This electronic thesis or dissertation has been downloaded from the King's Research Portal at https://kclpure.kcl.ac.uk/portal/

Exploration of genetic and iatrogenic risk factors associated with refractive error and myopia

Patasova, Karina

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 31. Oct. 2023

Exploration of genetic and iatrogenic risk factors associated with refractive error and myopia

Karina Patasova

Thesis submitted to the University of London in fulfilment of the requirements for the degree of Doctor of Philosophy

November 2021

Department of Ophthalmology

and

Department of Twin Research & Genetic Epidemiology

King's College London

Declaration

I, Karina Patasova declare that the publications presented in this thesis are my own. When information was obtained from other sources, I provide the references of the original material in the thesis.

Karina Patasova

November 17, 2021

Abstract

Refractive errors, and especially myopia, are one of the leading causes of preventable visual impairment and blindness worldwide. The prevalence of myopia has increased significantly over the last three decades and is projected to affect 50% of the global population by 2050 (1). Even though previous work identified several risk factors, the pathophysiology of refractive errors is not yet fully understood. My work, presented as a thesis incorporating publications, examines associations between genetic and environmental factors and the population's refractive error and myopia.

The molecular mechanisms underlying refractive development were examined in the functional analyses incorporating the results of a meta-analysis of refractive error genome-wide association studies (GWAS) in Europeans. The results of post-GWAS analyses suggested that refractive error was a genetically heterogeneous trait, driven by genes involved in the development of all the different structures of the eye. Genetic factors associated with myopia are associated with neurodevelopment, ion trafficking that regulates cell polarity, but also control circadian rhythm, and are involved in melanin pigmentation and maintenance of intraocular pressure. The loci identified in the meta-analysis currently explain approximately 20% of refractive error heritability and polygenic models that were created as part of this thesis were able to predict myopia with considerable accuracy among Europeans as well as Singaporean Asian populations. In Europeans, identified variants predicted myopia with the area under the curve (AUC) of 0.74-0.75, while in Singaporean adolescent children prediction model incorporating associated SNPs and time outdoors showed AUC values of 0.59 and 0.72 for high and moderate myopia, respectively.

These GWAS common variant findings should not disguise the importance of rare variants that influence myopia as other complex traits. This thesis describes how for the first time rare variants in several distinct loci were associated with refractive error in the general population. Through my work, I identified rare variant associations in eight loci

of which five were novel. The two most significant associations were identified for a gene encoding a transcription factor critical to the eye, retina and optic disc development and morphology *(SIX6),* and a novel candidate regulating photoreceptor specification and expression of many retinal genes *(CRX)*. Other candidates were implicated in neurogenesis and neural signalling and were reported as associated with major psychiatric disorders.

Although GWAS and rare variant gene-based association analyses advanced the understanding of the molecular mechanisms behind lifelong risk to diseases of refraction, they contributed relatively a modest amount of knowledge to the understanding of the age-related specifics of refractive error genes. The age of refractive correction is one of the indicators of myopia progression and associated with the degree of refractive error in adulthood. In this thesis, I describe the largest scale genome-wide time-to-event analyses of self-reported age of first spectacle wear, performed on a sub-sample of UK Biobank participants of European descent. Although the self-reported age of refractive correction was strongly genetically correlated with refractive error, my work identified six loci that were novel and previously not described in relationship to refractive error, as well as showed good replication for many other, previously known loci. My results indicate that the effects of gene loci over phenotypic traits were not necessarily linear. These analyses demonstrated that, despite a trend for genes with higher effects to cause correction during early childhood years, there was considerable variation in the times when genes exerted maximal effects, which was not always proportional to the magnitude of their effect over refractive error.

Refractive errors arise from a mixture of gene effects and environment, while the first is constant and the latter has been changing and driving increasing prevalence rates of myopia. Several relatively well-studied socio-economic and cultural factors have been implicated in refractive error and the changes of their prevalence throughout the past seven decades. However, the structure of morbidity and medication has also changed considerably during the same time frame, in parallel to the same cultural and societal

changes. The relationship between refractive errors and pharmacological treatments has been understudied. In this thesis, I describe the observations of the association between refractive error and medication-taking in a large population-based cohort. The most significant associations were found for 18 distinct pharmacological agents. Participants who reported taking anti-glaucoma preparations exhibited more negative, myopic refractive error compared to subjects who were not taking these medications, potentially due to the relationship between elevated intraocular pressure and myopic refraction. Interestingly, the group that was using pain control medications displayed a tendency towards a positive, more hypermetropic refractive error. Causal inference models show that multisite chronic pain, often associated with pain control medications may be at the origin of this relationship with refractive error.

The work presented in this thesis identified several novel factors associated with refractive error. The results described in this thesis, may enable the prediction of individual risk to refractive error and facilitate the development of optimised strategies for myopia prevention and treatment.

Acknowledgements

This thesis was undertaken with the financial support of the Fight for Sight grant; I am a very grateful recipient of this grant that provided many opportunities and financial support. I would like to thank my primary supervisor, Dr Pirro Hysi, for his excellent mentorship, unwavering support and guidance. Dr Pirro Hysi has encouraged me to pursue my research interests and develop critical thinking skills and I am indebted to him for the many hours of mentorship, research and career guidance. I would like to express my gratitude to my supervisor Mario Falchi who had taught me novel statistical methods and challenged my perceptions. My sincere thanks also go to Professor Chris Hammond who provided expertise on ophthalmology and helped me become a better researcher. I also would like to acknowledge lecturers within the Department of Ophthalmology, Omar Mahroo and Katie Williams who have been a vital source of knowledge and support. During my time at the Department of Twin Research and Genetic Epidemiology, I worked with other junior researchers – Mark Simcoe, Zihe Xu and Olatz Mompeo who provided technical support and were a source of ideas and inspiration. I would like to thank other researchers that helped me with data analysis and publications – Dr Anthony Khawaja, Professor Jugnoo Rahi, Dr Ameenat Solebo as well as staff at the Department of Twin Research, including Maria Bell and Diana Kozareva. I would like to express my gratitude to the participants of the UK Biobank and TwinsUk cohorts without whom research presented in this thesis wouldn't be possible. Lastly, a would like to thank my family for believing in me and encouraging me to join the PhD program.

Publications arising from this thesis

1. Lanca C*, Kassam I*, **Patasova K***, Foo LL, Li J, Ang M, Hoang Q, Teo YY, Hysi PG, and Saw SM. New polygenic risk score to predict high myopia in Singapore Chinese children. Traditional Vision Science and Technology. 2021 Jul 1;10(8):26–26.

2. Hardcastle AJ, Liskova P, Bykhovskaya Y, McComish BJ, Davidson AE, Inglehearn, CF, …, **Patasova K**, …, Iyengar SK, O'Brart D, Jorgenson E, Baird PN, Rabinowitz YS, Burdon KP, Hammond CJ, Tuft SJ, Hysi PG. A multi-ethnic genome-wide association study implicates collagen matrix integrity and cell differentiation pathways in keratoconus. Commun Biol. 2021 Mar 1;4(1):266.

3. **Patasova K**, Khawaja AP, Tamraz B, Williams KM, Mahroo OA, Freidin M, Solebo AL, Vehof J, Falchi M, Rahi JS, Hammond CJ, Hysi PG. Association Between Medication-Taking and Refractive Error in a Large General Population-Based Cohort. Invest Ophthalmol Vis Sci. 2021 Feb 1;62(2):15.

4. **Patasova K**, Hysi PG. The Slowly Emerging Uniqueness of the High Myopia Genetic Architecture. Ophthalmology. 2020 Dec;127(12):1625-1626.

5. Hysi PG, Choquet H, Khawaja AP, Wojciechowski R, Tedja MS, Yin J, Simcoe MJ, **Patasova K**, …, Jorgenson E, Hammond CJ. Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. Nat Genet. 2020 Apr;52(4):401-407.

Table of Contents

DECLARATION 2

ABSTRACT 3

ACKNOWLEDGEMENTS 6

PUBLICATIONS ARISING FROM THIS THESIS 7

TABLE OF CONTENTS 8

TABLE OF FIGURES 13

TABLES 15

ABBREVIATIONS 17

CHAPTER 1 | INTRODUCTION 22

1.1 OVERVIEW 23

1.2 EYE ANATOMY 23

1.3 DEFINITION AND NATURAL HISTORY OF MYOPIA 26

1.4 EPIDEMIOLOGY OF MYOPIA 29

1.4.1 GLOBAL PREVALENCE RATES OF MYOPIA IN CHILDREN 29

1.4.2 MYOPIA PREVALENCE IN URBAN VS RURAL AREAS 31

1.4.3 COHORT EFFECTS IN MYOPIA PREVALENCE 32

1.5 AETIOLOGICAL RISK FACTORS 35

1.5.1 GENETICS OF REFRACTIVE ERROR AND MYOPIA 35

1.5.1.1 HERITABILITY STUDIES 36

1.5.1.2 LINKAGE AND CANDIDATE GENE STUDIES 37

1.5.1.3 GENOME-WIDE ASSOCIATION STUDIES ON REFRACTIVE ERROR AND MYOPIA 38

1.5.1.4 NEXT-GENERATION SEQUENCING STUDIES ON MYOPIA 42

1.5.2 ENVIRONMENTAL FACTORS ASSOCIATED WITH REFRACTIVE ERROR AND MYOPIA 44

1.5.2.1 EDUCATIONAL ATTAINMENT 44

1.5.2.2 THE INTERPLAY BETWEEN COGNITIVE FUNCTIONS AND MYOPIA 46

1.5.2.3 NEAR-WORK 49

1.5.2.4 TIME OUTDOORS 50

1.5.2.5 BIRTH SEASON 52
1.5.2.6 URBANISATION 53
1.5.2.7 PARENTAL FACTORS ASSOCIATED WITH MYOPIA 53
1.5.2.8 OTHER FACTORS ASSOCIATED WITH MYOPIA. 54
1.6 PATHOLOGICAL CHANGES AND COMPLICATIONS ASSOCIATED WITH MYOPIA 57
1.7 REFRACTIVE ERROR CORRECTION METHODS 60
1.8 HEALTH DISPARITIES AND POTENTIAL FOR MYOPIA-RELATED LOSS OF VISION 60
CHAPTER 2 | AIMS OF THESIS 63
CHAPTER 3 | METHODS 67
3.1 INTRODUCTION 68
3.2 COHORTS' DESCRIPTION 68
3.2.1 UK BIOBANK COHORT 68
3.2.1.1 REFRACTIVE ERROR ASSESSMENT 68
3.2.1.2 DEMOGRAPHIC FACTORS 69
3.2.1.3 TOWNSEND DEPRIVATION INDEX 70
3.2.1.4 MEDICAL HISTORY AND MEDICATION USE 70
3.2.1.5 ARRAY DATA 71
3.2.1.6 WHOLE-EXOME SEQUENCING DATA 73
3.2.2 23ANDME 74
3.2.3 GERA 75
3.2.4 CREAM CONSORTIUM 76
3.2.5 TWINSUK 79
3.2.6 BDES 81
3.2.7 EPIC 82
3.2.8 SCORM 83
3.2.9 STARS 85
3.2.10 GUSTO 86

3.3 STATISTICAL ANALYSES 87
3.3.1 LINEAR REGRESSION 87
3.3.2 LINEAR MIXED REGRESSION MODELS 88
3.3.2.1 BAYESIAN MIXED MODELS 88
3.3.2.2 MIXED-EFFECT MODELS 89
3.3.3 GENE-BASED ANALYSES 90
3.3.4 COX PROPORTIONAL-HAZARDS REGRESSION MODEL 92
3.3.5 LD-SCORE REGRESSION 93
3.3.6 CONDITIONAL ANALYSES 93
3.3.7 GENE-SET ENRICHMENT ANALYSES 94
3.3.8 MENDELIAN RANDOMISATION ANALYSIS 95
3.3.9 SUMMARY DATA-BASED MENDELIAN RANDOMISATION ANALYSIS 97
CHAPTER 4 | COMMON GENETIC VARIATION INFLUENCING POPULATION'S REFRACTIVE ERROR ... 99
4.1 INTRODUCTION 100
4.2 STATISTICAL ANALYSES 102
4.3 RESULTS 107
4.3.1 META-ANALYSIS 107
4.3.2 POST-GWAS ANALYSES 109
4.3.2.1 FUNCTIONAL ANALYSES 109
4.3.2.2 GENE-SET ENRICHMENT ANALYSES 109
4.3.2.3 TRANSCRIPTION FACTOR ENRICHMENT ANALYSES 111
4.3.2.4 GENE EXPRESSION IN DIFFERENT BODY TISSUES 111
4.3.2.5 MENDELIAN RANDOMISATION ANALYSES 116
4.3.2.6 CONDITIONAL ANALYSES 119
4.3.2.7 MYOPIA PREDICTION MODEL 119
4.3.2.8 VALIDATION OF MYOPIA PREDICTION MODELS IN A COHORT OF SINGAPOREAN CHILDREN 121

4.4 DISCUSSION 127

CHAPTER 5 | RARE GENETIC VARIANTS ASSOCIATED WITH POPULATION'S REFRACTIVE ERROR ... 130

5.1 INTRODUCTION 131

5.2 METHODS 133

5.2.1 STUDY POPULATION 133

5.2.2 WHOLE EXOME SEQUENCING DATA 133

5.2.3 THE CREAM CONSORTIUM DATASET 133

5.3 STATISTICAL ANALYSES 134

5.3.1 GENE-BASED ASSOCIATION ANALYSES 134

5.3.2 SEQUENTIAL ANALYSES EVALUATING THE ROLE OF SINGLE VARIANTS IN GENE-BASED ASSOCIATIONS 135

5.4 RESULTS 136

5.5 DISCUSSION 145

CHAPTER 6 | GENETIC ASSOCIATIONS WITH THE AGE OF REFRACTIVE CORRECTION 150

6.1 INTRODUCTION 151

6.2 METHODS 153

6.2.1 STUDY POPULATION 153

6.2.2 GENETIC DATA 154

6.2.3 STATISTICAL ANALYSES 154

6.3 RESULTS 156

6.4 DISCUSSION 175

CHAPTER 7 | MEDICATION USE AND REFRACTIVE ERROR 180

7.1 INTRODUCTION 181

7.2 ASSOCIATION BETWEEN MEDICATION-TAKING AND REFRACTIVE ERROR IN A LARGE GENERAL POPULATION-BASED COHORT 184

CHAPTER 8 | DISCUSSIONS OF RESULTS ARISING FROM THIS THESIS, CONCLUSIONS AND POSSIBLE FUTURE DIRECTIONS 216

8.1 CONTRIBUTIONS TO RESEARCH AND IMPLICATIONS 217

8.2 LIMITATIONS 228

8.3 FUTURE DIRECTIONS 233

8.4 CONCLUSIONS 236

CHAPTER 9 | APPENDIX 238

CHAPTER 10 | REFERENCES 315

REFERENCES 316

Table of Figures

Figure 1 Sagittal section of the human eye.....25

Figure 2 The position of light rays focusing.....27

Figure 3 Gene-set enrichment analysis for Gene Ontology terms.....110

Figure 4 Genes from the regions associated with refractive error that are particularly expressed in eye tissues, compared to non-ocular tissue..... 112

Figure 5 Mendelian randomisation results on the causality of IOP over refractive error.....117

Figure 6 Mendelian randomisation results on the causality of fluid intelligence over refractive error.....118

Figure 7 Receiver operating characteristic (ROC) curves for myopia predictions, using information from 890 SNP markers identified in the meta-analysis.....120

Figure 8 Receiver Operating Characteristic (ROC) curve for identifying high myopia (≤ −5.00 D) and moderate myopia (≤ −3.00 D) versus no myopia controls using the polygenic risk score in SCORM..... 126

Figure 9 Spherical equivalent distribution in UK Biobank cohort (N = 50,893).....137

Figure 10 Manhattan plot displaying SKAT-O association results: each point represents one of the 19,293 genes tested for the association with the spherical equivalent in the UK Biobank cohort (N=50,893).....139

Figure 11 Sensitivity analyses result for *SIX6* gene.....143

Figure 12 Sensitivity analyses results for *CRX* gene.....144

Figure 13 Distribution of the age of first spectacle wear in UK Biobank Cohort.....158

Figure 14 Distribution of refractive error across the age of first spectacle wear in UK Biobank cohort (N = 81,675).....159

Figure 15 Manhattan plot displaying 44 genome-wide significant associations with the age of first spectacle wear in UK Biobank cohort (N=340,318).....162

Figure 16 Scatterplot displaying the correlation between the AFSW hazards ratios and spherical equivalent beta coefficients.....163

Figure 17 Smoothed log-hazard ratios as a function of age for a selection of six SNPs that showed evidence of deviation from the proportional hazards assumption that are located in regions of previously known association with refractive error.....168

Figure 18 Smoothed log-hazard ratios as a function of age for six SNPs located near or within novel genes. The dashed red line displays the estimated log-hazard ratio (beta) for each variant in the proportional hazards model..... 169

Tables

Table 1 Phenotyping and genotyping methods in CREAM non-UK European cohorts.....78

Table 2 SMR results using retinal eQTL data.....114

Table 3 Genetic correlation between spherical equivalent and other phenotypic traits.....115

Table 4 Multivariate linear regression models of polygenic risk score (PRS), parental myopia, time outdoors and books read per week, assessing the prediction accuracy of spherical equivalent and axial length....122

Table 5 Multivariate linear regression models of PRS, parental myopia, time outdoors and books read per week for spherical equivalent and axial length in SCORM (N = 1004).....123

Table 6 Cox proportional hazard regression models of time to moderate myopia and time to high myopia in SCORM (N = 1,004).....124

Table 7 Characteristics of the study participants.....136

Table 8 Top eight gene associations with refractive error.....140

Table 9 Top three gene associations with refractive error.....140

Table 10 Replication of four loci associated with refractive error using gene-based analyses performed in Beaver Dam (N = 1740) and CREAM Consortium dataset (N = 11,505).....141

Table 11 Demographic and refractive characteristics of the study participants.....156

Table 12 Genetic correlations between SPHE GWAS effects and genome-wide survival analyses.....160

Table 13 Replication of six novel loci associated with AFSW.....165

Table 14 Association with spherical equivalent of the same SNPs associated with age of first spectacle wear.....166

Table 15 Results for the test of deviation from the proportional hazards assumption font.....167

Table 16 SMR results for age of first spectacle wear and eQTL in retinal tissues (Ratnapriya *et al*. Nature Genetics 2019).....170

Table 17 SMR results for age of first spectacle wear and eQTL in the brain tissues (Wang *et al.* 2018 Science).....171

Table 18 SMR results for age of first spectacle wear and eQTL in the foetal brain tissues (Hannon *et al*. 2015; Nat Neuroscience).....172

Table 19 Genetic correlations between age first spectacle wear and 23 different traits.....173

Table 20 Gene-set enrichment analysis for Gene Ontology terms.....174

Table 21 SMR results using the GTEx-brain summary eQTL data.....310

Table 22 SMR results using summary eQTL data from the peripheral blood (Westra *et al.*).....314

Abbreviations

AFSW – age of first spectacle wear

ALIENOR – Age-Related Eye Study

ALSPAC – Avon Longitudinal Study of Parents and Children

ANZRAG – Australian and New Zealand Registry of Advanced Glaucoma

AREDS – Age-Related Eye Study

AUC – area under the curve

BATS – Brisbane Adolescent Twins Study

BDES – Beaver Dam Study

CNS – central nervous system

CNV – choroidal neovascularisation

COJO – conditional and joint association analysis

COPD – chronic obstructive pulmonary disease

CREAM – The Consortium for Refractive Error and Myopia

D – dioptre

DALY – disability-adjusted life years

DCCT – Diabetes Control and Complications Trial study

DNA – deoxyribonucleic acid

DZ twins – dizygotic twins

E3 – European Eye Epidemiology consortium

EAS – East Asia

EGCUT – Estonian Genome Center of the University of Tartu

EPIC – European Prospective Investigation into Cancer study

eQTL – expression quantitative trait loci

ERF – Erasmus Rucphen Family Study

FDM – form deprivation myopia

FDR – False Discovery Rate

FECD – Fuch's endothelial corneal dystrophy

FITSA – Finnish Twin Study on Aging

GERA – Genetic Epidemiology Research in Adult Health and Aging cohort

GUSTO – Growing Up in Singapore Towards Healthy Outcomes cohort

GWAS – genome-wide association study(-ies)

HEIDI – Heterogeneity Independent Instruments (HEIDI) test

HR – hazards ratio

HRC – Haplotype Reference Consortium

HWE – Hardy-Weinberg equilibrium

ICD10 – International Classification of Diseases, Tenth Revision

IOP – intraocular pressure

IQ – intelligence quotient

ISCED – International Standard Classification of Education

IVW – inverse-variance weighted test

KKH – KK Women's and Children's Hospital

KORA – Cooperative Health Research in the Region Augsburg

LASIK – Laser-Assisted in Situ Keratomileusis

LD – linkage disequilibrium

LMMS – linear mixed models

LOF – loss-of-function variants

MMD – myopic macular degeneration

MR – Mendelian randomisation

MZ twins – monozygotic twins

NICER – Northern Ireland Childhood Errors of Refraction study

NUH – National University Hospital

OMIM – Online Mendelian Inheritance in Man

OR – odds ratio

p – p-value

POAG – open-angle glaucoma

PRS – polygenic risk scores

QC – quality control

REHS – Raine Eye Health Study

ROSLA – raising of the school leaving age in the UK education reform

RPE – retinal pigment epithelium

RPGEH – Kaiser Permanente Research Program on Genes, Environment and Health

RR – risk ratio

RSI – Rotterdam Study I

SCORM – Singapore Cohort Study of the Risk Factors of Myopia

SKAT – Sequence Kernel Association Test

SKAT-O – Optimised SNP-set Kernel Association Test

SMR – Summary-data based Mendelian randomisation method

SNEC – Singapore National Eye Centre

SNP – single nucleotide polymorphism

SPHE – spherical equivalent

STARS – Strabismus, Amblyopia and Refractive Error Study

TEST – Twins Eye Study in Tasmania

trbl –Tribbles protein family

UCSF – University of California, San Francisco

UK – The United Kingdom of Great Britain and Northern Ireland

US – The United States of America

USD – US dollar

VCDR – vertical cup-to-disc ratio

VI – vision impairment

WES – whole-exome sequencing

WESDR – Wisconsin Epidemiologic Study of Diabetic Retinopathy

YFS – Young Finns Study

Chapter 1 | Introduction

1.1 Overview

Refractive errors, and especially myopia, are one of the common ocular disorders and a leading cause of preventable blindness worldwide. Myopia is not only associated with complications that lead to visual impairment and potentially loss of sight but also contributes to reduced quality of life and significant economic and social burden. Both genetic and environmental factors play a role in myopia development. However, despite extensive genetic and epidemiological research, exact biological processes that lead to disease development and progression remain elusive.

In this chapter, I will discuss what refractive errors and myopia are, provide an overview of the current epidemiology and disease aetiology, including the findings of observational as well as twin and genome-wide association studies (GWAS), and explain how their results advance our understanding of refractive error and myopia. Lastly, I will report on potential complications that arise even after refractive correction and the economic and social costs associated with refractive error.

1.2 Eye anatomy

The human eye has a complex anatomy and consists of multiple structures that can be described as individual elements of three distinct parts (2) (Figure 1). The outer region includes the cornea and sclera, connecting at the limbus. The cornea reflects and transmits light to the lens and retina, but also protects the deeper layers from infection and structural damage. The sclera creates a protective coat of connective tissue and helps to maintain the shape of the ocular globe. The visible part of the sclera is protected by a transparent mucous membrane, the conjunctiva. The middle region of the eye is comprised of the iris, the ciliary body and the choroid. The iris and ciliary body control the amount of light that reaches the retinal plane with the lens regulating the size of the pupil and the ciliary body determining the power and shape of the lens (2). With constriction of ciliary muscles, the lens which is placed between the iris and the vitreous

chamber, assumes a more spherical shape, resulting in increased optical power which brings near objects into focus. Conversely, when ciliary muscles relax, the lens flattens which helps to bring more distant objects into view (3). The choroid is a middle vascular layer in the wall of the eye which supplies nutrients and oxygen to the outer layers of the retina. The retina is the inner layer that covers two-thirds of the ocular globe. The concentration of retinal cells increases from the periphery to the central part of the retina recognisable by the macula containing the cone-enriched fovea (2). Two major types of cells, neurons and glial cells form the retinal tissue (4). The retina is responsible for the phototransduction cascade and relays electrical signals to the brain. Retinal neurons can be grouped into subclasses based on their function. In particular, photoreceptors, namely cones and rods, bipolar and ganglion cells transmit electrical signals vertically, while horizontal cells, including amacrine and interplexiform cells, modulate the impulse transmission (4).

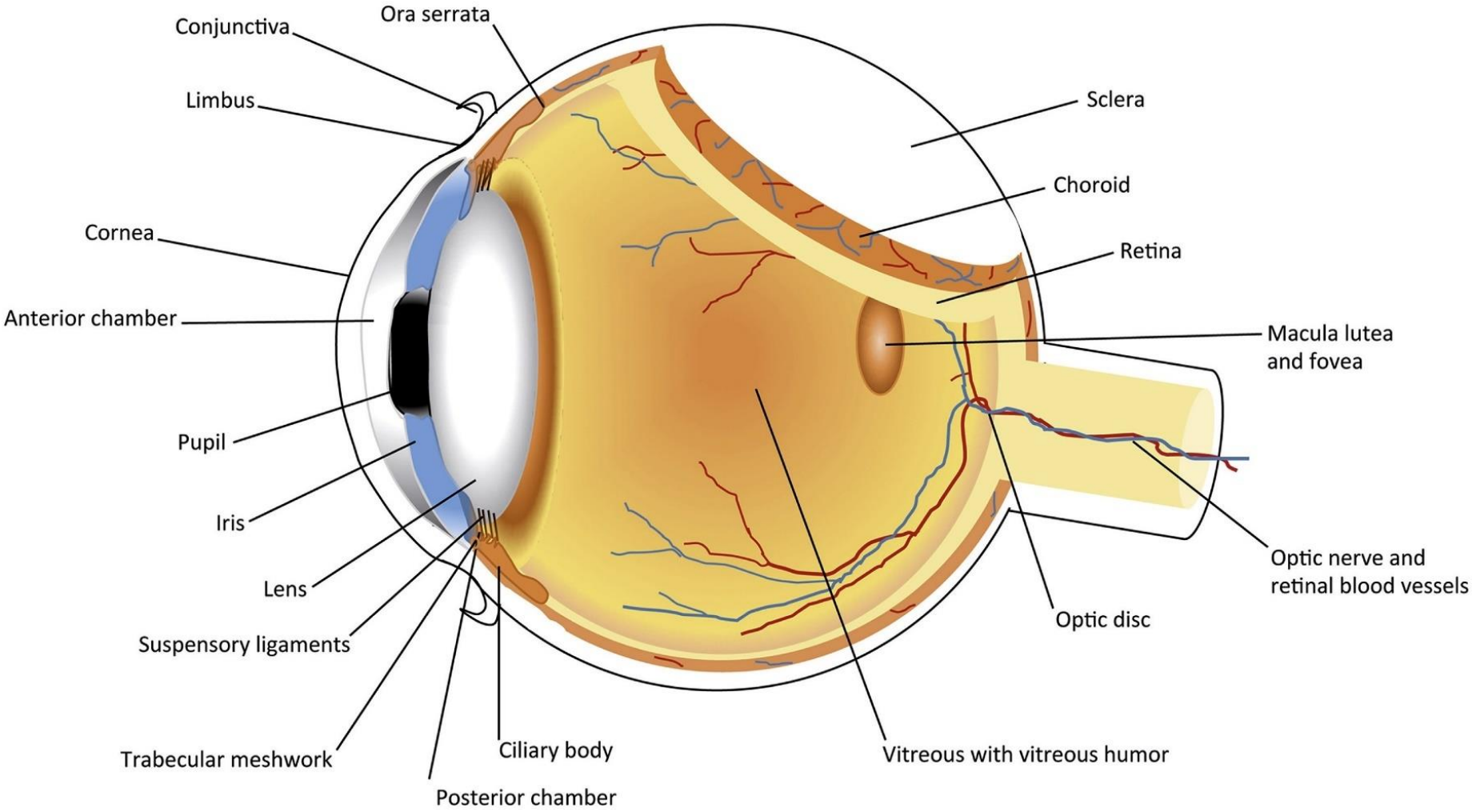


Figure 1 Sagittal section of the human eye. Adapted from Mohlin *et al.* "The link between morphology and complement in ocular disease" (2).

1.3 Definition and natural history of myopia

Refractive errors, including myopia, can be defined as the optical power of the lens required to overcome optical defocus and direct the light onto the retina (5) The degree of defocus is estimated by calculating spherical equivalent measured in dioptres (D), defined as the sum of the sphere power and half of the cylinder power. Spherical equivalent is commonly used to estimate the degree of farsightedness (hypermetropia) and near-sightedness (or myopia) and describe their severity (6).

The amount and degree of refractive error are determined by the optical power of the lens and cornea as well as the axial length of the eyeball (7). Axial myopia occurs when the excessive eye globe elongation causes the image of a viewed object to project in front of the retina (7) (Figure 2). Refractive or index myopias emerge as a consequence of lens and cornea changes rather than abnormal ocular growth. In the cornea, these changes involve progressive thinning and bulging of the tissue, as seen in keratoconus (8). Comparably, alterations in the refractive index of the lens caused by abnormal cell structures or proteins amplify light scattering and lead to blurred vision (9). However, the majority of myopia is caused by abnormal ocular growth and progressive eyeball elongation.

Myopia can be categorised based on spherical equivalent cut-offs. Because there is no universally accepted classification of myopia, near-sightedness is often defined as below a threshold variably set anywhere between 0 and -1.00 D (6). Similarly, the degree of myopic refractive error is described using different spherical equivalent thresholds with some authors utilising a binary division of low (≤ -1.00 D) and high myopia (≤ -5.00 D), and others distinguishing between mild (≤-0.5 D), moderate (≤ 3 D) and severe (≤ - 6 D) forms of myopia (10,11). The paediatric studies using non-cycloplegic refraction techniques (without pupil dilation) require lower spherical equivalent thresholds due to potential bias toward myopic refraction. Although cycloplegia is considered a gold

standard for measuring refraction in children and adolescents (6), it is not required for population estimates of refractive error in adults (12).

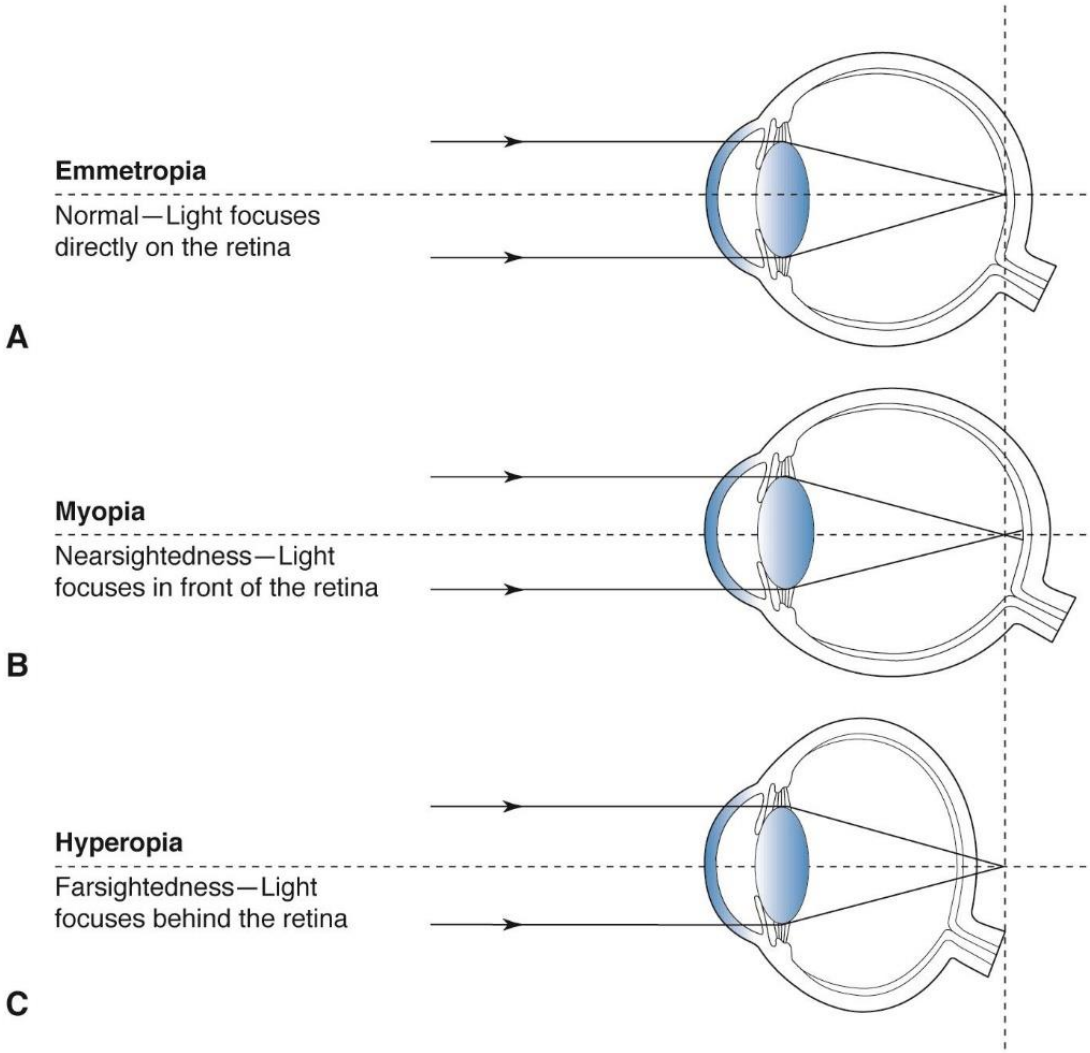


Figure 2 The position of light rays focusing. A, The normal eye. B, A myopic eye. C, A hypermetropic eye. Illustration developed by Scott E. Brodie, MD. © 2021 American Academy of Ophthalmology

Refractive errors, particularly axial myopia, are hypothesised to arise from a failure of homeostatic control of ocular growth (13). Emmentropisation is a dynamic developmental process that regulates the axial length and eye growth to match the optical power (14). Ocular growth, controlled through a visual feedback mechanism, facilitates emmentropisation and causes a gradual shift in refractive error distribution (14). In the neonatal period, the refractive error curve is wider and centred in hypermetropic ranges (15,16). The lengthening of the ocular globe drives reduction towards emmetropia in infancy and early childhood when the eye experiences a rapid axial length elongation from 18 mm at birth to 22-23 mm at the age of three (14). At this time, the cornea reaches its adult size and contributes 2/3 of the eye's refractive power (17,18). Subsequent refractive development depends on a gradual and controlled increase in axial length, rather than changes in cornea curvature (14). The optical power of spherical equivalent decreases proportionally to the eye size, which results in a gradual change from low hypermetropic to emmetropic refractive error (19) (Figure 2). Between the ages of six and twelve, the depth of anterior and vitreous chambers increases by 0.22 mm and 0.52 mm, respectively, which corresponds to an axial length change of 0.66 mm (20). Changes in other optical components neutralise an excessive myopic shift that could occur due to an increase in vitreous chamber depth (20). The flattening of the cornea and thinning of the crystalline lens lower the net refractive power and ensure a smooth transition from low hypermetropia to emmetropia (20). However, if predetermined scleral growth exceeds cornea/lens flattening rate or retinal modulating response, this could result in myopia (21) (Figure 2). In contrast, if the axial elongation is too slow for the flattening rate of the cornea/lens, the focal point moves behind the retinal plane, and hypermetropia may develop (21) (Figure 2). Nonetheless, barring individuals of East-Asian descent, the majority of the human population never develops refractive errors because the axial length and corneal and lens power match each other (22).

The age-dependent increase in ocular axial length, which drives the development of myopia, often coincides with the start of formal education (14). Scientific literature often distinguishes childhood, juvenile, school, early and late adult-onset myopias (6). Childhood-onset myopia is the most common form of near-sightedness. Myopia development at an earlier age in childhood is associated with a rapid disease progression and high-degree myopia in adulthood (23). Progressive high myopia can lead to vision-threatening complications such as choroidal neovascularisation, retinal breaks, myopic macular degeneration and optic disc changes in adolescents and adults (24,25). It has been suggested that 70% of children that require vision correction before puberty eventually progress to moderate myopia of -3.0 to -5.75 D, while 18% of early myopes develop refractive error of -6 D or greater by the age of sixteen (26). Spherical equivalent at the age of 6 –11 appears to be a single most consistent predictor for juvenile-onset myopia (27).

Although childhood myopia appears to be predictive of future severe myopic refraction, this correlation is not perfect. Despite generally reflecting myopia severity, self-reported age of refractive correction explains only 15% of the spherical equivalent variance in cross-sectional investigations (28). By contrast, genetic studies report high concordance of effects between self-reported age of myopia onset (29) and directly measured refractive error (30,31), indicating that the age of refractive correction is a reasonable proxy for the final refractive status in the large-scale population studies.

1.4 Epidemiology of myopia

1.4.1 Global prevalence rates of myopia in children

Refractive errors, especially if uncorrected, negatively impact the quality of life, limit employability and affect academic performance (32). Moreover, myopia, in particular, is associated with several ocular complications and could lead to visual impairment and blindness (33). Children with early-onset near-sightedness is a group at major risk, as they will have a longer duration of the disease, will exhibit more rapid myopia

progression and likely develop high degree myopia in the future. Given that near-sightedness is an irreversible and untreatable condition that starts in childhood and adolescence, the higher prevalence of myopia in small children often means a greater health burden in adulthood.

Myopia prevalence in children varies considerably across different regions reaching epidemic levels in some countries (34). A recently published literature review on the global prevalence estimates of near-sightedness in schoolchildren revealed that East Asian countries were disproportionally affected by myopia. In particular, the prevalence of myopia was twice higher in East Asians than in similar age Europeans (1). The highest prevalence rates were recorded among schoolchildren in Singapore, Taiwan and South Korea (35). In particular, myopia was observed among 86% and 80% of 15-year-olds living in Singapore and Taiwan, respectively (35). Similarly, in Hong Kong, the prevalence of myopia among adolescents reached 78% (35). Observational studies conducted in South Korea estimated that close to 73% of children aged 5-18 years were affected by myopia. Conversely, Japan exhibited moderate prevalence rates of school myopia compared to those in Chinese and South Korean populations (36), whereas the lowest rates were documented in Mongolian children (35).

Outside Asia, the prevalence rates of paediatric myopia tended to be lower. For example, in Australia, myopia was observed among 1.4% of six-year-olds and 12% of twelve-year-olds (37). However, the prevalence of near-sightedness varied between ethnic groups – children of East (39.5%) and South Asian (31.5%) descent was more likely to be myopic than their European (4.6%) and Middle Eastern (6.1%) peers (38). Likewise, in the United States, the prevalence of near-sightedness in paediatric populations of European ancestry varied between 4.5% and 28% among 6- and 12-year-olds, respectively (39). Another study of American adolescents aged between 5 and 17 years suggested that prevalence rates could be higher in non-European ethnicities; myopia was more prevalent in children of Asian (18.5%) and Hispanic (13%) descent and relatively low in their African (6.6%) and European (4.4%) counterparts (40). In the South

American continent and Africa, school myopia was not common, and the prevalence rates remained under 10 % (41).

In Europe, overall paediatric myopia's prevalence rates were lower than in Asia (60% vs 40%) (32) but varied across the continent. For example, in the United Kingdom, myopia was present in 2.8% of children aged six or seven and was observed in 17.7% of adolescents, as early as 12 and 13 years (42). Compared to the UK population, a teenage population in Sweden demonstrated a greater refractive tendency towards myopisation (43). In particular, myopia had a prevalence of 49.7% in 12- and 13-year-olds, whereby the proportion of bilateral near-sightedness in the same age group was 39% (43). Additionally, high myopia, defined as refractive error <= -5 D, was observed in 2.5% of Swedish children (43). Similarly, in France, the prevalence rates of school myopia reached 42.7% (SPHE <=-0.5 D) in adolescents aged 10-17 years, while 1.2% and 0.3% of teenagers had high (SPHE <=-6 D) and severe myopia (SPHE <=-10 D), respectively (44). In South Europe, the prevalence rates were lower, for example, in Greece more than 28% of school students were myopic, while only 14.1% of students had myopia in Bulgaria (45).

1.4.2 Myopia prevalence in urban vs rural areas

Previous studies reported higher myopia prevalence rates among children residing in urban settings compared to children from the same ethnic groups living in rural areas (35). Stronger differences between urban and rural environments observed in East Asian and South Asian countries reflected greater socio-economic disparities compared to developed economies (35). In particular, longitudinal surveys conducted on Taiwanese schoolchildren between 1995-2000 revealed an association between highly urbanised residential areas and greater prevalence rates of myopia (46). Specifically, the prevalence of school myopia was highest in urban regions (45.7%), followed by suburban (40%) and rural areas (29.5%). Moreover, among the surveyed Taiwanese adolescents living in urban settings was associated with a 2-fold increase in the risk of myopia (46).

Similarly, In China, the prevalence of childhood myopia was greater in urban than in rural regions. In urban areas, 6% of children aged 5 and 30% and 75% of adolescents aged 10 and 15 respectively, had myopia (47), whereby in rural settings, almost no five-year-olds were myopic, and only 37% of thirteen-years-olds, 43% of fifteen-year-olds and 54% of adolescents aged 17 had developed near-sightedness (48). Likewise, in South Asia, a similar trend was also observed. In urbanised residential areas of India, the prevalence of paediatric myopia was higher than that in rural parts of the country. Myopia was identified in 5%,7% and 10% of five-year-olds, ten-year-olds and fifteen-year-olds who lived in urban areas (49). By contrast, the prevalence of near-sightedness was lower in rural parts of India, specifically, 3%, 4% and 7% in children aged 7, 10 and 15, respectively (49). Similarly, in Nepal, 11% of eleven-year-olds, 17% of twelve-year-olds and 27% of fifteen-year-olds had myopia, while only 1.2% of rural children aged between 5 and 15 years were myopic (50).

Unsurprisingly, populations with a high prevalence of near-sightedness also exhibited rapid myopia progression rates. Asian children living in urban areas were observed to progress faster than their Western counterparts (51). The meta-analysis of myopia progression rates in urban adolescents found that nine-year-old Asian children on average, progressed by -0.82 D over one year, while in predominantly European populations progression rate was -0.55 D (51). From the findings of these studies, rapid urbanisation was one of the major cultural and economic contributors to the myopic transition and ongoing myopia epidemic in Asia (35).

1.4.3 Cohort effects in myopia prevalence

In the last thirty years, East-Asian populations experienced an unprecedented rise in paediatric myopia prevalence rates (52), with an estimated average of 23% increase per decade (35). In Taiwan alone, from the mid-80s to the early 2000's a prevalence of near-sightedness rose from just 6% and 37% to 21% and 61% in seven-year-old and twelve-year-old children, accordingly (52). In fifteen-year-old Taiwanese adolescents, the

prevalence rates increased from 64% to 84% in 2000, while among sixteen and eighteen-year-olds, the rates had risen from 74% to 84% (52). Likewise, a 2010 Singaporean study of Chinese children myopia was identified in 11% of the participants aged between 6 and 72 months (53). In the children aged 7-9 years, the prevalence varied between 29% and 53% (54). Similarly, in Korea, studies of age-specific prevalence indicated a significant generational shift. In particular, 50% of children aged between 5 and 11 and 78% of teenagers who were 12 and older had myopia, whereas the prevalence rates in high-school students were lower – 45% (55).

Previous observational work assessed the evidence for the myopic transition by examining changes in myopia prevalence in older age groups. The 2015 meta-analysis of Asian studies spanning 16 countries and regions found a considerable variation in the age-specific prevalence rates of myopia (56). The age-specific pooled estimates of this meta-analysis implicated strong generational effects (56). In particular, myopia was prevalent among 47.4% of Asians aged 20-29 years, whereas, among individuals aged 40 years or older, the pooled prevalence of myopia reached 30% (56). The increase of prevalence rates among younger individuals was more pronounced due to recent cohort effects in urban Asian populations exposed to rigorous schooling (56).

Although population-based studies indicated that myopia was less common in Europe, affecting about 3-5% of younger children and up to 20% of adolescents (35), there was an observable time trend towards myopisation. For example, a 2016 study revealed that incidence and frequency of near-sightedness were relatively low in the UK contemporary population of European ancestry compared to other countries (57). Still, the prevalence of near-sightedness almost doubled during the last 50 years in teenagers aged 10 – 16 years, and children also developed myopia at a younger age (57). In the 2016 NICER study, myopia was observed in 1.7% of children aged 6-7 years (57). Moreover, the proportion of myopes increased significantly during adolescence (12-13 years), more than 14.6 % ($p<0.001$) of children in this age group were myopic (57). There was a less dramatic increase between 12–13-year-olds and 18–20-year-olds with a 2 –

4% change in prevalence rates for a younger and older cohort accordingly (57). More children in the NICER study were myopic at the age of 12 – 13 than in the Sydney study (16% vs 4.4%, p<0.001) (57).

Similar to other regions with ongoing rapid myopisation, in Europe, a moderate increase in myopia prevalence coincided with the rise in education levels, as exposure to tertiary education changed from 5% to 50% (34). In particular, a marked cohort effect for myopia prevalence was observed across recent birth cohorts in the 2015 E3 consortium meta-analysis. These analyses revealed significant generational effects for growing myopia prevalence rates (58): approximately 18% of individuals born between 1910 and 1939 had developed myopia during their lifetime, whereas in the 1940 –1979 birth cohort, the prevalence of myopia was higher, 24% (58). The generational effects of increasing myopia in European populations were attributed to the additive effect of educational levels and birth cohort (58). The relationship between refractive error and educational attainment will be discussed in detail in paragraph 1.5.2.1

Studies from the United States also report an increase in the prevalence of near-sightedness over the past few decades (59). A study by Vitale *et al*. assessed myopia prevalence rates in the United States between 1971-1972 and 1999-2004 (59). The investigation found a significant increase in overall myopia prevalence from 25% in 1971-1972 to 41.6% in 1999-2004 (59). Moreover, the prevalence rates were higher in individuals of European (43.0% vs 26.3%) and African descent (33.5% vs 13.0%) in 1999-2004 compared to 1971-1972 as well as for all degrees of near-sightedness: for mild (17.5% vs 13.4%), moderate (22.4% vs 11.4%) and severe (1.6% vs 0.2%) myopia (59).

These findings supported the evidence of a generational increase in the prevalence rates of myopia and gradual myopisation of the population.

1.5 Aetiological risk factors

Myopia is a common ocular disease that develops due to a complex interplay between genetic and environmental factors. Twin and family studies suggest that refractive error is a highly heritable trait with 70%–90% of its variance explained by genetic factors, also known as heritability. Conversely, exposures that influence the prevalence of near-sightedness are diverse and range from perinatal to early development. This chapter will discuss genetic factors that contribute to refractive errors and myopia but also lifestyle and behavioural factors such as time spent outdoors, demographic characteristics, including education and socioeconomic status, and personal characteristics such as and mental capacity.

First, I will present the findings of heritability studies and provide arguments for studying genetic factors influencing refractive development. I will summarise evidence from the most recent genome-wide and genomic sequencing studies that shed a light on molecular mechanisms implicated in the pathogenesis of refractive disorders. Next, I will describe key environmental factors associated with refraction that interact with each other and drive global trends in the prevalence rates of near-sightedness.

1.5.1 Genetics of refractive error and myopia

Refractive errors are aetiologically heterogeneous, as both environmental exposures and hereditary factors are involved in their pathogenesis. Refractive error and especially severe myopia demonstrate clear familial aggregation. There is also strong evidence for the impact of genetic factors outweighing environmental influences and largely determining the trait variance. A number of population-based epidemiological investigations suggest a strong relationship between parental myopia and the risk of myopia in the offspring (60–63). Additionally, twin and family studies assessing the contribution of genetic factors to the phenotypic variance report high heritability of the trait (64,65). To date, research has identified hundreds of different chromosomal regions associated with myopia or related traits, but the knowledge about their role in

myopia development is limited. In this chapter, I will give an overview of the genetic findings on non-syndromic myopia, using current data from twin studies and genome-wide association studies and briefly discuss the results from candidate gene and linkage analyses.

1.5.1.1 Heritability studies

Heritability is the fraction of trait variation among individuals within a population that is explained by their genotypes (66). The heritability of refractive error (measured as spherical equivalent) has been assessed in many family and twin studies. The heritability estimates vary widely from a low of 10% reported by family investigations of Eskimos (67,68), to a high of 98% heritability observed in studies of female twins (65,69,70). This broad range can be partly attributed to differences in study design and methodologies (70), but also a considerable degree by unequal environments (71). Specifically, in populations where environmental triggers are low, the contribution of genetic variants leads to higher heritability estimates and, conversely, intense environmental exposures lead to lower heritability estimates (71).

A number of twin studies have been performed on refractive error and myopia in different decades and across various geographical regions. The models used in twin investigations estimated the proportion of observed variance explained by additive and dominant genetic effects (A), shared environment influences (C) and effects from the unshared or unique environment (E). Twin studies ascertained the contribution of genetic effects based on assumption that monozygotic twins shared all of their genes, whereas dizygotic twins shared 50% of their genetic material. Notably, the heritability estimates of refractive error and myopia as well as related traits from recent large twins' studies were consistently higher compared to the estimates reported by family studies (72,73). Twin investigations suggested high heritability of myopia ranging between 70% and 90%,(65,74–76) and found no significant effect of shared environment. Likewise, the pooled heritability estimates of several refractive biometric traits indicated

refractive error heritability of 71% (70). Similarly, to all-cause refractive error, myopia exhibited considerable variation in heritability (70), with literature attributing 60% – 80% of myopia variance to genetic factors (71).

1.5.1.2 Linkage and candidate gene studies

Before the era of genome-wide association studies, the identification of disease susceptibility genes relied on family linkage and candidate gene investigations. Several linkage studies sought the evidence of genetic marker co-segregation with myopia through pedigrees (71). At least 30 candidate regions, including MYP family genes (14,77–79) as well as a handful of other loci (80–85) were identified using genetic linkage in families with dominant inheritance myopia. Fine mapping of the loci helped to detect additional candidates, such as the *IGF1* gene located within a broader *MYP3* region (86). Although most loci discovered in linkage studies were rarely replicated independently, some of these rare variants were linked to common myopia, indicating a degree of genetic overlap between Mendelian and complex near-sightedness (87).

Candidate gene studies selected genetic targets based on existing knowledge about their biological, physiological and functional relevance (88). However, similarly to family-linkage investigations, candidate gene studies suffered from several methodological issues. In particular, genetic variance across populations made it difficult to differentiate between normal and disease-causing variation using the candidate gene approach (89). This method relied on existing literature and was also vulnerable to publication bias (88).

Candidate gene studies have identified several genes associated with myopia and ocular refraction. In particular, genes implicated in collagen syntheses, such as *COL1A1* and *COL2A1* (90,91), as well as candidates encoding a host of growth factors (92–94) and metalloproteinases (95,96) showed evidence of the association with myopia. Additionally, candidate gene studies were among the first to report the association between ocular development gene *PAX6* (97) and extreme near-sightedness (97). This association with a single *PAX6* variant was consistent across different candidate gene

investigations and replicated in the meta-analysis (98) and subsequent GWAS (99). Nonetheless, similarly to linkage analyses, most candidate gene studies lacked replication in independent samples.

1.5.1.3 Genome-wide association studies on refractive error and myopia

In the last decade, genome-wide association studies (GWAS) became the dominant method for genetic interrogation of complex traits and diseases, leading to the discovery of thousands of loci associated with many hundreds of different diseases and other phenotypic traits. GWAS identify genetically determined molecular mechanisms underlying the pathogenesis of common disorders (100). GWAS rely on the "Common Disease, Common Variant" hypothesis, according to which high-frequency genetic variations have lower magnitudes of effect but contribute most of the genetic predisposition to common diseases and traits (100). GWAS tests millions of genetic polymorphisms across the genome, using high-throughput genotyping arrays or chips, and examine associations between common traits and common genetic polymorphisms in the form of SNPs (100). These genetic associations are further interrogated in post-GWAS analyses to gain insights into disease pathogenesis and find core biological pathways. Since 2005 more than 3000 GWA investigations have been published, including over 30 reports on refractive error, myopia and related phenotypes (101)

Refractive errors and myopia were studied exploiting a wide array of genetic methodologies (71). However, greater success was achieved by conducting genome-wide association studies that discovered hundreds of genetic loci and replicated them across different populations and ethnicities. The first published GWAS were case-control studies of severe and pathological myopia performed in small cohorts of East Asian patients (66,102) which led to the discovery of at least ten myopia susceptibility loci (103–105). However, the majority of currently known loci were identified in GWAS that considered refractive error as a quantitative trait and analysed genetic data from large

population-based cohorts rather than a selected sample of clinic-based highly affected cases (100). The first GWAS utilising spherical equivalent as a quantitative proxy of refractive error was carried out in two European cohorts and yielded two novel genome-wide significant associations with one SNP located near the *RASGFR1* gene on *15q25.1* (30) and another placed near *GJD2* on *15q14* (106). These genetic signals were replicated in a larger multi-ethnic study which also reported an additional novel variant within the *RBFOX1* gene (107). However, these findings were surpassed by the Consortium on Refractive Error and Myopia (CREAM) meta-analysis which was performed using HapMap2 imputation and aggregating spherical equivalent data from 35 different cohorts of European and Asian ancestry (31). The meta-analysis validated associations with *RFBOX1*, *RASGRF1* and *GJD2* and enabled the detection of 23 new loci at a genome-wide significance level, including associations with *LAMA2, PRSS56, RDH5* and others (31).

In the same year, direct-to-consumer genomics company *23andMe* published a genome-wide survival analysis of myopia onset using the self-reported information from 55,177 participants of European descent (108). The analyses utilised the phenotype of self-reported myopia and the age of first spectacle wear as a proxy for the severity of myopic refraction (108). The study by *23andMe* identified 22 genomic hits of which 11 were completely novel. Surprisingly, the survival analyses found genetic associations that were highly comparable to the genetic effects seen in CREAM which were obtained using more precise refractive error measurements (108). The time-to-event analyses conducted by *23andMe* replicated 8 genomic loci reported by CREAM (109). Conversely, CREAM confirmed 16 genetic associations identified by *23andMe* (109). Despite these genome-wide analyses using different phenotypes and measurement scales – dioptres and hazards ratios, there was a linear relationship between the effects sizes for common significant SNPs obtained from both investigations (109). Remarkably, the locus-specific hazards for myopia onset before 30 correlated with subtle refractive error changes later in life (109).

Although CREAM and *23andMe* discovered a large number of susceptibility loci, cumulatively they explained only 3% of refractive error variance (31,108). A large meta-analysis combined the data from CREAM and *23andMe* with subsequent replication in UK Biobank and increased the number of validated loci to 161 (110). Moreover, based on the observed strong genetic correlation between Asian and European cohorts, the results of the meta-analysis suggested that the genetic architecture of refractive error was relatively similar between Asian and European populations (110). Collectively the genetic signals identified in the meta-analysis of CREAM and *23andMe* GWAS predicted 7.8% of phenotypic variance (110), which was twice as large as the fraction of variance attributed to previously reported genetic associations. However, further improvements in explained phenotypic variance could be achieved with larger sample sizes. The meta-analysis also assessed the predictive utility of polygenic risk scores (PRS) which was calculated for each individual using a sum of refractive error-associated alleles across 161 susceptibility loci weighted by their effect sizes. The model based on PRS and adjusted for age and sex was able to differentiate between myopia and hypermetropia, yielding a predictive value of 0.77 (110).

Previous experimental studies proposed the hypothesis that a retina-to-sclera signalling cascade was driving eye growth and influencing the refractive state (111). However, before the integration of GWAS with functional genomics data, there was limited information on the molecular drivers of retinal circuitry (71). A number of large GWAS identified common polymorphisms located within or near genes involved in retinal signalling (31,108,112).

In particular, among enriched gene pathways in CREAM analyses were extracellular matrix remodelling, neurotransmission, ion transport, metabolism of retinoic acid and ocular development (31). Similarly, *23andMe* found genetic evidence implicating visual cycle, neurogenesis, ocular growth and retinal ganglion cells (108). A pathway analysis combining the data from CREAM and *23andMe* revealed gene enrichment for a multitude of processes including cell-cell adhesion, cation channel activity, synaptic

transmission, calcium ion binding (113). Additional examination of protein-protein networks detected significant overrepresentation of the genes involved in cell cycle and growth, specifically TGF-beta/SMAD and MAPK pathways (113).

The 2018 meta-GWAS of CREAM and *23andMe* cohorts further expanded these findings, providing more in-depth insights into retinal signalling (110). The results of this meta-analysis underscored the role of TGF-beta signalling and found evidence of gene enrichment for the dopamine pathway which had been extensively studied in experimental models for its light-driven mediation of ocular growth (114). Additionally, many refractive error-associated polymorphisms were placed within genes that participated in anterior segment morphogenesis and angiogenesis (110). However, meta-gene sets affecting photoreceptor development and characteristics were ranked as the most important (110). Specifically, analyses found an overrepresentation of the genes involved in the detection of the light stimulus and linked to abnormal photoreceptor inner segment morphology and thin retinal outer nuclear layer (110). Together, these results indicated a prominent role of retinal cell physiology as well as light processing in the refractive development with strong evidence of involvement of all retinal cells, including retinal pigment epithelium, vascular endothelium and extracellular matrix (110).

Before the GWAS studies, the efforts to identify genes associated with refractive error and myopia relied on family and candidate gene studies which were underpowered and difficult to replicate (71). However, with the advent of GWAS, there were over a hundred genomic regions associated with refractive disorders (71) and the number of detected refractive error-associated genes exponentially increased with growing sample sizes (71). Nonetheless, as with other genome-wide investigations, there was a discrepancy between the number of declared susceptibility loci and the amount of phenotypic variance explained by the entire collection of genotyped SNPs (115). Previous studies suggested that about 35% of refractive error variation in unrelated individuals could be attributed to common polymorphisms (116). By contrast, genetic variants reported by

the currently largest meta-GWAS were able to collectively predict only about 8% of refractive error variance (110). However, previous research has demonstrated the most predictive and useful insights into the impact of common polymorphisms were uncovered after the study samples had reached hundreds of thousands of participants. Therefore, the increase in the number of identified loci and their predictive power could be achieved by genotyping large cohorts and meta-analysing their results. Additionally, larger GWAS would enable the detection of higher functional mechanisms indirectly influencing emmentropisation and refractive development. In *Chapter 4*, I describe what we could learn from post-GWAS analyses of the largest meta-GWAS of refractive error, that highlighted new biological mechanisms and expanded our understanding of myopia.

1.5.1.4 Next-generation sequencing studies on myopia

Until recently insights derived from large genetic investigations have only been available for common and low-frequency polymorphisms because SNP arrays used in traditional GWAS did not cover many directly genotyped rare and ultra-rare markers and the potential for their imputation is low, due to weak linkage disequilibrium with other markers (117). Generally, an inverse relationship was observed between variant effect size and their allele frequency in the population; markers with larger effects tended to be rarer (118). The impact of rare variance has been assessed in family-based and rare disease studies by combining genetic data from phenotypically similar probands and identifying associated groups of rare markers (117). However, contributions of rare polymorphisms to the variance of common traits have only been evaluated for selected phenotypes (119–121), as large datasets with phenotypic and exome information were not available until very recently.

Population-based whole-exome sequencing studies (WES) provided a superior alternative to linkage and candidate gene investigations of rare genetic variance (117). Until recently, rare variant analyses using next-generation sequencing were performed

on a selected number of common diseases, precisely developmental delay, diabetes and schizophrenia (117). However, this changed with the availability of data such as the UK Biobank that included close to 200,000 exomes matched to thousands of phenotypes (117).

To date, there was no published large population-based whole-exome sequencing study performed on refractive error or myopia. Previous myopia WES analysed only cases or families with a particular form of refractive disorders, such as myopic anisometropia, or X-linked near-sightedness (122–124). Currently, at least 17 non-syndromic myopia genes were discovered through next-generation sequencing, of which 11 were autosomal dominant, four autosomal recessive and two were X chromosome-linked genes (101). These genes were involved in mitochondrial function (*NDUFAF7, SCO2*) and DNA transcription (*ZNF644*, *CCDC111)*; they also participated in collagen synthesis (*P4HA2*), cell and retina-specific signalling (*BSG*, *NC5D*) and were implicated in the TGF-beta pathway (*LRPAP1*, *SLC39A5*, *LOXL3*). Additional 29 potentially pathogenic variants were identified by examining a large number of markers within retinal dystrophy genes in probands with early-onset high myopia (125).

Although several next-generation sequencing studies identified novel genes associated with high myopia, some of the mutations did not show a distinct cosegregation with the disease in the family (101). Previous exome sequencing work focused on multigenerational proband families and small case-control comparisons, reporting results with a high probability of false positives. Non-replication was common in these studies for reasons, such as power limitations, small samples, multiple testing within and between studies, publication bias and overestimation of true effect sizes (88). Moreover, genes that carried mutations increasing the risk of myopia in specific families, were not necessarily important for the overall population's risk. Given that previous attempts to find rare variants driving the population's refractive error were largely unsuccessful, to assess the impact of rare and ultra-rare genetic variance I exploited the

UK Biobank exome-sequencing project and present the findings of this investigation in *Chapter 5*.

1.5.2 Environmental factors associated with refractive error and myopia

The rise in myopia prevalence rates mirrors generational shifts in lifestyle and behaviours with children in many countries spending a considerable amount of time indoors engaged in near-work activities such as studying, reading and writing. The evidence has linked the current myopia epidemic to a wide range of behavioural and lifestyle factors that interact with each other and create a myopiagenic environment. This section will provide an overview of environmental exposures associated with refractive errors and myopia.

1.5.2.1 Educational attainment

The relationship between educational attainment and myopia is well documented and has been replicated by multiple observational studies over the last 150 years and in many parts of the world (126). In developed economies, myopia onset occurs during childhood, specifically in school years (127). Several cross-sectional and longitudinal studies observe a clear dose-response relationship between years of education and near-sightedness risk. Individuals with a university degree are four times more likely to become myopic than those with primary education (127). In this context, education attainment levels are a composite product of near-work activities taking place in childhood, adolescence and early adulthood (120), including reading, writing and time spent watching a computer screen. Near-work puts an additional strain on the developing eye by requiring extra refractive power from the accommodation of the lens, to help focus the image on the retina. In addition to near-work, educational activities often entail less time spent outdoors, which is another independent risk factor for myopia (39) and is discussed in more detail in paragraph 1.5.2.4.

Rapidly rising myopia prevalence rates were associated with increasing educational levels in Europe (58) and Asia(128). The connection between increasing populations education levels and the myopia epidemic was evident in East Asia, where the proportion of individuals with university degrees markedly changed from 6% in 70+-year-olds to 46% in 20–29-year-olds (128), while the prevalence rates of myopia increased from 10.5% in older cohorts to 78.5% in recent birth cohorts (128). In parallel, a trend towards the increasingly early-onset myopia was also observed. As a result of intense early education at the preschool level and limited time spent outdoors, 50% of Asian primary school children were myopic (129), whereas only 10% of similar age children had a myopic vision in the UK ALSPAC study (130).

Similarly, rising prevalence rates of myopia in Europe were also associated with increasing levels of education. Analysis of data from Northern and Western Europe have found a significant cohort effect for rising prevalence rates of near-sightedness in Europe (58). In particular, the study observed a 5.7% difference in prevalence rates between the older and younger cohorts (17.8% for the 1910-1939 birth cohort vs 23.5% 1940-1975 birth cohort) (58). Moreover, educational attainment was significantly associated with near-sightedness. Myopia was more common in individuals with higher education (36.6%) compared to the participants with primary (25.5%) and secondary (29.1%) education (58). Interestingly, individuals with higher education or those born in the 1960s had 2.43- and 2.62-times greater myopia prevalence rates than the reference group that included participants born in the 1920s who also completed primary education. In contrast, individuals born in the 1960s who also attained university degree had myopia prevalence rates four times higher than that of the reference group (58). Thus, a combination of a more recent birth year and higher education had an additive effect on increasing myopia prevalence rates in Europe (58).

