	rs224218	-0.042	9.53E-05	-0.042	9.03E-05
Mean Corneal Curvature	rs2277481	-0.049	5.71E-05	-0.047	3.68E-05
Mean Corneal Curvature	rs2277483	-0.049	4.63E-05	-0.049	4.49E-05
Mean Corneal Curvature	rs2360090	-0.052	9.89E-05	0.052	6.26E-05
Mean Corneal Curvature	rs2432614	-0.047	1.32E-04	0.048	9.83E-05
Mean Corneal Curvature	rs2474373	0.063	5.93E-05	0.063	2.75E-05
Mean Corneal Curvature	rs4567028	0.054	5.83E-05	0.054	6.08E-05
Mean Corneal Curvature	rs4743942	-0.081	4.63E-06	0.081	4.31E-06
Mean Corneal Curvature	rs4743942	-0.081	4.63E-06	0.081	4.31E-06
Mean Corneal Curvature	rs649009	-0.048	5.25E-05	0.048	5.28E-05
Mean Corneal Curvature	rs6938321	-0.066	7.02E-05	0.067	5.44E-05
Mean Corneal Curvature	rs7280798	-0.068	3.91E-05	0.067	5.52E-05
Mean Corneal Curvature	rs7868385	0.075	9.28E-06	0.075	8.76E-06

Table 5.6b) Data check II (corneal curvature). Effect sizes and P values for SNPs analysed with corneal curvature as a phenotype in a large dataset (Mach2qtl) and a relatively smaller one (PLINK). Both are similar apart from trivial differences, indicating no systematic error when moving from small to large sets. Direction of effect (either positive or negative) may differ (e.g. SNP rs4743942) due to a different choice of reference allele.

Mach Output				PLINK	
Trait	Marker	Effect	P value	Beta	P value
Mean Axial Length	rs10093643	0.139	5.25E-05	0.141	4.13E-05
Mean Axial Length	rs1044429	0.229	1.23E-05	0.227	1.26E-05
Mean Axial Length	rs10502036	0.183	1.55E-05	0.184	1.40E-05
Mean Axial Length	rs10808622	-0.165	7.85E-05	-0.167	5.96E-05
Mean Axial Length	rs10895714	0.183	1.54E-05	0.188	1.07E-05
Mean Axial Length	rs1200618	-0.202	2.97E-06	0.205	2.18E-06
Mean Axial Length	rs12044963	0.254	4.98E-05	0.253	4.96E-05
Mean Axial Length	rs12410731	-0.170	1.13E-05	0.168	1.44E-05
Mean Axial Length	rs13232210	-0.197	3.02E-05	0.198	2.53E-05
Mean Axial Length	rs1424687	-0.135	4.07E-04	-0.161	4.09E-05
Mean Axial Length	rs17676175	0.211	1.40E-05	0.216	8.79E-06
Mean Axial Length	rs17741042	-0.143	2.03E-05	0.147	1.25E-05
Mean Axial Length	rs1834197	0.161	4.02E-05	-0.162	3.78E-05
Mean Axial Length	rs1860872	-0.154	5.24E-05	0.156	3.73E-05
Mean Axial Length	rs1983365	-0.145	1.86E-05	-0.142	2.71E-05
Mean Axial Length	rs2294394	-0.157	8.15E-05	-0.156	8.33E-05
Mean Axial Length	rs2350106	0.138	3.60E-05	0.137	4.26E-05
Mean Axial Length	rs2505515	0.159	2.10E-05	-0.156	2.96E-05
Mean Axial Length	rs528641	-0.168	8.63E-05	0.166	8.86E-05
Mean Axial Length	rs5762814	0.214	4.52E-05	-0.212	4.72E-05
Mean Axial Length	rs5762857	-0.220	3.12E-05	-0.224	2.48E-05
Mean Axial Length	rs5771104	-0.128	1.55E-04	0.132	9.81E-05
Mean Axial Length	rs639622	0.268	4.05E-05	0.268	3.96E-05
Mean Axial Length	rs6735865	0.214	7.69E-05	0.213	7.32E-05
Mean Axial Length	rs722354	0.183	1.44E-05	0.189	7.44E-06

Table 5.6c) Data check III (axial length). Effect sizes and P values for SNPs analysed with axial length as a phenotype in a large dataset (Mach2qtl) and a relatively smaller one (PLINK). Both are similar apart from trivial differences, indicating no systematic error when moving from small to large sets. Direction of effect (either positive or negative) may differ (e.g. SNP rs9309123) due to a different choice of reference allele.

Mach Output				PLINK	
Trait	Marker	Effect	P value	Beta	P value
AveSphTF3	rs10104895	0.227	1.87E-05	-0.229	1.62E-05
AveSphTF3	rs10509491	-0.144	7.24E-05	0.144	7.42E-05
AveSphTF3	rs10515122	-0.206	2.37E-05	0.204	2.61E-05
AveSphTF3	rs10925945	-0.272	2.44E-04	0.294	8.47E-05
AveSphTF3	rs1108079	0.149	9.40E-05	-0.151	7.97E-05
AveSphTF3	rs11759825	-0.143	9.25E-05	-0.143	8.38E-05
AveSphTF3	rs1266922	-0.154	7.36E-05	-0.154	7.50E-05
AveSphTF3	rs1498748	0.189	7.99E-06	-0.195	4.95E-06
AveSphTF3	rs165075	0.438	5.44E-06	-0.405	3.25E-05
AveSphTF3	rs1823759	0.176	3.12E-05	0.177	2.94E-05
AveSphTF3	rs1865375	-0.147	6.90E-05	0.148	6.48E-05
AveSphTF3	rs1949356	0.150	6.65E-05	0.146	8.89E-05
AveSphTF3	rs2388780	-0.300	4.00E-05	-0.306	2.62E-05
AveSphTF3	rs250306	-0.184	8.11E-05	-0.184	8.03E-05
AveSphTF3	rs2635351	0.151	4.90E-05	0.150	5.52E-05
AveSphTF3	rs2836760	-0.251	3.79E-05	-0.252	3.48E-05
AveSphTF3	rs2839650	-0.156	2.17E-05	0.159	1.39E-05
AveSphTF3	rs2964132	0.159	1.68E-05	-0.159	1.72E-05
AveSphTF3	rs3904668	0.148	5.40E-05	0.150	4.59E-05
AveSphTF3	rs4243949	-0.301	7.25E-05	0.301	6.75E-05
AveSphTF3	rs4685567	-0.155	2.67E-05	-0.155	2.58E-05
AveSphTF3	rs4851079	-0.133	2.64E-04	0.146	5.30E-05
AveSphTF3	rs687848	-0.148	4.95E-05	0.148	5.25E-05
AveSphTF3	rs734826	0.155	2.82E-05	-0.153	3.08E-05
AveSphTF3	rs7861755	0.151	3.89E-05	0.151	3.40E-05
AveSphTF3	rs7895270	0.171	6.54E-05	0.172	6.12E-05
AveSphTF3	rs926002	0.149	4.16E-05	0.150	3.96E-05
AveSphTF3	rs9297026	-0.144	8.73E-05	-0.143	8.76E-05
AveSphTF3	rs998639	0.152	1.10E-04	0.154	8.81E-05

Table 5.6d) Data check IV (average spherical equivalent). Effect sizes and P values for SNPs analysed with average spherical equivalent (AveSphTF3) as a phenotype in a large dataset (Mach2qtl) and a relatively smaller one (PLINK). Both are similar apart from trivial differences, indicating no systematic error when moving from small to large sets. Direction of effect (either positive or negative) may differ (e.g. SNP rs2964132) due to a different choice of reference allele.

Statistics

		AveSphTF3	AxialLength	CornealCurvature
N	Valid	2246	1195	1159
	Missing	976	2027	2063
	Mean	-.385213	23.433874	7.826747
	Variance	1.49705	0.73785	0.06810

SPSS MyopiaStatus

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Non-Myope	1872	58.1	83.3	83.3
	Myope	374	11.6	16.7	100.0
Missing	System	976	30.3		
Total		3222	100.0		

MACH OUTPUT

FITTED MODELS (for covariate adjusted residuals)

==========

Trait	Raw Mean	Raw Variance
AveSphTF3	-0.38521	1.49705
MeanAxialLength	23.43387	0.73785
MeanCornealCurvature	7.82675	0.06810

MyopiaStatus SAMPLE-SIZE 374 cases 1872 controls

Figure 5.1) Data check V (phenotype file). A master phenotype file (from which phenotype files for analysis were generated) was analysed in SPSS (top). Information on mean and variances of each trait is given in the genome-wide association analysis programs Mach2qtl and Mach2dat (bottom). The two are compared to make sure no error was generated when creating input files for analysis from the master file. No error is apparent for the quantitative traits. Also the number of valid cases and controls in both the master file and the Mach output are the same. Since the means and variances are identical to the fifth decimal place (the last number printed in Mach output) this indicates the same number of subjects was present in both master phenotype file and input files. Similarly for myopia status, there are a total of 3,222 observations (including missing values) in both SPSS and Mach2dat files (1872 controls and 374 cases).

traits, deviations from normality were examined and it was observed that average spherical equivalent displayed large deviations, while both ocular biometric traits (axial length and corneal curvature) showed slight deviations. To guard against false positive signals in association results, P values were generated for traits both untransformed and transformed to a ranked normal distribution (using a Mach2qtl option --quantile normalisation). This type of transformation ranks data points in a trait and transforms these ranks into a normal distribution. Association signals which are driven by increased numbers of observations in a tail of an untransformed distribution relative to a normal distribution (and therefore have larger than expected deviations from a mean value) will be mediated against. However at the same time the relative order of trait values remains unchanged. Therefore a change in a trait value occurring with a particular allele is maintained (i.e. allele X displays a significantly higher mean than allele Y). In general, P values for SNPs tested for association with corneal curvature and axial length using either transformed or untransformed data were similar and any SNPs with low P values which will be investigated later on in this study are quoted with both P values. For average spherical equivalent, which displayed larger deviation from normality, there were more noticeable and frequent differences between P values generated with transformed and untransformed trait values (Table 5.7). Since the distribution of average spherical equivalent leads to false positive signals, for all further results, P values for a transformed quantile normal average spherical distribution were used. P values generated with an untransformed distribution are only shown for comparison.

The test statistic for myopia status is based on a proportion of affected/unaffected which uses a binomial distribution which in turn is normally distributed at sufficiently large numbers. It is noted however that there is a choice of two statistics generated by the genome-wide association analysis, the Wald statistic or the likelihood ratio test and in the case of empty cells (typically driven by low minor allele frequency) the Wald statistic gave more moderate P values. However the Wald statistic is known to give higher P values when a large value of a regression coefficient is observed (Norušis and SPSS Inc., 2003) leading to an increased chance of a false negative result. Given this drawback of the Wald statistic, the likelihood ratio test is used in all further results, but for SNPs with a low P value, the Wald generated P value is also quoted.

Trait	Marker	P value	P valueQT
AveSphTF3	rs705380	8.0E-08	7.7E-06
AveSphTF3	rs12122818	6.6E-08	1.1E-03
AveSphTF3	rs2429095	5.1E-12	1.3E-03
AveSphTF3	rs2673046	1.5E-09	4.9E-05

Table 5.7) Non-normality. Differences between P values for SNPs (all happen to be imputed) tested for association with transformed (P valueQT) and untransformed (P value) trait values of average spherical equivalent (AveSphTF3). The increase in untransformed P values is most prominent for SNPs with a minor allele frequency (MAF) below 1% (SNPs rs12122818, rs2429095, rs2673046) but still large for a SNP with a MAF of 4% (rs705380). The transformed trait gives more moderate P values in each case, indicating that the signals are been driven mainly by deviations from normality.

Furthermore all SNPs with a frequency of below 1% or above 99% (MAF < 1%) were removed from any further analysis and results. Low minor allele frequencies lead to biased mean values due to a greatly reduced number of observations for an allele. Also since imputed SNPs were tested for association with each trait, only SNPs with a r square value (an indication of imputation quality) of above 0.3 were retained for examination. 0.3 is the recommended value to flag badly imputed SNPs without losing many (1%) of well imputed SNPs (http://www.sph.umich.edu/csg/yli/mach/tour/imputation.html). Out of the 2,543,888 SNPs in the full dataset, 56,418 displayed a MAF of 1% or less and 60,353 displayed an r square of 0.3 or less. This left a total of 2,427,117 SNPs for further analysis.

5.2.5 Plots

To examine interesting signals, plots of their genomic regions were drawn. These plots were drawn using a modified R script (http://www.broadinstitute.org/science/projects/diabetes-genetics-initiative/plotting-genome-wide-association-results (Saxena et al., 2007)). Information except linkage disequilibrium (LD) measures was obtained from the site given above. LD measures

were generated by downloading SNP genotype data from the HapMap website and loaded in Haploview (http://www.broadinstitute.org/haploview/haploview (Barrett et al., 2005)). Parameters for LD data were as follows; CEU population, HWE P value cut off, 0.001, minimum percentage genotypes available, 75%, maximum number of Mendelian errors, 1, minimum minor allele frequency, 0.001. A Perl script was used to match LD measures to genome-wide analysis output.

Quantile quantile plots (QQplot) were drawn for each trait using a modified R script (Saxena et al., 2007). Briefly, an observed P value distribution transformed to a minus log with base ten scale and sorted from high to low was plotted against a distribution with the same number of observations as the observed set with data points evenly distributed, also on a minus log scale.

5.3 Results

Table 5.8 lists a brief description of number of valid observations, mean, variance, maximum, minimum and tests of normality for three quantitative traits and frequency of cases and controls for myopia status. These values give an indication of the size of the datasets available for analysis and their general statistical properties. For tests of normality used, when sample sizes are large even small deviations from normality can lead to low P values (Norušis and SPSS Inc., 2003). Myopia status was defined as cases with an average spherical equivalent of less than or equal to minus one dioptre and controls with more than minus one dioptre. This was indicated in Chapter 3 as being an efficient point to capture the most true myopes/non-myopes. Briefly 4% of non-myopes would be indicated as myopes and 11% of myopes would be indicated as non-myopes (specificity and sensitivity of 96% and 89%).

QQ plots (Figure 5.2) show a mild increase in P values for myopia status. For axial length, there are an elevated number of P values between 10e-5 and 10e-4 relative to higher P values, but this relationship reverses under 10e-5. Figure 5.3 shows association results for myopia status and rs9521666 (P = 2 X 10e-6) a genotyped SNP (MAF 21%) with the lowest P value, which is located on chromosome 13q34 in an intron (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=9521666) of the collagen type IV alpha 1 gene (COL4A1, OMIM # 120130). The next lowest P value for a genotyped SNP (rs34583, P = 6 X 10e-6) is shown in Figure 5.4. The genotyped SNP is located beside three imputed SNPs one of which reaches a P value of 6 x 10e-8 (SNP rs12817923), however the low P value may be driven by a reasonably low MAF (6%). These SNPs are located at 12q23.1. The nearest gene is ETS-domain-protein (ELK3, OMIN # 600247) at 12q.23, which is a transcription factor. The next lowest P value for a genotyped SNP and myopia status is 1 X 10e-5 (data not shown).

The two genotyped SNPs with the lowest P values for an association with axial length are shown in Figures 5.5 and 5.6. SNP rs1200618 reaches a P values of 3 X 10e-6 and is located at 11q22.3, while SNP rs12410731 (1 X 10e-5) is located at 1q41. The next lowest P value for a genotyped SNP was 1.2 X 10e-5 (rs1044429, data not shown). The lowest P value observed for an association with corneal curvature was for SNP rs4743942 (P = 5 X 10e-6) (Figure 5.7). This SNP is located at 9q22.32, over two

	AveSphTF3	Axial length	Corneal curvature
Valid	2246	1195	1159
Missing	976	2027	2063
Mean	-0.39	23.43	7.83
Variance	1.50	0.74	0.07
Minimum	-10.00	20.49	7.13
Maximum	7.25	26.56	8.67

	Tests of Normality					
		Kolmogorov-Smirnova		Shapiro-Wilk		
	Statistic	df	P value	Statistic	df	P value
AveSphTF3	0.16	2246	1.16E-165	0.81	2246	8.17E-46
Axial length	0.03	1195	0.009	0.99	1195	4.35E-05
Corneal curvature	0.03	1159	0.004	1.00	1159	0.010

		Frequency	Percent	Valid Percent
Valid	Non-myope	1872	58.1	83.3
	Myope	374	11.6	16.7
Missing	System	976	30.3	
Total	3222		100	100

Table 5.8) Descriptive statistics. Descriptive statistics of phenotype data (top), tests of normality for three quantitative traits (Kolmogorov-Smirnova and Shapiro-Wilk, middle panel) and frequency of cases and controls for myopia status (bottom panel). AveSphTF3 (average spherical equivalent), axial length (refers to the average axial length of right and left eyes), corneal curvature (refers to the mean corneal curvature of right and left eyes).

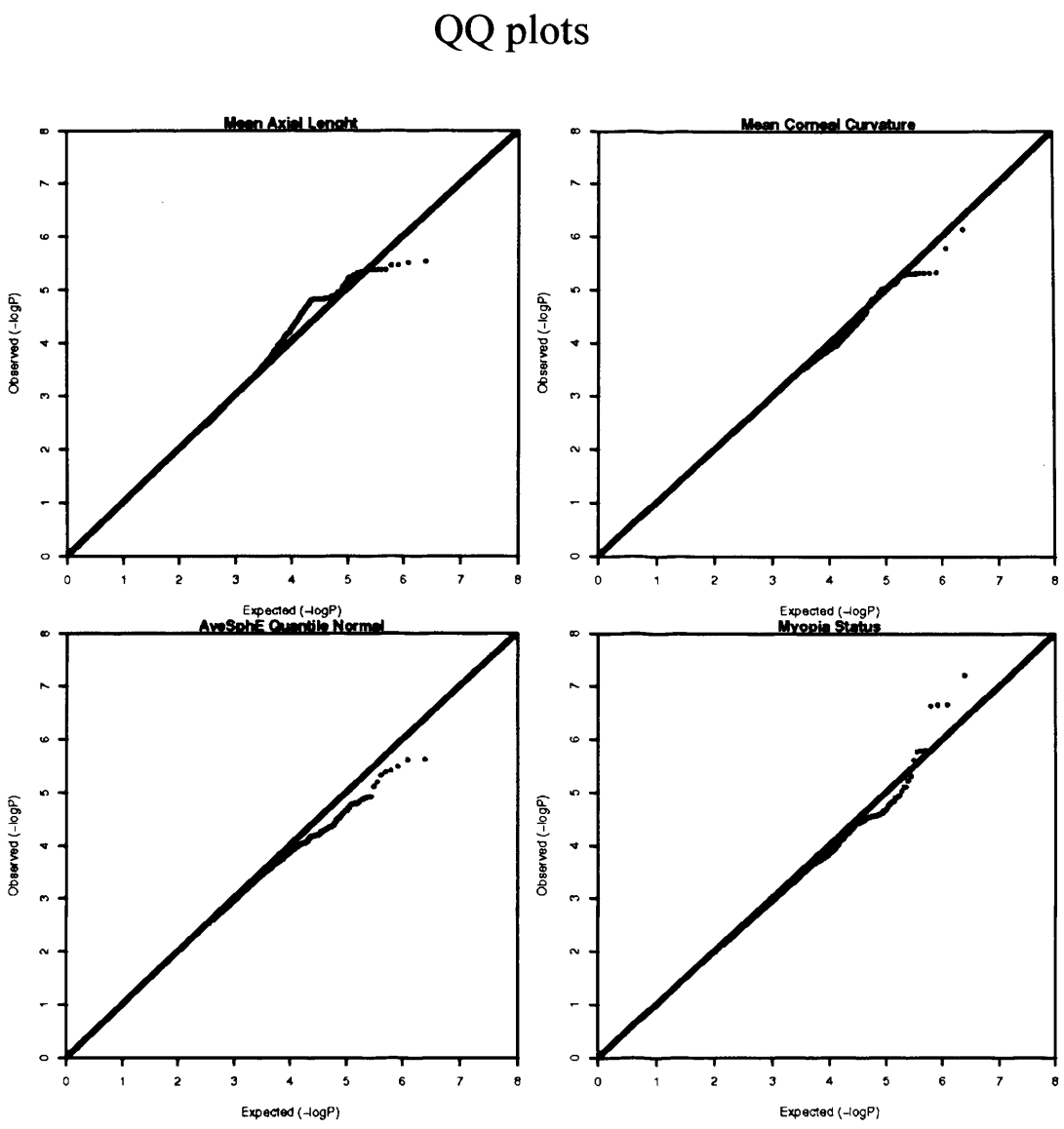
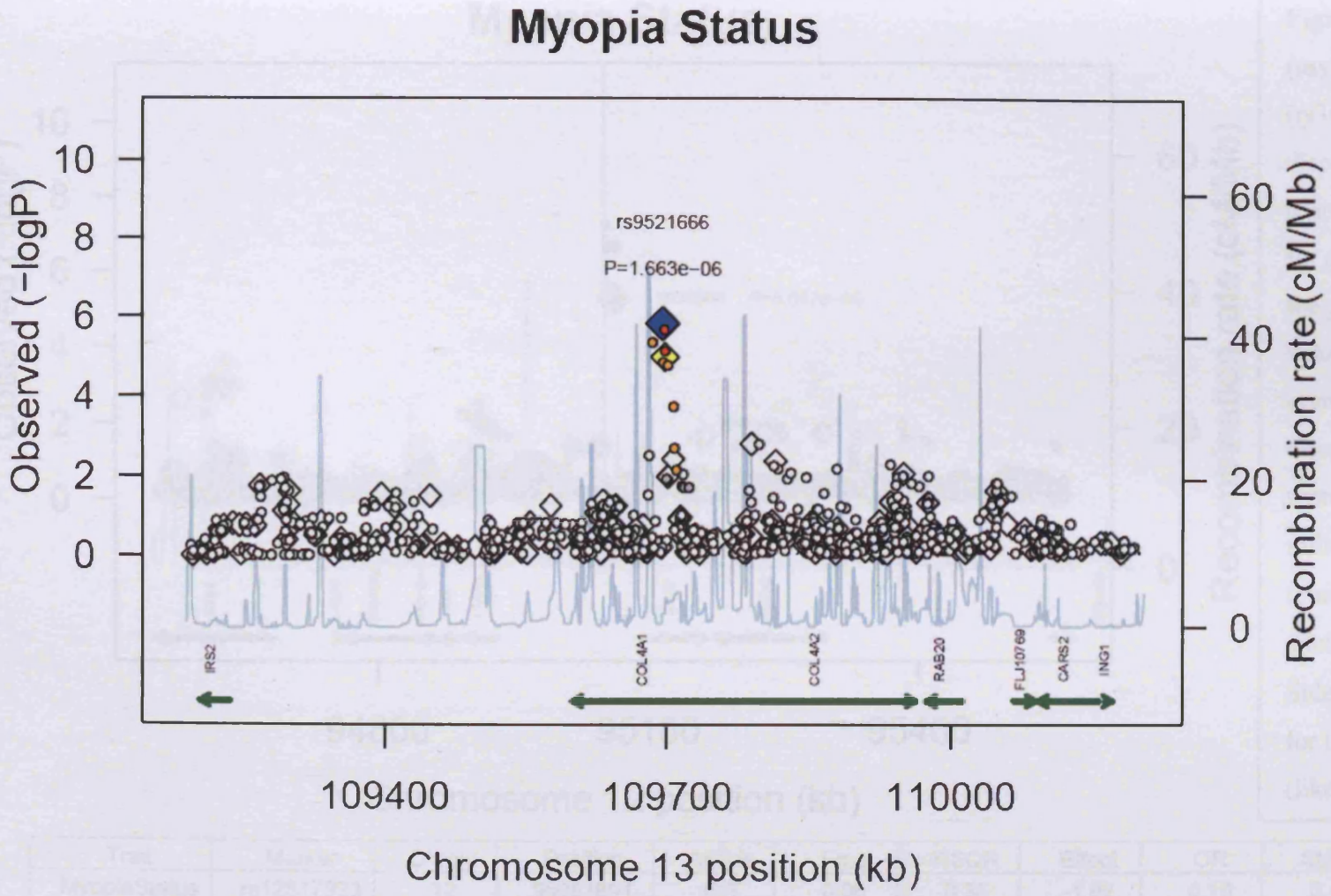


Figure 5.2) Quantile quantile plots. One plot for each trait was drawn in R using a modified script (Saxena et al., 2007). A blue line represents two distributions that have equal number of observations in each quantile (a quantile refers to a definite portion of a distribution). A black line describes a relative increase (above blue line) or decrease (below blue line) of observed P values relative to the expected distribution. Moving upwards along the blue line, P values decrease. Principal components (PC) were included in analyses of the four traits if they were associated. For the observed P values given in the Q-Q plots above, PC5 and PC6 were included in the analysis of myopia status, PC5 included with axial length, PC6 in an analysis of average spherical equivalent and PC10 included with corneal curvature.

unknown genes C9orf102 (http://www.ncbi.nlm.nih.gov/gene?term=C9orf102) and Loc375748 (http://www.ncbi.nlm.nih.gov/protein/EAW92637.1). The next lowest P value was 1 X 10e-5 for an association between a genotyped SNP and corneal curvature (data not shown). The genotyped SNP, rs1823759 displayed the lowest P value for a transformed average spherical equivalent (P = 4 X 10e-6) (Figure 5.8). This SNP is located at 18q12.1, over nucleolar-localised protein gene (NOL4, OMIM # 603577), which codes for a protein involved in nuclear localisation.



Trait	Marker	Chrom	Position	Alleles	Freq	RSQR	OR	Stderr	Wald P	LRPvalue
MyopiaStatus	rs9521666	13	109695445	A,G	0.79	0.994	0.64	0.09	1.13E-06	1.66E-06

Figure 5.3) Association result I (myopia status). A genotyped SNP (rs9521666) with the lowest P value is displayed in blue. Genotyped SNPs are represented as diamonds. LD is displayed for all SNPs (pairwise comparisons with the genotyped SNP with the lowest P value); colour coded white, yellow/grey, orange and red (with increasing linkage disequilibrium). All imputed SNPs are represented as circles. One megabase (a million bases) is shown. Below, Chrom (chromosome), Freq (frequency of one allele), RSQR (r square), OR (odds ratio), Stderr (standard error), Wald P (P value for the Wald statistic), LRPvalue (likelihood ratio P value).

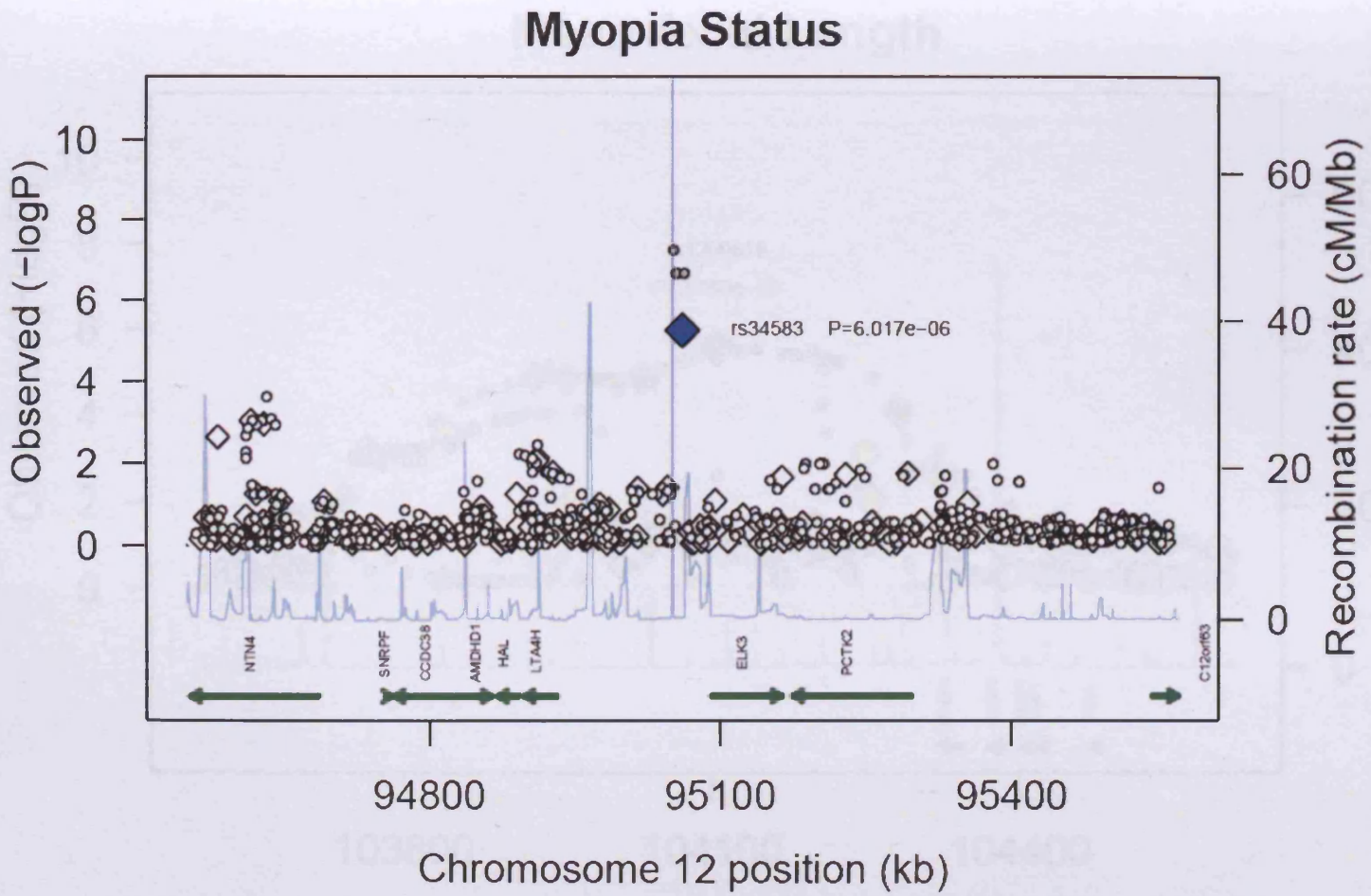


Figure 5.4) Association result II (myopia status). A genotyped SNP (rs34583) with the lowest P value is displayed in blue. Genotyped SNPs are represented as diamonds. LD is displayed for all SNPs (pairwise comparisons with the genotyped SNP with the lowest P value); colour coded white, yellow/grey, orange and red (with increasing linkage disequilibrium). All imputed SNPs are represented as circles. One megabase (a million bases) is shown. Below, Chrom (chromosome), Freq (frequency of one allele), RSQR (r square), OR (odds ratio), Stderr (standard error), Wald P (P value for the Wald statistic), LRPvalue (likelihood ratio P value).

Trait	Marker	Chrom	Position	Alleles	Freq	RSQR	Effect	OR	Stderr	Wald P	LRPvalue
MyopiaStatus	rs12817923	12	95051891	C,T	0.06	0.35	-1.92	0.15	0.40	1.68E-06	6.10E-08
MyopiaStatus	rs12424333	12	95054741	A,T	0.94	0.52	1.48	4.38	0.32	3.40E-06	2.18E-07
MyopiaStatus	rs4762272	12	95055704	A,G	0.06	0.52	-1.47	0.23	0.32	3.42E-06	2.21E-07
MyopiaStatus	rs34583	12	95061269	A,G	0.86	0.93	0.60	1.83	0.14	2.08E-05	6.02E-06

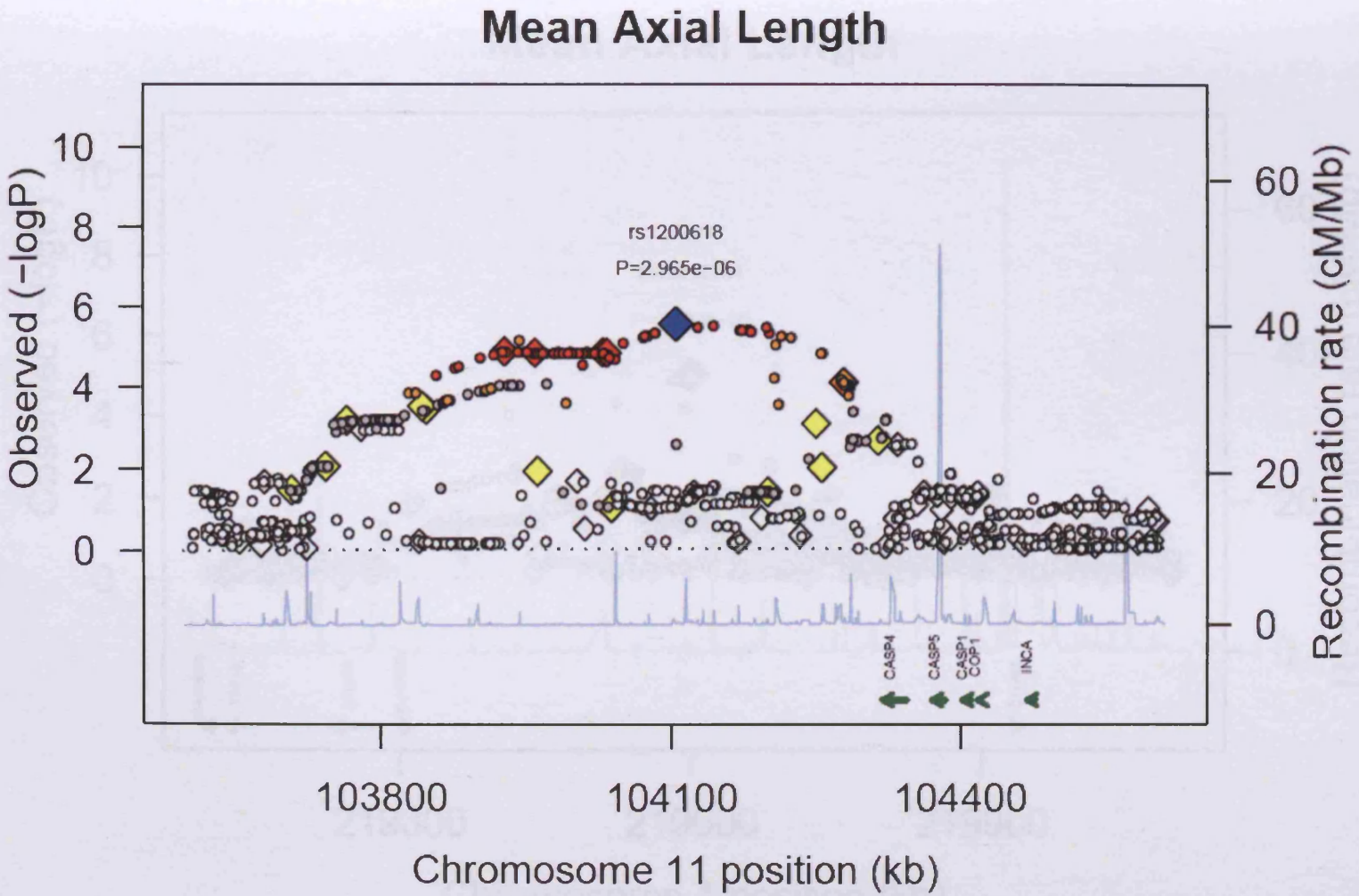


Figure 5.5) Association result III (axial length). A genotyped SNP (rs1200618) with the lowest P value is displayed in blue. Genotyped SNPs are represented as diamonds. LD is displayed for all SNPs (pairwise comparisons with the genotyped SNP with the lowest P value); colour coded white, yellow/grey, orange and red (with increasing linkage disequilibrium). All imputed SNPs are represented as circles. One megabase (a million bases) is shown. Below, Chrom (chromosome), Freq (frequency of one allele), RSQR (r square), Stderr (standard error), PvalueQT (P value for quantile normal trait).

Trait	Marker	Chrom	Position	Alleles	Freq	RSQR	Effect	Stderr	P value	PvalueQT
Mean Axial Length	rs1200618	11	104105095	A,G	0.197	0.999	-0.202	0.043	2.97E-06	4.74E-06

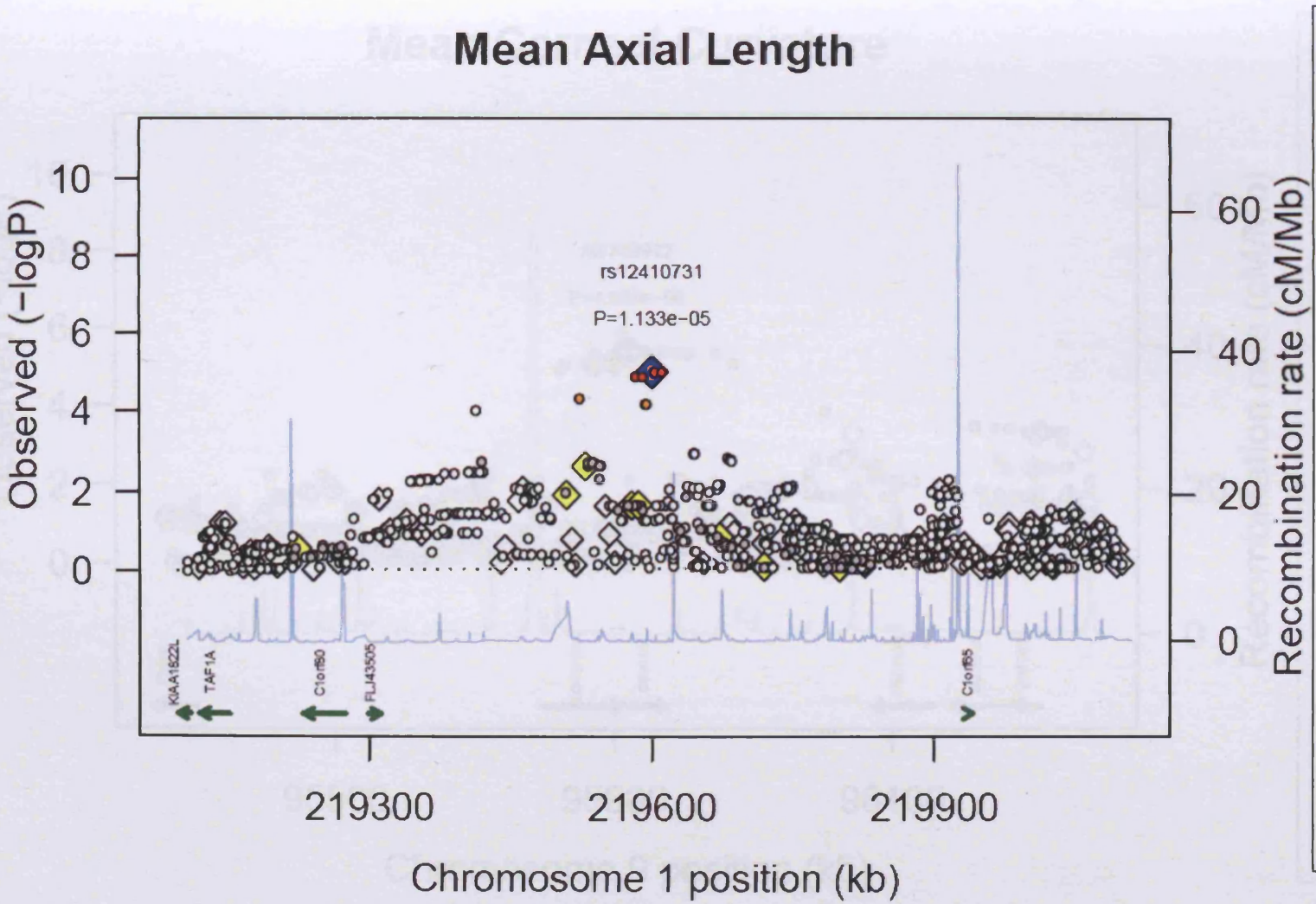


Figure 5.6) Association result IV (axial length). A genotyped SNP (rs12410731) with the lowest P value is displayed in blue. Genotyped SNPs are represented as diamonds. LD is displayed for all SNPs (pairwise comparisons with the genotyped SNP with the lowest P value); colour coded white, yellow/grey, orange and red (with increasing linkage disequilibrium). All imputed SNPs are represented as circles. One megabase (a million bases) is shown. Below, Chrom (chromosome), Freq (frequency of one allele), RSQR (r square), Stderr (standard error), PvalueQT (P value for quantile normal trait).

Trait	Marker	Chrom	Position	Alleles	Freq	RSQR	Effect	Stderr	P value	PvalueQT
Mean Axial Length	rs12410731	1	219599973	C,T	0.254	0.999	-0.170	0.039	1.13E-05	1.08E-05

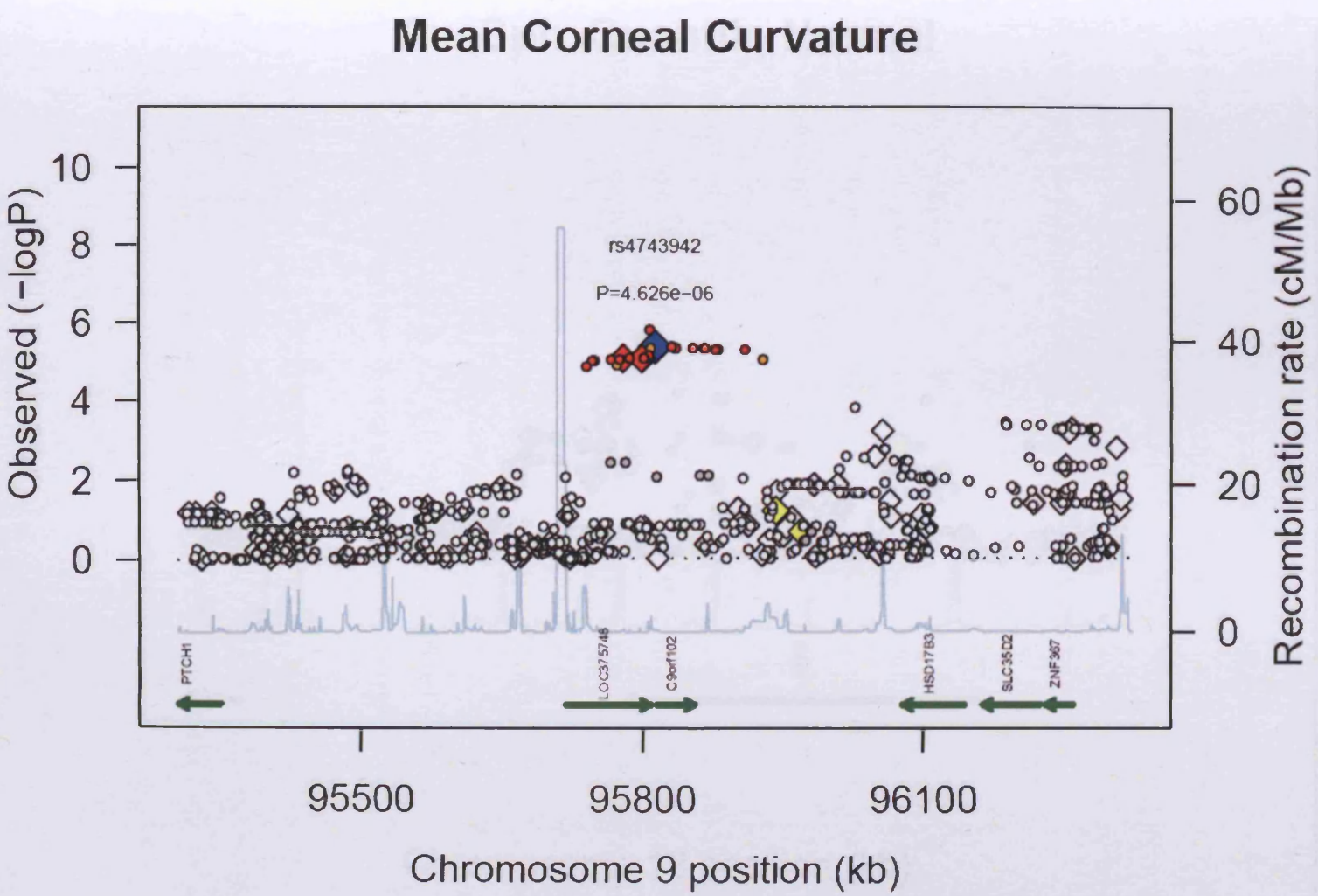


Figure 5.7) Association result V (corneal curvature). A genotyped SNP (rs4743942) with the lowest P value is displayed in blue. Genotyped SNPs are represented as diamonds. LD is displayed for all SNPs (pairwise comparisons with the genotyped SNP with the lowest P value); colour coded white, yellow/grey, orange and red (with increasing linkage disequilibrium). All imputed SNPs are represented as circles. One megabase (a million bases) is shown. Below, Chrom (chromosome), Freq (frequency of one allele), RSQR (r square), Stderr (standard error), PvalueQT (P value for quantile normal trait).

Trait	Marker	Chrom	Position	Alleles	Freq	RSQR	Effect	Stderr	P value	PvalueQT
Mean Corneal Curvature	rs4743942	9	95813581	A,G	0.0957	0.9999	-0.081	0.018	4.63E-06	3.09E-06

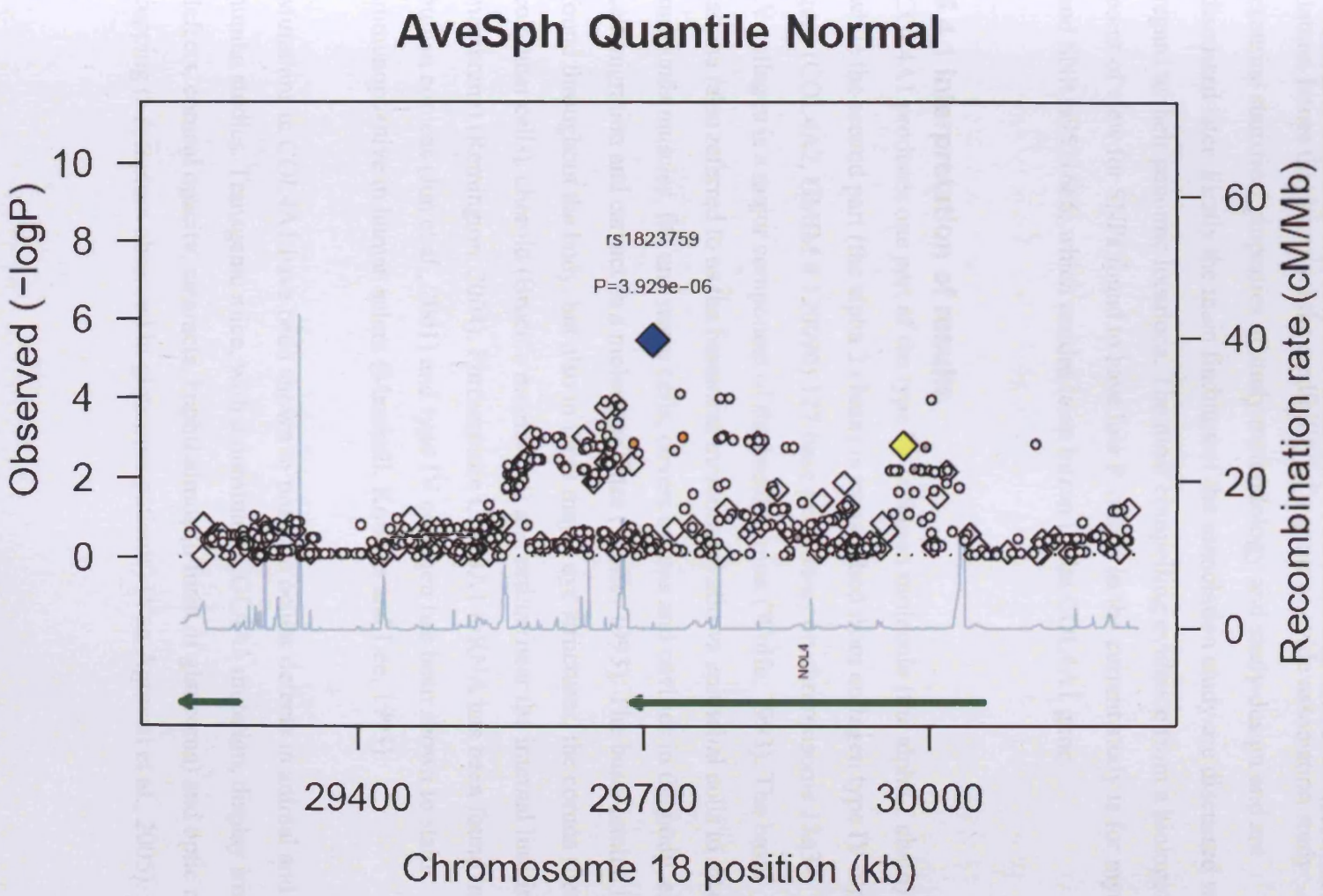


Figure 5.8) Association result VI (average spherical equivalent). A genotyped SNP (rs1823759) with the lowest P value is displayed in blue. Genotyped SNPs are represented as diamonds. LD is displayed for all SNPs (pairwise comparisons with the genotyped SNP with the lowest P value); colour coded white, yellow/grey, orange and red (with increasing linkage disequilibrium). All imputed SNPs are represented as circles. One megabase (a million bases) is shown. Below, Chrom (chromosome), Freq (frequency of one allele), RSQR (r square), Stderr (standard error), PvalueQT (P value for quantile normal trait).

Trait	Marker	Chrom	Position	Alleles	Freq	RSQR	Effect	Stderr	P value	PvalueQT
AveSphTF3	rs1823759	18	29709947	A,G	0.75	1.00	0.18	0.04	3.12E-05	3.93E-06

5.4 Discussion

The genetics of myopia, refractive error and a number of their ocular determinants (axial length and corneal curvature) have been investigated in the ALSPAC cohort via a genome-wide association study. A number of the SNPs with the lowest P values identified in this study were within or close to known genetic elements. A genome-wide association study is hypothesis free (i.e. each genetic marker is assumed to have an equal chance of harbouring a causative mutation). Therefore it is exploratory in nature. Issues that surround the credibility of a genome-wide association study examine statistical properties of study methodology and study design and are discussed later. Firstly the main findings of the association study are discussed with regard to their genomic locations. The most compelling evidence from a biological point of view for SNPs found to have low P values in the current study is for myopia and SNP rs9521666, which resides in an intron in the COL4A1 gene.

5.4.1 Interpretation of results

COL4A1 produces one part of the type IV collagen molecule (the alpha 1 chain) while the second part (the alpha 2 chain) is transcribed from collagen type IV alpha 2 gene (COL4A2, OMIM # 120090) 127 base pairs away on chromosome 13q34. Type IV collagen is a major component of the basal lamina (Wolfe, 1993). The basal lamina (also referred to as the basement membrane) allows epithelial cells to attach, surrounds muscles, fat, and nerve cells, covers organs and cavities in the body, allows cell migration and can act as a molecular filter (Wolfe, 1993). The basal lamina is found throughout the body, but also in three major eye structures, the cornea (beside columnar cells), choroid (Bruch's membrane) and retina (near the internal limiting membrane) (Remington, 2004). Furthermore COL4A1 mRNA has been found in human corneas (Jun et al., 2001) and type IV collagen has been shown to stain immunopositive in human sclera (Marshall, Konstas and Lee, 1993).

Mutations in COL4A1 have been shown to produce ocular defects in animal and human studies. Transgenic mice, with a dominant COL4A1 mutation, display iris defects, corneal opacity, cataracts, buphthalmos (a form of glaucoma) and optic nerve cupping (a symptom observed in glaucoma patients) (Van Agtmael et al., 2005). In

another study, mice with heterozygous deletions in COL4A1, displayed ocular anterior segment dysgenesis (ASD, abnormal development of the anterior region of the eye), optic nerve hypoplasia, buphthalmos, and varying levels of intraocular pressure (IOP) compared to control mice. It was also found that the degree of severity of symptoms varied with genetic background. Mutant mice with a C57BL/6J genetic background could be rescued by crossing with either 129/SvEvTac or CAST/EiJ mice. The authors found a locus on chromosome one that segregated with ASD rescue. The symptoms of rescued mice were not completely absent, but much milder. Rescued mice displayed slightly enlarged anterior chambers, a feature of myopia (Gould et al., 2007). That the COL4A1 ocular phenotype in mice is modified by other loci suggests that COL4A1 is part of a complex biological pathway that may exhibit some redundancy.

There is also evidence indicating COL4A1 in disorders of the human eye. Brain small vessel disease with Axenfeld-Rieger anomaly has clinical symptoms of cerebral vasculopathy, congenital cataract, congenital glaucoma, microcornea (where the cornea and anterior segment of the eye are smaller than normal), amblyopia and retinal detachment. This disorder has been shown to cosegregate with a dominant mutation in the COL4A1 gene (Sibon et al., 2007). Mutations in the COL4A1 gene leading to ocular defects (reduced visual acuity, amblyopia, retinal detachment, corneal opacity, changed IOP, microcornea, glaucoma, cataract and myopia among others) have been recently found in two other families (Coupry et al., 2010).

The above evidence of a role for COL4A1 in maintaining eye health is further supported by the general role of collagen in myopia development. Myopia elicits numerous changes in the collagen content of the eye. Electron microscopy examination of the sclera has shown differences between myopic and normal human eyes, including a reduction in size and dispersion of collagen molecules (Curtin et al., 1979). In the mammalian model of myopia (tree shrew), myopia is associated with reduced collagen content and size (McBrien, Cornell and Gentle, 2001). This is further supported by downregulation of collagen type I mRNA (Gentle et al., 2003). Furthermore myopia is known to be associated with glaucoma (Attebo et al., 1999), cataract (Younan et al., 2002) and less often microcornea (Sohajda et al., 2006), all of

which are features of disruption to the COL4A1 gene in humans (Sibon et al., 2007; Coupry et al., 2010).

Although such observations are compelling, the requirement of seeking replication to further investigate nominal P values is valuable. There is a possibility that evidence from animal studies is confounded by large amounts of genetic divergence between mice and humans. Similarly mutations that lead to gross changes in human phenotypes may not be responsible for variation in milder versions of the similar phenotypes.

In the current study evidence of association is observed in a region close to a gene but not within an exon (a DNA sequence that is expressed in a protein). SNP rs1200618 which shows the lowest P value in a test of association with axial length is located in an area with no known genes. However the region is an area of low recombination (Figure 5.5) and it can be hypothesized that the SNP is in linkage disequilibrium with a genetic element that has not been genotyped or imputed in the current study. Similarly SNP rs34583 (myopia status on chromosome 12, Figure 5.4) is located upstream of ELK3, a transcription factor. An imputed SNP (rs12817923) close to SNP rs34583 reaches a P value of 6 X 10e-8. It is suggested the threshold for declaring significance for a genome-wide association study is a P value 5 X 10e-8 (Risch and Merikangas, 1996; Hirschhorn and Daly, 2005). SNP rs12817923 is an imputed SNP and therefore depends partly on the available information on recombination and haplotype diversity used to infer genotypes at its location in the study sample. However it is possible to genotype this SNP in the sample to ascertain whether such strength of association will be observed after genotyping.

5.4.2 Future directions

In the current study the lowest P value for a genotyped SNP was 2 X 10e-6 with an effect size of 1.6 (rs9521666, Myopia status). Although the P value was low it did not reach the level expected for a SNP with a 5% chance of error. Rice et al. (Rice et al., 2008) point out that a genome-wide association test that involves 500,000 SNPs, at random, the expected number of effect sizes with a P value of less that 0.05 would be 25,000. Based on empirical observations and theoretical considerations (Dudbridge

and Gusnanto, 2008; Risch and Merikangas, 1996) a P value of 5.5 X 10e-8 would maintain a 5% family-wise error rate for a genome-wide association study. However, to observe a P value of low magnitude, an effect size of very large magnitude would need to be observed in a sample of moderate size. It is predicted that common variants underlying common disease will show only moderate effect sizes (Reich and Lander, 2001). This prediction has been found to be largely the case for genome-wide association studies of common diseases, with moderate to small effect sizes of 1.1 to 1.5 found in a majority of studies (Altshuler et al., 2008).

A P value is a statistic that measures the chance of finding an effect size at random *confounded by sample size*. As sample size increases, the precision of an effect size increases leading to more confidence that there is a statistical difference between groups based on the presence of an allele, which is reflected in a diminishing P value.

To observe a P value of 5 X 10e-8 for an allele with a frequency of 15% and a moderate effect size of 1.25 approximately 6,000 cases and 6,000 controls would be required. It is also noted that to achieve a P value that provides suggestive evidence of association would require 1,200 cases and 1,200 controls (Hirschhorn and Daly, 2005). In the current study there are comparable numbers of subjects for analysis of average spherical equivalent (N = 2,246). However for axial length and corneal curvature a smaller number of observations were available (N = 1,195 and N = 1,159 respectively). Similarly there was genotype and phenotype information available for 2,246 subjects to investigate genes underlying myopia but only 17% (n = 374) of these were cases.

The ALSPAC cohort was designed to investigate how genes and the environment influence health and development (Golding et al., 2001). High throughput genotyping is planned for the majority of subjects (approximately 10,000 genomes) which will allow the problem of moderate effect sizes and multiple testing to be partly circumvented. However in the case of refractive error the number of phenotypic measures is slightly lower (approximately 5,000) due to measures necessarily been taken when participants were at an age (15) when juvenile onset myopia is thought to have stabilised (Zadnik and Mutti, 2006). Similarly the number of phenotypic measures of axial length and corneal curvature are lower. However it is generally

accepted that simply increasing sample size *ad nauseum* is not a completely efficient method to provide evidence of association between a genetic element and a common disease. To avoid complications associated with spurious results generated by multiple testing and large recruitment and genotyping costs, efforts to map genes underlying common diseases have instead focussed on replication of findings in independent cohorts.