Work by Plotnikov and colleagues assessed the causal effect of education on refractive error using regression discontinuity analysis (131). In particular, the introduction of ROSLA (raising of the school leaving age in the UK) education reform was utilised as a

natural experiment aimed to evaluate the effect of additional compulsory schooling on refraction (131). The reform, first implemented in England and Wales in 1972, extended the age at which a child was allowed to leave compulsory education by one year (from 15 to 16 years) (131). The study sample comprised 21,548 participants of European descent from the UK Biobank cohort, and the refractive error of those born before and after the ROSLA reform was compared (131). The study found that spherical equivalent worsened by -0.29 D during the 1952-1972 transition period (131). The estimated causal effect of the ROSLA act on population refraction was -0.77 D in the direction of myopia (131). However, individuals with a greater genetic risk of myopia were less affected by the reform implementation than participants with lower genetic risk. If the findings of this study held true, genetically susceptible individuals would develop myopia independently of exposure to education (131). These results strengthened existing evidence from Mendelian randomisation studies, which suggested that remaining in the education resulted in a greater risk of myopia (131).

A 2018 study by Mountjoy *et al.* combined educational attainment and self-reported myopia GWAS from two independent samples to determine if more years in education were a causal risk factor for myopia (132). These analyses revealed that an additional year of education resulted in more myopic refraction of -0.27 D (132). Conversely, there was no evidence indicating that myopia could influence education(132).

1.5.2.2 The interplay between cognitive functions and myopia

Cognition, measured as a proxy for the "general mental ability for reasoning, problem-solving, and learning", is associated with myopia independently of educational attainment. One of the earliest studies conducted in 1958 by Nadell and Hirsch suggested that children with myopia had higher intelligence scores (133). Similar findings were reported by researchers from New Zealand, the USA, Israel, Check Republic, Denmark (133). Other epidemiological work demonstrated that myopic children tended to perform better academically irrespective of their intelligence, whilst

their hypermetropic peers had comparatively lower cognitive scores and worse academic performance (133). Some early investigations examining this association proposed a pleiotropic relationship between negative refraction and high cognitive function, whereby a single gene or several shared genes influence two or more unrelated traits.

The positive association between myopia and higher cognitive measurements in children has been characterised by several longitudinal studies. One such study assessed the link between nonverbal intelligence quotient (IQ) and myopia in a cohort of 1,204 Singaporean children (SCORM) aged 10 to 12 years. The study authors found that 67.9% of school children in the highest quartile for IQ had myopic refraction, while only 46% of students were myopic in the lowest quartile (134). Myopia was strongly associated with nonverbal IQ in the highest quartile compared to the lowest (OR = 2.4) (134). Nonverbal IQ score was also associated with axial length and vitreous chamber increase and a more myopic refractive error (134). The observed relationship between IQ and myopia was independent of reading or near-work activities, with cognitive scores more strongly associated with refraction than the number of books read per week (134).

A further study examining the relationship between cognitive measures and myopia in male teenagers found a strong association between myopic refraction and intelligence and years of education (135). However, high cognitive scores and education were equally important for myopia development, resulting in higher prevalence rates among more educated and higher IQ groups (135). The IQ of myopic male teenagers was on average seven points higher than that of their emmetropic peers (136), and later research revealed that this relationship was more pronounced for verbal than non-verbal components of cognitive function (137). Interestingly, the correlation with cognitive performance was weaker in adults and the elderly (138). For example, in the Gutenberg Health Study of 3819 adults aged 40 -79 years, both education and cognitive score correlated with the severity of myopia; however, only years of education significantly predicted near-sightedness (138). Similarly, Spierer *et al.* reported a

nominal association with cognition in a cohort of older women that became less significant after adjustment for education (139).

Despite intelligence being one of the most studied risk factors for myopia, the exact mechanism of the association is unknown, but several possible reasons have been suggested. One of the proposed explanations is behavioural patterns related to both traits (140). In particular, it was previously proposed that individuals who read more also perform better in standardized tests, especially those assessing verbal IQ (141). Intense and regular near-work activities such as reading were also linked to higher prevalence rates of myopia, potentially confounding the relationship with intelligence (142). Moreover, verbal intelligence tests that require linguistic skills primarily developed through reading consistently demonstrate a stronger correlation with refractive error than non-verbal tests (142). However, it can also be argued that intelligence determines reading habits; therefore, intelligent children are more likely to become myopic. Nonetheless, this notion is contradicted by research suggesting a significant relationship between myopia and spatial intelligence that doesn't rely on verbal skills or retained information. Additionally, correlation with spatial intelligence counters the hypothesis suggesting that quick and efficient reading contributes to higher intelligence scores observed in myopic individuals (143).

Another explanation for the correlation between myopia and cognitive ability is a biological mechanism linking cerebral and ocular development, which leads to higher observable intelligence scores in myopes. Some previous work suggested that IQ and refractive error could share a single pleiotropic genotype, also known as Eye Brain Gene (EBG), that brings about two different but related traits: neurocognitive development resulting in higher intelligence and overdevelopment of the eye leading to myopia (144). A study by Williams and colleagues evaluated the extent of shared genetic architecture between refractive error and IQ in a longitudinal twins' birth cohort (Twins Early Development Study) (145) and found that shared genetic effects explained 78% of the phenotypic correlation between refraction and IQ, estimated at −0.116. Polygenic risk

scores (PRS) contributed to a considerable proportion of the shared variance. In particular, in the PRS model constructed for intelligence, higher IQ scores showed a strong association with negative refraction. The Twins Early Development Study also demonstrated that shared genetic factors significantly influenced the covariance between IQ and myopia (145).

1.5.2.3 Near-work

Near-work, which includes activities performed at short working distance, may influence the association between education and myopia. In particular, increased accommodation associated with near-work activities was thought to mediate the effect of educational attainment. However, observational studies that assessed distance-weighted or time-based measures of near-work found a weak correlation between myopia and near-work, especially when this association was adjusted for other factors such as parental myopia (146,147). Conversely, it was also suggested that near work could induce accommodative lag and hyperopic defocus seen to promote ocular growth in animal models (148).

A systematic review by Huang *et al.* aimed to examine the association between near-work activities and myopia in children and quantify the effect of near-work on refractive error (149). The meta-analysis reviewed articles published between 1989 and 2014 and comprised twelve cohorts and fifteen cross-sectional studies from Asia, the US, Europe and the Middle East (149). The definition of myopia varied across the articles selected for the review, while time spent on near-work activities was self-reported by parents or students (149). In ten cross-sectional studies, near-work was associated with an increased risk of myopia in children aged 6-18 (149). In these investigations, schoolchildren reading more than two books per week were three times more likely to be myopic than their peers reading less than two books per week (84). Moreover, a reading distance of less than 30 cm and continuous reading of more than 30 minutes each day were associated respectively with 2.5 times and 1.5 greater the likelihood of

myopia (84). However, only two of the six cohort studies included reported statistically significant correlations. In particular, one investigation found that hours per week spent reading and playing video games were greater in myopes than emmetropes before and after the onset of myopia (150). Another prospective cohort study demonstrated that myopic participants carried out significantly more near-work than emmetropes, specifically in the younger cohort (151). Overall, the systematic review of the current evidence on myopia and near work suggested that children who performed more close distance work were 85% more likely to be myopic, and with each additional dioptre-hour of near-work per week, the likelihood of myopia increased by 2% (149). Only reading and not television watching or playing video games was associated with myopic refractive error (149). The authors of the review were unable to combine the results of myopia progression studies due to inconsistent definitions of the outcome (149). After the consolidation of available study findings, the evidence for myopia and near work association was of moderate strength and close-work activities were considered moderately important for myopic refractive error development (149).

1.5.2.4 Time outdoors

Few lifestyle factors with protective effects have been identified for myopia. Among potentially modifiable risk factors, the most conclusive evidence was found for exposure to outdoor light which was shown to be protective against myopia. One of the first studies investigating this association was the 2008 seminal publication reporting that more time outdoors, irrespective of physical activity, was associated with lower myopia prevalence rates in children (152). Since then, the findings of several longitudinal and randomised control trials supported the hypothesis that time outdoors decreased the risk of myopia onset (153,154). In 2013, Sherwin *et al.* conducted a meta-analysis systematically reviewing the findings of seven cross-sectional studies and reported that one additional hour spent outdoors resulted in myopia risk reduction by 13% (155). The review did not examine the efficacy of outdoor time against refractive error progression in already myopic eyes. However, subsequent studies found a significant effect of

increased time spent outdoors over the onset of myopia, but not its' progression (156). In particular, the data suggested a significant protective effect of increased time spent outdoors against incidental myopia (RR = 0.53; 95%CI: 0.338 – 0.85) (156). One hour of outdoor time lowered the chance of myopia by 45 – 50% (156) and reduced myopic shift by 0.30 D (156). However, while the increase of time outdoors was protective against myopic refraction, it did not influence disease progression in the already myopic eyes (156). Given these findings, outdoor time appears to have a stronger effect before myopia has developed, because no association was found between time spent outdoors and myopia progression (156).

There is some evidence that time outdoors could prevent myopia via multiple pathways (157). The proposed biological mechanisms include higher luminance, reduced peripheral defocus, chromatic spectrum light, its association with physical activity and altered circadian rhythms (157). The higher luminance theory gained the most support from research on human and animal studies (157). The brightness outdoors is 10-1000 times greater than indoor luminance and may explain the association between time outdoors and incidental myopia (157). In animal models, chicks and rhesus monkeys, brighter ambient lighting ranging 10 000 – 40 000 lux was observed to prevent form-deprivation myopia (FDM) in a dose-dependent manner but had a weaker effect on lens-induced myopia (157).

Longitudinal studies in humans reported the association between higher average daily light exposure and slower ocular growth (157). Interestingly, exposure to artificially increased light also lowered the risk of myopia (157). In particular, daily 30-minute light therapy sessions were shown to thicken choroid in adults, a measure associated with a reduced axial elongation in children (157). In another intervention study, the classroom brightness was increased from 100 to 500 lux, which resulted in a 6% lower risk of incidental myopia (157).

The evidence for other hypotheses explaining the relationship between time outdoors and myopia is inconclusive (157). Hypermetropic peripherical defocus was observed to cause ocular growth in animal models, while in humans, it was associated with myopia progression but not the onset (157). Nonetheless, whether the differences between peripheral defocus in outdoor and indoor environments could potentially mediate the relationship between time outdoors and myopia remains to be elucidated (157). Although other factors, such as circadian rhythms, spatial frequency and chromatic spectrum, could modulate the effect of time spent outdoors, there is a paucity of studies on humans supporting these hypotheses(157). The current research also provides no conclusive evidence backing the mediation by near-work or physical activity (157,158). Conversely, studies conducted in animals indicate that brighter light and reduced peripheral defocus exhibit a synergistic effect over eye growth, suggesting the combination of these two factors might be required for maximal efficacy of the interventions in humans (157).

1.5.2.5 Birth season

Throughout the years, an association between refractive error and the birth season has been hypothesised but remains controversial. These hypotheses hold that the season of birth affects the exposure to light in the early months or years of life, or it changes the age in which the child is first exposed to educational pressures since the school start months in the Northern hemisphere is fixed in late summer or early autumn. Some studies found an association between myopia with birth during the months of spring (159), others autumn (160,161), or the photoperiod (different months and seasons like spring and autumn have similar luminance) (160), but these findings remain to be replicated (162,163).

1.5.2.6 Urbanisation

Urbanisation is a broad term that encompasses different aspects of cultural and environmental changes to which the individuals can be exposed in an urban setting. Urbanisation is therefore a general and very non-specific risk factor.

Several observational studies linked geographical location to childhood myopia prevalence rates with higher prevalence reported in urban compared to rural areas (47,140,164). However, data on the impact of urbanisation is conflicting. It was previously suggested that air pollution may be driving higher myopia prevalence rates in urban areas (165–167), but the evidence in favour of this association was weak (167). Epidemiological work had also suggested that population density could potentially explain the observed association with urbanisation (167). In particular, schoolchildren living in the areas with higher population density were more likely to be myopic irrespective of other factors such as parental myopia and education, near-work, time spent outdoors and ethnicity (168–170), whereas the relationship between urban status and myopic refraction appeared to be confounded by education and time outdoors (171).

1.5.2.7 Parental factors associated with myopia

Several parental factors were linked to myopia in children, and these associations may be mediated through shared genetic factors or direct influences that parental choices have on the children’s environmental exposures. Parental genotypes influenced their offspring's traits even when the alleles were not transmitted (172), a phenomenon referred to as "genetic nurture" (172).

An array of maternal factors showed significant associations with the child's refractive error (173). Maternal age and height were implicated in the risk of myopia (24,173). Life course studies identified increasing maternal age and height as important pre-conceptional life stage factors for near-sightedness (174). Pregnancy in older women

was associated with a 2.3- and 2.1-fold increase in high and early-onset myopia risk, respectively (174), although this association was likely mediated by lower birth weight (24). Additionally, fertility treatments that moderately correlated with maternal age and inversely correlated with gestational age demonstrated a protective effect over myopia, decreasing the risk by almost 30% (173). Although women who underwent fertility treatments tended to have more factors predisposing to myopia, such as higher education and socioeconomic status, unexpectedly the relationship between fertility interventions and offspring myopic refraction remained significant even after accounting for potential confounders (173). The studies also found an association between maternal educational attainment and a child's refractive state. Higher maternal education correlated with myopia in the offspring (173), particularly in childhood but also persisting later in life (173). The relationship between maternal qualifications and a child's myopic refraction most likely reflected several other influences such as parental socioeconomic status, income, parenting, educational encouragement, and genetic factors shared between mother and offspring (173). Interestingly, paternal socio-economic factors, such as occupation, also appeared to influence a child's myopia risk. Non-manual paternal occupation increased the risk of overall, high-degree and early-onset myopias by 20%, 70%, and 80%, respectively, especially in early life (174).

1.5.2.8 Other factors associated with myopia.

Education, near work and reduced time outdoors, had been cited as major environmental stressors underlying pathogenesis of myopia, but several other less recognised lifestyle factors, such as diet, were also implicated in refractive error development. The hypothesis suggesting that dietary variations in protein, fat and cholesterol intake could be associated with refractive error and myopia was first introduced by Gardiner, who compared dietary habits of progressing and stable myopes and related increased intake of proteins and carbohydrates to stable myopia (175). More recent work reported rising prevalence rates of myopia in developing countries

that adopted western diet (176) characterised by excessive consumption of food with a high glycaemic load. A high glycaemic load imposed chronic hyperglycaemia which was thought to induce myopic axial elongation by disrupting the retinoid receptor pathways and triggering scleral growth (176). This mechanism has been supported by more recent evidence from animal experiments (177–180) and some observational studies (181). Notably, the findings of several genetic investigations of high myopia also implicated insulin pathways (182–184). Interestingly, hyperglycaemia and poor metabolic control, but not insulin dosage were associated with myopia in a retrospective cohort of patients with diabetes type I (185). Further cross-sectional assessments had confirmed that myopia was more prevalent than hypermetropia in diabetic populations (186), however refractive changes in patients with diabetes were attributed to increased lens thickness, rather than scleral growth. Specifically, short-term fluctuations in blood glucose were shown to affect the refraction of the lens by altering osmotic pressure and contributing to the accumulation of fructose and sorbitol in the lens material (187).

High glycaemic index diets have also been postulated to mediate the relationship between refractive error and weight. Early reports had documented a correlation between high body mass index and axial length (176,188). Obesity, in particular, was associated with 2.7 times greater odds of myopia in a cohort of Irish children with the relationship persisting even after other important factors such as sedentary lifestyle, daylight exposure and near-work were considered (189). This association with obesity was replicated in children of diverse ancestry (190,191) and partially explained by socioeconomic prevalence differences in paediatric myopia (191).

The rise in myopia prevalence rates across different regions was mirrored by a similar epidemiological trend described for cardiometabolic disease (CVD) (181). Particularly, rapid urbanisation and associated lifestyle changes were believed to contribute to the rising prevalence of abdominal obesity, metabolic syndrome and diabetes type 2 (181). It was argued that the rapid socioeconomic transition and changing environment linked the cardiometabolic disease to the gradual population myopisation through interaction

with epigenetic factors and an underlying genetic susceptibility (181). In particular, sedentary lifestyles associated with education as well as hyperglycaemic diet fuelled increasing insulin resistance, which was a pathophysiological mechanism implicated in CVD and myopia (181).

Similarly to morbidity structure, patterns of medication use also exhibited generational changes and could potentially influence the optic properties of the eye (192). The literature documented transient refractive errors induced by certain pharmacological agents (39). Furthermore, several classes of medications showed promise in arresting myopia progression (193). Evidence of varying strength was found for antimuscarinic agents, adenosine receptor antagonists and intraocular pressure-lowering drugs (193). Nonetheless, current pharmacological options are limited to atropine which slows down, but does not stop, myopia progression and has significant adverse effects (194). The investigations of systemic medication use in the general population in relation to refractive error could potentially help identify novel pharmacological candidates for myopia control and strengthen the evidence for medications that demonstrated effectiveness in animal models. Additionally, previous research illustrated the merit of examining medication-taking patterns to identify novel biological mechanisms involved in disease pathogenesis (195). Systemic medication use was shown to modulate the risk of glaucoma (196) and several useful biological insights were gained from these analyses (197), but for refractive error and myopia, the associations with commonly prescribed treatments haven't been examined. In *Chapter 7*, I explore the hypothesis that medication-taking in adulthood is related to spherical equivalent changes in the population. I also investigate whether refractive status at an earlier age could predict morbidity and medication use behaviours in adulthood.

1.6 Pathological changes and complications associated with myopia

The rising prevalence of myopia is a global public health issue, not only due to direct economic costs and substantial social burden but also because of vision-threatening ocular complications and potential blindness (198). Progressive ocular globe elongation that accompanies myopia leads to ocular pathology and changes in the fundus (198). Myopia that results in degenerative changes of the eye collectively described as "pathological", "malignant", or "degenerative" myopia (198). Pathological near-sightedness is defined as the following: "excessive axial elongation associated with myopia that leads to structural changes in the posterior segment of the eye (including posterior staphyloma, myopic maculopathy, and high myopia-associated optic neuropathy) and that can lead to vision loss" (6).

Optic disc abnormalities are an early complication of myopia that develops in childhood and adolescence and progresses with myopia degree. Tilted appearance, as well as a temporal crescent, are tell-tale characteristics of a myopic optical disc (199) and clinically important for glaucoma diagnosis (198). Correlation with a myopic shift in adolescents and refractive error and axial length in adults suggests that the disc shape changes arise from the stretching of the ocular globe and axial elongation (200). During a myopic shift, the temporal sclera recedes and flattens, which causes the optic nerve to move towards a temporal direction (200). A high prevalence of optic nerve head changes (37%), but not macular abnormalities (0.1%) in myopic teenagers, indicates a temporal pattern of complications' development. Based on this concept, optic disc tilt, chorioretinal thinning and disruption of RPE in the surrounding area develop first, followed by staphylomas and lacquer cracks and eventual formation of Fuchs' spot as well as chorioretinal atrophy in late adulthood (25).

Increased axial elongation associated with myopia leads to mechanical stretching and thinning of both choroid and pigment epithelium with accompanying degeneration and

vascular changes (198). The most common chorioretinal abnormalities found in patients with myopia include lacquer cracks and retinal breaks, Fuch's spot, choroidal neovascularisation, posterior staphyloma and chorioretinal atrophy (198).

Posterior staphylomas describe a bulging of the sclera in the posterior segment of the eye (fundus) with a curvature radius that is smaller than the radius of the adjoining eyewall (201). Staphylomas are the most characteristic finding of pathological myopia (202) and are associated with thinning and elongation of the local sclera which leads to deformity of the fundus. Posterior staphylomas are among other major causes contributing to the development and progression of myopic maculopathy also termed myopic macular degeneration (MMD) (203).

Lacquer's cracks, Fuch's spot and chorioretinal atrophy are lesions that form in the posterior segment of the eyeball and represent different stages of myopic maculopathy (204). Myopic maculopathy starts with the development of tessellated fundus characterised by thinning of retinal pigment epithelium which makes choroidal vessels visible (204). The tessellated fundus is a typical early sign of myopia and can be observed in the eyes of highly myopic children (205). The tessellation of the fundus is followed by the development of diffused chorioretinal atrophy and lacquer cracks indicating progressive thinning of the choroid and retina (204). Diffused chorioretinal atrophy appears as a faint yellowish lesion around the optic disc that spreads with age and axial elongation, eventually covering an entire area within the staphyloma (204). The disproportionate thinning of the choroid compared to surrounding tissues is a key feature of the diffused chorioretinal atrophy, whereby the RPE layer and outer retina are not destroyed (204). By contrast, lacquer cracks affect deeper layers of the eyewall and represent healed mechanical breaks in the Bruch’s membrane, that divides choroid and RPE (204). Lacquer cracks, seen as yellowish thick linear patterns in the macula, can progress to other lesions, including patchy chorioretinal atrophy (204). This type of chorioretinal atrophy is recognised by greyish-white and well-defined lesions (204). Patchy chorioretinal atrophy leads to complete obliteration of RPE, outer retina and

most of the choroid, making sclera visible through transparent retinal tissue (204). Lacquer cracks and patchy chorioretinal atrophy predispose to choroidal neovascularisation (CNV) (206). CNV describes the growth of new blood vessels originating from choroid that extend to pigment epithelium or subretinal space through the breaks in the Bruch's membrane (207). CNV is a major cause of blindness and the most common sight-threatening complication of severe myopia (208–210). Fuch's spot is a pigmented lesion that forms during the scarring phase of CNV.

Staphylomas and macular lesions are not common among children (204,211), as both increasing age and axial length affect their development and progression (204). The most common fundus findings in adult patients with high myopia include posterior staphylomas and chorioretinal atrophy (212,213). Among myopic adults aged 40 or older, staphylomas and chorioretinal atrophy were observed in 23% and 19.3% of eyes, respectively (212). Additionally, for each increase in myopic refractive error, the risk of staphyloma and chorioretinal atrophy subsequently increased by 48% and 56%, accordingly (212). Axial length was also associated with more frequent occurrence of staphylomas, chorioretinal atrophy, and lacquer cracks – for each additional millimetre of AL, the likelihood of developing these complications increased 2.08, 3.13 and 2.03 times, respectively (212).

Excessive axial elongation, increasing age and posterior staphylomas are the main risk factors predisposing to myopic macular degeneration (maculopathy), which is a significant cause of vision impairment. According to the 2000 epidemiological data, the global estimated prevalence of MMD-related visual impairment and blindness was 0.07% and 0.02%, with 4.2 million and 1.3 million people affected, respectively (214). However, MMD complication rates were dependent on the prevalence of high myopia. For example, in Europe where the prevalence of severe myopia was lower than in the East-Asia, macular degeneration contributed to approximately 23% of blindness participants over 75 years old (215), while in Taiwan myopic degeneration accounted for 25% of all visual impairment cases and was the second common cause of vision loss

(216). However, the impact of myopic degeneration is expected to increase, and the prevalence of MMD-related visual impairment is projected to grow eightfold by 2050 (214), because of decreasing spherical error and growing average axial length, increasing myopia prevalence in the population, and longer lifespan (214).

While the rising myopia prevalence is a relatively new phenomenon, seen in more recent generations, the impact of complications may only become apparent in the upcoming decades, as high myopes age and are more likely to develop MMD and other ocular disorders (214). Moreover, undercorrected myopia is still an issue due to undiagnosed myopia and eye care inaccessibility in some parts of the world and is one of the primary causes of vision impairment (198).

1.7 Refractive error correction methods

The correction of refractive error is achieved by changing the trajectory of visible light as it enters the eye. Several correction methods are available for myopes today, including spectacles, contact lenses and techniques that reshape corneas such as photorefractive keratectomy, laser in situ keratomileuses (LASIK) and laser-assisted subepithelial keratectomy (214). Spectacles are the most common myopia correction method, but they are not always practical or cosmetically acceptable (214). Conversely, contact lenses are more invasive and have a higher risk of intolerances and complications (214). Refractive surgery is the most invasive method and may have different drawbacks, depending on the technique (214).

1.8 Health disparities and potential for myopia-related loss of vision

Different myopia correction methods are available and accessible in Europe and the US. However, in emerging economies in South and East Asia, where the highest prevalence of myopia is recorded, a vision impairment due to uncorrected refractive error is still an issue (217). Factors such as reduced health expenditure and resource scarcity contribute

to low vision correction rates in patients with myopia and can indirectly increase the risk of vision impairment (VI) and pathological complications (217). Spectacle coverage describes a percentage of individuals who had their vision corrected to normal levels, among those who require spectacles and can be used to assess the availability of vision correction methods (217). It was previously estimated that distance-vision spectacle coverage varied between 2% and 93% across 27 countries, where the data on vision correction was available (217). In countries with low spectacle coverage, high rates of vision impairment and MMD-induced blindness were attributable to concordant effects of high myopia prevalence and age distribution as well as limited resources (214).

Vision impairment associated with either MMD or uncorrected myopia can significantly affect both quality of life and productivity (217). Few economic evaluations of the financial impact of uncorrected myopia exist to date but considering the increasing number of working-age individuals with myopic macular degeneration, a substantial burden on communities can be anticipated in the near future (217). Previous research demonstrated that uncorrected refractive error contributed to the global productivity loss of 202 billion USD that was also adjusted for country-specific employment rates as well as labour force participation (218). These estimates, however, included not only myopia but also hypermetropia and astigmatism and were based on refractive data from 2007 (217). However, more recent evaluations published in 2018 suggest that impaired vision from uncorrected myopia resulted in a global productivity loss at 244 billion USD, whereby MMD-related VI caused a 6 billion USD productivity loss (217). Moreover, the greatest potential burden due to reduced productivity, as a proportion of economic activity, was identified in East, South and Southeast Asia (217).

Uncorrected refractive disorders are one of the leading causes of VI and secondary blindness. The findings of the Global Burden of Disease Study suggest that in 2010 uncorrected refractive errors contributed to 101.2 million and 6.8 million cases of moderate/severe VI and blindness, respectively (219). Likewise, they also negatively influenced productivity, educational attainment and quality of life (218,220). There are

considerable disparities in the prevalence of uncorrected refractive disorders. In particular, the estimated prevalence rates of uncorrected refractive error were 10%, 15%, 21% in the Australian populations of European (221), Hispanic (222) and Chinese (223,224) ancestries, respectively. In addition, according to data from the 2013 Global Burden of Disease, uncorrected refractive errors contributed to 11.3 million disability-adjusted life years (DALYs) (225), which predominantly affected female patients (226,227).

Factors such as limited income and lack of financial decision-making authority, limited access to eye care and services, shortage of ophthalmologists impede refractive correction in less developed economies and lead to significant regional and gender disparities (226). Tackling these differences may help reduce the prevalence of uncorrected refractive disorders and subsequently prevent avoidable blindness.

Chapter 2 | Aims of Thesis

Refractive errors, and especially myopia, are among the leading causes of visual impairment and blindness worldwide. The prevalence of myopia has increased significantly over the past three decades, particularly in East Asia, but is also on the rise in Europe, Australia and the United States of America (214). Myopia is caused by progressive axial elongation that leads to irreversible changes in the sclera and retina and can cause blindness. Currently, there are no effective treatments against myopia onset and pharmacological options slowing disease progression are limited (228), and although research had implicated several lifestyle factors potentially influencing the development of refractive errors (14), their pathophysiology is not fully understood. Refractive error is a complex trait and to prevent or delay its' onset or to develop effective therapies, a better understanding of genetic and environmental factors is required.

Refractive errors, and myopia, develop as a result of a complex interplay between an individual's genetic liability and environmental exposures. Twin studies demonstrated high heritability of refractive disorders ranging between 60% and 90% (71). Previously, more than 140 distinct loci were reported in association with the refractive error in the population. However, known susceptibility variants predicted less than 8% of the variance (110). More loci remained to be associated with refractive error, which would help build better myopia prediction models and better characterise molecular mechanisms leading to refractive disorders. Additionally, elucidation of the age-specific effects of these genes could potentially help understand which genotypes lead to early-onset myopia. This information could be utilised to detect children at high risk before they develop refractive errors and use prevention strategies to inhibit myopia onset.

Although larger SNP-array based GWAS have narrowed the gap between trait heritability and variance explained by common SNPs, they will be unable to adequately assess the contribution of rare variants. Previously, rare genetic associations were examined using family linkage design and candidate gene investigations of familial forms of severe myopia. However, the results obtained from these studies were rarely replicated and generally did

not overlap with genetic factors influencing the population's refraction (100). The release of extensively phenotyped exome sequencing projects such as UK Biobank provides an opportunity to study the impact of rare markers on complex traits, such as refractive error.

Accumulating evidence suggests that the pathogenesis of refractive errors is intertwined with other chronic health conditions (181,229). The investigations of morbidity structure and patterns of medication use could potentially unveil biological mechanisms shared between refractive disorders and other diseases and pathologies.

This thesis will examine genetic and environmental factors influencing the development of refractive errors and myopia primarily using data from the UK Biobank cohort with additional collaborative analyses with a personal genomics company *23andMe* and Consortium for Refractive Error and Myopia (CREAM).

Aims:

1. To use genetic association to elucidate the biological mechanisms underlying refractive development and assess the utility of the polygenic risk prediction models to predict myopia in populations of diverse ethnicity.
2. To identify genetic risk factors that are associated with the age of refractive error onset, using the age of first spectacle wear.
3. To evaluate associations between systemic medication-taking and refractive error.

The work presented in this thesis incorporates publications. The methodology and statistical analyses that were used in studies included in the thesis are described in *Chapter 3*. In *Chapter 4*, I present the findings of the largest to date genome-wide meta-analysis of refractive error that analysed genetic data from more than half of a million participants of European descent. I report on the utility of identified loci for myopia prediction and discuss the evidence of implicated molecular mechanisms. In *Chapter 5,* I explore the influence of rare genetic variance on the population's spherical equivalent and shed a light on novel genetic associations, not documented in prior GWAS. *Chapter 6* includes the results of

genome-wide time-to-event analyses of the age of first spectacle wear which yielded several new genetic associations with age-at-onset of refractive disorders and showed that effects of many myopia genes varied with age. In *Chapter 7*, I examine the relationship between medication-taking behaviours and refractive error and using these insights uncover novel mechanisms involved in the pathogenesis of refractive disorders. *Chapter 8* draws together the findings of the investigations that contributed to this thesis and discusses the general implications of their results.

Chapter 3 | Methods

3.1 Introduction

This chapter will introduce cohorts and the methodology used throughout this thesis. The descriptions provided here will be general, with more contextual detail briefly given in the different chapters on the different research projects that are part of this thesis.

3.2 Cohorts' description

3.2.1 UK Biobank cohort

The UK Biobank is a prospective cohort study that includes 502,682 individuals, all United Kingdom residents, aged 40 and 69 years at recruitment (230). The participants were recruited between 2006–2010 via the UK National Health Service based on their living proximity to the 22 assessment centres (230). The volunteers provided extensive phenotypic and genotypic data. At recruitment, study subjects signed an electronic consent form, filled in an electronic questionnaire that included inquiries about socio-demographic, lifestyle and health-related factors and participated in physical assessments (230). The baseline physical assessments comprised blood pressure, handgrip strength, anthropometry, spirometry and heel bone density measures (230). The participants were also interviewed by the trained nurses who collected additional phenotypic data on lifestyle and environmental exposures (230). Blood, urine and saliva samples were collected and stored to allow the use in different assays such as proteomic, genetic and metabonomic analyses (230). After the recruitment stage was completed, additional assessments were introduced, including ophthalmologic examinations (230).

3.2.1.1 Refractive error assessment

A subsample of 117,279 (23%) volunteers from the UK Biobank cohort participated in a comprehensive ophthalmic examination that included the assessment of refractive error (231). Individuals that reported a current eye infection or underwent ocular

surgery in the preceding four weeks were not eligible for examination (231). The measurement of non-cycloplegic autorefraction was performed by Tomey RC 5000 auto refractometer (Tomey Corp., Nagoya, Japan) (231). The assessments were carried out on the right eye first, with up to 10 tests for each eye (231). The most representative results were selected based on autorefractor reliability score and used to calculate the mean spherical equivalent (sphere + ½ cylinder power) for each study participant (UK Biobank data field numbers: 5084-5085; 5086-5087) (231).

Following previously published methods, some of the study participants were removed from subsequent analyses (231). Individuals were excluded if they reported cataracts with mild myopia, LASIK, glaucoma, vitrectomy or any other eye surgery or bilateral and unilateral eye injury that resulted in vision loss. However, the refractive error measurements in the intact eye were still included (231).

3.2.1.2 Demographic factors

Education

The information on participants' education was collected using touch-screen questionnaires (Field number: 6138). The study subjects were asked to specify their qualifications by picking one of the six suggested response categories: "College/University", "A levels/AS levels or equivalent", "O levels/GCSEs or equivalent", "CSEs or equivalent", "NVQ or HND or HNC or equivalent", "Other professional qualifications e.g.: nursing, teaching". The years of education were quantified from self-reported data on professional qualifications following previously published methods (232). The response categories were transformed according to the International Standard Classification of Education (ISCED): none of the above (no qualifications) = 7 years of education; CSEs or equivalent= 10 years; O levels/GCSEs or equivalent = 10 years; A levels/AS levels or equivalent = 13 years; professional qualifications = 15 years; NVQ or HNC or equivalent = 19 years; college or university degree = 20 years of education.

3.2.1.3 Townsend deprivation index

The Townsend deprivation index measures a level of material deprivation in the geographical area, based on four census variables: unemployment, car and house non-ownership and household overcrowding (233). The Townsend deprivation index negatively correlates with social deprivation, with the higher scores representing higher levels of destitution.

The Townsend deprivation index was administered to the UK Biobank participants immediately before the enrolment (Field number: 189) (230). The deprivation scores were calculated based on the preceding national census output areas, and each subject was assigned a score corresponding to their postcode area (234).

3.2.1.4 Medical history and medication use

The information on medication use was collected via electronic questionnaires and face-to-face interviews with a trained nurse (Data field: 20003) (195). The data on treatments dosages and duration were not recorded (195). First, the participants were asked if they were taking any medications and then prompted to select the matching names from the questionnaire list (195). The use of over-the-counter drugs, vitamins and supplements was only documented in the touchscreen questionnaire, and this information was not mentioned again unless the participant forgot to answer the question about medication-taking. The interviewer, a trained nurse, recorded over the counter and prescription treatments by entering the trade and generic names into the Search facility and selecting from the list of possible matches. The nurse recorded only regular prescription medications defined as treatments taken weekly, monthly or every three months.

The matching of available generic and trade names to the active ingredients was performed by searching the electronic medicine compendium, available at https://www.medicines.org.uk/emc/. Only pharmaceutical drugs reported by at least 10 participants of European descent were selected for the subsequent analyses. The

analyses did not include topical treatments or herbal and homoeopathic medicines and were limited to eye drops, inhalers and medications taken systematically. When the treatments contained more than one ingredient, each one of them was added to the count of medications with the same specific ingredient. The polypharmacy was described as a self-reported intake of 5 more treatments at the time of the baseline assessment.

The medical history was ascertained from hospitalisation episode statistics, recorded according to the tree of International Classification of Diseases, Tenth Revision (ICD-10) codes (231) (Field number: 41270). The dataset comprised 2,779,598 records and 7,719,358 diagnoses with at least 1 record available for 395,978 participants (231). The ICD-10 list was compiled by the World Health Organisation and followed a hierarchical structure, covering 19,155 clinical terms (231). Each hospitalisation episode in the dataset was annotated with a primary diagnosis associated with an event and if necessary one or several secondary diagnoses (231). Disease outcomes were treated as binary traits and generated by combining primary and secondary diagnostic annotations (231). By default, study participants were considered to be unaffected by a given diagnostic term (ICD10 code) unless there was a reported diagnosis in any of the different fields (231).

3.2.1.5 Array data

Genotyping was performed on 488,377 UK Biobank cohort participants (230). The extracted DNA was genotyped using two similar and compatible platforms – Applied Biosystems UK BiLEVE Axiom Array capturing 807,411 genetic markers and Applied Biosystems UK Biobank Axiom Array covering 825,927 genetic markers. Approximately 95% of genetic markers were shared between the two platforms. Applied Biosystems UK Biobank Axiom Array specifically included genome-wide genetic variation as well as deletions and short insertions. The markers were selected based on their clinical relevance or potential role in disease pathogenesis. Additionally, the array covered a

wide range of indels, including minor allele frequencies and rare variants, markers with good genome-wide imputation coverage of common (MAF > 5%) and rare (MAF 1-5%) variants in European populations.

The genotyping of extracted DNA was performed centrally (230). The blood samples collected from participants during the visit to the assessment centre were processed in 106 sequential batches of approximately 4,700 samples each. A custom genotype calling pipeline and quality control filtering procedures were developed to handle the novel genotyping arrays as well as genotyping experiments. As a result of custom genotype calling 812,428 unique markers were called, including indels and biallelic SNPs and further filtering procedures were applied to improve their quality (230).

The quality control pipeline was developed to account for genotyping issues that might have been potentially caused by the use of slightly different arrays, batch effects as well as diverse ethnic background of the study participants. Approaches incorporating principal component analysis were used to control for population structure in both sample and marker quality control (230).

Genetic markers of poor quality were identified by running statistical tests checking for significant differences in experimental factors, including batches and arrays (230). As a result of the quality control procedures, 0.97% of calls were set to missing. The quality of samples was assessed based on the rate missingness and metrics of heterozygosity adjusted for population structure, which resulted in 0.2% of all samples being marked as potentially problematic. The discrepancies between reported and biological sex ascertained from X and Y chromosomes marker intensity were also checked. Based on measured X and Y chromosomes intensities, 652 genotyped individuals presented with sex chromosomes karyotypes that differed from XX or XY combinations. After genotype calling and quality control procedures, the final dataset covered a total of 805,426 markers and included 488,377 samples (230).

The phasing of autosomes was preceded by haplotype assessment for the whole cohort and a subsequent haploid imputation (230). In the pre-phasing step, genetic markers that did not pass quality control in more than one batch and had missingness greater than 5% or estimated allele frequency below 0.0001 were removed. Samples that were identified as outliers in relation to the population-adjusted heterozygosity and excessive missingness were also excluded. Consequently, 670,739 markers and 487,442 samples were used in autosomal phasing. The phasing of autosomes was carried out by the SHAPEIT3 program that was modified to process large datasets. The 1000 genomes reference panel was selected to phase the sample of European and non-European ancestry. Moreover, a new divisive clustering algorithm was developed to detect haplotype clusters and calculate Hamming distances between the haplotype pair within each cluster. Consequently, Hidden Markov Model only included haplotypes in each cluster to estimate surrogate family copying states (230).

Imputation was carried out using Haplotype Reference Consortium (HRC) data as a primary reference panel, but also merged 1000 Genomes phase 3 and UK10K reference panels. HRC data contained more than 39 million SNPs from broadly European haplotypes, whereas a merged panel had 87,696,888 bi-allelic markers (230). Only markers that were shared between HRC and 1000 genomes/UK10K datasets were selected for imputation; therefore, a final dataset covered 93,095,623 autosomal SNPs in conjunction with large structural variants and indels in a total of 487,442 study participants. Additionally, 3,963,705 genetic markers on the X chromosome were also imputed (230).

3.2.1.6 Whole-exome sequencing data

The whole-exome sequencing data were, at the time of writing, available for 200,643 UK Biobank participants (235). Individuals who had more complete data, such as baseline measurements, magnetic resonance imaging, hospital episodes/primary care records, were prioritised for the sequencing. The whole–exome study sample was

representative of the general UK Biobank cohort for sex, age and ancestry, but was enriched for hospital episode statistics codes and physical measures, such as eye and hearing assessments, electrocardiograms (235).

Exome sequencing was performed on IDT x Gen Exome Research Panel v1.0 modified to include supplemental probes (236). The panel targeted 39Mb of the human genome and covered 19,396 genes on autosomal and sex chromosomes (236). The targeted regions included 38,997,831 bases and 10M variants, with coverage surpassing 20X at 95.6% of sites. The targeted regions contained 8,086,176 SNPs and 370,958 indels, 1,596,984 multi-allelic variants. Approximately 84% of 8M sequenced SNPs were coding variants comprising 4,549,694 (53.8%) missense, 2,139,318 (25.3%) synonymous and 453,733 (5.4%) predicted loss-of-function variants (LOF) that altered at least one coding transcript. Almost 99% of all synonymous, missense and nonsynonymous markers had minor allele frequency below 1%. The median number of missenses, synonymous and loss of function variants per individual was 7,724, 8,586 and 702, respectively. Additionally, 18,045 genes with heterozygous LOF variants and 1492 genes with homozygous LOF markers were observed.

3.2.2 23andMe

Study participants

Participants from *23andMe*, a personal genomics company, contributed extensive self-reported data on lifestyle factors, family background and disease history (29,99). The study subjects signed informed consent and filled in electronic questionnaires following the approved protocol that was reviewed and received approval from Ethical & Independent Review Services (http://www.eandireview.com). Myopia cases were identified based on self-reported refractive status (29,99). Controls were participants that did not report myopia.

Genotyping

Genotyping was done on one of the four platforms (29,99). The V1 and V2 contained 560,000 genetic markers from the Illumina HumanHap550+ BeadChip of which 25,000 were customary to *23andMe*. The V3 platform included 950,000 SNPs from Illumina OmniExpress+ BeadChip that improved overlap with V2 (29,99). The V4 platform was a fully custom array covering 570,000 variants, constituting of low redundancy markers from V1 and V2 and lower frequency coding SNPs. Samples with a call rate below 98.5% were reanalysed. The GWAS was restricted to the sub-sample of the study participants of European descent as ascertained through local ancestry analysis. The ethnic categorisation was carried out by partitioning phased genomic information into windows of 100 variants and matching individual haplotypes to one of the 30 reference populations. The classifications were then entered into a hidden Markov model that controlled for assignment and switch errors and calculated probabilities for each population at each partition. The population reference data was obtained from publicly available databases, specifically 1000 Genomes, Human Diversity Project and HapMap, and the self-reported family history of four grandparents from the same country. A segmental identity-by-descent (IBD) estimation algorithm was used to identify a subset of unrelated subjects. The relatedness was defined as sharing more than 700 cM IBD (26,233). In male participants, autosomal non-pseudoautosomal regions were treated as pseudo-autosomal diploids (29,99).

3.2.3 GERA

Study Participants

The Genetic Epidemiology Research in Adult Health and Aging (GERA) cohort included 110,266 participants, all members of the Kaiser Permanente Research Program on Genes, Environment and Health (RPGEH) (237). This program recruited patients from Kaiser Permanente Northern California integrated health care delivery system and collected longitudinal health records, including eye examinations. The study sample comprised 34,998 non-Hispanic adults of European ancestry aged 35 and older who

participated in spherical equivalent assessments conducted between 2008 and 2014 (237). Refractive error was measured as mean spherical equivalent, using a standard formula equal to the sphere + ½ cylinder (294). Subjects with a history of cataract and refractive surgery as well as diagnosis of keratitis or corneal diseases were not advanced for the subsequent analyses (237). The study protocols received approval from the Institutional Review Board of the Kaiser Foundation Research Institute (237).

Genotyping

The DNA extraction from Oragene kits (DNA Genotek Inc., Ottawa, ON, Canada) was performed at Kaiser Permanente Northern California facilities while genotyping was handled by the Genomics Core Facility of the University of California, San Francisco (UCSF) (99). The genotyping panel based on Affymetrix Axiom arrays (Affymetrix, Santa Clara, CA, USA) covered 665,000 SNPs (99). The observed initial genotyping call rate exceeded 97%, while allele frequency discrepancies between males and females were less than 15%. SNPs with call rate < 90% and allele frequency < 1% were dropped, resulting in 98% of assayed genetic markers and 97% of samples passing QC (99). The QC procedures were followed by prephasing in SHAPEIT2 (238), and subsequent statistical imputation of additional variants from the multi-ethnic 1000 Genomes Project reference panel as implemented in IMPUTE2 v2.3.0 (239). Based on IMPUTE2 info r2, measuring the correlation between true and imputed genotype (239), markers with r2 <0.3 and minor allele count < 20 were removed.

3.2.4 CREAM Consortium

Study participants

The Consortium for Refractive Error and Myopia (CREAM) study sample comprised 41,793 adults of European descent aged 25 or older. Several measurements of refraction and spherical equivalent were used across consortium cohorts (110). Individuals with ocular conditions that could potentially alter the measurement of refractive error, such as LASIK and cataract surgery, vitrectomy and keratoconus or other systemic eye

diseases, were removed from the analyses (110). The studies conducted in the UK were not part of the meta-analysis described in this thesis due to sample overlap. Written informed consent was obtained from all CREAM participants and the studies received approval from relevant local review boards and medical ethics committees (110). Each cohort applied appropriate recruitment strategies and methods of ascertainment that varied across the studies (110). More detailed information about phenotyping methods is given in Table 1

Genotyping and imputation

Consortium cohorts were genotyped on one of the two platforms – the Illumina or Affymetrix (Table 1) (110). The genotypes were imputed against the 1000 Genomes Project (Phase I version 3, March 2012 release) ethnically matched reference panel with IMPUTE (240) or minimac (241).

Samples with discrepancies between clinical and genotyped sex and evidence of non-European ancestry were dropped (110). Pre-imputation QC criteria varied across the cohorts, but genotype call thresholds were set to values above 95% (110)

Exome chip genotyping

The European CREAM cohorts were genotyped separately using either the Illumina HumanExome-12 v 1.0 or v 1.1 array but were jointly recalled to improve calling accuracy for genetic markers with a minor allele frequency (MAF) ≤ 0.01. Recalling genotypes simultaneously across all study samples allowed to call rare variants with a more discrete distinction between allele calls and sensitivity for low-frequency loci. The data recalling was performed in GenomeStudio® v2011.1 (Illumina Inc., San Diego, CA, USA). Nine, predominantly European studies and one Singapore Indian investigation (SINDI) were aggregated into a single cohort (N = 11,505).

Cohort Name	Phenotyping method	Genotyping chip
ALIENOR	Autorefractor Luneau SPEEDY K	Illumina Human 610-Quad BeadChip
ANZRAG	Refractive details were obtained from clinical notes	Illumina Omni 1M/Illumina Omni Express
AREDS	Subjective Refraction	Illumina HumanOmni2.5-4v1_H array
BMES	Autorefractor Zeiss Humphrey-530	Illumina 670 Quad Custom Chip
BATS	Autorefractor Zeiss Humphrey-598	Illumina HumanHap 610-Quad array
Croatia-Korcula	Auto Ref/Keratometry NIDEK ARK30	Illumina 370CNV-Quad v1 BeadChip
Croatia-Split	Auto Ref/Keratometry NIDEK ARK30	Illumina 370CNV-Quad v3 BeadChip and OMNI
Croatia-Vis	Auto Ref/Keratometry NIDEK ARK30	Illumina HumanHap300 v1.0 BeadChip
DCCT	Subjective Refraction	Illumina Human-1M BeadChip
EGCUT	Eyeglasses prescriptions	Illumina Omni Express
ERF	Topcon RM-A2000 autorefractor	Illumina 6K, 318K, 350K, 610K, Affymetrix 250K
FECD	Subjective Refraction	HumanOmni2.5-Quad BeadChip
FITSA	TOPCON AT	Illumina HumanCoreExome
Framingham	Subjective Refraction	Affymetrix Mapping500K (Nsp & Sty) + 50K HumanHap supplement
Gutenberg Health Study 1	Zeiss Humphrey® Automated Refractor/Keratometer (HARK) 599™	Affymetrix Genome-Wide Human SNP Array 6.0
Gutenberg Health Study 2		
KORA	Nikon Retinomax and Eyeglasses prescriptions	Illumina Omni 2.5/Illumina Omni Express
OGP Talana	Autorefractor Topcon RK-8100	Affymetrix GeneChip Human Mapping 500K
Rotterdam Study I	Topcon RM-A2000 autorefractor	HumanHap 610-Quad BeadChip
Rotterdam Study II		Illumina HumanHap 550-Duo v3 BeadChip
Rotterdam Study III		Illumina HumanHap 610-Quad BeadChip
TEST	Autorefractor Zeiss Humphrey-598	Illumina HumanHap 610-Quad BeadChip
WESDR	WESDR	Illumina Omni1-Quad BeadChip
YFS	NIDEK AR-310AR autorefractor	Illumina 670K Custom Array

Table 1 Phenotyping and genotyping methods in CREAM non-UK European cohorts. Data extracted from Tedja et al (110).

3.2.5 TwinsUK

The TwinsUK registry was established in 1993 and includes 14,010 twins who are predominantly older age women of European descent recruited throughout the UK (242). The volunteers were recruited using successive media campaigns without prioritising specific diseases or traits. At the time of recruitment, participants were not aware of the future eye studies and provided informed consent under a protocol approved by a local research ethics committee (EC04/015). Twins were seen at the clinical facilities of the Department of Twin Research and Genetic Epidemiology, King's College London, where blood and urine samples were collected, and comprehensive phenotyping assessments of twins were carried out. The zygosity was ascertained using a standardised questionnaire completed by twins. When there was uncertainty about zygosity status, it was confirmed by genotyping.

Over 6,000 twin subjects participated in the ophthalmic examination, including the assessment of refractive error (243). Non-cycloplegic autorefraction (ARM-10; Takagi Seiko, Japan), was calculated as a mean spherical equivalent, taking the average reading of both eyes. If only one eye was available due to surgery or inability to perform the test in another eye, then the reading from the good eye was used to calculate spherical equivalent. The exclusion criteria were self-reported cataract, refractive or any other eye surgery and any eye pathologies that could alter refractive error (243).

Whole-genome sequencing

A subset of 2,377 participants was selected for whole-genome sequencing performed on an Illumina HiSeqX sequencer, utilising a 150-base paired-end single-index read format (244). The reads mapping was done against the hg38 reference panel, and the variants were called using ISIS Analysis Software (v. 2.5.26.13; Illumina). Sequencing data coverage > 30× was achieved for 2,377 samples (245). Genomes with a high ratio of heterozygous to homozygous variant genotypes were removed. Additionally,

genomic analyses were restricted to individuals of European ancestry which was verified with ADMIXTURE (246). All genomes were required to be predominantly European (70%). The prediction of family relationships, including parent-offspring trios and twin pairs, was carried out in RelateAdmix (247). Samples with inconsistencies between self-reported and predicted family relations were dropped, resulting in the final set of 1,960 subjects of whom 383 and 522 were DZ and MZ twins, respectively (244). Multiallelic variants were converted to biallelic variants, whereas indels were left-aligned. Missing genotypes were imputed with reference homozygous calls. A high-confidence genome region was determined by identifying positions with a call rate greater than 90% by using three sets of randomly sampled 100 genomes. Duplicated variants and variants outside a high-confidence genome region were removed (244).

3.2.6 BDES

Study participants

The Beaver Dam Study (BDES) – an epidemiological investigation of ocular diseases that included adult residents of Beaver Dam, Wisconsin (248). At the baseline, examination participants completed a personal history questionnaire and participated in an eye exam, including the assessment of refractive error (249). Non-cycloplegic refraction was measured with the Humphrey 530 refractor (Humphrey Instruments Inc., San Leandro, CA) and for each study, subject mean spherical equivalent was calculated by taking the average of the right and left eyes refractive error (249). Following quality control procedures, individuals with significant spherical error differences between left and right eyes as well as one participant from each first and second-degree relative pair and subjects with missing phenotype and covariate data were excluded, resulting in the final sample size of 874 participants (249).

Exome chip genotyping

A subsample of 2,032 from BDES was genotyped using Illumina Infinium HumanExome-12 v1.1 BeadChip (Illumina, Inc., San Diego, CA) exome array (249). A total of 1,908 individuals were successfully genotyped with a call rate > 95%, however, following quality control procedures, some of them were subsequently removed due to unresolved sex discrepancies, Mendelian errors as well as duplicated samples (249). Moreover, 12 additional individuals were excluded because of non-European ancestry as ascertained from principal component analysis (249).

Genetic variants were filtered if they were nonautosomal, exhibited poor call rates (< 98%) or showed evidence of deviation from Hardy–Weinberg Equilibrium (p-value < 1.0×10^{-6}); monomorphic, indel and duplicated markers were also removed, which resulted in 1,871 subjects and 95,376 genetic markers being advanced for imputation against HRC reference panel (249). The imputation was performed using SHAPEIT version 2 and

Minimac version 3 (238,250–252). A total of 1,024,985 predominantly exonic markers passed the imputation threshold of R2 threshold ≥ 0.6 and were selected for association analyses (249).

3.2.7 EPIC

Study participants

The European Prospective Investigation into Cancer study is a pan-European project investigating the course of several chronic diseases (253,254). EPIC-Norfolk, one of the EPIC UK studies, recruited and conducted baseline assessments on 25,639 volunteers between 1993 and 1997 (253,254). The participants were residents of Norwich and surrounding areas selected via the National Health Service registry. Ophthalmic examinations, which were part of third visits, formed EPIC-Norfolk Eye Study, conducted between 2004 and 2011 (253,254). The study sample comprised 8,623 individuals for whom spherical refraction was available (253,254). Autorefraction was performed by a Humphrey Auto-Refractor 500 (Humphrey Instruments, San Leandro, California, USA). All study subjects gave written informed consent (253,254). The study received ethics approval from East Norfolk & Waveney NHS Research Governance Committee (2005EC07L) and the Norfolk Local Research Ethics Committee (05/Q0101/191) (253,254).

Genotyping and imputation

A total of 7,117 participants were genotyped based on the Affymetrix UK Biobank Axiom Array (253,254). SNPs exclusion criteria were observable batch effects across genotyping plates, significant deviation from Hardy-Weinberg equilibrium (p-value< 10-07), low call rates and poor clustering. Samples with poor genotyping rates, sex inconsistencies, abnormal heterozygosity or identity-by-descent values, as well as third-degree or closer relatives, were dropped (253,254). The data pre-phasing was done in

SHAPEIT (version 2) (251), followed by imputation against the Phase 3 build of the 1000 Genomes reference (255).

3.2.8 SCORM

Study participants

Singapore Cohort Study of the Risk Factors of Myopia (SCORM) is a prospective longitudinal cohort of Singaporean children aged 7-9 at recruitment, who were recruited from three schools between 1999 and 2001 (256). The study received ethical approval from Ethics Committee, Singapore Eye Research Institute and followed the tenants of Helsinki (256). The written informed consent was obtained from students' parents. The schools were sampled based on the National examination results to ensure the inclusion of children with different overall academic performances (256). Two schools located in the north-eastern and south-eastern regions of Singapore joined the study in 1999, and in 2001 the cohort was expanded to include an additional school located in the western part of the city (256). All 2186 eligible children were invited to participate in the study; of them, 81 were excluded due to serious illnesses, or ocular disorders such as syndromic myopia and congenital cataracts or allergy to eye drops, resulting in the final sample of 1979 individuals (256).

Annual eye examinations were performed on schoolchildren during a 7-year follow-up period with the final visit in 2007 (256). The assessments of cycloplegic refraction were conducted at each visit. Pupil dilation was induced by the topical administration of 3 eye drops of cyclopentolate hydrochloride at 5-minute intervals (256). Cycloplegic autorefraction measurements (model RK5; Canon, Inc., Ltd., Japan) were obtained at least 30 minutes after the last drop instillation (256). Spherical equivalent was calculated using a standard formula equal to sphere plus half of the cylinder (256). The study participants were categorised into groups of moderate (−3.00 D to > −5.00 D) and high myopes (≤ −5.00 D) based on their refractive status (256). The assessments of axial length were performed using contact ultrasound biometry (probe frequency of 10 Mhz;

Echoscan model US-800; Nidek Co., Ltd., Tokyo, Japan) following instillation of 1 drop of 0.5% proparacaine (256).

Additional data on lifestyle factors were reported by students' parents who completed an eight-page questionnaire two weeks before the baseline assessments (257). The parents were asked to specify the number of books their children read per week and if parents wore spectacles/lenses to correct near-sightedness (257). Time spent outdoors was self-reported by study subjects who completed the Sydney Myopia Study questionnaire. Time outdoors was defined as the sum of leisure and sports activities and was documented separately for school weekdays and weekends (258).

Genotyping

A total of 1004 children were genotyped using the Illumina HumanHap 550 or 550 Duo Beadarrays with more than half of genotyped subjects (57.4%) participating in the final follow-up examination (259). Following quality control procedures rare SNPs (MAF< 0.05), and variants that demonstrated high missingness (call rates < 95%) and violated Hardy-Weinberg equilibrium (p-value (HWE p-value < 10^{-6}) were dropped from the analyses (260). Additionally, variants that exhibited substantial allele coding and frequency deviations (> 0.20) from East Asia (EAS) reference population in the 1000 Genomes reference panel were also excluded (260). The ancestry was ascertained by combining participants' genotypes with data from the 1000 Genomes panel which was based on 2504 individuals from 26 populations (260). Multidimensional scaling analysis was conducted on the sample of 3508 participants and included 568,974 HapMap3 SNPs that passed quality control filters (MAF < 0.05; call rate < 0.01 and HWE p-value < 10^{-6}) (260). The SNPs were pre-phased with SHAPEIT version 2.837 based on family trio information. The imputation was performed against the 1000 Genomes reference panel using the "PBWT, no pre-phasing" pipeline (260). Non-monomorphic, biallelic imputed variants with MAF > 0.05, INFO > 0.5 and HWE p-value > 10^{-6} were selected for the inclusion (260).

3.2.9 STARS

Study participants

Strabismus, Amblyopia and Refractive Error Study (STARS) is a population-based survey of ocular diseases in Chinese families with children aged 6 and 12 months (53). Eligible participants were recruited using household registry records obtained from the Ministry of Home Affairs (53). The selection of the families residing in the Southwest region of Singapore was performed using a disproportionate stratified random sampling method based on the six-month age increments (53). Following the exclusion criteria, enrolees with chronic medical conditions or those who had not lived at the registered address for at least five months or changed the living address were excluded from the study (53). The invitations to participate in the survey were sent by mail and were followed by door-to-door visits from trained interviewers administering comprehensive questionnaires (53). The questionnaires contained inquiries about demographic and lifestyle factors, medical and ocular history of the family (53). Of 4162 children selected from the household registry and invited to participate, 3009 (72.3%) were enrolled in the study (53). Eye examinations were conducted between May 2006 and November 2008 in two speciality eye hospitals: Singapore National Eye Centre (SNEC) and the Jurong Medical Centre hospitals (53).

Comprehensive eye tests were performed by trained ophthalmologists, orthoptists or optometrists and included among others the assessment of refractive error (53). The measurement of cycloplegic objective refraction was carried out based on the participants' ability to complete ocular examination (53). Refractive error in subjects aged 24 to 72 months was measured using an autorefractor (Autorefractor RK-F1; Canon, Tokyo, Japan), while for participants aged 12-24 months autorefraction was performed using a hand-held device (Retinomax K-Plus 2; Nikon, Tokyo, Japan) (53). Streak retinoscopy was performed for children younger than 12 months or in instances of autorefractor test failure (53).

Genotyping

The genotyping of DNA samples was performed using Illumina Human610 Quad Beadchips (Illumina Inc., San Diego, USA)(261). The samples with call-rates < 95%, evidence of excessive heterozygosity, cryptic relatedness or deviation in population structure were removed (261). Quality controls filters for genetic variants included MAF < 1%, per-SNP call rate < 95% and Hardy–Weinberg equilibrium < 10^{-6}. To ascertain the population structure the principal component analysis was carried out by the EIGENSTRAT program (261). The genotypes that passed the quality control procedures were advanced for imputation against 1000 genomes phase 1 cosmopolitan panel haplotypes (261).

3.2.10 GUSTO

Study participants

Growing Up in Singapore Towards Healthy Outcomes (GUSTO) cohort comprised offspring of ethnically diverse mothers aged 18 and above who attended the first trimester dating ultrasound scan clinic at National University Hospital (NUH) and KK Women's and Children's Hospital (KKH) from June 2009 to September 2010 (262). All recruited women were citizens or permanent residents of Chinese, Indian and Malay ancestry and homogenous ethnic background (262). The mothers were included in the study sample if they delivered at NUH or KKH and lived in Singapore in the next five years (262). The exclusion criteria were chemotherapy, psychotropic treatments or diagnosis of diabetes type I. Additionally, only pregnant women who agreed to donate birth tissues were enrolled in the study (262). Of screened 3751 families 2034 met the study requirements (262). The study was approved by the National Health Group's Domain Specific Review Board and SingHealth Centralised Institutional Review Board. The participants provided written informed consent (262).

The offspring cohort included 1176 children born 30 November 2009 and 1 May 2011 (262). The participants were assessed through infancy and childhood (262). Cycloplegic autorefraction was first performed on study subjects at age 36 months and continued in follow-up visits (262). The information on myopia risk factors, specifically time spent outdoors, near-work, maternal education and parental myopia, was collected using questionnaires completed by the parents (262). However, the study analyses described in this thesis were restricted to data from the year nine examinations conducted on Chinese children (262).

Genotyping

The participants were genotyped based on the Infinium OmniExpressExome array (260). Following quality control procedures SNPs with MAF < 0.05, missingness > 5% and HWE p-value < 10^{-6} were removed from the analyses (260). Additionally, variants that demonstrated different allele coding or frequency that deviated more than >0.2 from the EAS reference population in the 1000 Genome reference panel were also dropped (260). A total of 529,083 SNPs were advanced for subsequent analyses and pre-phased with SHAPEIT version 2.837 including trio family information (260). The SNPs were imputed to the 1000 Genome reference panel using the "PBWT, no pre-phasing" pipeline in Sanger Imputation Service (260). Non-monomorphic (MAF >0) biallelic variants with MAF > 0.05, HWE p-value > 10^{-6} and INFO > 0.5 were selected for the inclusion (260).

3.3 Statistical analyses

3.3.1 Linear regression

Linear regression was a standard method to examine genetic associations with continuous traits in GWAS (263). The standard approach for testing genetic associations was to include genotype coded as a dosage of the less frequent allele. Most studies utilised an additive model, fitting linear regression for each locus using the number of

minor alleles at the given SNP as a covariate while adjusting for other factors such as age and sex (Formula 1). Therefore, SNP beta coefficients denoted the degree of change in phenotype per each copy of the minor allele. Following convention, a Bonferroni-correction of 1 million independent linkage-disequilibrium blocks, leading to a significance threshold of 5×10^{-8} (264) was commonly applied to all genome-wide analyses.

$$\mathbf{y} = \mathbf{G}\beta_G + \mathbf{X}\beta_X + \mathbf{e}$$

Formula 1 Linear regression model. *y* represents phenotype, G is the genotype with genetic effect β_G, X corresponds to the fixed-covariate matrix with covariate effects β_X, and e is the residual error (265).

Throughout this thesis, I performed the statistical analyses of quantitative traits using linear regression. For each SNP at the locus, linear regressions testing associations with phenotypes (such as the mean spherical equivalent, used to characterise refractive error) were fitted for each study participant, also adjusting for age, sex and 10 principal components. The effects of SNPs were modelled using additive dosages.

3.3.2 Linear mixed regression models

3.3.2.1 Bayesian mixed models

False-positive findings in GWAS could be partially attributed to spurious associations arising from populations structure and cryptic relatedness among participants of the study cohort (266). Linear mixed models (LMMS) had been increasingly used for the analyses of binary and continuous traits in GWAS because they were able to robustly control for population stratification and admixture through the incorporation of random effects for the genetic relationship matrix (266). In the LMM-based methods, populations structure was included as fixed effects, while genetic kinship was modelled as the variance-covariance structure and used as a random effect (Formula 2) (266).

$$\boldsymbol{y} = \boldsymbol{x}_{snp}\beta_{snp} + \boldsymbol{X}_c\boldsymbol{\beta}_c + \boldsymbol{g} + \boldsymbol{e}$$

Formula 2. Mixed linear model. *y* represents phenotype. χ_{snp} is the variant genotype with its effect β_{snp} , while X_c are fixed covariates with their corresponding coefficients β_c. *g* is a relatedness matrix and *e* represents residual error (267).

In this thesis, I used LMMs to assess genetic associations with refractive error. Models were included mean spherical equivalent as an outcome and SNPs at the locus as independent predictors while simultaneously adjusting for age, sex and 10 principal components as implemented in BOLT-LMM (268). This program implements an algorithm for fast Bayesian mixed-model association which allowed to correct for populations structure and relatedness in very large samples with reduced computational costs and better efficiency (268). Using this approach each chromosome was left out in turn and Bayesian models were fitted on the remaining SNPs, while left-out variants were later tested applying a retrospective hypothesis test (268). This method enables association analyses for millions of variants (268).

3.3.2.2 Mixed-effect models

To assess the associations between refractive error and pharmaceutically active ingredients, I utilised linear mixed-effects models (LMEs). Linear mixed-effects models were an extension of simple linear models that incorporated both fixed effects which were parameters associated with the entire population and random effects which measured trait variance within non-independent groups (Formula 3) (269).

$$\begin{aligned} y_{ij} \quad = \quad & \beta_1 x_{1ij} + \cdots + \beta_p x_{pij} \\ & + b_{i1} z_{1ij} + \cdots + b_{iq} z_{qij} + \varepsilon_{ij} \end{aligned}$$

Formula 3 y_{ij} represents the value of the outcome variable for the *j*th of n_i observations in the *i*th of M groups. Fixed-effect coefficients are shown as $\beta_1, \ldots, \beta p$ with their regressors as $x1_{ij}, \ldots, x_{pij}$ for observation *j* in group *i*. $b_{i1}, \ldots, b_{iq}$ correspond to the random-effect coefficients for group *i*, while $z_{1ij}, \ldots, z_{qij}$ represent the random-effect regressors (270).

The models were built using *lme4* and *lmertest* packages in R. Refractive error followed leptokurtic distribution but based on previously published recommendations (263),

untransformed mean spherical equivalent as an outcome and independent medications were entered as independent predictors in the model. Moreover, the models included recruitment age and age squared, sex, years of education, Townsend deprivation index, and polypharmacy as additional covariates. Following previously published methods (271), participants' birth decade was used as a random effect grouping factor in LMEs methods to account for potential non-linear generational effects (58). Mixed-effects modelling allowed to control for the autocorrelation of observations within birth cohorts. The model coefficients for each medication quantified the difference in spherical equivalent between treatment users and non-users. The associations were considered statistically significant if their probabilities were below Bonferroni multiple testing threshold, calculated by diving desired alpha level (α = 0.05) by the number of analysed medications.

3.3.3 Gene-based analyses

Population-based investigations of the impact of rare genetic variation on complex traits required methods different from traditional GWAS approaches, specifically because statistical power to find a significant association with rare variants dropped as minor allele frequency decreased in the population (272). Thus, in the discovery analyses, the variants were typically grouped together (273). The most common approach was to analyse the cumulative effects of rare markers in the genomic region, usually at the gene level (274). The gene-based analyses were broadly classified into the burden and non-burden tests (274). Burden methods aggregated rare variants into a single burden variable and then regressed the trait on the burden variable to assess the cumulative effect of grouped polymorphisms (274). Notably, burden tests were performed under an implicit assumption that all variants at the locus were causal and had the same direction and similar magnitude of the effects, thus this approach led to the loss of power when assumptions were violated (275,276). By contrast, Kernell-based methods, such as sequence kernel association tests (SKAT), instead of aggregating markers, combined individual variant-score statistics for each SNP in the SNP-set (277). SKAT was

powered to detect associations in genomic regions containing both deleterious and protective mutations or a large fraction of noncausal variants (274). Conversely, SKAT-O, an optimised version of the SKAT test, performed well in both scenarios (274). The SKAT-O test combined burden and Kernel-based association methods by calculating a weighted average of burden and SKAT statistics with weights estimated from the data to maximize the power (274). In WES where some genomic regions included a high number of causal variants with the same direction of effects while others harboured causal markers with opposing effects or noncausal variants, SKAT-O was able to detect associations missed by burden and SKAT tests (274).

I applied SKAT-O test (274) in gene-based analyses of refractive error presented in this thesis (278). The associations with genes were considered statistically significant if their probabilities were below the Bonferroni multiple testing correction adjustment level, calculated by dividing the desired alpha level ($\alpha = 0.05$) by the number of tested genes. Based on previously published methods (279), the analyses incorporated heterogeneous variant annotations that demonstrated an increased power to detect causal associations in gene-wide association studies. The analyses were restricted to the synonymous and protein-altering genetic variants such as nonsynonymous, start gain/loss, stop gain/loss and splice mutations with allele frequency below 1%. UTR variants as well as markers with unknown or inconclusive molecular functions such as intronic polymorphisms were not selected for inclusion. The variant annotation was performed in ANNO statistical package that defined splicing sites as 3 bases into exon or 5 bases into intron with mutations in these regions characterised as "Normal_Splice_Site". However, the mutations that affected the functionally important intron "GU...AG" region were annotated as "Essential_Splice_Site".

3.3.4 Cox proportional-hazards regression model

The time-to-event analysis referred to a set of methods that analysed a length of time until the occurrence of the event of interest (280). The Cox-proportional hazards regression model was a commonly used time-to-event analysis technique that allowed to simultaneously assess multiple covariates (280). Cox regression modelled hazard function was performed under the assumption that all study subjects had a common time-dependent baseline hazard function (280). The ratio of the hazard rates for different individuals in the study was expected to be constant over time, specifically that the effect of the covariates did not vary with time and was equal at all time points (280). Following the proportionality of hazard assumption, the exponentiated model coefficients for each covariate described HR for every unit change of the respective covariate value (Formula 4) (280). Thus, HR above 1 represented an increase in the hazard of the event, while HR<1 indicated the opposite.

$$\lambda(t; X_i, G_i) = \lambda_0(t)\exp\left(X_i^T \beta + G_i \gamma\right)$$

Formula 4. Cox proportional hazards regression model. The model specifies the hazard function for the time to event T_i associated with genotype G_i and vector of covariates X_i. $\lambda_0(t)$ is a baseline hazard function, β corresponds to the effect of covariates and Υ represents genetic effect (281)

I assessed genetic associations with the rate of onset of refractive correction in the Cox proportional-hazards model adjusted for the age and sex of the study participants. The statistical probabilities for each SNP were calculated using the likelihood ratio test. The hazard ratios and p-values were respectively computed in *gwasurvivr* (282) and *SPACox* (281) R packages. The hazard ratios (HR) showed a multiplicative change in the rate of onset of refractive correction per copy of the tested allele. The genetic associations were considered statistically significant if their probabilities were below the conventional genome-wide significance level. Additionally, the assumption of hazards proportionality was tested in the *survival* R package (244) (https://cran.r-project.org/web/packages/survival/).

3.3.5 LD-score regression

Both polygenicity, as well as SNP ascertainment bias arising from population stratification and admixture, could potentially inflate the distribution of statistics in GWAS (283). Linkage disequilibrium score regression quantified the contribution of genuine polygenic effects and confounding bias by inspecting the relationship between GWAS test statistics and linkage disequilibrium (283). In particular, genetic markers in linkage disequilibrium with causal variants demonstrated inflation of the statistics proportional to the collinearity with causal SNPs (283). The more phenotype variation was explained by the index variant, the higher was the probability that it tagged causal SNP (283). In contrast to genuine polygenic effects, variation due to cryptic relatedness and population stratification did not correlate with LD. Thus, genomic inflation from SNP ascertainment bias was quantified by regressing GWAS $\chi 2$ statistics against LD score and subtracting the regression intercept from 1 (283). Values of regression intercept that were close to 1 indicated a small contribution of confounding bias from populations stratification and admixture and that the increase in mean $\chi 2$ was driven by polygenicity (283). However, a moderate degree of inflation in the LD-score regression intercept was expected in the studies with very large sample sizes (283).

In addition to assessing genomic inflation, I used LD-score regression to examine bivariate genetic correlations between refractive error and potentially related traits whose summary data were obtained from the publicly available GWAS repository *LD Hub* (284). *LD Hub* automated cross-trait LD score regression analysis pipeline which was an extension of simple linkage disequilibrium score regression and only required summary-level statistics (285).

3.3.6 Conditional analyses

I further refined genetic associations obtained from the meta-GWAS in conditional and joint analyses as implemented in the COJO module of the GCTA software (286). Approximate joint and conditional analyses ascertained independent genetic signals

from meta-analysis summary statistics by inferring linkage structure from the reference sample of unrelated UK Biobank participants of European descent. The genetic locus was defined as a genomic region containing adjacent associated SNPs within a 1 Mbp distance from each other. Statistical significance and collinearity thresholds were set to p-value = 5×10^{-08} and $r^2 = 0.9$, respectively. The phenotypic variance explained by all SNPs was estimated by converting marginal effects to joint effects using meta-GWAS summary information and individual-level genotype data from the reference sample.

3.3.7 Gene-set enrichment analyses

Gene-set enrichment analysis was an analytical approach used in genome-wide association studies to uncover relevant biological mechanisms implicated by genetic associations (287). In contrast to traditional GWAS, which assessed the significance of the associations between phenotype and individual SNPs, gene-set enrichment analyses evaluated the evidence of the association between phenotype and set of genes grouped based on common function, making pathways a unit of analysis (287). These analyses examined the combined effects of a large number of disrupted genes that affect specific cell types or biological functions (287). However, the power of these analyses depended on the accurate detection of the genes contributing to the specific pathophysiology of the studied trait or disease (287).

Many pathway-based analyses implemented modified algorithms originally developed to study the pathways using gene expression data in which differentially expressed genes were annotated and grouped based on the same annotation terms (288,289). The annotation terms were obtained from GSEA (*www.gsea-msigdb.org*) which hosted molecular signature collections such as gene-sets from Gene Ontology (290) and KEGG pathway databases (291).

In this thesis, I explored the shared functionality of associated genes through gene-set enrichment analyses, as implemented in MAGENTA software (287). The publicly available gene-set were tested for the association with refractive error and the age of

first refractive correction. By default, MAGENTA employed two thresholds of gene-level significance at "95%" and "75%". The pathway-level p-values were computed by comparing the number of observed significant genes to the number of expected significant genes within the pathway. The threshold of 75% was used in this thesis to define statistical significance because it captured weaker signals of the association and was suited for highly polygenic traits (287).

3.3.8 Mendelian randomisation analysis

Mendelian randomisation (MR) utilised genetic markers to ascertain whether the observed correlation between the exposure and the outcome is consistent with the causal effect (292). MR method was based on a concept of a random assortment of parental alleles during meiosis resulting in their random distribution in the population (292). Individuals were randomly and naturally assigned at birth to either inherit polymorphisms predisposing to higher levels of exposure or not inherit these variants (292). Because these susceptibility markers were not typically related to potential confounders, the differences in the disease risk among the carriers and non-carriers of these exposure-associated alleles could be attributed to the differences in the exposure levels (292).

Mendelian randomisation relied on three assumptions: "1) the genetic variant is associated with the exposure, 2) but not confounders; and 3) the genetic variant influences the outcome only through the exposure" (292). The second and the third conditions were related to pleiotropy which refers to a genetic marker affecting the outcome through pathways distinct from the exposure variable (292). The first assumption was tested by assessing the strength of the association between SNPs, referred to as instrumental variables, and the exposure (292). By contrast, the second and third conditions could not be empirically tested and required further sensitivity analyses (292).

Increasingly, MR analyses were performed using summary statistics obtained from GWAS available in the public domain. Several MR methods had been developed including inverse-variance weighted (IVW), weighted median and Egger tests (293). MR investigations selected lead SNP within associated genomic regions as instrumental variables for their analyses, thus the reliability of these studies depended on the validity of instrumental variables (293). The estimate ratio for the causal effect of exposure over the outcome is calculated by dividing the estimate of the SNP association with the outcome by the estimate of the same SNP association with the exposure (294). The IVW worked under the assumption that all selected markers were valid instrumental variables and aggregated the ratio estimates for all SNPs, weighting their linear combination by the inverse of their variances (Formula 5) (295).

$$\hat{\beta}_{IVW} = \frac{\sum_k X_k Y_k \sigma_{Yk}^{-2}}{\sum_k X_k^2 \sigma_{Yk}^{-2}}$$

Formula 5 Inverse-variance weighted test. The pooled ratio estimate of the causal effect of exposure *X* on trait *Y* using genetic markers k shown here as Y_k/X_k. The standard error of the ratio estimate is represented by term $\sigma_{Y\,k}/X_k$ (295).

The weighted median test allowed the majority of markers to be valid and took the median of the ratio estimates weighted by their variance (296). Conversely, the MR-Egger method assumed that plurality of SNPs was valid as instrumental variables (297). Egger although similar to IVW was adapted to test bias arising from pleiotropy with the slope from the regression representing the estimate of the causal effect (Formula 6) (297).

$$\hat{\beta}_{Yj} = \theta_{0E} + \theta_{1E}\hat{\beta}_{Xj} + \epsilon_{Ej}$$

Formula 6 MR-Egger test. $\hat{\beta}_{Y\,j}$ is genetic association between the trait and variant *j*. $\hat{\beta}_{X\,j}$ is genetic association between exposure and variant *j*. The parameters ϑ_{0E} and ϑ_{1E} represent the intercept and the slope, respectively, and ϵ_{Ej} is the residual term (298).

In particular, the intercept of the MR-Egger regression that significantly deviated from zero indicated a presence of the directional pleiotropy (297). This approach was used to evaluate the robustness of MR findings (297) and conduct sensitivity analyses complementary to the IVW test which yielded more precise estimates (294).

3.3.9 Summary data-based Mendelian randomisation analysis

Summary-data based Mendelian randomisation method (SMR) analyses integrated summary statistics from large published GWAS and expression quantitative trait loci (eQTL) studies and detect genes whose expression levels are associated with the trait due to pleiotropic effects (299). Additionally, SMR was used to conduct a transcriptome-wide association study and to test associations between gene expression and complex traits (299). The estimates of SNP effects on trait variance and gene expression were obtained from large-scale GWAS and eQTL studies, respectively (299). Only probes with eQTL p-value < 5×10^{-08} were selected for inclusion and the associations between the trait and the probe were tested at the top *cis*-eQTL (299).

Nonetheless, significant associations in the SMR test did not necessarily imply that the variant was causal because the top *cis*-eQTLs could be in the linkage with two independent causal markers one of which affected gene expression while another influenced trait variance (299). Heterogeneity independent instruments (HEIDI) test implemented in SMR examined multiple SNPs in the *cis*-eQTL region and differentiated pleiotropic signals from the linkage (299). Based on this test, for SNPs demonstrating p-value > 0.05, the hypothesis of a single causal variant associated with both gene expression and trait variance could not be rejected (299). However, because HEIDI was unable to perfectly separate pleiotropy from linkage disequilibrium in cases of extreme collinearity, the results of SMR analyses should be interpreted under the model of pleiotropy and not causality (299). In this thesis, I used SMR to analyse the expression of refractive error and the age of refractive correction genes incorporating eQTL data

from peripheral blood samples (300) and summary statistics from methylation and eQTL analyses in brain tissues (301) and retinal samples (302).

Chapter 4 | Common genetic variation influencing population's refractive error

4.1 Introduction

Refractive errors, but especially myopia, are of the leading causes of preventable blindness and vision impairment worldwide, especially in the communities without access to effective methods of refractive correction (303). Refractive disorders are predominantly driven by complex cultural and lifestyle factors in genetically predisposed individuals. Given the sharp rise in global prevalence rates of myopia in the past four decades (59,304) and an array of behavioural (305) and other exposures (174) linked to near-sightedness, there is growing evidence suggesting that refractive disorders develop largely due to environmental influence. Although it appears that environmental exposures affect the fine balance between components of refraction, which is necessary to achieve emmentropisation, the ocular growth is largely under genetic control and hereditary factors determine the severity of refractive errors, specifically myopia (126). Moreover, observational reports suggest that parental myopia is predictive of early-onset myopic refraction in children (61), which in conjunction with high heritability estimates from twin studies (76), indicates that genetic factors dictate a large proportion of refractive error variance.

To date, a host of genomic tools were exploited to study refractive disorders in populations as well as families. At least 25 myopia loci were discovered via linkage studies, and numerous genes within identified linkage intervals were examined (101). These genetic regions were targeted based on prior information about their functionality or genetic linkage evidence (14). However, candidate gene studies have demonstrated poor replication, likely due to several factors such as genotypic and phenotypic heterogeneity and methodological issues that contributed to reduced statistical power (306).

In contrast to candidate-region investigations, genome-wide association studies used a hypothesis-free approach which was shown to be more suitable for genetic interrogation of complex polygenic traits such as refractive error. However, although,

previous refractive error GWAS reported over a hundred associated genomic regions, these loci explained less than 8% of the trait phenotypic variance (110) and provided an incomplete picture of molecular mechanisms involved in the pathogenesis of refractive disorders.

This thesis chapter describes the results of the largest to date genome-wide refractive error meta-analysis which was published in *Nature Genetics* journal under the title "**Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia**". The meta-GWAS analysed genetic data and quantitative spherical equivalent measurements and myopia cases from several large international cohorts including UK Biobank, *23andMe* and the Genetic Epidemiology Research on Adult Health and Aging (GERA). The subsequent replication and meta-analysis were conducted using data from the CREAM Consortium. Several research groups had contributed to this meta-analysis and provided GWAS summary statistics of their respective cohorts.

My role included conducting post-GWAS analyses and assessing the impact of identified genetic loci on myopia risk as well as giving a functional characterisation of the associated variants to discern affected biological pathways and tissues. I performed conditional analyses that aimed to identify polymorphisms independently associated with refractive error. Subsequently, I used independent SNPs to predict myopia risk and to calculate the amount of phenotypic variance explained by selected variants. The utility of the PRS model for myopia prediction in the East-Asian adolescents was assessed in an independent sample of Singaporean children and the results of these analyses were published in the journal of *Translational vision science and technology* under the title **''New Polygenic Risk Score to Predict High Myopia in Singapore Chinese Children"**. Additionally, I assessed the relationship between genotypes and gene expression using summary-data Mendelian Randomisation tests. I also evaluated the causality of the association between refractive error and intelligence in Mendelian Randomisation analyses and assessed correlation with other traits using LD-score

regression A detailed description of the meta-analysis methods, statistical analysis, summary of the results and implications of the findings will be provided in this chapter.

4.2 Statistical analyses

Meta-analysis

The two-stage meta-analysis aggregated refractive error GWAS results from 542,934 subjects of European descent (99). The discovery analyses incorporated data from the first and the second sub-samples of UK Biobank participants, the GERA and *23andMe* cohorts. The final meta-analysis included information from all available cohorts – UK Biobank, GERA, *23andMe* and CREAM consortium studies (99). The effect sizes and p-values were calculated from the weighted sum of individual statistics using METAL statistical software (307). Weights were computed based on the squire root of the effective population sample size. No genomic control correction was applied in these meta-analyses (99). The effective population size was quantified for each genetic locus and was equal to the total sample size for investigations that used linear or mixed linear models (99). The sample size was calculated based on the formula: 2/ (1/Number of cases + 1/Number of controls) for case-control studies, as described elsewhere (308). Consequently, due to two case-control investigations included in the metanalysis, the estimated effective sample was 379,227. Genetic variants with minor allele frequency > 1%, present in at least 70% of meta-analysis participants and not missing in any of the cohorts, were advanced for the meta-analysis (99).

Gene annotation

The genomic locus was defined as one or multiple jointly associated polymorphisms located within 1 MB distance from each other (99). The most strongly associated variants in the region were annotated with the nearest transcript coding gene symbol (99). Likewise, the genomic loci were annotated with the gene whose transcript coding region containing or was adjacent to the lead variant (99).

LD-score regression

Genomic inflation was estimated using two programs – R package gap (https://cran.r-project.org) and LD-score regression (283), distinguishing between effects from polygenetic and inflation arising from uncontrolled admixture and population structure. Additionally, LD-score regression was used to assess bivariate genetic correlations between refractive error and selected traits whose summary statistics were obtained from the publicly available GWAS repository (284). LD-score regression methods are described in *Chapter 3* paragraph 3.3.5.

The relationships between refractive error loci and other traits were further explored using two datasets – OMIM (Online Mendelian Inheritance in Man), an online compendium of genes and Mendelian disorders, and GWAS Catalog, a curated database of published associations between SNPs or genes with phenotypes (99).

Conditional analyses

To detect SNPs independently associated with spherical equivalent within identified loci, I utilised the meta-analysis summary results in conditional and joint analyses as implemented in GCTA COJO (286). The total phenotypic variance explained by conditional SNPs was also estimated in GCTA. Statistical significance and collinearity thresholds were set to p-value = 5×10^{-08} and r2 = 0.9, respectively. The methodology of conditional analyses is presented in more detail in *Chapter 3* paragraph 3.3.6

Mendelian Randomisation analyses

Mendelian randomisation analyses (MR), investigating causality between spherical equivalent and intraocular pressure, were performed in the R *MendelianRandomisation* package, as described in *Chapter 3* paragraph 3.3.8. I additionally evaluated the causal effect of mental capacity over refractive error. The MR instrumental variables (IV) were independent genetic variants associated with IOP and intelligence and were selected from previously published GWAS (309,310).

Gene expression analysis

The myopia-relevant tissues and cell types were selected based on the analyses integrating expression data and summary statistics from all five meta-analysis cohorts (99). The expression data were obtained from several sources: GTEx release v7 (https://gtexportal.org/home/datasets), transcriptome study of adult and foetal trabecular meshwork, cornea and ciliary body (311) and retinal transcriptome and eQTL data from healthy donors (302). The patterns of transcript expression across different tissues were compared by visualising hierarchical clustering by means as implemented in R *hclust*.

SMR analyses tested whether the causal effect of the variant over the trait of interest was mediated by a specific gene. GCTA SMR utilised summary statistics from GWAS and eQTL investigations to predict complex traits targeted by the gene (299). I analysed the expression of refractive error genes using eQTL data from peripheral blood samples (300) and summary statistics from methylation and eQTL analyses in brain tissues (301) and retinal tissues from healthy donors (302).

Gene-set enrichment analyses

I investigated the shared functionality of refractive error genes in gene ontology (GO) enrichment analyses performed in MAGENTA software as described in *Chapter 3* paragraph 3.3.7 (312).

Prediction models

To evaluate the predictive value of identified loci for myopia, I trained and tested the regression models in two independent samples (99). In particular, the training of the models was carried out in the subsample of UK Biobank participants for whom direct measurements of mean spherical measurements were available, while the testing was performed in the EPIC-Norfolk cohort not included in the meta-analysis (99).

Participants' age, sex and lead variants associated with mean spherical equivalent were used as predictors, while the outcome in the models were different definitions of myopia based on sliding refractive error thresholds (99). The thresholds equal to ≤ −0.75 D, ≤−3.00 D and ≤ -5.00 D were applied to define "any myopia", "mild myopia" and "high myopia", respectively (99). The more stringent threshold of -6.00 D wasn't used for severe myopia because it would significantly reduce the number of cases in the testing set. The controls were participants with refractive error ≥−0.5 D (99).

Receiver operating characteristic (ROC) curves were drawn for each myopia definition to contrast the sensitivity and specificity across the range of values for their predictive ability. The logistic regression and predictive analyses were performed in R *glm* and *ROCR* packages (99).

Validation of myopia prediction models in a cohort of Singaporean children

Polygenic Risk Scores

The selected variants with MAF ≥ 1% in Chinese children were retained for the inclusion if they were in high linkage disequilibrium (LD) with refractive-error-associated SNPs or were themselves directly associated with refraction in European cohorts. Conversely, polymorphisms that were not strongly associated with refractive error in European GWAS were not included. Observed redundant genetic signals in the study participants of Chinese descent were amended with LD-pruning using PLINK (265). The calculation of minor allele frequencies and LD between pairs of genetic markers was carried out in the STARS cohort that did not take part in the testing of the prediction models.

The meta-analysis results were presented in terms of z-scores, due to the input summary statistics being on different units and scales (i.e., spherical error effects in dioptre, the risk to myopia in Odds Ratios, etc.). In order to convert Z-scores from the discovery meta-analysis to betas the following formula was used:

$$\beta = Zscore / (N.eff * p * (1-p))$$

where Zscore is Z-score from the original refractive error GWAS meta-analysis, N.eff is the effective population size and p is a minor allele frequency at the locus.

A total number of 655 SNPs were taken forward and included in PRS models that calculated the risk scores for each SCORM participant. The polygenic risk scores represented the sum of the risk alleles weighted by effect sizes for the same SNPs. PRS scores were calculated using PRSice version 2.3.1. statistical programme. To ascertain whether the inclusion of additional variants in PRS models would improve prediction accuracy, the study exploited recently developed genome-wide Bayesian methods – SbayesS, SbayesR (313) and SbayesRS. These particular approaches combined the GWAS summary statistics with LD reference panel (313) and were shown to improve prediction accuracy relative to traditionally used methods such as LS-pred, clumping and thresholding (314).

Prediction models

The associations between spherical equivalent (SPHE) and axial length (AL) measurements, taken at the last SCORM visit, and PRS were assessed in the multivariate regression model. To decipher the impact of other factors, a model containing age, sex, school and maternal education and 10 genotype principal components (PCs) as independent predictors was built PCs were calculated in GCTA (286) using the genetic relationship matrix based on 5,285,015 imputed SNPs in SCORM. The principal components measured the subtle genetic correlation between individuals and were added to the model to control for cryptic relatedness and populations stratification. After that PRS was combined with the basic model and incremental gain in the total explained trait variance (R2) was determined. Full multivariate model in addition to basic model covariates included near-work and time outdoors. The strength of the association was evaluated based on the effect sizes, standard errors, 95% confidence intervals and R2 statistics. The replication of the SCORM results was sought in the GUSTO cohort.

The associations between high/moderate and PRS were tested in logistic regression models. Moderate and high myopias were defined as spherical equivalent equal to or worse than -3.00 D and -5.00 D, respectively, and included as binary outcomes in their respective logistic models. Participants with refractive error > -0.5 at the last assessment were treated as a non-myopic group.

Time-to-event analyses assessed the rate of onset of high or severe myopia throughout eight consecutive yearly visits. The analyses defined the time of onset as the visit number at which participants developed the disease. Subjects that did not become myopic throughout the study were censored. The associations between PRS and the progression to high or moderate myopia was tested in the multivariate Cox proportional hazards model.

The performance of predictive models was assessed by area under the curve (AUC) which measured discriminative ability to distinguish between high or moderate myopes and non-myopes (260). Predictors of myopia risk were progressively added to compare the discriminative ability of different models. The first model included age, time outdoors and parental myopia; in the second model PRS was analysed with the first model covariates; the third model only had PRS as a covariate, and the fourth tested parental myopia. To determine whether models had significantly different AUC DeLong's test was conducted using the *pROC* package in R.

Lastly, the odds of high or moderate myopia were compared between individuals in the PRS top 25^{th} and bottom 75^{th} percentiles, as well as between top and bottom 50^{th} percentiles, using the $\chi 2$ test. All analyses were performed in R version 3.6.0.

4.3 Results

4.3.1 Meta-analysis

Although refractive error meta-GWAS was not the result of my own work, I am reporting it because is a necessary introduction to my own follow-up post-GWAS analyses. The

discovery meta-analysis aggregated results from GWAS of directly measured spherical equivalent from UK Biobank (N=102,117) and GERA (N=34,998) and genetic analyses of self-reported myopia from *23andMe* (N= 191,843) (99). As anticipated the large sample size of the discovery GWAS resulted in a nominally large genomic inflation factor ($\lambda = 1.94$). The intercept and the ratio (intercept-1)/ (mean (X2) − 1) were equal to 0.17 and 0.097, respectively, which was consistent with expectations of polygenicity (99).

The 438 distinct genomic loci were identified as associated with refractive error at the genome-wide significance level. Among these associations, 308 genomic regions were not reported by previous studies. Interestingly, fourteen novel signals were mapped to chromosome X. Moreover, the direction and the magnitude of the effects were consistent across all metanalysis cohorts. The statistically strongest association was found with rs12193446 (p-value= 9.87×10^{-328}) located within *LAMA2*, whose mutations caused muscular dystrophy (315). *LAMA2* also demonstrated a strong association refractive error in several prior GWAS (31,108). SNPs located within other *LAMA2* receptors encoding genes, specifically *ITGA7* (rs17117860; p-value= 8.57×10^{-9}), *DAG1* (rs111327216, p-value= 1.67×10^{-8}), whose mutations resulted in muscular dystrophy (316,317), significantly correlated with refraction in the discovery meta-analysis (99).

The results of the discovery meta-analysis were compared to the CREAM consortium GWAS findings (99). The replication was restricted to non-UK European CREAM consortium subjects to prevent potential sample overlap between UK Biobank and CREAM UK arms (99). Despite marked power differences, 55 lead SNPs from discovery meta-analysis were associated below the level of Bonferroni significance in the replication study (99). The association probabilities for the other 142 variants were below the threshold of false discovery rate, while additional 192 polymorphisms showed nominal associations in CREAM (99). There was a strong correlation between effect sizes observed in discovery and replication analyses. A subsequent meta-analysis, incorporating data from all five cohorts, including CREAM, increased the number of associated loci to 449; of them, 336 regions were not previously reported (99).

4.3.2 Post-GWAS analyses

4.3.2.1 Functional analyses

Functional analyses of associated independent SNPs revealed 367 frameshift or missense variants (99). Several of the associated loci contained non-synonymous mutations, for example, rs1048661 within *LOXL1*, that caused a p.Arg141Leu alteration and was associated with glaucoma and pseudoexfoliation syndrome(318), and *ARMS2* intron variant rs10490924 that resulted in the alanine-to-serin substitution (p.Ala69Ser), linked to increased susceptibility to age-related macular degeneration (319). Other markers exhibited deleterious effects such as variants located within *RGR* (rs1042454; p-value = 6.89×10^{-68}) previously reported as associated with refraction (110,320) and night blindness (321) and in *FBN1* gene, whose mutations caused Marfan syndrome and iridogoniodysgenesis (322).

4.3.2.2 Gene-set enrichment analyses

To explore the functional link between identified polymorphisms and refractive phenotypes, gene-set enrichment analyses were performed which elucidated properties shared by the genes that were significant in the final meta-analysis. The enrichment analysis for Gene Ontology terms revealed genes involved in gene expression, including nucleic acid binding transcription factor activity and regulation of RNA polymerase II transcription (p-value = 1×10^{-6}) (Figure 3). Additionally, significant enrichment was found for genes participating in sensory organ and ocular development (p-value = 6.10×10^{-6}), specifically retina formation in the camera-type eye (p-value = 1.3×10^{-05}). Moreover, genes associated with the refractive error were involved in broader CNS functions, such as synaptic signalling (p-value = 4×10^{-05}) regulation of neuron death and apoptotic process (p-value = 0.0001), but also circadian rhythmicity (p-value = 1.10×10^{-4}) and diurnal variation of gene expression, a mechanism that was not reported in previous GWAS (Figure 3).

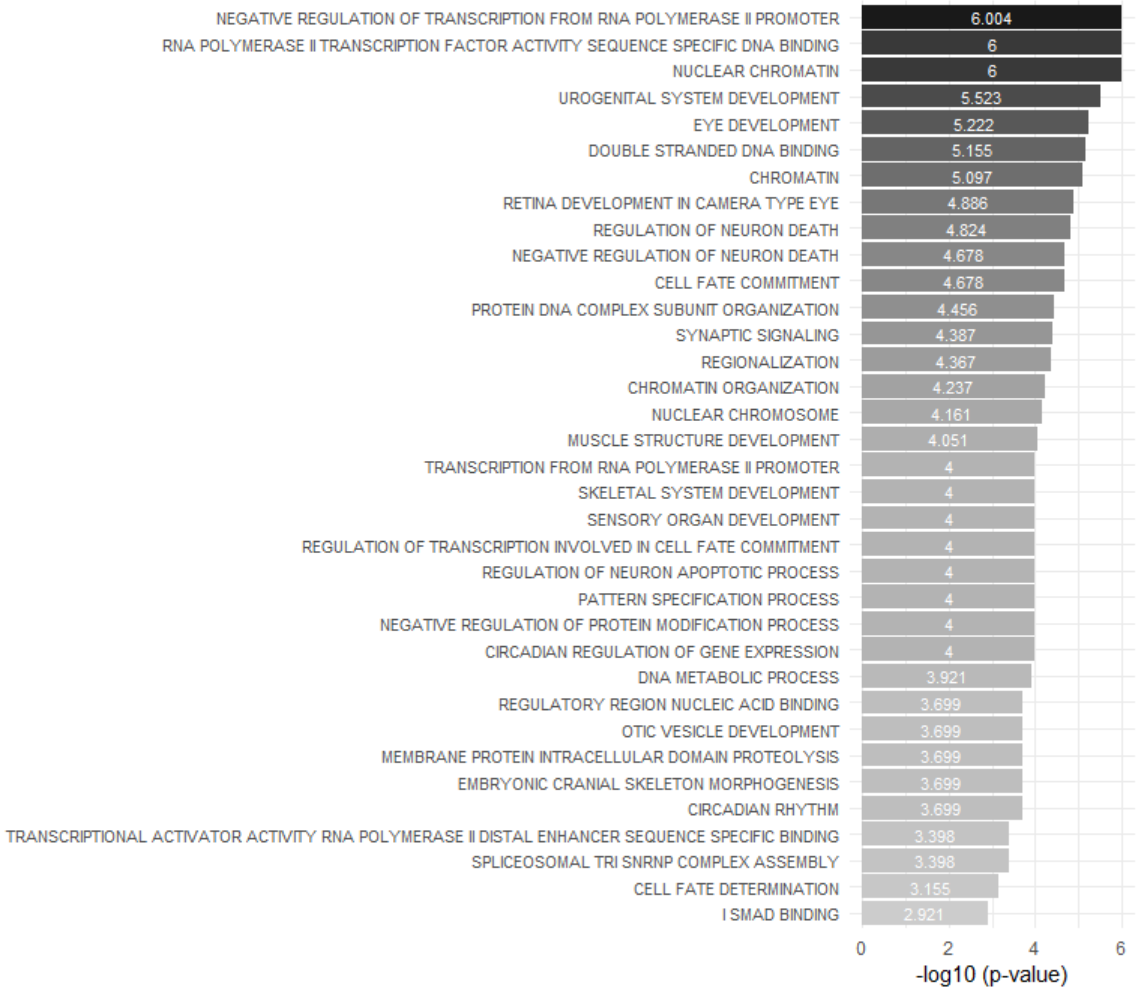


Figure 3 Gene-set enrichment analysis for Gene Ontology terms. The analysis tested the observed versus expected number of gene scores above the 75th centile. The -log(p-values) of the FDR test are displayed on X-axis.

4.3.2.3 Transcription factor enrichment analyses

Over-representation of sites targeted by *RREB1, GATA4* and *EP300* genes was observed in the transcription factor binding enrichment analyses. Moreover, the binding sites of other highly enriched transcription factors were involved in ocular development and morphogenesis. Particularly, *MAF*, *PITX2*, *FOXC1* and *CRX* were implicated in autosomal cataracts, anterior segment dysgenesis and cone-rod dystrophy, respectively (99). Moreover, the transcription factors of *CRX* and *PAX4,* also participated in melatonin synthesis and circadian rhythm (323). All significantly enriched gene sets were not reported in prior refractive error GWAS, but similar mechanisms were previously suggested (324).

4.3.2.4 Gene expression in different body tissues

Significant polymorphisms were placed near or within genes expressed in a number of body tissues, including the nervous system which supported evidence of CNS involvement in refractive development. Moreover, several of the associated genes were highly expressed in different eye structures, specifically in the cornea, trabecular meshwork (311), ciliary body as well as the retina (Figure 4)(325).

Additionally, observed genetic associations most strongly correlated with gene expression levels in the retina and CNS basal ganglia, although these correlations were not significant after multiple testing adjustments. Interestingly, concomitant associations with refraction and eQTL transcriptional regulation effects in peripheral blood and brain tissues were respectively identified for 159 and 97 genes (Appendix Tables 21-22). There was also evidence linking refractive error-related genes to brain methylation levels. Moreover, SMR-based integration analyses of retina-specific eQTLs with refractive error meta-GWAS identified 41 genes whose expression levels were associated with the refractive state (Table 2).

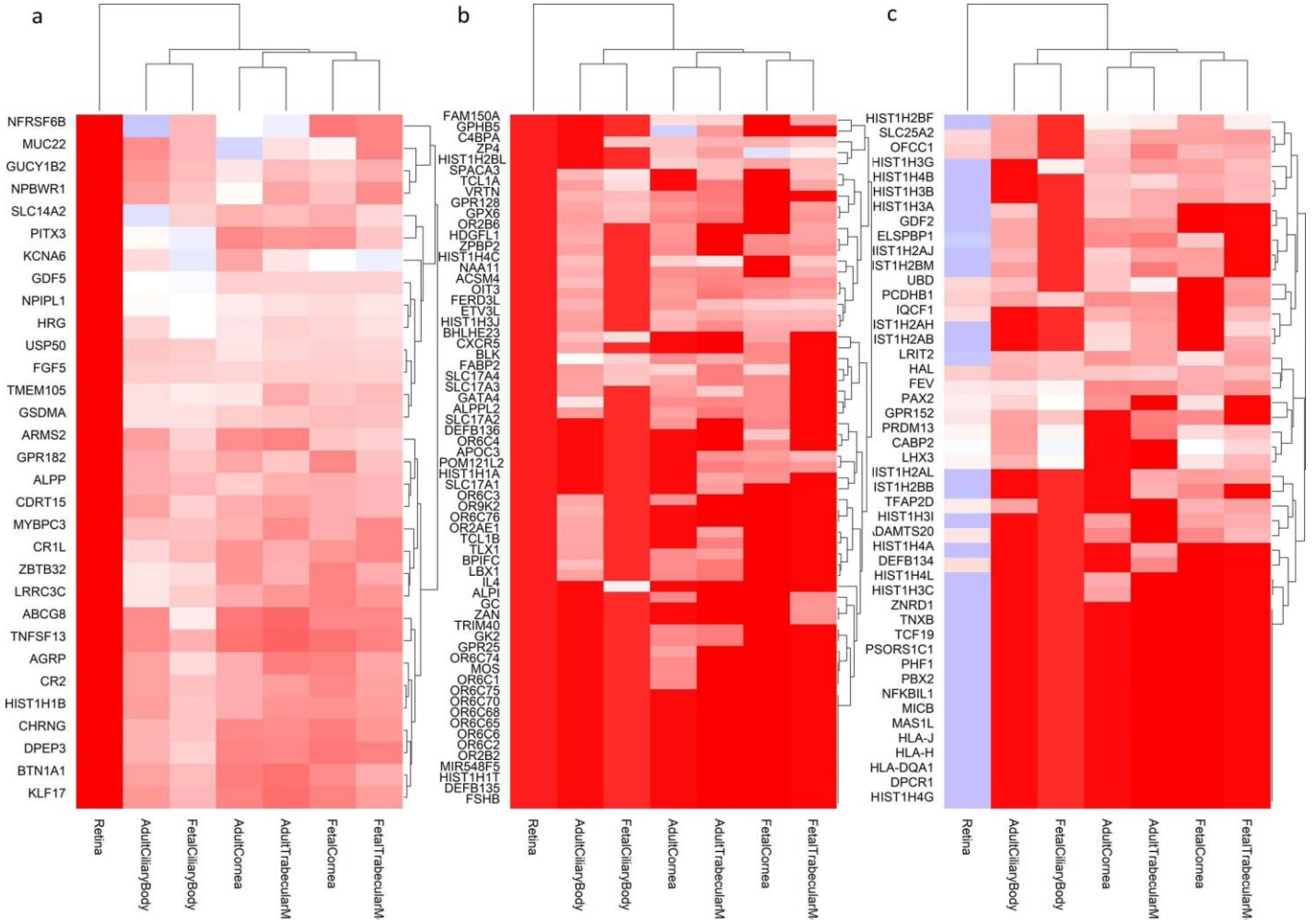


Figure 4 Genes from the regions associated with refractive error that are particularly expressed in eye tissues, compared to non-ocular tissue. a) genes that are expressed more in other ocular tissues (foetal and adult) but much less in the adult retina. b) genes that are highly expressed in the retina and other ocular tissues, and c) genes that are expressed in the retina, but less in the other ocular tissues tested (99).

ProbeID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ENSG00000257176	*ERGIC2*	rs10843390	T	C	0.27	-0.06	0.01	4.28×10^{-06}	1.23	0.06	1.94×10^{-86}	-0.05	0.01	7.58×10^{-06}	0.001
ENSG00000010626	*LRRC23*	rs2040352	A	C	0.28	-0.05	0.01	3.02×10^{-06}	1.04	0.07	4.66×10^{-55}	-0.05	0.01	7.69×10^{-06}	7.12×10^{-09}
ENSG00000141219	*C17orf80*	rs12951304	C	T	0.47	0.04	0.01	2.78×10^{-06}	-1.01	0.06	1.08×10^{-60}	-0.04	0.01	6.57×10^{-06}	1.94×10^{-24}
ENSG00000138111	*MFSD13A*	rs12762305	C	T	0.46	0.01	0.002	2.73×10^{-06}	0.95	0.06	1.39×10^{-54}	0.01	0.002	7.11×10^{-06}	2.41×10^{-15}
ENSG00000151789	*ZNF385D*	rs259453	A	T	0.19	-0.06	0.01	2.70×10^{-06}	1.12	0.08	2.48×10^{-47}	-0.06	0.01	8.09×10^{-06}	2.35×10^{-18}
ENSG00000142655	*PEX14*	rs2480777	T	C	0.33	-0.01	0.002	1.60×10^{-06}	0.97	0.06	6.34×10^{-56}	-0.01	0.002	4.42×10^{-06}	2.8×10^{-29}
ENSG00000187260	*WDR86*	rs12533730	T	C	0.25	-0.02	0.005	1.27×10^{-06}	0.83	0.07	1.22×10^{-30}	-0.03	0.01	8.02×10^{-06}	2.02×10^{-06}
ENSG00000123427	*EEF1AKMT3*	rs3782130	C	G	0.33	-0.03	0.007	1.15×10^{-06}	0.87	0.06	1.98×10^{-41}	-0.04	0.01	4.74×10^{-06}	7.75×10^{-16}
ENSG00000168297	*PXK*	rs7618604	A	G	0.17	-0.03	0.005	8.90×10^{-07}	0.95	0.09	1.12×10^{-26}	-0.03	0.01	8.01×10^{-06}	1.24×10^{-07}
ENSG00000131711	*MAP1B*	rs72761109	T	C	0.31	-0.04	0.008	8.64×10^{-07}	1.13	0.07	1.19×10^{-64}	-0.04	0.01	2.29×10^{-06}	9.12×10^{-06}
ENSG00000119139	*TJP2*	rs7850573	A	G	0.42	-0.02	0.005	8.06×10^{-07}	0.67	0.07	7.80×10^{-26}	-0.03	0.01	7.96×10^{-06}	8.92×10^{-15}
ENSG00000255021	*SLC6A6, LSM3*	rs4685109	G	C	0.36	0.03	0.005	7.70×10^{-07}	-1.08	0.06	3.01×10^{-82}	-0.02	0.01	1.69×10^{-06}	4.96×10^{-26}
ENSG00000146833	*TRIM4*	rs2082744	C	T	0.41	0.03	0.006	7.39×10^{-07}	-0.93	0.06	8.32×10^{-48}	-0.03	0.01	2.78×10^{-06}	1.95×10^{-15}
ENSG00000223496	*EXOSC6*	rs7192268	T	C	0.49	-0.01	0.002	6.07×10^{-07}	-1.22	0.05	2.46×10^{-128}	0.01	0.00	1.03×10^{-06}	8.49×10^{-13}
ENSG00000261556	*SMG1P7*	rs7192268	T	C	0.49	-0.01	0.002	6.07×10^{-07}	-1.04	0.06	5.00×10^{-73}	0.01	0.00	1.52×10^{-06}	4.39×10^{-11}
ENSG00000263503	*MAPK8IP1P2*	rs2732651	T	C	0.16	-0.08	0.02	6.01×10^{-07}	1.33	0.08	1.75×10^{-64}	-0.06	0.01	1.69×10^{-06}	0.96
ENSG00000125247	*TMTC4*	rs1283228	A	G	0.43	-0.03	0.01	5.68×10^{-07}	0.68	0.07	6.22×10^{-24}	-0.04	0.009	7.42×10^{-06}	1.26×10^{-17}
ENSG00000088448	*ANKRD10*	rs7987467	A	G	0.41	-0.05	0.01	5.43×10^{-07}	-0.94	0.06	5.35×10^{-49}	0.05	0.01	2.11×10^{-06}	5.59×10^{-17}
ENSG00000204172	*AGAP9*	rs4342964	T	C	0.21	-0.02	0.004	4.79×10^{-07}	-1.01	0.08	3.60×10^{-34}	0.02	0.004	3.28×10^{-06}	5.39×10^{-10}
ENSG00000254929	*ZNF488 ,ANXA8*	rs4342964	T	C	0.21	-0.02	0.004	4.79×10^{-07}	-0.84	0.08	6.40×10^{-23}	0.02	0.01	7.35×10^{-06}	3.98×10^{-09}
ENSG00000111671	*SPSB2*	rs2269359	A	G	0.28	-0.05	0.01	4.71×10^{-07}	0.78	0.07	1.93×10^{-27}	-0.06	0.01	4.89×10^{-06}	4.65×10^{-07}
ENSG00000206579	*XKR4*	rs11785942	T	C	0.31	-0.04	0.008	4.67×10^{-07}	0.82	0.07	1.87×10^{-29}	-0.05	0.01	4.22×10^{-06}	6.67×10^{-10}
ENSG00000144455	*SUMF1*	rs4685744	T	C	0.49	-0.02	0.003	$4.17 \times 10_{-07}$	-0.63	0.07	1.39×10^{-21}	0.03	0.01	7.78×10^{-06}	3.09×10^{-12}
ENSG00000134744	*GLIS1*	rs7355078	T	C	0.05	-0.08	0.02	3.18×10^{-07}	1.81	0.14	7.66×10^{-39}	-0.05	0.01	1.94×10^{-06}	4.86×10^{-06}
ENSG00000080644	*CHRNA3*	rs12914385	T	C	0.38	-0.05	0.009	2.20×10^{-07}	-0.72	0.07	5.06×10^{-26}	0.07	0.01	3.30×10^{-06}	1.49×10^{-14}
ENSG00000196502	*SULT1A1*	rs151321	C	T	0.46	0.05	0.01	2.15×10^{-07}	0.72	0.07	1.17×10^{-26}	0.08	0.02	3.08×10^{-06}	9.43×10^{-10}
ENSG00000214078	*CPNE1*	rs6060524	T	G	0.10	-0.06	0.01	2.08×10^{-07}	-1.39	0.12	3.58×10^{-30}	0.04	0.01	2.29×10^{-06}	1.33×10^{-08}
ENSG00000140299	*BNIP2*	rs7183831	A	C	0.41	-0.04	0.007	2.06×10^{-07}	1.04	0.06	4.16×10^{-68}	-0.03	0.01	6.43×10^{-07}	7.22×10^{-18}
ENSG00000169738	*DCXR*	rs6502047	A	G	0.12	-0.05	0.01	2.05×10^{-07}	1.34	0.10	2.11×10^{-39}	-0.04	0.01	1.37×10^{-06}	0.01
ENSG00000245532	*NEAT1*	rs674485	G	A	0.27	0.05	0.009	1.91×10^{-07}	0.85	0.08	6.02×10^{-29}	0.06	0.01	2.37×10^{-06}	4.23×10^{-19}
ENSG00000092820	*EZR*	rs3102966	G	T	0.37	0.02	0.004	1.54×10^{-07}	0.75	0.07	5.74×10^{-29}	0.03	0.01	2.04×10^{-06}	4.04×10^{-09}
ENSG00000229152	*ANKRD10-IT1*	rs7332021	A	G	0.41	-0.05	0.01	1.47×10^{-07}	-0.7	0.07	1.67×10^{-24}	0.07	0.02	2.96×10^{-06}	8.48×10^{-11}
ENSG00000106536	*POU6F2*	rs11973905	T	C	0.21	-0.03	0.005	1.36×10^{-07}	0.7	0.08	2.73×10^{-19}	-0.04	0.01	5.48×10^{-06}	7.80×10^{-09}
ENSG00000177994	*C2orf73*	rs12713256	G	A	0.23	0.05	0.009	1.34×10^{-07}	-0.80	0.08	8.96×10^{-26}	-0.06	0.01	2.45×10^{-06}	3.68×10^{-08}

ProbeID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ENSG00000112706	*IMPG1*	rs1326170	T	C	0.32	-0.03	0.006	1.20×10^{-07}	0.64	0.07	6.34×10^{-18}	-0.05	0.01	6.42×10^{-06}	9.10×10^{-08}
ENSG00000100225	*FBXO7*	rs1029301	A	G	0.37	-0.02	0.004	1.02×10^{-07}	1.23	0.06	1.11×10^{-99}	-0.02	0.004	2.43×10^{-07}	0.0002
ENSG00000149591	*TAGLN*	rs488962	T	C	0.23	-0.04	0.007	6.18×10^{-08}	-1.04	0.07	6.71×10^{-45}	0.04	0.01	4.38×10^{-07}	5.74×10^{-31}
ENSG00000166033	*HTRA1*	rs11200645	T	C	0.34	-0.05	0.010	5.97×10^{-08}	-0.61	0.08	1.24×10^{-15}	0.09	0.02	7.22×10^{-06}	2.21×10^{-08}
ENSG00000125991	*ERGIC3*	rs2277862	T	C	0.14	-0.07	0.01	5.46×10^{-08}	1.25	0.10	3.37×10^{-35}	-0.06	0.01	6.45×10^{-07}	5.85×10^{-09}
ENSG00000119718	*EIF2B2*	rs735452	C	T	0.48	0.04	0.007	5.03×10^{-08}	-0.67	0.06	2.06×10^{-24}	-0.05	0.01	1.54×10^{-06}	3.26×10^{-16}

Table 2 SMR results using retinal eQTL data. Columns containing "ProbeID" is the ID for the eQTL probe, "Gene" is the gene that the probe eQTL effects act on, "topSNP" is the SNP associated with the largest effects on the trait (Spherical Equivalent) from the meta-analysis; "A1", "A2" and "Freq" are the effect allele, other allele, and effect allele frequency for the 'topSNP',respectively. "b_GWAS", "se_GWAS" and "p_GWAS" are the effect size, standard error, and p-value for the "topSNP". "b_eQTL", "se_eQTL" and "p_eQTL" are the effect size, standard error, and p-value for the "probeID". "b_SMR", "se_SMR", and "p_SMR" are the effect size, standard error, and p-value results from the SMR analysis. "p_HEIDI" is the p-value from the HEIDI (Heterogeneity in Dependent Instruments) test.

Refractive error shared susceptibility variants with several other traits, including cognitive ability, education and a number of ocular disorders (Table 3). In particular, on the genetic level refraction strongly correlated with childhood IQ ($r_g = -0.27$, p-value $= 4.76 \times 10^{-9}$) and adulthood IQ ($r_g = -0.25$, p-value $= 1.56 \times 10^{-39}$), education ($r_g = -0.24$, p-value $= 3.36 \times 10^{-54}$), cataracts ($r_g = -0.31$, p-value $= 4.70 \times 10^{-10}$) and intraocular pressure ($r_g = -0.14$, p-value $= 1.04 \times 10^{-12}$) (Table 3). Further interrogation of observed genetic correlation with educational attainment revealed one genome-wide significant genetic interaction with a variant located within the *TRPM1* gene (rs536015141, p-value $= 2.35 \times 10^{-9}$), encoding protein expressed in ON-bipolar cells and implicated in stationary night-blindness (326). Previous research suggested a non-linear relationship between refractive error and education (327), explaining the paucity of identified genetic interactions.

Phenotype	**PMID**	**r_g**	**se**	**z**	**p-value**
Years of schooling 2016	27225129	-0.24	0.02	-15.50	3.361×10^{-54}
Fluid intelligence score	UKBB	-0.25	0.02	-13.16	1.56×10^{-39}
Intelligence	28530673	-0.26	0.02	-11.24	2.56×10^{-29}
Years of schooling (proxy cognitive performance)	25201988	-0.27	0.03	-9.54	1.472×10^{-21}
Years of schooling 2013	23722424	-0.26	0.03	-9.32	1.192×10^{-20}
IOP	29785010	-0.14	0.02	-7.12	1.042×10^{-12}
Eye problems/disorders: Cataract	UKBB	-0.31	0.05	-6.23	4.69×10^{-10}
Childhood IQ	23358156	-0.27	0.05	-5.86	4.76×10^{-09}
Diagnoses – main ICD10: H26 Other cataract	UKBB	-0.26	0.04	-5.74	9.36×10^{-09}
Vertical cup-to-disc ratio (VCDR)	28073927	-0.03	0.03	-0.93	0.4
Hair Colour	29662168	-0.01	0.02	-0.75	0.5

Table 3 Genetic correlation between spherical equivalent and other phenotypic traits. "Phenotype" is the trait tested for genetic correlation. "PMID" is the PubMed ID for the study providing summary statistics of the phenotype. "r_g" is the genetic correlation value and "se", "z", "p-value" are the respective standard error, Z-score and p-value.

4.3.2.5 Mendelian Randomisation analyses

The relationship between refraction and intraocular pressure was explored by building Mendelian Randomisation models (MR). MR analyses showed a causal association between elevated intraocular pressure and myopic refraction. Particularly, for every 1mm IOP increase spherical equivalent decreased by 0.05–0.09 Dioptres and populations' refractive error shifted towards more myopic ranges (Figure 5). Conversely, there was no evidence of a causal relationship between intelligence and refractive error, but instead, MR results suggested genetic pleiotropy (Figure 6)

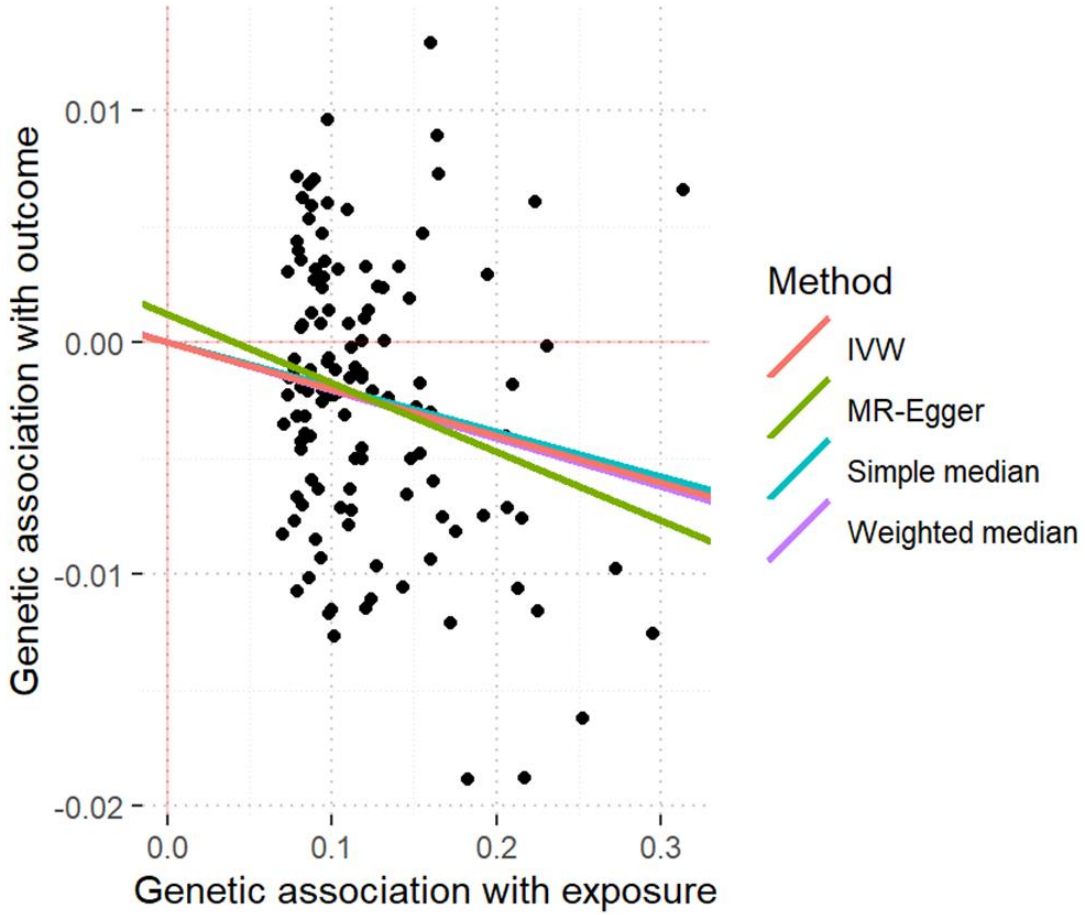


Figure 5 Mendelian randomisation results on the causality of IOP over refractive error. Single points in the graph represent coordinates determined by the effect of each specific SNP over IOP (x-axis, mmHg) and spherical equivalent (y-axis, Dioptre units).

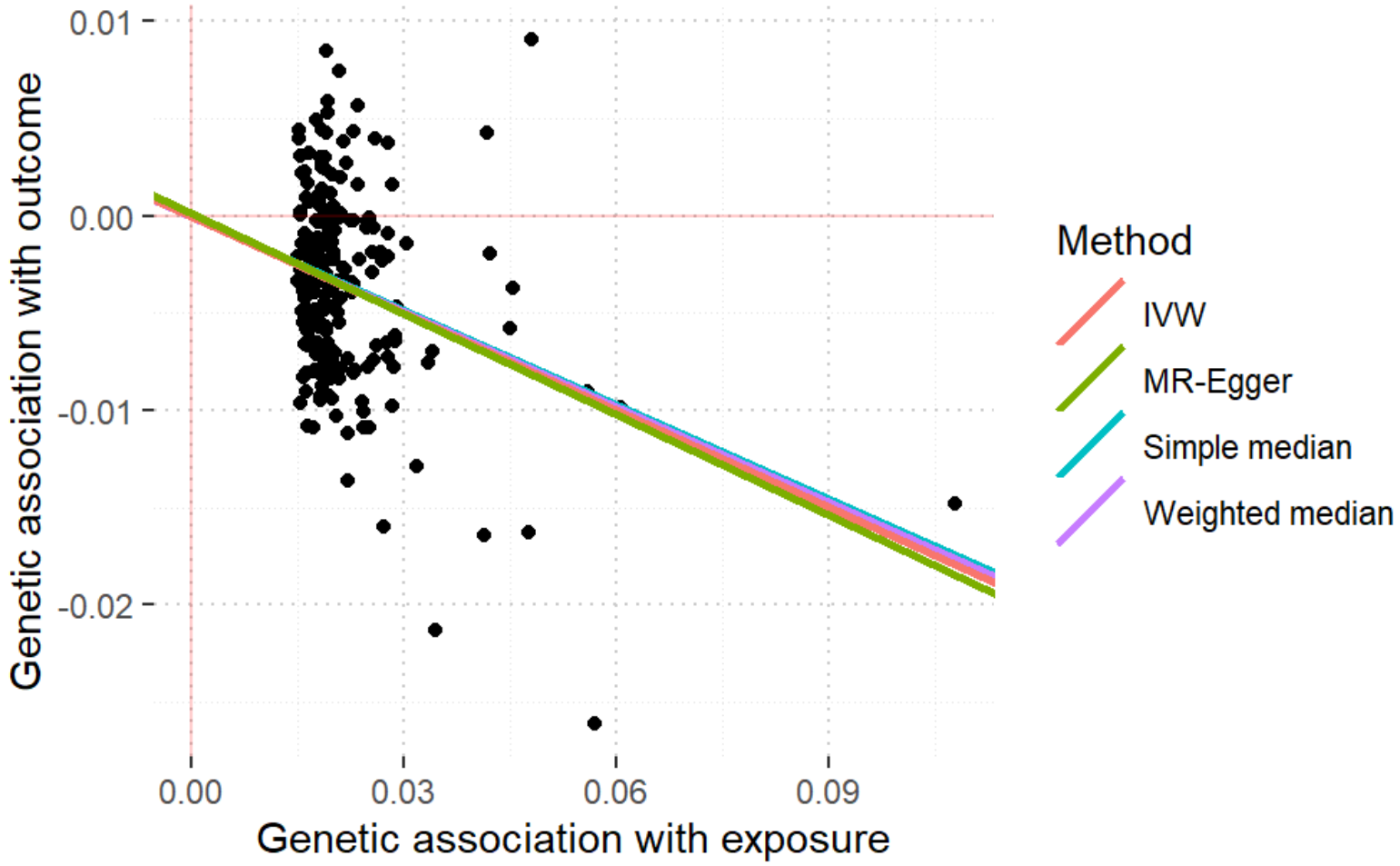


Figure 6 Mendelian randomisation results on the causality of fluid intelligence over refractive error. Single points in the graph represent coordinates determined by the effect of each specific SNP over intelligence score and spherical equivalent (y-axis, Dioptre units).

4.3.2.6 Conditional analyses

In order to identify secondary genetic associations at significant loci, I conducted conditional analyses. The origin of association signals was traced to 904 independent variants associated with refractive error at genome-wide significance level; of them, 890 were present in the EPIC-Norfolk study, a separate cohort that was not included in the meta-analysis (99). After fitting the linear model to spherical error and estimating the variance explained by all available markers, I ascertained that these SNPs cumulatively predicted 12.1% of spherical equivalent phenotypic variance and 18.4 % of its heritability in the EPIC-Norfolk cohort. By contrast, novel associations explained 4.6 % of refractive error variance significantly improving heritability reported for previously published associations (110)

4.3.2.7 Myopia prediction model

Using 890 significant variants from the conditional analysis for which genotypes were available for the EPIC-Norfolk participants, I constructed PRS and assessed the performance of associated SNPs in a regression-based myopia prediction model. The model, in addition to including associated alleles, was also adjusted for age and sex. Despite being limited to a smaller number of SNPs, the prediction model distinguished between myopia cases and controls. In particular, it performed well for mild myopic refraction, yielding an AUC of 0.67. Even better results were achieved for moderate and severe myopia with AUC of 0.74 and 0.75, respectively (Figure 7). Nonetheless, the models' predictive capacity did not markedly improve for more severe near-sightedness, suggesting that common susceptibility variants identified in the meta-analysis represented a genetic risk of common myopia observed in the general population, while results of recent genetic investigations provided evidence for the uniqueness of high myopia's genetic architecture (328).