Replication refers to testing whether an association signal between a SNP and trait of interest shows evidence of a relationship when examined in an independent cohort. This strategy is an example of measuring an effect *under changing conditions* (it is noted that the null hypothesis has not changed, nor the method, genome-wide association, used to test it, in a replication study, but rather the data collection and storage and genotyping platforms may differ between cohorts). An association signal can be driven by artefacts introduced during measurement of the trait and genotyping or researcher error during data collection and storage. It is has been noted that seeking replication in an independent cohort can be an opportunity to sieve out such artefacts (McCarthy et al., 2008) as there will be differences in the study protocols and possibly between genotyping assays. It is noted in the same paper that replication of an association signal in an independent cohort should be considered more valid if the association signal is observed for the same genetic marker or haplotype or proxy to either. Seeking replication in an independent cohort also can allow researchers to undertake a meta-analysis. When data collected in different genome-wide association studies are comparable, data can be pooled to increase power to detect a moderate effect size (McCarthy et al., 2008). This strategy was recently employed to identify a common polymorphism that underlies normal variation in central corneal thickness in five Caucasian genome-wide association studies (Lu et al., 2010) (approximately 5,000 subjects).

The aim of the current study was to find reliable evidence of a relationship between a genetic element and refractive error, myopia and components important in the development of both, namely axial length and corneal curvature. There is suggestive evidence of association for a number of SNPs. Therefore replication is being sought in a number of cohorts with comparable genome-wide data. These cohorts are TwinsUK at the Department of Twin Research & Genetic Epidemiology, Kings College

London, UK, the Australian Twin Study at the Queensland Institute of Medical Research, Brisbane, Australia and the Lions Eye Institute, University of Western Australia, Perth, Australia and a number of isolate populations coordinated at the MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK. It is noted that the mean age of participants was 15 years and therefore the time required for adult onset myopia to develop is absent. Therefore it is reasonable to state that a genome-wide association study for juvenile onset and early onset myopia and refractive error has been undertaken, along with a genome-wide study of normal variation in axial length and corneal curvature in a cohort of young adults.

Chapter 6

Birth Order and Myopia

6.1 Introduction

The aim of this study was to investigate a relationship between birth order (the order of pregnancies preceding the current pregnancy, the first pregnancy confers a birth order of one) and myopia present at age eleven. Data in this chapter is from two diverse sources; a) ALSPAC, when children were aged 11 and b) the International High Myopia Genetics Consortium (IHMGC), a high myopia cohort. The study seeks to identify if an association is present in either or any of the two cohorts studied and in the ALSPAC cohort to further examine factors that may be related to a birth order myopia relationship.

Events before birth are known to influence health and development. An example of a relationship between factors before birth and health later in life is maternal phenylketonuria. Mothers are homozygous for a gene causing phenylketonuria (PKU) a condition that results in an inability to convert phenylalanine to tyrosine (King et al., 2006). Unless phenylalanine is removed from the diet brain damage can occur. Infants of PKU mothers can be exposed to high concentrations of phenylalanine in the womb, irrespective of their own genotype, and can suffer brain damage as a result (King et al., 2006). The maternal environment (in this case the genotype of the mother) determines the risk to infants of PKU mothers, a risk that is modified by the removal of phenylalanine from the maternal diet.

Another example of the importance of the maternal environment on the development of an infant is given by cigarette smoking by a mother and birthweight of an infant. It has been noted that infants of mothers who smoke have a lower birthweight (Gordis, 2009). In this case it is the behavior of mothers, not their genotype that determines the birthweight of the baby. Birthweight is a characteristic that shows a degree of environmental variation. It has been noted that the maternal genotype and the non-genetic maternal environment are responsible for a significant portion of variability in the birthweight of infants (Falconer and Mackay, 1996). In a study on the inheritance of birthweight it was observed that siblings born in succession display more similarity

in birthweight than siblings with one sibling intervening or two siblings intervening (Morton, 1955). It has been noted that a maternal environment changes with time and the difference in birthweight could be attributable to the sharing of such temporal effects (Morton, 1955; Falconer and Mackay, 1996).

The refractive power of the eye displays a large variability in the first few months after birth (-10 to + 5 dioptres (D)). This variability decreases with age with an estimate of a standard deviation of 3.2 D at four weeks after birth to 0.85 D 130 to 260 weeks after birth (Goss, 2006). It has been noted that premature infants (babies born before the normal length of pregnancy) show an increased amount of myopia (Goss, 2006). This observation indicates that the refractive state of the eye is influenced by factors present during pregnancy.

6.1.1 Factors at birth and myopia

In the current study, a hypothetical relationship between myopia and factors present at birth is critical to interpretation of the analysis. An association study was undertaken to examine the relationship between birth order and myopia. For an association to indicate causality there should be a plausible biological explanation (Woodward, 2005). It is suggested that temporal factors present during pregnancy influence the refractive state of the eye at birth (as noted above). In the current study measures of refraction were taken after a significant amount of time had passed since birth (when participants were age 11). Therefore a relationship between factors present at birth and myopia later in life is examined.

A relationship between axial length and refractive error of children aged six and physical measurements at birth has been examined (Saw et al., 2004a). It was found that axial length varied by birthweight, head circumference at birth, birth length and gestational age (the length of time of a pregnancy). It is suggested that variations in axial length due to measures taken at birth could be explained partly by the infants' genotypes. However measures such as gestational age and birthweight are under control at least in part by the maternal environment. Furthermore myopia is determined in part by axial length. High myopic patients often display an abnormally long axial length (Edwards, 1998b) and there is a strong negative correlation between

axial length and refractive error (Wildsoet, 1998; Goss, 2006). The suggestion that factors present during pregnancy influence axial length can be extended to myopia.

Although it is clear from studies of premature infants and myopia that the refractive state of the eye is influenced by factors present at birth, there may be variability in the degree to which such factors exhibit an effect on refractive error development later in life. For example, longitudinal studies that find an increased percentage of myopes in premature babies, note that the amount of myopia decreases with age. It has been noted that approximately 50% of premature infants who displayed myopia shortly after birth, were emmetropic by the age of seven (Goss, 2006).

Myopia develops frequently from the ages of 5 to 15. Studies report prevalences of less than 5% myopia before the age of 5 to more than 15-20% by the age of 15 (Maul et al., 2000; Zhao et al., 2000; Goss, 2006). It is suggested that a number of factors are responsible for the development of myopia at this time. There is evidence that the development of myopia depends upon a person's genotype. It has been observed that children with two myopic parents display myopia more often than children with no parental history of myopia (Mutti et al., 2002; Low et al., 2010). Furthermore refractive error is heritable, with estimates of between 50% to 90% of variability in the trait estimated to be due to additive genetic factors (Hammond et al., 2001; Chen et al., 2007a; Klein et al., 2009). There is also evidence that myopia can develop due to environmental influence. It has been observed in animal models of myopia that the removal of patterned vision induces large amounts of myopia (Smith, 1991). The development of this myopia is reversible, after restoration of normal vision. Furthermore it has been found that myopia occurs more often in groups who are exposed to high levels of nearwork (Zadnik and Mutii, 1998).

Myopia that is present shortly after birth may be influenced by the maternal environment (both a maternal genotype and non-genetic factors). Its absence by the age of seven may be influenced by the genotype of the infant, environmental exposure and factors present during pregnancy. Some studies have investigated a possible mechanism that links factors present during pregnancy and myopia in later life. The roles of genes and the environment in a relationship between myopia and birthweight have been investigated in twins. Monozygotic twins and dizygotic twins share the

same maternal environment. Monozygotic twins display identical genotypes unlike dizygotic twins who share 50% of their genetic information. It was hypothesized that a gene underlying myopia development may also influence birthweight. The birthweights of monozygotic twins with myopia were compared to the birthweights of monozygotic twins without myopia. If genes responsible for myopia in the twins were also responsible for birthweight, a difference between those with and without myopia in terms of their birthweight may have been observable. Similarly such a difference in birthweight would be observable between dizygotic twins disconcordant for the presence of myopia, although the difference may have been reduced due to less genetic information being shared in dizygotic twin groups. In that study no differences were observed between either pairs of twins disconcordant for myopia in terms of their birthweight (Dirani, Islam and Baird, 2009a).

6.1.2 Confounders and a birth order-myopia relationship

It has been observed that subjects who developed myopia between the ages of seven and eleven display higher birth orders (Peckham et al., 1977). This observation suggests that factors present at birth are able to influence myopia development into late childhood. Similarly this study is concerned with the relationship between birth order and myopia at age eleven. Although an association between myopia and birth order has been previously observed, a third factor may influence the relationship. For example, there is a relationship between number of children and risk of breast cancer in mothers. This can be partly explained by age of the mother. As age of mothers increases there is an increased risk of breast cancer and more time for a large number of children. Age can explain, in part, a relationship between number of children and prevalence of breast cancer in mothers. Age is termed a confounder (Woodward, 2005).

The ALSPAC cohort was designed to investigate how genes and the environment influence health and development (Golding et al., 2001). The study has collected numerous measures on health and development. Although the current study is mainly concerned with a relationship between birth order and myopia, a number of other measures taken shortly after birth were analysed. If a relationship between birth order and myopia is explained partly by birthweight for example, it would be necessary to

include information on birthweight in the analysis. Strategies for dealing with confounding include using knowledge of a prior biological mechanism and testing for an association with and without the confounder present. Since little is known about the mechanism of a relationship between birth order and myopia, the latter strategy was employed to deal with confounding. For example if birthweight could explain, in part, a relationship between birth order and myopia, inclusion of information on birthweight in the analysis would take into account its effect.

Information on a number of measures taken a considerable time after birth was also available. It has been noted that children with acquired myopia at age eleven display higher birth orders. It is also noted that myopia is associated with environmental factors (Zadnik and Mutii, 1998). Therefore a relationship between birth order and myopia may be explained in part by exposure to an environmental factor in the time in between birth and the development of myopia. To identify confounding from environmental sources after birth, measures on environmental exposures that have previously been associated with myopia were included in the current analysis. Finally it is noted that myopia is partly influenced by genetics. Although no direct information of subjects' genotypes was included in the study, the number of myopic parents of a subject was available. Therefore confounding from an increased number of genes shared in a family was investigated.

Interpretation of the effect of a confounder on a relationship may identify other factors relevant to the relationship and its biological pathway. For example if it was observed that myopes displayed higher levels of birth order but that after taking into account gestational age, myopes displayed similar levels of birth order, gestational age would be implicated in a myopia-birth order relationship. It could be suggested that birth order predisposes a subject to a certain gestational age and that in turn predisposes an individual to myopia. Investigations of the mechanism of a relationship between birth order and other disorders have been undertaken. Number of older siblings is associated with allergic disorders (Forastiere et al., 1997) with subjects who have an increased number of older siblings displaying allergic disorders less often. It has been noted that levels of immunoglobulin E are associated both with birth order and sensitivity to allergies (Karmaus, Arshad and Mattes, 2001), implicating it in a relationship between birth order and allergic disorders. Interpretation of the effect of a

number of confounders is more difficult. However it is possible to conclude after adjusting for a number of confounders, if a relationship is still present, that a relationship is not driven by one of the confounders.

To summarize; birth order has been previously associated with myopia in school-age children (Peckham et al., 1977; Rudnicka et al., 2008). This study set out to investigate an association between birth order and myopia in two cohorts with particular attention on the effects of confounding from a number of myopia risk factors and pregnancy related measures.

6.2 Methods

6.2.1 Study populations

Refractive measures were collected from ALSPAC participants at a clinic when they were age eleven. These measures were non-cycloplegic autorefraction measures. After exclusion of cases with no refraction data or no recorded birth order sample size was 5,795.

The International High Myopia Genetics Consortium (IHMGC) is a collaborative of high myopia studies (Young, 2009). Families are recruited to the IHMGC if they contain at least one high myope. Five research groups are currently part of the IHMGC located in Cardiff University in the United Kingdom, Duke University Medical Centre in the United States, National Eye Clinic, Kennedy Institute in Denmark, University of Melbourne in Australia, and Toulouse University in France. Measures from the IHMGC were taken at optometric practises and obtained for this study previously by post. Subjects displayed a mean age of 41 years. After exclusion of families where birth order could not be established and only children, the sample size was 647.

6.2.2 Measurements of refractive error and birth order

Refractive error was recorded as average spherical equivalent (average of right and left spherical equivalents). Non-cycloplegic autorefraction was undertaken at age eleven in ALSPAC subjects. Cycloplegia was not used to balance compliance and accuracy. It has been observed that non-cycloplegic autorefraction leads to the overestimation of myopia (Zhao et al., 2004; Fotedar et al., 2007) (see Chapter 3 for more). In this study myopia was investigated which therefore necessitates truncating the refractive error distribution at a suitable point. Interpretation of the performance of such a decision can be measured by the degree of accuracy of the resulting classification into myopes/non-myopes. To achieve this, calibration with a more accurate measure can be undertaken. Such measures were not available in the current study. However a validation of non-cycloplegic measures of ALSPAC participants was undertaken when children were age 7 (Williams et al., 2008b). In that study, refractive measures taken by non-cycloplegic autorefraction and cycloplegic retinoscopy were compared. It was observed that the optimal point (highest area under

the receiver operating curve) to infer the presence of myopia (< - 0.5 D refractive error measured by the more accurate technique) was -1.5 D using non-cycloplegic autorefraction. In this thesis a second validation of non-cycloplegic autorefraction was undertaken in the ALSPAC cohort when participants were age 15. It was found that -1 D was optimal to classify myopia. It is suggested that the optimal point to infer myopia in the current study is between -1.5 D and -1 D. Subjects with a refractive error below -1.5 D were classed as myopes. In the first validation study it was observed that 2 D was the optimal point to infer hyperopia. An upper limit of 2 D was set to infer hyperopia in the current study following Negrel et al. (Negrel et al., 2000). Subjects were classed as emmetropes if their refractive error was between -1.5 D and 2 D. Finally, the majority of ALSPAC participants displayed mild amounts of myopia by age 11; therefore no classification on severity of myopia was made.

In the case of the IHMGC cohort, either objective or subjective refraction was obtained via postal prescription from participating optometrists. Subjects were classed as myopes and emmetropes if they displayed an average spherical equivalent of less than -0.50 D or between -0.50 to 1.00 D respectively.

In the ALSPAC cohort, birth order (or parity) was collected via a questionnaire filled out by subjects' mothers at gestational week eighteen. Mothers were asked how many times they had been previously pregnant. A birth order of fourth born or above was infrequent (a total of 1% of subjects) and these groups were not included in the analysis. In the IHMGC cohort, birth order was not obtained directly but inferred from a pedigree file. Birth order was derived from date of birth of participants and age of siblings (i.e. if a sibling was older than the participant, higher birth order was inferred). In families where all siblings were listed birth order could be reliably estimated. The process was automated via a macro written in Visual Basic (version 6.5, Microsoft Corporation) in Microsoft Excel (2003, Microsoft Corporation). The number of observations per group diminished as birth order increased (i.e. families of larger sizes were less frequent). In this study the percentage myopia is compared across levels of birth order. It is important that within strata the percentage is reliably estimated which depends upon the number of observations per group. Under a hypothesis that an effect of birth order on risk of myopia linearly changes depending on birth order, groups were collapsed to examine if a birth order effect was evident in

groups with low numbers. For example, a birth order of three or more was collapsed into one group to give more stable frequencies for higher levels of birth order. Furthermore in both cohorts, birth order was collapsed into two groups, either first born or not first born to examine whether there was a difference between subjects with a birth order of one and those with higher birth orders in terms of number of myopes.

6.2.3 Confounders

The ALSPAC cohort is designed to examine how genetics and the environment influence health and development (Golding et al., 2001). A number of phenotypic measures were available for analysis, some of which have been previously associated with myopia development. In the ALSPAC cohort, to examine whether an association between birth order and risk of myopia could be explained by other factors, myopia risk factors were identified and included in the analysis. Number of myopic parents was established via parental questionnaire twelve weeks into pregnancy. Parents indicated "I can't see clearly at a distance" in either eye to indicate myopia. Other options were "I can't see clearly close up", "always very good sight" or "I can't see much at all". Whether mothers smoked during the first trimester was recorded via questionnaire at gestational week eighteen. Choices were "no" or "yes, cigarettes". Social class of each parent was assessed via a questionnaire given to mothers 32 weeks into pregnancy and via the Standard Occupational Classification which uses information on job title of a subject's parents. The social class of each parent was either I, II, IIInm (non-manual), IIIm (manual), IV or V. Classes IIIm to V were collapsed into one group because of diminishing numbers of cases. When subjects were six months old, mothers were asked questions about breastfeeding. Choices were "I am still breastfeeding", "I breastfed but have stopped now" "how old was the baby when you stopped" and "I never breastfed". This led to categories of the following duration of breastfeeding; never, less than one month, one to three months, three to six months or more than six months. Information on time spent reading was recorded from parental questionnaire when subjects were age eight. Mothers were asked how many hours subjects read for pleasure in school holidays. Choices given were "not at all", "less than one", "one to two" or "three or more". Amount of time spent outdoors was recorded via a parental questionnaire at age eight. Mothers indicated whether

subjects spent "not at all" (none), "one", "one to two" or "three or more" hours outside in the summer weekday. Activity was measured directly via an Actigraph accelerometer, worn on the body over a period of seven days when subjects were aged 11 to 12. The accelerometer is described in detail elsewhere (Riddoch et al., 2007). Briefly, frequency and intensity of vertical movements are recorded over defined time periods (one minute). Subjects were split into quartile groups according to average number of counts per minute. Whether subjects had siblings living at home was established by parental questionnaire given to mothers when subjects were age 11, with the choice of either "yes" or "no". Birthweight, maternal age at delivery and gestation were obtained from medical records of the birth. Low birthweight was defined as less than 2,500 grams (UNICEF, 2004). For gestation, subjects were classed depending on what tertile of the gestational period their mothers fell into, either less than thirty-nine, forty or more than forty weeks. For maternal age, subjects' mothers were classed in ten year intervals for convenience and for stable frequencies. These groups were sixteen to twenty-five, twenty-six to thirty-five and thirty-six to forty-five years old.

6.2.4 Statistical analysis

All statistics (Mann Whitney U test, Kruscal-Wallis test and logistic regression) were undertaken with SPSS (version 16.0, SPSS Inc., Chicago).

6.3 Results

Subjects from the ALSPAC cohort displayed different average spherical equivalents depending on birth order (P = 0.013, Table 6.1). The percentage number of myopes was 5%, while 6% of first born children were myopic, 4% of third born and 2% of fourth born subjects. A significant decrease in number of myopes in birth order groups above one was observed (Figure 6.1 and Table 6.2). To investigate whether there was an appreciable difference in number of myopes between groups of a birth order of two or more, the first born group was temporarily excluded. No significant difference between the number of myopes was observed (data not shown). This suggests an increase in risk of myopia is only manifest for first born individuals; therefore further analysis of confounding variables was undertaken after subjects were re-categorised into first born or not first born groups. A number of risk factors were independently associated with myopia and the first born group (Tables 6.4a and 6.4b). To test whether any of these variables could explain a relationship between birth order and myopia, they were included as covariates in an adjusted analysis. After adjustment subjects in the first born group displayed more myopes (Table 6.5).

A relationship between myopia and birth order was examined in the IHMGC cohort. An increased number of myopes were observed in the IHMGC cohort (Table 6.3). No relationship between number of myopes and birth orders above one was observed (data not shown).

Birth Order	N	Mean Rank (AveSph)	P-value
1	2841	2933.71	0.013
2	2154	3091.02	
3	812	3043	
4	211	3068.92	

Table 6.1) Average spherical equivalent by birth order (ALSPAC). Average spherical equivalent (AveSph) grouped according to birth order (Kruscal-Wallis test).

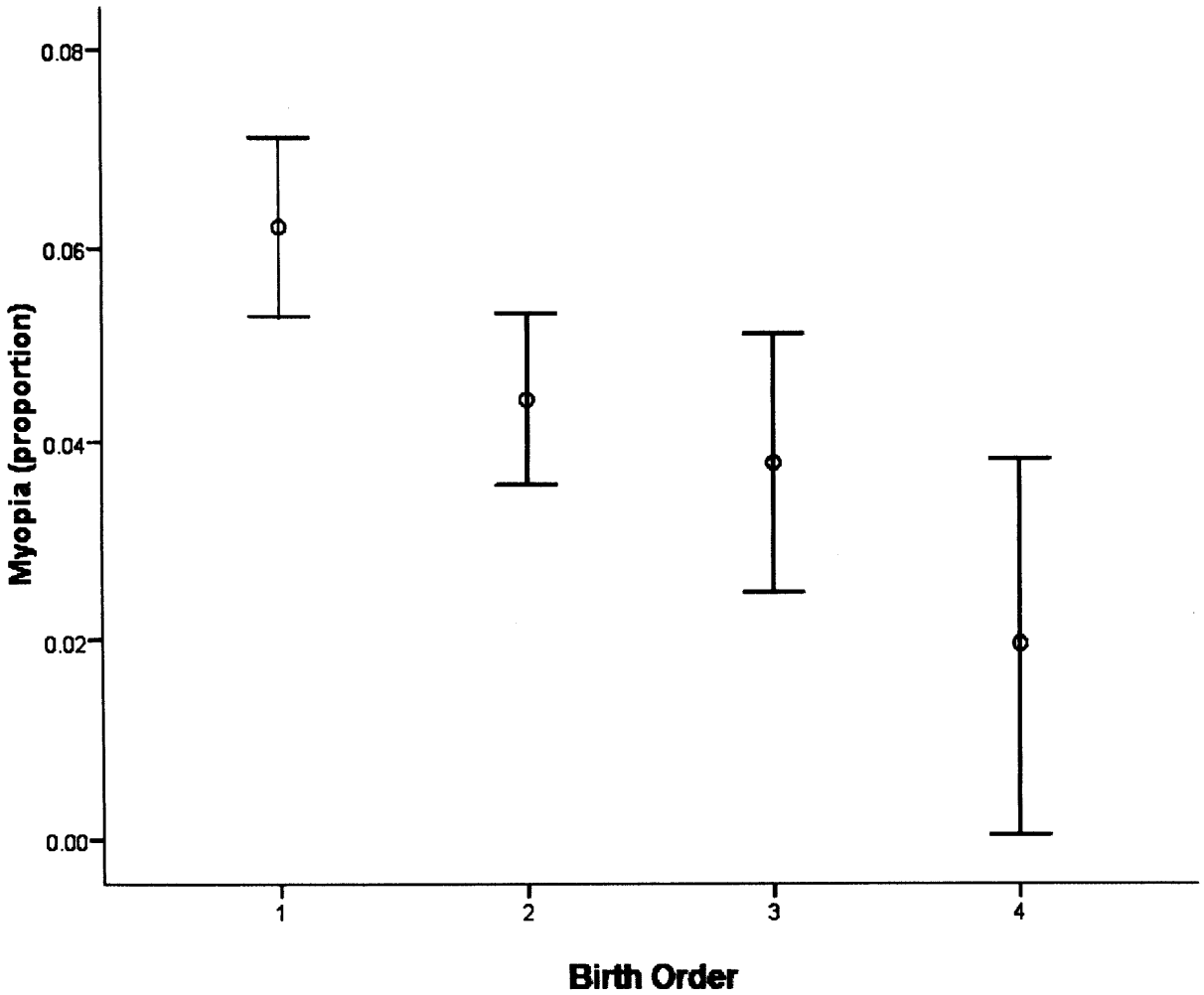


Figure 6.1) Proportion of myopia by birth order (ALSPAC). Estimates and 95% confidence intervals of proportions of myopes for varying levels of birth order.

Birth Order	P value	OR (95% CI)
1	0.002	Reference group
2	0.008	0.7 (0.54-0.91)
3	0.010	0.6 (0.4-0.89)
4	0.018	0.3 (0.11-0.81)

Table 6.2) Logistic regression I (ALSPAC). Odds ratios (OR) and 95% confidence intervals (95% CI) for myopia versus emmetropia depending upon birth order.

Birth Order	P value	OR (95% CI)
1	0.024	Reference group
2	0.023	0.58 (0.37-0.93)
3 or more	0.027	0.5 (0.27-0.92)

Table 6.3) Logistic regression II (IHMGC). Odds ratios (OR) and 95% confidence intervals (95% CI) for myopia versus emmetropia depending upon birth order.

Covariates	Myopia	
	P-value	OR (95% CI)
Number of parents with myopia (0)*	1.01E-05	
1	1.88E-04	1.61 (1.25-2.06)
2	3.19E-05	2.22 (1.53-3.24)
Hours outside during summer weekday (<2 vs. <3+)*	0.001	1.54 (1.2-1.98)
Activity in quartiles (1st* vs. 2nd-4th)*	7.28E-05	1.68 (1.3-2.17)
Number of hours reading during holidays (0)*	3.29E-11	
1	0.148	1.66 (0.84-3.31)
1-2	0.004	2.79 (1.4-5.57)
3+	2.20E-06	6.03 (2.87-12.69)
Siblings living at home (no vs. yes)*	0.001	1.77 (1.23-2.51)
Average social class of parent (I)*	1.76E-04	
II	0.006	0.63 (0.45-0.88)
III	4.00E-04	0.53 (0.37-0.75)
IIIm-V	3.46E-05	0.41 (0.27-0.63)
Gender (female vs. male)*	0.038	1.27 (1.01-1.59)
Birthweight (<= 2500 g vs. >2500 g)	0.442	1.19 (0.76-1.86)
Breastfeeding duration (never)	0.061	
< 1 month	0.518	0.86 (0.55-1.35)
1-3	0.953	0.99 (0.64-1.52)
3-6	0.839	0.96 (0.61-1.49)
>6	0.078	1.36 (0.97-1.92)
Gestation (weeks) in tertiles (1st)	0.125	
(2nd)	0.117	0.79 (0.6-1.06)
(3rd)	0.531	1.09 (0.83-1.42)
Maternal age 16 -25 years	0.238	
Maternal age 26 -35	0.091	1.3 (0.96-1.76)
Maternal age 36 -45	0.432	1.21 (0.75-1.96)
Mother smokes during first trimester (yes vs. no)	0.423	1.13 (0.84-1.52)

Table 6.4a) Multiple logistic regression I (myopia risk factors, ALSPAC).
Covariates examined in univariable binary logistic regression with myopia/emmetropia status as a dependent variable (* $P < 0.05$). OR (odds ratio), 95% CI (95% confidence intervals). Reference group is listed first i.e. for the covariate 'Hours outside during summer weekday (<2 vs. <3+)', <2 is the reference group.

Covariates	Birth Order (First vs. not first)	
	P-value	OR (95% CI)
Number of parents with myopia (0)*	4.92E-10	
1	1.40E-05	1.26 (1.13-1.39)
2	6.61E-09	1.75 (1.45-2.12)
Hours outside during summer (<2 vs. 3+)	0.054	1.11 (1-1.23)
Activity in quartiles (1st vs. 2nd-4th)	0.09	1.11 (0.98-1.25)
Number of hours reading during holidays (0)*	8.13E-17	
1	0.08	1.21 (0.98-1.5)
1-2	1.65E-07	1.81 (1.45-2.26)
3+	2.30E-08	2.34 (1.74-3.15)
Siblings living at home (no vs. yes)*	1.17E-40	3.75 (3.09-4.56)
Average social class of parent (I)*	1.02E-14	
II	0.805	1.02 (0.86-1.21)
III	0.297	1.1 (0.92-1.31)
IIIm-V	3.09E-07	0.61 (0.5-0.74)
Gender (female vs. male)	0.495	0.97 (0.88-1.06)
Birthweight (<= 2500 g vs. >2500 g)*	3.69E-11	2 (1.63-2.45)
Breastfeeding duration (never)*	5.87E-18	
< 1 month	2.68E-14	1.97 (1.66-2.35)
1-3	8.61E-10	1.73 (1.45-2.06)
3-6	2.20E-06	1.54 (1.29-1.84)
>6	0.033	1.17 (1.01-1.36)
Gestation (weeks) in tertiles (1st)*	0.017	
(2nd)	0.56	1.04 (0.92-1.16)
(3rd)	0.005	1.18 (1.05-1.33)
Mother smoked during first trimester (yes vs. no)	0.056	1.13 (1-1.29)

Table 6.4b) Multiple logistic regression II (birth order associations, ALSPAC). Associations between each covariate and number of first born versus not first born cases (* $P < 0.05$). OR (odds ratio), 95% CI (95% confidence intervals). Reference group is listed first i.e. for the covariate 'Hours outside during summer weekday (<2 vs. <3+)', <2 is the reference group.