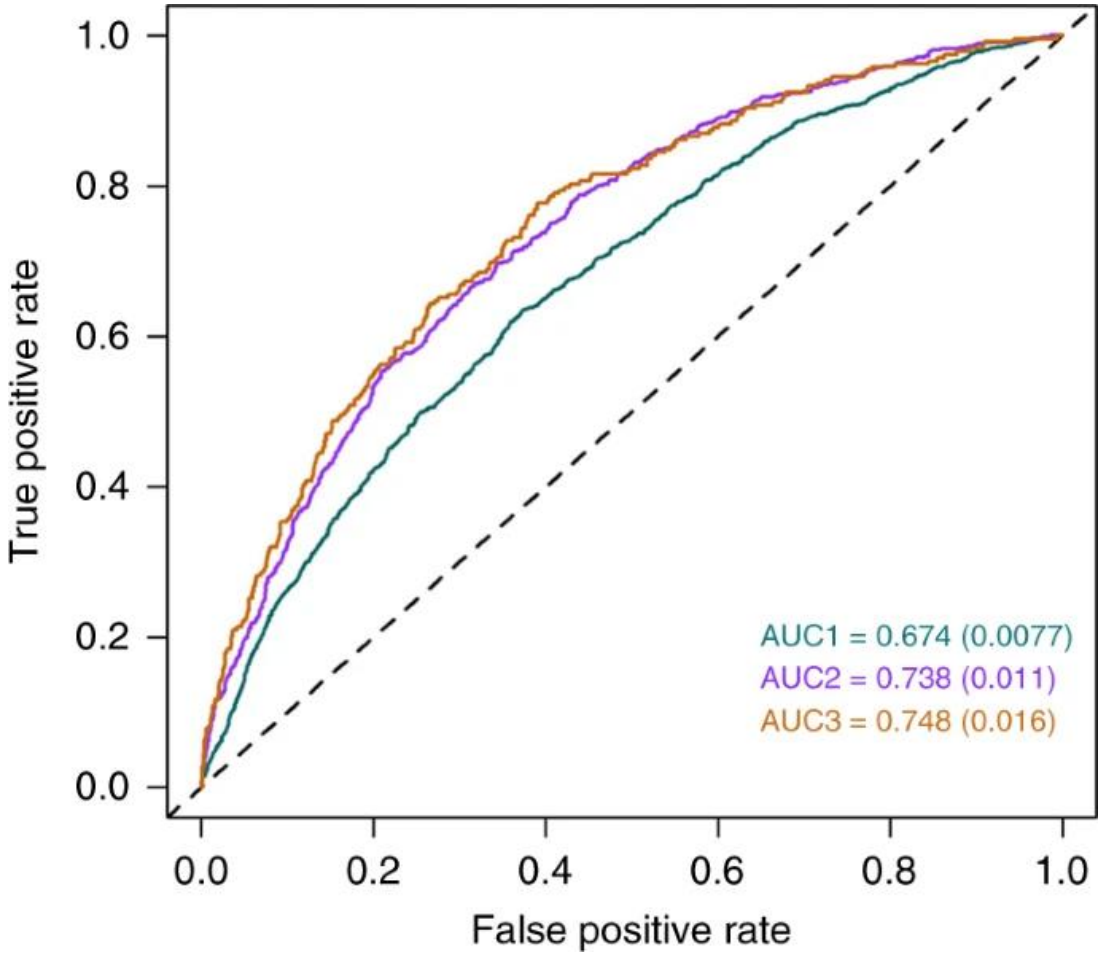


Figure 7 Receiver operating characteristic (ROC) curves for myopia predictions, using information from 890 SNP markers identified in the meta-analysis. The three different colors represent three different curves for each of the different definitions of myopia: green, all myopia (defined as <−0.75 D); magenta, moderate myopia (<−3.00 D); and brown, severe myopia (<−5.00 D).

4.3.2.8 Validation of myopia prediction models in a cohort of Singaporean children

Cycloplegic auto-refraction and PRS based on 655 SNPs discovered in Europeans, were available for 1,004 Chinese children in SCORM. The basic model which only included age and sex, maternal education, school and 10 PCs explained 4.0% and 10.8% of spherical equivalent (SPHE) and axial length (AL) variance, respectively.

The inclusion of PRS in the basic model markedly improved its' predictive capacity, demonstrating an incremental increase in adjusted R^2 for SPHE (0.041; ANOVA p-value< 0.001) and AL (0.022, ANOVA p-value< 0.001). Although the addition of parental myopia and time outdoors also contributed to statistically significant improvement over a basic model, the change in adjusted R^2 was smaller than when PRS was considered in the model (Table 4).

A multivariate model containing age, sex, mother's education, outdoor time, number of books read per week and 10 PCs as independent covariates performed better with PRS, predicting between 11.9% and 15.7% of SPHE and AL variance. By contrast, a multivariate model that did not include PRS, but instead assessed the impact of parental myopia, had worse accuracy in relation to SPHE (R^2 = 8.3%) and AL (R^2 = 13.5%) (Table 5).

	Spherical equivalent (D)				Axial length (mm)			
	Adjusted R^2 full model	Incremental adjusted R^2	Incremental 95% CI	ANOVA p-value	Adjusted R^2 full model	Incremental adjusted R^2	Incremental 95% CI	ANOVA p-value
Basic model	0.04	-	-	-	0.106	-	-	-
Basic model + PRS	0.082	0.041	0.01;0.073	4.95×10^{-07}	0.13	0.022	-0.001;0.046	1.49×10^{-04}
Basic model + parental myopia	0.051	0.011	-0.006;0.027	7.16×10^{-03}	0.112	0.004	-0.006;0.014	0.065
Basic model + time outdoors	0.076	0.035	0.006;0.064	3.36×10^{-06}	0.133	0.026	0.0004;0.051	5.06×10^{-05}
Basic model + books read per week	0.04	0	0;0	0.466	0.106	0	0;0	0.849

Table 4 Multivariate linear regression models of polygenic risk score (PRS), parental myopia, time outdoors and books read per week, assessing the prediction accuracy of spherical equivalent and axial length. The basic model included age, sex, maternal education, school and 10 principal components. A basic model predicting axial length was additionally adjusted for height. Columns ''Adjusted R^2 full model'' listed total phenotypic variance explained by all covariates in the model. Field '' Incremental adjusted R^{2}'' showed additional phenotypic variance explained by including the PRS, parental myopia, time outdoors and books read per week in the model with corresponding 95% confidence interval and p-values displayed in columns ''Incremental 95% CI'' and "ANOVA p-value", respectively.

	Spherical equivalent (D)			Axial length (mm)		
	Univariate	Multivariate	n	Univariate	Multivariate	n
Polygenic risk score[1]			568			548
beta	-0.21	-0.22		0.18	0.17	
SE	0.04	0.04		0.04	0.04	
95% CI	-0.30, -0.13	-0.30, -0.14		0.09, 0.26	0.09, 0.25	
p-value	6.21×10^{-07}	1.49×10^{-7}		4.70×10^{-05}	5.39×10^{-05}	
R^2	0.041	0.119		0.038	0.157	
Parental myopia[1]			568			548
beta	-0.32	-0.23		0.3	0.15	
SE	0.09	0.09		0.09	0.09	
95% CI	-0.49, -0.15	-0.41, -0.05		0.13, 0.47	-0.03, 0.33	
p-value	2.10×10^{-04}	0.01		6.05 x 10-4	0.1	
R^2	0.022	0.083		0.029	0.135	
Time outdoors[2]			568			548
beta	0.16	0.21		-0.13	-0.18	
SE	0.04	0.04		0.04	0.04	
95% CI	0.08, 0.24	0.13, 0.29		-0.21, -0.05	-0.26, -0.10	
p-value	1.08×10^{-04}	8.77×10^{-07}		2.09×10^{-03}	1.94×10^{-05}	
R^2	0.024	0.119		0.025	0.157	
Books read per week[3]			568			548
beta	0.01	0.03		0.02	0.02	
SE	0.05	0.05		0.05	0.05	
95% CI	-0.09, 0.11	-0.07, 0.12		-0.08, 0.13	-0.08, 0.11	
p-value	0.88	0.59		0.68	0.75	
R^2	0	0.119		0.008	0.157	

Table 5 Multivariate linear regression models of PRS, parental myopia, time outdoors and books read per week for spherical equivalent and axial length in SCORM (N = 1004). The fields "beta", "SE", "95% CI" and "p-value" denote respectively the regression coefficient (slope), the standard errors of the coefficient estimation, the coefficients lower and upper bounds of the 95% confidence intervals, and p-values for the association. Field "R^2" indicates total phenotypic variance explained by the model and column ''n" shows the sample size included in the analyses. [1]Multivariate models were adjusted for baseline age, sex, mother's education, school, time outdoors, number of books read per week and 10 genotyping principal components. AL at the last visit was additionally adjusted for baseline height. [2]Multivariate models were adjusted for baseline age, sex, mother's education, school, polygenic risk score, number of books read per week and 10 genotyping principal components. AL at the last visit was additionally adjusted for baseline height. [3]Multivariate models were adjusted for baseline age, sex, mother's education, school, polygenic risk score, time outdoors and 10 genotyping principal components. AL at the last visit was additionally adjusted for baseline height.

Additionally, PRS appeared to predict the onset of high (HR = 1.49; R^2 = 12.8%) and moderate degree myopia (HR = 1.39; R^2 = 12.5%) and was associated with their increased hazard (Table 6). Similarly, parental myopia predisposed to earlier onset of

moderate (HR = 1.34; R^2 = 8.9 %) and severe (HR= 1.61; R^2 = 10.6 %) near-sightedness but was slightly less accurate (Table 6).

	Time to moderate myopia (-3.00 D)			**Time to high myopia (-5.00 D)**		
	Univariate	Multivariate	n	Univariate	Multivariate	n
Polygenic risk score[1]			672			672
HR	1.33	1.39		1.42	1.49	
SE	0.06	0.06		0.08	0.08	
95% CI	1.18, 1.49	1.24, 1.56		1.21, 1.67	1.26, 1.75	
p-value	1.19×10^{-06}	1.25×10^{-08}		1.97×10^{-05}	1.84×10^{-06}	
R^2	0.034	0.125		0.026	0.128	
Parental myopia[1]			672			672
HR	1.52	1.34		1.8	1.61	
SE	0.12	0.13		0.19	0.21	
95% CI	1.19, 1.93	1.03, 1.74		1.25, 2.61	1.07, 2.41	
p-value	7.60×10^{-04}	0.03		1.71×10^{-03}	0.02	
R^2	0.018	0.089		0.016	0.106	
Time outdoors[2]			672			672
HR	0.89	0.79		0.71	0.58	
SE	0.06	0.06		0.09	0.1	
95% CI	0.79, 1.00	0.70, 0.90		0.59, 0.85	0.47, 0.70	
p-value	0.04	2.80×10^{-04}		1.41×10^{-04}	8.30×10^{-08}	
R^2	0.006	0.125		0.023	0.128	
Books read per week[3]			672			672
HR	1.09	1.08		1.16	1.15	
SE	0.05	0.06		0.06	0.06	
95% CI	0.98, 1.22	0.97, 1.21		1.04, 1.30	1.02, 1.31	
p-value	0.1	0.17		7.32×10^{-03}	0.03	
R^2	0.003	0.125		0.007	0.128	

Table 6 Cox proportional hazard regression models of time to moderate myopia and time to high myopia in SCORM (N = 1,004). The fields "HR", "SE", "95% CI" and "p-value" denote respectively the regression hazard ratio (HR), the standard errors of the HR estimation, the HR lower and upper bounds of the 95% confidence intervals, and p-values for the association. Field "R^2" indicates total phenotypic variance explained by the model and column ''n" shows the sample size included in the analyses. [1]Multivariate models were adjusted for baseline age, sex, mother's education, school, time outdoors, number of books read per week and 10 genotyping principal components. [2]Multivariate models were adjusted for baseline age, sex, mother's education, school, polygenic risk score, number of books read per week and 10 genotyping principal components.[3]Multivariate models were adjusted for baseline age, sex, mother's education, school, polygenic risk score, time outdoors and 10 genotyping principal components. AL at the last visit was additionally adjusted for baseline height.

The AUC was the indicator I used to evaluate the predictive value of PRS for high/moderate myopia compared to healthy (non-myopia) status. By convention, an AUC of 1 indicated a perfect ability to distinguish between affected and non-affected individuals whereas a value of 0.5 suggested a bad performance of the model which is at this point randomly predicting the disease status as case or control. A prediction model that was adjusted for age and included parental myopia and time outdoors as main predictors demonstrated an AUC of 0.72 and 0.60 for high and moderate myopia, accordingly. However, the inclusion of PRS significantly improved model performance in relation to high-degree myopic refractive error producing an AUC of 0.77 (Figure 8). Added PRS also predicted moderate myopia with an AUC of 0.62, but this result was not significantly different from a base prediction model. Interestingly, the PRS appeared to be more predictive (an AUC of 0.64 and 0.57 for high and moderate myopia, respectively) than parental near-sightedness (AUC of 0.61 and 0.56 for high and moderate myopia, respectively), when they were compared using two separate models. Overall, the best predictive model contained both the PRS and parental myopia in conjunction with age and time spent outdoors.

The prediction accuracy of the full multivariate model was similar in the GUSTO cohort which included 339 Chinese children and was used for replication of SCORM study results. In particular, in GUSTO the PRS was associated with more myopic refractive error (beta = −0.24) but explained a smaller fraction of SPHE variance (R^2 = 4.6%) compared to SCORM analyses (beta = −0.22; R^2 = 11.9%). The effect of the PRS over AL was of similar size in both studies (beta = 0.17 vs beta = 0.19 in SCORM vs GUSTO) and predicted between 15.7% – 16.7% of its' variance when other factors such as time outdoors, near-work, sex and height were considered. Additionally, in the GUSTO study, the inclusion of the PRS significantly improved prediction accuracy showing an incremental R^2 of 4.9% and 3.3% for SPHE and AL, respectively.

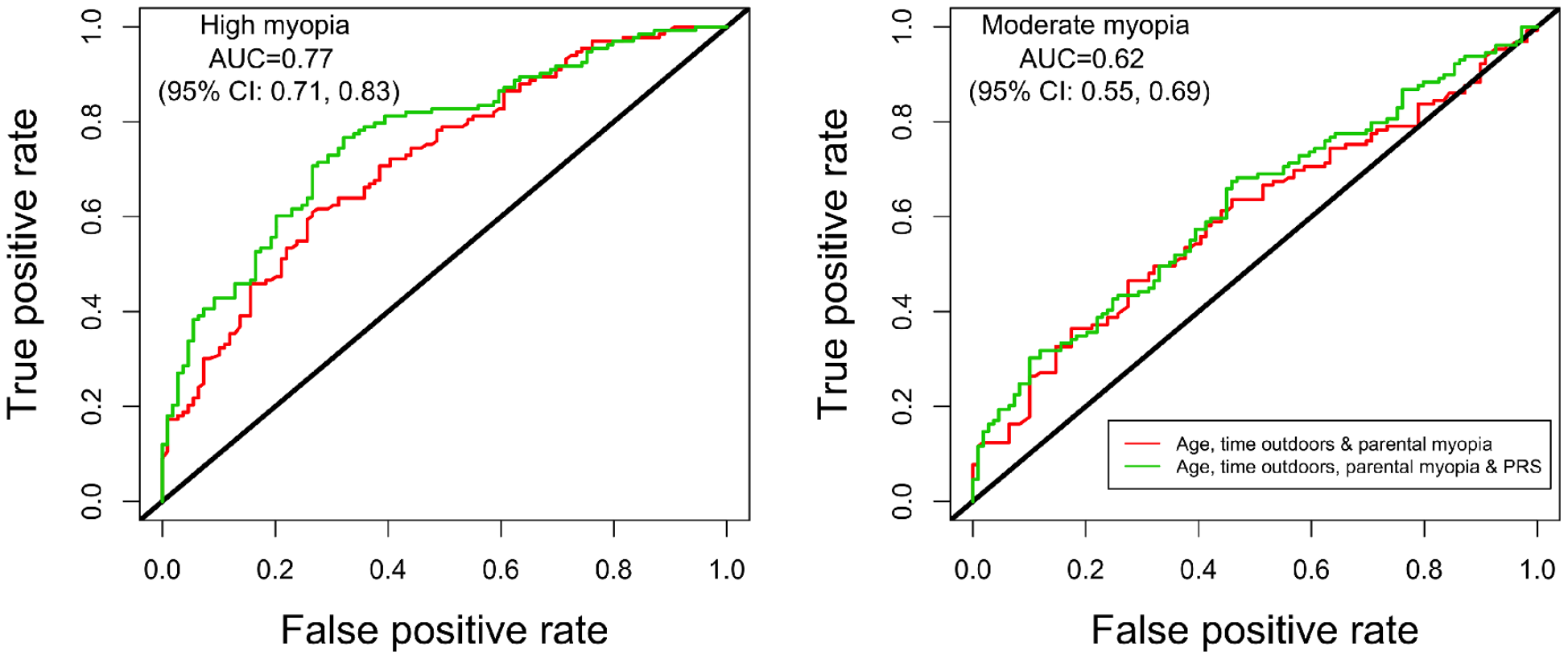


Figure 8 Receiver Operating Characteristic (ROC) curve for identifying high myopia (≤ −5.00 D) and moderate myopia (≤ −3.00 D) versus no myopia controls using the polygenic risk score in SCORM. ROC curve for high myopia (≤ −5.00 D, n = 133) and moderate myopia (≤ −3.00 D, n = 129) versus no myopia controls (n = 109) with PRS, age, time outdoors, and parental myopia as predictors.

4.4 Discussion

The refractive error meta-analysis described in this thesis found evidence for two major sets of mechanisms underlying the pathogenesis of refractive disorders (99). The first mechanism implicated IOP, ocular development and physiology as well as different eye structures (99). The second supported the notion of pleiotropy between CNS and refractive error and the role of circadian rhythm control in emmentropisation and refractive development (99). These two prominent mechanisms influenced all anatomical factors related to refractive power, light response and processing, photoconductance and cerebral function (99).

Many polymorphisms that showed significant associations with populations refractive error were located near or within retinal genes, supporting the hypothesis of light mediating ocular growth and emmentropisation and highlighting the importance of retina in the light-induced signalling (329). The findings were also corroborated by transcriptome-wide SMR analyses that revealed high levels of gene expression in retinal tissues. More recent work described a universal set of transcriptome changes, contained within well-defined retinal networks that couldn't be bypassed and played a fundamental role in ocular growth (330). These essential networks that could not be avoided in ocular growth inhibition included cell signalling and circadian entrainment (330). Although light transduction and retinal signalling were suggested pathways in several previous GWAS (20,82), the meta-analysis presented in this thesis was the first to provide genetic evidence implicating circadian rhythmicity in human refractive development. Together, these results implicate retinal light response and processing as the main drivers of refractive errors (99).

Evidence from animal experiments indicated that in addition to local ocular mechanisms, emmentropisation was under control from CNS (331). Consistent with these findings, the meta-analysis observed strong genetic correlations between spherical equivalent and intelligence as well as educational attainment and learning

disability (99). Mendelian randomisation further supported evidence of pleiotropy but not causality between mental capacity and refractive error, implicating shared genetic factors in neurocognitive and refractive development (145).

Although meta-analysis upheld prior findings of enrichment for cell-level properties, such as cation and non-cation ion channels, reported by smaller studies, it also provided evidence for broader physiological processes, including light processing and transduction (99). Moreover, for the first time higher regulatory mechanisms, contributing to smaller proportions of spherical equivalent variance, were identified. Specifically, the study results indicated significant enrichment for binding sites of the transcription factors involved in circadian rhythm regulation (99). It was previously suggested that rising myopia prevalence rates in contemporary societies could be partially driven by the increasing exposure to ambient lighting that desynchronised endogenous ocular circadian rhythms from the natural light/dark cycle (332). Accumulating evidence mostly from animal experiments supported the hypothesis that dysregulation of circadian rhythms contributed to the development of ametropias. Interestingly, both animal and human eyes experienced diurnal fluctuations in dimensions, including axial length and choroidal thickness. In animals with induced myopia or hypermetropia, the changes in these phase relationships coincided with the development of refractive errors (333), whereas there was a paucity of data on whether diurnal oscillations in ocular dimensions were affected in children with ametropias. In addition to hypotheses linking circadian rhythmicity to diurnal variations in axial length, choroidal thickness as well as IOP, the interaction between dopamine and retinal circadian clocks was postulated to affect refractive development and myopia (324).

In addition to describing biological mechanisms driving the refractive error development in the general population, the meta-analysis identified over 400 common polymorphisms that explained a substantial proportion of spherical equivalent variability and heritability (99). Despite being limited to a smaller number of significant variants the PRS -based prediction model performed well in a replication sample of

adults from the EPIC-Norfolk cohort, yielding an AUC of 0,74 and 0.75 for moderate and high degree myopia, respectively.

Moreover, a subsequent SCORM study demonstrated the utility of the PRS for myopia prediction in East-Asian children who are considered a high-risk group (1). In particular, the PRS based on SNPs discovered in Europeans explained a substantial fraction of spherical error and axial length variation in Singaporean adolescents but was also associated with the altered risk to time to myopia. Importantly, identified refractive error susceptibility loci in conjunction with age, time outdoors and parental myopia greatly improved prediction of severe myopia in Asian children. Generally, the full multivariate model with PRS was superior to the model that was adjusted for the same covariates but included parental myopia. The PRS alone incrementally explained 2.2% and 4.1% of adolescent spherical error and axial length, showing that parental myopia was a less effective proxy for genetic factors, as suggested previously (334). These findings indicate that parental myopia and inherited genetic risk could be exerting independent effects on a myopic refractive error in the offspring. In particular, parental myopia was hypothesised to capture the myopiagenic family environment (329). Predictably, the results of the SCORM study were lower than the variance predicted by reported genetic scores in European adults (7.9%) (110). This suggests that SNPs selected based on their effects in the European population were likely to perform better in predicting spherical error in Europeans.

Overall, the PRS based on SNPs reported for Europeans demonstrated a good accuracy of myopia prediction in the sample of schoolchildren of East-Asian descent. Consistent with expectations of imperfect transferability of the PRS across ancestry divergent populations (335), the prediction AUC generated in the SCORM cohort were lower than those seen in studies where training and target samples were both of European ancestry. Future large-scale GWAS in East Asians will help to close this gap and improve the PRS used for myopia prediction in the East-Asian populations.

Chapter 5 | Rare genetic variants associated with population's refractive error

5.1 Introduction

In addition to interrogating the associations with common genetic variance, this thesis explored the impact of rare polymorphisms. Several previous investigations sought to uncover genetic factors driving refractive changes in the population, however, the heritability that could be attributed to the reported loci was modest. Much of the genetic contribution to the population's refraction remained unexplained, and some more up-to-date reports suggested that lower-frequency markers could play an influential role in the development of more severe forms of refractive error (321).

Rare genetic variation is known to have a role in human diseases. Mendelian disorders in particular and rare forms of common diseases were caused by highly penetrant rare markers (336). Natural evolution against deleterious mutations makes them more likely to have low frequency in the population due to purifying selection which tends to suppress fitness-reducing alleles. However, mutations rates remain sufficiently large that purifying selection cannot efface all deleterious mutations, and polymorphisms with a modest effect on fitness may reach allele frequencies of 1% or occasionally more, particularly if the effect is recessive (336). Empirical data show that deleterious changes such as loss-of-function mutations which affected the generation of functional proteins were extremely rare. It also suggests that in addition to Mendelian disorders and familial analogues of chronic conditions, rare variants may influence complex diseases, but their effects were less penetrant (336). The appreciable number of rare markers was identified for complex cardiometabolic traits (244) and neurodevelopmental disorders (337,338) of which severity they impact. Nonetheless, the general consensus was that common SNPs created the background liability to complex diseases, whereas environmental exposures and rare polymorphisms gave an extra impetus that pushed individuals beyond the liability threshold of disease (336).

Exome sequencing has changed the understanding of rare disorders, uncovering hundreds of causal variants associated with these conditions. By contrast, the genetic

investigations of complex traits and diseases relied on GWAS, which analysed common polymorphisms and used genotyping arrays not designed to capture rare genetic variation. Recent genome-wide studies discovered hundreds of distinct loci harbouring genes involved in refraction development (99). Cumulatively those variants accounted for approximately 18% of total heritability (99), as discussed in *Chapter 4*, while twin studies estimate the refractive error heritability between 50%-90% (70). The missing heritability in refractive error GWAS could be attributed to several causes, such as confounding arising from linkage disequilibrium, statistical power limitations, epistasis and heritability explained by rare genetic variants that are usually not identifiable by traditional genetic association studies. Statistical power to detect associations with individual genetic variants is proportional to the magnitude of the risk they individually confer, but more crucially, to their frequency in the population.

Improved sequencing facilitated a more complete assessment of the impact of low-frequency and rare genetic variation and investigation of their role in complex traits and diseases. UK Biobank cohort currently provides one of the largest exome-sequencing resources. I exploited exome-sequencing data from the UK Biobank project and evaluated the contribution of rare genetic variance to the refractive error heritability in a sub-sample of 50,893 unrelated participants for whom measurements non-cycloplegic spherical equivalent were available. A detailed description of the study methodology, results and implications will be given in this chapter. The results of these analyses are currently under the review in the journal of *PLOS ONE* and will be published in the near future.

5.2 Methods

5.2.1 Study population

The UK Biobank is a cohort of half of a million volunteers for whom extensive demographic, phenotypic and biomarker data is available. A detailed description of UK Biobank cohort is provided in *Chapter 3* paragraph 3.2.1. Approximately 23% of the UK Biobank study subjects participated in comprehensive eye examination, described in Chapter 3 paragraph 3.2.The study was conducted with the approval of the North-West Research Ethics Committee (ref. 06/MRE08/65), following the principles of the Declaration of Helsinki. Study participants provided written informed consent.

5.2.2 Whole exome sequencing data

Exome sequencing was performed on 200,629 participants from the UK Biobank cohort as described in *Chapter 3* paragraph 3.2.1.6 (339). Briefly, the panel targeted 39 Mb of the human genome and covered 19,396 genes on autosomal and sex chromosomes, including 4,735,722 variants within the targeted regions. About 98% of the sequenced coding variants had an allele frequency below 1%.

5.2.3 The CREAM consortium dataset

CREAM was established in 2011 as a collaboration between studies with data on refractive error which had performed genome-wide association analysis based on SNP arrays. For the current study, 10 participating studies with available exome chip data were included. These studies included: Singapore Indian Eye Study (SINDI), Age-Related Eye Study (AREDS), Rotterdam Study I (RSI), Erasmus Rucphen Family Study (ERF), Estonian Genome Center of the University of Tartu (EGCUT), Finnish Twin Study on Aging (FITSA), Ogliastra, Croatia-Korcula, TwinsUK, Raine eye health study (REHS) and Beaver Dam eye study (BDES).

All cohorts had been genotyped on either the Illumina HumanExome-12 v 1.0 or v 1.1 array. All cohorts were jointly recalled to obtain a larger sample size of rare variants (here defined as variants with a minor allele frequency (MAF) ≤ 0.01), as recalling genotypes simultaneously across all samples increases the ability to call rare variants with a more discrete distinction between allele calls and sensitivity for low-frequency (high-intensity) loci. All data was recalled using GenomeStudio® v2011.1 (Illumina Inc., San Diego, CA, USA). Nine of these predominantly European CREAM cohorts were combined in a single cohort (N = 11,505), henceforth referred to as the CREAM cohort, for analysis. Legalities required that BDES (N = 1740) be analysed in the United States while the CREAM cohort had to be analysed in Europe, preventing BDES from being analysed jointly with the CREAM cohort.

5.3 Statistical analyses

5.3.1 Gene-based association analyses

To minimise confounding arising from population structure, the study sample was restricted to 50,893 unrelated UK Biobank participants of European descent. The ancestry and relatedness information was calculated based on the genetic data made available from the UK Biobank (340). Individuals with European descent were identified by projecting UK Biobank participants onto the coordination of 1000 Genome Project principal components. The genetic data was used to identify related individuals by calculating kinship coefficients for all pairs and third-degree or closer relatives were excluded.

Gene-based analyses were conducted in the optimised SNP-set kernel association test SKAT-O test (341) implemented in the rvtests package (278), as described in *Chapter 3* paragraph 3.3.3. The spherical equivalent measurements were the dependent variable and the weighed allelic burden the independent variables. All analyses were adjusted for age and sex. The analyses incorporated several variant annotations that previous works have shown to boost the power and accuracy of detecting causal associations in

gene-based analyses (279). Variants in protein-coding regions of genes including synonymous and non-synonymous, stop gain/loss, start gain/loss or splice-site mutations with minor allele frequency below 1% were selected for inclusion. The splicing sites were defined as 3 bases into exon and 8 bases into an intron. Mutations in these regions were annotated as "Normal_Splice_Site" unless they affected the functionally important "GU...AG" region of the intron which was annotated as "Essential_Splice_Site". UTR variants and polymorphisms with unknown or inconclusive molecular consequences such as intronic variants were excluded. Variants were identified and annotated using the ANNO package (https://github.com/zhanxw/anno). The replication of significant genes was sought using the results of the gene-based analysis performed in the predominantly European CREAM cohort (31) and the BDES cohort (342,343). Replication was considered successful if the association probabilities were below the selected Bonferroni multiple testing correction level.

I additionally performed gene-based analyses that included deleterious variants annotated with ANNOVAR and selected based on SIFT scores, ranging between 0.0 to 1.0. Markers with respective SIFT scores of 0.0 – 0.05 are considered to be deleterious. Accordingly, rare variants (MAF< 1%) with respective SIFT scores of <=0.05 were advanced for gene-based analyses, carried out in SKAT-O test (341) implemented in the rvtests package (278). The spherical equivalent measurements were used as the dependent variable and the weighed allelic burden as the independent variables. All analyses were adjusted for age and sex.

5.3.2 Sequential analyses evaluating the role of single variants in gene-based associations

To elucidate which variants were driving observed associations with candidate genes, I performed sequential sensitivity analyses by progressively removing markers from the gene-based analyses. The associations with target genes were assessed using the SKAT-

O test adjusted for age, sex and lead common variants within the same locus. The lead common variants were selected from previously published refractive error GWAS (99).

5.4 Results

The final study sample included 50,893 unrelated UK Biobank participants of European descent; 54% were women with a median age of 57 years (±8 years). Detailed information about the study participants' demographic and refractive characteristics can be found in Table 7 and characteristics of the spherical equivalent distribution in Figure 9.

Age (mean (SD))	56.8 (7.9)
Sex (N, %)	
Women	27,221 (53.5)
Men	23,672 (46.5)
SE (mean(SD))	-0.3 (2.7)
Refractive status (N,%)	
Emmetropia	23,193 (32.9)
Hyperopia	13,952 (33.8)
Myopia	13,748 (33.3)

Table 7 Characteristics of the study participants

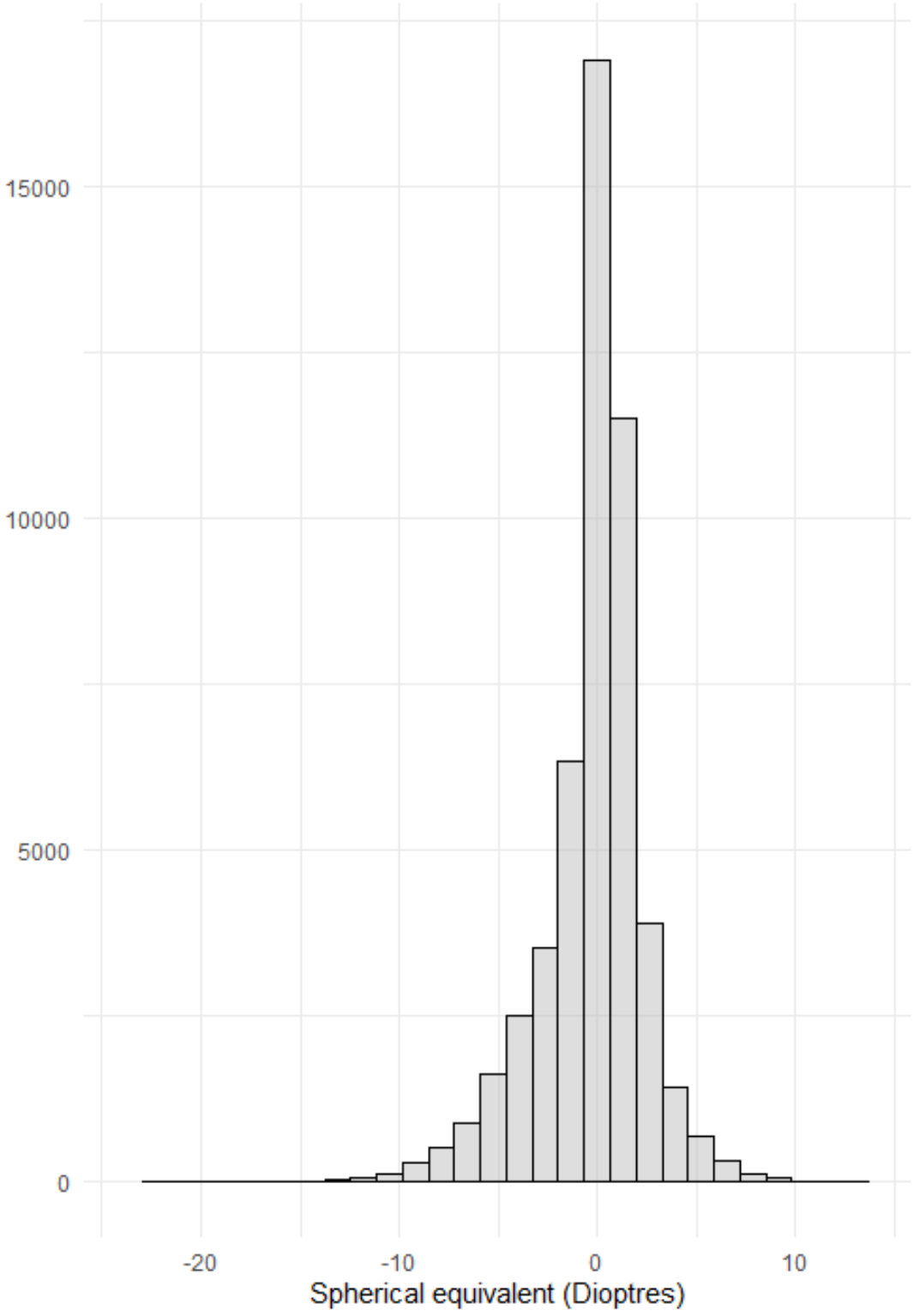


Figure 9 Spherical equivalent distribution in UK Biobank cohort (N = 50,893). The distribution of the spherical equivalent (x-axis) in the samples; the number of participants for each spherical equivalent bin is given in the y-axis.

SKAT-O tests were performed using 2,923,839 rare (minor allele frequency, MAF < 0.01) variants in 19,293 genes (Figure 10). The statistically strongest association was observed between spherical equivalent (SPHE) and *SIX6* gene (p-value = 2.15 x10^{-10}). The second Bonferroni-significant association was found with *CRX* (p = 6.65 x 10^{-08}). This finding was novel and not described in prior refractive error GWAS. Suggestive statistical evidence of association was found for the *RPSAP52* (p = 1.65 x 10^{-05}), *PCCA* (p = 1.82 x 10^{-05}), *MIR4683* (p = 2.81 x 10^{-05}), *SELENOM* (p = 3.52 x 10^{-05}), *NAPA* (p = 4.55 x 10^{-05}) and *VWA8* (p = 5.68 x 10^{-05}) genes, whose association however did not meet the criteria of statistical significance after multiple testing correction (Table 8). Additionally, I performed gene-based analyses, including deleterious variants with minor allele frequency < 0.01. These analyses revealed Bonferroni-significant associations with *SIX6* (p = 4.05 x 10^{-06}), *ZFP1* (p = 7.65 x 10^{-08}) and *C1GALT1* (p = 3.06 x 10^{-07}) (Table 9).

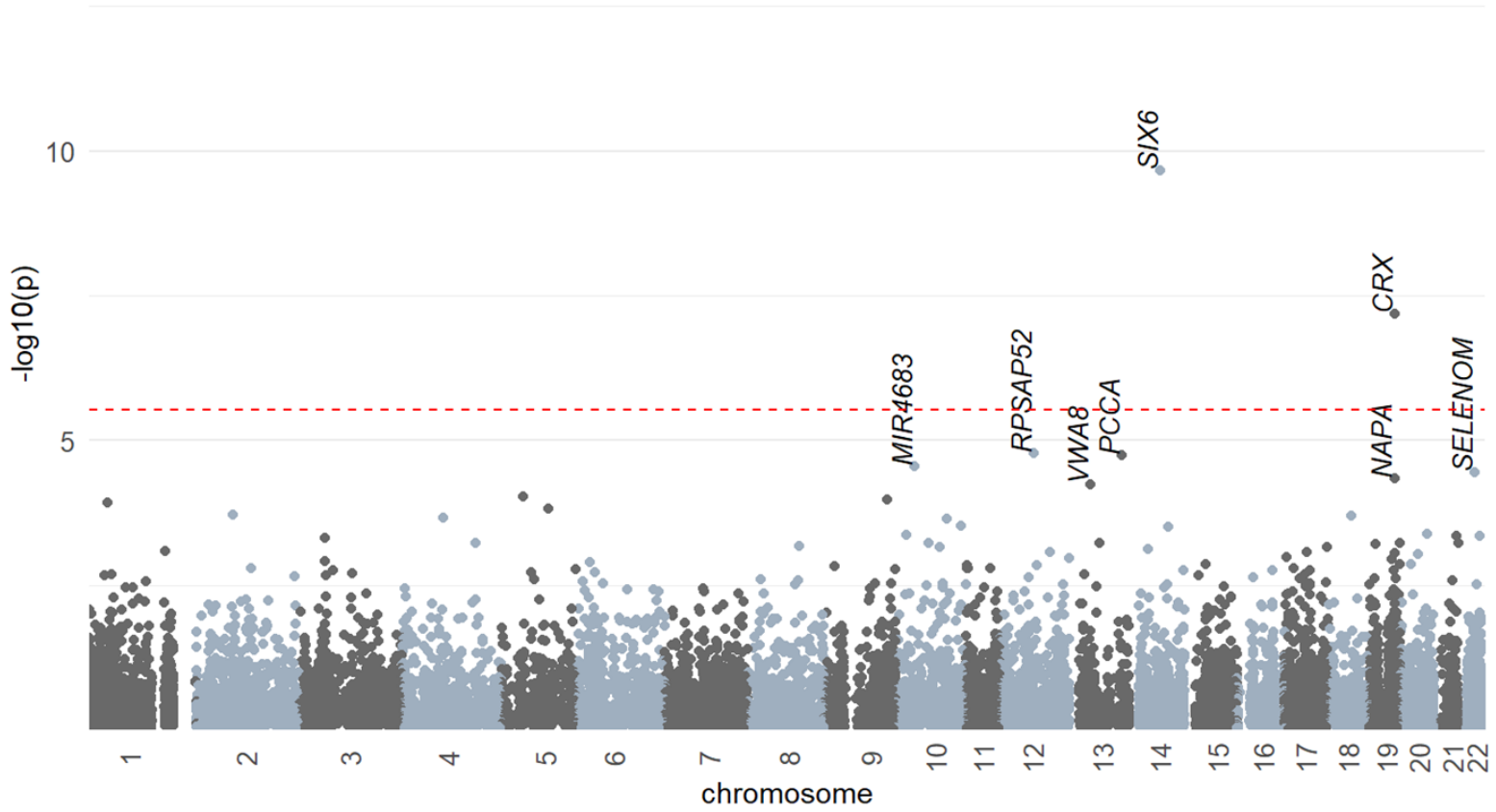


Figure 10 Manhattan plot displaying SKAT-O association results: each point represents one of the 19,293 genes tested for the association with the spherical equivalent in the UK Biobank cohort (N=50,893). The plot shows —log10 transformed p-values for each gene plotted against the chromosomal location. The red dashed line indicates the Bonferroni significance threshold ($p < 2.59 \times 10^{-6}$). Regions are named with symbols of genes that were most strongly associated with refractive error.

Gene symbol	Genetic coordinates	Number of variants	Top SNP at the locus	SIFT score	A1	A1 freq	Beta	p-value	Bonferroni- corrected p-value
SIX6	14:60509145-60512849	67	rs146737847	0.02	A	0.007	-0.86	2.15×10^{-10}	3.65×10^{-06}
CRX	19:47821936-47843324	92	rs61748438	1	A	0.004	0.71	6.65×10^{-08}	0.001
RPSAP52	12:65758020-65826974	6	12:65825320:G:T	0.14	T	3.93×10^{-05}	-6.04	1.65×10^{-05}	0.28
PCCA	13:100089014-100530437	215	rs61749895	0.51	T	0.0097	-0.35	1.82×10^{-05}	0.31
MIR4683	10:35641171-35641252	4	rs1363993147	0.19	G	9.82×10^{-06}	-11.70	2.81×10^{-05}	0.48
SELENOM	22:31104776-31107568	50	22:31105965:T:C	0.14	C	1.97×10^{-05}	-8.28	3.52×10^{-05}	0.6
NAPA	19:47487636-47515063	90	19:47488308:C:T	0.01	T	9.82×10^{-06}	-8.99	4.55×10^{-05}	0.77
VWA8	13:41566834-41961109	583	13:41891421:C:T	1	T	9.82×10^{-06}	-14.55	5.68×10^{-05}	0.96

Table 8 Top eight gene associations with refractive error. Gene-based analyses included synonymous and potential loss of function variants. Column "Gene symbol" lists the symbols of the genes associated with spherical equivalent. Fields "Genetic coordinates", "p-value" and "Bonferroni corrected p-value" include genetic coordinates (reference genome hg38) of the tested genes, and denote p-values and Bonferroni-corrected p-values for the respective associations. The column "Number of variants" includes the number of tested genetic variants in each respective gene. The field "Top SNP at the locus" lists variants with the statistically strongest associations at the locus. The columns "A1" denotes the effect allele for which betas (Beta) were calculated. The field "A1 freq" shows the frequency of the effect alleles in the study sample. The column "SIFT score" includes SIFT scores for top SNPs at the locus.

Gene symbol	Genetic coordinates	Number of variants	Top SNP at the locus	SIFT score	A1	A1 freq	Beta	P-value	Bonferroni- corrected p-value
SIX6	14:60509145-60512849	25	rs146737847	0.02	A	0.0076	-0.86	2.15×10^{-10}	3.94×10^{-06}
ZFP1	16:75148523-75172234	36	rs748117598	0.025	A	3.93×10^{-05}	-6.51	7.65×10^{-08}	0.001
C1GALT1	7:7182543-7248615	26	rs765265066	0.009	T	5.89×10^{-05}	-5.89	3.06×10^{-07}	0.006

Table 9 Top three gene associations with refractive error. Gene-based analyses were restricted to deleterious mutations. Column "Gene symbol" lists the symbols of the genes associated with spherical equivalent. Fields "Genetic coordinates", "p-value" and "Bonferroni corrected p-value" include genetic coordinates (reference genome hg38) of the tested genes, and denote p-values and Bonferroni-corrected p-values for the respective associations. The column "Number of variants" includes the number of tested genetic variants in each respective gene. The field "Top SNP at the locus" lists variants with the statistically strongest associations at the locus. The columns "A1" denotes the effect allele for which betas (Beta) were calculated. The field "A1 freq" shows the frequency of the effect alleles in the study sample. The column "SIFT score" includes SIFT scores for top SNPs at the locus.

No other exome sequencing datasets of comparable size were available for replication. However, two smaller cohorts were genotyped for a selection of rare exonic variants using SNP chip arrays, in both of which data from only four of the novel candidate genes, including *SIX6*, *CRX* and two genes associated with SPHE at suggestive levels in the UK Biobank analysis were available. In addition, only 2 and 3 variants were present in the exome chip data for the *SIX6* and *CRX* genes respectively, that had demonstrated the statistically strongest relationship with spherical equivalent in the discovery cohort. None of these genes was associated at statistically significant levels with SHPE in the pooled exome chip cohort from the Consortium for the Refractive Error and Myopia (CREAM, N = 11,505), but a strong association for the *NAPA* gene was observed (SKAT-O p =3.73 x 10^{-05}) in the smaller Beaver Dam Eye Study (Table 10).

		Beaver Dam (N = 1740)		*CREAM (N = 11,505)*	
Gene symbol	**Genetic coordinates**	**Number of variants**	**p-value**	**Number of variants**	**p-value**
PCCA	13:100089014-100530437	6	0.6	3	0.7
SIX6	14:60509145-60512849	2	0.2	2	0.1
NAPA	**19:47487636-47515063**	**1**	**3.73 x 10^{-05}**	6	0.2
CRX	19:47821936-47843324	2	0.8	3	0.4

Table 10 Replication of four loci associated with refractive error using gene-based analyses performed in Beaver Dam (N = 1740) and CREAM Consortium dataset (N = 11,505). Column "Gene symbol" lists the symbols of the genes associated with spherical equivalent. Fields "Genetic coordinates", " p-value" include genetic coordinates of the tested genes and denote p-values for the respective associations in Beaver Dam and predominantly European CREAM replication cohorts. The column "Number of variants" includes a number of tested genetic variants in each respective gene. The associations that had p-values below Bonferroni multiple testing correction threshold are shown in bold letters (0.05/4 = 0.01)

To identify independent variants driving gene-based associations at the *SIX6* and *CRX* loci, sensitivity analyses were performed by progressively removing SNPs from gene-based analyses. The removal of rare variants from the gene-based SKAT-O analyses revealed a decrease in the statistical significance of the analyses. The results of these analyses suggested that association with the *SIX6* gene was most strongly influenced by the rs146737847 variant, whose removal resulted in the loss of statistical significance in the study samples (Figure 11). Moreover, missense marker rs146737847 introduced p.Glu129Lys substitution in the *SIX6* gene and was predicted to be deleterious and affect protein function (SIFT score= 0.02). Similarly, exonic marker rs61748438 was identified as a single lead variant in the *CRX* locus (Figure 12). This particular missense variant initiated p.Val66Ile alteration in the *CRX* gene, and was predicted to be benign, showing respective SIFT score of 1. The removal of other functionally important variants within this gene also resulted in a progressive decrease in statistical significance. This gradual decrease may suggest that although gene-based association at both loci is mostly due to the presence of a few lead variants, additional lower frequency variants within these genes may also contribute to associations with spherical equivalent, but the abilities to fully evaluate their role at a general population level may be constrained by sample size and statistical power limitations.

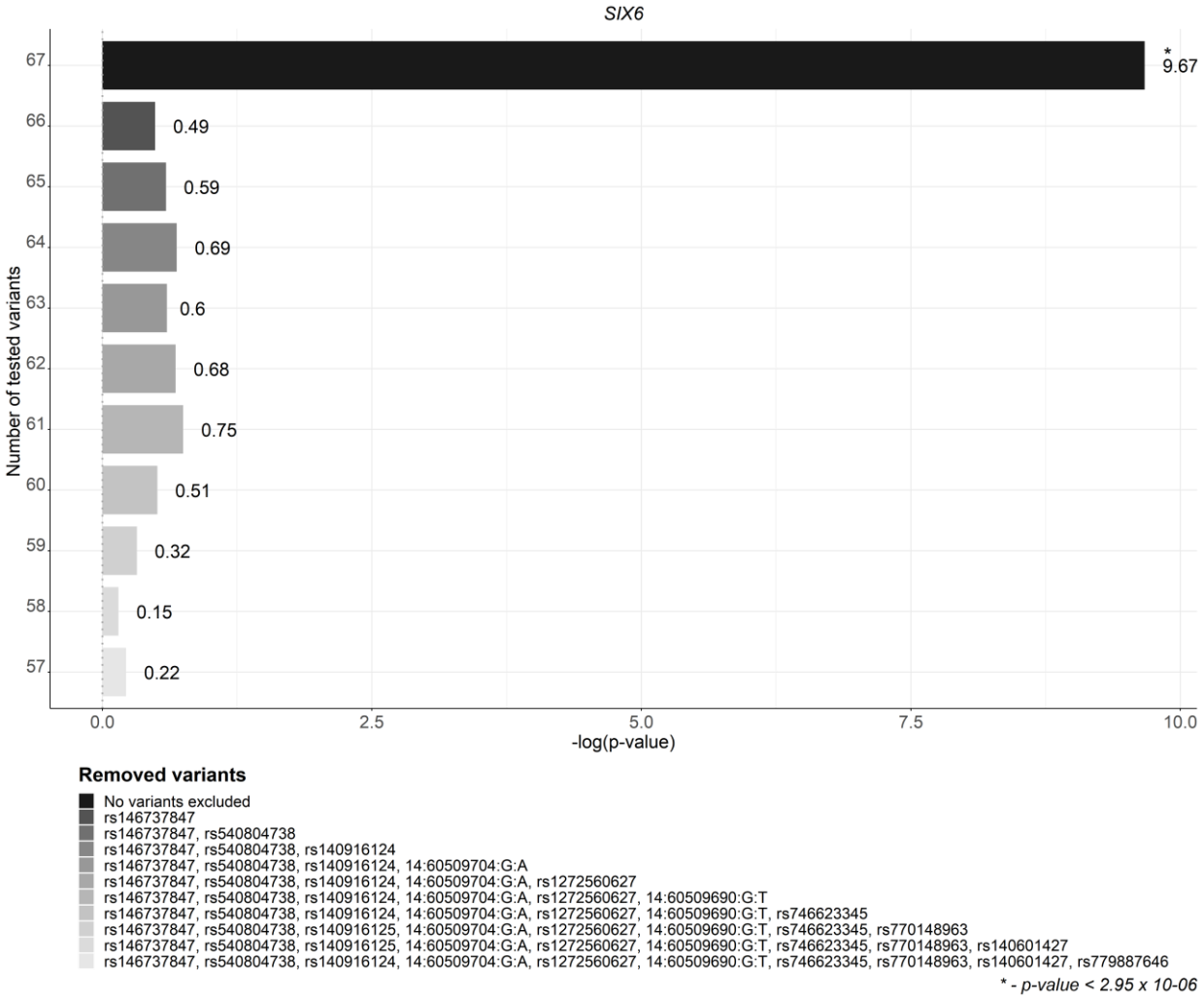


Figure 11 Sensitivity analyses result for *SIX6* gene. Y-axis shows the number of *SIX6* variants included in gene-based analyses, testing associations with SPHE. The model was adjusted for age, sex and the best common variant within the same locus. The -log(p-values) from SKAT-O tests are displayed on X-axis.

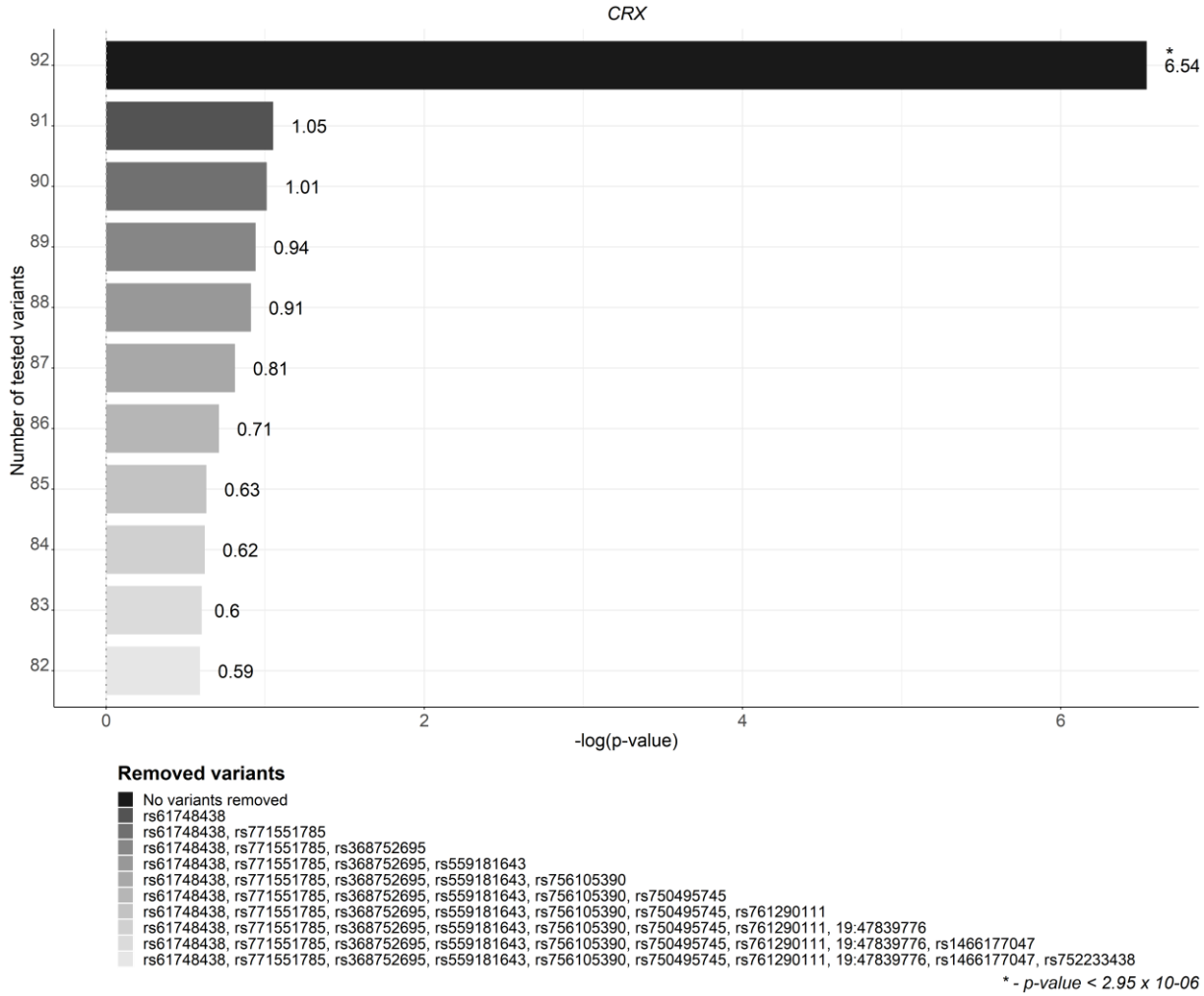


Figure 12 Sensitivity analyses results for *CRX* gene. Y-axis shows the number of *CRX* variants included in gene-based analyses, testing associations with SPHE. The model was adjusted for age, sex and the best common variant within the same locus. The -log(p-values) from SKAT-O tests are displayed on X-axis.

5.5 Discussion

Here I report significant associations between spherical equivalent and rare variants located within *SIX6* and *CRX*, but also *RPSAP52*, *PCCA*, *MIR4683*, *SELENOM*, *NAPA* and *VWA8* genes. In this study, the strongest association was observed with the *SIX6* gene, located on *14q23.1* and which encodes a homeobox protein involved in ocular development (344), morphogenesis and visual perception (345). The SIX Homeobox 6 (*SIX6*) is part of a group of evolutionarily conserved genes, known eye transcription factors (346), which regulate the proliferation of specific retinal cells during optic disc development (346) and retain their importance in the mature retina (346). *SIX6* is implicated both in the early stages of eye formation and subsequent differentiation of retinal progenitor cells (RPC). Interestingly, previous works have shown that deleterious missense variant rs146737847 (Glu129Lys; SIFT score = 0.02) adversely affects the *SIX6* gene function (347) and is also associated with primary open-angle glaucoma potentially through its known effect over the vertical cup-disc ratio (348). While observational correlation between glaucoma and myopia status is well known (349,350) there is little evidence of large-scale shared allelic risk between spherical equivalent and vertical cup-to-disc ratio (99). The associations observed with both spherical equivalent and glaucoma phenotypes for the rs146737847 suggest that genetic pleiotropy may explain a considerable proportion of the phenotypic correlation between these two common ocular conditions.

The Cone-Rod homeobox gene, or *CRX,* located on *19q13.33*, encodes a photoreceptor-specific transcription factor (351). Although a previous association with refractive error was detected for the broader chromosomal location, this is the first time that direct evidence links this gene with spherical equivalent or myopia. This gene is a master regulator of photoreceptor development (352) and differentiation (353). Certain mutations in this gene cause several retinal disorders, including cone-rod dystrophy, retinitis pigmentosa, adult-onset macular dystrophy and Leber congenital amaurosis

(351,354). Additionally, the Cone-Rod Homeobox (*CRX*) Transcription Factor regulates the expression of over 700 genes in the retina, including downstream effects over rhodopsin and cone arrestin (355). *CRX* expression in the retinal cells was inhibited by light stimulation, a mechanism previously implicated in myopia development (356,357).

Additional Bonferroni-significant associations were identified with deleterious markers located within *ZFP1* and *C1GALT1* genes. *ZFP1* encoded a member of zinc finger protein family that attached to DNA via zinc-mediated secondary structures, zinc fingers, participating in transcriptional activity (358). *ZFP1* gene harboured polymorphisms that were linked to advanced age-related macular degeneration (359) and brain connectivity (360). Deleterious mutations within zinc finger family genes impacted gene expression in the retina and RPE and were implicated in eye development and axial elongation (361). Similarly, polymorphisms within or near *C1GALT1* were reported as associated with refractive error in previous GWAS (99). The gene coded for protein generating common core 1 O-glycan which was precursor for several of extended mucin-type O-glycans on the cell surface and secreted glycoproteins. Core 1 O-glycan was a main component of mucin transmembrane at ocular surface and contributed to the maintenance of epithelial barrier (362).

Suggestive associations with rare variants located within other genes and refractive error were also detected. In particular, the study analyses implicate the *NAPA* and *PCCA* genes. Common polymorphisms at genomic loci encompassing these genes are associated with refractive error (99) and the age of refractive correction (110), but this is the first time that, rare variants within their coding regions are associated with SPHE. *PCAA* encoded the biotin-binding region of mitochondrial Propionyl-CoA carboxylase involved branched and odd chain fatty-acid and cholesterol catabolism (363). The protein product of the *NAPA* gene is a member of the soluble NSF attachment protein family aiding the fusion and docking of vesicles to target membranes (364). *NAPA* also participates in synaptic activity and plays a role in neurogenesis (364). Notably, this

particular gene was the only candidate that achieved replication in an independent dataset.

Several novel potential candidates for which I found suggestive evidence of association are implicated in cognitive development and learning difficulty disorders. In particular, Ribosomal Protein SA Pseudogene 52 (*RPSAP52*) is associated with brain structure variations in TWAS (365) and described in genetic investigations of cognitive impairment, neurodevelopmental and neurodegenerative disorders (366). The polymorphisms within the *RPSAP52* gene were associated with schizophrenia in founder populations (367) and associated with biomarkers of Alzheimer's disease, such as cerebrospinal fluid beta-amyloid 1-42 levels (368). Similarly, mutations in microRNA *MIR4683* may be associated with epilepsy in children (369). Another interesting candidate *VWA8* encodes a mitochondrial ATPase, whose precise function is not fully understood (370). Genome-wide association studies demonstrated that variation in *VWA8* may influence susceptibility to autism (371) and bipolar disorder (372), and also educational attainment and mathematical ability (373). *SELENOM,* another novel candidate, encoded a selenoprotein that is highly expressed in the brain and that is thought to be essential for normal neurocognitive development (374).

For this study, I used some of the largest ever sample sizes analysed to date to assess the role of rare variants in refractive error. However, this study has several limitations. Gene-based analyses assumed a simple dominant model of inheritance was used, while recessive or compound heterozygosity models of inheritance may also play a role in refractive disorders. The analyses were also restricted to the coding regions of the genome. However, non-coding areas of the genome also proved to be important for several other diseases (375–377) and could potentially provide a new direction for additional myopia work. In my study, chip array information was used in two replication datasets and found relatively little evidence for replication. However, the arrays only include a small number of variants within the exome of the genes of interest and none of which was a particularly strong contributor to the overall gene-based association in

the discovery data. Given that 95,376 typed SNPs were advanced for imputation in the BDES, low-density of the genotyping array could lead to suboptimal imputation accuracy, especially for rare markers which are harder to impute. It has been demonstrated that significant discrepancies between imputed and observed genotypes could potentially cause inflation of the association results, especially for variants with high uncertainty (378). Consequently, although the association with the *NAPA* was significant in BDES cohort, this particular gene contained only one variant. Therefore, I cannot exclude possibility that this particular association was a potentially false-positive finding. Additionally, CREAM studies were genotyped separately but recalled jointly to improve calling accuracy for rare makers. However, this was a conservative approach which significantly reduced the number of analysed variants. Performing gene-based analyses separately in each CREAM cohort and meta-analysing results from each study would have increased statistical power. These are general limitations of array-based studies, and to fully validate the results described in this thesis, future work on large scale exome sequencing on independent cohorts will be needed. Lastly, despite the large sample size, statistical power for rare variants is often limited due to the very low allele frequencies. Power will benefit from additional sequencing data from the several national cohorts and biobanks whose data will become available in the future.

This study demonstrates that variants with significantly large effects on refractive error are extremely rare. The study analyses show strong evidence of association between population refractive error and rare variants located within the protein-coding regions of the *SIX6* gene, which plays an important role in eye morphogenesis and is implicated in several ocular disorders, including myopia. Additional significant association is identified for the *CRX* gene, a transcription factor crucial for the development of photoreceptors, as the origin of an important association signal. Current UK Biobank WES presented in this thesis suggests that high-quality whole-exome sequencing provides a superior alternative to array-based methods that have power limitations and are prone to bias arising from population admixture (379). Beyond novel associations,

the incorporation of rare variants in existing myopia risk prediction models that currently rely on common polymorphisms will improve their accuracy and augment our understanding of refractive disorders.

Chapter 6 | Genetic associations with the age of refractive correction

6.1 Introduction

The aetiology of refractive errors is believed to be complex and multifactorial (14). In particular, postnatal ocular growth is guided by visual stimuli that induce a signalling cascade in the retina. The initiated in the retina signals pass through the retinal pigment epithelium (380) as well as the choroid (381) and regulate scleral extracellular matrix remodelling whose alterations were shown to promote axial elongation and myopia (382).

At birth, the length of the human eye measures at 17 mm and continuously grows through childhood and adolescence, reaching 21 – 22 mm at adulthood (383). By the age of 5, only 2% of children develop myopic refraction. By contrast, myopia incidence increases sevenfold during adolescence with peak growth observed between 9 and 14 years (384). Childhood-onset myopias progress and become worse during puberty, but stabilise by 30 (108). Conversely, myopia with adult-onset tends to be less severe (385) and the age of the distance correction generally reflects the severity of refractive disorders. For example, of those who started wearing glasses between 3-10 years, 25.5%, 38.5% and 35.9% respectively develop mild, moderate and severe myopic refractive error, whereas in the group with late-onset refractive error almost all individuals have mild myopia (385).

Environmental factors, such as educational attainment (132) and time spent outdoors, vastly influence the development and progression of myopia. Their effects depend on lifestyles and cultural trends, but they typically affect whole cohorts across countries and societies (58) sharing similar living environments. Within a society at any given time, the environmental exposures are stable and relatively homogenously distributed, and heritable factors explain over half of the spherical equivalent and risk to refractive error (70). Several genetic studies conducted in the general population have identified DNA variations associated with the risk of refractive error (110,386) and age of myopia onset (260). Genes associated with the age at first refractive error correction usually overlap

with those associated with spherical equivalent (387), and both predispose to pathological myopia (388). Yet, the timing of the individual genes' effects is not evenly distributed throughout the childhood years or lifetime. Different genes have varying strength of effect and association throughout the years and among the genetic factors associated with spherical equivalent, some genes predispose to earlier refractive correction than others (389). Additionally, there is considerable genetic pleiotropy in the eye and the same genetic factors may be independently be associated with several endophenotypes (390), each with a potential to alter the age in which correction of refractive errors is needed.

Prior investigations reported a strong genetic correlation between final refractive status and myopia age-at-onset which shared roughly 93% of their genetic susceptibility (110). Moreover, genetic variants associated with both traits were concordant in the direction of the effects and displayed similar effect sizes (110). For these common markers, a 10% decrease in ln(HR) of the age of refractive correction resulted in a 0.15 dioptre decrease in spherical equivalent (110). To date, a handful of genetic investigations examined genetic determinants of the age of myopia onset. In 2013 among the first Kiefer et al reported 22 significant genetic associations that predisposed to earlier myopia in Europeans (108). Following this study, Tideman and colleagues compared the genetic effects of 39 myopia-associated loci in children and adults of mixed ancestry, finding associations between axial length and three loci in the youngest age group (389). Additional four associations were detected in the subsample aged 10-25, while in older adults 20 distinct genomic regions were associated (389). The results of this study indicated that some refractive error genes exhibited the greatest effects in childhood while others were expressed in adolescence or during the entire lifespan of myopia development (389). By contrast, in a study of Asian children, 16 refractive error loci demonstrated evidence of myopia onset in early childhood, whilst later-onset associations arising between 8 and 15 years were discovered for 11 genes (391).

I further explored genetic factors contributing to the early-onset refractive disorders by conducting genome-wide time-to-event analyses of the age of the first spectacle wear in the large subsample of UK Biobank participants. This study aimed to further analyse the relationship between the age of refractive correction and population spherical equivalent. The following chapter reports on the study methods, statistical analyses, results and implications. The results of these analyses are currently under the review in the journal of *Human Molecular Genetics* and will be published in the near future.

6.2 Methods

6.2.1 Study population

The UK Biobank cohort is a large population-based longitudinal study of half of a million volunteers from across the United Kingdom. A detailed description of the cohort is provided in *Chapter 3* paragraph 3.2.1. At the baseline assessment, UK Biobank study participants completed electronic questionnaire which contained several eyesight-related inquires, including the questions about the age of first spectacle wear (The UK Biobank field number: 2217) and reasons for refractive correction (Field number: 6147)(230). About 23% of all UK Biobank volunteers also undertook ophthalmic examination (231), as described in Chapter 3 paragraph 3.2.1.1. Participants who had eye surgery, infection, bilateral eye injury before the assessment or self-reported cataracts with mild myopia were excluded from the subsequent analyses (231). To minimise confounding arising from population genetic structure, the study sample was restricted to individuals of European ancestry, as ascertained by using genetic information. Ancestry was defined based on Principal Component Analyses of the participants' genotypes, pre-computed and calculated by the UK Biobank working group.

6.2.2 Genetic data

Genotyping was performed on 488,377 subjects from the UK Biobank cohort as described before in *Chapter 3* paragraph 3.2.1.5 (230) using two similar and mutually compatible genotyping platforms (Applied Biosystems UK BiLEVE Axiom Array and the UK Biobank Axiom Array), which although not fully identical, shared approximately 95% of genetic markers. The study analyses used a subset of unrelated UK Biobank participants of European descent, for whom information about the refractive error and the age of refractive correction was available. Specifically, spherical equivalent analyses were conducted in N = 102,117 subjects, the all-causes age of spectacle wear in N = 340,318 subjects, age of spectacle wear in individuals with myopia in N = 24,363 and individuals with hypermetropia in N = 24,711 subjects.

Phasing and further genomic imputation were conducted as described in *Chapter 3* paragraph 3.2.1.5 (230). Briefly, imputation was carried out using Haplotype Reference Consortium (HRC) data as a primary reference panel and merged 1000 Genomes phase 3 and UK10K reference panels. Only markers shared between HRC, and 1000 genomes/UK10K datasets were selected for imputation; therefore, a final dataset covered 93,095,623 autosomal SNPs in conjunction with large structural variant indels (230).

6.2.3 Statistical analyses

Descriptive analyses were carried out using the *epiDisplay* package in R. I calculated frequencies and percentages and means and standard errors for categorical and continuous variables. For the time-to-event genetic association analyses, I built Cox proportional hazards regression model adjusted for age and sex, as described in *Chapter 3* paragraph 3.3.4. The likelihood ratio test was used to compute p-values for each SNP in the model. I used two R packages, *gwasurvivr* (282) and *SPACox* (281), to calculate hazard ratios (HR) and their corresponding p-values. HR represented the multiplicative change in the rate of onset of refractive correction per copy of the tested allele. The

genetic variants with p-values below the customary genome-wide significance level of 5 x 10^{-08} were considered statistically significant. The proportionality of the hazards for significant associations was assessed using the *survival* package in R (https://cran.r-project.org/web/packages/survival/). The replication of the novel genetic associations was sought using time-to-event results previously published by Kiefer *et al.* (108). Replication was considered significant if the association probabilities were below the Bonferroni multiple testing correction level (observed p-value multiplied by the number of tests no higher than 0.05). The genomic inflation arising from sample stratification and uncontrolled admixture was tested using LD score regression (283).

Data from 45,771 research volunteers recruited among the customer base of the *23andMe* genomics company (Sunnyvale, CA, USA) were used for replication. More detailed information can be found in *Chapter 3* paragraph 3.2.2 (392), but briefly, the phenotypic status was ascertained through an online medical history questionnaire or an eyesight questionnaire. Participants were genotyped and additional SNP genotypes were imputed against the 1000 genomes data and the imputed genotypes from individuals of European ancestry were used for Cox proportional hazards models.

I assessed genetic correlations between AFSW and other phenotypic traits using LD-score regression (285) and the summary statistics from GWAS Catalog (284).

The shared functionality of associated genes was further explored through gene-set enrichment analyses, as implemented in MAGENTA software (312). The relationship between genotypes and gene expression was modelled using Mendelian Randomisation tests implemented in the SMR program (299), using expression data from GTEx release v8 (https://gtexportal.org/home/datasets), the Atlas of the Developing Human Brain (393) (BrainSpan 11) and retinal cis-eQTL data from healthy donors (394).

6.3 Results

The final study sample included 340,318 UK Biobank participants of European ancestry who reported the AFSW in the electronic questionnaire; of them, 46% (N = 156,388) were men. The mean age at recruitment was 58 years (± 7.5 years). The more detailed information on the study subjects' demographic characteristics and refraction can be found in Table 11.

Variables	**Statistics**
Age (mean, SD)	57.7 (7.5)
Sex (%, N):	
female	54% (N = 183,930)
male	46% (N = 156,388)
SPHE (mean, SD)	-0.3 (2.9)
Reasons for wearing glasses/contact lenses (%, N):	
Myopia	11.1% (N= 37,762)
Hypermetropia	5.6 % (N = 19,178)
Presbyopia	9.1% (N = 31,137)
Astigmatism	0.7% (N = 2246)
Strabismus	0.2% (N = 783)
Amblyopia	0.3% (N = 1173)
Other eye condition	0.2% (N = 788)
Unknown	72.6% (N = 247,251)

Table 11 Demographic and refractive characteristics of the study participants. The values shown here are calculated for the subset of the UK Biobank (N = 340,318) that responded to the questionnaire on the reason for their spectacle or contact lens correction and not for the entire cohort.

The age of the first spectacle wear followed bimodal distribution with the first mode between 1 – 35 years, peaking at the age of thirteen, and the second mode between the ages of 36 – 72 years with a peak at the age of 43 (Figure 13). Participants who started wearing glasses/contact lenses before 35 tended to be more myopic, while subjects with AFSW over 35 years were more likely to have hypermetropia (Figure 14). Of those reporting AFSW, 11% (N = 37,762) specified myopia as a reason for wearing glasses/contact lenses. Presbyopia (9.1%; N = 31,137) and hypermetropia (5.6% (N =

19,178) were the second and the third most commonly self-reported reasons for refractive correction. Moreover, 0.7%, 0.3%, 0.2% and 0.2% of participants wore glasses/contact lenses to correct astigmatism, amblyopia, strabismus or other eye conditions, respectively. The remaining 72.6% of subjects did specify the reason for wearing glasses.

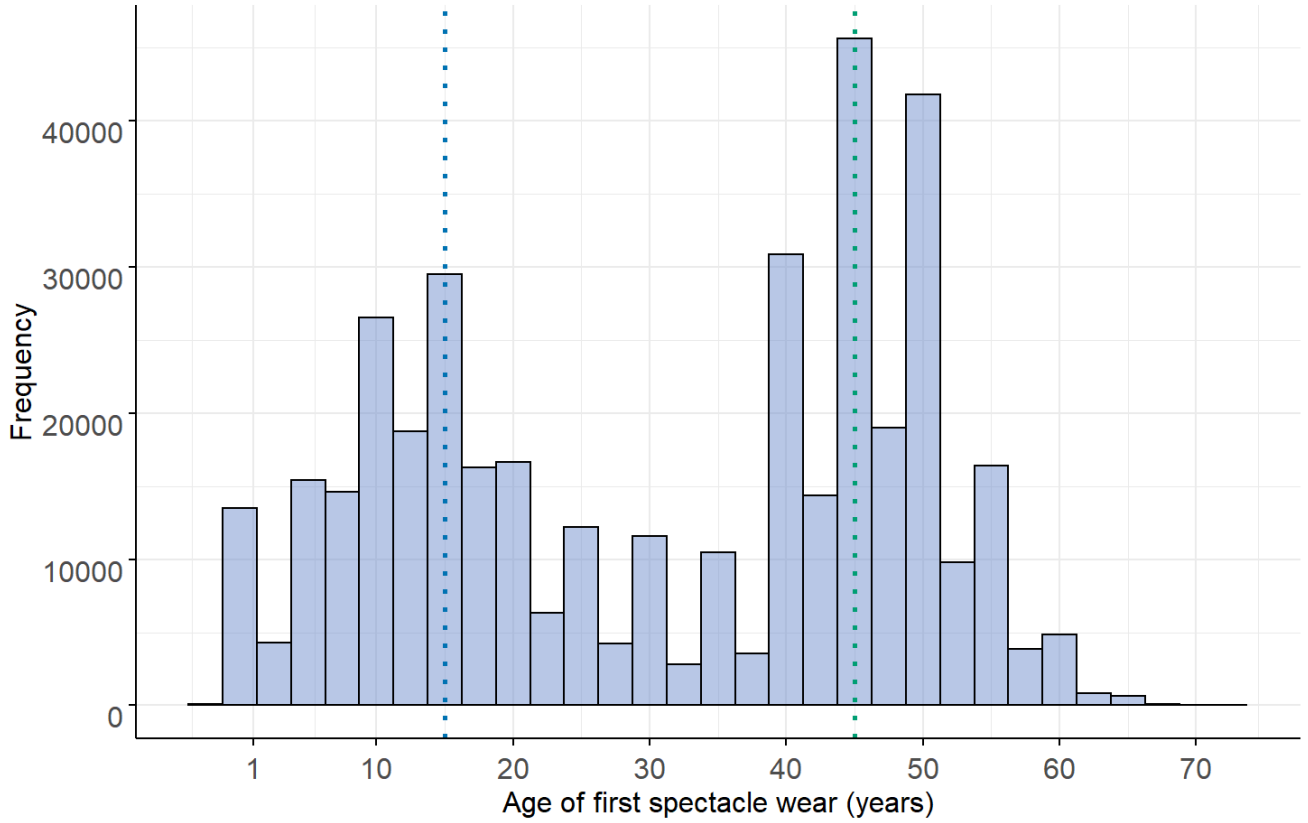


Figure 13 Distribution of the age of first spectacle wear in UK Biobank Cohort. The dash lines represent medians.

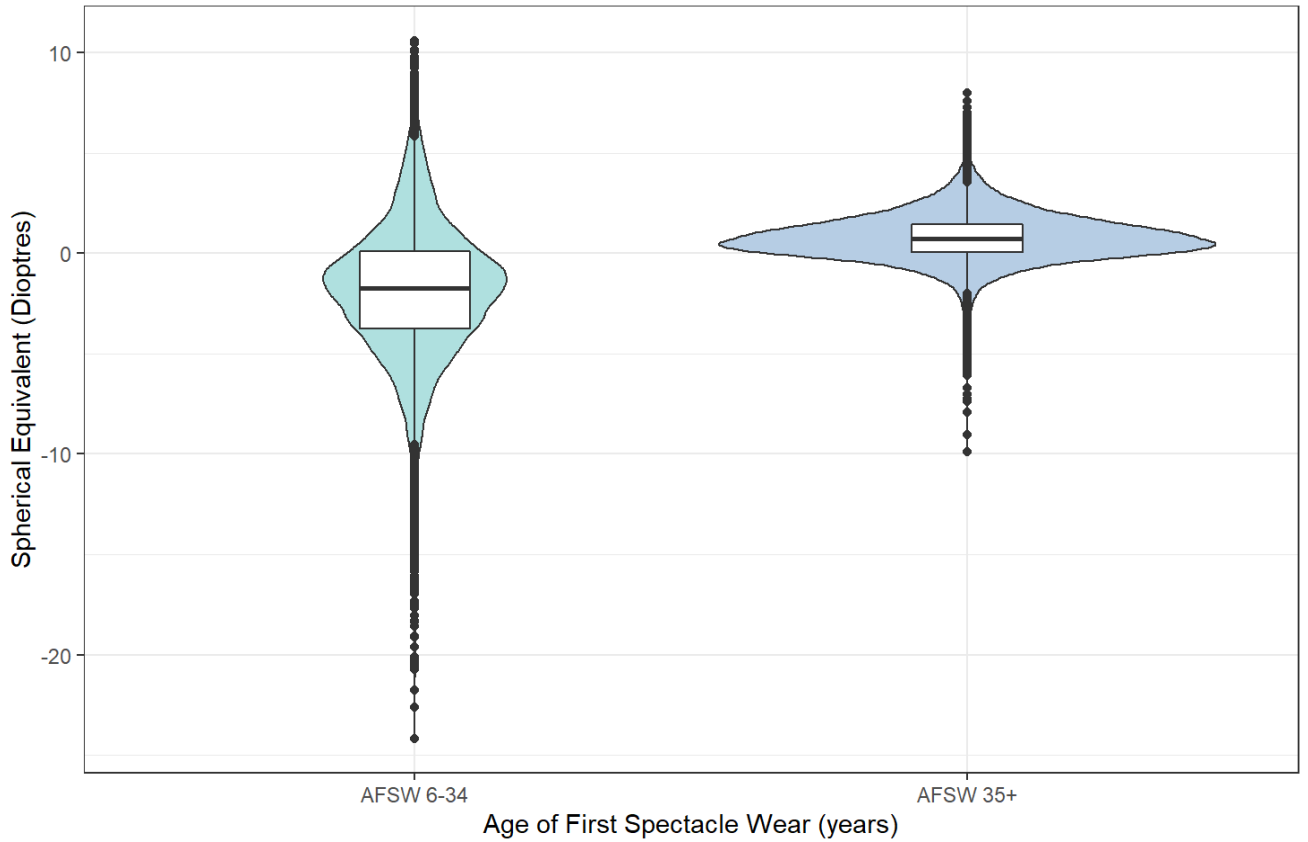


Figure 14 Distribution of refractive error across the age of first spectacle wear in UK Biobank cohort (N = 81,675). The horizontal axes display two main modes of the age of spectacle wear. The shape of the violin plots shows the distribution of refractive error within the respective group. The central horizontal bar denotes the mean spherical equivalent (vertical axis) across all participants in the respective categories; the box displays the interquartile range, and the whiskers represent the observations within 1.5 × Interquartile Range.

The large study sample size (N= 340,318) resulted in a high genomic inflation factor (λ = 1.23) in the time to lens or spectacle correction. However, the low intercept of the linkage disequilibrium regression 0.93, and (intercept-1)/ (mean (X2) − 1) ratio (-0.19, SE= 0.02), reassuringly indicate a conservative control for potential confounding in this study.

I first assessed the degree of similarity between the genomic architectures of the spherical equivalent, self-reported age of first lens or spectacle correction for myopia, self-reported correction for hypermetropia and self-reported first correction for any reason. Consistent with previous reports, I found a strong genetic correlation between the first myopia correction and spherical equivalent (rg=-0.97, Table 12). Additionally, it was also noted that the age of the first correction in participants with myopia alone was strongly genetically correlated with the age of the first correction of any refractive error (rg = 0.89, Table 12) and less so with the age of the first correction among hypermetropic subjects (rg = -0.65, Table 12). Spherical equivalent and all-cause AFSW shared most of their heritability and were significantly correlated (rg = -0.68, $p = 9.6 \times 10^{-171}$). Because of the strong correlation and the phenotypic information was available for considerably more individuals with all-cause AFSW information than any other phenotypic definition, I focused this work on the analysis of the all-causes AFSW, because of the expectation of superior statistical power.

	AFSW all	**AFSW, myopia only**	**AFSW, hyperopia only**
Spherical Equivalent	-0.683	-0.968	0.808
AFSW all		0.889	-0.085
AFSW, myopia only			-0.651

Table 12 Genetic correlations between SPHE GWAS effects and genome-wide survival analyses. Each value represents the pairwise genetic correlation (rg) observed between the trait shown in the table headers and rows.

Current genome-wide association study for all-cause AFSW found a significant association with 44 independent genomic regions (Figure 15), many of which were previously reported in relation to refractive errors (21). The statistically strongest association was observed between AFSW and a locus in the *TSPAN10* gene (rs7405453, HR = 1.03, $p = 1.71 \times 10^{-35}$). The second strongest association was discovered at another locus previously associated with spherical equivalent (rs4736886, near the *ZMAT4* gene, $p = 3.36 \times 10^{-27}$). Interestingly both genes that show the most significant associations with AFSW, although known for associations with refractive error, have relatively low effects sizes over the spherical equivalent. Only further down the list of genome-wide associations with AFSW do I find the loci at genes usually considered as the strongest risk factors for refractive error, such as *GJD2*, *LAMA2* and *PRSS56* ($p = 1.63 \times 10^{-12}$, $p=6.27 \times 10^{-24}$ and 1.31×10^{-18}).

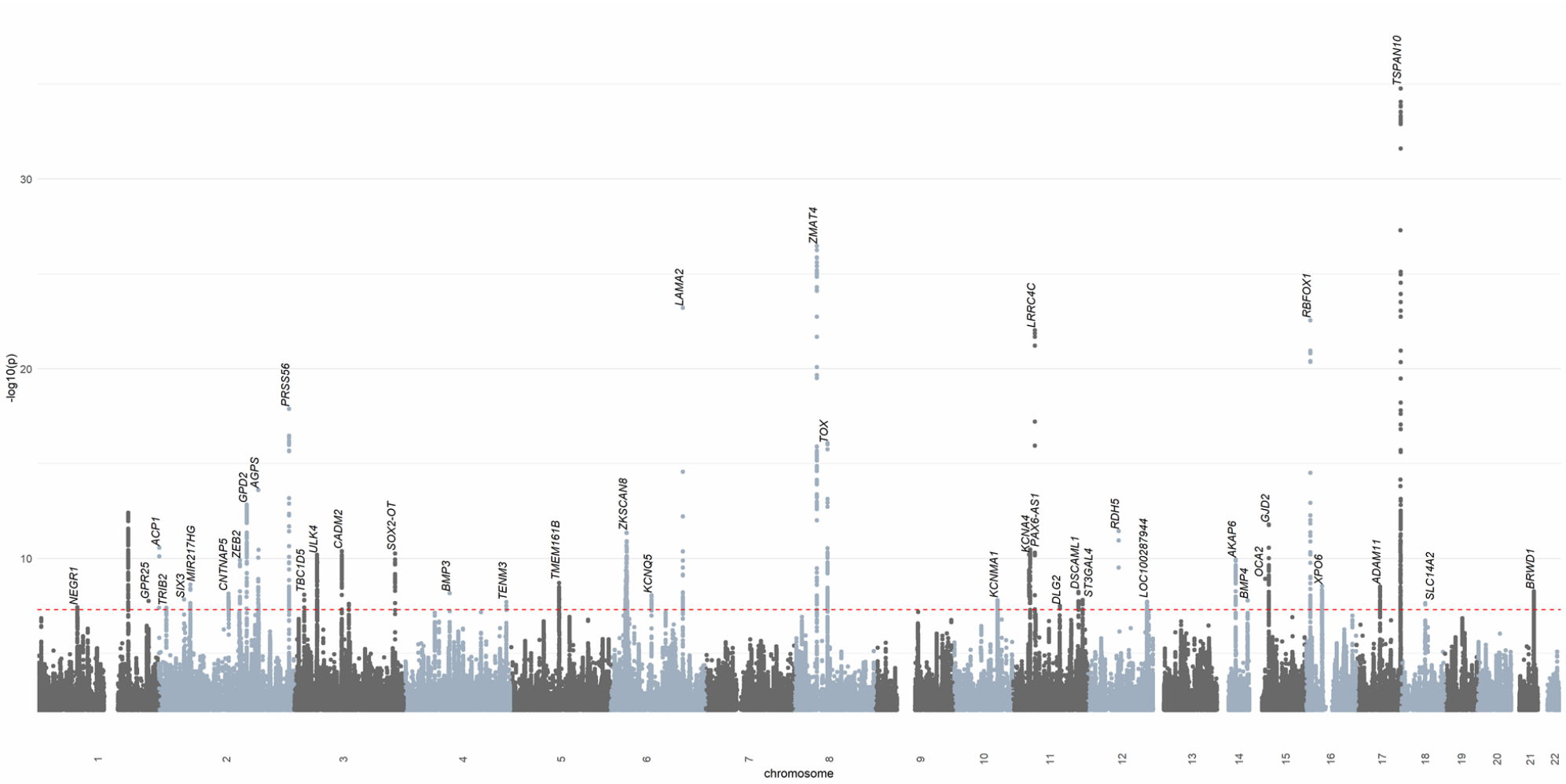


Figure 15 Manhattan plot displaying 44 genome-wide significant associations with the age of first spectacle wear in UK Biobank cohort (N=340,318). The plot shows log10 transformed p-values for each marker plotted against the chromosomal location. The red dashed line indicates the genome-wide significance threshold (p-value < 5.00×10^{-8}). Regions are named with symbols of the transcript-coding genes nearest to the most strongly associated variant in the region.

Although the effect sizes of association with spherical equivalent were generally correlated with their effect over the AFSW for the same alleles, there were notable exceptions. For example, the SNP alleles at the known *BMP3*, *ZMAT4* and *TSPAN10* genes predisposed to much earlier correction compared to the final spherical equivalent status in adulthood than most other loci. Conversely, alleles in the *SOX2-OT* gene seem to confer a low risk towards myopia, but at a much later age than the general regression line averaging all loci (Figure 16).

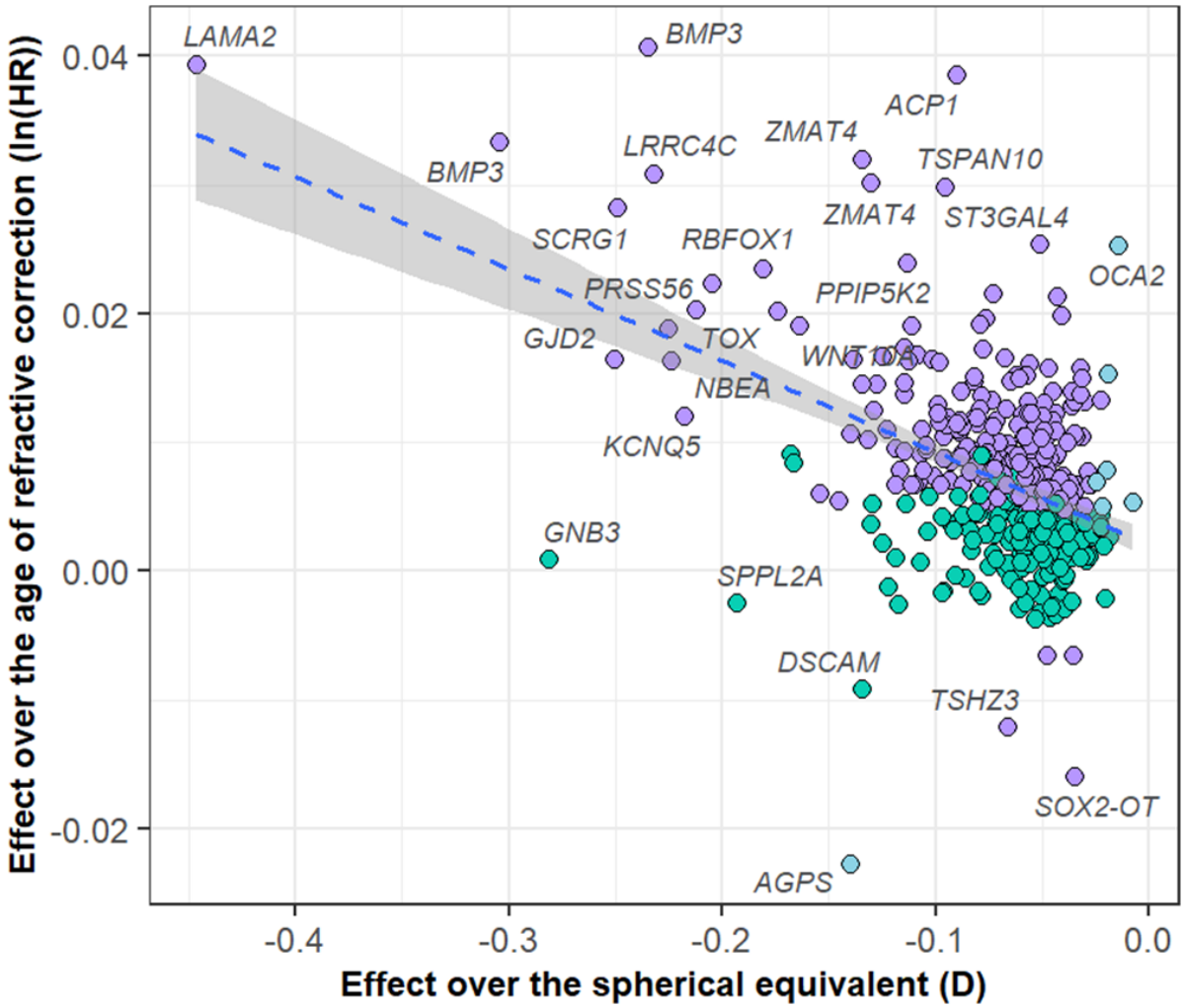


Figure 16 Scatterplot displaying the correlation between the AFSW hazards ratios and spherical equivalent beta coefficients. Hazard ratios shown here as (ln (HR)) represent the multiplicative change in the rate of first spectacle wear per copy of the myopia risk allele calculated in the full sample of 340 318 UK Biobank participants, which was taken as reference. The results are shown for the most strongly associated SNPs in the region. The SNPs in purple exhibiting stronger effects over the age of myopia onset, SNPs in turquoise are associated with spherical equivalent but not AFSW, and SNPs in blue are associated with AFSW but not spherical equivalent.

The genome-wide associations with six additional loci were observed that were not described in any previous GWAS for refractive error (99). The study analyses revealed new associations with polymorphisms within the genomic sequence of the *NEGR1* gene (rs1204700722, HR = 1.01, p = 3.72 x 10^{-08}), a member of an immunoglobulin superfamily cell adhesion molecule supergroup, implicated in neuronal growth and connectivity (395), where previous studies have identified association with depression and affective disorders (367,396). A novel significant association was also detected at a chromosome 2 intergenic region between the *TRIB2* and *LOC1005064* gene sequences (rs10164589, HR = 1.012, p = 3.95 x 10^{-08}). The *TRIB2* gene is a pseudokinase family member that regulates intracellular cell signalling through ubiquitination and scaffolding (397). Additionally, I found an association for a locus on chromosome 3 (rs6577621, HR = 1.01, p = 8.15 x 10^{-09}) in a region located between the *TBC1D5* gene, a regulator of GTPase-activating proteins (398) and the *SATB1* gene, which participates in chromatin remodelling (399). Finally, I identified an association for polymorphisms located within the *ADAM11* gene (rs55882072, HR = 1.01, p = 3.06 x 10^{-09}) a metalloprotease that regulates cell and matrix communications (400) and markers within *BRWD1* gene (401) (rs8131965, HR = 0.98, p = 5.41 x 10^{-09}).

Four out of six novel regions replicated in a slightly smaller but independent cohort (108) (Table 13), at a Bonferroni multiple testing correction level (p-value < 0.05/6 = 0.01, (Table 13). Crucially, all six loci demonstrated the same direction of effects as in the discovery GWAS. Specifically, *NEGR1*, *TRIB2*, *TBC1D5*, *LOC100287944*, *ADAM11* risk alleles were associated with earlier-age myopia, while *BRWD1* showed significant association with later-age refractive error correction (Table 13). Most SNPs were associated, at various levels of statistical significance, with spherical equivalent in the refracted subgroup of European UK Biobank participants (Table 14).

Chr	BP	SNP	Gene	A1	A2	Freq.	Discovery HR	95% CI		Discovery p-value	Replication HR	Replication p-value
1	**72720383**	**rs1194277 ***	***NEGR1***	**G**	**T**	**0.69**	**1.007**	**1.008**	**1.018**	**3.72×10^{-08}**	**1.01**	**0.14**
2	**13042958**	**rs10164589**	***TRIB2***	**T**	**G**	**0.48**	**1.013**	**1.008**	**1.018**	**3.96×10^{-08}**	**1.02**	**0.001**
3	**18192988**	**rs6577621**	***TBC1D5***	**G**	**A**	**0.45**	**1.014**	**1.009**	**1.019**	**8.15×10^{-09}**	**1.02**	**0.01**
12	106927958	rs7295942	*LOC100287944*	C	T	0.75	1.015	1.01	1.021	1.96×10^{-08}	1.01	0.38
17	**42847438**	**rs55882072**	***ADAM11***	**C**	**G**	**0.72**	**1.015**	**1.01**	**1.021**	**3.06×10^{-09}**	**1.03**	**0.0002**
21	**40575426**	**rs8131965**	***BRWD1***	**G**	**A**	**0.64**	**0.986**	**0.981**	**0.991**	**5.41×10^{-09}**	**0.97**	**0.00007**

Table 13 Replication of six novel loci associated with AFSW. Replication was carried out using the results of a genome-wide myopia onset survival study by Kiefer *et al*. (2013 PLoS Genetics 2013). The field "SNP" includes the polymorphic variants with the strongest associations (Discovery p-value) for each region, for which the Chromosome number (Chr) and genomic position (BP) are displayed. "A1" lists the alleles at each SNP locus for which the effect sizes (Discovery HR as hazard ratios), 95% confidence intervals of the estimates (95% CI) and frequencies (Freq.) are reported, and the field "A2" lists alleles alternative to effect allele. "Gene" includes the symbol of transcript-coding gene nearest to the most strongly associated variant in the region. The columns "Replication HR" and "Replication p-value" display hazard ratios and p-values for the genetic associations in Kiefer *et al*. genome-wide survival analyses. The associations with replication p-value below the threshold of multiple testing correction (p = 0.01) are shown in bold font. * The rs1194277 SNP, the second-best associated SNP in the AFSW analysis, was used as a replacement for rs1204700722, which was not available in the *23andMe* dataset

SNP	A1	A2	Beta	SE	p-value
rs1204700722	TA	T	-0.03	0.01	0.01
rs10164589	T	G	-0.04	0.01	$1.5x10^{-03}$
rs6577621	G	A	-0.03	0.01	$2.4x10^{-03}$
rs7295942	C	T	-0.03	0.01	0.01
rs55882072	C	G	-0.02	0.01	0.12
rs8131965	G	A	0.02	0.01	0.081

Table 14 Association with spherical equivalent of the same SNPs associated with age of first spectacle wear. The effects over spherical equivalent were estimated in the directly refracted UK Biobank subsample (N = 102,117). The field "SNP" includes the polymorphic variants with the strongest associations (p-value) for each region."A1" lists the alleles at each SNP locus for which the effect sizes (Beta), standard errors (SE) and p-values (p-value) are displayed. The field "A2" contains alleles alternative to effect allele.

The Cox proportional hazards model assumes that the effects of the tested SNPs have a constant, linear relationship with age. Proportionality of the hazards analyses showed that this assumption held true for many loci, for example, *BMP4*, *TMEM161B*, *XPO6* (Table 15). By contrast, many loci exhibited non-linear effects with age, including *TSPAN10*, *OCA2* loci and, interestingly, *PAX6*, a gene known to harbour variants that cause microphthalmia and severe eye malformation (402) (Table 15), with effects peaking around adolescence.

Gene	SNP	p-value
TSPAN10	**rs7405453**	**7.80×10^{-80}**
LAMA2	**rs12193446**	**8.04×10^{-38}**
LRRC4C	**rs11602008**	**1.51×10^{-31}**
RBFOX1	**rs58514548**	**1.44×10^{-27}**
RDH5	**rs3138142**	**5.93×10^{-22}**
GJD2	**rs524952**	**8.62×10^{-19}**
BMP3	**rs112000479**	**2.73×10^{-18}**
ULK4	**rs11709010**	**4.56×10^{-16}**
PRSS56	**rs1550094**	**3.60×10^{-15}**
TOX	**chr8:60179048**	**5.72×10^{-14}**
GPD2	**rs2352096**	**3.64×10^{-13}**
KCNA4	**11:30017064**	**4.84×10^{-12}**
ZMAT4	**rs4736886**	**2.74×10^{-11}**
KCNQ5	**rs1935527**	**2.43×10^{-10}**
1:163963062-164321002	**rs55634905**	**7.45×10^{-10}**
OCA2	**rs1800407**	**5.07×10^{-09}**
ZEB2	**rs13382811**	**9.03×10^{-09}**
SLC14A2	**rs28408641**	**3.77×10^{-08}**
PAX6-AS1	**rs1806152**	**9.36×10^{-08}**
AKAP6	**rs2300861**	**5.65×10^{-06}**
TBC1D5	**rs6577621**	**2.03×10^{-05}**
KCNMA1	**rs1074181**	**2.75×10^{-05}**
SIX3	**rs9309119**	**0.0002**
CNTNAP5	**rs35920642**	**0.0003**
CADM2	**rs62263912**	**0.001**
BRWD1	rs8131965	0.002
MIR217HG	rs1432564	0.004
DSCAML1	rs11217076	0.007
ADAM11	rs55882072	0.007
BMP4	rs7154152	0.007
TENM3	rs71605082	0.01
GPR25	rs3738250	0.01
TMEM161B	rs1377996460	0.02
DLG2	rs12791570	0.04
TRIB2	rs10164589	0.06
ZKSCAN8	rs35128564	0.1
SOX2-OT	rs9842371	0.13
AGPS	rs538506852	0.2
ACP1	rs11553742	0.2
WWP1P1	rs34070949	0.5
XPO6	rs866722267	0.5
NEGR1	rs1204700722	0.5
ST3GAL4	rs59379014	0.6
LOC100287944	rs7295942	0.8

Table 15 Results for the test of deviation from the proportional hazards assumption. The column "SNP" list variants that were significantly associated with the age of first spectacle wear and the field "p-value" lists their respective p-values for the proportional hazards' assumption. Column "Gene" includes the symbols of the transcript-coding genes nearest to the most strongly associated variant in the region The SNPs deviating from proportional hazards assumption after correction for 44 tests are displayed in bold font.

For example, the *LAMA2* variant had a stronger effect over AFSW hazard at an early age, peaking around 16 and a more inhibited effect after the age of 40, similar to the effects of other well-known refractive error genes such as *GJD2*, *ZMAT4*, *RDH5* and interestingly *PRSS56*, a gene also known to be associated with eye structural

malformations (Figure 17) (403). Among novel loci, *TBC1D5* exerted its influence at an earlier age (Figure 18), whereas *NEGR1* and *TRIB2*, and *ADAM11* were expressed over the entire lifespan with the strongest effects over AFSW observed in adolescence (Figure 18).

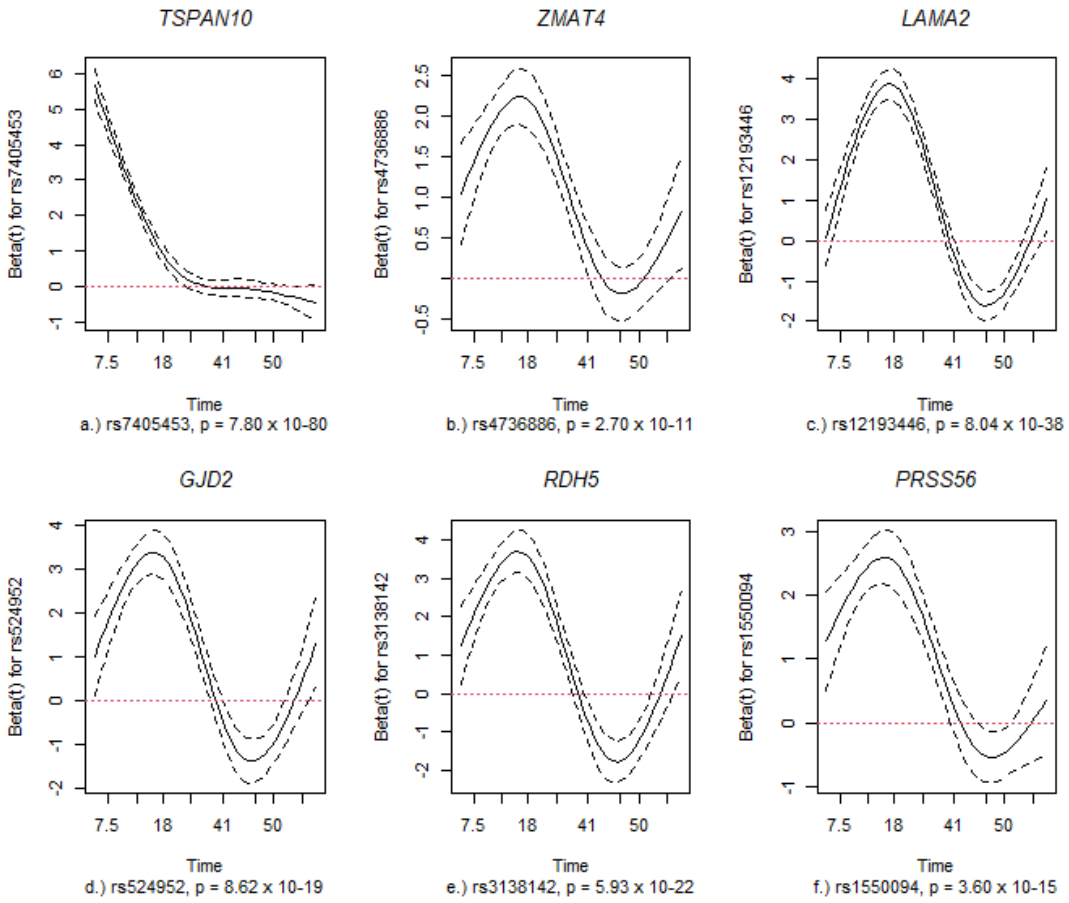


Figure 17 Smoothed log-hazard ratios as a function of age for a selection of six SNPs that showed evidence of deviation from the proportional hazards assumption that are located in regions of previously known association with refractive error. The solid curve denotes a spline fit to beta estimated at different values and the dashed curved line represent 95% confidence intervals for the curve. The p-value from the proportional hazard assumption test is provided in the caption.

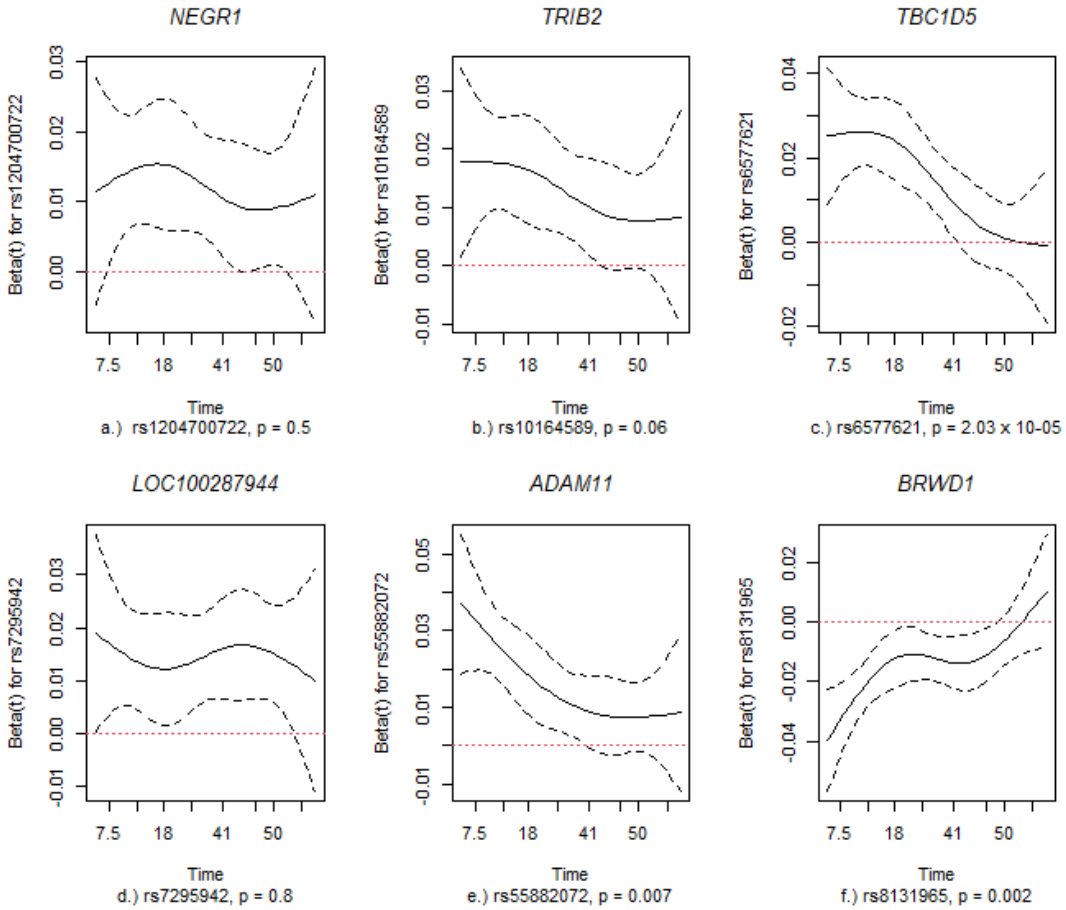


Figure 18 Smoothed log-hazard ratios as a function of age for six SNPs located near or within novel genes. The dashed red line displays the estimated log-hazard ratio (beta) for each variant in the proportional hazards model. The solid curve denotes a spline fit to beta estimated at different values and the dashed curved line represent 95% confidence intervals for the curve. The p-value from the proportional hazard assumption test is provided in the caption.

Our associations with AFSW showed significant enrichment in different body tissues, particularly in the nervous system and retina (Table 16-18), particularly the brain prefrontal cortex, especially in late infancy. Consistent with a higher than expected expression in cerebral tissues, AFSW genes showed strong genetic correlations with neurological traits. Specifically, earlier age of refractive correction correlated with increased cognitive ability (r_g= 0.41, p = 9.73 x 10^{-33}) and neuroticism (r_g= 0.49, p= 0.0039), decreased sleep duration (r_g=-0.24, p=0.04) as well as elevated probability of insomnia (r_g=0.29, p=0.01), and more years of education (years of schooling, r_g= 0.35, p =4.48 x 10^{-67}, Table 19). Moreover, AFSW demonstrated significant genetic correlations several socio-economically influenced traits, particularly increased forced vital capacity (r_g=0.26, p=5.34 x 10^{-05}) and lower probability of asthma (r_g=-0.44, p=0.03, Table 19).

Probe ID	Gene	Top SNP	A1	A2	Freq	GWAS beta	GWAS se	GWAS p-value	eQTL beta	eQTL se	eQTL p-value	SMR beta	SMR se	SMR p-value	HEIDI p-value
ENSG00000135441	*BLOC1S1*	rs3138141	A	C	0.25	-0.018	0.003	1.01 x 10^{-10}	0.73	0.08	5.69 x 10^{-19}	-0.02	0.005	1.69 x 10^{-07}	0.84
ENSG00000182255	*KCNA4*	rs516484	A	C	0.25	-0.016	0.003	5.75 x 10^{-09}	-0.87	0.08	6.23 x 10^{-29}	0.02	0.004	2.42 x 10^{-07}	0.06
ENSG00000197062	*ZSCAN26*	rs1736895	A	G	0.42	0.013	0.002	9.98 x 10^{-08}	-0.74	0.07	1.79 x 10^{-27}	-0.02	0.004	1.73 x 10^{-06}	0.1
ENSG00000196993	*NPIPB9*	rs151174	T	C	0.4	-0.014	0.002	4.25 x 10^{-08}	0.56	0.07	3.48 x 10^{-17}	-0.02	0.005	4.34 x 10^{-06}	0.81

Table 16 SMR results for age of first spectacle wear and eQTL in retinal tissues (Ratnapriya *et al*. Nature Genetics 2019). Column "eQTL summary data" includes the name of eQTL datasets that were used in the analyses. The field "Probe ID" lists the ID for the eQTL probe. The column "Gene" is the gene that the probe eQTL effects act on, "Top SNP" contains the SNP associated with the largest effects on the trait (AFSW); "A1", "A2" and "Freq" are the position (Build 37), effect allele, another allele, and effect allele frequency for the "Top SNP" respectively. The fields "GWAS beta", "GWAS se" and "GWAS p-value" display the effect size, standard error, and p-value for the "Top SNP". Columns "eQTL beta", "eQTL se" and "eQTL p-value" include the effect size, standard error, and p-value for the "Probe ID", "SMR beta", "SMR se", and "SMR p-value" are the effect size, standard error, and p-value results from the SMR analysis. The fields "HEIDI p-value" include the p-value from the HEIDI (Heterogeneity In Dependent Instruments) test.

Probe ID	Gene	Top SNP	A1	A2	Freq	GWAS beta	GWAS se	GWAS p-value	eQTL beta	eQTL se	eQTL p-value	SMR beta	SMR se	SMR p-value	HEIDI p-value
ENSG00000237412	*PRSS56*	rs2741299	G	T	0.82	-0.021	0.003	1.21×10^{-11}	-0.78	0.05	1.27×10^{-54}	0.03	0.004	5.13×10^{-10}	0.02
ENSG00000189298	*ZKSCAN3*	rs213236	C	T	0.4	0.014	0.002	2.52×10^{-08}	-0.71	0.03	2.34×10^{-95}	-0.02	0.004	7.41×10^{-08}	0.01
ENSG00000235109	*ZSCAN31*	rs213236	C	T	0.4	0.014	0.002	2.52×10^{-08}	0.6	0.04	1.73×10^{-64}	0.02	0.004	1.20×10^{-07}	0.01
ENSG00000233232	*NPIPB7*	rs62034319	G	T	0.41	-0.013	0.002	4.44×10^{-08}	-0.6	0.04	5.21×10^{-56}	0.02	0.004	2.34×10^{-07}	0.74
ENSG00000135930	*EIF4E2*	rs1729257	A	G	0.22	0.016	0.003	4.92×10^{-08}	0.6	0.04	1.44×10^{-41}	0.03	0.005	4.25×10^{-07}	0
ENSG00000185658	*BRWD1*	rs10427502	A	G	0.37	0.013	0.003	9.43×10^{-08}	-0.6	0.04	1.85×10^{-53}	-0.02	0.004	4.59×10^{-07}	0.01
ENSG00000196502	*SULT1A1*	rs4149406	G	A	0.38	-0.013	0.003	1.29×10^{-07}	-0.64	0.04	2.23×10^{-59}	0.02	0.004	5.11×10^{-07}	0.14
ENSG00000197165	*SULT1A2*	rs4149406	G	A	0.38	-0.013	0.003	1.29×10^{-07}	0.63	0.04	1.30×10^{-58}	-0.02	0.004	5.20×10^{-07}	0.02
ENSG00000216901	*AL022393.7*	rs34024998	C	G	0.38	0.013	0.002	2.66×10^{-07}	0.69	0.04	3.46×10^{-81}	0.02	0.004	6.74×10^{-07}	0.005
ENSG00000218016	*ZNF192P2*	rs34024998	C	G	0.38	0.013	0.002	2.66×10^{-07}	-0.64	0.04	1.31×10^{-68}	-0.02	0.004	7.92×10^{-07}	0.004
ENSG00000227207	*RPL31P12*	rs2568960	G	A	0.6	0.012	0.002	7.31×10^{-07}	-0.68	0.04	2.99×10^{-71}	-0.02	0.004	1.82×10^{-06}	0.24
ENSG00000262049	*RP13-1032I1.7*	rs12450195	C	T	0.39	-0.017	0.002	1.40×10^{-11}	0.25	0.04	1.12×10^{-10}	-0.07	0.015	3.07×10^{-06}	0.15
ENSG00000112763	*BTN2A1*	rs12190859	A	T	0.22	0.015	0.003	2.52×10^{-07}	-0.54	0.05	3.49×10^{-24}	-0.03	0.006	4.29×10^{-06}	0.01

Table 17 SMR results for age of first spectacle wear and eQTL in the brain tissues (Wang *et al*. 2018 Science). The field "Probe ID" lists the IDs for the eQTL probe. The column "Gene" is the gene that the probe eQTL effects act on, "Top SNP" contains the SNP associated with the largest effects on the trait (AFSW); "A1", "A2" and "Freq" are the position (Build 37), effect allele, another allele, and effect allele frequency for the "Top SNP" respectively. The fields "GWAS beta", "GWAS se" and "GWAS p-value" display the effect size, standard error, and p-value for the "Top SNP". Columns "eQTL beta", "eQTL se" and "eQTL p-value" include the effect size, standard error, and p-value for the "Probe ID", "SMR beta", "SMR se", and "SMR p-value" are the effect size, standard error, and p-value results from the SMR analysis. The fields "HEIDI p-value" include the p-value from the HEIDI (Heterogeneity in Dependent Instruments) test.

Probe ID	Gene	Top SNP	A1	A2	Freq	GWAS beta	GWAS Se	GWAS p-value	eQTL beta	eQTL se	eQTL p-value	SMR beta	SMR se	SMR p-value	HEIDI p-value
cg18240062	*NPLOC4*	17:79530332	A	G	0.51	0.017	0.002	$5.36x10^{-13}$	-0.03	0.004	$5.08x10^{-44}$	-0.622	0.1	$5.97x10^{-10}$	$2.48x10^{-04}$
cg19825371	---	17:79454945	G	A	0.41	0.014	0.002	$1.47x10^{-09}$	-0.09	0.004	$4.40x10^{-124}$	-0.166	0.03	$1.82x10^{-08}$	0.08
cg11644478	*PSMG1*	21:40617483	A	G	0.35	0.014	0.003	$1.52x10^{-08}$	0.03	0.002	$3.96x10^{-72}$	0.464	0.09	$2.12x10^{-07}$	0.3
cg23266546	---	6:28207363	T	C	0.61	-0.013	0.002	$2.48x10^{-08}$	0.05	0.004	$2.88x10^{-29}$	-0.281	0.06	$1.39x10^{-06}$	0.4
cg13177375	---	6:28866528	C	T	0.12	0.019	0.004	$1.09x10^{-07}$	0.06	0.004	$6.36x10^{-46}$	0.321	0.07	$1.45x10^{-06}$	0.04
cg11144229	*TSPAN10*	17:79617871	A	G	0.63	0.013	0.003	$1.59x10^{-07}$	0.05	0.003	$2.10x10^{-47}$	0.259	0.06	$2.17x10^{-06}$	$2.65x10^{-05}$
cg11890956	*PSMG1*	21:40542757	TA	T	0.34	0.013	0.003	$4.64x10^{-07}$	0.04	0.002	$7.28x10^{-88}$	0.287	0.06	$2.79x10^{-06}$	0.1.
cg19292749	*ARL16; HGS*	17:79641736	C	T	0.23	-0.015	0.003	$9.18x10^{-08}$	0.01	$7x10^{-04}$	$2.56x10^{-31}$	-1.87	0.4	$3.01x10^{-06}$	0.02
cg01336390	---	6:29938258	A	C	0.3	0.013	0.003	$3.00x10^{-07}$	0.05	0.004	$2.01x10^{-36}$	0.267	0.06	$4.64x10^{-06}$	0.15

Table 18 SMR results for age of first spectacle wear and eQTL in the foetal brain tissues (Hannon *et al*. 2015; Nat Neuroscience). The field "Probe ID" lists the ID for the eQTL probe. The column "Gene" is the gene that the probe eQTL effects act on, "Top SNP" contains the SNP associated with the largest effects on the trait (AFSW); "A1", "A2" and "Freq" are the position (Build 37), effect allele, other allele, and effect allele frequency for the "Top SNP" respectively. The fields "GWAS beta", "GWAS se" and "GWAS p-value" display the effect size, standard error, and p-value for the "Top SNP". Columns "eQTL beta", "eQTL se" and "eQTL p-value" include the effect size, standard error, and p-value for the "Probe ID", "SMR beta", "SMR se", and "SMR p-value" are the effect size, standard error, and p-value results from the SMR analysis. The fields "HEIDI p-value" include the p-value from the HEIDI (Heterogeneity in Dependent Instruments) test

Trait	PMID	r_g	se	p-value
Years of schooling 2016	27225129	0.35	0.02	4.48×10^{-67}
Intelligence	28530673	0.41	0.03	9.73×10^{-33}
Forced vital capacity	28166213	0.26	0.06	5.34×10^{-05}
Age of first birth	27798627	0.4	0.11	0.0001
Neuroticism	24828478	0.49	0.17	0.004
Insomnia	27992416	0.29	0.11	0.01
Asthma	17611496	-0.44	0.2	0.03
Years of schooling	25201988	0.3	0.14	0.03
Sleep duration	27494321	-0.24	0.11	0.04
Age of smoking initiation	20418890	-0.4	0.21	0.05

Table 19 Genetic correlations between age first spectacle wear and 23 different traits. Columns "Trait", "PMID", describe the name of the tested trait and PubMed repository study number. Fields "r_g", "se" and "p-value" list Pearson's correlation coefficient, standard error, z-score as well as p-values of the tested correlations.

Gene-set enrichment analyses showed that similar to the findings of other published refractive error GWAS (99), AFSW-associated genes were involved in nervous system development (Table 20) and other processes, such as cell signalling and intracellular communications (Table 20). Gene Ontology enrichment analysis results also supported previous conclusions that genes involved in refractive error influence RNA polymerase transcription and gene expression (Table 20) (99).

Gene Ontology Term	Gene set size	Effective set	Expected number of genes	Significant genes	FDR
CGMP Metabolic Process	23	18	5	12	0.0001
Protein Heterotetramerisation	38	23	6	14	0.0002
Chromatin Assembly or Disassembly	177	100	25	42	0.0004
Negative Regulation of Gene Expression Epigenetic	112	65	16	29	0.0004
Cyclic Nucleotide Metabolic Process	57	52	13	24	0.001
Morphogenesis of an Epithelial Sheet	43	37	9	18	0.001
Negative Regulation of Nitrogen Compound Metabolic Process	1517	1224	306	344	0.001
Negative Regulation of Gene Expression	1493	1219	305	340	0.002
Histone Exchange	53	41	10	19	0.002
Positive Regulation of Biosynthetic Process	1805	1510	378	411	0.002
Regulation of Cell Growth	391	316	79	101	0.002
Negative Regulation of Cell Growth	170	150	38	53	0.003
Sensory Perception of Light Stimulus	212	161	40	55	0.004
Regulation of Mitochondrion Organization	218	167	42	57	0.005
ATP Dependent Chromatin Remodelling	74	66	17	26	0.01
Negative Regulation of Hematopoietic Progenitor Cell Differentiation	24	18	5	10	0.01
Regulation of Type 2 Immune Response	26	24	6	12	0.01
Protein N Linked Glycosylation	75	69	17	27	0.01
Regulation of Pathway Restricted Smad Protein Phosphorylation	60	44	11	19	0.01
Regulation of Small GTPase Mediated Signal Transduction	278	218	55	70	0.01
Ribonucleotide Catabolic Process	28	27	7	13	0.01
Regulation of Chondrocyte Differentiation	46	36	9	16	0.01
Positive Regulation of Gene Expression Epigenetic	78	61	15	24	0.01
Cilium Organization	188	137	34	47	0.01

Table 20 Gene-set enrichment analysis for Gene Ontology terms. Column "Gene Ontology Term" lists Gene Ontology terms that were significantly enriched. Fields "Gene set size" and "Effective set" show the number of genes per gene-set and the number of genes included in the GSEA analysis. Columns "Expected number of SNPs" and "Significant genes" provide the expected and observed number of genes with a corrected gene p-value above the 75-percentile enrichment cut-off. "FDR" reports estimated false discovery rate. All analyses were at a 75-percentile gene cut-off and the table is truncated at FDR < 0.01 to only show the most significantly enriched gene sets.

6.4 Discussion

AFSW is a heterogeneous phenotype that is contributed to by the presence and age of onset of several different forms of refractive error. Observationally and genetically, this phenotype is strongly correlated with presence and age of developing myopia, the most common form of refractive error in the general population, although other forms of refractive error are also correlated with it. The study presented in this thesis demonstrated that AFSW survival analysis is a powerful statistical method that could be used to augment the existing information available from directly measured refractive error. I found evidence that refractive error and AFSW were strongly correlated and shared most of their heritability and genetic risk loci. Additionally, I have identified six novel regions associated with the age of refractive correction and replicated four of them. One of the new genes, *TRIB2,* was previously reported for several different ocular traits and disorders. However, other novel AFSW-associated genes were linked to neurological and neurodevelopmental traits, for which genetic correlation with the refractive error was previously described (110).

Neuronal Growth Regulator 1 *(NEGR1)* nested within newly identified locus on chromosome 1 coded for a GPI-anchored cell adhesion molecule in the immunoglobulin superfamily cell adhesion molecule supergroup (404). Accumulating evidence implicated *NEGR1* in a wide array of psychiatric diseases. Neuronal Growth Regulator 1 was abundant in neuronal somata and dendritic synaptic vesicles of different adult and developing brain regions where it participated in the establishment and remodelling of the neural circuit (405,406). Specifically, *NEGR1*, a synaptic adhesion molecule spanned synaptic cleft and controlled neurite outgrowth and synapse formation (407). Large scale genetic investigations associated polymorphisms present in *NEGR1* with increased susceptibility to Alzheimer's disease (408), major depressive disorder (409) and schizophrenia (410) which shared common polygenic traits contributing to impaired neural signalling. Moreover, the genetic variations within Neuronal Growth Regulator 1 have been linked to low white matter integrity, implicated in the pathogenesis of many psychiatric conditions (411). Partial deletions in *NEGR1* were documented to cause

learning and behavioural difficulties (412), while rare complete deletions resulted in moderate cognitive disability and severe language impairment (413).

Additionally, genome-wide time-to-event analyses revealed an association between all-cause AFSW and *ADAM11* in A Disintegrin And Metalloprotease family. ADAM family comprised over 30 different membrane-spanning proteins that regulated receptor-mediated signals and participated in cell-cell and cell-matrix communications (414). In contrast to other family members, *ADAM11* encoded protein did not contain a catalytic motif, but behaved as an adhesion molecule and was predominantly expressed in CNS (415). A Disintegrin And Metalloprotease 11 was critical for normal neural function and its defects manifested as deficits in motor function, spatial learning (416) and abnormal nociception (415). Functional studies reported widespread gene involvement in neuron-glial and neuron-neuron communications during early development and adulthood (417). Moreover, *ADAM11* was also described in genetic investigations of familial epilepsy (418) and cerebral white matter hyperintensity (419) and educational attainment GWAS (373).

Similarly, another novel AFSW-associated locus *TBC1D5/SATB1* was implicated in several neurological and neurodevelopmental phenotypes. GTPase activating protein *TBC1D5* enhanced retromer activity which had a role in Parkinson's and Alzheimer's diseases, whereas Satb1 coded for a nuclear scaffold protein involved in long-range chromatin remodelling and influencing the expression of neurodevelopmental genes (420,421). Both genes were abundant in developing the somatosensory cortex (420) as well as the postnatal brain, where they demonstrated associations with brain morphology, cortical thickness and general cognitive function (422). Moreover, multiple genetic studies related polymorphisms within genetic sequences of *TBC1D5/SATB1* locus to increased risk of schizophrenia (423), Tourette syndrome (424), autism (425), major depression (426,427) and neuroticism (427,428).

An additional genomic region on chromosome 21 that showed significant association with the age of refractive correction contained the *BRWD1* gene encoding chromatin regular in the WD-repeat protein family. The *BRWD1* was located within Down Syndrome Critical Regions (DSCRs) and was involved in neural differentiation and signal transduction processes (429). Moreover, the gene was among other histone regulators implicated in neural development and plasticity (430). In particular, aberrant histone regulation in CNS disrupted normal neurodevelopment and resulted in subsequent behavioural abnormalities and neurological diseases later in life (430). Several GWAS reported common polymorphisms within *BRWD1* that influenced populations cognitive function (431,432) and were predisposed to bipolar disorder (372).

TRIB2 was a new gene associated with the age of refractive correction, but not reported for any other refractive phenotypes. The gene encoded a member of the Tribbles protein family (trbl) that modulated signal transduction pathways (433). Animal model experiments suggested the Tribble family could participate in eye development and trbl gene mutations were likely suppressing ocular overgrowth phenotypes(434). Tribbles Pseudokinase 2 was reported as associated with several ocular disorders including autoimmune uveitis (435), AMD (436), age-related cataracts (437) and POAG (438). Additionally, *TRIB2,* expressed in the human trabecular meshwork (439), showed associations with glaucoma-related parameters, specifically the optic nerve cup area (440). These findings were consistent with published results indicating genetic pleiotropy between myopia and optic nerve changes (99).

Four out of six novel loci associated with AFSW were replicated at robust multiple testing correction levels. The *TRIB2*, *TBC1D5* and *ADAM11* genes were significantly associated with myopia correction at an earlier age, while *BRWD1* showed association with myopia correction at older ages. Although *NEGR1* and *LOC100287944* were not significantly associated with the age of myopia onset in a replication dataset, they demonstrated the same direction of the effects as in the discovery GWAS, and it is possible that a lack of

statistical significance in replication analyses could be due to sample size and power limitations.

The strongest genetic association was identified with a variant located within *TSPAN10*. This gene showed a moderate association with refractive error (99) but was strongly associated with corneal astigmatism (390) as well as with strabismus and amblyopia (441), which manifest early in childhood. Notably, the association between *TSPAN10* strabismus was independent of refractive error (441). Because the current study sample wasn't limited to individuals with myopia and hyperopia, the observed association between early AFSW and *TSPAN10* may have reflected contributions from other ocular disorders such as astigmatism, strabismus or amblyopia.

This study also found strong associations with markers located near or within *ZMAT4*, *LAMA2* and *GJD2* genes. Similarly to previously published results, I found that *LAMA2* and *GJD2* had an early effect that increased with age (389). In particular, these genes were observed to have the strongest effect on myopia in 10 – 25-year-olds but were also expressed during the entire age span of myopia development (389).

The results of this study confirm the strong correlation between AFSW and myopia. These results also demonstrate that AFSW is a complex phenotype that is likely to capture pleiotropic genetic effects that influence phenotypic traits other than myopia. SNP loci associated with AFSW appear to exert their effects at different time, although it is not clear whether the effect size changes over time of these loci are due in part to that pleiotropy or can simply be explained by their effects over myopia.

A potential limitation of this investigation is that the phenotype used in the analyses was based on self-reported data and not on clinical evaluations. Although self-reported data is occasionally prone to recall bias that could affect the results, its wider availability compared to directly measured refractive error may lead to statistical power gains. Other potential limitation includes the generalisability of these results. The effect sizes we report were largely consistent in the two large European population cohorts in which they were initially estimated and replicated. However, both cohorts are likely to be enriched for myopic participants. Findings in these cohorts may not be generalisable to

other general population cohorts, and particularly they may not apply to more diverse populations.

The time-to-event analyses presented in this thesis identified genome-significant associations with 44 independent loci, most of which were documented in refractive error and myopia GWAS. The study results demonstrate that the effects of many of these regions strongly correlate with myopic refraction but vary with age, which to date was reported for a handful of spherical equivalent genes. Additionally, associations with six novel regions were uncovered of which four were successfully replicated in an independent cohort. The results of this investigation support the role of neural development and signalling in the pathogenesis of myopia. The findings of this study further our knowledge on the genetic basis of refractive disorders and demonstrate the value of large-scale population-based genetic studies.

Chapter 7 | Medication use and refractive error

7.1 Introduction

Myopia on the most common ocular conditions worldwide, affecting 30% of the world population but its prevalence is still rising, especially in East Asia (127,442), where the proportion of urban-dwelling adolescents with myopia gradually increased from 56% in 2006 to 65% in 2015 (443). In the last 15 years, a 20% increase in myopia prevalence was observed among Chinese children and adolescents and is expected to reach 83% by 2050 (444). Although East Asian countries recorded the highest prevalence rates, a gradual myopic shift was also seen in other regions of the world. For example, in Australia and Northern Ireland cross-sectional investigations documented a significant increase in myopia prevalence among 12-year-old children (445,446). The studies conducted in the US also found evidence of increasing prevalence (59), particularly in young adults (447). Likewise, the data from adult populations suggested a significant myopic shift (448). Interestingly, a meta-analysis of observational studies of common ocular diseases in adults across Europe uncovered a marked generational effect for the rising myopia prevalence rates which increased from 18% to 24% in participants born before and after 1940, respectively.-(58). Although younger birth cohorts were more educated, the reasons for generational effects were multifactorial (58). The relationship between the rising prevalence of myopia and education reflected a network of lifestyle exposures related to educational attainment, including near-work activities and limited time spent outdoors, pleiotropic effects between intelligence and myopia, and socioeconomic factors affecting educational opportunities (58).