Covariates	P-value	OR (95% CI)
First Born vs. not first	0.016	1.5 (1.08-2.08)
Number of parents with myopia (0)	0.001	
1	0.021	1.5 (1.06-2.11)
2	3.2E-04	2.35 (1.47-3.73)
Hours outside during summer (<2 vs. 3+)	0.023	1.46 (1.05-2.02)
Activity in quartiles (1st vs. 2nd-4th)	0.001	1.74 (1.26-2.39)
Number of hours reading during holidays (0)	1.0E-05	
1	0.344	1.56 (0.62-3.93)
1-2	0.053	2.5 (0.99-6.29)
3+	0.001	5.13 (1.89-13.89)
Siblings living at home (no vs. yes)	0.013	1.78 (1.13-2.81)
Average social class of parent (I)	0.092	
II	0.936	0.98 (0.63-1.52)
III	0.654	0.9 (0.55-1.46)
IIIm-V	0.024	0.45 (0.23-0.9)
Breastfeeding duration (never)	0.231	
<1 month	0.261	0.69 (0.36-1.32)
1-<3 months	0.824	0.93 (0.51-1.71)
3-<6 months	0.911	1.04 (0.56-1.9)
6 or more	0.389	1.25 (0.75-2.09)

Table 6.5) Multiple logistic regression III (adjusted analysis, ALSPAC). Multiple logistic regression of myopia on membership of the first born group. OR (odds ratio), 95% CI (95% confidence intervals). Reference group is listed first i.e. for the covariate 'Hours outside during summer weekday (<2 vs. <3+)', <2 is the reference group.

6.4 Discussion

In this study a relationship between birth order and myopia later in life was examined. It was observed that myopes were found in a group with a birth order of one more often than emmetropes in two cohorts. Furthermore it was observed that after taking into account a number of pregnancy related factors, environmental exposures and family history of myopia, a relationship between birth order and myopia was evident.

An increase in risk of myopia was found for subjects displaying a birth order of one (first born). It can be inferred from these results that the rate of myopia progression in children who are first born is higher than children who have birth orders of 2 or more. The observation that subjects with a higher birth order display more myopia has been made previously in other cohorts (Peckham et al., 1977; Rudnicka et al., 2008). This poses a question as to what mechanism could account for a relationship between myopia and birth order. Examination of other factors can lead to identification of parts of the mechanism.

6.4.1 Potential biases

Inclusion of factors that explain a relationship between birth order and myopia could help to indicate an underlying mechanism in the relationship. Inclusion of other factors also helps control for bias that may be present due to non-random sampling or study design. For example, it could be that those with a birth order of one in the study are predominantly from families where myopia is present in a number of family members, while subjects with a birth order above one come from families with no history of myopia. It may be that a number of subjects in the first born group display myopia more often because of genetic influences, while subjects with a birth order above one do not have myopia causing genes at high frequencies and display lower numbers of myopes. Comparison of number of myopes in the first born group against the number of myopes in groups with birth orders above one would indicate an increased risk of myopia for subjects with a birth order of one.

The ALSPAC study was designed with random sampling. It is a birth cohort, started in 1991, with 85% of births (14,000) in a defined geographical region, enrolled in the study (Golding, 2004). Approximately 8,000 participants attend annual clinics to

record measures on physical and psychological health and development (Golding, 2004). In this study approximately 6,000 measures of refractive error of subjects who attended an ALSPAC clinic at age 11 were included. Therefore a proportion of subjects of the birth cohort were included in the current analysis.

A reason for missing data may be due to withdrawal (when a subject leaves the cohort before the end date). The reasons for withdrawal may be unrelated to the study (for example a participant's parent may take a job outside the study area that requires long amounts of travel to participate). It is also possible that withdrawal is due to factors related to the study. For example a study of alcohol related injuries may find a weaker relationship because many of the subjects who are injured when under the influence of alcohol may not present at the study due to hospitalisation, while subjects who consume alcohol and remain injury free will be present. In this study the amount of myopia presenting in subjects is mild or moderate that requires correction by the use of prescription lenses with little associated risk of morbidity. Therefore it is suggested any withdrawals in the study sample was not related to ocular health.

Withdrawals may also occur because a participant is distressed by a particular method of measurement (such as a fear of needles). Measures of refraction were undertaken using an autorefractor without cycloplegic, which causes minimal discomfort to subjects during measurement. Therefore a reduction in the number of participants in the current study may not be due to a subject's discomfort during measurement. Participation in an ALSPAC clinic involves a number of activities including measures of a wide range of phenotypes. Therefore any withdrawal that may be related to the study would be related to a number of measures and not specifically biased towards ocular health. It is suggested that the current study sample is not biased in study design or due to sampling. The IHMGC cohort is similarly protected from bias in study design and sampling. Subjects were drawn from families that display high myopia and members of each family were included. Since participants would share a similar prevalence of the disease, it is suggested that they display similar levels of exposure to myopia risk factors. Furthermore birth order levels contained individuals from the same families. It is suggested that family members would display similar exposures to environmental influences compared to individuals randomly selected from the general population. Confounding due to study design and sampling can occur

if differences between exposures and outcome (presence/absence of disease) are present. It is suggested that participants in the IHMGC show similar environmental exposure and outcomes.

6.4.2 Confounding

An effect of a third variable on a relationship between birth order and myopia may indicate a mechanism underlying the relationship. In the current study it is observed that first born children are drawn from families with two myopic parents more often than children with birth orders of two or above (an odds ratio of 1.75, $P < 0.001$, Table 6.4b). Children with two myopic parents are at an increased risk of developing myopia, therefore a relationship between birth order and myopia may be explained by an increased number of myopic parents in the first born group. It is also observed in the current study that children with a birth order of one spend more time reading during holidays than children with a birth order of two or more. An association between exposure to nearwork and myopia has been noted in several studies (Zadnik and Mutii, 1998), therefore a relationship between birth order and myopia may partly be explained by time spent reading. Observing these two risk factors together (two myopic parents and increased time spent reading) indicates that subjects exposed to these factors would display an increased number of myopes. It is observed that the first born group display more myopes before taking into account these factors (Table 6.2). Furthermore it is observed that after taking into account number of myopic parents and time spent reading, the number of myopes in the first born group is still relatively high (Table 6.5).

It is observed that membership of the first born group was associated with a number of measures taken around the time of pregnancy. First born children were more often in the low birthweight group (an odds ratio of 2, $P < 0.001$, Table 6.4b). First born children were more often found in the third gestational tertile (an odds ratio of 1.2, $P < 0.001$, table 6.4b) compared to pregnancies of a shorter gestational period. It has been noted that premature infants, particularly those with a low birthweight display an increased number of myopes in the first few years of life (Goss, 2006). A relationship between gestational age and myopia after birth is found to be absent in some cases by the age of seven (Goss, 2006), indicating a return to normal vision for subjects who

have a short gestational period and myopia after birth. Neither birthweight nor gestational age were found to be associated with myopia at age eleven in the current study (Table 6.4a) indicating that a relationship between gestation and myopia that may have been present after birth is not present when subjects have reached the age of eleven. It is suggested that gestational age and birthweight do not explain a relationship between birth order and myopia at age eleven.

For a variable to be considered a confounder it should be related to the disease and to the exposure (Woodward, 2005). An association between birth order and a number of risk factors for myopia was investigated to identify variables that may be confounding a relationship between birth order and myopia (Table 6.4a and b). A number of confounding variables were identified and included when examining an association between birth order and myopia. It was observed that a factor associated with birth order predisposed subjects in the first born group to display increased amounts of myopia (for example family history and time spent reading). It was also possible that a factor associated with birth order protected subjects in the first born group from the development of myopia. However it is observed that the covariates identified tended to predispose subjects in the first born group to increased amounts of myopia (Tables 6.4-5). In other words, subjects displaying a birth order of one, showed a similar spectrum of risk factors to those in the myopia group.

Not all of the myopia risk factors examined in the current study were associated with birth order. Myopic subjects spent less time outdoors and were less active than their emmetropic counterparts (Table 6.4a), while there was little evidence of association between birth order and time outdoors and activity (Table 6.4b). Birth order was associated with measures taken around the time of pregnancy (birthweight, gestation and breastfeeding, Table 6.4a) but myopia at age eleven showed little or no evidence of association with these measures. Therefore it is observed that subjects in the first born group share a subset of risk factors that predispose to myopia development. Furthermore it is noted that membership of the first born group predisposes individuals to a number of myopia risk factors which in turn may lead to myopia. It is suggested that a relationship between birth order and myopia is mediated partly by environmental risk factors such as time spent reading and risk factors that indicate a genetic influence (number of myopic parents).

The relationship between birth order and myopia was examined using a number of related variables. The interpretation of the effect of a number of confounders is more difficult than considering each factor separately (Woodward, 2005). After adjustment a relationship between birth order and myopia was evident (an odds ratio of 1.5, $P < 0.016$, Table 6.5). It is possible to conclude that birth order has an effect on myopia, regardless of the other factors investigated. Although subjects displaying a birth order of one are exposed to higher levels of nearwork and in turn that may predispose a subject to myopia, the relationship between birth order and myopia cannot be fully explained by nearwork or any of the other factors examined. It is suggested that a relationship between birth order and myopia is mediated via a different mechanism than those underlying the relationship between nearwork, number of myopic parents and other factors investigated in the current study.

In summary a relationship between birth order and myopia was found in two cohorts from the UK. It is observed that there are an increased number of myopes among individuals with a birth order of one. Furthermore first born subjects spend more time reading and more often are found to have parents who are myopic. Nevertheless it is observed that a relationship between birth order and myopia is not absent after taking into account these and other myopia risk factors.

Chapter 7

Season of Birth and Myopia

7.1 Introduction

This study is different from previous chapters of the thesis in that the analysis does not include data from the ALSPAC cohort. In this chapter data was analysed from a large adult population drawn from optometric practises in the UK (Farbrother et al., 2004a). The aim of this study was to examine a relationship between season of birth (the season at which birth occurs) and myopia later in life.

It has been suggested from research in animal models of myopia that light plays a role in the development of myopia. Myopia can be induced by the removal of patterned vision in the mammalian model (monkey) of myopia to a large degree (form deprivation myopia). Removal of patterned vision can be achieved by suturing eyelids together. The following myopia is correlated with length of loss of patterned vision (Smith, 1991). It has been noted that form deprivation myopia does not occur after eyelid closure when animals are reared in the dark (Smith, 1998), indicating that a signal that controls the development of myopia is been mediated by the presence of light.

Further evidence suggests a more complicated role of light in the development of myopia in animal models. Form deprivation myopia is also observed in chicks where similar to eyelid closure in monkeys, covering the eye with a translucent material, and therefore loss of patterned vision, results in a rapid and large amount of myopia (Wallman, 1991). It has been observed that chicks fitted with translucent diffusers and exposed to high intensities of light display a reduced amount of myopia (Ashby and Schaeffel, 2010), indicating that light intensity is important in the modulation of experimentally induced refractive error.

7.1.1 Light, season of birth and myopia

In the current study, a relationship between season of birth and adult myopia is examined in a large clinical cohort. Season of birth is an exposure, which is hypothesized to be related to the development of myopia. Exposures can be classified into at least two groups (Gordis, 2009). Microenvironmental exposure relates to factors that depend on exposures that act at an individual level. Nearwork is associated with myopia (Zadnik and Mutii, 1998). Differences between subjects in terms of their reading habits in a cohort would vary at the individual level, some individuals read often and others read less often. The amount of exposure to nearwork would depend on other individual factors such as the type of school subjects attend, the type of employment that parents are engaged in, an aptitude for sport and so on. Macroenvironmental exposures affect populations or regions where exposure to a risk factor occurs for most individuals. An example of a macroenvironmental exposure is air pollution (Gordis, 2009).

In this study season of birth is a macroenvironmental exposure with groups of individuals similarly exposed to a particular season of birth. Since season of birth is a construct used to describe a portion of time (i.e. three months on the yearly calendar), it is considered that season of birth is a factor that determines exposure to a risk factor. For example, it is noted that the number of hours of daylight depends upon the time of year. Each season also displays relatively different numbers of daylight hours. Evidence from animal models indicates that light has a role in the development of refractive error (Smith, 1991; Ashby and Schaeffel, 2010) and it can be hypothesized that differences between refractive errors attributable to differences between seasons can be explained by variation in the number of hours of daylight.

Observations on a relationship between geography, light and refractive errors in humans have indicated that variation in environmental light (lighting conditions that affect the broader population) may play a role in the development of myopia. Mildefart (Midelfart, 2002) noted that in certain countries, such as Norway, large differences between light exposures are present in the general population. It was noted that since part of Norway is located in high latitudes, there are many hours of daylight during the summer months. Similarly, during the winter, long periods of darkness occur. However the south of Norway is located at latitudes similar to central Europe

and the number of hours of daylight is less during the summer, with more hours during the winter compared to northern latitudes. It was also noted that the prevalence of myopia in medical students from northern Norway was 20% higher than medical students from southern Norway (Midelfart, 2002). Similarly, it has been noted that the prevalence of myopia is higher in northern Finland than other regions. Refractive error, determined by questionnaire, was obtained from young male adults, serving in the military. Current place of residence was found to be representative of place of birth. The study found that subjects living above the Arctic Circle displayed a trend towards an increased prevalence of myopia (Vannas et al., 2003).

7.1.2 Causality; Season of birth and myopia

In this study an association between season of birth and myopia in adults is examined. An association study is the first step to establishing a causal relationship between an exposure and disease. However an association can be observed between a factor that is not causal because there is a relationship between the causal factor, risk factor and exposure (Woodward, 2005). For example a study may find an association between coffee and pancreatic cancer. It may be that coffee and pancreatic cancer are associated but not causally related. Individuals who drink a lot of coffee tend to be smokers. Pancreatic cancer in coffee drinkers may be due to the high frequency of smokers in the group (Gordis, 2009). Similarly if an association is observed between season of birth and myopia it may not be that birth in a particular season influences myopia development. It may be that an individual born in a particular season tends to be exposed to another variable. For example individuals born in summer may be exposed to a seasonal infection that leads to fever. An exposure to periods of raised body temperature may be a hypothetical cause of myopia later in life rather than an effect of season of birth.

An important aim of epidemiological investigation is to obtain a reduction in morbidity and mortality of a disease (Gordis, 2009). In some cases once a cause has been identified, steps to prevent occurrence of the disease are readily available. For example, in the 18th century approximately 400,000 individuals died each year from smallpox (Gordis, 2009). The cause of the disease was unknown. Edward Jenner made a connection between resistance to smallpox in diary maids and prior infection

with a milder disease, called cowpox. This led to the first vaccination that in turn led to the eradication of small pox late in the 20th century (Gordis, 2009). In the current study a connection between season of birth and myopia may not translate to an amenable preventative solution. Moreover the purpose of this investigation is to identify an association and estimate the size of the effect on risk of adult myopia. A causal relationship is not investigated directly and therefore the value of a relationship in terms of preventative strategies is not important.

However a primary concern of epidemiology is causality and in the current study an association is interesting because it can inform on a possible causal relationship. For an association to indicate causality there should be a plausible biological explanation. A high prevalence of myopia has been reported in premature infants shortly after birth. The amount of myopia decreases with age; by the age of one, many previously myopic subjects display emmetropia (Goss, 2006). This indicates that factors at around the time of pregnancy can influence refractive development. In the current study a relationship between season of birth and myopia in adults is investigated. Season of birth describes the first few months of a subject's growth and development. It has been noted that shortly after birth, refractive error is found to display large amounts of variability in humans (Goss, 2006). It has been noted that refractive errors across individuals become more similar a number of years after birth and it has been suggested that a wide variability shortly after birth is partly due to varying degrees of maturity of the eye after birth (Goss, 1991). It is suggested that the eye continues to undergo development after birth. It has also been noted that refractive error on average tends to be mildly myopic after birth but a hyperopic shift has taken place by the ages of 2 to 5 towards emmetropia (Goss, 2006).

In this study an exposure around the time of birth is examined as a potential risk factor for myopia development. It has been noted that the magnitude of refractive error produced by form deprivation in the animal model of myopia is related to the age of eyelid closure; an earlier age of eyelid closure leads to larger amounts of myopia (Smith, 1991). Furthermore it has been noted that similar experiments to induce form deprivation in adult monkeys do not produce large changes in refractive error compared to those observed after eyelid closure in monkeys early in life (Smith, 1991). It has been noted that the critical period for form deprivation is shortly after

birth lasting for up to 2 to 3 years (Smith, 1991). This suggests that the eye is more sensitive to a signal that induces myopia development if exposure occurs when the eye is still developing.

Myopia is thought to occur due to an imbalance of the refractive indices of components of the eye (axial length, corneal curvature and the crystalline lens). There is a correlation between axial length and refractive error, with longer axial lengths associated with lower refractive errors (Wildsoet, 1998; Goss, 2006). It is also noted that many high myopes display an axial length that is abnormally long (axial length displays a mean value of 24 mm (Millodot and Laby, 2002), but high myopes often show an axial length of over 26 mm (Edwards, 1998b)). It is also noted that myopia that develops during schooling is accompanied by axial elongation (Goss, 2006). Form deprivation studies in the animal model of myopia demonstrate that myopia can be induced via changes in the external environment. Furthermore form deprivation myopia does not occur in total darkness, indicating that light is necessary to mediate a signal that is responsible for myopia development. It has been noted that form deprivation leads to an increase in axial length (Smith, 1991). Since high levels of light intensity reduce the progression of form deprivation myopia (Ashby and Schaeffel, 2010) it is suggested that there is a relationship between light and axial length.

A positive correlation between radius of corneal curvature and refractive error and between radius of corneal curvature and vitreous chamber elongation has been noted, which imply that flatter corneas occur when the eye is longer (Wildsoet, 1998). Other studies have observed that corneal power does not change to a large degree from early childhood (Rosenfield, 2006), which implies that a relationship between corneal power and refractive error is determined early on in childhood. Furthermore it has been noted that corneal power in some myopes is significantly smaller than their emmetropic counterparts (Rosenfield, 2006) and that radius of corneal curvature is reduced in subjects displaying juvenile onset myopia. Since corneal power has been noted to be relatively stable by early childhood it is suggested that the reduced corneal power and shortened radius of corneal curvature observed for some myopes may occur early in childhood. It has been noted that continuous amounts of light lead to a flattening of the cornea in poultry (Howland, 2010). Furthermore a correlation

between light intensity and corneal refractive power in chicks has been noted. Early on in life, chicks were reared in continuous light with varying intensity. As light intensity increased the flatness of the chick cornea increased and the amount of hyperopia grew similarly (Cohen et al., 2008). It is suggested that light plays a role in the development of corneal power in the chick.

In summary a relationship between season of birth and adult myopia is investigated in the current study. The aim of the study is to identify if an association exists as a starting point for further epidemiological analysis. A possible mechanism that relates season of birth and myopia is given by animal models of form deprivation myopia, which indicates light is an important source of variation in refractive development. This plus the observation that the human eye is still developing shortly after birth suggests that factors related to season of birth may have a role in myopia later in life.

7.2 Methods

The study population (Farbrother et al., 2004a) consisted of 90,884 subjects who attended optometrists for a sight test in the UK. After removal of subjects below eighteen years of age and systematic errors, there were 74,459 subjects. Information on subjective spectacle prescriptions, gender, date of birth and date of eye test were available. The subjects were aged 18-100 years (the mean ± standard deviation age was 50 ± 17 years in males and 50 ± 19 years in females). 59% of the study population were female.

Average spherical equivalent (mean of right and left spherical equivalent) was used. Subjects displaying greater than -0.75 D were classified as non-myopes. Myopia level was categorised by severity; mild (-0.75 to -2.99 D), moderate (-3.00 to -5.99 D) and high (< -6.00 D). Statistics (Chi square test, Kruscal-Wallis test and logistic regression) were generated in SPSS (version 12.0.2, SPSS Inc., Chicago).

Season of birth was defined as winter (December, January and February), spring (March, April and May), summer (June, July and August) and autumn (September, October and November). Daylight hours in the UK were also obtained, from (http://aa.usno.navy.mil/data/docs/RS_OneYear.php). Photoperiod (number of daylight hours) categories follow a definition by Mandel et al., who found that photoperiod and myopia were associated (Mandel et al., 2008). The categories for daylight hours were defined as follows; Photoperiod 1 (November, December and January), photoperiod 2 (February, March and October), photoperiod 3 (April, August and September) and photoperiod 4 (May, June and July).

7.3 Results

An association between myopia and photoperiod was not observed ($\chi^2 = 8.6$, df = 9, P = 0.475, Table 7.2). An association between season of birth and myopia was observed ($\chi^2 = 20.5$, df = 9, P = 0.015, Table 7.3). Multivariable logistic regression analysis was undertaken with age and gender as covariates. An association between photoperiod category four and mild myopia (OR 0.94, 95% CI 0.89 – 0.99, P = 0.019) was evident (Table 7.4). An association between the high myopia group and birth in summer (an odds ratio of 1.17, 95% CI 1.05 – 1.30, P = 0.006) and in autumn (an odds ratio of 1.16, 95% CI 1.04 – 1.30, P = 0.007) was found (Table 7.5). No association between the degree of high myopia and season of birth (Kruscal Wallis test, P = 0.41) was observed (Figure 7.1). Both age and sex were significantly associated with myopia (Table 7.4 and 7.5).

Photoperiod category	Daylight hours
	UK
1	9.31 - 10.15
2	10.16 - 13.03
3	13.04 - 15.71
4	15.72 - 18.01

Table 7.1) Daylight hours and photoperiod. Daylight hours in the UK in photoperiod categories 1-4.

	Myopia severity category		
Photoperiod category	Mild	Moderate	Severe
1	20.4%	12.1%	3.8%
2	20.0%	12.0%	3.7%
3	20.0%	12.3%	3.8%
4	19.7%	11.8%	4.1%

Table 7.2) Myopia by photoperiod category. Percentage number of myopes in three severity categories for each photoperiod category.

	Myopia severity category		
Season	Mild	Moderate	Severe
Winter	20.0%	12.0%	3.6%
Spring	19.9%	12.0%	3.6%
Summer	19.8%	11.9%	4.1%
Autumn	20.4%	12.3%	4.1%

Table 7.3) Myopia by season of birth. Percentage number of myopes in three severity categories by season of birth.

	Mild myopia		Moderate myopia		High myopia	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Photoperiod category						
1	Reference group	0.940	Reference group	0.218	Reference group	0.278
2	0.95 (0.90 – 1.00)	0.052	0.95 (0.89 – 1.02)	0.169	0.94 (0.84 – 1.05)	0.246
3	0.97 (0.92 – 1.02)	0.258	1.01 (0.94 – 1.08)	0.803	1.00 (0.90 – 1.11)	0.964
4	**0.94 (0.89 – 0.99)**	0.019	0.96 (0.89 – 1.02)	0.191	1.04 (0.94 – 1.16)	0.446
Age	**0.962 (0.961 – 0.963)**	<0.001	**0.954 (0.953 – 0.956)**	<0.001	**0.967 (0.964 – 0.969)**	<0.001
Sex	0.99 (0.95 – 1.03)	0.581	**1.16 (1.11 – 1.22)**	<0.001	**1.36 (1.25 – 1.47)**	<0.001

Table 7.4) Multiple logistic regression I (photoperiod). Odds ratios (OR) and 95% confidence intervals (95% CI) for myopia versus non-myopia. Significant ($P < 0.05$) associations are in bold type. Numbers are correct to two decimal places except when three decimal places are needed to define 95% CI i.e. for an association between age and mild myopia the estimate and 95% CI are 0.96 (0.96 – 0.96) when correct to two decimal places.

	Mild myopia		Moderate myopia		High myopia	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Season						
winter	Reference group	0.481	Reference group	0.435	Reference group	0.002
spring	0.99 (0.93 – 1.04)	0.585	0.98 (0.92 – 1.05)	0.622	1.00 (0.89 – 1.12)	0.973
summer	0.99 (0.94 – 1.05)	0.779	1.00 (0.93 – 1.07)	0.903	**1.17 (1.05 – 1.30)**	0.006
autumn	1.03 (0.97 – 1.08)	0.356	1.04 (0.97 – 1.11)	0.284	**1.16 (1.04 – 1.30)**	0.007
Age	**0.962 (0.961 – 0.963)**	<0.001	**0.954 (0.953 – 0.956)**	<0.001	**0.967 (0.965 – 0.969)**	<0.001
Sex	0.99 (0.95 – 1.03)	0.596	**1.17 (1.11 – 1.22)**	<0.001	**1.36 (1.25 – 1.47)**	<0.001

Table 7.5) Multiple logistic regression II (season of birth). Odds ratios (OR) and 95% confidence intervals (95% CI) for myopia and season of birth. Bold type indicates significant ($P < 0.05$) associations. Numbers are correct to two decimal places except when three decimal places are needed to define 95% CI i.e. for an association between age and mild myopia the estimate and 95% CI are 0.96 (0.96 – 0.96) when correct to two decimal places.

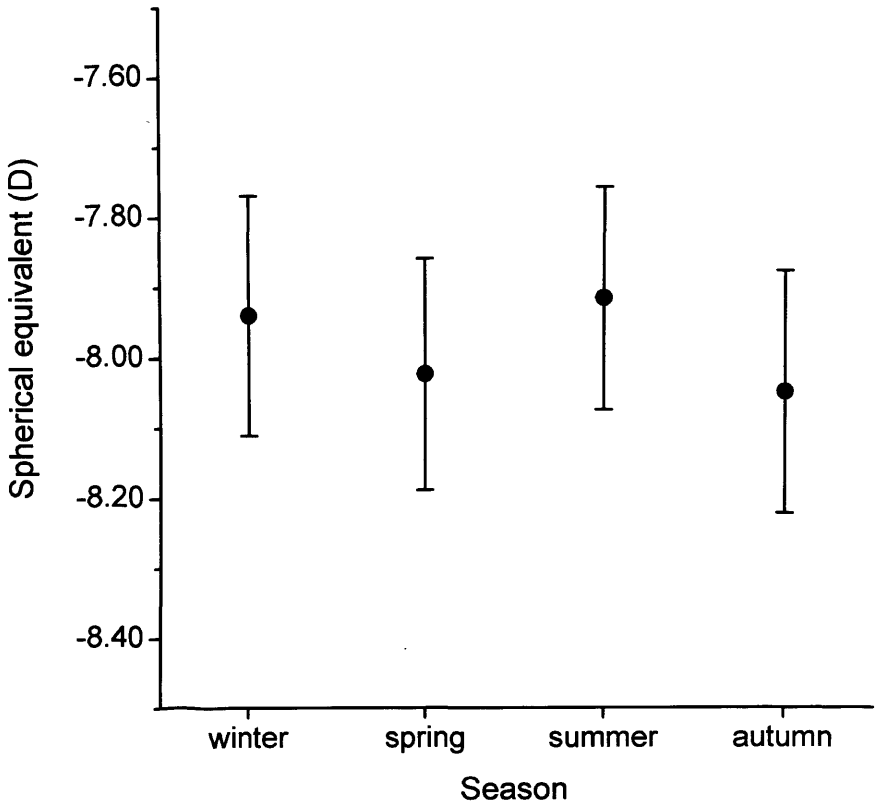


Figure 7.1) High myopia by season of birth. Mean spherical equivalent values (black dots) with 95% confidence intervals in high myopes as a function of season of birth.

7.4 Discussion

In this study a relationship between season of birth and myopia was examined. It was observed that birth during summer and autumn is associated with an increased chance of high myopia in adults (Table 7.5). A non-significant increase in the number of mild and moderate myopes born in the autumn was also noted (Tables 7.3 and 7.5). These finding are similar to others in an independent cohort (Mandel et al., 2008) where an increased prevalence in moderate to severe myopia was noted for subjects born in the summer months.

7.4.1 Alternative explanation

Season of birth refers to the first few months of life, shortly after birth. Evidence from animal models indicates that light is necessary for myopia that is induced by removal of patterned vision in the early visual experience. It is noted that myopia induced by form deprivation is more potent shortly after birth (Smith, 1991). Other studies have investigated a relationship between light and myopia. It has been noted that subjects who were exposed to ambient lighting (either a night light or a room light) before the age of two displayed a higher number of myopes by the age of 12 than subjects with little or no exposure (Quinn et al., 1999). The study found a strong association between ambient lighting at night and a dose response between the amount of ambient light and the number of myopes observed. It was also noted that a relationship between night light use in children and myopia could be explained by a mechanism analogous to form deprivation myopia via eyelid closure, where some amount of light enters the eye and a degraded image is transmitted to the retina (Quinn et al., 1999). The study provided evidence that a relationship between ambient lighting at night and myopia was evident and that the association may indicate causality. There are a number of principals that can help guide whether an association indicates a causal relationship (Woodward, 2005; Gordis, 2009), which the study highlighted. There was a strong association between the risk factor and disease. A dose response in the number of myopes was observed with increasing exposure to ambient light at night. There was evidence that exposure to the risk factor preceded disease onset; although measures of refractive error were obtained across an age range of 2 to 16, night light use was recorded for the first two years after birth. There was a plausible biological explanation.

Another principle to guide whether an association indicates causality relates to the findings of other studies that examined a relationship between use of ambient light and myopia. If an association indicates a causal relationship, it is expected that other investigations find similar results in different research environments. If similar results are found in independent studies the effect on association from sources of bias such as study design, study population and random sampling is limited (Woodward, 2005). An association between ambient light use at night and myopia was examined in a number of cohorts after the report of the initial finding was made. An association was not found between ambient light use and myopia in a number of other cohorts (Gwiazda et al., 2000; Zadnik et al., 2000). Instead an association between night light use and number of myopic parents was observed. Myopia is more common among subjects who have two myopic parents (Drack, 1998; Mutti et al., 2002). An association between number of myopic parents and night light use suggests that a relationship between ambient lighting at night and myopia can be explained by an increased number of myopic parents in the group reporting increased night light use. In other words a relationship between ambient lighting at night and myopia may have been confounded by number of myopic parents.