The rising myopia prevalence rates coincided with rapid urbanisation. A marked increase in the prevalence of myopia was documented in immigrants from India and Malaysia who moved to industrialised societies like Singapore, possibly indicating that maladaptation to a different environment, lifestyle and diet could affect the expression of myopia (449). Notably, the increase in prevalence rates of cardiovascular disease was also attributed not only to the ageing of the population but an accelerated process of

urbanisation and associated insulin resistance. Therefore, it was proposed that the increasing prevalence of myopia could be also a consequence of the socio-economic transition and adoption of a westernised lifestyle. Westernised lifestyle comprised of behavioural patterns that prioritised near-work due to increasing educational demands but were also associated with habitual sedentariness and high glycaemic diet that led to insulin resistance and morbidity. Interestingly, myopia was more prevalent among patients with diabetes compared to non-diabetic controls and was associated with inadequate metabolic control (450). Additionally, there was evidence indicating that insulin resistance could potentially play a role in myopisation by inducing chronic hyperinsulinemia which promoted axial elongation (181). Accordingly, the exploration of the population's morbidity structure, such as the glycaemic profile of the population, but also other diseases, could have theoretically been contributing to the growing myopia prevalence. Similarly, the patterns of medication use also varied across generations and could be potential factors behind the development of refractive disorders. Although mechanisms of adverse drugs reactions are not representative of the morbidity in a wider population, investigation of the population's medication-taking patterns could reveal exceptional mechanisms influencing refractive errors through pathways that may not be physiological. Systemic medication associations were reported for several ocular diseases and traits such as IOP and glaucoma (451,452), where selective serotonin reuptake inhibitors, calcium channel and angiotensin receptor blockers were shown to alter disease risk and progression (453). However, although transient refractive errors are definitely induced by certain pharmacological agents as previously documented (192), the relationship between chronic medication-taking and refractive disorders in the general population hasn't been thoroughly examined.

To further study biological mechanisms involved in the aetiology of refractive errors, I examined the relationship between systemic medication use and mean spherical equivalent in a large subsample of UK Biobank participants. The study aimed to assess

whether childhood-onset refractive errors could predict future medication-taking behaviour and use these insights to decipher novel disease mechanisms.

7.2 Association Between Medication-Taking and Refractive Error in a Large General Population-Based Cohort

This chapter is presented as a published paper and is an exact copy of the following journal publication:

Karina Patasova; Anthony P. Khawaja; Bani Tamraz; Katie M. Williams; Omar A. Mahroo; Maxim Freidin; Ameenat L. Solebo; Jelle Vehof; Mario Falchi; Jugnoo S. Rahi; Chris J. Hammond; Pirro G. Hysi. Association Between Medication-Taking and Refractive Error in a Large General Population-Based Cohort. Invest. Ophthalmol. Vis. Sci., 62, 15–15. https://doi.org/10.1167/iovs.62.2.15. Epub February 2021

Clinical and Epidemiologic Research

Association Between Medication-Taking and Refractive Error in a Large General Population-Based Cohort

Karina Patasova,[1,2] Anthony P. Khawaja,[3] Bani Tamraz,[4] Katie M. Williams,[1–3,5,6] Omar A. Mahroo,[1–3,5,6] Maxim Freidin,[2] Ameenat L. Solebo,[7] Jelle Vehof,[1,2,8] Mario Falchi,[2] Jugnoo S. Rahi,[6,9] Chris J. Hammond,[1,2] and Pirro G. Hysi[1,2,7]

[1]Section of Ophthalmology, School of Life Course Sciences, King's College London, United Kingdom
[2]Department of Twin Research and Genetic Epidemiology, School of Life Course Sciences, King's College London, United Kingdom
[3]NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and the UCL Institute of Ophthalmology, London, United Kingdom
[4]Department of Clinical Pharmacy, University of California San Francisco, San Francisco, California, United States
[5]Department of Ophthalmology, St Thomas' Hospital, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom
[6]Institute of Ophthalmology, University College London, London, United Kingdom
[7]UCL Great Ormond Street Hospital Institute of Child Health, London, United Kingdom
[8]University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[9]Ulverscroft Vision Research Group, University College London, London, United Kingdom

Correspondence: Pirro G. Hysi, Section of Ophthalmology, School of Life Course Sciences, St Thomas' Hospital, King's College London, Westminster Bridge Road, London SE1 7EH, UK; pirro.hysi@kcl.ac.uk.

Received: October 3, 2020
Accepted: January 25, 2021
Published: February 16, 2021

Citation: Patasova K, Khawaja AP, Tamraz B, et al. Association between medication-taking and refractive error in a large general population-based cohort. *Invest Ophthalmol Vis Sci.* 2021;62(2):15. https://doi.org/10.1167/iovs.62.2.15

Purpose. Refractive errors, particularly myopia, are common and a leading cause of blindness. This study aimed to explore associations between medications and refractive error in an aging adult cohort and to determine whether childhood-onset refractive errors predict future medication use to provide novel insights into disease mechanisms.

Methods. The study compared the spherical equivalent values measured in 102,318 UK Biobank participants taking the 960 most commonly used medications. The strengths of associations were evaluated against the self-reported age of spectacle wear. The causality of refractive error changes was inferred using sensitivity and Mendelian randomization analyses.

Results. Anti-glaucoma drugs were associated with 1 to 2 diopters greater myopic refraction, particularly in subjects who started wearing correction in the first two decades of life, potentially due to the association of higher intraocular pressure since early years with both myopia and, later in life, glaucoma. All classes of pain-control medications, including paracetamol, opiates, non-steroidal antiinflammatory drugs, and gabapentinoids, were associated with greater hyperopia (+0.68–1.15 diopters), after correction for deprivation, education, and polypharmacy and sensitivity analyses for common diagnoses. Oral hypoglycemics (metformin, gliburonide) were associated with myopia, as was allopurinol, and participants using bronchodilators (ipratropium and salbutamol) were more hyperopic.

Conclusions. This study finds for the first time, to our knowledge, that medication use is associated with refractive error in adults. The novel finding that analgesics are associated with hyperopic refraction, and the possibility that multisite chronic pain predisposes to hyperopia, deserves further research. Some drugs, such as antihyperglycemic or bronchodilators, may directly alter refractive error. Intraocular pressure appears causative for myopia.

Keywords: myopia, medications, intraocular pressure, chronic pain, refractive error

Refractive errors arise from mismatches between the light-converging power of the cornea and lens and the axial length of the eye. They are common[1] and increasingly a cause of visual impairment in many communities.[2] Of the different forms of refractive error, myopia is by far the most prevalent. The prevalence of myopia is rising worldwide, especially in Asia,[3] but also in Europe[4] and the rest of the world.[5] At the current pace of growth, half of the world population is expected to develop myopia by 2050.[6]

Factors affecting myopia prevalence are wide ranging from perinatal to early developmental.[7] Educational attainment[8] and behavioral factors, such as time spent outdoors,[9,10] are important drivers of generational shifts in the prevalence of myopia.[11] Nevertheless, cultural and socioeconomic changes are often intercorrelated spatially and temporally. Disentangling effects of each social and cultural factor from others is challenging. For example, the extra time spent indoors coincides with an environment that

Copyright 2021 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Investigative Ophthalmology & Visual Science

Downloaded from iovs.arvojournals.org on 06/30/2021

is changing its dioptric structure[12] and ambient light intensity, as well as concurrent cultural changes, such as diet, also associated with myopia in children.[13] Several other environmental and cultural factors also vary across generations and may contribute to myopia risk.[14] Morbidity structure and patterns of medication use are also likely to have altered through generations, potentially affecting the eye and its optical properties.[15]

Documented drug-induced refractive errors are mostly described in relation to transient myopia or other idiosyncratic ocular manifestations in the eye.[15,16] It is currently unknown whether subtler, chronic changes in the spherical equivalent (SpE) are associated with commonly prescribed medications in the broader community. The investigation of systemic medication use in relation to spherical equivalent might help identify potentially modifiable risk factors. This study aims to investigate potential associations between commonly used medications and refractive error. It explores the hypotheses that medication taken in adulthood is associated with altered SpE measurements and whether refraction status in earlier ages is associated with morbidity and medication-taking patterns in later life. This would give us insights into the etiology of these conditions and provide opportunities for their early detection and prevention.

Methods

Study Population

The UK Biobank cohort includes 502,682 participants, all UK residents selected from the UK National Health Service register.[17] Participants provided extensive reports on their lifestyle and environmental exposures, either by filling in touch-screen questionnaires or through face-to-face interviews.[17] All UK Biobank data were acquired cross-sectionally, although previous medical histories were also reported and coded according to the International Classification of Diseases, 10th Revision (ICD10; standard UK Biobank field number 41270)[17] or retrospectively recorded as answers to questionnaires.

Approximately 23% (N = 117,279) of all participants underwent a comprehensive ophthalmological examination.[18] Noncycloplegic autorefraction was performed cross-sectionally at the time of recruitment, using a Tomey RC 5000 device (Tomey Corp., Nagoya, Japan), and for each participant the spherical equivalent was calculated as SpE = sphere + $\frac{1}{2}$ cylinder power (UK Biobank field numbers 5084-5085, 5086-5087) for each eye separately. The spherical equivalent measure that we have used in this study represents the average SpE of the left and right eyes. Measurements from one eye were used if data from the fellow eye were unavailable. Following previously published recommendations,[19] we excluded participants with a previous history of eye surgery, self-reported cataract with mild myopia, or bilateral or unilateral eye injury resulting in vision loss (SpE measurements in the intact, healthy eye were included). To minimize confounding arising from the population genetic structure, we limited the study sample to individuals of European ancestry, as ascertained by using genetic information.

In addition, the UK Biobank participants retrospectively reported the age at first refractive correction (hereafter referred to as the age of spectacle wear, AOSW). The information about AOSW was collected during the imaging visit, which included refractive error assessment. Information on medication use was also collected retrospectively, using touch-screen questionnaires and face-to-face interviews.[20] The names of the treatments, but not the dosage or the duration of treatment, were recorded. Participants were asked to answer the question "Do you regularly take any prescription medications?" and were prompted to specify what they were taking.[20] The interviewer recorded the name of both prescription and over-the-counter medications.[20] The regular prescription medications were defined as regular treatments taken daily, weekly, monthly, or quarterly, for drugs such as depot injections.[20] Drugs taken transiently (for example, a single-week course of antibiotics or analgesics taken 2 days before the interview) and recently discontinued treatments were not recorded.[20] Names of medications were selected from listed options that included both generic and trade names.

For our analyses, we matched available generic and trade names to the active ingredients. Active ingredients of brand names were retrieved using information from the electronic medicine compendium (available at https://www.medicines.org.uk/emc/). We selected pharmaceutical drugs that had at least 10 users of European ancestry in our dataset and excluded topical dermatological treatments, as well as herbal and homoeopathic remedies. We specifically included medication taken systemically (orally and parenterally), inhalers, and eye drops. When more than one active ingredient was present in one medication, each was counted as if separately prescribed and added to the count of medications containing the same specific active ingredient alone. Polypharmacy was defined as the reported intake of 5 or more medications during the reporting period,[21] even though these medications were not necessarily taken at the exact same time. In the subsequent paragraphs, we will use the term "medication" interchangeably for "active ingredient."

We included polypharmacy, physical activity, social and economic conditions, and other potential confounders as relevant covariates in our analyses. Physical activity of the study participants was ascertained through touch-screen questionnaires asking about all other forms of exertion, whether occupational or leisure. The length, intensity, and typical duration of daily walking and of moderate or vigorous physical activity (UK Biobank field numbers 864, 874, 894), the 24-recall of those activities (UK Biobank field numbers 104920, 104910, 104900), and the frequency of physical activity that lasted for at least 10 minutes (field numbers 864, 884, 904) were documented.

Years of education were calculated from self-reported data on academic and professional qualifications (field number 6138) according to the International Standard Classification of Education categories.[22] The Townsend deprivation index (field number 189) was calculated by aggregating data from the national census, based on previously described methods.[23] The Townsend deprivation index was used as a proxy for a participant's socioeconomic status.[24] The index is based on four components of material deprivation: (1) unemployment, (2) car non-ownership, (3) home non-ownership, and (4) household overcrowding.[24]

The study received the approval of the institutional and/or national research committee and adhered to the tenets of the Declaration of Helsinki. Informed consent was obtained from all study participants at recruitment.

Statistical Analyses

Descriptive analyses were performed using the epiDisplay package in R (R Foundation for Statistical

Downloaded from iovs.arvojournals.org on 06/30/2021

Computing, Vienna, Austria). We calculated means, standard errors, medians, and interquartile ranges for normally and non-normally distributed continuous variables. Categorical data were described by computing frequencies. Results were considered statistically significant if they passed the threshold of 0.05.

The associations between refractive error and medications were tested in linear mixed models (LMMs). We used lme4[25] and lmertest[26] packages to generate the models. SpE was not fully normally distributed; however, following convention, we used untransformed refractive error as the outcome in our model, with the self-reported intake of each of the pharmaceutically active ingredients as predictors. Additionally, the models were adjusted for recruitment age and age squared, sex, number of years spent in formal education, Townsend deprivation index, and polypharmacy. The birth cohort was defined as the participant's birth decade and was used as a random-effect grouping factor in our mixed models to account for potential nonlinear intergenerational effects.[4] The LMM coefficients for each treatment represent adjusted the difference in SpE between medication users and non-users. Associations were considered significant if their probabilities were below the Bonferroni multiple testing correction adjustment level (alpha/number of medications, alpha = 0.05). We performed three main analyses in our study, including all participants with available data (n = 481 medications), subjects who started wearing spectacles before the age of 35 (n = 228 medications), and those who started after the age of 35 (n = 241 medications). The Bonferroni corrected P values were generated using the p.adjust function in R and were 0.0001, 0.000219, and 0.000207, respectively.

Mendelian Randomization

Association between any two variables does not necessarily imply that one causes the other. Exploration of causation in many cases is difficult, as it may require observations spanning over a long period of time and may be subject to bias due to the co-occurrence of confounding factors. An established bias-reducing procedure is randomized clinical trials (RCTs). The key in RCTs is to randomly assign participants to each of the study arms, regardless of the confounding.

Mendelian randomization (Supplementary Fig. S1) is an alternative randomization technique, conceptually similar to RCTs. We know from Mendel's laws of heredity that alleles segregate randomly after each meiosis and that they are assorted to offspring independently of the traits for which they code. If several polymorphisms are associated with a hypothetical trait (trait 1), they will be randomly be reshuffled as they are transmitted from one generation to the next, assigning each child at the moment of conception with an almost unique combination of risk alleles. The net sum of the risk that these alleles confer with respect to trait 1 is referred to as a polygenic risk score. Because the alleles are reassigned after each different conception, the PRS predicts only the hypothetical trait 1 and no other trait, unless strongly correlated with trait 1.

If trait 1 predisposes to another phenotype, such as spherical equivalent in our study, the individuals who randomly received a higher dose of genetic risk will be exposed at birth to higher levels of trait 1 (which in this context is referred to as "exposure"), and in response to this exposure these individuals will have, on average, higher levels of a second phenotype (trait 2, or the "outcome"), which should not be observationally correlated with the exposure. When the exposure causally influences the outcome, each individual genetic factor (in this context referred to as "instrumental variables") selected to increase the exposure will also proportionally increase the likelihood of the outcome. With increasing polygenic risk toward the exposure, the likelihood of its outcome also changes proportionally. Mendelian randomization (MR) statistically tests the linearity of the correlation of effect sizes of the genetic associations of the same instruments with the exposure and outcome phenotypes. As an example, we can use as instrumental variables the genetic variants significantly associated with intraocular pressure (IOP, the exposure) to examine if the association of IOP with the outcome (SpE) is causal or vice versa.

To further explore relationships between refractive error and medication intake, we built MR models to test for potential causal relationships between SpE and 12 different phenotypes. MR instrumental variables included independent genome-wide significant genetic variants ($P < 5 \times 10^{-8}$, $r^2 < 0.1$) associated with the specific selected traits. For every medication significantly associated with refractive error, we aimed to perform two experiments. The first assessed the relationship between refractive error and underlying disease potentially confounding the association with the treatments. The second tested causality between SpE and medication classes that included drugs associated with refraction. The analyses were performed using the R MendelianRandomization package. Given that the outcomes were tested on refractive error data that were partially obtained from the UK Biobank (about 25% of the effective sample size), whenever the instrument variables were identified through prior UK Biobank analyses we also corrected for overlap of samples between analyses by setting the parameter psi to the observational correlation between exposure and outcome phenotypes. This correction effectively treats analyses as one-sample MRs and may be overly conservative, given the partial potential sample overlap. These results are provided as sensitivity analyses in relevant cases. Common and independent genetic variants associated with selected phenotypes (instrumental variables) were extracted from publicly available genome-wide association study (GWAS) summary statistics (https://www.ebi.ac.uk/gwas/). Three separate, yet complementary, MR tests were used (MR-Egger, simple median, and inverse-variance weighted). These tests are usually applied together to jointly estimate causality; that is, they are not independent tests requiring multiple testing correction. In special circumstances (significant sample overlap) a maximum likelihood test is added to test whether causality remains significant even after adjustment.

Secondary Analyses

Secondary analyses were conducted to address the potential confounding effect of comorbidities and involved removing individuals with a previous medical history of the most prevalent diagnoses among the category receiving a specific medication. Sequentially, and for each ICD10 entry, individuals reporting the same codes were dropped from the models, and associations were also tested in the subsample of the study participants who did not report any ICD10 codes.

Secondary associations also assessed the potential confounding effect of other factors. We generated nine models to test relationships between SpE and selected treatments, adjusted for age, age squared, sex, and socioeco-

Downloaded from iovs.arvojournals.org on 06/30/2021

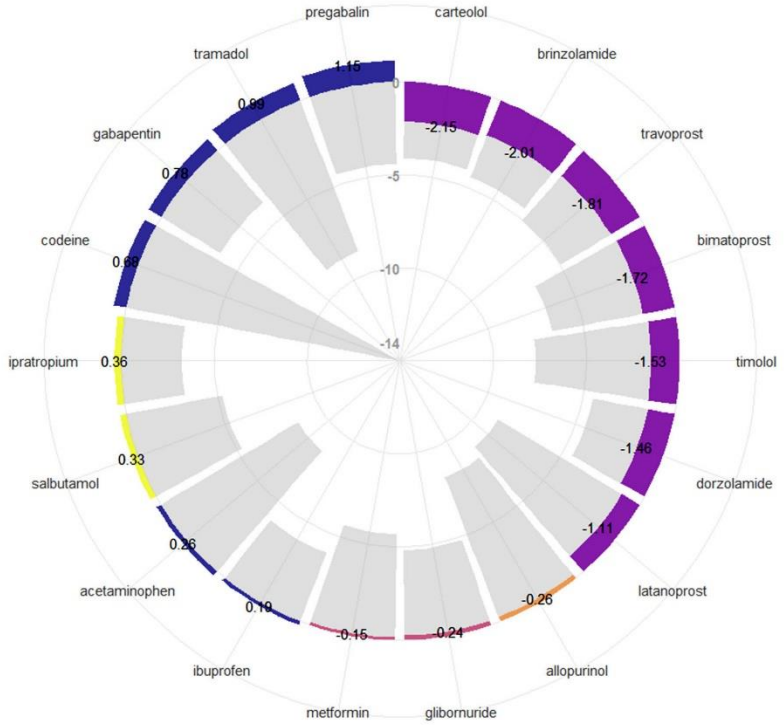


Figure 1. Circular bar plot showing the associations between refractive error and 18 different medications. *Colored bars* represent the LMM betas, and *gray bars* show the *P* values of the associations. The values refer to the difference in spherical equivalent between the group of subjects under the specific medication and the control group (rest of the cohort that does not report taking that same medication).

nomic factors, as well as nine different measures of physical activity.

Results

The study sample included 102,318 UK Biobank participants of homogeneous European ancestry; 47% (47,874) were men, and median age was 59 years (interquartile range [IQR], 51–64). A more detailed description of the sample's demographic characteristics can be found in Supplementary Table S1. SpE followed a characteristic distribution where most individual observations are seen both closer to the mean value and to the extremes, giving it a characteristic shape of a sharper peak with heavy and wider tails compared to the normal distribution (also called leptokurtic distribution; Supplementary Fig. S2) also seen in other populations.[14] The AOSW among the subset of participants for whom this information was available followed a bimodal distribution (Supplementary Fig. S3), with the first mode occurring between 1 and 35 years, peaking around the age of 13, and the second mode peaking around 47 years, when the average spherical equivalent is hyperopic (Supplementary Fig. S4).

We focused on the subsample of the UK Biobank participants (N = 102,318) of European ancestry for whom SpE was directly measured. In this sample, 481 active ingredients were taken orally, parenterally, or through nasal sprays or eye drops by a minimum of 10 participants. The most commonly reported medications used during the surveyed period were paracetamol (also known as acetaminophen, 17%), followed by simvastatin (13%) and acetylsalicylic acid (aspirin, 13%) (Supplementary Fig. S5). The median number of reported medications per participant was two (IQR, 0–4), but 18% (18,558) of the subjects were taking more than five medications. The number of medications taken was significantly associated with demographic and socioeconomic factors. Age and Townsend deprivation index had a statistically strong positive relationship with polypharmacy, whereas education was negatively associated with the concurrent use of multiple medications (Supplementary Table S2).

Eighteen different medications showed statistically significant associations with refractive error after adjustment for multiple testing (Table, Fig. 1). The seven strongest associations were from medications of different classes commonly used to control IOP and glaucoma. In particular, three prostaglandin agonists (latanoprost, $P = 2.98 \times 10^{-20}$; bimatoprost, $P = 1.47 \times 10^{-11}$; travoprost, $P = 2.45 \times 10^{-8}$), two carbonic anhydrase inhibitors (dorzolamide, $P = 1.58 \times 10^{-11}$; brinzolamide, $P = 2.68 \times 10^{-9}$), and two beta blockers (timolol, $P = 9.57 \times 10^{-12}$; carteololl $P = 7.95 \times 10^{-5}$) were associated with a lower (i.e., more myopic) SpE (Table). Subjects taking these medications had, on average, SpE values that were 1 to 2 diopters lower than those of the other participants. We stratified analyses by AOSW in a subset of 90,550 participants for whom this information was available. The associations between IOP-lowering medication and SpE were driven by participants who reported spectacle correction earlier in life (AOSW< 35 years) (Table), particularly in the first two decades (Fig. 2), well before the

Investigative Ophthalmology & Visual Science

Downloaded from iovs.arvojournals.org on 06/30/2021

Downloaded from iovs.arvojournals.org on 06/30/2021

TABLE. Results of LMMs Testing the Association Between Refractive Error and Medication Intake

Medications Tested	Medication Action/Class	All Study Participants (N = 102,318)				Participants with AOSW Between 5 and 35 Years (n = 38,960)				Participants with AOSW Over 35 Years (n = 48,240)			
		β*	SE*	P†	Bonferroni-Corrected P†	β*	SE*	P†	Bonferroni-Corrected P†	β*	SE*	P†	Bonferroni-Corrected P†
Latanoprost	IOP-lowering (prostaglandin analogue)	−0.99	0.11	2.99×10^{-20}	1.44×10^{-17}	−1.11	0.18	1.38×10^{-9}	3.1×10^{-7}	−0.2	0.07	0.007	1
Timolol	IOP-lowering (β blocker)	−1.05	0.15	9.57×10^{-12}	4.60×10^{-9}	−1.53	0.27	1.92×10^{-8}	4.3×10^{-6}	−0.13	0.1	0.19	1
Bimatoprost	IOP-lowering (prostaglandin analog)	−1.27	0.19	1.48×10^{-11}	7.10×10^{-9}	−1.72	0.33	1.94×10^{-7}	4.4×10^{-5}	−0.07	0.12	0.57	1
Dorzolamide	IOP-lowering (carbonic anhydrase inhibitor)	−1.4	0.21	1.58×10^{-11}	7.58×10^{-9}	−1.46	0.35	3.40×10^{-5}	7.7×10^{-3}	−0.36	0.15	0.02	1
Brinzolamide	IOP-lowering (carbonic anhydrase inhibitor)	−1.61	0.27	2.68×10^{-9}	1.29×10^{-6}	−2.01	0.49	3.75×10^{-5}	8.5×10^{-3}	−0.32	0.18	0.07	1
Travoprost	IOP-lowering (prostaglandin analog)	−1.3	0.23	2.45×10^{-8}	1.18×10^{-5}	−1.81	0.41	1.07×10^{-5}	0.02	−0.18	0.15	0.25	1
Carteolol	IOP-lowering (β blocker)	−2.15	0.55	7.95×10^{-5}	0.04	−3.8	1.02	0.0002	0.05	−0.98	0.33	0.002	0.61
Codeine	Analgesic (opiate)	0.26	0.04	2.65×10^{-9}	1.27×10^{-6}	0.68	0.09	1.32×10^{-15}	3.0×10^{-13}	0.02	0.03	0.58	1
Pregabalin	Analgesic and anticonvulsant (gabapentinoid)	0.59	0.14	1.63×10^{-5}	0.008	1.15	0.28	3.88×10^{-5}	0.09	0.12	0.09	0.16	1
Tramadol	Analgesic (opiate)	0.35	0.08	2.41×10^{-5}	0.012	0.99	0.17	1.98×10^{-9}	4.5×10^{-7}	−0.05	0.05	0.4	1
Acetaminophen	Analgesic (aniline analgesic)	0.09	0.02	0.0001	0.07	0.26	0.04	2.59×10^{-9}	5.9×10^{-7}	0.02	0.01	0.3	1
Gabapentin	Analgesic and anticonvulsant (GABA analog)	0.25	0.11	0.03	1	0.78	0.21	0.0001	0.41	−0.04	0.07	0.63	1
Ibuprofen	Analgesic (NSAID)	0.05	0.03	0.04	1	0.19	0.05	0.0001	0.25	0	0.02	0.95	1
Metformin	Oral hypoglycemic (biguanide)	−0.17	0.05	0.0008	0.37	−0.05	0.1	0.6	1	−0.15	0.03	2.14×10^{-6}	0.0005
Glibornuride	Oral hypoglycemic (sulfonylurea)	−0.26	0.09	0.003	1	−0.08	0.16	0.6	1	−0.24	0.05	1.50×10^{-5}	0.004
Allopurinol	Gout (xanthine oxidase inhibitor)	−0.22	0.08	0.003	1	−0.32	0.15	0.036	1	−0.26	0.04	5.14×10^{-9}	1.23×10^{-6}
Salbutamol	Bronchodilator (β2 adrenergic receptor agonist)	0.1	0.04	0.006	1	0.33	0.07	5.42×10^{-6}	0.012	0.01	0.02	0.80	1
Ipratropium	Bronchodilator (anticholinergic bronchodilator)	0.45	0.16	0.006	1	0.6	0.31	0.06	1	0.36	0.1	0.0005	0.1

Models were adjusted for sex, age, age squared, years of education, Townsend deprivation index, and number of medications taken concomitantly. The results are shown for medications that passed Bonferroni correction in at least one of the groups.

*Average difference in spherical equivalent between participants under the medication and those not reporting taking the same medication (the LMM coefficient).

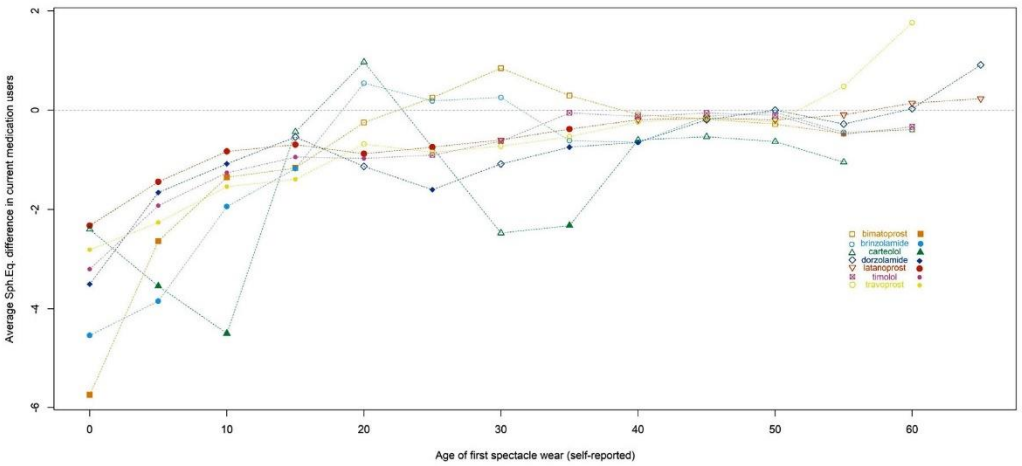


FIGURE 2. The differences in spherical equivalent between subjects receiving IOP-lowering medication and subjects who were not receiving them, by age of first spectacle wear. The AOSW shown on the *x*-axis is the first year of each of the 5-year periods in which the individuals started correcting their refractive error (e.g., 0–5, 5–10), and the *y*-axis denotes the adjusted difference in spherical equivalent observed in each group. The medications are specified in the figure legend, where a *symbol* to the left of a medication name denotes results that are not statistically significant (see Methods), and a *solid symbol* to the right indicates a statistically significant difference.

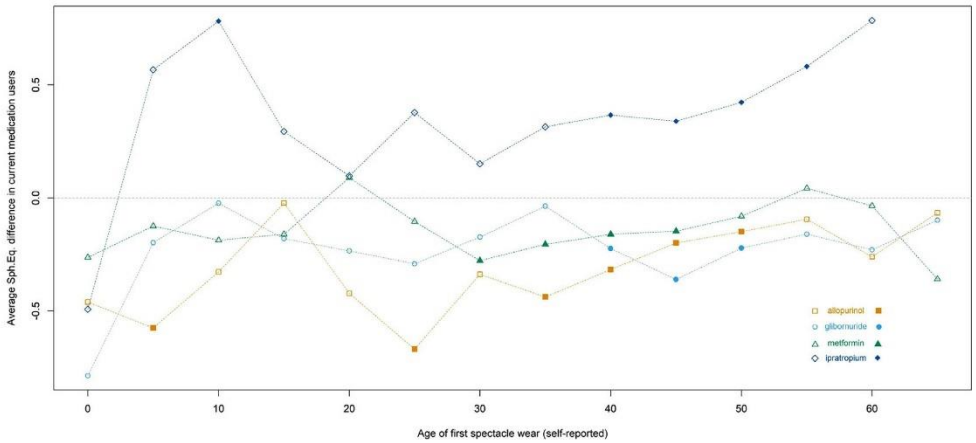


FIGURE 3. The differences in spherical equivalent between subjects receiving antihyperglycemic, anti-uricemic or ipratropium medication and subjects who were not receiving them, by age of first spectacle wear. The AOSW shown on the *x*-axis is the first year of each of the 5-year periods in which the individuals started correcting their refractive error (e.g., 0–5, 5–10), and the *y*-axis denotes the adjusted difference in spherical equivalent observed in each group. The medications are specified in the figure legend, where a *symbol* to the left of a medication name denotes results that are not statistically significant (see Methods), and a *solid symbol* to the right indicates a statistically significant difference.

age in which these medications are typically prescribed. For the group that started wearing glasses or contact lenses at a later age (AOSW > 35 years), there was no statistically significant difference in SpE between individuals prescribed IOP-lowering medication and those not.

Two other groups of drugs showed associations with SpE and AOSW. First, three drugs usually prescribed for metabolic disturbances showed a statistically significant association with lower (i.e., more myopic) SpE among participants who started wearing corrective lenses or glasses later in life (Table). Two of them were oral antihyperglycemic agents, metformin ($\beta = -0.15$, SE = 0.03, $P = 2.14 \times 10^{-6}$) and glibornuride ($\beta = -0.24$, SE = 0.05, $P = 1.5 \times 10^{-5}$). Allopurinol was also significantly associated with myopia ($\beta = -0.26$, SE = 0.04, $P = 5.14 \times 10^{-9}$) among individuals whose refraction was corrected at or after the fourth decade (Table). Ipratropium was associated with a positive SpE (hyperopia), particularly among participants with late AOSW ($\beta = 0.36$, SE = 0.1, $P = 0.0005$), becoming progressively stronger with increasing AOSW. These associations were particularly significant among individuals who started refractive correction after the fourth decade of life (Fig. 3), but because we lacked suitable instrumental variables (single nucleotide polymorphisms [SNPs] significantly

Investigative Ophthalmology & Visual Science

Downloaded from iovs.arvojournals.org on 06/30/2021

associated with their intake in addition to the diseases they are indicated for) we were unable to draw firm conclusions about the direction of causality. We also found a positive and significant association with salbutamol ($\beta = 0.33$, SE = 0.07, $P = 5.42 \times 10^{-6}$), another bronchoactive medication (Table), but only in the group with early AOSW.

Unexpectedly, our analyses revealed several significant associations between refractive error and a pharmacologically heterogeneous group of drugs for which the main clinical indication is pain management and analgesia. They were all associated with higher mean SpE (i.e., hyperopia). Participants receiving codeine ($\beta = 0.68$, SE = 0.09, $P = 1.32 \times 10^{-15}$) and tramadol ($\beta = 0.99$, SE = 0.17, $P = 1.97 \times 10^{-9}$), either alone or in combination with other non-opioid analgesics, were significantly more hyperopic than participants with no known history of opiate use (Table). Other analgesics, such as tramadol ($P = 1.97 \times 10^{-9}$), codeine ($P = 1.32 \times 10^{-15}$), paracetamol ($P = 2.58 \times 10^{-9}$), ibuprofen ($P = 0.0001$), pregabalin ($P = 3.88 \times 10^{-5}$), and gabapentin ($P = 0.0002$), were also associated with hyperopia in the group who wore spectacles early in life (AOSW between 5 and 35 years) (Table, Supplementary Fig. S6), before the typical age of presbyopia.

Other analgesic medications were nominally associated but below our strict multiple testing correction significance threshold. Typically, analgesics were used in combination with other medications, and the strength of their association was blunted by our correction for the total number of medications taken. For example, two additional nonsteroidal antiinflammatory medications (naproxen and diclofenac) were significantly associated with SpE when the number of total medications was removed from the model (Bonferroni-corrected $P = 0.02$ and $P = 0.04$, respectively).

There was no statistically significant relationship between the strength of these associations and comorbidity (Supplementary Figs. S7, S8, S9, S10, S11, S12) or physical activity (Supplementary Table S3), and associations with analgesic drugs remained significant.

Mendelian Randomization Models of Causality

The likelihood of taking any particular medication is correlated with pre-existing disease.[20] Any observed SpE differences among the users of certain medications may be driven by the presence of underlying conditions or the primary effect of the medication on eye structures. To explore causality, we used MR techniques to assess the relationship between the primary disease for which these medications were indicated and SpE (Supplementary Table S4). In addition, because the same medication is often prescribed to treat several different diseases, we used MR to assess the relationship between different medication classes and refractive error (Supplementary Table S4). We used SNPs significantly associated with the primary diseases or medication taking as instruments and their effect over SpE[27] for the outcome.

MR analyses results supported previously published suggestions that higher IOP caused lower (more myopic) SpE[27] (MR–Egger $P = 0.02$), with no significant directional pleiotropy (intercept $P = 0.256$). This remained statistically significant after correcting for overlap of study samples (maximum likelihood method adjusted for sample overlap $P = 0.001$). MR also suggested that this relationship is not mediated by IOP-lowering medications.[20]

MR analyses using instruments from previous GWASs[28] revealed strong evidence of pleiotropy, and not causation, between type 2 diabetes and SpE (MR–Egger $P < 0.001$, intercept $P = 0.001$). However, we found statistical evidence suggesting that oral antidiabetic medications[20] may directly and causally contribute to refractive error (MR–Egger $P = 0.003$, intercept $P = 0.166$).

We found no evidence for a causal relationship between chronic obstructive pulmonary disease[29] (the main therapeutic indication for ipratropium) and SpE (MR–Egger $P = 0.23$). Neither gout[30] nor uric acid levels[31] had any statistically significant causal influence on refractive error, although the instruments selected for these analyses were weak and potentially unreliable (only nine and one instruments available for these two traits, respectively).

Our analyses found a causative relationship between multisite chronic pain[32] and myopia (Supplementary Table S4). Although phenotypically the traits are not correlated (Pearson's $r = 0.05$) and SNPs that are associated with multisite chronic pain are generally not associated with SpE at statistically significant levels (Supplementary Table S5), individuals with progressively higher centiles of multisite chronic pain PRS distribution have increasingly higher average measurements of SpE (Supplementary Fig. S13). MR models find that this relationship is statistically significant and that multisite chronic pain directly contributed to hyperopia (MR–Egger $P = 0.02$, intercept $P = 0.15$). This relationship remained significant even after correcting for overlap of samples in which exposure and outcome effects were measured (maximum-likelihood method $P < 0.001$) (Supplementary Table S4, Supplementary Fig. S14). These findings suggest that there is a causal link between sensitivity to pain and refractive error and not specific diseases that causally influence SpE. We found no significant causal relationship between nonsteroid antiinflammatory drugs and SpE (MR–Egger $P = 0.18$). Likewise, salicylic acid and its derivatives were not causally associated with SpE (MR–Egger $P = 0.849$), and rheumatoid arthritis[33] was not causal to SpE (MR–Egger $P = 0.13$).

Discussion

Here we analyzed associations between medication and spherical equivalent in a large general population-based cohort. Previous large studies have found significant correlations cross-sectionally between IOP and negative spherical equivalent in children,[34,35] but not with the speed of myopia progression.[35] Our work suggests that IOP-lowering medications are associated with lower SpE (more myopic refraction) and the use of analgesics with higher SpE (more hyperopic refraction). These associations are particularly strong among individuals who develop refractive errors early in life. Previous studies have linked glaucoma and myopia, particularly high myopia, which is usually of very early onset (before the age of 10 years).[36] Lower (more myopic) SpE is associated with higher IOP in observational studies.[34,35] Only some recent studies suggest that IOP may be one of the contributing causes of myopia.[27] The association between IOP-lowering treatment and SpE could, therefore, be a consequence of elevated IOP.

The link between pain medication and hyperopia (longsightedness) is interesting, as it was unexpected. These associations remained statistically significant even after addressing potential sources of confounding and regardless of the subjects' comorbidities or current activity levels. The

Investigative Ophthalmology & Visual Science

Downloaded from iovs.arvojournals.org on 06/30/2021

mechanism behind these associations is not entirely clear. However, given the heterogeneity of opiates, nonsteroidal antiinflammatory drugs, paracetamol, and gabapentin and their very different modes of action, it is most likely that any direct effects on the eye do not cause associations.

Our MR analyses suggest a causal effect of multisite and chronic pain over hyperopic refractive error. Chronic pain is a complex trait that is not fully understood. Among the potential causes of hyperalgesia is a lower threshold for excitation of sensory neurons, possibly linked to genetic variants affecting gated and ungated cation channels.[37] Interestingly, refractive error is associated with genetic polymorphisms in the ungated cation channels and, more generally, with possible changes in the efficiency of visual sensory signal transduction.[38] In addition, pain induces activation of autonomic responses, including efferent sympathetic activity, leading to pupil dilation[39] and accommodation relaxation. This may possibly affect mechanisms driving the development of refractive error including, for example, dynamic changes in the choroid, or the magnitude and direction of blurring of the retinal image. However, the evidence provided here is purely probabilistic, and more work is needed to investigate possible mechanisms behind these novel findings.

Medications aimed at metabolic diseases such as diabetes (metformin and glibornuride) and hyperuricemia (allopurinol) were associated with a more negative spherical equivalent. These associations were particularly significant among individuals who started refractive correction after the fourth decade of life (Fig. 3), but, because we lacked suitable instrumental variables (SNPs significantly associated with their intake in addition to the diseases they are indicated for), we were unable to draw firm conclusions about the direction of causality.

We found evidence suggesting that oral antihyperglycemics may have some direct effects on lowering the spherical equivalent and potentially cause late-onset myopia. Given that the effects were more pronounced in the later decades in life when diabetes and antidiabetic therapies are started, myopia is unlikely to be due to axial length changes. Hyperglycemia thickens the crystalline lens[40] and increases its refractive index.[41] Autonomic dysfunction of sympathetic and parasympathetic activity in diabetes and changes in pupil size and lens power might also be relevant.[42] This evidence is probabilistic and does not refute other alternatives that we may have lacked sufficient power to test in full. For example, both diabetes and poor metabolic control of glucose have been associated with myopia development,[43] so we cannot fully exclude earlier life interactions between glucose and emmetropization.

Similarly, we hypothesize that the mechanisms through which allopurinol is associated with lower SpE may involve growing lens opacity as opposed to changes in axial length. Exposure to higher levels of uric acid may cause osmolar changes that, over time, alter the refraction index of the lens through mechanisms similar to those described before for diabetes.[44] Gout[45] (but not allopurinol[44]) may be associated with age-related cataract, which can cause SpE changes,[46] but our MR analyses were underpowered for this condition and could not detect any relationship between hyperuricemia and SpE. Also, the relationship between SpE and ipratropium, a drug commonly prescribed against chronic obstructive pulmonary disease, remains unexplained. Ipratropium is an anticholinergic drug,[47] similar to atropine, mild mydriatic effects of which may add to the likelihood of hyperopic correction later in life, especially in individuals with a degree of pre-existing hyperopia and presbyopia. A similar argument can also be made for salbutamol,[48] an adrenergic agonist with similar effects on the lens and pupil.

Strong statistical evidence for one mechanism cannot conclusively exclude the importance of concomitant mechanisms. For example, several medications significantly associated with lower SpE are sulfonamides (dorzolamide, brinzolamide, and glibornuride). Drugs in the sulfonamide class,[49] particularly dorzolamide,[50] have been linked to episodes of transient myopia, and the relative overrepresentation of these drugs in our results raises the possibility of the presence of a small direct myopia-inducing effect of these drugs.

The results of this study should be interpreted within the context and limitations of the design of the study from which they were obtained. Generalizing them to other populations should therefore warrant caution. For example, our sample included volunteers of European descent and had a higher percentage of university/college graduates and less deprivation than the population average. This study also did not have any information on medication dosage and length of the treatment. Incorporation of this information would have improved the power of the models. Our MR models had uneven power due to the differing number and quality of available instrumental variables. Finally, although previous research demonstrated good validity of self-reported medication intake,[51,52] recall bias and misclassification are potential issues, especially among older participants. False-positive results could arise from residual confounding, despite our attempts to minimize this.

To our knowledge, our work is the largest cohort-based epidemiological study exploring associations of commonly used medications with refractive error in the general population. We identified several classes of medications associated with spherical equivalent changes, some of which may be attributable to the medication and not merely the underlying conditions. More research is needed to replicate the novel associations between analgesic use and hyperopic refractive error and to clarify the mechanisms connecting them to spherical equivalent.

Acknowledgments

Supported by a project grant from Fight for Sight (Ref 5037/5038). The funder was not involved in the study design, data collection, analyses, or writing of this journal article. All authors worked independently of the funder.

Disclosure: **K. Patasova,** None; **A.P. Khawaja,** None; **B. Tamraz,** None; **K.M. Williams,** None; **O.A. Mahroo,** None; **M. Freidin,** None; **A.L. Solebo,** None; **J. Vehof,** None; **M. Falchi,** None; **J.S. Rahi,** None; **C.J. Hammond,** None; **P.G. Hysi,** None

References

1. Vos T, Abajobir AA, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet.* 2017;390(10100):1211–1259.
2. Fricke T, Holden B, Wilson D, et al. Global cost of correcting vision impairment from uncorrected refractive error. *Bull World Health Organ.* 2012;90(10):728–738.
3. Pan C-W, Dirani M, Cheng C-Y, Wong T-Y, Saw S-M. The age-specific prevalence of myopia in Asia: a meta-analysis. *Optom Vis Sci.* 2015;92(3):258–266.

Downloaded from iovs.arvojournals.org on 06/30/2021

4. Williams KM, Bertelsen G, Cumberland P, et al. Increasing prevalence of myopia in Europe and the impact of education. *Ophthalmology*. 2015;122(7):1489–1497.
5. Morgan IG, French AN, Ashby RS, et al. The epidemics of myopia: aetiology and prevention. *Prog Retin Eye Res*. 2018;62:134–149.
6. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*. 2016;123(5):1036–1042.
7. Rahi JS, Cumberland PM, Peckham CS. Myopia over the lifecourse: prevalence and early life influences in the 1958 British birth cohort. *Ophthalmology*. 2011;118(5):797–804.
8. Williams KM, Kraphol E, Yonova-Doing E, Hysi PG, Plomin R, Hammond CJ. Early life factors for myopia in the British Twins Early Development Study. *Br J Ophthalmol*. 2019;103(8):1078–1084.
9. French AN, Morgan IG, Mitchell P, Rose KA. Risk factors for incident myopia in Australian schoolchildren: the Sydney Adolescent Vascular and Eye Study. *Ophthalmology*. 2013;120(10):2100–2108.
10. Tideman JWL, Polling JR, Hofman A, Jaddoe VW, Mackenbach JP, Klaver CC. Environmental factors explain socioeconomic prevalence differences in myopia in 6-year-old children. *Br J Ophthalmol*. 2018;102(2):243–247.
11. Williams KM, Hysi PG, Nag A, Yonova-Doing E, Venturini C, Hammond CJ. Age of myopia onset in a British population-based twin cohort. *Ophthalmic Physiol Opt*. 2013;33(3):339–345.
12. Flitcroft DI. The complex interactions of retinal, optical and environmental factors in myopia aetiology. *Prog Retin Eye Res*. 2012;31(6):622–660.
13. Lim LS, Gazzard G, Low Y-L, et al. Dietary factors, myopia, and axial dimensions in children. *Ophthalmology*. 2010;117(5):993–997.e4.
14. Williams KM, Verhoeven VJM, Cumberland P, et al. Prevalence of refractive error in Europe: the European Eye Epidemiology (E3) Consortium. *Eur J Epidemiol*. 2015;30(4):305–315.
15. Pan C-W, Saw S-M, Wong T-Y. Epidemiology of myopia. In: Spaide RF, Ohno-Matsui K, Yannuzzi LA , eds. *Pathologic Myopia*. New York: Springer; 2014:25–38.
16. Murphy RM, Bakir B, O'Brien C, Wiggs JL, Pasquale LR. Drug-induced bilateral secondary angle-closure glaucoma: a literature synthesis. *J Glaucoma*. 2016;25(2):e99–e105.
17. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–209.
18. Chua SYL, Thomas D, Allen N, et al. Cohort profile: design and methods in the eye and vision consortium of UK Biobank. *BMJ Open*. 2019;9(2):e025077.
19. Cumberland PM, Bao Y, Hysi PG, et al. Frequency and distribution of refractive error in adult life: methodology and findings of the UK Biobank Study. *PLoS One*. 2015;10(10):e0139780.
20. Wu Y, Byrne EM, Zheng Z, et al. Genome-wide association study of medication-use and associated disease in the UK Biobank. *Nat Commun*. 2019;10(1):1891.
21. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. *BMC Geriatr*. 2017;17(1):230.
22. Cheesman R, Coleman J, Rayner C, et al. Familial influences on neuroticism and education in the UK Biobank. *Behav Genet*. 2020;50(2):84–93.
23. Hill WD, Hagenaars SP, Marioni RE, et al. Molecular genetic contributions to social deprivation and household income in UK Biobank. *Curr Biol*. 2016;26(22):3083–3089.
24. Townsend P, Phillimore P, Beattie A. *Health and Deprivation: Inequality and the North*. London: Croom Helm; 1988.
25. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. Available at: http://arxiv.org/abs/1406.5823 . Accessed February 4, 2021.
26. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw*. 2017;82(1):1–26.
27. Hysi PG, Choquet H, Khawaja AP, et al. Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. *Nat Genet*. 2020;52(4):401–407.
28. Zhao W, Rasheed A, Tikkanen E, et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat Genet*. 2017;49(10):1450–1457.
29. Sakornsakolpat P, Prokopenko D, Lamontagne M, et al. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet*. 2019;51(3):494–505.
30. Nakayama A, Nakaoka H, Yamamoto K, et al. GWAS of clinically defined gout and subtypes identifies multiple susceptibility loci that include urate transporter genes. *Ann Rheum Dis*. 2017;76(5):869–877.
31. Shin S-Y, Fauman EB, Petersen A-K, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014;46(6):543–550.
32. Johnston KJA, Adams MJ, Nicholl BI, et al. Genome-wide association study of multisite chronic pain in UK Biobank. *PLoS Genet*. 2019;15(6):e1008164.
33. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*. 2014;506(7488):376–381.
34. Jiang WJ, Wu JF, Hu YY, et al. Intraocular pressure and associated factors in children: the Shandong children eye study. *Invest Ophthalmol Vis Sci*. 2014;55(7):4128–4134.
35. Li S-M, Iribarren R, Li H, et al. Intraocular pressure and myopia progression in Chinese children: the Anyang Childhood Eye Study. *Br J Ophthalmol*. 2019;103(3):349–354.
36. Vision NRC (US) C on. Appendix C: review of the progression literature. Available at: https://www.ncbi.nlm.nih.gov/books/NBK235051/ . Accessed August 15, 2018.
37. Fillingim R, Wallace M, Herbstman D, Ribeiro-Dasilva M, Staud R. Genetic contributions to pain: a review of findings in humans. *Oral Dis*. 2008;14(8):673–682.
38. Tedja MS, Wojciechowski R, Hysi PG, et al. Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nat Genet*. 2018;50(6):834–848.
39. Bouffard MA. The pupil. *Contin Lifelong Learn Neurol*. 2019;25(5):1194.
40. Charman WN, Adnan, DA. Gradients of refractive index in the crystalline lens and transient changes in refraction among patients with diabetes. *Biomed Opt Express*. 2012;3(12):3033–3042.
41. Fledelius HC, Miyamoto K. Diabetic myopia–is it lens-induced? An oculometric study comprising ultrasound measurements. *Acta Ophthalmol (Copenh)*. 2009;65(4):469–473.
42. Pittasch D, Lobmann R, Behrens-Baumann W, Lehnert H. Pupil signs of sympathetic autonomic neuropathy in patients with type 1 diabetes. *Diabetes Care*. 2002;25(9):1545–1550.
43. Jacobsen N, Jensen H, Lund-Andersen H, Goldschmidt E. Is poor glycaemic control in diabetic patients a risk factor of myopia? *Acta Ophthalmol (Copenh)*. 2008;86(5):510–514.
44. Luo C, Chen X, Jin H, Yao K. The association between gout and cataract risk: a meta-analysis. *PLoS One*. 2017;12(6):e0180188.

Investigative Ophthalmology & Visual Science

Downloaded from iovs.arvojournals.org on 06/30/2021

45. Li Y-J, Perng W-T, Tseng K-Y, Wang Y-H, Wei JC-C. Association of gout medications and risk of cataract: a population-based case-control study. *QJM*. 2019;112(11):841–846.
46. Panchapakesan J, Rochtchina E, Mitchell P. Myopic refractive shift caused by incident cataract: the Blue Mountains Eye Study. *Ophthalmic Epidemiol*. 2003;10(4):241–247.
47. Baigelman W, Chodosh S. Bronchodilator action of the anticholinergic drug, ipratropium bromide (Sch 1000), as an aerosol in chronic bronchitis and asthma. *Chest*. 1977;71(3):324–328.
48. Barisione G, Baroffio M, Crimi E, Brusasco V. Beta-adrenergic agonists. *Pharmaceuticals*. 2010;3(4):1016–1044.
49. Grinbaum A, Ashkenazi I, Gutman I, Blumenthal M. Suggested mechanism for acute transient myopia after sulfonamide treatment. *Ann Ophthalmol*. 1993;25(6):224–226.
50. Tsai J-C, Chang H-W. Refractive change after dorzolamide use in patients with primary open-angle glaucoma and ocular hypertension. *J Ocul Pharmacol Ther*. 2001;17(6):499–504.
51. Drieling RL, LaCroix AZ, Beresford SAA, Boudreau DM, Kooperberg C, Heckbert SR. Validity of self-reported medication use compared with pharmacy records in a cohort of older women: findings from the Women's Health Initiative. *Am J Epidemiol*. 2016;184(3):233–238.
52. Hafferty JD, Campbell AI, Navrady LB, et al. Self-reported medication use validated through record linkage to national prescribing data. *J Clin Epidemiol*. 2018;94:132–142.

Investigative Ophthalmology & Visual Science

Downloaded from iovs.arvojournals.org on 06/30/2021

Supplementary Information:

Supplementary Table 1. Characteristics of the study participants (N = 102, 318). The fields "N,%" and "mean, se", "median, IQR" denote a number of the study participants, percentage, means, standard errors, medians and interquartile range respectively. Spherical equivalent cut-off points that were used to define emmetropia, hyperopia and myopia are provided in brackets

Demographic characteristics	
Age (N= 102,218) (median, IQR)	59 (51-64)
Sex (N = 102,218) (N, %)	
women	54,444 (53)
men	47,874 (47)
Townsend deprivation index (N = 102, 218) (median, IQR)	-2.01 (-3.53-0.45)
Years of education (N = 101, 470) (mean, se)	13 (10-20)
Refractive error-related variables	
Age of first spectacle wear (N = 93,357) (mean, se)	32 (16.2)
Spherical equivalent (N =102,318) (mean, se)	-0.3 (2.7)
Emmetropia (-0.5 < SpE< 0.5 Diopters) (N, %)	45,322 (44)
Hyperopia (SpE >=1 Diopters) (N, %)	28,944 (28)
Myopia (SpE <= -1 Diopters) (N, %)	28,052 (27)

Supplementary Table 2. Associations between polypharmacy and socioeconomic and demographic factors. The fields "beta", "se", "OR", "p-value" denote the regression coefficient (ln(OR), the coefficients standard errors of estimation, odds ratios and the association p-values from the logistic regression model

Factor	Beta	OR	se	p-value
Sex (male)	-0.005	0.99	0.016	0.73
Age	0.04	1.03	0.003	1.92×10^{-36}
Years of education	-0.03	0.97	0.002	1.30×10^{-60}
Townsend deprivation index	0.12	1.13	0.008	1.09×10^{-52}
Birth cohort (1940 - 1949)	-0.74	0.47	0.07	1.68×10^{-25}
Birth cohort (1950 - 1959)	-1.01	0.36	0.08	1.39×10^{-35}
Birth cohort (1960 - 1970)	-1.38	0.25	0.1	3.95×10^{-42}

Supplementary Table 3. Results of mixed linear model regression analyses, testing association between refractive error and analgesic medications. Models were adjusted for sex, age, age^2, years of education, Townsend deprivation index, polypharmacy and measures of physical activity. Columns "Medication" and "Physical activity measures" list the names of the tested drugs as well as physical activity measures that were entered into a model as covariates to adjust for potential confounding. The fields "Beta", "se", "95% CI" and "p-value" denote respectively the regression coefficient (slope), the standard errors of the coefficient estimation, the coefficients lower and upper bonds of the 95% confidence intervals, and p-values for each tested medication. The results are shown for medications that passed the sensitivity analyses. The analyses were performed in the sub-sample of study participants with AOSW between 5 and 35 years (N = 38,960)

Physical activity measures	Medication	Beta	se	95% CI	p-value
Frequency of walking (days per week)	tramadol	0.97	0.17	0.64 - 1.30	8.08×10^{-09}
	codeine	0.64	0.09	0.47 - 0.81	1.07×10^{-13}
	acetaminophen	0.26	0.04	0.18 - 0.35	2.05×10^{-09}
	ibuprofen	0.19	0.05	0.09 - 0.28	9.58×10^{-05}
Frequency of moderate physical activity (days per week)	tramadol	0.90	0.17	0.56 - 1.23	1.99×10^{-07}
	codeine	0.63	0.09	0.46 - 0.80	1.09×10^{-12}
	acetaminophen	0.26	0.04	0.17 - 0.34	6.71×10^{-09}
	ibuprofen	0.19	0.05	0.09 - 0.28	0.0001
Frequency of vigorous physical activity (days per week)	tramadol	0.94	0.17	0.60 - 1.27	3.32×10^{-08}
	codeine	0.69	0.09	0.52 - 0.86	3.80×10^{-15}
	acetaminophen	0.26	0.04	0.17 - 0.35	4.48×10^{-09}
	ibuprofen	0.18	0.05	0.09 - 0.28	0.0002
Duration of daily walking (minutes)	tramadol	1.00	0.19	0.63 - 1.37	1.04×10^{-07}
	codeine	0.64	0.10	0.43 - 0.84	9.74×10^{-10}
	acetaminophen	0.28	0.05	0.19 - 0.37	2.08×10^{-09}
	ibuprofen	0.23	0.05	0.12 - 0.32	1.01×10^{-05}
Duration of daily moderate physical (minutes)	tramadol	0.95	0.21	0.54 - 1.36	5.26×10^{-06}
	codeine	0.66	0.11	0.44 - 0.89	4.77×10^{-09}
	acetaminophen	0.28	0.05	0.18 - 0.38	3.26×10^{-08}
	ibuprofen	0.18	0.05	0.07 - 0.29	0.0009
Duration of daily vigorous physical (minutes)	tramadol	1.14	0.25	0.65 - 1.62	4.23×10^{-06}
	codeine	0.52	0.13	0.27 - 0.77	5.52×10^{-05}
	acetaminophen	0.28	0.06	0.17 - 0.39	4.85×10^{-07}

Physical activity measures	Medication	Beta	se	95% CI	p-value
	ibuprofen	0.20	0.06	0.08 - 0.31	0.00097
24-hour recall of light physical activity	tramadol	0.95	0.21	0.53 - 1.37	9.54×10^{-06}
	codeine	0.56	0.12	0.33 - 0.80	2.95×10^{-06}
	acetaminophen	0.22	0.05	0.11 - 0.32	3.87×10^{-05}
	ibuprofen	0.20	0.06	0.09 - 0.31	0.0004
24-hour recall of moderate physical activity	tramadol	0.96	0.21	0.53 - 1.38	8.57×10^{-06}
	codeine	0.58	0.11	0.37 - 0.79	1.10×10^{-07}
	acetaminophen	0.22	0.05	0.11 - 0.32	3.15×10^{-05}
	ibuprofen	0.20	0.06	0.09 - 0.31	0.0004
24-hour recall of vigorous physical activity	tramadol	0.97	0.21	0.55 - 1.39	6.78×10^{-06}
	codeine	0.57	0.12	0.34 - 0.81	1.99×10^{-06}
	acetaminophen	0.22	0.05	0.11 - 0.32	3.94×10^{-05}
	ibuprofen	0.20	0.06	0.09 - 0.31	0.0004

Supplementary Table 4. Results of Mendelian randomisation analyses testing the causal association between refractive error and 12 different traits. Column "Exposure" lists the names of the tested traits. Field "PMID" contains PubMed database publication numbers for the studies that provided MR exposure instruments. The studies that used UK Biobank data are marked with "*". The instruments for the outcome –refractive error, were selected from the most recent GWAS meta-analysis (PMID: 32231278). Number of instrumental variables (IVs) is the number of SNPs used in the analyses, and can be found in the column "Number of IVs". Column "Test" includes the names of the tests that were used to test causality. The fields "Beta", "se", "95% CI" and "p-value" denote respectively the MR coefficient estimated by each method, the standard errors of the coefficient estimation, the coefficients lower and upper bonds of the 95% confidence intervals and p-values for each tested trait. The units of the coefficients denote the number of Diopters (positive or negative) change when the exposure changes by 1 unit of the exposure (if the exposure is quantitative), or the difference of Diopters between cases and controls if the exposure is a categorical variable.

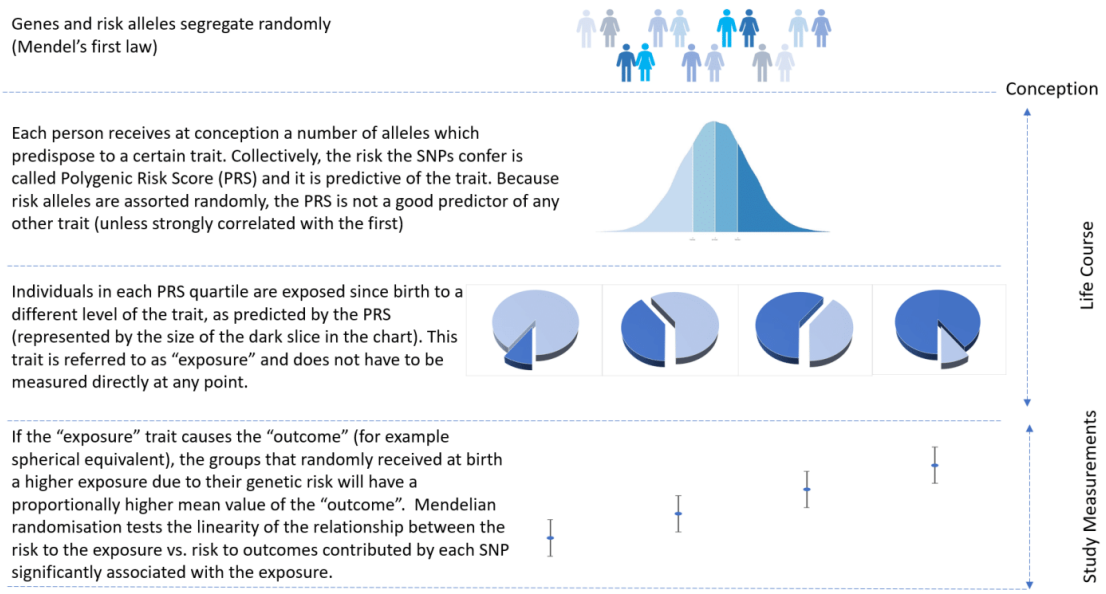
Exposure	PMID	Number of IVs	Test	Beta	Se	95 % CI	p-value
Intraocular pressure	29785010*	73	Simple median	-0.04	0.02	-0.088 - -0.009	0.016
			IVW	-0.045	0.013	-0.075 - -0.023	< 0.001
			MR-Egger	-0.087	0.038	-0.16 - -0.01	0.02
			(intercept)	-0.005	0.004	-0.004 - 0.01	0.256
			Maximum-likelihood method	-0.046	0.014	-0.073 - -0.02	0.001
IOP-lowering drugs	31015401*	14	Simple median	-0.043	0.007	-0.056 - -0.029	< 0.001
			IVW	-0.043	0.011	-0.064 - -0.022	< 0.001
			MR-Egger	-0.053	0.036	-0.124 - 0.018	0.15
			(intercept)	0.002	0.006	-0.01 - 0.014	0.78
Diabetes	30054458*	139	Simple median	-0.022	0.006	-0.033 - -0.01	< 0.001
			IVW	-0.022	0.008	-0.037 - -0.007	0.004
			MR-Egger	-0.071	0.017	-0.105 - -0.038	< 0.001
			(intercept)	0.004	0.001	0.001 - 0.006	0.001
Drugs used in diabetes	31015401*	56	Simple median	-0.024	0.0001	-0.024 - -0.024	< 0.001
			IVW	-0.028	0.007	-0.042 - -0.015	< 0.001
			MR-Egger	-0.049	0.016	-0.081 - -0.017	0.003
			(intercept)	0.003	0.002	-0.001 - 0.006	0.166
Chronic multisite pain	31194737*	38	Simple median	0.195	0.04	0.116 - 0.274	< 0.001
			IVW	0.201	0.044	0.115 - 0.287	< 0.001
			MR-Egger	0.505	0.218	0.079 - 0.932	0.02
			(intercept)	-0.005	0.004	-0.012 - 0.002	0.153
			Maximum-likelihood method	0.204	0.046	0.114 - 0.295	< 0.001

Exposure	PMID	Number of IVs	Test	Beta	Se	95 % CI	p-value
Anti-inflammatory and antirheumatic products, non-steroids	31015401*	6	Simple median	0.006	0.032	-0.057 - 0.07	0.844
			IVW	-0.02	0.072	-0.16 - 0.12	0.779
			MR-Egger	-0.541	0.411	-1.345 - 0.264	0.188
			(intercept)	0.025	0.019	-0.013 - 0.063	0.199
Salicylic acid and derivatives	24390342*	10	Simple median	0.011	0.016	-0.021 - 0.042	0.51
			IVW	0.026	0.015	-0.003 - 0.054	0.08
			MR-Egger	0.007	0.041	-0.074 - 0.088	0.873
			(intercept)	0.001	0.003	-0.004 - 0.007	0.621
Rheumatoid arthritis	24390342*	30	Simple median	0.008	0.004	0.000 - 0.016	0.05
			IVW	0.008	0.003	0.001 - 0.014	0.022
			MR-Egger	0.008	0.005	-0.002 - 0.017	0.134
			(intercept)	0	0.001	-0.001 - 0.002	0.958
Adrenergics inhalants	31015401*	54	Simple median	0.001	0.007	-0.013 - 0.015	0.91
			IVW	-0.003	0.01	-0.022 - 0.016	0.77
			MR-Egger	-0.031	0.028	-0.086 - 0.024	0.265
			(intercept)	0.003	0.002	-0.002 - 0.007	0.28
Chronic obsructive pulmonary disease	30804561*	47	Simple median	-0.015	0.021	-0.056 - 0.026	0.487
			IVW	-0.018	0.023	-0.064 - 0.027	0.434
			MR-Egger	-0.081	0.068	-0.215 - 0.052	0.23
			(intercept)	0.003	0.003	-0.002 - 0.008	0.321
Gout	27899376	9	Simple median	-0.011	0.014	-0.038 - 0.017	0.435
			IVW	-0.005	0.013	-0.031 - 0.02	0.675
			MR-Egger	0.067	0.03	0.007 - 0.126	0.027
			(intercept)	-0.013	0.005	-0.024 - -0.003	0.011
Uric acid	24816252	1	Simple median	-0.6678	0.2466	-0.914 - -0.421	0.007
			IVW	-0.369	0.18	-0.549 - -0.189	0.04
			MR-Egger	0.216	0.306	-0.09 - 0.522	0.481
			(intercept)	-0.009	0.0036	-0.013 - -0.005	0.024

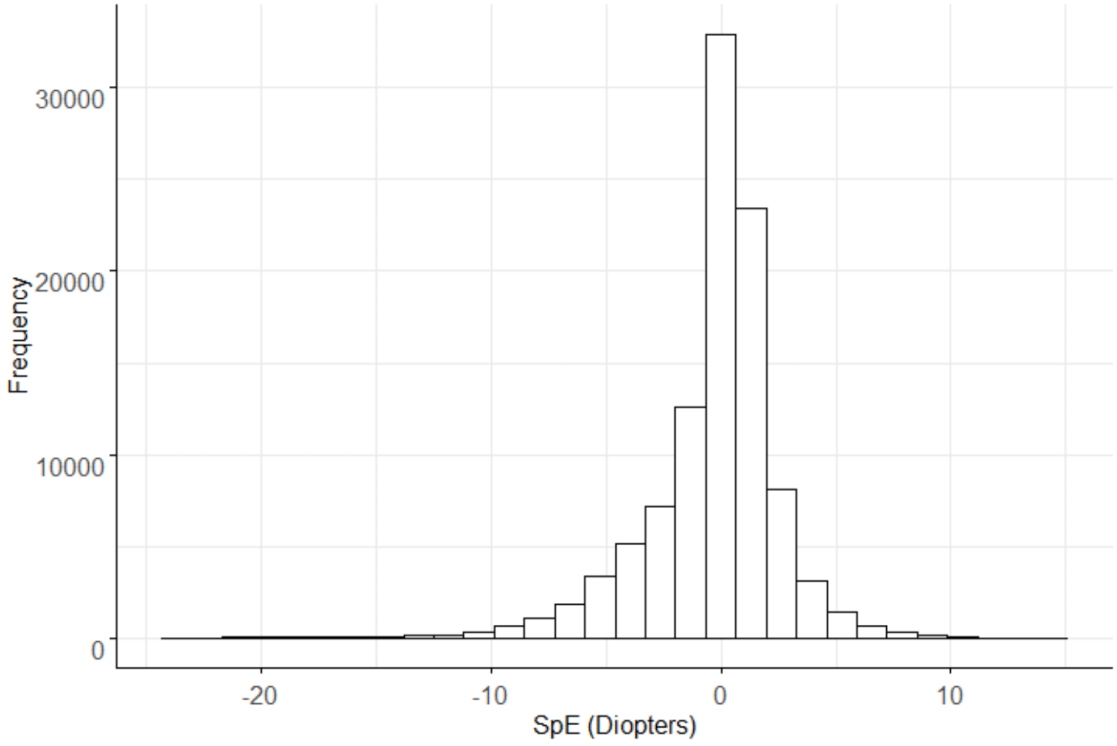
Supplementary Table 5. Association between SE and SNPs that have previously been shown to be associated with multisite chronic pain, in the UK Biobank. "Beta" denotes the linear regression coefficient, "SE" its standard errors, and "P-value" the association probabilities. The field "SNPs" contains the name of the associated polymorphisms, followed by the reference allele for which the effects are being reported in this table.

SNPs	Beta	se	P-value
rs10888692_C	0.017	0.012	0.15
rs197422_C	0.036	0.012	0.0031
rs59898460_T	0.010	0.012	0.40
rs12071912_C	0.012	0.013	0.34
rs1443914_T	0.005	0.012	0.68
rs12435797_G	0.039	0.015	0.01
rs2006281_C	0.027	0.012	0.03
rs2386584_T	0.015	0.012	0.21
rs285026_G	-0.019	0.012	0.11
rs11871043_T	0.019	0.012	0.12
rs11079993_G	0.014	0.012	0.25
rs62098013_G	0.019	0.012	0.11
rs4852567_A	0.008	0.013	0.53
rs2424248_G	0.021	0.017	0.23
rs7628207_T	0.014	0.016	0.39
rs28428925_G	-0.006	0.017	0.72
rs6770476_C	-0.019	0.013	0.14
rs34811474_G	0.009	0.014	0.55
rs13135092_A	0.101	0.021	8.18×10^{-07}
rs13136239_G	0.024	0.012	0.05
rs6869446_T	-0.003	0.012	0.80
rs1976423_A	0.006	0.012	0.61
rs17474406_G	0.049	0.036	0.17
rs1946247_T	0.014	0.017	0.42
rs11751591_G	0.018	0.016	0.28
rs6907508_A	0.025	0.018	0.18
rs6926377_A	0.019	0.013	0.15
rs10259354_G	0.004	0.013	0.79
rs7798894_A	0.010	0.014	0.45
rs6966540_T	0.021	0.012	0.08
rs12537376_A	-0.007	0.012	0.55
rs11786084_G	0.033	0.012	0.007
rs10992729_C	0.019	0.013	0.14
rs6478241_A	0.003	0.012	0.81
chr9:140251458_G_A_G	0.019	0.018	0.29
rs2183271_T	0.004	0.012	0.74
rs11599236_T	0.031	0.012	0.009
rs12765185_T	-0.003	0.013	0.79
rs61883178_C	-0.004	0.016	0.82

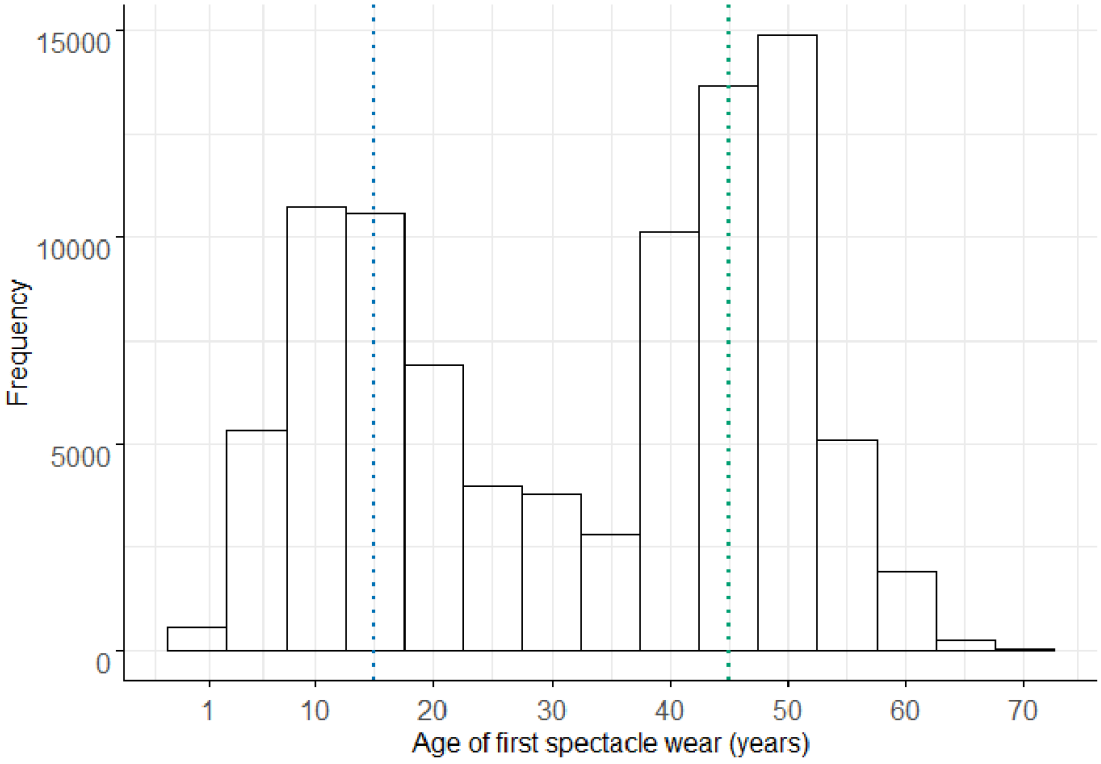
Supplementary Figure 1. The general principles of Mendelian Randomization. From top to bottom: the risk alleles are re-shuffled after each meiosis, they are assorted independently and randomly so they create unique risk profiles predicting only one trait (the "exposure"). When increased levels of "exposure" lead the specific outcome, progressively higher levels of the outcome phenotype will be observed with proportionally higher levels of randomly assigned levels of genetic risk.



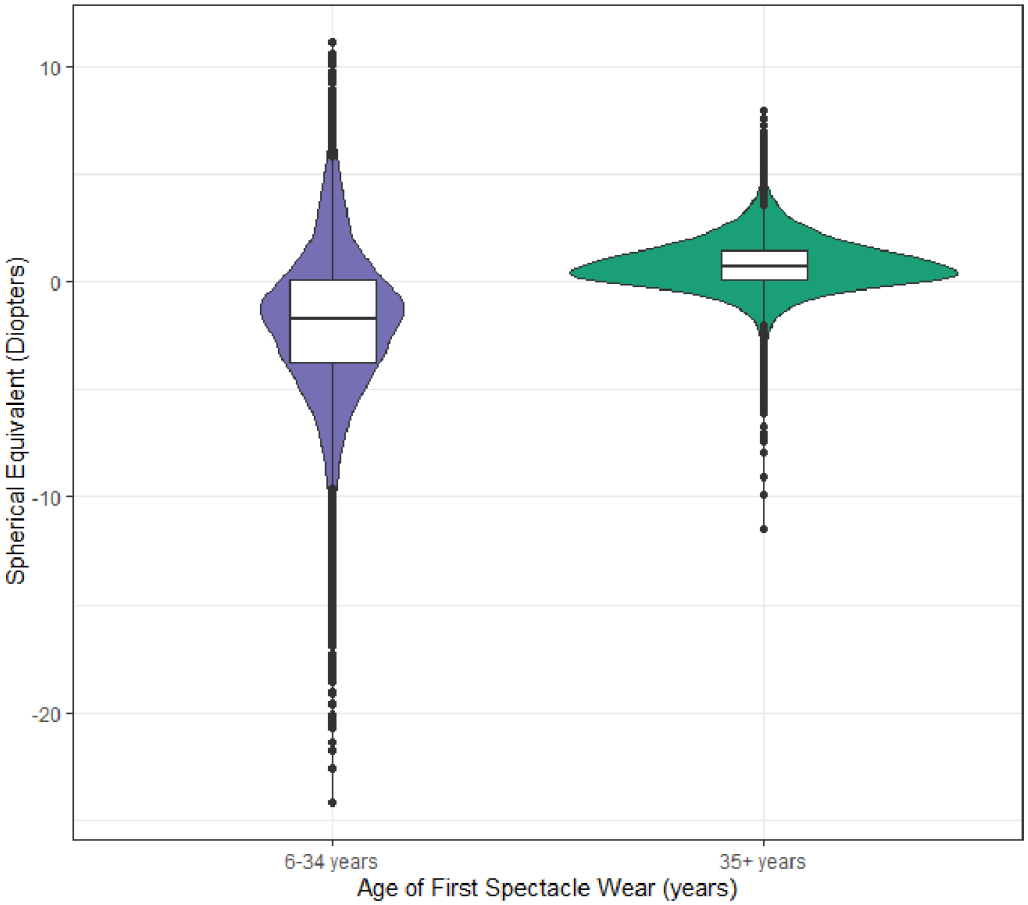
Supplementary Figure 2. Distribution of spherical equivalent (in Diopters) in UK Biobank cohort. The spherical equivalent is shown on the x-axis and the corresponding number of UK Biobank participants included in our work, for each spherical equivalent bracket, is given in the y-axis.



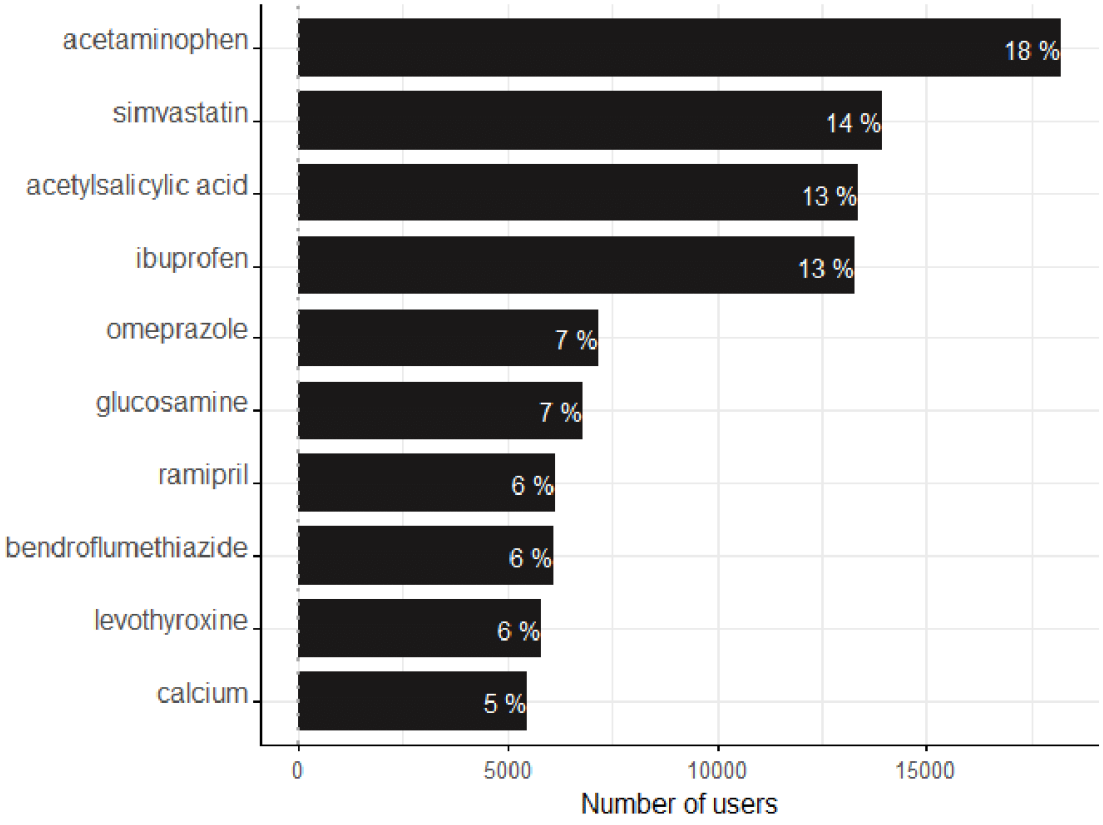
Supplementary Figure 3. Distribution of the age of first spectacle wear in UK Biobank cohort. The age of first spectacle wear is shown on the x-axis and the corresponding number of UK Biobank participants included in our work, for each spherical equivalent bracket, is given in the y-axis. The dashed lines represent the peaks of the modes of the distribution.



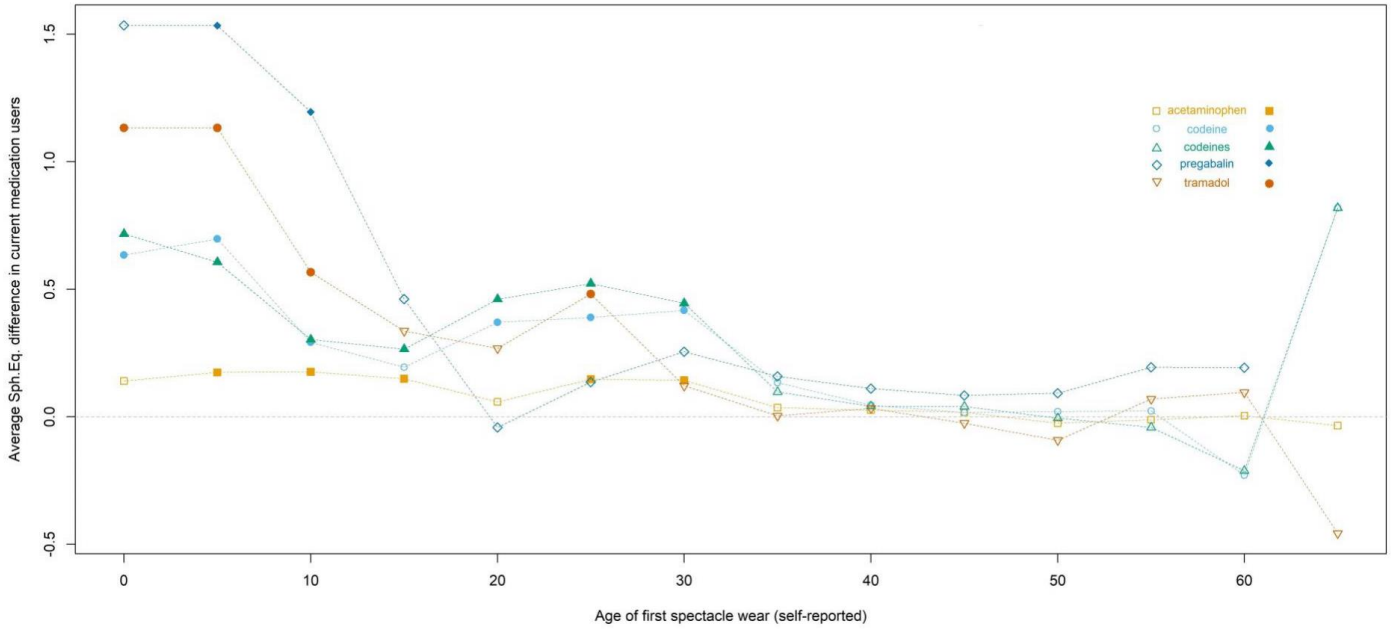
Supplementary Figure 4. Distribution of refractive error across the age of first spectacle wear in UK Biobank cohort (N = 87, 200). In the horizontal axes are shown the two main modes of the age of spectacle wear. The shape of the violin plots denotes the distribution of the spherical equivalent within the respective group. The central horizontal bar represents the mean spherical equivalent measurement (vertical axis) across all individuals in the respective categories; the box shows the interquartile range and the whiskers represent the observations in the 1.5 × IQR.



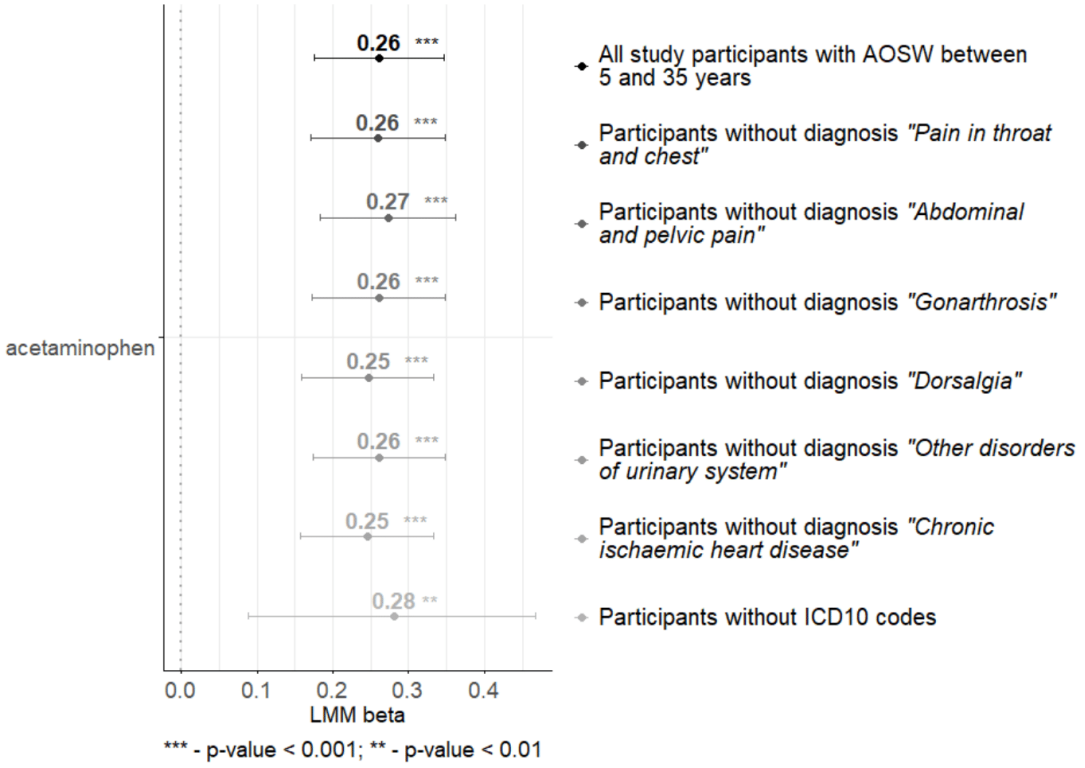
Supplementary Figure 5. Medication use in UK Biobank cohort. Information is limited to the ten most commonly reported drugs in the subset used in our study (N = 481). The numbers inside the bars represent the percentage of the study participants that reported taking medications.



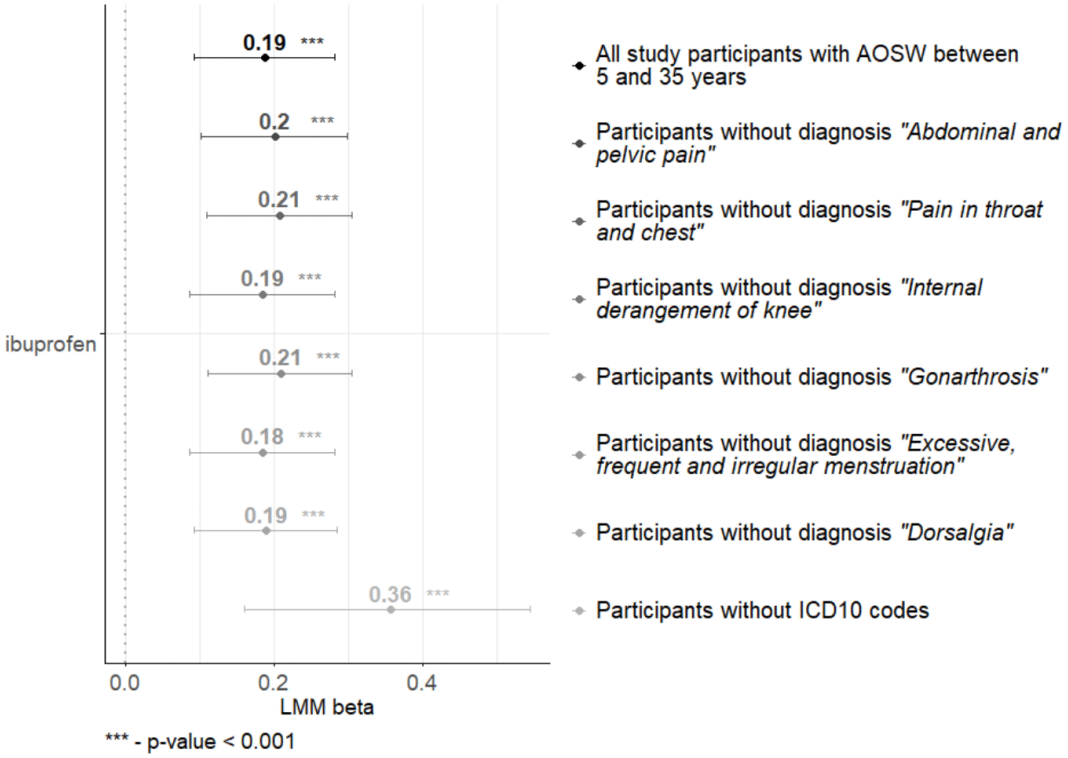
Supplementary Figure 6. The differences in spherical equivalent between subjects receiving analgesic medication and subjects who were not receiving them, by age of first spectacle wear. The AOSW shown in the x-axis is the first year of each of the 5-year periods in which the individuals started correcting their refractive error (i.e. 0-5, 5-10, etc.) and the y-axis denotes the adjusted difference in spherical equivalent observed in each group. The medications are specified in the Figure legend, with the symbol to the left of their names denoting results that are not statistically significant (see Methods) and the solid symbol to their right statistically significant differences.



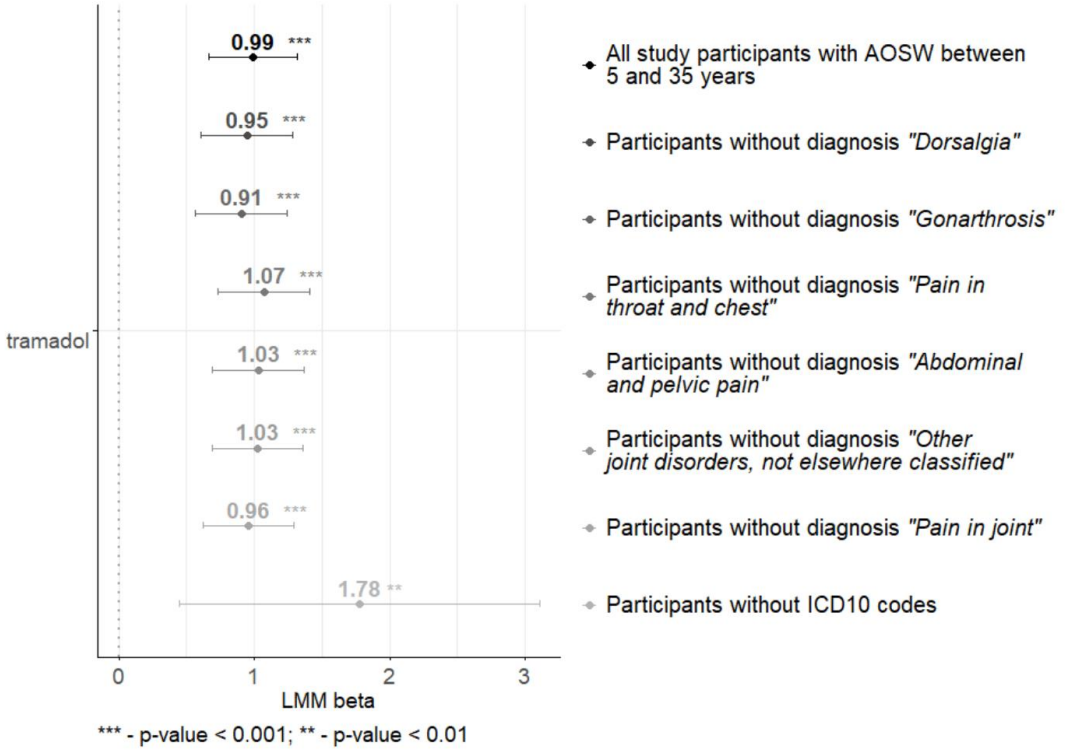
Supplementary Figure 7. Association between refractive error and acetaminophen use. Sensitivity analyses models' betas and their 95% confidence intervals are represented by error bar plots. Significance shown by the symbols ***, ** and * denote associations with p< 0.001, p < 0.01 and p < 0.05 respectively. Grey colour gradient represents the decreasing sample size. The legend keys follow the order of error bar plots. The resampling procedure was based on ICD10 codes – individuals with selected diagnoses were sequentially removed from the analyses



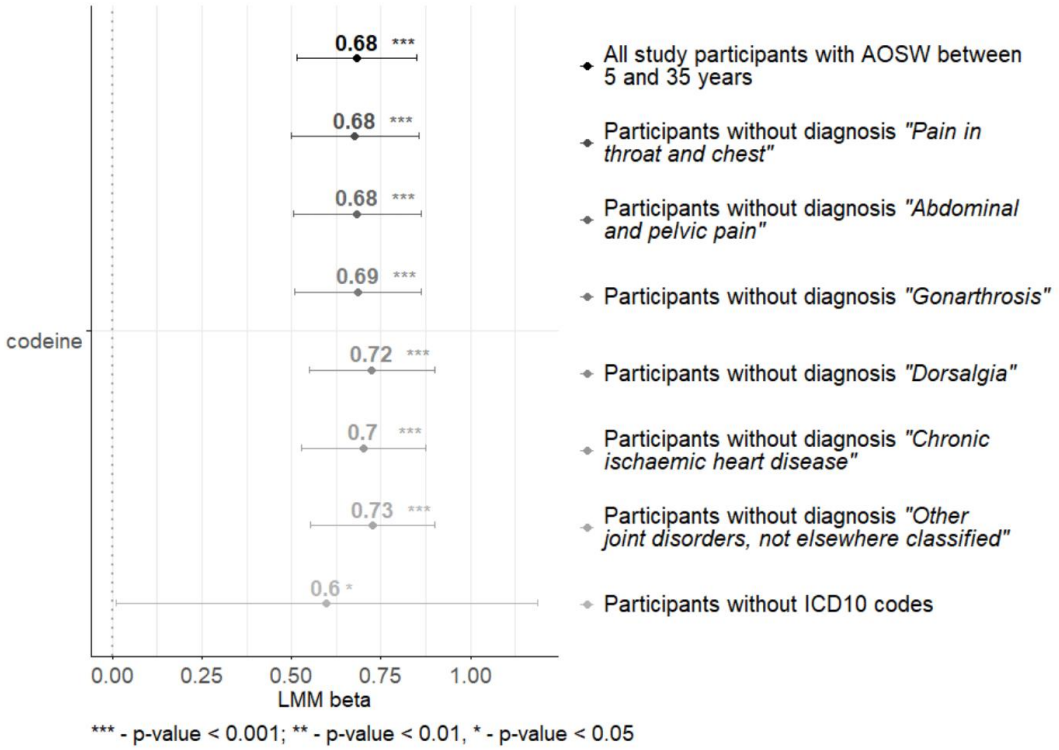
Supplementary Figure 8. Association between refractive error and ibuprofen use. Sensitivity analyses models linear model effect estimates, and their 95% confidence intervals are represented by error bar plots. Significance shown by the symbols ***, ** and * denote associations with p< 0.001, p < 0.01 and p < 0.05 respectively. Grey colour gradient represents the decreasing sample size. The resampling procedure was based on ICD10 codes – individuals with selected diagnoses were sequentially removed from the analyses



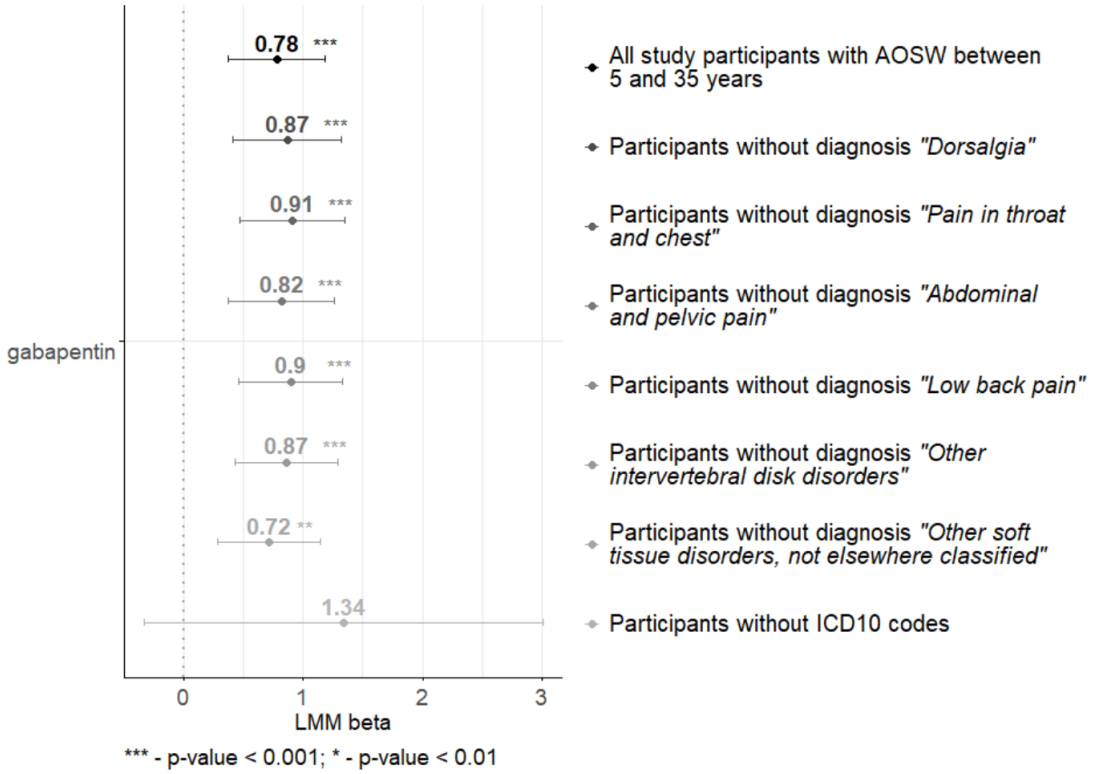
Supplementary Figure 9. Association between refractive error and tramadol use. Sensitivity analyses models' betas and their 95% confidence intervals are represented by error bar plots. Significance shown by the symbols ***, ** and * denote associations with p< 0.001, p < 0.01 and p < 0.05 respectively. Grey colour gradient represents the decreasing sample size. The resampling procedure was based on ICD10 codes – individuals with selected diagnoses were sequentially removed from the analyses



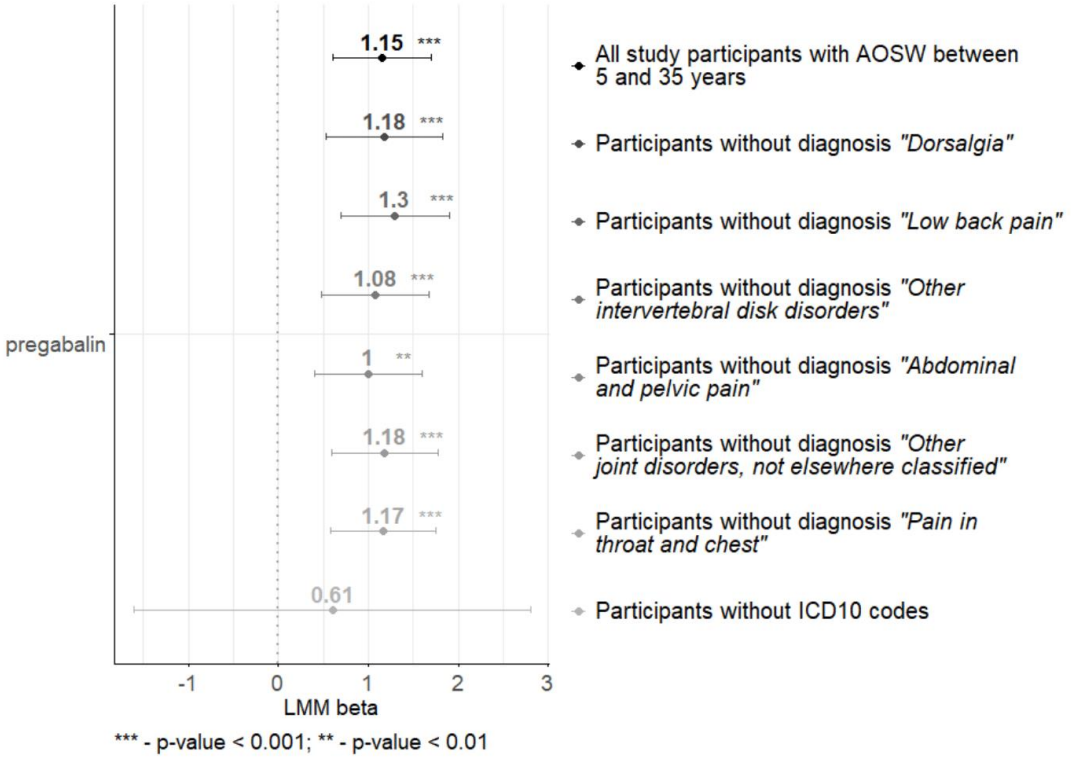
Supplementary Figure 10. Association between refractive error and codeine use. Sensitivity analyses models' betas and their 95% confidence intervals are represented by error bar plots. Significance shown by the symbols ***, ** and * denote associations with p< 0.001, p < 0.01 and p < 0.05 respectively. Grey colour gradient represents the decreasing sample size. The resampling procedure was based on ICD10 codes – individuals with selected diagnoses were sequentially removed from the analyses



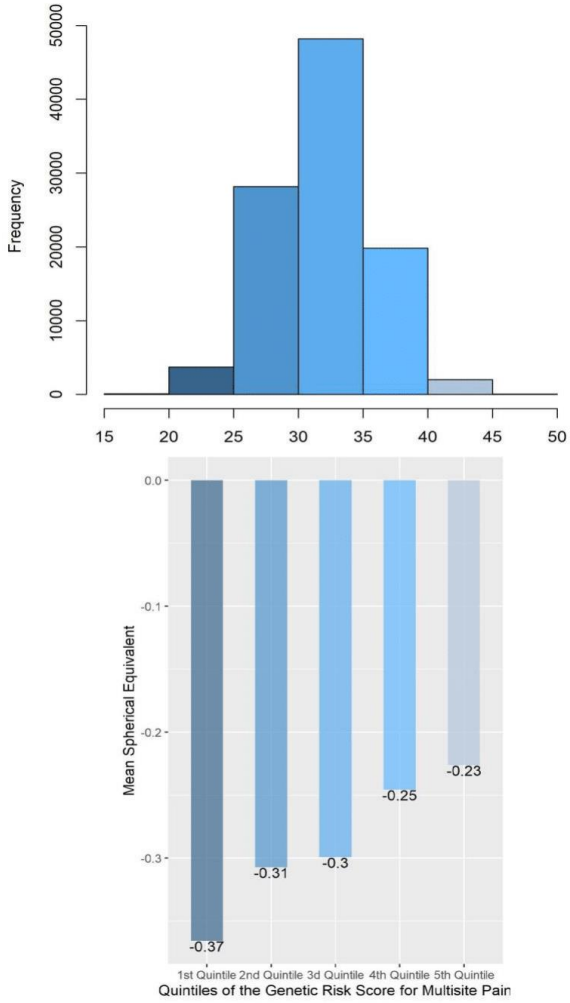
Supplementary Figure 11. Association between refractive error and gabapentin use. Sensitivity analyses models' betas and their 95% confidence intervals are represented by error bar plots. Significance shown by the symbols ***, ** and * denote associations with p< 0.001, p < 0.01 and p < 0.05 respectively. Grey colour gradient represents the decreasing sample size. The resampling procedure was based on ICD10 codes – individuals with selected diagnoses were sequentially removed from the analyses



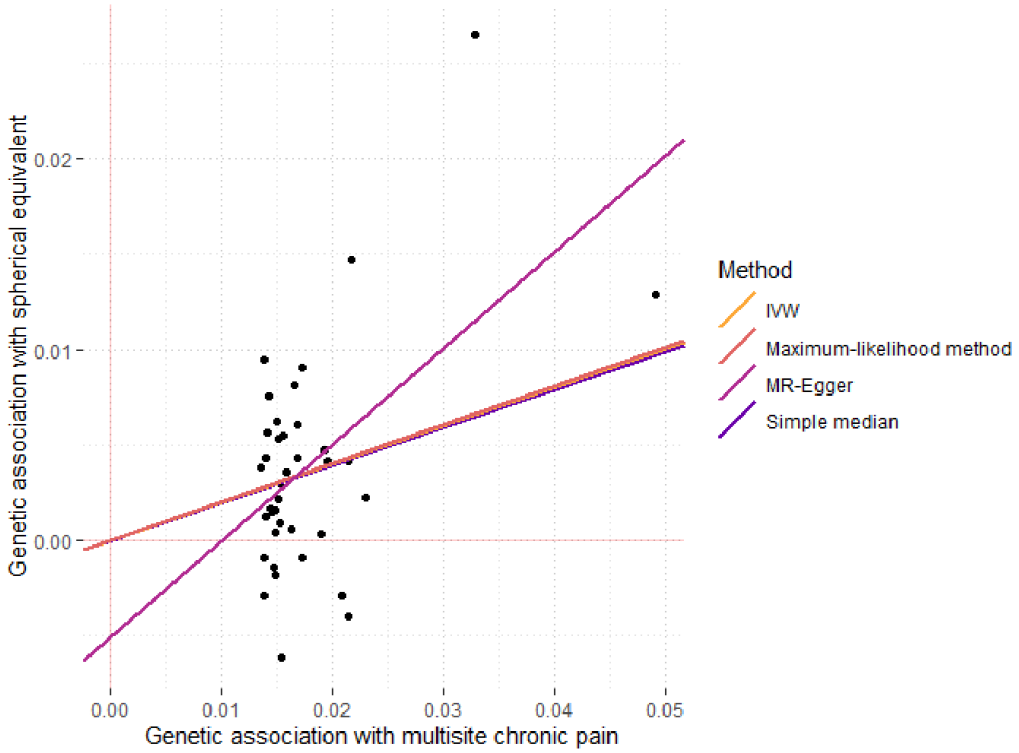
Supplementary Figure 12. Association between refractive error and pregabalin use. Sensitivity analyses models' betas and their 95% confidence intervals are represented by error bar plots. Significance shown by the symbols ***, ** and * denote associations with p< 0.001, p < 0.01 and p < 0.05 respectively. Grey colour gradient represents the decreasing sample size. The resampling procedure was based on ICD10 codes – individuals with selected diagnoses were sequentially removed from the analyses



Supplementary Figure 13. The mean spherical equivalent observed in each quintile of the chronic multisite pain Polygenic Risk Score. Although chronic multisite pain and spherical equivalent are only weakly correlated (r=0.05), increasingly hypermetropic average levels of spherical equivalent are reliably observed for rising PRS values. The result of the test statistics for this relationship are reported in Supplementary Table 4.



Supplementary Figure 14. Results of Mendelian Randomisation tests for causality of multisite pain on spherical equivalent. Each of the points in the figure represents one instrumental variable (SNP) that has been selected on the basis of significant ($p < 5x10^{-08}$) association with multisite pain. The Mendelian randomization tests the relationship of the effect sizes of the associations of these instruments with multisite chronic pain ("exposure, x-axis) compared to the spherical equivalent (the "outcome", y-axis). The differently coloured lines represent the results of each of the tests used, as described in the Methods.



Chapter 8 | Discussions of results arising from this thesis, conclusions and possible future directions

8.1 Contributions to research and implications

The genetic work described in this thesis contributed significantly to the understanding of the genetic architecture of refractive disorders (110). Meta-analysis presented in this thesis built on previous genetic research and identified association with 438 discrete genomic regions. Cumulatively, associated genetic polymorphisms predicted 18.4% of refractive error heritability compared to only 8% of variance explained by previously published results (110). The prediction model incorporating the PRS markedly improved the genetic prediction of common near-sightedness in European adults demonstrating a high AUC of 0.74 – 0.75 for moderate and severe myopia, respectively (99).

The PRS-model based on SNPs associated with refractive error in Europeans also accurately predicted high and moderate myopia in Asian schoolchildren. The PRS alone outperformed a prediction model that included conventional measurement of parental myopia. The combination of genetic risk alleles and parental myopia with other factors such as time spent outdoors significantly improved the prediction of severe myopia risk in Asian teenagers, yielding one of the highest AUC reported for this age group. These results indicated the potential clinical value of the current PRS model which could be used to detect high-risk children and guide myopia prevention strategies.

In addition to building and training a PRS-based myopia prediction model that was independently validated in the sample of Singaporean schoolchildren, I performed post-GWAS analyses that facilitated the discovery of major mechanisms driving the pathogenesis of refractive disorders. In particular, I found evidence indicating significant pleiotropy between CNS and refractive error. Through the work that I did on eQTL, it was revealed that genes involved in refractive error were significantly enriched in CNS tissues. Strong genetic correlations with multiple neurocognitive traits and subsequent MR analyses suggested shared genetic effects with cognitive functions, and higher CNS regulatory mechanisms, each contributing to smaller proportions of refractive error

variance. Work in which I contributed to, for the first time provided statistical support linking circadian rhythm genes to refractive development in humans. In the functional analyses, which I performed, significant enrichment was seen in gene-sets associated with circadian entrainment. Gene enrichment for binding sites of the transcription factors involved in circadian rhythm regulation was also observed. Particularly, fifteen different refractive error genes identified in the meta-GWAS of Europeans (99) were previously reported as associated with circadian entrainment. For example, *PROX1, CREB1, PPARGC1A, RORB, RORA, TCF4* and *SUV39H1* regulated circadian rhythm, while *ZFHX3*, *PARGC1A*, *NRIP1, SIX3* and *DRD1* modulated diurnal gene expression or influenced circadian behaviour and sleep/wake cycle. Interestingly, some of these genes, such as the homeobox-containing class (454), coordinate retinal progenitor cell proliferation and cell fate specification (455), but also play a role in pituitary gland development (456,457) and are required for superchiasmatic nucleus circadian output (458). Recent evidence suggests that retinal development is in part regulated by circadian clock-related pathways. In particular, during retinal embryonic phase as well as postnatally circadian clocks regulate precursor proliferation and differentiation, by ensuring appropriate timing of these processes. In addition to guiding retinal development, the clock appears to integrate environmental cues and modulate a number of pathological processes in the retina and other ocular compartments, including myopisation (454).

The hypothesis that disordered rhythms and clock genes contributed to the development of refractive disorders gained the most support from experimental research (332). Growing evidence implicated diurnal and circadian rhythm in postnatal eye growth and refractive development (324). In both animals and human eyes, anatomical and physiological features such as axial length, choroidal thickness and IOP, underwent diurnal oscillations. Systematically, these natural rhythms were coordinated by the ''master clock'' within the suprachiasmatic nucleus, which received signals from

photosensitive ganglion cells via activation of photopigment melanopsin. Retina also contained an endogenous pacemaker (clock) that regulated retinal physiology by adapting vision to the day-night cycle and modulating retinal signalling to the circadian system. Circadian biology was controlled by light exposure, and dysregulation of circadian rhythm caused by patterns of ambient lighting was shown to perturb normal diurnal oscillations of ocular parameters in animal models. Alternatively, retinal neurotransmitter dopamine synchronised intrinsic retinal rhythms to the day-night cycle but was also involved in retinal signalling and refractive development, providing another explanation for the relationship between circadian rhythmicity and refractive development. Circadian clocks are based on transcriptional-translational feedback loops utilising clock genes and their proteins expressed by the retina (324).

Work, which I led and contributed to, characterised molecular mechanisms driving the refractive development in the general population and identified over 400 common variants, explaining a considerable proportion of spherical equivalent variability and heritability (99). Identified refractive error susceptibility loci greatly improved risk prediction for common myopia and could be utilised in the near future to create personalised strategies for myopia prevention. Additionally, this study uncovered two major sets of mechanisms indistinguishable from previous smaller GWAS (99). The meta-analysis found evidence of elevated IOP causing myopic changes which carried significant implications for future research as well as myopia prevention and treatment strategies (99). Moreover, in addition to local ocular mechanisms, meta-analysis implicated broader CNS functions (99). The results of this investigation demonstrated that disturbances in circadian rhythm could lead to refractive errors in humans – a hypothesis which was previously supported largely by data from animal experiments and limited investigations of ocular growth in human eyes, but not by genetic research (324).

Chapter 5 of this thesis presented the results of refractive error WES conducted on a subsample of the UK Biobank participants of European descent. The study examined genes with a focus on rare mutations within protein-coding regions of the genome. Gene-based analyses revealed eight genes associated with spherical equivalent, five of which were not documented in previous refractive error GWAS. The statistically strongest association was found for *SIX6*, the SIX/sine oculis family homeodomain transcription factor implicated in early eye development and highly expressed in adult ciliary body, sclera, choroid, optic nerve head as well as the retina (346).

The genomic signal at the *SIX6* locus was narrowed down to a single potentially causal marker rs146737847, whose amino acid change (Glu129Lys) affected the homeobox DNA-binding domain. This evolutionarily conserved variant was independently associated with VCDR and increased the susceptibility to POAG in European (348) and Asian populations (459). However, UK Biobank WES was the first instance relating rs146737847 to refractive error variance.

One of the new candidates *CRX* demonstrated a second Bonferroni-significant relationship, while others showed suggestive evidence of association with refraction. *CRX* was nested within a broader genomic locus reported in the refractive error meta-analysis which was part of this thesis (99). However, UK Biobank WES was the first genetic investigation that provided direct evidence for the association between cone-rod homeobox gene and refractive error. *CRX* encoded a photoreceptor-specific transcription factor expressed in the inner nuclear layer, where it engaged with other transcription factors and promoted photoreceptor cell differentiation and maintenance (460). It has been suggested that CRX was regulating more than 700 genes in photoreceptor-specific cells, including its downstream genes rhodopsin (*RHO*) and cone arrestin (*ARR4*) (461). The expression of cone-rod homeobox protein in retinal neurocytes exhibited a luminescence-dependent decrease upon light exposure with similar downstream effects observed in rhodopsin and arrestin, whose upregulated

expression was implicated in myopia pathogenesis (462). With the evidence of light driving emmentropisation, Wu *et al.* speculated that *CRX* upregulation in the retinal neurocytes cultured in darkness pointed at the potential gene involvement in myopia formation (355), however, until now there was no evidence demonstrating its' effects in humans. The association between spherical error and the *CRX* also upheld the previous findings implicating circadian rhythm in refractive development and ocular growth. In addition to photoreceptors and RPE cells, the cone-rod homeobox gene was abundant in the pineal gland, where it exhibited daily expression rhythms (463) and expedited nocturnal melatonin synthesis (464).

The broad implications of refractive error WES were that variants with large effects over the population's spherical equivalent appeared to be extremely rare. Rare variants with strong effects generally cause Mendelian disorders and rare markers with small and intermediate effects were difficult to capture because of the limited statistical power of traditional analytical methods at that end of the minor allele frequency range (465). Previous genome-wide rare variant analyses that aggregated several related phenotypes and used samples of hundreds of thousands of sequenced participants, had reported more than 30 loci harbouring rare markers enriched for cardiometabolic traits (119). It is still possible that increasing sample sizes from analysis of the UK Biobank full cohort, or through collaborations of international cohorts will enable future discovery of additional rare variants associated with refractive error. This investigation also illustrated the advantages of high-quality whole-exome sequencing over array methods that showed power limitations and bias arising from population admixture (379). Recently, published large-scale UK Biobank analyses indicated that many genetic associations obtained using sequencing techniques could not be identified with chip-based methods. In many cases, polymorphisms with MAF < 1% were below the range of allele frequencies that could be confidently imputed. Moreover, the unique variant

composition of the UK Biobank exome-sequencing dataset was not accessible by chip but vital for the detection of novel associations with rare variants (117).

The genetic determinants of the age of first spectacle wear were studied in genome-wide time-to-event analyses performed on 340,318 participants of European descent. The study revealed genetic similarity and a strong genetic correlation between refractive error and the age of the first wear, but also the age of myopia onset (108). The findings of this investigation suggested that the effects of significant genetic associations strongly correlated with myopic refractive error, but changed with age, which was reported for few known refractive error genes. For instance, *ZMAT4*, *BMP3* and *TSPAN10* showed much earlier onset of myopia than other significant loci, with effects peaking in adolescence and diminishing after 30, whereas *SOX2-OT* conferred a lower risk of myopic refraction at a later age. Among 44 significant genomic regions, six were novel and not reported by prior GWAS. Most of the novel loci had no reported ocular effects but participated in higher CNS functions.

Understanding the genetic basis of age of first spectacle wear was critical for the early screening programs and gaining insight into the aetiology and progression of refractive disorders. Genetic determinants reported by case-control studies are predictive of measures of the age-at-diagnosis. Alleles associated with the later age of onset may influence pathophysiology through several mechanisms, some of which are not necessarily shared with the early-onset refractive error. Currently, only one study investigated genetic variants associated with the age of refractive correction (29). The publication presented in the thesis expanded the number of refractive error loci with known age-specific effects from previously published 20 to 44. In the future, these genotypes can be utilised to detect children who are more likely to develop myopia at a younger age and should be targeted for earlier and more aggressive myopia prevention. The key benefit of using genetic information instead of screening for low-level

hypermetropia is that genotyping can be performed at a very young age before the reduction of childhood hypermetropic refraction starts.

The biological mechanisms underlying the development of refractive disorders were further investigated utilising the data on self-reported medication-taking in the UK Biobank cohort. The strongest associations were observed between refractive error and different classes of IOP lowering medications, specifically three prostaglandin agonists, two carbonic anhydrase inhibitors and beta-blockers. The participants who reported taking anti-glaucoma preparations had spherical equivalent 1-2 Dioptres lower than that of other study subjects.

The results of animal studies suggested that ocular hypotensive drugs, such as brimonidine and timolol, could potentially prevent myopia (466) and attenuate its progression in guinea pigs (467) and chicks (468). Additionally, recent studies indicated that antimuscarinic antagonists, such as atropine, inhibited myopia via α- adrenoceptors (469) and not muscarinic acetylcholine receptors as previously believed (470). A report by Carr *et al.* testing the hypothesis that α-adrenoceptors could affect ocular growth in chicks revealed that intravitreal injection of brimonidine prevented axial elongation and inhibited experimentally induced myopic refraction (468). In the UK Biobank study, presented in the thesis, brimonidine was nominally associated with myopic refractive error (p-value= 0.002), although this relationship did not reach the level of Bonferroni adjustment significance. Nonetheless, observed in animal studies inhibition of lens and form-deprivation myopia by brimonidine (468) and latanoprost (471) treatments, respectively, was attributed to the reduction of IOP. Intraocular pressure has been proposed as contributing factor to accommodation-induced myopic refractive error. Elevated IOP was linked to ocular axial elongation and was reported as an independent susceptibility factor for myopia (472). Additionally, in observational studies participants with myopia exhibited significant IOP elevation compared to subjects with emmetropic vision (473). However, prospective investigations found no significant difference in the

intraocular pressure between children with myopic and normal refraction, although they documented a marked increase of IOP in the incident sub-sample following the onset of near-sightedness (237,474,475). By contrast, experimental studies indicated that axial elongation and myopia could be directly induced by intraocular pressure raised by intravitreal fluid injection (476). Consistently with these findings, it was observed that axial length decreased following the lowering IOP after a trabeculectomy which provided additional evidence for the causal relationship between intraocular pressure and axial myopia (477). Previous research also suggested that accommodation-induced IOP changes could affect myopia progression (478) Similarly, in the UK Biobank investigation observed correlation between myopic refractive error and anti-glaucoma medications was a consequence of increased IOP.

Another interesting, yet unexpected finding reported in this thesis was the association between self-reported use of pain management medications and hypermetropic refraction. Individuals receiving analgesics were significantly more hypermetropic than subjects that did not report regularly taking painkillers. Importantly, the associations were found for a heterogeneous group of pain control medications, including opioids, gabapentinoids and NSAIDs which suggested confounding by the underlying condition. Subsequent Mendelian Randomisation analyses uncovered a causal relationship between multisite chronic pain and hyperopic refraction, which was a completely novel finding, not described elsewhere. The pathophysiology of chronic pain is not completely understood, but studies implicated multiple biological and psychological processes related to immune response, metabolism, cellular growth and protein degradation (479). However, neurotransmission was overrepresented in all chronic pain conditions (480) Particularly, genetic evidence underscored the involvement of neural signalling, precisely imbalances in ascending nociceptive and descending inhibitory signalling (479). Genes associated with chronic pain were enriched for neuro projection guidance, CNS neuron differentiation, DCC-mediated attractive signalling, as well as neurogenesis

(481), with some of these processes also implicated in the pathogenesis of refractive error (99) and AFSW (108). The lower excitation of neurons in chronic pain conditions was linked to polymorphisms regulating cation and noncation channels (482). Interestingly, spherical error was influenced by genes encoding cation channels, and more broadly, related to the efficiency of visual signal transduction (99).

Another significant association was found for self-reported intake of medications aimed at metabolic diseases. Specifically, UK Biobank participants who were taking antihyperglycemic agents on average displayed spherical equivalent 0.15 – 0.24 lower than the non-user group. Subsequent Mendelian randomisation analyses indicated that anti-diabetes preparations could be causally associated with negative refraction. However, it was still possible that some of the observed correlations were owed to the association with diabetes and hypoglycaemia. Poor metabolic control and the high glycaemic load have been previously suggested as potential contributing factors to myopia development (185). Observational investigations found that the prevalence of myopia was considerably higher in patients with diabetes compared to the background population (185,483,484). Moreover, near-sightedness in individuals with diabetes was described as late-onset and with a lower degree of refractive error than in the general population (532). It has been hypothesised that myopic refractive changes seen in patients with diabetes could be lens-induced. In particular, hyperglycaemia appeared to thicken the crystalline lens and increase refractive index which caused the osmotic difference between the lens tissue and aqueous humour (185,187,485). Alternatively, pupillary autonomic neuropathy that often-preceded systemic dysfunction of the sympathetic and parasympathetic systems could also play part in the association between poor metabolic control and myopia (486).

Chronic inflammation underlies the pathogenesis of diabetes (487), chronic pain (482) and COPD (488) and could provide an alternative explanation for observed associations between different pharmacological classes and refractive error. A number of reports

highlighted the role of inflammation in myopia onset and progression. For example, in longitudinal cohort of patients with juvenile chronic arthritis (JCA) myopic refractive error was detected in greater fraction of affected individuals compared to an age-matched control group (229,489). Similarly, transient high myopia was documented in systemic lupus erythematosus (490–492) and attributed to anterior displacement of the iris lens diaphragm (491). Overall, patients with inflammatory disorders exhibited higher incidence of myopia than disease-free individuals. Myopic refraction was more common among patients with diabetes compared to the general population (38% vs 27%) and diabetes type I and II were among other morbidity factors predisposing to myopia (176,185,186,483,484). Chronic uveitis, that has a potential to induce acute or constitutive near-sightedness, was observed in 10–20% of individuals with JCA (493–495) and potentially contributed to myopia progression in these patients. Elevated expression levels of inflammatory cytokines, such as IL-1β and -6, TGF-β, TNF-α, IL-6 and TNF-α, were found in the serum of diabetes patients (487) and the aqueous humor of individuals affected by JCA (496).

The elevated levels of inflammatory cytokines supported the notion that chronic inflammatory or systemic acute state associated with these diseases possibly contributed to the occurrence of near-sightedness. In particular, TNF-α, TGF-β and IL-6, activated NF-κB, a transcription factor important for regulation of inflammatory response. Interestingly, TGF-β signalling via NF-κB targeted MMP2 (497) which was shown to cleave collagens, therefore accelerating myopia progression (497–499). Consistent with this hypothesis, TGF-β and MMP2 transcription levels were increased in the myopic eyes, whereas expression of *COL1* was supressed. Additionally, these changes were overset by administration of atropine which decreased expression of TGF-β and MMP2 and promoted COL1 expression in the retina and sclera (229). Atropine also downregulated other inflammatory cytokines, including IL-6 and TGF- α and c-Fos, whose expression was increased in the myopic eyes (229). Similarly, refractive error

shifted towards hyperopic values after treatment with immunosuppressive agent, such as cyclosporine, while the inflammatory stimulators induced myopic shift (229).

Nevertheless, despite proposed molecular mechanism implicating tumour necrosis factors and interleukins, it seems unlikely that association between ocular inflammation and near-sightedness could explain the rising prevalence rates in East and South-East Asia(167). There is no correlation between international prevalence of myopia and that of allergic rhino conjunctivitis in children and adolescents (500). The possibility that eye rubbing due to allergies can lead to myopic refraction through the damage to cornea, as seen in keratoconus (501), was not confirmed by subsequent investigations (502). The influence of time spent outdoors was also not explored in the observational studies reporting associations between inflammatory diseases and near-sightedness (167).

Another plausible mechanism was that lower vitamin D levels, naturally seen in children who spent less time outdoors, could be confounding association between myopia and inflammatory and autoimmune conditions (503). Studies in both humans and animals provided compelling evidence suggesting that through actions on systemic inflammation and insulin secretion and resistance, vitamin D supplementation decreased the incidence of diabetes type I and II and could be potentially utilised in the disease treatment and prevention (503). Interestingly, children and young adults with myopia exhibited significantly lower vitamin D levels than nonmyopic individuals (median 67.6 vs. 72.5 nmol) (504,505). However, these results were initially reported by cross-sectional investigations and later detailed longitudinal survival and Mendelian randomisation analyses (158) revealed no causal relationship between vitamin D and myopia. Vitamin D3 was biomarker of time spent outdoors whose residual effects confounded association with near-sightedness (506). In MR analyses, the effect of genetically determined vitamin D levels over degree of spherical error was indistinguishable from zero (507).

I investigated the impact of several disorders associated with chronic inflammation on the observed relationship between hypermetropic refractive error and analgesics. Dorsalgia, soft tissue and joint disorders as well as gonarthrosis were the most common diagnoses among UK Biobank participants who reported taking pain control medications. Nonetheless, removal of the study subjects with these diseases did not impact the significance of the association with pain management medications. Consistent with other MR studies (508), there was no statically significant relationship, neither pleiotropic or causal, between spherical error and rheumatoid arthritis.

UK Biobank study described in the thesis was the first attempt to systematically examine the relationship between medication intake and refractive error in a population-based cohort and utilise patterns of medication use and underlying morbidity to elucidate mechanisms involved in refractive development. This study provided evidence supporting the causal role of IOP in myopia development, but also for the first time identified chronic pain as a potential risk factor for hypermetropia. Significant associations were also identified for medications aimed at diabetes, gout and COPD; however, these relationships require further investigation.

8.2 Limitations

Limitations to this work should be acknowledged and were discussed in detail in each chapter. One of the most important methodological issues that have to be addressed is the distinction between correlation and causation. Different sensitivity and Mendelian Randomisation analyses were used in order to tackle the issue of causation. Although Mendelian randomisation methods were utilised to discriminate between causal associations and pleiotropic correlations, strong statistical evidence for one mechanism could not conclusively exclude the influence of other contributing mechanisms. Furthermore, some of the utilised MR models displayed uneven power due to the number or quality of the instruments, which was the reason for not finding conclusive

evidence for some of the detected associations. Moreover, it must be underscored that statistical evidence of causal association doesn't necessarily imply a true biological relationship. Thus, additional functional analyses are required to further understand some of the correlations, including the relationship between multisite chronic pain and hyperopic refractive error.

This thesis presented the findings of a genome-wide meta-analysis of refractive error and time-to-event analysis of the age of first spectacle wear. Although the insights from these studies contributed to a better understanding of the pathophysiology of refractive disorders the limitations of their methodology must be acknowledged. It is worth mentioning, that genome-wide association studies did not necessarily identify causal variants but instead captured genomic signals, which were driven by a local correlation of multiple SNPs due to linkage disequilibrium, complicating the discovery of actual causal variants (264). Moreover, most association signals reported by GWAS were mapped to in non-coding regions of the genome, thus the etiological and functional implications of these associations were less straightforward (509,510) compared to protein-altering mutations. Consequently, additional post-GWAS analyses were often needed to discern causal markers and their targeted genes, including functional and evolutionary analyses as well as resequencing and remapping of multi-ethnic and admixed populations (264). Furthermore, despite technological advancements and the development of new statistical methods, functional characterisation still had to be based on a hypothesis about the underlying mechanisms (264).