In the current study an association between season of birth and myopia was observed. It is suggested that light levels may explain the observation that an increased number of myopes were found to have been born in the summer months. The rationale for this suggestion is similar to that proposed to explain a relationship between the use of ambient light at night and myopia. Another principle useful to interpret the results of an association study is that there should be no other convincing alternative explanation (Woodward, 2005). A relationship between ambient light at night and myopia could be explained by an increased number of myopic parents for subjects exposed to increased levels of ambient light use at night. Similarly there are other explanations for an association between season of birth and myopia in adults. Myopia is associated with socio-economic status (Wong et al., 2000; Shimizu et al., 2003; Zadnik and Mutti, 2006). It has been noted that myopia is least frequent in lower income groups and increases in higher income groups (Zadnik and Mutti, 2006). If birth during the summer months is associated with socio-economic status then an association between season of birth and myopia may be explained, in part, by a

relationship between socio-economic status and myopia. A relationship between birth during the summer months, socio-economic status and myopia has been investigated. Mandel et al. found an increased number of myopes were born during the summer months (Mandel et al., 2008). It was observed that no association was found between time of birth of a subject's siblings and a subject's refractive error. It was noted that if family planning could explain a relationship between birth during the summer months and myopia then siblings would be born during a similar period of time and an association between time of birth of siblings and refractive error would be maintained. The ability of family history of myopia to explain a relationship between birth during the summer and myopia was also investigated. Sibling refractive error was included in an adjusted analysis of a relationship between birth during the summer months and myopia. It was noted that a relationship between birth during the summer months and myopia was independent of siblings' refractive errors (Mandel et al., 2008). It is indicated that a relationship between birth during the summer months and myopia in that study was not completely explained by familial factors.

7.4.2 Caveats

Evidence from animal models indicates a role for light in the development of form deprivation myopia; however it is possible that the findings do not translate well to human refractive error. It has been noted that human myopia develops more frequently from the ages of eight to fifteen, while evidence from animal models indicates that form deprivation is possible at an earlier sensitive period, shortly after birth while the eye is developing (Goss, 2006). Furthermore a large disruption to the normal visual experience necessary to produce form deprivation myopia does not occur often in human populations. It has been noted that an example of a large loss of normal vision analogous to form deprivation in terms of magnitude is cataract (Goss, 2006). Cataract is a partial or complete loss of transparency of the crystalline lens. Cataract results in a gradual loss of vision. Symptoms include dimming of illumination and diminution of optical image (Millodot and Laby, 2002). In patients with congenital cataract, which develops early in life, axial length is found to be greater than normal (Goss, 2006). It is noted that form deprivation myopia leads to significant increases in axial length (Smith, 1991). It has also been noted that ocular conditions that lead to loss of visual experience such as neonatal eyelid closure and

ptosis also can lead to high myopia (Goss, 2006). Ptosis is a drooping of the upper eyelid that can lead to partial loss of the visual field (Millodot and Laby, 2002).

Another observation that can be made regarding a relationship between light and myopia and animal models is that it has been noted that an increase in light intensity leads to a reduction of form deprivation (Cohen et al., 2008; Ashby and Schaeffel, 2010). In human studies on a relationship between ambient night light use (Quinn et al., 1999) and variation in natural light (Vannas et al., 2003; Mandel et al., 2008; McMahon et al., 2009) increased amounts of light have been noted to be associated with an increase in the number of myopes in certain groups. It is suggested that the effect of light on the development of myopia is different in animal models and human studies which further suggests a more complex set of interactions between the visual experience and variation in exposure to light.

Furthermore, it has been noted that myopes engage in less sporting activity independent of time spent reading (Mutti et al., 2002) (i.e. myopes engaged in a certain amount of reading participate in less sports than non-myopes who display similar amounts of exposure to reading). It has also been noted that children engaged in more sports and outdoor activities develop myopia less often (Jones et al., 2007) independent of nearwork. It is suggested that there is a relationship between activity and myopia. Studies have also found that time engaged in indoor sport was not associated with myopia, but that subjects that spent time engaged in outdoor activities displayed less numbers of myopes (Rose et al., 2008a; Dirani et al., 2009b), indicating time outside as an important factor in a relationship between time engaged in sporting activity and myopia. It is suggested that a mechanism that requires time spent outside is involved in myopia development in children.

In this study a relationship between season of birth and myopia is examined. An increased number of high myopes were found among subjects born in the summer months. It is hypothesized that light plays a role in a relationship between season of birth and myopia. A relationship between an increase in the amount of light during summer and predisposition to myopia is tenuous. However since disruption of the visual experience leads to increased axial length and high myopia in humans, and in animal models a similar relationship between loss of visual stimulus and myopia is

observed, a relationship between season of birth and myopia via amount of light is hypothetically possible. An effect observed in the current study is small (Table 7.5), indicating that only a proportion of individuals who are born during the summer months display high myopia later in life. It is suggested that a relationship between season of birth and myopia depends upon other factors that predispose to myopia and it is possible that season of birth is related to a causal factor and is related to myopia but is a consequence of neither.

It is noted that season of birth is a macroenvironmental exposure. There are variations in season of birth in terms of light exposure. During winter less hours of daylight are observed than summer, when days are long. However there are within season variations that lead to changes in the amount and intensity of light. Summer can have periods of rain and a reduced exposure to light levels. In the current study an association between photoperiod (number of daylight hours) and high myopia was not observed. It is suggested that photoperiod in the UK is modulated by the changeable weathers conditions of North West Europe. An association between high myopia and birth during the summer months was observed in this study. It is suggested that the summer months are the sunniest and represent a time when exposure to periods of uninterrupted daylight would be at its highest. In an independent cohort a strong, dose dependent association between photoperiod and myopia was observed (Mandel et al., 2008). The study examined a relationship between photoperiod and myopia among subjects enrolled in military duty in Israel. An association was found between birth during the summer months and myopia. Israel affords brighter summers than the UK. It suggested that season of birth measures indirectly an exposure to a risk factor that may display a relationship with myopia. There may be an accumulation of small errors in the measure that lead to a reduction in the strength of a relationship between season of birth and myopia. For example some summers may be associated with longer exposure to intense sunlight. Season of birth refers to the time of birth only and does not inform on individual changes that may be associated with birth at that time.

In summary a relationship between season of birth and adult myopia was investigated in this study. An increased number of high myopic subjects were observed among subjects born during the summer and autumn months. Age or gender could not account for the relationship.

Chapter 8

General Discussion

This chapter discusses the findings of this thesis, the motivations behind their examination, their value towards better understanding myopia in the ALSPAC cohort, some of their caveats and future work. This thesis seeks to better understand the genetics and epidemiology of myopia in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort. Myopia is an inability to see objects in the distance clearly. It has been noted that myopia clusters in families (1889; Drack, 1998; Zadnik and Mutti, 2006) and that groups classified according to exposure to certain risk factors display more myopia than others, including individuals exposed to nearwork (Zadnik and Mutii, 1998; Zadnik and Mutti, 2006). It can be inferred from findings in myopia research that the disorder may be under the influence of both genetic and environmental influences. This thesis seeks to examine these influences and their effect on myopia in the ALSPAC cohort.

ALSPAC was designed to better understand how genes and the environment influence a person's health and development (Golding et al., 2001). Between 1991 and 1992, 85% of pregnancies (approximately 14,000) were enrolled in the study from the Avon region in South West England (Golding, 2004). The health and development of participants has been followed since and the study is ongoing at present. The development of refractive errors has been examined in the ALSPAC cohort (Williams et al., 2008a; Williams et al., 2008c) and attempts to better understand the genes that play a role in refractive error and environmental risk factors that may predispose an individual to the development of refractive error are ongoing in the ALSPAC cohort. As such this thesis is part of those efforts.

8.1 Measurement

In this thesis, the reliability of non-cycloplegic autorefraction to identify myopia was assessed, using subjective refractions collected from optometrists. Subjective refractions are thought to be more accurate than non-cycloplegic autorefraction as measurements are made based on patient feedback, in the presence of a trained optometrist. Furthermore a number of measures of refractive error are taken using different instruments, similar to changing the conditions of measurement, allowing for irregularities due to factors unrelated to refractive error to be identified such as accommodative anomalies. Subjective refractions of a subset of individuals of the larger ALSPAC cohort were collected and the reliability of non-cycloplegic autorefraction to classify individuals according to the presence or absence of myopia was assessed. This allows for an inference to be made on the reliability of non-cycloplegic autorefraction on the larger ALSPAC cohort. In the validation study (Chapter 3) it was observed that non-cycloplegic autorefraction has an optimal sensitivity and specificity of 89% and 96% respectively to classify individuals according to the presence or absence of myopia in the ALSPAC cohort when participants were age 15.

Epidemiological and genetic studies are enhanced by accurate definition of a phenotype (Haines and Pericak-Vance, 1998; Woodward, 2005). In Chapter 5 a genome-wide association study of myopia was undertaken in the ALSPAC cohort when participants were age 15. The validation study allows for a more accurate definition of the presence of myopia to be inferred (at a refractive error of less than -1 dioptres (D)). Often a refractive error of less than -0.5 D is used to indicate the presence of myopia (Edwards, 1998b; Negrel et al., 2000). However it is shown in the validation study that using -0.5 D or less on a refractive error distribution produced by non-cycloplegic autorefraction would give a false positive rate (the percentage number of subjects incorrectly indicated as myopes) of 23% compared to only 4% if -1 D is used to identify myopia. It has been noted that misclassification can result in a reduction in the ability of a study to identify a true positive (reliable) result (Haines and Pericak-Vance, 1998). Therefore an advantage of undertaking calibration of non-cycloplegic autorefraction is an increased chance of success in epidemiological and genetic analysis of myopia in the ALSPAC cohort.

8.2 Heritability

A second step to identify components of a disease is determination of whether genetic influences play a role (Haines and Pericak-Vance, 1998; Gordis, 2009). A heritability study allows estimation of the proportion of total phenotypic variation that is attributable to genetic factors (Klug, 2009). It is possible to identify whether a trait is determined in part by genetic factors by examining previous studies that asses the contribution of genes to the development of the disease (Haines and Pericak-Vance, 1998). However a review of previous heritability studies may not guarantee that the importance of genes in the development of a disease in a particular cohort will be established. As noted in Visscher (Visscher et al., 2008, p.256) "... the heritability in one population does not, in theory, predict the heritability of the same trait in another population."

A study that indicates a heritability for a trait will indicate that a proportion of phenotypic variation is attributable to genetic factors but does not necessarily mean that in another population, genetic factors play a role in the development of a disease. This is due to two factors, pointed out in Visscher (Visscher et al., 2008), a) variation in genetic factors and b) variation in environmental factors. Estimates of heritability are determined in part by gene frequencies (Hartl and Clark, 1997). Therefore differences in gene frequencies between populations may mean that heritability estimates vary from one group to another. The amount of phenotypic variability may change due to non-genetic factors such as the environment. If the environment accounts for a large proportion of phenotypic variation then a smaller proportion of variation may be due to genetic factors. In theory, a genetically identical population which has not been exposed to the same environmental effects may show a larger heritability although the absolute contribution due to genetic factors has not changed. This is largely due to heritability being measured as a percentage of the total phenotypic variation, which itself can change irrespective of genetic factors. As noted by Visscher (Visscher et al., 2008, p256), 'A consequence of the definition of heritability is that it depends on the population, because both the variation in additive and non-additive genetic factors, and the environmental variance, are population specific.'

Therefore, there is a valid reason why a review of previous studies on the heritability of refractive error may not guarantee that genetic factors account for a large proportion of variation in the development of myopia in the ALSPAC cohort; the majority of myopes may have developed the condition largely due to differing environmental exposure (myopia prevalence varies in groups classified according to various environmental factors such as amount of nearwork and socioeconomic status (Zadnik and Mutii, 1998; Zadnik and Mutti, 2006) or/and differences in gene frequencies are present (estimates of the heritability of refractive error vary from 18% to 27% in a genetically isolated population on the island of Sardinia (Biino et al., 2005) compared to above 50% in family and twin studies (Hammond et al., 2001; Klein et al., 2009)).

In this thesis, refractive error was found to be a heritable trait in the ALSPAC cohort, with an estimate of 50%. Therefore it can be concluded that a proportion of the variation in refractive error in the ALSPAC cohort is attributable to genetic factors and the trait will be amenable to genetic analysis. Furthermore it was observed that a common environment shared between siblings could explain 18% of variation in refractive error. Therefore environmental influences shared between siblings play a role in the development of myopia as measured in subjects from the ALSPAC cohort. It can be expected from the results of the heritability study in this thesis, that refractive error in the ALSPAC cohort is determined by both genetic and environmental factors.

8.3 Genetics

Chapters 2-4 detail collection of refractive error (Chapter 2), definition of the phenotype of interest (Chapter 3) and determination that refractive error displays a genetic component (Chapter 4). A third step to identifying the component of a disease is the identification of a genomic region that may harbour a causal mutation (Haines and Pericak-Vance, 1998). In this thesis a genome-wide association analysis was undertaken of myopia, average spherical equivalent, axial length and corneal curvature. It was hypothesized that changes in refractive error leading to myopia may be determined via changes in axial length or corneal curvature. Axial length is correlated with refractive error (Wildsoet, 1998; Goss, 2006) and many myopes display an increase in axial length (Goss, 2006). Similarly corneal curvature is found to be reduced in some cases of myopia (Gonzalez Blanco et al., 2008) and a difference between corneal power between myopes and emmetropes has been observed (Rosenfield, 2006).

Genome-wide association studies have been used to identify hundreds of alleles involved in many common disorders (Manolio and Collins, 2009). Advantages of genome-wide association studies include small genomic regions that are identified and, via recruitment of large numbers of unrelated individuals, identification of alleles that lead to a small change in a phenotype. A number of studies of the genetics of myopia have been undertaken using linkage analysis (Young et al., 1998a; Young et al., 1998b; Hammond et al., 2004) which has identified a number of cytogenetic locations which may harbour a mutation that leads to myopia. Regions identified by linkage analysis are significantly longer than regions identified in genome-wide association studies, being in the order of megabases (1 million base pairs, Mb) compared to genome-wide association studies which are in the kilobase (1000 base pairs, kb) range. Fine mapping of regions indicated by linkage analysis has yet to indicate a causal mutation that leads to myopia; this may be due in part to the size of the region identified.

It has been noted that genome-wide association studies have led to the identification of many genes that may play a role in a disease without any prior biological hypothesis indicating their involvement. Examples of such findings include age-

related macular degeneration (Altshuler et al., 2008). Age-related macular degeneration (AMD) or age-related maculopathy leads to degeneration of photoreceptors of the macula (an oval area of the retina where best visual acuity is obtained (Millodot and Laby, 2002)) and severe reduction of vision (Gorin, 1998; Millodot and Laby, 2002). A genome-wide association study found an association between a SNP in the complement factor H gene on chromosome 1q31 and AMD (Klein et al., 2005). Complement factor H is a key regulator of the complement pathway. Complement proteins can mediate immune responses such as phagocytosis (ingestion of invading pathogens by immune cells such as leukocytes) and bacteriolysis (destruction of bacteria cells) (King et al., 2006).

An advantage of a genome-wide association study is that it is hypothesis free (i.e. no prior information is required regarding a biological mechanism linking a genomic region to a disease). In this thesis approximately 500,000 directly genotyped single nucleotide polymorphisms (SNPs) were examined for an association with myopia, average spherical equivalent and their ocular determinants (axial length and corneal curvature), along with another 2 million SNP genotypes estimated via imputation. A genotyped SNP with the lowest P value in an association with the presence of myopia was observed in an intron of the type IV collagen gene (COL4A1) on chromosome 13q34. Mutations in this gene that cause ocular symptoms, a number of which are similar to myopia, in mice (Van Agtmael et al., 2005; Gould et al., 2007) and humans (Sibon et al., 2007; Coupry et al., 2010) have been documented. Therefore COL4A1 represents a plausible gene in the development of myopia. However it is noted that genes associated with gross defects of an organ may not be responsible for less severe changes of a phenotype and findings in animal models may not translate to the human case (Jobling et al., 2004b).

A number of genome-wide association analyses have identified SNPs that are located in genes that have previously been related to the disease being investigated (Altshuler et al., 2008). For example it has been noted that in genome-wide association analyses of low density lipoprotein, high density lipoprotein and triglyceride levels, a majority of SNPs identified were located at loci with known functions related to the phenotypes (Altshuler et al., 2008). Elevated levels of low density lipoproteins are a feature of familial hypercholesterolemia a common disease with a prevalence of 1/500

among American, Japanese and European populations (King et al., 2006). In familial hypercholesterolemia cholesterol builds up on arterial walls leading to atherosclerosis (hardening of the arteries) (King et al., 2006).

The SNP in COL4A1 lies in an intron. Intronic DNA is transcribed to RNA but is lost before translation to a protein. However the intron/exon boundary contains sequences of DNA that are involved in excision and splicing mechanisms (King et al., 2006). Splicing refers to the processing of messenger RNA (which contains a copy of a gene sequence destined for translation into a protein) by removal of intron sequences (Klug, 2009). Alternative splicing refers to the combination of different exons within a gene to form various proteins with different functions (Jobling et al., 2004b). At least 40% of human genes undergo alternative splicing. For example the calcitonin/calcitonin gene-related peptide gene (CT/CGRP gene) is spliced in such a way to produce a messenger RNA transcript of the first four exons in thyroid cells, but in the brain and nervous system CT/CGRP transcripts contain exons five and six. The two different proteins produced from the same gene in different locations vary in length and function (a peptide of 32 amino acids in the thyroid functions in regulating calcium compared to a peptide of 37 amino acids in the nervous system that is an active hormone in a wide range of tissues) (Klug, 2009). Therefore it is possible that a SNP within an intron leads to a change in function of a protein.

Although a genome-wide association signal identifies a smaller genomic region compared to linkage analysis it may not identify a causal mutation (Altshuler et al., 2008). Genome-wide association studies incorporate hundreds of thousands of SNPs across the entire genome and are not limited to regions of known function. It has been noted that the human genome consists of approximately 1% coding regions, the majority of DNA sequence being taken up by intronic DNA (a piece of DNA that is transcribed to RNA but subsequently lost before translation to a protein (King et al., 2006), 24%) and intergenic DNA (a segment of DNA located between two genes, 75%) (Venter et al., 2001). Therefore variation in some of the markers examined in a genome-wide association study leads to changes in a protein and in turn displays an effect on a phenotype; however variation in many more genotypes may not lead to a functional change.

Although many single nucleotide polymorphisms (SNPs) in the human genome are not located within an exon of a gene, many are inherited together on a single stretch of DNA (which may or may not encompass a gene) termed a haplotype. Haplotypes are stretches of DNA containing a combination of alleles (Jobling et al., 2004b). Haplotypes may be broken up by recombination (the exchange of DNA between a pair of homologous chromosomes during meiosis (Jobling et al., 2004b)). It has been noted that a substantial proportion of the human genome is made up of blocks of significant length (approximately 44 kilobases in European populations, 22 kilobases in African populations) (Gabriel et al., 2002). Within haplotypes which are not separated by recombination, SNPs tend to display a degree of linkage disequilibrium (co-inheritance of alleles). Within haplotype blocks, linkage disequilibrium tends to be high with consecutive alleles being inherited together (Jobling et al., 2004b). It has been noted that a small number of haplotypes (3 to 5) represent a large majority of variation in a population (90%) within a block (Gabriel et al., 2002). However, across sites of recombination linkage disequilibrium is reduced leading to increased haplotype diversity. The block like structure of the human genome indicates that a number of SNPs within a block represent similar information and that a reduced number of SNPs can be used to indicate most of the variation present in a population.

SNP rs9521666 may be in linkage disequilibrium with a coding or regulatory variant within an exon. Other association signals observed in this thesis outside known genes may be explained by linkage disequilibrium. A SNP (rs34583) with the second lowest P value (approximately 10e-6) in an association with myopia status is located approximately 40 kilobases upstream of ETS-domain-protein, a transcription factor located on chromosome 12q.23. A regulatory sequence is a DNA sequence that regulates the expression of other genes (King et al., 2006) (such as a transcription factor). A transcription factor is a protein that can bind to a stretch of DNA on a chromosome and regulate the transcription of a gene or number of genes. For example it has been noted that a C/T (cytosine to thymine) sequence variant (a transcription factor binding site) for lactase persistence (an ability to digest lactose in adulthood), resides approximately 14 kilobases upstream of the lactase gene. The C/T variant disrupts a consensus binding site of the transcription factor AP-2 (Jobling et al., 2004b).

In this thesis the strongest association signals (10e-6) for axial length were found for SNP rs1200618 located on chromosome 11q22.3, approximately 300 kilobases away from the nearest gene and SNP rs12410731 located on chromosome 1q24 also approximately 300 kilobases away from the next nearest gene. It has been noted that a number of genome-wide association studies have found evidence of association in gene deserts. Gene deserts are regions of DNA that contain no known gene. It has been noted that approximately 20% of the genome is defined by stretches of DNA of at least 500 kilobases where no known gene can be found (Venter et al., 2001). Gene deserts are not uniformly spread out through the genome and there are areas that are rich in genes such as chromosome 17, 19 and 22 (Venter et al., 2001). It has been noted that a SNP associated with myocardial infarction located on chromosome 9p21 is 150 kilobases from the nearest gene and a variant on chromosome 8q24 associated with cancers of the prostate, breast and colon is 300 kilobases from the nearest gene (Altshuler et al., 2008; Frazer et al., 2009).

8.4 Epidemiology

Epidemiological analysis can be useful to help identify components of disease progression. In this thesis a relationship between birth order (the relative order of pregnancies of an individual, first born, second born and so on) and myopia was examined in the ALSPAC cohort when participants were approximately 11 years of age. It was found that, after adjustment for a number of myopia risk factors that are hypothesized to be either biological or environmental, first born individuals displayed an increased number of myopes. It can be hypothesized that a relationship between first born individuals and myopia represents a biological pathway or that first born individuals are exposed to a spectrum of myopia risk factors (such as nearwork) that predisposes the group to an increased risk of myopia. If a relationship between birth order and myopia is mediated through a biological pathway, there are a number of hypotheses that can be made about such a pathway's origin. Other studies on a relationship between birth order and susceptibility to allergic disorders have noted that a biological molecule (immunoglobulin E) is associated with both birth order and sensitivity to allergies (Karmaus et al., 2001). It is possible that the levels of a biological molecule that are different in first born individuals lead to susceptibility to myopia later on in life.

A relationship between birth order and myopia could also be explained by a predisposition of first born children to other environmental myopia risk factors. In this thesis it was observed that after adjustment by three environmental myopia risk factors (reading, time spent outdoors/activity and socioeconomic status), a relationship between birth order and myopia was still evident. It is suggested that if a relationship between birth order and myopia is explained by an environmental risk factor, then a new risk factor would be a likely candidate.

A relationship between season of birth and high myopia was examined in an adult cohort. It was observed that subjects born in the summer months were at an increased risk of high myopia in adult life. In animal studies a relationship between light and myopia has been identified (Smith, 1991; Ashby and Schaeffel, 2010) and in human studies a relationship between light and myopia has also been identified (Mandel et al., 2008). It is possible that a relationship between season of birth and myopia is

mediated by light. It is also noted that season of birth may be related to a different causal factor, in turn which is related to the development of myopia but not a consequence of either.

It is noted that the risk factors examined in this thesis may interact with biological susceptibility to myopia. In support of this, risk of myopia was only moderately increased in first born subjects or subjects born in the summer months, indicating that only a subset of individuals were at an increased risk associated with exposure.

8.5 Future directions

An advantage of undertaking calibration of non-cycloplegic autorefraction is an increased chance of success in epidemiological and genetic analysis of myopia in the ALSPAC cohort. In the future more data on refractive error measured by subjective refraction would allow a more precise measure of the bias due to non-cycloplegic autorefraction to be obtained. Furthermore a calibration study of non-cycloplegic autorefraction taken when subjects were 18 (if such measures were taken) would be expected to show a smaller bias (it has been observed that non-cycloplegic autorefraction has a negligible offset for older individuals (Krantz et al., 2010)). Both these strategies would be expected to reduce the amount of uncertainty in measures of non-cycloplegic autorefraction measures of refractive error.

Estimates of heritability can be used to indicate a number of different source of phenotypic variation. A large heritability study of sibling pairs could investigate an appreciable difference in the heritability of refractive error among sibling pairs that varied by zero, one or two intervening siblings. This idea stems from work on the heritability of birth weight by Morton (Morton, 1955). Such a study could shed more light on the possibility of maternal effects and the influence of refractive error development. If heritability was significantly different between such sibling pairs, then evidence for temporary maternal factors and myopia development would be obtained. There is already evidence from the Framingham eye study that resemblance between sibling-sibling pairs varies according to time between births. The Framingham Offspring Eye Study Group found an increased risk of myopia for individuals who had a myopic sibling (The Framingham Offspring Eye Study Group, 1996) with the risk of myopia more than doubling when siblings were born within two years of each other compared to within 10 years.

Inclusion of more measures from the ALSPAC cohort would increase the power to detect a genetic factor underlying myopia development. The ability to detect a genetic factor decreases as the effect size decreases. In other words more measures will be needed to identify a significant difference between groups (defined by genotypic classes). There may be many different loci involved in the pathology of common

disease. For example 18 genetic variants have been associated with type 2 diabetes (Frazer et al., 2009) while only 4% of disease risk has been explained. 40 loci have been associated with human height with only 5% of phenotypic variation explained. Similarly it has been noted that the majority of alleles that have shown evidence of association with common diseases display an estimated increase of risk by a factor of 1.1 to 1.5 (Altshuler et al., 2008). If myopia is similar to other complex diseases for which alleles underlying disease progression have been identified it is expected that a genome-wide association analysis will lead to the identification of many alleles (a number of loci have been identified for common myopia via linkage analysis) with moderate effect size.

More fine mapping or sequencing of regions identified by an initial genome-wide scan would help identify causal variants. A genome-wide study, while being able to identify a small region of DNA where a causative mutation may reside, may not identify a SNP that leads to a change in a phenotype directly rather a SNP of interest may be in linkage disequilibrium with a causal variant. Therefore regions identified in a genome-wide association study, may need to be mapped in greater detail or sequenced to identify a causal mutation (Altshuler et al., 2008). This strategy was undertaken in a genome-wide association study that identified a mutation in complement factor H and age-related macular dystrophy (AMD) (Klein et al., 2005). Approximately 110,000 single nucleotide polymorphisms were scanned across the genome of a number of cases with AMD and controls, with one SNP located in an intron of the complement factor H gene on chromosome 1 identified with a P value of approximately 1 X 10e-7. Patterns of linkage disequilibrium were explored and a region of reduced recombination was identified. This region was re-sequenced and a polymorphism was found in exon 9 of the complement factor H gene that resulted in a protein coding change (a tyrosine-histidine change) that was present on 97% of chromosomes in high risk patients.

Fine mapping or sequencing of non-coding regions identified in a genome-wide association study would help identify causal variants. It has been noted that other genome-wide association studies have found association signals using single nucleotide polymorphisms representing upstream loss of DNA (deletion) in regulator

elements such as the IRGM gene and Crohn's disease (Altshuler et al., 2008). Furthermore it has been noted that 5% of the human genome is evolutionarily conserved and less than one third of this relates to protein coding genes (Altshuler et al., 2008), indicating that the majority of conserved sequence is not located within a gene.

Sequencing of regions identified (not just genotyping more SNPs) would help identify causal variants that are in linkage disequilibrium with a SNP but are due to other structural variation. SNPs are the most abundant genetic marker in the human genome (Wang et al., 1998) but there are other genetic elements that may play a role in disease development. Structural variants include insertions (addition of one or more bases), deletions (loss of a section of DNA), inversions (reversal of a segment of DNA within a chromosome) and copy number variants (identical sequences of DNA repeated on the chromosomes of some individuals but not on others) (King et al., 2006; Frazer et al., 2009) among others. Many rare structural variants are found at an increased frequency in patients with schizophrenia compared to healthy controls (Walsh et al., 2008). It has been noted that SNPs used in a genome-wide association study may be in reasonable linkage disequilibrium with structural variants (Frazer et al., 2009) and Craddock et al. (Craddock et al., 2010) noted that copy number variants may be well represented by SNPs.

Epidemiology risk factors for a disease may act on disease progression independent of genetic factors. In this thesis an association was found between birth order and season of birth and myopia. It also possible that genes and the environment interact to modulate risk of disease progression. Risk of myopia was only moderately increased in first born subjects or subjects born in the summer months, indicating that only a subset of individuals were at an increased risk associated with exposure. Identification of other factors that modulate risk of these factors would help identify groups most at risk. Analysis of birth order and myopia in younger and older age groups would establish if the association was consistent. If the association is due to exposures that occur during school years an association between birth order and myopia before the age of seven would not be expected (and it could possibly get stronger by age 15). If the association between birth order and myopia is due mainly to factors at birth, a

strong association may be present at a young age before other exposure to myopia risk factors has occurred.

8.6 Summary

In summary this thesis set out to better understand the genetics and epidemiology of myopia in the ALSPAC cohort. It was observed that the use of non-cycloplegic autorefraction leads to a negative offset in measures of refractive error which can be partly corrected by calibration with a more accurate measure. The heritability of refractive error was assessed and it was observed that both genetics and the environment play a role in the development of refractive error. A genome-wide association study was also undertaken of myopia, average spherical equivalent and a number of their ocular determinants (axial length and corneal curvature). Evidence of association was found for a number of genomic locations and these traits and efforts to replicate the findings and to increase the number of subjects in the study via extra genotyping is ongoing. A relationship between birth order and myopia in later life was examined in the ALSPAC cohort. It was observed that first born individuals, after adjustment for a number of myopia risk factors, displayed an increased number of myopes. A relationship between season of birth and myopia in adult life was also examined in a cohort from the UK. It was observed that subjects born in the summer months were at an increased risk of high myopia.

Appendix A

Gene Mapping

A.1 Cytogenetic locations

Table A.1) lists a summary of studies aimed to reveal a cytogenetic location for high myopia ordered by chromosomal region. Each study used genome-wide microsatellite linkage analysis except for cytogenetic region 11q24.1 which was identified by a genome-wide single nucleotide polymorphism (SNP) association study and 21q22.3 which was identified by case-control analysis of SNPs located in a region previously prioritised by an unpublished genome-wide scan. Table A.2) lists a summary of studies aimed to reveal a cytogenetic location for common myopia ordered by chromosomal region. Each study used genome-wide microsatellite linkage analysis. Ukn (unknown due to information of study based on abstract or article is written in foreign language). NA (not applicable, in this case no OMIM name has been assigned). AD (autosomal dominant), AR (autosomal recessive). Criteria for independent replication: similar finding by different research group and different study population. Studies marked with an ampersand (&) were independently replicated in a high/common myopia cohort, studies marked with an asterisk (*) have been replicated in different study samples by the research group that made the original discovery.

High myopia

Chromosomal region	OMIM name (number)	Study population (Inheritance pattern if indicated)	Independently replicated	Year	Reference
2q37.1	MYP12 (609995)	Large multigenerational family (AD)	Yes [&]	2005	(Paluru et al., 2005); (Chen et al., 2007b)
4q22-q27	MYP11 (609994)	Large multigenerational family (AD)	No	2005	(Zhang et al., 2005)
5p15.33-p15.2	MYP16 (612554)	High myopic families (AD)	No	2008	(Lam et al., 2008a)
7p15	MYP4/MYP17 (608367/608367)	High myopic families	Yes [&]	2008	(Ciner et al., 2008; Paget et al., 2008a)
10q21.1	MYP15 (612717)	Large multigenerational family (Hutterite) (AD)	No	2007	(Nallasamy et al., 2007)
11q24.1	NA	Case (high myopes) control (general population)	No	2009	(Nakanishi et al., 2009b)
12q21	MYP3 (603221)	Large multigenerational family (AD)	Yes	1998	(Young et al., 1998a; Nurnberg et al., 2008; Li et al., 2009b)
14q22.1-q24.2	NA	Multigenerational family (AR)	No	2009	(Yang et al., 2009)
15q12-13	NA	Ukn (AD)	No	2007	(Yu et al., 2007)
17q21-22	MYP5 (608474)	Large multigenerational family (AD)	No	2003	(Paluru et al., 2003)
18p11.31	MYP2 (160700)	High myopic families (AD)	Yes	1998	(Young et al., 1998b; Heath et al., 2001; Lam et al., 2003b)
21q22.3	NA	Case (high myopes) control (general population)	No	2009	(Nishizaki et al., 2008)
Xq23-25	MYP13 (300613)	Large multigenerational family (X-linked recessive)	*	2006	(Zhang et al., 2006; Zhang et al., 2007a)
Xq28	MYP1 (310460)	Large multigenerational family (X-linked recessive)	Yes	2004	(Schwartz et al., 1990; Young et al., 2004)

Table A.1) A summary of studies aimed to reveal a cytogenetic location for high myopia.

Common myopia

Chromosomal region	OMIM name (number)	Cohort	Study population	Independent replication	Year	Reference
1p36	MYP14 (610320)	Myopia Family Study	Myopic families (Ashkenazi Jewish)	*	2006	(Wojciechowski et al., 2006; Wojciechowski, Bailey-Wilson and Stambolian, 2009a)
2q37.1	MYP12 (609995)	Genes in Myopia Study (GEM)	Myopic families	Yes [&]	2007	(Paluru et al., 2005); (Chen et al., 2007b)
3q26	MYP8 (609257)	Twin Eye Study	Dizygotic twins	*	2004	(Hammond et al., 2004; Andrew et al., 2008)
4q12	MYP9 (609258)	Twin Eye Study	Dizygotic twins	No	2004	(Hammond et al., 2004)
4q21	Close to MYP11	Myopia Family Study	Myopic families	No	2009	(Wojciechowski et al., 2009b)
5q	NA	Myopia Family Study	Myopic families (Old Order Amish)	No	2009	(Wojciechowski et al., 2009b)
7p15	MYP4/MYP17 (608367/608367)	Myopia Family Study	Myopic families (African American)	Yes [&]	2008	(Ciner et al., 2008; Paget et al., 2008a)
8p23	MYP10 (609259)	Twin Eye Study	Dizygotic twins	Yes	2004	(Hammond et al., 2004; Stambolian et al., 2005)
11p13	MYP7 (609256)	Twin Eye Study	Dizygotic twins	No	2004	(Hammond et al., 2004)
12q24	NA	Myopia Family Study	Myopic families	No	2009	(Wojciechowski et al., 2009b)
22q12	MYP6 (608908)	Myopia Family Study	Myopic families (Ashkenazi Jewish)	Yes*	2004	(Stambolian et al., 2004; Stambolian et al., 2006; Klein et al., 2007)

Table A.2) A summary of studies aimed to reveal a cytogenetic location for common myopia.

A.2 Candidate gene analysis

Studies listed in 'candidate gene analysis' were carried out later than the year 2000 using a number of different methodologies; case-control association, family based association, quantitative trait association, linkage or co-segregation of mutations with phenotype. Studies also used varied genetic makers; microsatellites, tagging SNPs or markers identified by direct sequencing. There are four divisions made: candidate genes implicated via earlier mapping studies, genes important in the maintenance of the extracellular matrix (ECM), signalling proteins and genes implicated in ocular health and development. A number of positive associations have been reported but no underlying mutation that causes myopia has been identified.

Genes implicated via earlier mapping studies

Gene name	Gene symbol	Cytogenetic location	Linkage peak (OMIM number)	Type of myopia	Evidence of association	Number of studies	Reference
Testis-expressed gene on Xq28	TEX28	Xq28	MYP1 (310460)	High	Yes	1	(Metlapally et al., 2009b)
Laminin, alpha-1	LAMA1	18p11.31	MYP2 (160700)	High	No	1	(Sasaki et al., 2007)
Lipin 2	LPIN2	18p11.3	MYP2 (160700)	High	No	1	(Zhou and Young, 2005)
Transforming growth factor Beta induced factor	TGIF	18p11.3	MYP2 (160700)	High	Mixed	5	(Lam et al., 2003a; Scavello et al., 2004; Hasumi et al., 2006; Pertile et al., 2008; Wang et al., 2009b)
Decorin	DCN	12q21.3	MYP3 (603221)	High	Mixed	2	(Wang et al., 2006; Zhang et al., 2009)
Dermatan sulfate proteoglycan 3	DSPG3	12q21	MYP3 (603221)	High	No	2	(Wang et al., 2009a)
Lumican	LUM	12q21.3	MYP3 (603221)	High	Mixed	6	(Paluru et al., 2004; Wang et al., 2006; Majava et al., 2007; Wang et al., 2009b; Zhang et al., 2009; Lin et al., 2010)
Collagen I, alpha-1 polypeptide	COL1A1	17q21.31-q22	MYP5 (608474)	Common and high	No	5	(Inamori et al., 2007; Liang et al., 2007; Metlapally et al., 2009a; Nakanishi et al., 2009a; Vatavuk et al., 2009)
Paired box gene 6	PAX6	11p13	MYP7 (609256)	Common and high	Mixed	6	(Hewitt et al., 2007; Mutti et al., 2007a; Simpson et al., 2007; Tsai et al., 2008; Han et al., 2009; Ng et al., 2009)
Sry-related HMG-box gene 2	SOX2	3q26.3-q27	MYP8 (609257)	Myopia and hypermetropia	No	1	(Simpson et al., 2007)
Retinal pigment epithelium-derived rhodopsin homolog	RRH	4q	MYP11 (609994)	High	No	1	(Zhang et al., 2005)

Table A.3) A summary of genes located under a linkage peak identified in previous gene mapping studies.

Extracellular matrix

Gene name	Gene symbol	Cytogenetic location	Prior evidence	Type of myopia	Evidence of association	Number of studies	Reference
Collagen II, alpha-1 polypeptide	COL2A1	12q13.11-q13.2	Mutations in the COL2A1 gene cause Stickler syndrome, a disorder that is associated with high myopia (Ahmad et al., 1991).	Any myopia	Yes	2	(Mutti et al., 2007a; Metlapally et al., 2009a)
Fibromodulin	FMOD	1q32.1	Double null mice for lumican and fibromodulin show significantly larger axial length than normal mice (Chakravarti et al., 2003).	High	No	2	(Paluru et al., 2004; Lin et al., 2009b)
Matrix metalloproteinase 1	MMP1	11q22-q23	Increased expression of MMPs in sclera of form deprived animals (Rada and Brenza, 1995).	High and common	Mixed	3	(Hall et al., 2009; Nakanishi et al., 2010; Wojciechowski et al., 2010)
Matrix metalloproteinase 2	MMP2	16q13	MMP2 activity is increased in form deprived eyes of the tree shrew (Guggenheim and McBrien, 1996).	High and common	Mixed	2	(Nakanishi et al., 2010; Wojciechowski et al., 2010)

Table A.4a) A summary of genes important in the maintenance of the extracellular matrix (ECM). These include components of the ECM (collagen), molecules that support the ECM (laminin), enzymes which interact with its components (matrix metalloproteinases and their inhibitors) and leucine rich proteins (opticin, fibromodulin, lumican and nyctalopin). Decorin, dermatan sulfate proteoglycan 3, lumican, collagen I alpha-1 polypeptide and laminin have been tested as candidate genes but are not listed here because they are listed under genes implicated in earlier mapping studies.

Extracellular matrix continued

Gene name	Gene symbol	Cytogenetic location	Prior evidence	Type of myopia	Evidence of association	Number of studies	Reference
Matrix metalloproteinase 3	MMP3	11q23	MMP3 mRNA levels are reduced after monocular deprivation in the tree shrew (Siegwart and Norton, 2002).	High and common	Mixed	3	(Liang et al., 2006; Hall et al., 2009; Nakanishi et al., 2010)
Matrix metalloproteinase 9	MMP9	20q11.2-q13.1	MMP9 is a zinc metalloproteinase similar to MMPs 1-3; all of latter have been implicated in myopia development.	Common	Yes	1	(Hall et al., 2009)
Nyctalopin	NYX	Xp11.4	Mice with mutations in the mouse ortholog of NYX show a faster myopic shift under form deprivation than wild type mice (Pardue et al., 2008).	High	Yes	1	(Zhang et al., 2007b)
Opticin	OPTC	1q32	Opticin forms part of the extracellular matrix which is important in sclera remodelling. Changes in the sclera are apparent in the development of myopia (Norton and Rada, 1995).	High	Mixed	2	(Majava et al., 2007; Wang et al., 2009a)
Tissue inhibitor of metalloproteinase 1	TIMP1	Xp11.3-p11.23	Eyes recovering from monocular deprivation have more TIMP1 than normal eyes in the tree shrew (Siegwart and Norton, 2001).	High	No	1	(Liang et al., 2006)

Table A.4b) A summary of genes important in the maintenance of the extracellular matrix (ECM). See above for description.

Signalling proteins

Gene name	Gene symbol	Cytogenetic location	Prior evidence	Type of myopia	Evidence of association	Number of studies	Reference
Bone morphogenetic protein 2 kinase	BMP2K	?	Retinas of form deprived chicks show downregulation of bone morphogenetic protein 2 (McGlinn et al., 2007).	High	Yes	1	(Liu et al., 2009)
Cholinergic receptor, muscarinic 1	CHRM1	11q13	Muscarinic antagonists inhibit form deprived myopia (Cottriall and McBrien, 1996).	High	Yes	1	(Lin et al., 2009a)
Early growth response 1	EGR1	5q31.1	ERG1 null mice have longer eyes than wild type (Schippert et al., 2007).	High	No	1	(Li et al., 2008)
Fibroblast growth factor 2 (basic)	FGF2	4q25-q27	FGF2 intravitreal injections reduce form deprived myopia in chicks (Rohrer and Stell, 1994).	High	No	1	(Lin et al., 2009b)
Glutamate receptor metabotropic 6 gene	GRM6	5q35	GRM6 mutations are found in patients with congenital night blindness. Myopia is sometimes found in such cases (Dryja et al., 2005; O'Connor et al., 2006).	High	Yes	1	(Xu et al., 2009)

Table A.5a) A summary of genes coding for signalling proteins. Such proteins include growth factors (fibroblast growth factor 2, hepatocyte growth factor, insulin-like growth factor 1, transforming growth factor beta 1, transforming growth factor beta 2 and transforming growth factor beta induced factor), transcription factors (early growth response 1, paired box gene 6), neurotransmitter receptors (cholinergic receptor muscarinic 1, glutamate receptor metabotropic 6) and nuclear receptors (retinoic acid receptor alpha and beta). Transforming growth factor beta induced factor and paired box gene 6 are listed under genes implicated in earlier mapping studies.

Signalling proteins continued

Gene name	Gene symbol	Cytogenetic location	Prior evidence	Type of myopia	Evidence of association	Number of studies	Reference
Insulin-like growth factor 1	IGF	12q22-q24.1	Intravitreal injection of insulin stimulates myopia in chicks (Feldkaemper, Neacsu and Schaeffel, 2008).	High	Yes	1	(Metlapally et al., 2010)
Retinoic acid receptor alpha	RARA	17q21.1	Retinoic acid synthesis in the retina is correlated with vitreous chamber length in form deprived marmosets (Troilo et al., 2006).	High and common	No	1	(Veerappan et al., 2009)
Retinoic acid receptor beta	RARB	3p24	Dietary retinoic acid increases eye length in chicks (McFadden et al., 2006).	High	No	1	(Ding et al., 2010)
Transforming growth factor, beta 1	TGFB1	19q13.1	TGFB1 messenger RNA and protein is reduced in form deprived chick eyes (Honda et al., 1996).	High	Mixed	4	(Lin et al., 2006; Hayashi et al., 2007; Wang et al., 2009b; Zha et al., 2009)
Transforming growth factor, beta 2	TGF-beta2	1q41	TGFB2 mRNA levels are decreased in monocularly deprived eyes of the tree shrew (Jobling et al., 2004a).	High	Yes	1	(Lin et al., 2009b)
Hepatocyte growth factor	HGF	7q21.2	The mouse ortholog of HGF is located under a linkage peak for eye weight (Zhou and Williams, 1999).	High and common	Mixed	5	(Han et al., 2006; Schache et al., 2009; Wang et al., 2009b; Yanovitch et al., 2009; Veerappan et al., 2010)

Table A.5b) A summary of genes coding for signalling proteins. See above for description.

Ocular development and health

Gene name	Gene symbol	Cytogenetic location	Prior evidence	Type of myopia	Evidence of association	Number of studies	Reference
Myocilin	MYOC	1q24.3-q25.2	Mutations in MYOC cause open angle glaucoma (Stone et al., 1997). There is an increased prevalence of myopia in glaucoma patients (Mitchell et al., 1999).	High	Mixed	4	(Leung et al., 2000; Tang et al., 2007; Vatavuk et al., 2009; Zayats et al., 2009)

Table A.6) A summary of genes important to the development and health of the eye. These include myocilin (mutations in which are responsible for a juvenile form of open angle glaucoma) and PAX6 (a transcriptional regulator of oculogenesis). PAX6 is listed under genes implicated in earlier mapping studies.

Appendix B

Myopia Risk Factors

B.1 Risk factors

A list of a number of risk factors and references which investigate an association between a risk factor and myopia or refractive error are given below. Except for two references (intelligence (Williams et al., 1988) and birth order (Peckham et al., 1977)) all studies were carried out after 1999. Each group (marked in bold type) is used loosely and does not necessarily represent the nature of an association between risk factor and disease. For example intelligence is listed under behavioural; however it may be that associations between myopia and intelligence are due to changes in behaviour (i.e. more time spent reading) or due to a biological predisposition towards myopia *and* greater intelligence test scores. Similarly, family history is listed under familial as it may be that an association between family history and myopia is due to genetic or environmental factors. Over twenty risk factors are listed.

Type	Risk Factor	Reference
Familial	Family history	(Mutti et al., 2002; Williams et al., 2008a; Low et al., 2010)
Behavioural	Nearwork	(Mutti et al., 2002; Ip et al., 2008b; Low et al., 2010)
	Outdoor activity/activity	(Mutti et al., 2002; Deere et al., 2009; Dirani et al., 2009b)
	School achievement	(Mutti et al., 2002; Saw et al., 2007)
	Intelligence	(Williams et al., 1988; Saw et al., 2004b; Williams et al., 2008a)
Physical	Age	(Attebo et al., 1999; Maul et al., 2000; Pokharel et al., 2000; Zhao et al., 2000)
	Gender	(Zhao et al., 2000; Goh et al., 2005; Giordano et al., 2009)
	Cataract	(Bourne et al., 2004; Saw et al., 2008)
	Height	(Wong et al., 2001; Saw et al., 2002b; Wu et al., 2007)
	Weight	(Wong et al., 2001; Saw et al., 2002b; Wu et al., 2007)
	Intraocular pressure	(Lam et al., 1998; Attebo et al., 1999)
Socio-economic	Occupation	(Shimizu et al., 2003)
	Income	(Wong et al., 2000; Saw et al., 2002c; Shimizu et al., 2003)
	Urban environment	(Xu et al., 2005; Ip et al., 2008a; Zhang et al., 2010)
	Parental education	(Dandona et al., 2002; Murthy et al., 2002; Goh et al., 2005)
	Education	(Wong et al., 2000; Cheng et al., 2003; Shimizu et al., 2003)
Light	Natural light	(Vannas et al., 2003; Mandel et al., 2008; McMahon et al., 2009)
	Night light use	(Quinn et al., 1999; Gwiazda et al., 2000; Zadnik et al., 2000)
Birth	Birth order	(Peckham et al., 1977; Rudnicka et al., 2008)
	Birthweight	(Goss, 2006; Varghese et al., 2009)
	Breast feeding	(Chong et al., 2005; Rudnicka et al., 2008)
	Gestational age	(Goss, 2006; Varghese et al., 2009)

Table B.1) A list of a number of myopia risk factors

Appendix C

List of Publications

From this Thesis

Papers

McMahon G, Zayats T, Chen Y P, Prashar A, Williams C, and Guggenheim J A (2009) Season of birth, daylight hours at birth, and high myopia. *Ophthalmology* 116: 468-473.

Conferences

McMahon G, Northstone K, Zayats T, Guggenheim J A, and Williams C (2009) Birth order and myopia are associated in two UK cohorts. *Association for Research in Vision and Ophthalmology (ARVO)*: Fort Lauderdale, 3-7 May.

McMahon G, StPourcain B, Crawford M, Carmichael D, Northstone K, Guggenheim J A, and Williams C (2010) Data collection and analysis of subjective refractions in the Avon Longitudinal Study of Parents and Children (ALSPAC). *13th International Myopia Conference (IMC)*: Tubingen, 26-29 July.

Williams C, **McMahon G**, StPourcain B, Northstone K, Guggenheim J A (2010) A genome-wide association study in the Avon Longitudinal Study of Parents and Children (ALSPAC). *13th International Myopia Conference (IMC)*: Tubingen, 26-29 July.

StPourcain B, Whitehouse A, Warrington N, Golding J, Steer C, Kemp J, **McMahon G**, Timpson N J, Evans D M, Ring S M, Deloukas P, Palmer L, Pennell C, Davey Smith G (2010) Genome-wide meta-analysis of pragmatic communication skills. *60th Annual Meeting of the American Society of Human Genetics (ASHG)*: Washington DC, 2-6 November.

Paternoster L, Toma A M, Zhurov A I, Kemp J, Davey Smith G, Glaser B, **McMahon G**, Deloukas P, Ring S M, Timpson N, Richmond S, Evans D M (2010) The identification of SNPs associated with facial morphological traits in a genome-wide association study. *60th Annual Meeting of the American Society of Human Genetics (ASHG)*: Washington DC, 2-6 November.

During this Thesis

Papers

Zayats T, Yanovitch T, Creer R C, **McMahon G**, Li Y J, Young T L, and Guggenheim J A (2009) Myocilin polymorphisms and high myopia in subjects of European origin. *Mol Vis* 15: 213-222.

Flanagan J M, **McMahon G**, Brendan Chia S H, Fitzpatrick P, Tighe O, O'Neill C, Briones P et al. (2010) The role of human demographic history in determining the distribution and frequency of transferase-deficient galactosaemia mutations. *Heredity* 104: 148-154.

References

Ahmad N N, Ala-Kokko L, Knowlton R G, Jimenez S A, Weaver E J, Maguire J I, Tasman W et al. (1991) Stop codon in the procollagen II gene (COL2A1) in a family with the Stickler syndrome (arthro-ophthalmopathy). *Proc Natl Acad Sci U S A* 88: 6624-6627.

Almasy L, and Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 62: 1198-1211.

Altshuler D, Daly M, and Lander E (2008) Genetic mapping in human disease. *Science* 322: 881-888.

Andrew T, Maniatis N, Carbonaro F, Liew S H, Lau W, Spector T D, and Hammond C J (2008) Identification and replication of three novel myopia common susceptibility gene loci on chromosome 3q26 using linkage and linkage disequilibrium mapping. *PLoS Genet* 4: e1000220.

Antonarakis S E (1994) Genome linkage scanning: systematic or intelligent? *Nat Genet* 8: 211-212.

Ashby R S, and Schaeffel F (2010) The effect of bright light on lens-compensation in chicks. *Invest Ophthalmol Vis Sci* 51: 5247-5253.

Attebo K, Ivers R, and Mitchell P (1999) Refractive errors in an older population: the Blue Mountains Eye Study. *Ophthalmology* 106: 1066-1072.

Aylsworth A S (1998) Defining Disease Phenotypes. In: Haines J L, and Pericak-Vance M A [eds.] *Approaches to gene mapping in complex human diseases*. New York: Wiley-Liss, pp. 53-76.

Barrett J C, Fry B, Maller J, and Daly M J (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263-265.

Bartlett J D, Jaanus S D, and Blaho K E (2001) *Clinical ocular pharmacology*. 4th ed. Boston: Butterworth-Heinemann.

Bear J C (1991) Epidemiology and genetics of refractive anomalies. In: Grosvenor T P, and Flom M C [eds.] *Refractive anomalies : research and clinical applications*. Boston ; London: Butterworth-Heinemann, pp. 57-80.

Biino G, Palmas M, Corona C, Prodi D, Fanciulli M, Sulis R, Serra A et al. (2005) Ocular refraction: heritability and genome-wide search for eye morphometry traits in an isolated Sardinian population. *Hum Genet* 116: 152-159.

Bird A C (1998) Retinitis Pigmentosa. In: Traboulsi E I [ed.] *Genetic diseases of the eye*. New York: Oxford University Press, pp. 325-346.

Bland J, and Altman D (2010) Statistical methods for assessing agreement between two methods of clinical measurement. *Int J Nurs Stud* 47: 931-936.

Botstein D, and Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 33 Suppl: 228-237.

Bourne R, Dineen B, Ali S, Noorul Huq D, and Johnson G (2004) Prevalence of refractive error in Bangladeshi adults: results of the National Blindness and Low Vision Survey of Bangladesh. *Ophthalmology* 111: 1150-1160.

Campbell C E, Benjamin W J, and Howland H C (2006) Objective Refraction: Retinoscopy, Autorefraction, and Photorefraction. In: Benjamin W J, and Borish I M [eds.] *Borish's clinical refraction C18*. (2nd edn.) St. Louis, Miss.: Butterworth Heinemann/Elsevier, pp. 682-764.

Chakravarti S, Paul J, Roberts L, Chervoneva I, Oldberg A, and Birk D E (2003) Ocular and scleral alterations in gene-targeted lumican-fibromodulin double-null mice. *Invest Ophthalmol Vis Sci* 44: 2422-2432.

Chen C Y, Scurrah K J, Stankovich J, Garoufalis P, Dirani M, Pertile K K, Richardson A J et al. (2007a) Heritability and shared environment estimates for myopia and associated ocular biometric traits: the Genes in Myopia (GEM) family study. *Hum Genet* 121: 511-520.

Chen C Y, Stankovich J, Scurrah K J, Garoufalis P, Dirani M, Pertile K K, Richardson A J et al. (2007b) Linkage replication of the MYP12 locus in common myopia. *Invest Ophthalmol Vis Sci* 48: 4433-4439.

Cheng C, Hsu W, Liu J, Tsai S, and Chou P (2003) Refractive errors in an elderly Chinese population in Taiwan: the Shihpai Eye Study. *Invest Ophthalmol Vis Sci* 44: 4630-4638.

Cheng D, Schmid K L, and Woo G C (2007) Myopia prevalence in Chinese-Canadian children in an optometric practice. *Optom Vis Sci* 84: 21-32.

Chew S-J, Hoh S-T, Tan J, and Cheng H-M (1998) Muscarinic Antagonists for Myopia Control. In: Tokoro T [ed.] *Myopia updates : proceedings of the 6th International Conference on Myopia*. Tokyo ; New York: Springer, pp. 155-162.

Chong Y S, Liang Y, Tan D, Gazzard G, Stone R A, and Saw S M (2005) Association between breastfeeding and likelihood of myopia in children. *JAMA* 293: 3001-3002.

Choong Y F, Chen A H, and Goh P P (2006) A comparison of autorefraction and subjective refraction with and without cycloplegia in primary school children. *Am J Ophthalmol* 142: 68-74.

Churchill G A, and Doerge R W (1994) Empirical Threshold Values for Quantitative Trait Mapping. *Genetics* 138: 963–971.

Ciner E, Wojciechowski R, Ibay G, Bailey-Wilson J E, and Stambolian D (2008) Genomewide scan of ocular refraction in African-American families shows significant linkage to chromosome 7p15. *Genet Epidemiol* 32: 454-463.

Ciuffreda K J (2006) Accommodation, the Pupil, and Presbyopia. In: Benjamin W J, and Borish I M [eds.] *Borish's clinical refraction.* (2nd edn.) St. Louis, Miss.: Butterworth Heinemann/Elsevier, pp. 93-144.

Cohen Y, Belkin M, Yehezkel O, Avni I, and Polat U (2008) Light intensity modulates corneal power and refraction in the chick eye exposed to continuous light. *Vision Res* 48: 2329-2335.

Cottriall C L, and McBrien N A (1996) The M1 muscarinic antagonist pirenzepine reduces myopia and eye enlargement in the tree shrew. *Invest Ophthalmol Vis Sci* 37: 1368-1379.

Coupry I, Sibon I, Mortemousque B, Rouanet F, Mine M, and Goizet C (2010) Ophthalmological features associated with COL4A1 mutations. *Arch Ophthalmol* 128: 483-489.

Craddock N, Hurles M, Cardin N, Pearson R, Plagnol V, Robson S, Vukcevic D et al. (2010) Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* 464: 713-720.

Curtin B J, Iwamoto T, and Renaldo D P (1979) Normal and staphylomatous sclera of high myopia. An electron microscopic study. *Arch Ophthalmol* 97: 912-915.

Czepita D, Zejmo M, and Mojsa A (2007) Prevalence of myopia and hyperopia in a population of Polish schoolchildren. *Ophthalmic Physiol Opt* 27: 60-65.

Dandona R, Dandona L, Srinivas M, Sahare P, Narsaiah S, Muñoz S, Pokharel G et al. (2002) Refractive error in children in a rural population in India. *Invest Ophthalmol Vis Sci* 43: 615-622.

Deere K, Williams C, Leary S, Mattocks C, Ness A, Blair S N, and Riddoch C (2009) Myopia and later physical activity in adolescence: a prospective study. *Br J Sports Med* 43: 542-544.

Deng L, Gwiazda J, and Thorn F (2010) Children's refractions and visual activities in the school year and summer. *Optom Vis Sci* 87: 406-413.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P et al. (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380: 152-154.

Ding Y, Chen X, Yan D, Xue A, Lu F, Qu J, and Zhou X (2010) Association analysis of retinoic acid receptor beta (RARbeta) gene with high myopia in Chinese subjects. *Mol Vis* 16: 855-861.

Dirani M, Chamberlain M, Shekar S N, Islam A F, Garoufalis P, Chen C Y, Guymer R H et al. (2006) Heritability of refractive error and ocular biometrics: the Genes in Myopia (GEM) twin study. *Invest Ophthalmol Vis Sci* 47: 4756-4761.

Dirani M, Islam F M, and Baird P N (2009a) The role of birth weight in myopia - the genes in myopia twin study. *Ophthalmic Res* 41: 154-159.

Dirani M, Shekar S N, and Baird P N (2008) Adult-onset myopia - the Genes in Myopia (GEM) twin study. *Invest Ophthalmol Vis Sci* 49: 3324-3327.

Dirani M, Tong L, Gazzard G, Zhang X, Chia A, Young T L, Rose K A et al. (2009b) Outdoor activity and myopia in Singapore teenage children. *Br J Ophthalmol* 93: 997-1000.

Drack A V (1998) Inheritance of Refractive Errors. In: Traboulsi E I [ed.] *Genetic diseases of the eye*. New York: Oxford University Press, pp. 163-176.

Dryja T P, McGee T L, Berson E L, Fishman G A, Sandberg M A, Alexander K R, Derlacki D J et al. (2005) Night blindness and abnormal cone electroretinogram ON responses in patients with mutations in the GRM6 gene encoding mGluR6. *Proc Natl Acad Sci U S A* 102: 4884-4889.

Dudbridge F, and Gusnanto A (2008) Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* 32: 227–234.

Edwards M H (1998a) A Longitudinal Study of the Development of Myopia in Chinese Children: A Life Table Treatment. In: Tokoro T [ed.] *Myopia updates : proceedings of the 6th International Conference on Myopia.* Tokyo ; New York: Springer, pp. 48-52.

Edwards M H (1998b) Myopia:definitions, classifications and economic implications. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork.* Oxford: Butterworth-Heinemann, pp. 1-12.

Erickson P (1991) Optical Components Contributing To Refractive Anomalies. In: Grosvenor T P, and Flom M C [eds.] *Refractive anomalies : research and clinical applications.* Boston ; London: Butterworth-Heinemann, pp. 199-218.

Falconer D S, and Mackay T F C (1996) *Introduction to quantitative genetics.* 4th ed. Harlow: Longman.

Farbrother J E, Kirov G, Owen M J, and Guggenheim J A (2004a) Family aggregation of high myopia: estimation of the sibling recurrence risk ratio. *Invest Ophthalmol Vis Sci* 45: 2873-2878.

Farbrother J E, Kirov G, Owen M J, Pong-Wong R, Haley C S, and Guggenheim J A (2004b) Linkage analysis of the genetic loci for high myopia on 18p, 12q, and 17q in 51 UK families. *Invest Ophthalmol Vis Sci* 45: 2879-2885.

Farrer L A, and Cupples L A (1998) Determining the Genetic Component of a Disease. In: Haines J L, and Pericak-Vance M A [eds.] *Approaches to gene mapping in complex human diseases*. New York: Wiley-Liss, pp. 93-130.

Feldkaemper M P, Neacsu I, and Schaeffel F (2008) Insulin acts as a powerful stimulator of axial myopia in chicks. *Invest Ophthalmol Vis Sci* 50: 13-23.

Flanagan J M, McMahon G, Brendan Chia S H, Fitzpatrick P, Tighe O, O'Neill C, Briones P et al. (2010) The role of human demographic history in determining the distribution and frequency of transferase-deficient galactosaemia mutations. *Heredity* 104: 148-154.

Forastiere F, Agabiti N, Corbo G M, Dell'Orco V, Porta D, Pistelli R, Levenstein S et al. (1997) Socioeconomic status, number of siblings, and respiratory infections in early life as determinants of atopy in children. *Epidemiology* 8: 566-570.

Fotedar R, Rochtchina E, Morgan I, Wang J J, Mitchell P, and Rose K A (2007) Necessity of cycloplegia for assessing refractive error in 12-year-old children: a population-based study. *Am J Ophthalmol* 144: 307-309.

Frazer K, Murray S, Schork N, and Topol E (2009) Human genetic variation and its contribution to complex traits. *Nat Rev Genet* 10: 241-251.

Frazer K A, Ballinger D G, Cox D R, Hinds D A, Stuve L L, Gibbs R A, Belmont J W et al. (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449: 851-861.

Funarunart P, Tengtrisorn S, Sangsupawanich P, and Siangyai P (2009) Accuracy of noncycloplegic refraction in primary school children in southern Thailand. *J Med Assoc Thai* 92: 806-811.

Gabriel S, Schaffner S, Nguyen H, Moore J, Roy J, Blumenstiel B, Higgins J et al. (2002) The structure of haplotype blocks in the human genome. *Science* 296: 2225-2229.

Garner L F, Yap M K H, Smith G, Frith M J, and Kinnear R F (1998) Changes in the Crystalline Lens in Schoolchildren Aged 6 to 16 Years. In: Tokoro T [ed.] *Myopia updates : proceedings of the 6th International Conference on Myopia*. Tokyo ; New York: Springer, pp. 86-91.

Gentle A, Liu Y, Martin J E, Conti G L, and McBrien N A (2003) Collagen gene expression and the altered accumulation of scleral collagen during the development of high myopia. *J Biol Chem* 278: 16587-16594.

Ghose S, Nayak B K, and Singh J P (1986) Critical evaluation of the NR-1000F Auto Refractometer. *Br J Ophthalmol* 70: 221-226.

Gilmartin B (1998) Autonomic correlates of near-vision in emmetropia and myopia. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork*. Oxford ; Boston: Butterworth-Heinemann, pp. 117-146.

Giordano L, Friedman D S, Repka M X, Katz J, Ibironke J, Hawes P, and Tielsch J M (2009) Prevalence of refractive error among preschool children in an urban population: the Baltimore Pediatric Eye Disease Study. *Ophthalmology* 116: 739-746.

Goh P, Abqariyah Y, Pokharel G, and Ellwein L (2005) Refractive error and visual impairment in school-age children in Gombak District, Malaysia. *Ophthalmology* 112: 678-685.

Golding J (2004) The Avon Longitudinal Study of Parents and Children (ALSPAC)--study design and collaborative opportunities. *Eur J Endocrinol* 151 Suppl 3: U119-123.

Golding J, Pembrey M, and Jones R (2001) ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* 15: 74-87.

Gonzalez Blanco F, Sanz Fernandez J C, and Munoz Sanz M A (2008) Axial length, corneal radius, and age of myopia onset. *Optom Vis Sci* 85: 89-96.

Gordis L (2009) *Epidemiology*. 4th ed. Philadelphia: Elsevier/Saunders.

Gorin M B (1998) Genetics of Age-Related Maculopathy. In: Traboulsi E I [ed.] *Genetic diseases of the eye*. New York: Oxford University Press, pp. 407-434.

Goss D A (1991) Childhood Myopia. In: Grosvenor T P, and Flom M C [eds.] *Refractive anomalies : research and clinical applications*. Boston ; London: Butterworth-Heinemann, pp. 81-103.

Goss D A (2006) Development of the Ametropias. In: Benjamin W J, and Borish I M [eds.] *Borish's clinical refraction*. (2nd edn.) St. Louis, Miss.: Butterworth Heinemann/Elsevier, pp. 56-92.

Gould D B, Marchant J K, Savinova O V, Smith R S, and John S W (2007) Col4a1 mutation causes endoplasmic reticulum stress and genetically modifiable ocular dysgenesis. *Hum Mol Genet* 16: 798-807.

Grosvenor T (1998) Myopia control procedures. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork*. Oxford ; Boston: Butterworth-Heinemann, pp. 173-192.

Grosvenor T P (1991) Changes In Spherical Refraction During The Adult Years. In: Grosvenor T P, and Flom M C [eds.] *Refractive anomalies : research and clinical applications*. Boston ; London: Butterworth-Heinemann, pp. 131-145.

Grosvenor T P (2002) *Primary care optometry*. 4th ed. Boston: Butterworth-Heinemann.

Grosvenor T P, and Goss D A (1999) *Clinical management of myopia*. Boston ; Oxford: Butterworth-Heinemann.

Guggenheim J A, Hill C, and Yam T F (2003) Myopia, genetics, and ambient lighting at night in a UK sample. *Br J Ophthalmol* 87: 580-582.

Guggenheim J A, and McBrien N A (1996) Form-deprivation myopia induces activation of scleral matrix metalloproteinase-2 in tree shrew. *Invest Ophthalmol Vis Sci* 37: 1380-1395.

Gulcher J (2007) Microsatellite Markers for Linkage And Association Studies. In: Weiner M P, Gabriel S B, and Stephens J C [eds.] *Genetic variation : a laboratory manual.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, pp. 253-262.

Gurwitz D, Fortier I, Lunshof J E, and Knoppers B M (2009) Research ethics. Children and population biobanks. *Science* 325: 818-819.

Gwiazda J, Ong E, Held R, and Thorn F (2000) Myopia and ambient night-time lighting. *Nature* 404: 144.

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S et al. (1994) The 1993-94 Genethon human genetic linkage map. *Nat Genet* 7: 246-339.

Haim M, Fledelius H C, and Skarsholm (1988) X-linked myopia in Danish family. *Acta Ophthalmol (Copenh)* 66: 450-456.

Haines J L, and Pericak-Vance M A (1998) Overview of Mapping Common and Genetically Complex Human Disease Traits. In: Haines J L, and Pericak-Vance M A [eds.] *Approaches to gene mapping in complex human diseases.* New York: Wiley-Liss, pp. 1-16.

Hall N F, Gale C R, Ye S, and Martyn C N (2009) Myopia and polymorphisms in genes for matrix metalloproteinases. *Invest Ophthalmol Vis Sci* 50: 2632-2636.

Hammond C J, Andrew T, Mak Y T, and Spector T D (2004) A susceptibility locus for myopia in the normal population is linked to the PAX6 gene region on chromosome 11: a genomewide scan of dizygotic twins. *Am J Hum Genet* 75: 294-304.

Hammond C J, Snieder H, Gilbert C E, and Spector T D (2001) Genes and environment in refractive error: the twin eye study. *Invest Ophthalmol Vis Sci* 42: 1232-1236.

Han W, Leung K H, Fung W Y, Mak J Y, Li Y M, Yap M K, and Yip S P (2009) Association of PAX6 polymorphisms with high myopia in Han Chinese nuclear families. *Invest Ophthalmol Vis Sci* 50: 47-56.

Han W, Yap M K, Wang J, and Yip S P (2006) Family-based association analysis of hepatocyte growth factor (HGF) gene polymorphisms in high myopia. *Invest Ophthalmol Vis Sci* 47: 2291-2299.

Hartl D L (1983) *Human genetics.* New York: Harper & Row.

Hartl D L, and Clark A G (1997) *Principles of population genetics*. 3rd ed. Sunderland, Mass.: Sinauer Associates.

Hasumi Y, Inoko H, Mano S, Ota M, Okada E, Kulski J K, Nishizaki R et al. (2006) Analysis of single nucleotide polymorphisms at 13 loci within the transforming growth factor-induced factor gene shows no association with high myopia in Japanese subjects. *Immunogenetics* 58: 947-953.

Hayashi T, Inoko H, Nishizaki R, Ohno S, and Mizuki N (2007) Exclusion of transforming growth factor-beta1 as a candidate gene for myopia in the Japanese. *Jpn J Ophthalmol* 51: 96-99.

Heath S, Robledo R, Beggs W, Feola G, Parodo C, Rinaldi A, Contu L et al. (2001) A novel approach to search for identity by descent in small samples of patients and controls from the same mendelian breeding unit: a pilot study on myopia. *Hum Hered* 52: 183-190.

Hewitt A W, Kearns L S, Jamieson R V, Williamson K A, van Heyningen V, and Mackey D A (2007) PAX6 mutations may be associated with high myopia. *Ophthalmic Genet* 28: 179-182.

Hirai H, Saishin M, and Yamamoto K (1998) Longitudinal and Cross-Sectional study of Refractive Changes in Pupils from 3 to 17 Years of Age. In: Tokoro T [ed.] *Myopia updates : proceedings of the 6th International Conference on Myopia*. Tokyo ; New York: Springer, pp. 81-85.

Hirschhorn J, and Daly M (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6: 95-108.

Honda S, Fujii S, Sekiya Y, and Yamamoto M (1996) Retinal control on the axial length mediated by transforming growth factor-beta in chick eye. *Invest Ophthalmol Vis Sci* 37: 2519-2526.

Howland H C (2010) Light regimes and eye growth in animal models: the older literature. In: *13th International Myopia Conference*. Tubingen, 26th-29th July.

Inamori Y, Ota M, Inoko H, Okada E, Nishizaki R, Shiota T, Mok J et al. (2007) The COL1A1 gene and high myopia susceptibility in Japanese. *Hum Genet* 122: 151-157.

International HapMap Consortium et al. (2005) A haplotype map of the human genome. *Nature* 437: 1299-1320.

Ip J M, Huynh S C, Robaei D, Rose K A, Morgan I G, Smith W, Kifley A et al. (2007) Ethnic differences in the impact of parental myopia: findings from a population-based study of 12-year-old Australian children. *Invest Ophthalmol Vis Sci* 48: 2520-2528.

Ip J M, Rose K A, Morgan I G, Burlutsky G, and Mitchell P (2008a) Myopia and the urban environment: findings in a sample of 12-year-old Australian school children. *Invest Ophthalmol Vis Sci* 49: 3858-3863.

Ip J M, Saw S M, Rose K A, Morgan I G, Kifley A, Wang J J, and Mitchell P (2008b) Role of near work in myopia: findings in a sample of Australian school children. *Invest Ophthalmol Vis Sci* 49: 2903-2910.

Jobling A I, Nguyen M, Gentle A, and McBrien N A (2004a) Isoform-specific changes in scleral transforming growth factor-beta expression and the regulation of collagen synthesis during myopia progression. *J Biol Chem* 279: 18121-18126.

Jobling M, Hurles M, and Tyler-Smith C (2004b) *Human evolutionary genetics : origins, peoples & disease*. New York: Garland Science.

Jones L A, Sinnott L T, Mutti D O, Mitchell G L, Moeschberger M L, and Zadnik K (2007) Parental history of myopia, sports and outdoor activities, and future myopia. *Invest Ophthalmol Vis Sci* 48: 3524-3532.

Jones R W, Ring S, Tyfield L, Hamvas R, Simmons H, Pembrey M, and Golding J (2000) A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). *Eur J Hum Genet* 8: 653-660.

Jorge J, Queiros A, González-Méijome J, Fernandes P, Almeida J, and Parafita M (2005) The influence of cycloplegia in objective refraction. *Ophthalmic Physiol Opt* 25: 340-345.

Jun A S, Liu S H, Koo E H, Do D V, Stark W J, and Gottsch J D (2001) Microarray analysis of gene expression in human donor corneas. *Arch Ophthalmol* 119: 1629-1634.

Karmaus W, Arshad H, and Mattes J (2001) Does the sibling effect have its origin in utero? Investigating birth order, cord blood immunoglobulin E concentration, and allergic sensitization at age 4 years. *Am J Epidemiol* 154: 909-915.

King R C, Stansfield W D, and Mulligan P K (2006) *A dictionary of genetics*. 7th ed. Oxford: Oxford University Press.

Kirkup L, and Frenkel R B (2006) *An introduction to uncertainty in measurement using the GUM (guide to the expression of uncertainty in measurement)*. Cambridge ; New York: Cambridge University Press.

Klein A P, Duggal P, Lee K E, Klein R, Bailey-Wilson J E, and Klein B E (2007) Confirmation of linkage to ocular refraction on chromosome 22q and identification of a novel linkage region on 1q. *Arch Ophthalmol* 125: 80-85.

Klein A P, Suktitipat B, Duggal P, Lee K E, Klein R, Bailey-Wilson J E, and Klein B E (2009) Heritability analysis of spherical equivalent, axial length, corneal curvature,

and anterior chamber depth in the Beaver Dam Eye Study. *Arch Ophthalmol* 127: 649-655.

Klein R J, Zeiss C, Chew E Y, Tsai J Y, Sackler R S, Haynes C, Henning A K et al. (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* 308: 385-389.

Klug W S (2009) *Concepts of genetics*. 9th ed. San Francisco ; London: Pearson Benjamin Cummings.

Konstantopoulos A, Yadegarfar G, and Elgohary M (2007) Near work, education, family history, and myopia in Greek conscripts. *Eye* 22: 542-546.

Krantz E, Cruickshanks K, Klein B, Klein R, Huang G, and Nieto F (2010) Measuring refraction in adults in epidemiological studies. *Arch Ophthalmol* 128: 88-92.

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 22: 139-144.

Lam C Y, Tam P O, Fan D S, Fan B J, Wang D Y, Lee C W, Pang C P et al. (2008a) A genome-wide scan maps a novel high myopia locus to 5p15. *Invest Ophthalmol Vis Sci* 49: 3768-3778.

Lam D, Fan D, Lam R, Rao S, Chong K, Lau J, Lai R et al. (2008b) The effect of parental history of myopia on children's eye size and growth: results of a longitudinal study. *Invest Ophthalmol Vis Sci* 49: 873-876.

Lam D S, Lee W S, Leung Y F, Tam P O, Fan D S, Fan B J, and Pang C P (2003a) TGFbeta-induced factor: a candidate gene for high myopia. *Invest Ophthalmol Vis Sci* 44: 1012-1015.

Lam D S, Tam P O, Fan D S, Baum L, Leung Y F, and Pang C P (2003b) Familial high myopia linkage to chromosome 18p. *Ophthalmologica* 217: 115-118.

Lam D S C, Fan D S P, Saw S-M, and Chew S-J (1998) Risk Factors for Refractive Errors in Preschool Children. In: Tokoro T [ed.] *Myopia updates : proceedings of the 6th International Conference on Myopia*. Tokyo ; New York: Springer, pp. 68-69.

Larsson E K, Rydberg A C, and Holmstrom G E (2003) A population-based study of the refractive outcome in 10-year-old preterm and full-term children. *Arch Ophthalmol* 121: 1430-1436.

Leal S M (1997) Tests for dectecting linkage and linkage heterogeneity. In: Pawlowitzki I-H, Edwards J H, and Thompson E A [eds.] *Genetic mapping of disease genes*. San Diego ; London: Academic Press, pp. 97-114.

Leung Y F, Tam P O, Baum L, Lam D S, and Pang C C (2000) TIGR/MYOC proximal promoter GT-repeat polymorphism is not associated with myopia. *Hum Mutat* 16: 533.

Li T, Xiao X, Li S, Xing Y, Guo X, and Zhang Q (2008) Evaluation of EGR1 as a candidate gene for high myopia. *Mol Vis* 14: 1309-1312.

Li Y, Willer C, Sanna S, and Abecasis G (2009a) Genotype imputation. *Annu Rev Genomics Hum Genet* 10: 387-406.

Li Y J, Guggenheim J A, Bulusu A, Metlapally R, Abbott D, Malecaze F, Calvas P et al. (2009b) An international collaborative family-based whole-genome linkage scan for high-grade myopia. *Invest Ophthalmol Vis Sci* 50: 3116-3127.

Liang C L, Hung K S, Tsai Y Y, Chang W, Wang H S, and Juo S H (2007) Systematic assessment of the tagging polymorphisms of the COL1A1 gene for high myopia. *J Hum Genet* 52: 374-377.

Liang C L, Wang H S, Hung K S, Hsi E, Sun A, Kuo Y H, and Juo S H (2006) Evaluation of MMP3 and TIMP1 as candidate genes for high myopia in young Taiwanese men. *Am J Ophthalmol* 142: 518-520.

Lin H J, Kung Y J, Lin Y J, Sheu J J, Chen B H, Lan Y C, Lai C H et al. (2010) Association of the lumican gene functional 3'-UTR polymorphism with high myopia. *Invest Ophthalmol Vis Sci* 51: 96-102.

Lin H J, Wan L, Tsai Y, Chen W C, Tsai S W, and Tsai F J (2009a) Muscarinic acetylcholine receptor 1 gene polymorphisms associated with high myopia. *Mol Vis* 15: 1774-1780.

Lin H J, Wan L, Tsai Y, Liu S C, Chen W C, Tsai S W, and Tsai F J (2009b) Sclera-related gene polymorphisms in high myopia. *Mol Vis* 15: 1655-1663.

Lin H J, Wan L, Tsai Y, Tsai Y Y, Fan S S, Tsai C H, and Tsai F J (2006) The TGFbeta1 gene codon 10 polymorphism contributes to the genetic predisposition to high myopia. *Mol Vis* 12: 698-703.

Lin L L, Shih Y F, Tsai C B, Chen C J, Lee L A, Hung P T, and Hou P K (1999) Epidemiologic study of ocular refraction among schoolchildren in Taiwan in 1995. *Optom Vis Sci* 76: 275-281.

Liu H P, Lin Y J, Lin W Y, Wan L, Sheu J J, Lin H J, Tsai Y et al. (2009) A novel genetic variant of BMP2K contributes to high myopia. *J Clin Lab Anal* 23: 362-367.

Lopes M C, Andrew T, Carbonaro F, Spector T, and Hammond C J (2008) Estimating heritability and shared environmental effects for refractive error in twin and family studies. *Invest Ophthalmol Vis Sci* 50: 126-131.

Low W, Dirani M, Gazzard G, Chan Y, Zhou H, Selvaraj P, Au Eong K et al. (2010) Family history, near work, outdoor activity, and myopia in Singapore Chinese preschool children. *Br J Ophthalmol* 94: 1012-1016.

Lu Y, Dimasi D, Hysi P, Hewitt A, Burdon K, Toh T, Ruddle J et al. (2010) Common genetic variants near the Brittle Cornea Syndrome locus ZNF469 influence the blinding disease risk factor central corneal thickness. *PLoS Genet* 6: e1000947.

Lyhne N, Sjolie A K, Kyvik K O, and Green A (2001) The importance of genes and environment for ocular refraction and its determiners: a population based study among 20-45 year old twins. *Br J Ophthalmol* 85: 1470-1476.

Majava M, Bishop P N, Hagg P, Scott P G, Rice A, Inglehearn C, Hammond C J et al. (2007) Novel mutations in the small leucine-rich repeat protein/proteoglycan (SLRP) genes in high myopia. *Hum Mutat* 28: 336-344.

Mandel Y, Grotto I, El-Yaniv R, Belkin M, Israeli E, Polat U, and Bartov E (2008) Season of birth, natural light, and myopia. *Ophthalmology* 115: 686-692.

Manolio T A, and Collins F S (2009) The HapMap and genome-wide association studies in diagnosis and therapy. *Annu Rev Med* 60: 443-456.

Manolio T A, Collins F S, Cox N J, Goldstein D B, Hindorff L A, Hunter D J, McCarthy M I et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461: 747-753.

Marsh-Tootle W L, and Frazier M, G. (2006) Infants, Toddlers, and Children. In: Benjamin W J, and Borish I M [eds.] *Borish's clinical refraction.* (2nd edn.) St. Louis, Miss.: Butterworth Heinemann/Elsevier, pp. 1395-1460.

Marshall G E, Konstas A G, and Lee W R (1993) Collagens in the aged human macular sclera. *Curr Eye Res* 12: 143-153.

Maul E, Barroso S, Munoz S, Sperduto R, and Ellwein L (2000) Refractive Error Study in Children: results from La Florida, Chile. *Am J Ophthalmol* 129: 445-454.

Mavracanas T, Mandalos A, Peios D, Golias V, Megalou K, Gregoriadou A, Delidou K et al. (2000) Prevalence of myopia in a sample of Greek students. *Acta Ophthalmol Scand* 78: 656-659.

McBrien N, and Millodot M (1985) Clinical evaluation of the Canon Autoref R-1. *Am J Optom Physiol Opt* 62: 786-792.

McBrien N A, and Adams D W (1997) A longitudinal investigation of adult-onset and adult-progression of myopia in an occupational group. Refractive and biometric findings. *Invest Ophthalmol Vis Sci* 38: 321-333.

McBrien N A, Cornell L M, and Gentle A (2001) Structural and ultrastructural changes to the sclera in a mammalian model of high myopia. *Invest Ophthalmol Vis Sci* 42: 2179-2187.

McBrien N A, Jobling A I, and Gentle A (2009) Biomechanics of the sclera in myopia: extracellular and cellular factors. *Optom Vis Sci* 86: E23-30.

McCarthy M, Abecasis G, Cardon L, Goldstein D, Little J, Ioannidis J, and Hirschhorn J (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 9: 356-369.

McFadden S A, Howlett M H, Mertz J R, and Wallman J (2006) Acute effects of dietary retinoic acid on ocular components in the growing chick. *Exp Eye Res* 83: 949-961.

McGlinn A M, Baldwin D A, Tobias J W, Budak M T, Khurana T S, and Stone R A (2007) Form-deprivation myopia in chick induces limited changes in retinal gene expression. *Invest Ophthalmol Vis Sci* 48: 3430-3436.

McMahon G, Zayats T, Chen Y P, Prashar A, Williams C, and Guggenheim J A (2009) Season of birth, daylight hours at birth, and high myopia. *Ophthalmology* 116: 468-473.

Metlapally R, Ki C S, Li Y J, Tran-Viet K N, Abbott D, Malecaze F, Calvas P et al. (2010) Genetic association of insulin-like growth factor-1 polymorphisms with high-grade myopia in an international family cohort. *Invest Ophthalmol Vis Sci* 51: 4476-4479.

Metlapally R, Li Y J, Tran-Viet K N, Abbott D, Czaja G R, Malecaze F, Calvas P et al. (2009a) COL1A1 and COL2A1 genes and myopia susceptibility: evidence of association and suggestive linkage to the COL2A1 locus. *Invest Ophthalmol Vis Sci* 50: 4080-4086.

Metlapally R, Michaelides M, Bulusu A, Li Y J, Schwartz M, Rosenberg T, Hunt D M et al. (2009b) Evaluation of the X-linked high-grade myopia locus (MYP1) with cone dysfunction and color vision deficiencies. *Invest Ophthalmol Vis Sci* 50: 1552-1558.

Michaelides M, Johnson S, Bradshaw K, Holder G E, Simunovic M P, Mollon J D, Moore A T et al. (2005) X-linked cone dysfunction syndrome with myopia and protanopia. *Ophthalmology* 112: 1448-1454.

Midelfart A (2002) Association between light exposure during the night and myopia in children. *Arch Ophthalmol* 120: 406-407.

Millodot M, and Laby D M (2002) *Dictionary of ophthalmology*. Oxford ; Boston: Butterworth-Heinemann.

Mitchell P, Hourihan F, Sandbach J, and Wang J J (1999) The relationship between glaucoma and myopia: the Blue Mountains Eye Study. *Ophthalmology* 106: 2010-2015.

Morton N (1955) The inheritance of human birth weight. *Ann Hum Genet* 20: 125-134.

Murthy G, Gupta S, Ellwein L, Muñoz S, Pokharel G, Sanga L, and Bachani D (2002) Refractive error in children in an urban population in New Delhi. *Invest Ophthalmol Vis Sci* 43: 623-631.

Mutti D O, Cooper M E, O'Brien S, Jones L A, Marazita M L, Murray J C, and Zadnik K (2007a) Candidate gene and locus analysis of myopia. *Mol Vis* 13: 1012-1019.

Mutti D O, Hayes J R, Mitchell G L, Jones L A, Moeschberger M L, Cotter S A, Kleinstein R N et al. (2007b) Refractive error, axial length, and relative peripheral refractive error before and after the onset of myopia. *Invest Ophthalmol Vis Sci* 48: 2510-2519.

Mutti D O, Mitchell G L, Moeschberger M L, Jones L A, and Zadnik K (2002) Parental myopia, near work, school achievement, and children's refractive error. *Invest Ophthalmol Vis Sci* 43: 3633-3640.

Nakanishi H, Hayashi H, Yamada R, Yamashiro K, Nakata I, Shimada N, Ohno-Matsui K et al. (2010) Single nucleotide polymorphisms in promoter region of matrix metalloproteinase 1, 2, and 3 in Japanese with High Myopia. *Invest Ophthalmol Vis Sci* 51: 4432-4436.

Nakanishi H, Yamada R, Gotoh N, Hayashi H, Otani A, Tsujikawa A, Yamashiro K et al. (2009a) Absence of association between COL1A1 polymorphisms and high myopia in the Japanese population. *Invest Ophthalmol Vis Sci* 50: 544-550.

Nakanishi H, Yamada R, Gotoh N, Hayashi H, Yamashiro K, Shimada N, Ohno-Matsui K et al. (2009b) A genome-wide association analysis identified a novel susceptible locus for pathological myopia at 11q24.1. *PLoS Genet* 5: e1000660.

Nallasamy S, Paluru P C, Devoto M, Wasserman N F, Zhou J, and Young T L (2007) Genetic linkage study of high-grade myopia in a Hutterite population from South Dakota. *Mol Vis* 13: 229-236.

Nayak B K, Ghose S, and Singh J P (1987) A comparison of cycloplegic and manifest refractions on the NR-1000F (an objective Auto Refractometer). *Br J Ophthalmol* 71: 73-75.

Negrel A, Maul E, Pokharel G, Zhao J, and Ellwein L (2000) Refractive Error Study in Children: sampling and measurement methods for a multi-country survey. *Am J Ophthalmol* 129: 421-426.

Ng T K, Lam C Y, Lam D S, Chiang S W, Tam P O, Wang D Y, Fan B J et al. (2009) AC and AG dinucleotide repeats in the PAX6 P1 promoter are associated with high myopia. *Mol Vis* 15: 2239-2248.

Nishizaki R, Ota M, Inoko H, Meguro A, Shiota T, Okada E, Mok J et al. (2008) New susceptibility locus for high myopia is linked to the uromodulin-like 1 (UMODL1) gene region on chromosome 21q22.3. *Eye* 23: 222-229.

No authors listed. (1889) The Heredity of Myopia. *Science* 14: 84.

Norton T T, and Rada J A (1995) Reduced extracellular matrix in mammalian sclera with induced myopia. *Vision Res* 35: 1271-1281.

Norušis M J, and SPSS Inc. (2003) *SPSS 12.0 statistical procedures companion.* Upper Saddle River, N.J.: Prentice Hall.

Nurnberg G, Jacobi F K, Broghammer M, Becker C, Blin N, Nurnberg P, Stephani U et al. (2008) Refinement of the MYP3 locus on human chromosome 12 in a German family with Mendelian autosomal dominant high-grade myopia by SNP array mapping. *Int J Mol Med* 21: 429-438.

O'Connor E, Allen L E, Bradshaw K, Boylan J, Moore A T, and Trump D (2006) Congenital stationary night blindness associated with mutations in GRM6 encoding glutamate receptor MGluR6. *Br J Ophthalmol* 90: 653-654.

Olson J M, Witte J S, and Elston R C (1999) Genetic mapping of complex traits. *Stat Med* 18: 2961-2981.

Paget S, Julia S, Vitezica Z G, Soler V, Malecaze F, and Calvas P (2008a) Linkage analysis of high myopia susceptibility locus in 26 families. *Mol Vis* 14: 2566-2574.

Paget S, Vitezica Z, Malecaze F, and Calvas P (2008b) Heritability of refractive value and ocular biometrics. *Exp Eye Res* 86: 290-295.

Paluru P, Ronan S M, Heon E, Devoto M, Wildenberg S C, Scavello G, Holleschau A et al. (2003) New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest Ophthalmol Vis Sci* 44: 1830-1836.

Paluru P C, Nallasamy S, Devoto M, Rappaport E F, and Young T L (2005) Identification of a novel locus on 2q for autosomal dominant high-grade myopia. *Invest Ophthalmol Vis Sci* 46: 2300-2307.

Paluru P C, Scavello G S, Ganter W R, and Young T L (2004) Exclusion of lumican and fibromodulin as candidate genes in MYP3 linked high grade myopia. *Mol Vis* 10: 917-922.

Pardue M T, Faulkner A E, Fernandes A, Yin H, Schaeffel F, Williams R W, Pozdeyev N et al. (2008) High susceptibility to experimental myopia in a mouse model with a retinal on pathway defect. *Invest Ophthalmol Vis Sci* 49: 706-712.

Park J H, Wacholder S, Gail M H, Peters U, Jacobs K B, Chanock S J, and Chatterjee N (2010) Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet* 42: 570-575.

Payami H, Zhu M, Montimurro J, Keefe R, McCulloch C C, and Moses L (2005) One step closer to fixing association studies: evidence for age- and gender-specific allele

frequency variations and deviations from Hardy-Weinberg expectations in controls. *Hum Genet* 118: 322-330.

Peckham C S, Gardiner P A, and Goldstein H (1977) Acquired myopia in 11-year-old children. *Br Med J* 1: 542-545.

Pertile K K, Schache M, Islam F M, Chen C Y, Dirani M, Mitchell P, and Baird P N (2008) Assessment of TGIF as a candidate gene for myopia. *Invest Ophthalmol Vis Sci* 49: 49-54.

Pokharel G, Negrel A, Munoz S, and Ellwein L (2000) Refractive Error Study in Children: results from Mechi Zone, Nepal. *Am J Ophthalmol* 129: 436-444.

Price A L, Patterson N J, Plenge R M, Weinblatt M E, Shadick N A, and Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904-909.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A, Bender D, Maller J et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559-575.

Quinn G E, Shin C H, Maguire M G, and Stone R A (1999) Myopia and ambient lighting at night. *Nature* 399: 113-114.

Rada J A, and Brenza H L (1995) Increased latent gelatinase activity in the sclera of visually deprived chicks. *Invest Ophthalmol Vis Sci* 36: 1555-1565.

Rajan U, Saw S-M, Lau C, O'Brien L, Chan T-K, Lam D S C, and Chew S-J (1998) Prevalence of Myopia in Schoolchildren and Risk Factors for Its Progression. In: Tokoro T [ed.] *Myopia updates : proceedings of the 6th International Conference on Myopia.* Tokyo ; New York: Springer, pp. 69-80.

Reich D E, and Lander E S (2001) On the allelic spectrum of human disease. *Trends Genet* 17: 502-510.

Remington L A (2004) *Clinical anatomy of the visual system.* 2 ed. Missouri: Butterworth-Heinemann.

Rice T, Schork N, and Rao D (2008) Methods for handling multiple testing. *Adv Genet* 60: 293-308.

Riddoch C J, Mattocks C, Deere K, Saunders J, Kirkby J, Tilling K, Leary S D et al. (2007) Objective measurement of levels and patterns of physical activity. *Arch Dis Child* 92: 963-969.

Risch N, and Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273: 1516-1517.

Rohrer B, and Stell W K (1994) Basic fibroblast growth factor (bFGF) and transforming growth factor beta (TGF-beta) act as stop and go signals to modulate postnatal ocular growth in the chick. *Exp Eye Res* 58: 553-561.

Rose K A, Morgan I G, Ip J, Kifley A, Huynh S, Smith W, and Mitchell P (2008a) Outdoor activity reduces the prevalence of myopia in children. *Ophthalmology* 115: 1279-1285.

Rose K A, Morgan I G, Smith W, Burlutsky G, Mitchell P, and Saw S M (2008b) Myopia, lifestyle, and schooling in students of chinese ethnicity in singapore and sydney. *Arch Ophthalmol* 126: 527-530.

Rosenfield M (1998) Accommodation and myopia. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork.* Oxford ; Boston: Butterworth-Heinemann, pp. 91-116.

Rosenfield M (2006) Refractive Status of the Eye. In: Benjamin W J, and Borish I M [eds.] *Borish's clinical refraction.* (2nd edn.) St. Louis, Miss.: Butterworth Heinemann/Elsevier, pp. 3-34.

Rotsos T, Grigoriou D, Kokkolaki A, and Manios N (2009) A comparison of manifest refractions, cycloplegic refractions and retinoscopy on the RMA-3000 autorefractometer in children aged 3 to 15 years. *Clin Ophthalmol* 3: 429-431.

Rudnicka A R, Owen C G, Richards M, Wadsworth M E, and Strachan D P (2008) Effect of breastfeeding and sociodemographic factors on visual outcome in childhood and adolescence. *Am J Clin Nutr* 87: 1392-1399.

Sasaki S, Ota M, Meguro A, Nishizaki R, Okada E, Mok J, Kimura T et al. (2007) A single nucleotide polymorphism analysis of the LAMA1 gene in Japanese patients with high myopia. *Clin Ophthalmol* 1: 289-295.

Saw S, Chan Y, Wong W, Shankar A, Sandar M, Aung T, Tan D et al. (2008) Prevalence and risk factors for refractive errors in the Singapore Malay Eye Survey. *Ophthalmology* 115: 1713-1719.

Saw S, Cheng A, Fong A, Gazzard G, Tan D, and Morgan I (2007) School grades and myopia. *Ophthalmic Physiol Opt* 27: 126-129.

Saw S, Chua W, Hong C, Wu H, Chan W, Chia K, Stone R et al. (2002a) Nearwork in early-onset myopia. *Invest Ophthalmol Vis Sci* 43: 332-339.

Saw S, Chua W, Hong C, Wu H, Chia K, Stone R, and Tan D (2002b) Height and its relationship to refraction and biometry parameters in Singapore Chinese children. *Invest Ophthalmol Vis Sci* 43: 1408-1413.

Saw S, Tong L, Chia K, Koh D, Lee Y, Katz J, and Tan D (2004a) The relation between birth size and the results of refractive error and biometry measurements in children. *Br J Ophthalmol* 88: 538-542.

Saw S, Wu H, Seet B, Wong T, Yap E, Chia K, Stone R et al. (2001) Academic achievement, close up work parameters, and myopia in Singapore military conscripts. *Br J Ophthalmol* 85: 855-860.

Saw S M, Gazzard G, Koh D, Farook M, Widjaja D, Lee J, and Tan D T (2002c) Prevalence rates of refractive errors in Sumatra, Indonesia. *Invest Ophthalmol Vis Sci* 43: 3174-3180.

Saw S M, Gazzard G, Shih-Yen E C, and Chua W H (2005) Myopia and associated pathological complications. *Ophthalmic Physiol Opt* 25: 381-391.

Saw S M, Tan S B, Fung D, Chia K S, Koh D, Tan D T, and Stone R A (2004b) IQ and the association with myopia in children. *Invest Ophthalmol Vis Sci* 45: 2943-2948.

Saxena R, Voight B F, Lyssenko V, Burtt N P, de Bakker P I, Chen H, Roix J J et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316: 1331-1336.

Scavello G S, Paluru P C, Ganter W R, and Young T L (2004) Sequence variants in the transforming growth beta-induced factor (TGIF) gene are not associated with high myopia. *Invest Ophthalmol Vis Sci* 45: 2091-2097.

Schache M, Chen C Y, Dirani M, and Baird P N (2009) The hepatocyte growth factor receptor (MET) gene is not associated with refractive error and ocular biometrics in a Caucasian population. *Mol Vis* 15: 2599-2605.

Schildkraut J M (1998) Examining Complex Genetic Interactions. In: Haines J L, and Pericak-Vance M A [eds.] *Approaches to gene mapping in complex human diseases.* New York: Wiley-Liss, pp. 379-410.

Schippert R, Burkhardt E, Feldkaemper M, and Schaeffel F (2007) Relative axial myopia in Egr-1 (ZENK) knockout mice. *Invest Ophthalmol Vis Sci* 48: 11-17.

Schwartz M, Haim M, and Skarsholm D (1990) X-linked myopia: Bornholm eye disease. Linkage to DNA markers on the distal part of Xq. *Clin Genet* 38: 281-286.

Shimizu N, Nomura H, Ando F, Niino N, Miyake Y, and Shimokata H (2003) Refractive errors and factors associated with myopia in an adult Japanese population. *Jpn J Ophthalmol* 47: 6-12.

Sibon I, Coupry I, Menegon P, Bouchet J P, Gorry P, Burgelin I, Calvas P et al. (2007) COL4A1 mutation in Axenfeld-Rieger anomaly with leukoencephalopathy and stroke. *Ann Neurol* 62: 177-184.

Siegwart J T, Jr., and Norton T T (2001) Steady state mRNA levels in tree shrew sclera with form-deprivation myopia and during recovery. *Invest Ophthalmol Vis Sci* 42: 1153-1159.

Siegwart J T, Jr., and Norton T T (2002) The time course of changes in mRNA levels in tree shrew sclera during induced myopia and recovery. *Invest Ophthalmol Vis Sci* 43: 2067-2075.

Simpson C L, Hysi P, Bhattacharya S S, Hammond C J, Webster A, Peckham C S, Sham P C et al. (2007) The Roles of PAX6 and SOX2 in Myopia: lessons from the 1958 British Birth Cohort. *Invest Ophthalmol Vis Sci* 48: 4421-4425.

Smith E L (1998) Environmentally induced refractive errors. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork.* Oxford ; Boston: Butterworth-Heinemann, pp. 57-90.

Smith E L I (1991) Experimentally Induced Refractive Anomalies In Mammals. In: Grosvenor T P, and Flom M C [eds.] *Refractive anomalies : research and clinical applications.* Boston ; London: Butterworth-Heinemann, pp. 246-267.

Sohajda Z, Hollo D, Berta A, and Modis L (2006) Microcornea associated with myopia. *Graefes Arch Clin Exp Ophthalmol* 244: 1211-1213.

Spencer C C, Su Z, Donnelly P and Marchini J (2009) Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet* 5: e1000477.

Sperduto R D, Seigel D, Roberts J, and Rowland M (1983) Prevalence of myopia in the United States. *Arch Ophthalmol* 101: 405-407.

Stambolian D, Ciner E B, Reider L C, Moy C, Dana D, Owens R, Schlifka M et al. (2005) Genome-wide scan for myopia in the Old Order Amish. *Am J Ophthalmol* 140: 469-476.

Stambolian D, Ibay G, Reider L, Dana D, Moy C, Schlifka M, Holmes T et al. (2004) Genomewide linkage scan for myopia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 22q12. *Am J Hum Genet* 75: 448-459.

Stambolian D, Ibay G, Reider L, Dana D, Moy C, Schlifka M, Holmes T N et al. (2006) Genome-wide scan of additional Jewish families confirms linkage of a myopia susceptibility locus to chromosome 22q12. *Mol Vis* 12: 1499-1505.

Stephens J C, and Bamshad M (2007) Population Choice as a Consideration for Genetic Analysis Study Design. In: Weiner M P, Gabriel S B, and Stephens J C [eds.] *Genetic variation : a laboratory manual.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, pp. 11-18.

Stone E M, Fingert J H, Alward W L, Nguyen T D, Polansky J R, Sunden S L, Nishimura D et al. (1997) Identification of a gene that causes primary open angle glaucoma. *Science* 275: 668-670.

Tang W C, Yip S P, Lo K K, Ng P W, Choi P S, Lee S Y, and Yap M K (2007) Linkage and association of myocilin (MYOC) polymorphisms with high myopia in a Chinese population. *Mol Vis* 13: 534-544.

Terwilliger J D (1998) Linkage Analysis, Model-Based. In: Armitage P and Colton T [eds.]. *Encyclopedia of biostatistics.* Chichester: Wiley, pp. 2279-2291.

The Framingham Offspring Eye Study Group (1996) Familial aggregation and prevalence of myopia in the Framingham Offspring Eye Study. *Arch Ophthalmol* 114: 326-332.

The Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72: 971-983.

Traboulsi E I, Zhu D, and Maumenee I H (1998) Aniridia. In: Traboulsi E I [ed.] *Genetic diseases of the eye.* New York: Oxford University Press, pp. 99-114.

Troilo D, Nickla D L, Mertz J R, and Summers Rada J A (2006) Change in the synthesis rates of ocular retinoic acid and scleral glycosaminoglycan during experimentally altered eye growth in marmosets. *Invest Ophthalmol Vis Sci* 47: 1768-1777.

Tsai Y Y, Chiang C C, Lin H J, Lin J M, Wan L, and Tsai F J (2008) A PAX6 gene polymorphism is associated with genetic predisposition to extreme myopia. *Eye (Lond)* 22: 576-581.

UNICEF (2004) Low Birthweight: Country, regional and global estimates. UNICEF, New York.

Van Agtmael T, Schlotzer-Schrehardt U, McKie L, Brownstein D G, Lee A W, Cross S H, Sado Y et al. (2005) Dominant mutations of Col4a1 result in basement membrane defects which lead to anterior segment dysgenesis and glomerulopathy. *Hum Mol Genet* 14: 3161-3168.

Vannas A E, Ying G S, Stone R A, Maguire M G, Jormanainen V, and Tervo T (2003) Myopia and natural lighting extremes: risk factors in Finnish army conscripts. *Acta Ophthalmol Scand* 81: 588-595.

Varghese R, Sreenivas V, Puliyel J, and Varughese S (2009) Refractive status at birth: its relation to newborn physical parameters at birth and gestational age. *PLoS One* 4: e4469.

Vatavuk Z, Skunca Herman J, Bencic G, Andrijevic Derk B, Lacmanovic Loncar V, Petric Vickovic I, Bucan K et al. (2009) Common variant in myocilin gene is associated with high myopia in isolated population of Korcula Island, Croatia. *Croat Med J* 50: 17-22.

Veerappan S, Pertile K K, Islam A F, Schache M, Chen C Y, Mitchell P, Dirani M et al. (2010) Role of the hepatocyte growth factor gene in refractive error. *Ophthalmology* 117: 239-245 e231-232.

Veerappan S, Schache M, Pertile K K, Islam F M, Chen C Y, Mitchell P, Dirani M et al. (2009) The retinoic acid receptor alpha (RARA) gene is not associated with myopia, hypermetropia, and ocular biometric measures. *Mol Vis* 15: 1390-1397.

Venter J C, Adams M D, Myers E W, Li P W, Mural R J, Sutton G G, Smith H O et al. (2001) The sequence of the human genome. *Science* 291: 1304-1351.

Visscher P M, Hill W G, and Wray N R (2008) Heritability in the genomics era--concepts and misconceptions. *Nat Rev Genet* 9: 255-266.

Vitale S, Sperduto R D, and Ferris F L, 3rd (2009) Increased prevalence of myopia in the United States between 1971-1972 and 1999-2004. *Arch Ophthalmol* 127: 1632-1639.

Wallman J (1991) Retinal Factors In Myopia And Emmetropization: Clues From Research On Chicks. In: Grosvenor T P, and Flom M C [eds.] *Refractive anomalies : research and clinical applications*. Boston ; London: Butterworth-Heinemann, pp. 268-286.

Wallman J, Gottlieb M D, Rajaram V, and Fugate-Wentzek L A (1987) Local retinal regions control local eye growth and myopia. *Science* 237: 73-77.

Wallman J, Turkel J, and Trachtman J (1978) Extreme myopia produced by modest change in early visual experience. *Science* 201: 1249-1251.

Walsh T, McClellan J, McCarthy S, Addington A, Pierce S, Cooper G, Nord A et al. (2008) Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320: 539-543.

Wang D G, Fan J B, Siao C J, Berno A, Young P, Sapolsky R, Ghandour G et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280: 1077-1082.

Wang I J, Chiang T H, Shih Y F, Hsiao C K, Lu S C, Hou Y C, and Lin L L (2006) The association of single nucleotide polymorphisms in the 5'-regulatory region of the lumican gene with susceptibility to high myopia in Taiwan. *Mol Vis* 12: 852-857.

Wang P, Li S, Xiao X, Guo X, and Zhang Q (2009a) An evaluation of OPTC and EPYC as candidate genes for high myopia. *Mol Vis* 15: 2045-2049.

Wang P, Li S, Xiao X, Jia X, Jiao X, Guo X, and Zhang Q (2009b) High myopia is not associated with the SNPs in the TGIF, lumican, TGFB1, and HGF genes. *Invest Ophthalmol Vis Sci* 50: 1546-1551.

Wang W Y, Barratt B J, Clayton D G, and Todd J A (2005) Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 6: 109-118.

Weber J L, and May P E (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44: 388-396.

Wildsoet C F (1998) Structural correlates of myopia. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork.* Oxford ; Boston: Butterworth-Heinemann, pp. 31-56.

Williams C, Miller L, Gazzard G, and Saw S M (2008a) A comparison of measures of reading and intelligence as risk factors for the development of myopia in a UK cohort of children. *Br J Ophthalmol* 92: 1117-1121.

Williams C, Miller L, Northstone K, and Sparrow J M (2008b) The use of non-cycloplegic autorefraction data in general studies of children's development. *Br J Ophthalmol* 92: 723-724.

Williams C, Northstone K, Howard M, Harvey I, Harrad R A, and Sparrow J M (2008c) Prevalence and risk factors for common vision problems in children: data from the ALSPAC study. *Br J Ophthalmol* 92: 959-964.

Williams S, Sanderson G, Share D, and Silva P (1988) Refractive error, IQ and reading ability: a longitudinal study from age seven to 11. *Dev Med Child Neurol* 30: 735-742.

Wojciechowski R, Bailey-Wilson J E, and Stambolian D (2009a) Fine-mapping of candidate region in Amish and Ashkenazi families confirms linkage of refractive error to a QTL on 1p34-p36. *Mol Vis* 15: 1398-1406.

Wojciechowski R, Bailey-Wilson J E, and Stambolian D (2010) Association of matrix metalloproteinase gene polymorphisms with refractive error in Amish and Ashkenazi families. *Invest Ophthalmol Vis Sci* 51: 4989-4995.

Wojciechowski R, Congdon N, Bowie H, Munoz B, Gilbert D, and West S K (2005) Heritability of refractive error and familial aggregation of myopia in an elderly American population. *Invest Ophthalmol Vis Sci* 46: 1588-1592.

Wojciechowski R, Moy C, Ciner E, Ibay G, Reider L, Bailey-Wilson J E, and Stambolian D (2006) Genomewide scan in Ashkenazi Jewish families demonstrates evidence of linkage of ocular refraction to a QTL on chromosome 1p36. *Hum Genet* 119: 389-399.

Wojciechowski R, Stambolian D, Ciner E, Ibay G, Holmes T N, and Bailey-Wilson J E (2009b) Genomewide linkage scans for ocular refraction and meta-analysis of four populations in the Myopia Family Study. *Invest Ophthalmol Vis Sci* 50: 2024-2032.

Wolfe S L (1993) *Molecular and cellular biology.* Wadsworth Publishing Company.

Wong T, Foster P, Hee J, Ng T, Tielsch J, Chew S, Johnson G et al. (2000) Prevalence and risk factors for refractive errors in adult Chinese in Singapore. *Invest Ophthalmol Vis Sci* 41: 2486-2494.

Wong T Y, Foster P J, Johnson G J, Klein B E, and Seah S K (2001) The relationship between ocular dimensions and refraction with adult stature: the Tanjong Pagar Survey. *Invest Ophthalmol Vis Sci* 42: 1237-1242.

Woodward M (2005) *Epidemiology : study design and data analysis*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.

Wright A, Teague P W, Bruford E A, and Carothers A D (1997) Problems in dealing with linkage heterogeneity in autosomal recessive forms of retinitis pigmentosa. In: Pawlowitzki I-H, Edwards J H, and Thompson E A [eds.] *Genetic mapping of disease genes*. San Diego ; London: Academic Press, pp. 255-274.

Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661-678.

Wu H, Gupta A, Newland H, Selva D, Aung T, and Casson R (2007) Association between stature, ocular biometry and refraction in an adult population in rural Myanmar: the Meiktila eye study. *Clin Experiment Ophthalmol* 35: 834-839.

Wu H M, Seet B, Yap E P, Saw S M, Lim T H, and Chia K S (2001) Does education explain ethnic differences in myopia prevalence? A population-based study of young adult males in Singapore. *Optom Vis Sci* 78: 234-239.

Xu L, Li J, Cui T, Hu A, Fan G, Zhang R, Yang H et al. (2005) Refractive error in urban and rural adult Chinese in Beijing. *Ophthalmology* 112: 1676-1683.

Xu X, Li S, Xiao X, Wang P, Guo X, and Zhang Q (2009) Sequence variations of GRM6 in patients with high myopia. *Mol Vis* 15: 2094-2100.

Yang Z, Xiao X, Li S, and Zhang Q (2009) Clinical and linkage study on a consanguineous Chinese family with autosomal recessive high myopia. *Mol Vis* 15: 312-318.

Yanovitch T, Li Y J, Metlapally R, Abbott D, Viet K N, and Young T L (2009) Hepatocyte growth factor and myopia: genetic association analyses in a caucasian population. *Mol Vis* 15: 1028-1035.

Younan C, Mitchell P, Cumming R G, Rochtchina E, and Wang J J (2002) Myopia and incident cataract and cataract surgery: the blue mountains eye study. *Invest Ophthalmol Vis Sci* 43: 3625-3632.

Young T L (2009) Molecular Genetics of Human Myopia: An Update. *Optom Vis Sci* 86: E8-E22.

Young T L, Atwood L D, Ronan S M, Dewan A T, Alvear A B, Peterson J, Holleschau A et al. (2001) Further refinement of the MYP2 locus for autosomal dominant high myopia by linkage disequilibrium analysis. *Ophthalmic Genet* 22: 69-75.

Young T L, Deeb S S, Ronan S M, Dewan A T, Alvear A B, Scavello G S, Paluru P C et al. (2004) X-linked high myopia associated with cone dysfunction. *Arch Ophthalmol* 122: 897-908.

Young T L, Ronan S M, Alvear A B, Wildenberg S C, Oetting W S, Atwood L D, Wilkin D J et al. (1998a) A second locus for familial high myopia maps to chromosome 12q. *Am J Hum Genet* 63: 1419-1424.

Young T L, Ronan S M, Drahozal L A, Wildenberg S C, Alvear A B, Oetting W S, Atwood L D et al. (1998b) Evidence that a locus for familial high myopia maps to chromosome 18p. *Am J Hum Genet* 63: 109-119.

Yu Z Q, Li Y B, Huang C X, Chu R Y, Hu D N, Shen Z H, and Huang W (2007) [A genome-wide screening for pathological myopia suggests a novel locus on chromosome 15q12 - 13]. *Zhonghua Yan Ke Za Zhi* 43: 233-238.

Zadnik K (1997) The Glenn A. Fry Award Lecture (1995). Myopia development in childhood. *Optom Vis Sci* 74: 603-608.

Zadnik K, Jones L A, Irvin B C, Kleinstein R N, Manny R E, Shin J A, and Mutti D O (2000) Myopia and ambient night-time lighting. CLEERE Study Group. Collaborative Longitudinal Evaluation of Ethnicity and Refractive Error. *Nature* 404: 143-144.

Zadnik K, and Mutii D O (1998) Prevalence of myopia. In: Rosenfield M, and Gilmartin B [eds.] *Myopia and nearwork.* Oxford ; Boston: Butterworth-Heinemann, pp. 13-30.

Zadnik K, and Mutti D O (1987) Refractive error changes in law students. *Am J Optom Physiol Opt* 64: 558-561.

Zadnik K, and Mutti D O (2006) Incidence and Distribution of Refractive Anomalies. In: Benjamin W J, and Borish I M [eds.] *Borish's clinical refraction.* (2nd edn.) St. Louis, Miss.: Butterworth Heinemann/Elsevier, pp. 35-55.

Zadnik K, Satariano W, Mutti D, Sholtz R, and Adams A (1994) The effect of parental history of myopia on children's eye size. *JAMA* 271: 1323-1327.

Zayats T, Yanovitch T, Creer R C, McMahon G, Li Y J, Young T L, and Guggenheim J A (2009) Myocilin polymorphisms and high myopia in subjects of European origin. *Mol Vis* 15: 213-222.

Zha Y, Leung K H, Lo K K, Fung W Y, Ng P W, Shi M G, Yap M K et al. (2009) TGFB1 as a susceptibility gene for high myopia: a replication study with new findings. *Arch Ophthalmol* 127: 541-548.

Zhang F, Zhu T, Zhou Z, Wu Y, and Li Y (2009) Association of lumican gene with susceptibility to pathological myopia in the northern han ethnic chinese. *J Ophthalmol* 2009: 514306.

Zhang M, Li L, Chen L, Lee J, Wu J, Yang A, Chen C et al. (2010) Population density and refractive error among Chinese children. *Invest Ophthalmol Vis Sci* 51: 4969-4976.

Zhang Q, Guo X, Xiao X, Jia X, Li S, and Hejtmancik J F (2005) A new locus for autosomal dominant high myopia maps to 4q22-q27 between D4S1578 and D4S1612. *Mol Vis* 11: 554-560.

Zhang Q, Guo X, Xiao X, Jia X, Li S, and Hejtmancik J F (2006) Novel locus for X-linked recessive high myopia maps to Xq23-q25 but outside MYP1. *J Med Genet* 43: e20.

Zhang Q, Li S, Xiao X, Jia X, and Guo X (2007a) Confirmation of a genetic locus for X-linked recessive high myopia outside MYP1. *J Hum Genet* 52: 469-472.

Zhang Q, Xiao X, Li S, Jia X, Yang Z, Huang S, Caruso R C et al. (2007b) Mutations in NYX of individuals with high myopia, but without night blindness. *Mol Vis* 13: 330-336.

Zhao J, Mao J, Luo R, Li F, Munoz S, and Ellwein L (2002) The progression of refractive error in school-age children: Shunyi district, China. *Am J Ophthalmol* 134: 735-743.

Zhao J, Mao J, Luo R, Li F, Pokharel G P, and Ellwein L B (2004) Accuracy of noncycloplegic autorefraction in school-age children in China. *Optom Vis Sci* 81: 49-55.

Zhao J, Pan X, Sui R, Munoz S, Sperduto R, and Ellwein L (2000) Refractive Error Study in Children: results from Shunyi District, China. *Am J Ophthalmol* 129: 427-435.

Zhou G, and Williams R W (1999) Eye1 and Eye2: gene loci that modulate eye size, lens weight, and retinal area in the mouse. *Invest Ophthalmol Vis Sci* 40: 817-825.

Zhou J, and Young T L (2005) Evaluation of Lipin 2 as a candidate gene for autosomal dominant 1 high-grade myopia. *Gene* 352: 10-19.