Another issue with detecting causal associations in GWAS was the reporting of numerous susceptibility loci. For example, it was determined that for such highly polygenic traits as height, close to a hundred thousand SNPs had to be identified in order to explain 80% of heritability, which suggested that a substantial fraction of all genes was implicated in the complex trait variance (511). Addressing this, a proposed "omnigenic" model argued that due to significant interconnection of various gene

networks all genes expressed in the disease-related cells were likely to have nontrivial effects on the risk of the disease, whose heritability could be explained almost entirely by the gene effects outside core pathways (511). Therefore, it was suggested that genes associated with the continuum between monogenic and polygenic forms of the disease were more likely to be implicated in the core biological pathways and should be prioritised for future genetic investigations (512)

Nonetheless, despite difficulties in interpreting GWAS hits, improvements in fine-mapping and prioritisation of variants for functional follow-up contributed to a better understanding of the biological function and causality of observed genetic associations (264). For example, the improved density of SNP arrays as well as imputation panels helped to achieve high mapping resolution of common SNP associations in GWA studies (264). Moreover, genetic signals could be further refined in trans-ethnic meta-analyses by exploiting discrepancies between populations and toning down the associations with proxies (264). Additionally, the integration of GWAS findings with publicly available databases of regulatory elements across different tissues and cell types, such as GTEx, expedited the identification of functionally relevant genes (264).

Although thousands of susceptibility loci had been identified in GWAS, it had been suggested that common polymorphisms captured one-third to two-thirds of the heritability of most complex traits (513). For example, additive genetic effects contributing to the distribution of the population's refractive error, explaining 18% of its variance (99). However, in paediatric cohorts, SNP-based heritability of cycloplegic refractive error reached 35% which implied that genetic associations identified by traditional GWAS accounted for less than half of heritability reported by twin studies (162). There were several reasons for missing heritability. One probable explanation was that GWAS were underpowered to detect associations with variants of modest effect due to conservative Bonferroni correction threshold (514). The statistical power in

genome-wide investigations could be improved with larger sample sizes as well as the implementation of alternative statistical approaches (515–517) and study designs.

Discrepancies between SNP-based and twin heritability arose from genetic influence not captured by GWAS – rare variants, epigenetic effects, genetic nurturing, gene-environment interactions and epistasis (518). Evidence of nonadditive genetic effects, also known as epistasis, was challenging to assess using GWAS due to a lack of statistical power and methodological issues (519) Previous estimates indicated that epistasis most likely contributed a small fraction of genetic variance, thus very large sample sizes were needed to detect significant gene-gene interactions (519). Moreover, some epistatic effects were diluted by the loss of information due to imperfect linkage between genotyped and causal markers (520).

Meta-GWAS of refractive error examined gene-by-environment interactions with education, but despite a well-powered sample, only one locus showed significant interaction with educational attainment (99). This paucity of significant gene-environment interactions with education indicated that this relationship could be compounded by several other factors and possibly non-linear nature, as proposed by previous investigations (327). Recently novel methods inferring gene-environment interactions from genetic variants influencing phenotype variance were developed (521) and could potentially improve future studies investigating the interplay between genes and the environment.

Although empirical evidence indicated that much of complex trait heritability could be attributed to common genetic variation (513), rare and ultra-rare polymorphisms were also expected to contribute (264). However, SNP array-based GWAS are currently not sufficiently powered to detect associations with rare markers which were analysed in whole-exome and whole-genome sequencing studies.

I examined the contribution of the rare variants to the refractive error in the UK Biobank whole-exome sequencing study. The study analyses used a simple dominant model of inheritance, however previous publications demonstrated the value of recessive and compound heterozygosity models that included gene-gene and gene-variant interactions (522) and could provide a novel interesting direction for future myopia research. Moreover, the gene-based analyses were limited to the coding regions of the genome, even though non-coding genetic variation was shown to play a role in several complex disorders (375), and thus could be potentially relevant for myopia. The replication was sought in two independent cohorts using chip array data which proved to be less powerful than whole-exome sequencing methods. One of the general limitations of customised chip arrays was that they relied on imputation to infer genotypes of untyped rare variants by using reference panels. Imputation accuracy was typically lower for markers with MAF <1% (523). The ability to correctly impute rare markers was largely limited because rare variants were observed few times in the reference panel. This made it harder to determine the haplotype background and decreased the set of reference haplotypes available for matching. Imputation quality also depended on a number of factors, including the size of the reference panel, haplotype accuracy and the density of genotyping array (523). Given the low density of genotyping array in the BDES study, a substantial discordance between imputed and observed genotypes could lead to an inflation of the association results (378). Moreover, CREAM studies were genotyped separately and then recalled jointly to increase the calling accuracy. However, this approach also decreased power. Alternatively, the meta-analysis of gene-level rare variant association studies allowed to retain genetic information and had the advantage of boosting power (524) . Nevertheless, these issues were common limitations of array-based studies, and full validation of the results could be achieved using large exome sequencing cohorts. Moreover, the limited power of rare variants could be increased with the integration of additional sequencing data from national biobanks and cohorts.

8.3 Future directions

Following the work presented in this thesis, there are several potential directions for future myopia research. As discussed in this chapter, refractive error was genetically heterogeneous and influenced by genetic polymorphisms involved in the development of eye structures, but also participating in the regulation of circadian rhythm as well as pigmentation. The genetic loci reported by refractive error GWAS meta-analysis explained 18% of spherical equivalent heritability and were highly predictive of common myopia. However, further work is required to validate these findings in the populations of other ancestral and demographic groups. While the utility of PRS to predict myopia risk in large cohorts of European ancestry was previously examined in several publications, few studies assessed the generalisability of PRS in East-Asian populations where the prevalence and the burden of myopia are among the highest. I evaluated the utility of the PRS based on European SNPs to predict high and moderate myopia in the sample of Singaporean children, where selected variants in combination with parental myopia and time outdoors demonstrated a good accuracy of prediction. Nevertheless, the PRS alone had AUC between 0.57 and 0.64, indicating that the accuracy of genetic prediction could be further improved. This can be achieved with large-scale GWAS of refractive error and myopia in Asians which would help build better PRS and ascertain their performance before translation of these findings to clinical practice.

Polygenic risk scores for myopia could be further improved with the discovery of new susceptibility loci. The power estimates suggest that approximately 13,808 genetic markers contribute to spherical equivalent heritability of which 10.3% is explained by 543 SNPs with relatively large effects and the remaining 20.8% by common variants with lesser effects (99). Therefore, it is expected that the proportion of explained variance will continue to increase with even larger sample sizes but start to plateau around 1 million participants. For this reason, future genetic research should also focus on other

sources of missing heritability, including epistasis, genetic nurturing, rare variants and gene-environment interactions.

UK Biobank WES presented in the thesis demonstrated that variants with large effects over refractive error were extremely rare. Nonetheless, it is still possible that with the release of the third tranche of the UK Biobank exome-sequencing project covering additional 300,000 participants, new rare variants will be identified. However, replication of these results will remain a challenge due to the unavailability of large independent samples with high-quality whole-exome sequencing data.

Additionally, to novel genetic associations with refractive error, this thesis reported genetic polymorphisms influencing the age of refractive correction. Genome-wide AFSW time-to-event analyses conducted in UK Biobank detected 44 distinct loci that exhibited effects over the age of refractive correction hazard. Some of the known myopia risk genes were expressed at an earlier age and were more inhibited after the second decade of life. However, the study did not explore the causes of observed age-dependent effects. Interestingly, to date, only one publication examined the influence of lifestyle factors over the genetic risk of refractive error onset. ALSPAC study found nominal evidence suggesting that interaction with near-work could alter the age-dependent effects of some refractive error genes (391). Nonetheless, the gene-environment interactions with near-work or time spent outdoors were scarce for the vast majority of loci predisposing to childhood refractive errors, most likely due to limited statistical power and differences in environmental risk exposures between study cohorts (391).

Identification of lifestyles and genetic associations augments a better understanding of disease pathophysiology and facilitates precision medicine. The assessment of individual myopia risk presents a significant challenge due to the vast number of genetic polymorphisms and behavioural factors that influence susceptibility to the disease. Notably, prediction models that incorporated parental myopia and measures of

biometry and visual acuity achieved higher AUC (AUC 0.82-0.93) (27,28,525) than models limited to the life-course factors and lacking data on hereditary factors and ocular parameters (AUC 0.68) (173). This suggested that genetic, lifestyle and developmental influences were important for individual myopia risk assessment. Therefore, the improvement of current prediction models could be achieved with the inclusion of additional susceptibility markers and refinement of associations with known environmental risk exposures.

Genetic associations contributed to another aspect of precision medicine, pharmacogenomics. Pharmacogenomics studied the effects of certain genotypes and other "omics" on the variability of human drug response. Previous reports documented over 3,000,000 polymorphisms associated with altered pharmacologic response in individuals or populations (526), including variants affecting the efficacy of treatments for pathological myopia (527). Whole-genome analyses of genetic networks were used in pharmacogenomics to expedite the discovery of new pharmacological compounds and the repurposing of existing drugs (528). Given that pharmacological options for myopia are currently limited to atropine and methylxanthine which either cause significant adverse effects or show limited efficacy, there is a need for novel anti-myopia treatments. The work of this thesis demonstrates that investigations of populations medication-use patterns could lead to the discovery of novel disease pathogenesis mechanisms as well as help identify pharmacological agents that could be repurposed for myopia (528). However, future interrogation of novel drug targets and pharmacological candidates for myopia control will largely depend on the discovery and characterization of the cell-specific signalling mechanisms driving emmentropisation and refractive error development (528).

8.4 Conclusions

In this thesis, I explored the genetic influences of refractive errors and myopia. The work presented in this thesis suggested that refractive error was a highly polygenic and genetically heterogeneous trait influenced by common variants located within the genes that regulated anatomical eye development, circadian rhythm and pigmentation. Further evidence was found for the gene involved in retinal signalling, and more broadly in higher CNS functions. Moreover, for the first time, a causal link between elevated IOP and myopic refraction was reported. Refractive error exhibited a moderately strong genetic correlation with age of first spectacle wear and overall genetic polymorphisms associated with myopia predisposed to an earlier age of refractive error. Nonetheless, some known myopia genes, such as *LAMA2*, *BMP3* and *TSPAN10* exerted their effects particularly early in life, while other genes were expressed during the entire lifespan. Time-to-event analyses of the age of first spectacle wear identified six novel loci, demonstrating the utility of self-reported age of lens correction as a proxy for refractive error in GWA studies. The thesis also examined the impact of rare genetic variance on the population's refractive error and found associations with eight genes five of which were not described in previous genetic work. The top signals were detected at *SIX6* and *CRX* loci, while suggestive evidence was observed for six other candidates. *SIX6* encoded transcription factor involved in the optic disc as well as retinal development and morphology. The newly identified CRX gene was a master regulator of photoreceptor cell proliferation and specification and exhibited light-mediated effects on ocular growth. Other novel candidates were implicated in neurogenesis, neural signalling and were linked to learning disorders.

Additionally, biological mechanisms underlying the pathogenesis of refractive errors were explored by examining the relationship between medication-taking and spherical equivalent in a large population-based cohort. Eighteen different pharmacological treatments showed significant associations with refractive error in the UK Biobank

cohort. Significant evidence was found for anti-glaucoma preparations, medications aimed at metabolic diseases and pain management treatments. Participants who reported taking antiglaucoma medications had spherical equivalent that was on average 1-2 Dioptres lower compared to non-users. However, the observed correlation with antiglaucoma agents was influenced by a causal association between elevated IOP and myopic refraction. Similarly, in the group of participants who required refractive correction at a later age, self-reported use of oral antihyperglycemic agents correlated with lower refractive error. There was evidence that this relationship was causal and not confounded by underlying diseases. Another interesting finding was an unexpected association between different classes of analgesics and hypermetropic refractive error. Notably, subsequent analyses revealed that the observed association with pain control medications was confounded by causal association with chronic pain which contributed to hypermetropic refraction at adulthood. The UK Biobank study was the first attempt to systematically assess the relationship between populations medication use patterns and refractive error. This investigation identified several classes of medications associated with refractive changes and found evidence supporting the causal role of IOP and chronic pain in refractive development.

The work presented advances current knowledge about the pathophysiology of refractive errors, reporting several major mechanisms driving the ocular growth and emmentropisation. Nonetheless, known common refractive error susceptibility variants cumulatively explained approximately 20% of phenotypic variance and predicted common but not severe myopia, leaving room for improvement. It is expected that additional loci will be discovered with increasing sample sizes. Whole exome and genome sequencing studies will further bridge the gap between estimated heritability and reported genetic associations. Additional biological pathways could be identified by investigating the relationship between refractive error and other morbidities.

Chapter 9 | Appendix

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
Nat Genet. Author manuscript; available in PMC 2020 September 30.

Published in final edited form as:
Nat Genet. 2020 April ; 52(4): 401–407. doi:10.1038/s41588-020-0599-0.

Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia

Pirro G. Hysi[1,2,3,†,*], **Hélène Choquet**[4,†], **Anthony P. Khawaja**[5,6,†], **Robert Wojciechowski**[7,8,†], **Milly S Tedja**[9,10,†], **Jie Yin**[4], **Mark J. Simcoe**[2], **Karina Patasova**[1], **Omar A. Mahroo**[1,5], **Khanh K Thai**[4], **Phillippa M Cumberland**[3,12], **Ronald B. Melles**[13], **Virginie J.M. Verhoeven**[9,10,11], **Veronique Vitart**[14], **Ayellet Segre**[15], **Richard A. Stone**[16], **Nick Wareham**[6], **Alex W Hewitt**[17], **David A Mackey**[17,18], **Caroline CW Klaver**[9,10,19,20], **Stuart MacGregor**[21], **The Consortium for Refractive Error and Myopia**[22], **Peng T. Khaw**[5], **Paul J. Foster**[5,23], **The UK Eye and Vision Consortium**[22], **Jeremy A. Guggenheim**[24], **23andMe Inc.**[22], **Jugnoo S Rahi**[3,5,12,25,*], **Eric Jorgenson**[4,*], **Christopher J Hammond**[1,2,*]

[1]King's College London, Section of Ophthalmology, School of Life Course Sciences, London, UK [2]King's College London, Department of Twin Research and Genetic Epidemiology, London, UK [3]University College London, GOSH Institute of Child Health, London, UK [4]Division of Research, Kaiser Permanente Northern California, Oakland, California, USA [5]NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK [6]Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge School of Clinical Medicine, Cambridge, UK [7]Department of Biophysics, Johns Hopkins University, Baltimore, MD, USA [8]Wilmer Eye Institute, Johns Hopkins School of Medicine, Baltimore, MD, USA [9]Department of Ophthalmology, Erasmus Medical Center, Rotterdam, The Netherlands [10]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands [11]Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*pirro.hysi@kcl.ac.uk.
†These authors jointly led this work
*These authors jointly supervised this work

Author Contributions

P.G.H., J.S.R., E.J. and C.J.H. conceived and designed the study. P.T.K., P.J.F., J.S.R. contribute to the collection of data. P.G.H., H.C., A.P.K., R.W., M.S.T., J.Y., K.K.T., P.M.C., V.V., J.A.G and E.J. performed statistical analysis. A.P.K., M.J.S., K.P., K.K.T., AS and J.A.G. performed post-GWAS follow-up analyses. P.G.H., H.C., A.P.K., R.W., J.S.R., E.J., C.J.H. wrote the manuscript with help from O.A.M., P.M.C., R.B.M., V.J.M.V., A.S., R.A.S., N.W., A.W.H., D.A.M., C.C.W.K., S.M., P.T.K., P.J.F. and J.A.G. who helped with the interpretation of the results.

Competing interests

23andMe is a consumer genomics company.

Data availability

Summary statistics from the cohorts participating in the meta-analysis can be downloaded from ftp://twinr-ftp.kcl.ac.uk/Refractive_Error_MetaAnalysis_2020/ and public repositories such as the GWAS Catalogue (https://www.ebi.ac.uk/gwas/downloads/summary-statistics). These freely downloadable summary statistics are calculated using all cohorts described in this manuscript, except for the 23andMe participants. This is due to a non-negotiable clause in the 23andMe data transfer agreement, intended to protect the privacy of the 23andMe research participants.

To fully recreate our meta-analytic results, all bona fide researchers can obtain the 23andMe summary statistics by emailing 23andMe (dataset-request@23andme.com) and subsequently meta-analyzing them along the freely accessible summary statistics for all the other cohorts.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Netherlands [12]Ulverscroft Vision Research Group, UCL Great Ormond Street Institute of Child Health, University College London, UK [13]Kaiser Permanente Northern California, Department of Ophthalmology, Redwood City, CA, USA [14]MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, The University of Edinburgh, United Kingdom [15]Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear, Boston, MA, USA [16]Department of Ophthalmology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA [17]Department of Ophthalmology, Royal Hobart Hospital, Hobart, Tasmania [18]Centre for Ophthalmology and Visual Science, University of Western Australia, Lions Eye Institute, Perth, WA, WA 6009, Australia [19]Department of Ophthalmology, Radboud University Medical Center, Rotterdam [20]Institute of Molecular and Clinical Ophthalmology Basel, Switzerland [21]QIMR Berghofer Medical Research Institute, Brisbane, Australia [22]Names and affiliations of the consortium members are listed in the Supplementary Note [23]Division of Genetics and Epidemiology, UCL Institute of Ophthalmology, London, UK. [24]Cardiff University, School of Optometry & Vision Sciences, UK [25]Department of Ophthalmology and NIHR Biomedical Research Centre Great Ormond Street Hospital NHS Foundation Trust

Abstract

Refractive errors, in particular myopia, are a leading cause of morbidity and disability world-wide. Genetic investigation can improve understanding of the molecular mechanisms underlying abnormal eye development and impaired vision. We conducted a meta-analysis of genome-wide association studies involving 542,934 European participants and identified 336 novel genetic loci associated with refractive error. Collectively, all associated genetic variants explain 18.4% of heritability and improve the accuracy of myopia prediction (AUC=0.75). Our results suggest that refractive error is genetically heterogeneous, driven by genes participating in the development of every anatomical component of the eye. In addition, our analyses suggest that genetic factors controlling circadian rhythm and pigmentation are also involved in the development of myopia and refractive error. These results may make possible predicting refractive error and the development of personalized myopia prevention strategies in the future.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Introduction

Refractive errors occur when converging light rays from an image do not clearly focus on the retina. They are the seventh most prevalent clinical condition[1] and the second leading cause of disability in the world[2]. The prevalence of refractive error is rapidly increasing, mostly driven by a dramatic rise in the prevalence of one of its forms, myopia (near-sightedness). Although the causes of such a rise over a short time are likely due to environmental and cultural changes from the mid-20th century[3], refractive errors are highly heritable[4]. Several studies[5,6] have previously sought to identify genes controlling molecular mechanisms leading to refractive error and myopia. However, the variance and heritability

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that can be attributed to known genetic factors is modest[7] and our knowledge of pathogenic mechanisms remains partial. Here, we conduct a meta-analysis combining data from quantitative spherical equivalent and myopia status from large and previously unpublished genome-wide association studies (GWAS) of more than half a million subjects from the UK Biobank, 23andMe and the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohorts, with subsequent replication and meta-analysis with data previously reported from the Consortium for Refractive Error and Myopia (CREAM).

Results

Association Results.

Analyses were restricted to subjects of European ancestry (Extended Data Figure 1) and combined results from quantitative measures of spherical equivalent and categorical myopia status. Spherical equivalent quantifies refractive error; a negative spherical equivalent, below a certain threshold defines myopia. We used results obtained from GWAS of directly measured spherical equivalent in 102,117 population-based UK Biobank participants[8], and 34,998 subjects participating in the GERA Study[9] and combined them with results of analyses of self-reported myopia in 106,086 cases and 85,757 controls from the customer base of 23andMe, Inc. (Mountain View, CA), a personal genomics company[10]. Additionally, we included results from an analysis on the refractive status inferred using demographic and self-reported information on age at first use of prescription glasses among the UK Biobank participants not contributing to the quantitative GWAS (108,956 likely myopes to 70,941 likely non-myopes, see Online Methods). All analyses were adjusted for age, sex and main principal components. To obtain an overall association with refractive error, we meta-analyzed the results from all studies by using the z-scores from the GWAS of the spherical equivalent and the negative values of z-scores from the case-control studies (23andMe and UK Biobank), since myopia is negatively correlated with spherical equivalent. As expected, the large total sample size of the discovery meta-analysis (N=508,855) led to a nominally large genomic inflation factor (λ=1.94). The LD score regression intercept was (1.17), and the (intercept-1)/(mean(chî2)-1) ratio of 0.097 is fully in line with the expectations of polygenicity[11].

We found associations for 438 discrete genomic regions (Figure 1, Supplementary Table 1), defined by markers contiguously associated at conventional level of GWAS significance[12,13] of $p<5\times10^{-08}$, separated by more than 1 Mbp from other GWAS-associated markers, as recommended elsewhere[14]. Among them, 308 loci, including 14 on chromosome X, were not described in previous GWAS studies of refractive error[7]. The observed effect sizes were consistent across all the studies (Supplementary Table 1 and Supplementary Data 1). The association with refractive error was statistically strongest for rs12193446 ($p=9.87\times10^{-328}$), within *LAMA2*, a gene previously associated with refractive error[5,6], mutations of which cause muscular dystrophy[15]. Consistent with these *LAMA2* properties, polymorphisms located within the genes coding for both major *LAMA2* receptors, *DAG1*[16] ($p=1.67\times10^{-08}$ for rs111327216) and *ITGA7*[17] ($p=8.57\times10^{-09}$ for rs17117860) which are also known causes of muscular dystrophy[18,19], were significantly associated with refractive error in the discovery meta-analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We compared our discovery meta-analysis findings with GWAS results from 34,079 participants in the CREAM consortium, who were part of a previously reported meta-analysis[7]. To avoid any potential overlap with the UK Biobank participants, only non-UK European CREAM participants were used for replication. Despite the vast power differential, 55 of the SNPs that showed the strongest association in their respective regions in the discovery meta-analysis were significant after Bonferroni correction in the replication sample. A further 142 had a false discovery rate (FDR) < 0.05 and 192 were nominally significant at P < 0.05 (Supplementary Table 2). The effect sizes observed in the discovery and replication samples were strongly correlated (Pearson's r=0.91, Extended Data Figure 2). Meta-analysis of all five cohorts (discovery and replication) expanded the number to 449 associated of regions of variable length and number of SNPs (Extended Data Figure 3), of which 336 regions were novel (Supplementary Table 3).

Most of the 449 refractive error-associated regions contained at least one gene linked to severe ocular manifestations in the Online Mendelian Inheritance In Man (OMIM) resource or other genes with interesting link to eye disease (Supplementary Table 4). Although most loci identified through our meta-analyses were novel, several of them hosted genes that harbor mutations leading to myopia or other refractive error phenotypes (Supplementary Data 2). Several genes significantly associated with refractive error were linked to Mendelian disorders affecting corneal structure, some of which code for transcription factors involved in corneal development[20] (Supplementary Table 5). Mutations in these genes cause corneal dystrophies (*SLC4A11*, $p=5.81\times10^{-11}$ for rs41281858, *TCF4*, $p=4.14\times10^{-08}$, rs41396445; *LCAT*, $p=1.26\times10^{-10}$, rs5923; and *DCN*, $p=3.67\times10^{-09}$, rs1280632), megalocornea (*LTBP2*, $p=1.91\times10^{-24}$, rs73296215) and keratoconus (*FNDC3B*, $p=1.89\times10^{-14}$, rs199771582, previously described[7]). Eleven refractive error-associated genes were linked to anomalies of the crystalline lens (Supplementary Table 6), including genes linked to autosomal dominant cataracts (*PAX6* previously linked to myopia[21], $p=8.31\times10^{-11}$, rs1540320; *PITX3*, $p=1.05\times10^{-10}$, rs7923183; *MAF*, $p=5.50\times10^{-09}$, rs16951312; *CHMP4B*, $p=9.95\times10^{-11}$, rs6087538; *TDRD7*, $p=4.79\times10^{-08}$, rs13301794) and lens ectopia (*FBN1*, $p=3.30\times10^{-24}$, rs2017765; *ADAMTSL4*, $p=8.19\times10^{-14}$, rs12131376). Some of the genes affected several eye components. For example, *LTBP2 variants* are also associated with congenital glaucoma[22], and *COL4A3* (rs7569375, $p=1.14\times10^{-08}$) causes Alport syndrome, which manifests with abnormal lens shape (lenticonus) and structural changes in the retina.

Association was also observed within or near 13 genes known to harbor mutations causing microphthalmia (Supplementary Table 7), including *TENM3* ($p=2.48\times10^{-11}$, rs35446926); *OTX2* ($p=6.15\times10^{-11}$, rs928109); *VSX2*, ($p=4.60\times10^{-10}$, rs35797567); *MFRP*, ($p=2.85\times10^{-16}$, rs10892353) and the previously identified[6] *TMEM98*, ($p=3.49\times10^{-43}$, rs62067167). Association was also found for *VSX1* ($p=4.59\times10^{-08}$ for rs6050351), a gene that is closely regulated by VSX2[23] and believed to play important roles in eye development[24]. Many of the genes nearest associated SNPs have been linked to inherited retinal disease (Supplementary Table 8), including 32 genes linked to cone-rod dystrophies, night blindness and retinitis pigmentosa, and age-related macular degeneration (*HTRA1/ ARMS2*). Among genes in novel regions associated with refractive error, *ABCA4* ($p=3.20\times10^{-10}$ for rs11165052), and ARMS2/HTRA1 ($p=5.72\times10^{-23}$ for rs2142308) are

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

linked to macular disorders and numerous others to retinitis pigmentosa, retinal dystrophy and other retinal diseases, such as *FBN2*, ($p=8.63\times10^{-11}$, rs6860901) , *TRAF3IP1* ($p=5.71\times10^{-16}$, rs7596847), CWC27 ($p=1.84\times10^{-18}$, rs1309551). Significant association was found near other genes of interest such as *DRD1* ($p=4.51\times10^{-16}$, rs13190379), a dopamine receptor. Together, these results are consistent with previous suggestions of light transmission and transduction in refractive error[7,25].

Wnt signaling has previously been implicated in experimental myopia[26]. We found significant association near several *Wnt* protein-coding genes (*WNT7B*, a gene previously associated with axial length[27], $p=1.42\times10^{-26}$ for rs73175083; *WNT10A*, previously associated with central corneal thickness[28], $p=1.65\times10^{-17}$ for rs121908120 and *WNT3B*, $p=8.52\times10^{-16}$ for rs70600), suggesting that organogenesis through *Wnt* signaling is likely to be involved in refractive error. Significant association were found at genes coding for key canonical (e.g. rs13072632 within the *CTNNB1* gene, $p=7.30\times10^{-27}$; *AXIN2*, rs9895291, $p=1.40\times10^{-08}$) and non-canonical *Wnt* pathway members (*NFATC3*, rs147561310, $p=1.493\times10^{-12}$) or at genes coding for both (*RHOA*, rs7623687, $p=1.81\times10^{-11}$ or the previously described[7] *TCF7L2*, rs56299331, $p=9.38\times10^{-46}$; Supplementary Table 9).

Similar to previous published analyses[25], we found associations for genes involved in sodium, potassium, calcium magnesium and other cation transporters (Supplementary Table 10). The involvement of genes related to glutamatergic synaptic transmission was also notable (Supplementary Table 11). Glutamate is a first synapse transmitter released by photoreceptors towards bipolar cells and is the main excitatory neurotransmitter of the retina, and expression of genes participating in glutamate signaling pathways is significantly altered in myopia models[29]. These associations support the involvement in refractive error pathogenesis of neurotransmission and neuronal depolarization and hyperpolarization that was also suggested before[7]. Associations with *POU6F2* gene intronic variants (rs2696187, $p=1.11\times10^{-11}$) also suggests involvement of factors related to development of amacrine and ganglion cells[30]. Other genes at refractive error-associated loci were annotated to infantile epilepsy, microcephaly, severe learning difficulty, or other inborn diseases affecting the central nervous system (CNS) in OMIM (Supplementary Table 12).

Polymorphisms in genes linked to oculocutaneous albinism (OCA) were significantly associated with refractive error (Supplementary Table 13), although typically association was found for SNPs not strongly associated with other pigmentation traits[31]. Strong association with refractive error was found near the *OCA2* gene causing OCA type 2 ($p=1.37\times10^{-15}$, rs79406658), OCA3 (*TYRP1*, $p=1.18\times10^{-11}$, rs62538956), OCA5 (*SLC39A8*, $p=4.03\times10^{-17}$, rs13107325), OCA6 (*C10orf11*, $p=1.73\times10^{-16}$, rs12256171). In addition, significant association was found near genes linked to ocular albinism (OA) on chromosome X (*TBL1X* and *GPR143*[32], $p=2.20\times10^{-18}$, rs34437079) and Hermansky-Pudlak Syndrome albinism (*BLOC1S1*, $p=2.4610^{-22}$, for rs80340147; note that this gene forms a conjoint read-through transcript the *BLOC1S1-RDH5* with *RDH5*). Other associated markers were located within genes involved in systemic pigmentation also previously associated with refractive error[7], such as *RALY* ($p=3.14\times10^{-18}$, rs2284388), *TSPAN10* ($p=2.22\times^{-50}$, rs9747347), as well as melanoma (*MCHR2*, $p=2.37\times10^{-15}$ for rs4839756).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Functional properties of the associated markers

Among the significantly associated markers, 367 unique markers were frameshift or missense variants (Supplementary Table 14). Several are non-synonymous, such as the Arg141Leu mutation (rs1048661) within *LOXL1*, a gene that causes pseudoexfoliation syndrome and glaucoma[33] and Ala69Ser (rs10490924) in *ARMS2*, associated with increased susceptibility to age-related macular degeneration[34]. Other associated variants with predicted deleterious consequences were located in several genes, such as RGR (p=6.89×10–68, rs1042454), a gene previously associated with refractive error[7,10] and also retinitis pigmentosa[35], and within the *FBN1* gene, near clusters of mutations that cause Marfan Syndrome and anterior segment dysgenesis[36].

Because the functional link between other associated variants and development of refractive error phenotypes is less obvious, we next performed gene-set enrichment analyses to identify properties that are significantly shared by genes identified by the meta-analysis. An enrichment analysis of Gene Ontology processes (Supplementary Table 15) found enrichment for genes participating in RNA Polymerase II transcription regulation ($p=1\times10^{-06}$) and nucleic acid binding transcription factor activity ($p=1.10\times10^{-06}$), suggesting that many of the genetic associations we identified interfere with gene expression. "Eye development" ($p=6.10\times10^{-06}$) and "Circadian regulation of gene expression" ($p=1.10\times10^{-04}$) were also significantly enriched.

A transcription factor binding site (TFBS) enrichment analysis identified significant (FDR < 0.05) over-representation of sites targeted by *GATA4*, *EP300*, *RREB1*, for which association was observed in the meta-analyses (Supplementary Table 16). Binding sites of transcription factors involved in eye morphogenesis and development such as *MAF* (whose mutations cause autosomal cataract), *FOXC1* and *PITX2* (anterior segment dysgenesis) or *CRX* (cone-rod dystrophy) were also enriched. *CRX* and *PAX4*, binding sites were also significantly enriched; these transcription factors are two of the regulators of circadian rhythm and melatonin synthesis[37] alongside *OTX2*, for which SNP significant association was observed in our refractive error meta-analysis ($p=6.15\times10^{-11}$ for rs928109). All of these enriched gene-sets are observed for the first time in a GWAS analysis, although the presence of some of the mechanisms that relate them to refractive error and myopia were hypothesized before[38].

Many of the variants associated with refractive error in our analyses were located within or near genes that are expressed in numerous body tissues (Extended Data Figure 4), and in particular from the nervous system, consistent with our evidence of extraocular, central nervous system involvement in refractive error. Within the eye, these genes were particularly strongly expressed in eye tissues such as cornea, ciliary body, trabecular meshwork[39] and retina[40] (Extended Data Figure 5, Supplementary Table 17). A stratified LD score regression applied to specifically expressed genes (LDSC-SEG)[41] revealed the results of the GWAS are most strongly correlated with genes expressed in the retina and basal ganglia in the central nervous system but these correlations are not significant after multiple testing correction (Extended Data Figure 6 and Supplementary Table 18). It is possible that the strength of these correlations was constrained by the fact that in most cases, available expression levels

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

were measured in adult samples, while refractive error and myopia are primarily developed in younger ages.

A Summary data-based Mendelian Randomization (SMR) analysis[42] integrating GWAS with eQTL data from peripheral blood[43] and brain tissues[44] found concomitant association with refractive error and eQTL transcriptional regulation effects for 159 and 97 genes respectively (Supplementary Tables 19 and 20). A similar analysis integrating GWAS summary data with methylation data from brain tissues found association with both refractive error and changes in methylation for 134 genes (Supplementary Table 21).

Genetic effects shared between refractive error and other conditions

Examining the GWAS Catalog[45], some of the genetic variants reported here were previously associated with refractive error, and with other traits, in particular intraocular pressure, intelligence and education; the latter two are known myopia risk factors (Supplementary Table 22). We used LD score regression to assess the correlation of genetic effects between refractive error and other phenotypes from GWAS summary statistics (Supplementary Table 23). refractive error genetic risk was significantly correlated with intelligence, both in childhood[46] (r_g=−0.27, p=4.76×10^{-09}) and adulthood (fluid intelligence score r_g=−0.25, p=1.56×10^{-39}), educational attainment (defined as the number of years spent in formal education, r_g=−0.24, p=3.36×10^{-54}), self-reported cataract (r_g=−0.31, p=4.70×10^{-10}) and intraocular pressure (IOP, r_g=−0.14, p=1.04×10^{-12}).

Higher educational attainment appears to cause myopia as demonstrated by Mendelian randomization (MR) studies[47]. A gene by environment interaction GWAS for spherical equivalent and educational attainment (using age at completion of formal full-time education as a proxy) was conducted in 66,242 UK Biobank participants. Despite the relatively well-powered sample, only one locus yielded evidence of statistically significant interaction (rs536015141 within *TRPM1*, p=2.35×10^{-09}, Supplementary Table 24), suggesting that the true relationship between refractive error and education is compounded by several factors and may not be linear in nature, as suggested recently[48]. TRPM1 is localized in rod ON bipolar cell dendrites, and rare mutations cause congenital stationary night blindness[49], often associated with high myopia.

To further explore the nature of the relationship between refractive error and IOP, we built MR models using genetic effects previously reported for IOP[50]. On average, every 1 mmHg increase in IOP predicts a 0.05–0.09 diopters decrease in spherical equivalent (Supplementary Table 25, Extended Data Figure 7). We also built a MR model to assess the relationship between intelligence and spherical equivalent, but statistical evidence in this case points towards genetic pleiotropy rather than causation (Supplementary Table 26). This suggests that both myopia and intelligence are often influenced by the same factors, but without direct causal path linking one to the other. We found no significant genetic correlations between refractive error and the glaucoma endophenotype vertical cup to disc ratio (r_g=−0.01, p=0.45), or hair pigmentation (r_g=−0.03, p=0.35). Therefore, refractive error and pigmentation may have different allelic profiles with limited sharing of genetic risk.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Conditional analysis and risk prediction

We subsequently carried out a conditional analysis[51] on the meta-analysis summary results and found a total of 904 independent SNPs significantly associated with refractive error. 890 of these markers were available in the EPIC-Norfolk Study, an independent cohort that did not participate in the refractive error meta-analysis (Extended Data Figure 8). These markers alone explained 12.1% of the overall spherical equivalent phenotypic variance in a regression model or 18.4% (SE=0.04) of the spherical equivalent heritability. Newly associated markers found in our meta-analysis, but not in the previous large GWAS[7], explain 4.6% (SE=0.01) of the spherical equivalent phenotypic variance in EPIC-Norfolk Study, which is an improvement of one third compared to heritability explained by previously associated markers[7].

Predictive models, based on the above-mentioned 890 SNPs, along with age and sex, were predictive of myopia (versus all non-myopia controls) with areas under the receiving operating characteristic curve (AUC) of 0.67, 0.74 and 0.75 (Figure 2), depending on the severity cutoff for myopia (≤−0.75D, ≤−3.00D and ≤−5.00D respectively). The performance of the predictions appears not to improve for myopia definitions of −3.00D or worse, suggesting that the information extracted from our meta-analysis is more representative of the genetic risk for common myopia seen in the general population, than for more severe forms of myopia, which may have a distinctive genetic architecture.

Further exploration of refractive error genetic architecture

Using information from over half a million population-based participants SNPs identified in these analyses still only explain 18.4% of the spherical equivalent heritability. We next assessed how many common SNPs are likely to explain the entire heritable component of refractive error, and what sample sizes are likely to be needed in the future to identify them, using the likelihood-based approach described elswhere[52]. We estimate that approximately 13,808 (SE=969) polymorphic variants are likely to be behind the full refractive error heritability. Similar to other quantitative phenotypic traits that are previously published[52], our analyses estimate that 10.3% (SE=1.0%) of the phenotypic variance is likely explained by a batch of approximately 543 (SE=81) common genetic variants of relatively large effect size and a further 20.8% (SE=0.9%) of the entire phenotypic variance explained by the remainder. With increased sample sizes, we project that the proportion of variance explained will continue to improve fast but will start plateauing for sample sizes above one million, after which further increases in sample size will likely yield ever diminishing additional phenotypic variance (Extended Data Figure 9).

Discussion

Our results provide evidence for at least two major sets of mechanisms in the pathogenesis of refractive error. The first affect intraocular pressure, eye structure, ocular development and physiology, and the second are CNS-related, including circadian rhythm control. Contributors to refractive error include all anatomical factors that alter refractive power relative to eye size, light transmittance, photoconductance and higher cerebral functions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The findings implicate almost every single anatomical components of the eye, which along with the central nervous system participate in the development of refractive error. The healthy cornea contributes to 70% of the optical refractive power of the eyes[53] and genes involved in corneal structure, topography and function may directly contribute to refractive error through direct changes in the corneal refraction. Our results show that several genes involved in lens development also contribute to refractive error in the general population. It is unclear if their contribution is mediated through alterations in biomechanical properties that affect eyes' ability to accommodate, changes to the lens refractive index, or alterations in light transmission properties that impair the ability to focus images on the retina.

Many retinal genes are implicated in the development of refractive error, reflecting the role of light in mediating eye growth and the importance of the retina's role in light transduction and processing[7]. Associations with refractive error at genes coding for gated ion channels and glutamate receptors point to the photoreceptor-bipolar cell interface as a potentially key factor in refractive error. Rare mutations in several of our associated genes cause night blindness, implicating the rod system in the pathophysiology of refractive error, but many also affect cone pathways. The *TRPM1* gene, important for rod ON bipolar cell polarity[54], is also implicated in the gene-education interaction analysis. Associations observed for the *VSX1* and VSX2, its negative regulator, genes implicate the cone bipolar cells[55].

The association with genes involved in pigmentation, including most of the OCA-causing genes, raises questions about the relationship between melanin, pigmentation and eye growth and development. These associations are unlikely to be influenced by any cryptic population structure in our samples, which our analyses were designed to control. None of the major pigmentation-associated SNPs[31] was directly associated with refractive error and there was no significant correlation of genetic effects between refractive error and pigmentation.

The mechanisms linking pigmentation with refractive error are unclear. Foveal hypoplasia[56] and optic disc[57] dysplasias are common in all forms of albinism[58]. Although melanin synthesis is disrupted in albinism, both melanin and dopamine are synthesized through shared metabolic pathways. Disc and chiasmal lesions in albinism are often attributed to dopamine[59], but we found limited evidence supporting an association with refractive error for genetic variants involved in dopamine signaling. The scarcity of association with refractive error for genes involved in dopamine-only pathways contrasts with the abundance of association for genes involved in pigmentation and melanin synthesis. This may suggest that melanin metabolism is connected to refractive error through other mechanisms that are independent from the metabolic pathways it shares with dopamine production. Melanin reaches the highest concentrations in the retinal pigment epithelium at the outmost layer of the retina, and anteriorly, in the iris and variations in pigmentation may affect the intensity of the light reaching the retina. Light exposure is a major protective factor for development of myopia[60,61] It is possible that pigmentation plays a role in light signal transmission and transduction.

Animal model experiments suggest that in addition to local ocular mechanisms, emmetropization (the process by which the eye develops to minimize refractive error) is

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

strongly influenced by the CNS[62]. The strong correlation of genetic risks between refractive error and intelligence and association found for genes linked to severe learning disability support an involvement of the CNS in emmetropization and refractive error pathogenesis.

Results from gene-set enrichment analysis demonstrate an interesting evolution with increasing sample sizes. While smaller previous studies were sufficiently powered to discover enrichment of low, cell-level properties, such as cation channel activity and participation in the synaptic space structures[25], significantly more powered recent studies have found additional evidence for enrichment and involvement of more integrated physiological functions, such as light signal processing in retinal cells and others[7]. Beyond the identification of a much larger number of genes and explaining significantly higher proportions of heritability, our results, based in a considerably more statistically powered sample, uphold the previous findings and support the involvement of the same molecular and physiological mechanisms that were previously described.

In line with expectations from a higher power of association to discover genes and gene sets individually responsible for even smaller proportions of the refractive error variance[63], we find evidence for even higher regulatory mechanisms, that act more holistically over the eye development or integrate eye growth and homeostasis with other processes of extraocular nature. For example, we found evidence that binding sites of transcription factors involved in the control of circadian rhythm are significantly enriched among genes associated with refractive error. Circadian rhythm is important in emmetropization and its disruption leads to myopia in animal knock-out models[38], potentially through dopamine-mediated mechanisms, or changes in IOP and diurnal variations.

Most of the loci identified through our meta-analysis are not subject to particularly strong and systematic evolutionary pressures (Extended Data Figure 10). The variability in minor allele frequencies observed across loci associated with refractive error may therefore be the result of genetic drift. However, given the variety of the different visual components whose disruptions can result in refractive error, this variability may also be the result of overall balancing forces which encourage high allelic diversity of genes involved in refractive error, providing additional buffering capacity to absorb environmental pressures[48] or genetic disruptions on any of the individual components of the visual system.

Our results cast light on potential mechanisms that contribute to refractive error in the general population and have identified the genetic factors that explain a considerable proportion of the heritability and phenotypic variability of refractive error. This allows us to improve significantly our ability to make predictions of myopia risk and generate novel hypotheses on how multiple aspects of visual processing affect emmetropization, which may pave the way to personalized risk management and treatment of refractive error in the population in the future.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Online Methods

Study Participants

The UK Biobank—The UK Biobank is a multisite cohort study of UK residents aged 40 to 69 years who were registered with the National Health Service (NHS) and living up to 25 miles from a study center. Detailed study protocols are available online (http://www.ukbiobank.ac.uk/resources/ and http://biobank.ctsu.ox.ac.uk/crystal/docs.cgi). It was conducted with the approval of the North-West Research Ethics Committee (ref 06/MRE08/65), in accordance with the principles of the Declaration of Helsinki, and all participants gave written informed consent.

Two separate groups of UK Biobank participants were included in these analyses. The first included participants whose refractive error was directly measured (non-cycloplegic autorefraction using the Tomey RC 5000 Auto Refkeratometer, Tomey Corp., Nagoya, Japan). Direct measurements of refractive errors were available for 22.7% of the UK Biobank sample. To ensure reliable and accurate refractive error data, previously published QC criteria were applied[64]. The spherical equivalent was calculated as spherical refractive error (UK Biobank codes 5084 and 5085) plus half the cylindrical error (UK Biobank 5086 and 5087) for each eye.

The second UK Biobank group included participants without direct measurement of refractive error. These participants refractive error status was inferred using questionnaire and other indirect data. Available demographic and clinical information were used to obtain an estimate about the individual's likely myopia status. A Support Vector Machine (SVM) model, with age, sex, age of first spectacle wear and year of birth as prediction parameters was used to infer participants' myopia status. Initial training took place in 80% randomly selected UK Biobank participants of European descent for whom direct spherical equivalent and refractive error status were available. Then the performance was assessed in the remaining 20% of UK Biobank participants of European descent for whom direct spherical equivalent and refractive error status were available. Finally, the SVM predictions in the remaining individuals with no direct spherical error measurements available using the model developed for the training data.

All UK Biobank genotypes were obtained as described elsewhere[65]. The UK Biobank team then performed imputation from a combined Haplotype Reference Consortium (HRC) and UK10K reference panel. Phasing on the autosomes was carried out using a modified version of the SHAPEIT2[66] program modified to allow for very large sample sizes. Only HRC-imputed variants were used for the purpose our analyses of the UK Biobank participants. The variant-level quality control exclusion metrics applied to imputed data for GWAS included the following: call rate < 95%, Hardy–Weinberg equilibrium $P < 1 \times 10^{-6}$, posterior call probability < 0.9, imputation quality < 0.4, and MAF < 0.005. The Y chromosome and mitochondrial genetic data were excluded from this analysis. In total, 10,263,360 imputed DNA sequence variants were included in our analysis. Non-European ancestry and participants with relatedness corresponding to third-degree relatives or closer, samples with excess of missing genotype calls or heterozygosity were excluded. In total, genotypes were available for 102,117 participants of European ancestry with spherical equivalent data.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Association models in the first UK Biobank subset used the average of spherical equivalent as the outcome and allele dosages at each genetic locus as predictors. Mixed linear regressions, adjusting for age, sex and the first 10 principal components, implemented in the Bolt-LMM software[67] were used.

For the second UK Biobank subset, for which no direct spherical equivalent measurement was available, the mixed linear model was built with the predicted myopia status as outcome and using the same covariates as for the previously described linear regression analysis on spherical equivalent. Odds Ratios were obtained from the beta regression coefficient using the equation:

$$\ln\left(\mathrm{OR}\right) = \frac{\beta}{\mu(1-\mu)}$$

where μ is the fraction of the cases in the sample (μ=0.606). Genotypes with MAF <0.01 and MAC< 400 were removed from analyses in this group.

23andMe—Participating subjects were all volunteers from the 23andMe (Mountain View, CA, USA) personal genomics company customer base. All participants provided informed consent and answered surveys online according to the approved 23andMe human subjects protocol, which was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (http://www.eandireview.com). The participants were identified as myopia cases if they self-reported a diagnosis of myopia or suffering from symptoms of myopia (see Supplementary Notes for more detail).

DNA extraction and genotyping were performed on saliva samples by CLIA-certified and CAP-accredited clinical laboratories of Laboratory Corporation of America. Samples were genotyped on one of four genotyping platforms and batches (Illumina HumanHap550, BeadChip, SNPs, Illumina OmniExpress, plus a variable number of custom SNP assays). Only samples with more than 98.5% genotyping success rate were included. Ethnic categorization was conducted using a support vector machine (SVM) which classified individual haplotypes into one of the 31 reference populations derived from public datasets (the Human Genome Diversity Project, HapMap, and 1000 Genomes), as well as 23andMe customers who have reported having four grandparents from the same country. Genotypes were imputed against the September 2013 release of 1000 Genomes Phase1 reference haplotypes using a Beagle haplotype graph-based phasing algorithm for the autosomal and Minimac2[68] for X Chromosome loci.

Association test results were computed by linear regression assuming additive allelic effects using imputed allele dosages. Covariates for age, gender, the first ten principal components to account for residual population structure were also included into the model.

The Genetic Epidemiology Research in Adult Health and Aging (GERA) cohort —GERA is part of the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH) and has been described in detail elsewhere[69]. It comprises adult men and women consenting members of Kaiser Permanente Northern California (KPNC), an

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

integrated health care delivery system, with ongoing longitudinal records from vision examinations. For this analysis, 34,998 adults (25 years and older), who self-reported as non-Hispanic white, and who had at least one assessment of spherical equivalent obtained between 2008 and 2014 were included. All study procedures were approved by the Institutional Review Board of the Kaiser Foundation Research Institute. Participants underwent vision examinations, and most subjects had multiple measures for both eyes. Spherical equivalent was assessed as the sphere + cylinder/2. The spherical equivalent was selected from the first documented assessment, and the mean of both eyes was used. Individuals with histories of cataract surgery (in either eye), refractive surgery, keratitis, or corneal diseases were excluded from further analyses.

DNA samples from GERA individuals were extracted from Oragene kits (DNA Genotek Inc., Ottawa, ON, Canada) at KPNC and genotyped at the Genomics Core Facility of the University of California, San Francisco (UCSF). DNA samples were genotyped using the Affymetrix Axiom arrays (Affymetrix, Santa Clara, CA, USA). SNPs with initial genotyping call rate ≥97%, allele frequency difference ≤0.15 between males and females for autosomal markers, and genotype concordance rate >0.75 across duplicate samples were included. In addition, SNPs with genotype call rates <90% were removed, as well as SNPs with a minor allele frequency (MAF) < 1%.

Imputation pre-phasing of genotypes was done using Shape-IT v2.r727[19][66], variants were imputed from the cosmopolitan 1000 Genomes Project reference panel (phase I integrated release; http://1000genomes.org) using IMPUTE2 v2.3.0[70]. Variants with an imputation IMPUTE r^2 < 0.3 were excluded, and analyses were restricted to SNPs that had a minor allele count (MAC) ≥20.

For each SNP locus, linear regressions of each individual's spherical equivalent were performed with the following covariates: age at first documented spherical equivalent assessment, sex, and genetic principal components using PLINK v1.9 (www.cog-genomics.org/plink/1.9/). Data from each SNP were modeled using additive dosages to account for the uncertainty of imputation. The top 10 ancestry PCs were included as covariates, as well as the percentage of Ashkenazi ancestry to adjust for genetic ancestry, as described previously[69].

The Consortium for Refractive Error And Myopia (CREAM)—All participants selected for this study were of European descent, 25 years of age or older. refractive error was represented by measurements of refraction and spherical equivalent (Spherical equivalent = spherical refractive error +1/2 cylinder refractive error) was the outcome variable for CREAM. Participants with conditions that could alter refraction, such as cataract surgery, laser refractive procedures, retinal detachment surgery, keratoconus, or ocular or systemic syndromes were excluded from the analyses. Recruitment and ascertainment strategies varied by study and were previously published elsewhere[71].

The genotyping process has been described elsewhere[71]. Samples were genotyped on different platforms, and study-specific QC measures of the genotyped variants were implemented before association analysis. Genotypes were imputed with the appropriate

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ancestry-matched reference panel for all cohorts from the 1000 Genomes Project (Phase I version 3, March 2012 release). Quality control criteria used for SNP and sample inclusions These metrics were similar to those described in a previous GWAS analyses and detailed information for each cohort is described elsewhere[71].

To avert sample overlap, cohorts from the United Kingdom (1985BBC, ALSPAC-Mothers, EPIC-Norfolk, ORCADES and Twins UK) were excluded from the GWAS meta-analysis. Association analyses were performed as described elsewhere[71]For each individual cohort, a single-marker analysis for the phenotype of SphE (in diopters) was carried out with linear regression with adjustment for age, sex and up to the first five principal components. For all non-family-based cohorts, one of each pair of relatives was removed. In family-based cohorts, mixed model-based tests of association were used to adjust for within-family relatedness.

The European Prospective Investigation into Cancer (EPIC) Study—The EPIC-EPIC is one of the UK arms of a broad pan-European prospective cohort study designed to investigate the etiology of major chronic diseases[72]. This study was conducted following the principles of the Declaration of Helsinki and the Research Governance Framework for Health and Social Care. The study was approved by the Norfolk Local Research Ethics Committee (05/Q0101/191) and East Norfolk & Waveney NHS Research Governance Committee (2005EC07L). All participants gave written, informed consent. Refractive error was measured in both eyes using a Humphrey Auto-Refractor 500 (Humphrey Instruments, San Leandro, California, USA). Spherical equivalent was calculated as spherical refractive error plus half the cylindrical error for each eye.

The EPIC-Norfolk participants were genotyped using the Affymetrix UK Biobank Axiom Array (the same array as used in UK Biobank); 7,117 contributed to the current study. SNP exclusion criteria included: call rate < 95%, abnormal cluster pattern on visual inspection, plate batch effect evident by significant variation in minor allele frequency, and/or Hardy-Weinberg equilibrium $P < 10^{-7}$. Sample exclusion criteria included: DishQC < 0.82 (poor fluorescence signal contrast), sex discordance, sample call rate < 97%, heterozygosity outliers (calculated separately for SNPs with minor allele frequency >1% and <1%), rare allele count outlier, and impossible identity-by-descent values. Individuals with relatedness corresponding to third-degree relatives or closer across all genotyped participants were also removed from further analyses. Following these steps all participants were of European descent. Data were pre-phased using SHAPEIT[66] version 2 and imputed to the Phase 3 build of the 1000 Genomes project[74] (October 2014) using IMPUTE[70] version 2.3.2.

The relationship between allele dosage and mean spherical equivalent was analyzed using linear regression adjusted for age, sex and the first 5 principal components. Analyses were carried out using SNPTEST version 2.5.1.

Statistical analyses

We conducted two meta-analyses. For the initial meta-analysis (discovery), we used summary statistic results from the UK Biobank 1st and 2nd subset, the GERA and 23andMe Studies.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

For the final meta-analysis, we used all available information (UK Biobank 1 and 2, the GERA, 23andMe and CREAM Consortium).

For all meta-analyses we applied a Z-score method, weighted by the effective population sample size, as implemented in METAL[75]. No genomic control adjustment was applied during the meta-analysis.

The effective population size was calculated per each locus and as was equal to the total sample size if a linear regression or linear mixed model were used. For case-control studies the effective population was calculated as:

$$\mathrm{N.eff} = 2/(\frac{1}{N.cases} + \frac{1}{N.controls})$$

as recommended before[76], where N.eff is the effective sample size, N.cases is the number of cases considered to have myopia and N.controls is the number of subjects considered not to have myopia. Following this method, we calculated that for the full-sample analysis of 542,934 subjects, due to the presence of two case-control cohort, our effective sample sizes was 379,227.

Only SNPs with minor allele frequency of at least 1%, which were available from at least 70% of the maximum number of participants across all studies and that were not missing in more than one strata (cohorts), were considered further.

Conditional analyses were conducted using the conditional and joint analysis on summary data (COJO) as implemented in the GCTA program[77] to identify independent effects within associated loci as well as the calculation of the phenotypic variance explained[78] by all polymorphisms associated with the trait after the conditional analyses. The threshold of significance was set at 5×10^{-08} and the collinearity threshold was set at r^2=0.9.

Genomic inflation was assessed using the package 'gap' in R (https://cran.r-project.org/) and to distinguish between the effect of polygenicity and those arising from sample stratification or uncontrolled population admixture, the LD score regression intercepts were calculated using the program LD Score (https://github.com/bulik/ldsc).

Bivariate genetic correlations between refractive error and other complex traits whose summary statistics are publicly available were assessed following previously described methodologies[79], using the program LD Score (https://github.com/bulik/ldsc).

To assess the potential value of the loci associated with refractive error to predict myopia, regression-based models were trained and tested separately in two separate groups. The training set comprises the European UK Biobank participants for whom the spherical equivalent measurements were available. The models were tested in the EPIC-Norfolk cohort, which was not part of any of the analyses through which the genetic associations were identified.

The model in included age, sex, and the major genetic variants associated with refractive error after the conditional analysis. Three different definitions of myopia were used based on

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sliding spherical equivalent thresholds: M_1 ≤−0.75D, M_2 ≤−3.00 D and M_3 ≤−5.00D. These three different definitions of myopia were chosen to correspond to the generally accepted definitions of "any myopia", "moderate myopia" and "high myopia". For the latter, we opted for the −5.00D, because definitions based on the more stringent threshold of ≤ −6.00D would have not allowed for a sufficient number of cases in the testing set. For the purpose of these analyses, a "control" was any subject who did not have myopia, defined by a mean spherical equivalent ≥−0.5D.

A Receiver Operating Characteristic (ROC) curve was drawn for each case and the Area Under the Curve (AUC) was calculated. R programming language and software environment for statistical computing (https://cran.r-project.org/) was used for both the logistic regression models ('glm') and to evaluate the performance of the model ('ROCR').

Polymorphisms associated at a GWAS level ($P<5\times10^{-08}$) were clustered within an "associated genomic region", defined as a contiguous genomic region where GWAS-significant markers were within 1 million base pairs from each other. Significant polymorphisms were annotated with the gene inside whose transcript-coding region they are located, or alternatively, if located between two genes, with the gene nearest to it. The associated genomic regions were collectively annotated with the gene overlapping, or nearest the most significantly associated variant within that region.

The known relationships between identified genetic loci and other phenotypic traits were derived from two datasets: the Online Mendelian Inheritance In Man (OMIM, https://omim.org), which is a continuously curated catalog of human genes and phenotypic changes their polymorphic forms cause in humans and the GWAS Catalog[80] which is a curated catalog of previous GWAS association of SNPs or genes with other phenotypic traits.

The R (https://cran.r-project.org) package MendelianRandomization v3.4.4 was used for Mendelian randomization analyses.

Disease-relevant tissues and cell types were identified by analyzing gene expression data together with summary statistics from the meta-analysis of refractive error in all five cohorts, as described elsewhere[81]. Expression data was obtained from the following sources: 1) the GTEx release v7 (https://gtexportal.org/home/datasets) 2) Fetal and adult corneal, trabecular meshwork and ciliary body RNA sequencing data previously described [82] and 3) data from the subset of subjects with presumed healthy adult retinas (AMD=1) from datasets described elsewhere[83].

As the transcription data were heterogeneous and in different units, expression levels for all tissues were rank-transformed. Hierarchical clustering was used to help visualize similarities and differences of patterns of transcript expression across different tissues ('hclust' package in R).

SMR (Summary data–based Mendelian randomization) uses GWAS variants as instrumental variables and gene expression levels or methylation levels as mediating traits, in order to test whether the causal effect of a specific variant on the phenotype-of-interest acts via a specific gene[84]. The SMR tests were performed used three different: the summary statistics of eQTL

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

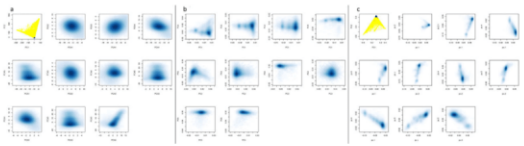
associations in the untransformed peripheral blood samples of 5,311 subjects[85], as well as eQTL effects and *cis*- methylation analysis (*cis*-mQTL), both in brain tissues[86].

The Gene-Set Enrichment Analysis (GSEA) was implemented in the MAGENTA software[87]. We used the versions from September 2017.

Results of three statistical tests for natural selection were imported from the 1000 Genomes Selection Browser[88].

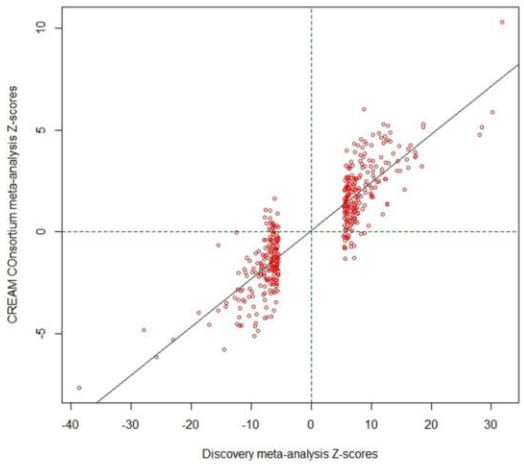
Extended Data

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



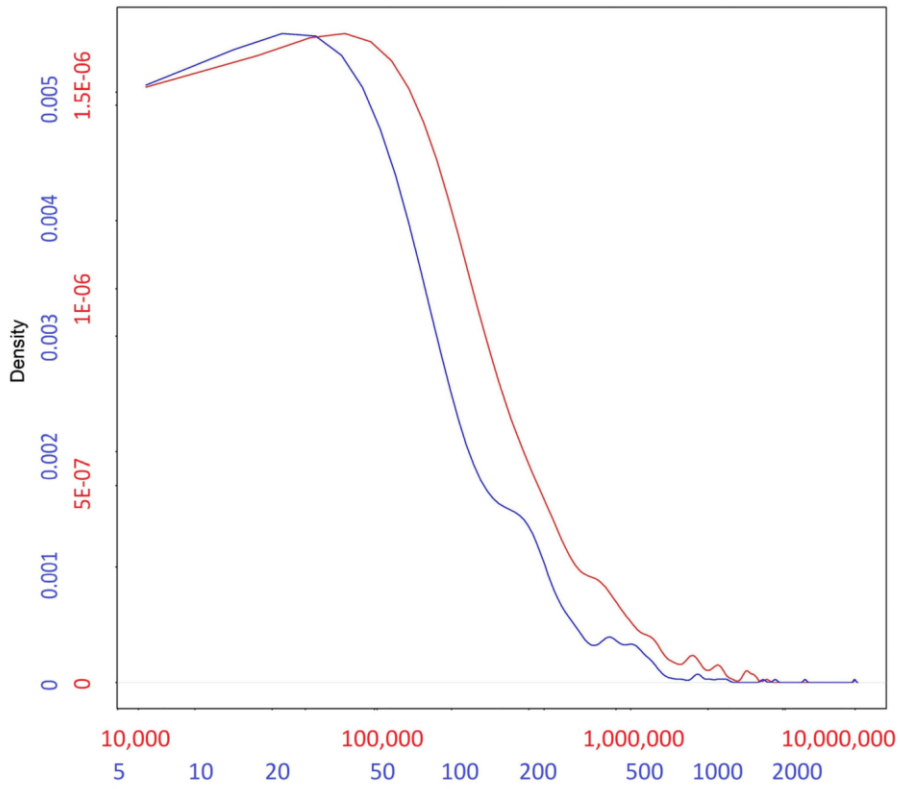
Extended Data Fig. 1:
Principal components plots of the subjects in the main participating cohorts. a) UK Biobank (including the 102,117 subjects with direct refraction measurement and the imputed 108,956 likely myopes to 70,941 likely non-myopes, for a total of 179,897 subjects) , B) Genetic Epidemiology Research on Adult Health and Aging (GERA, N=34,998), C) 23andMe (106,086 cases and 85,757 controls, or 191,843 subjects in total).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



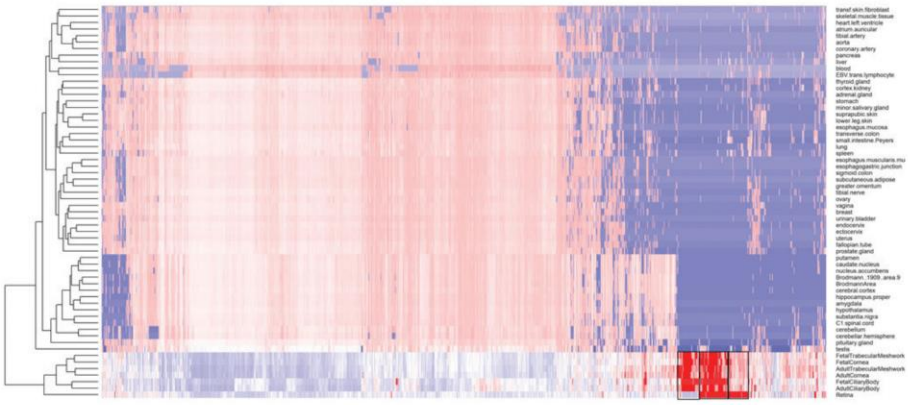
Extended Data Fig. 2:
Correlation of effect sizes between the discovery cohort meta-analysis. Effect sizes are from two analyses, discovery (UK Biobank analysis on spherical equivalent + GERA, spherical equivalent + 23andMe, self-reported myopia cases and controls + UK Biobank inferred myopia cases and controls, for a total of N=508,855 subjects) and the replication from the non-British CREAM Consortium participants (N=34,079), used as replication. The z-scores for the discovery are on the y-axis and those from the CREAM cohort in the x-axis.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



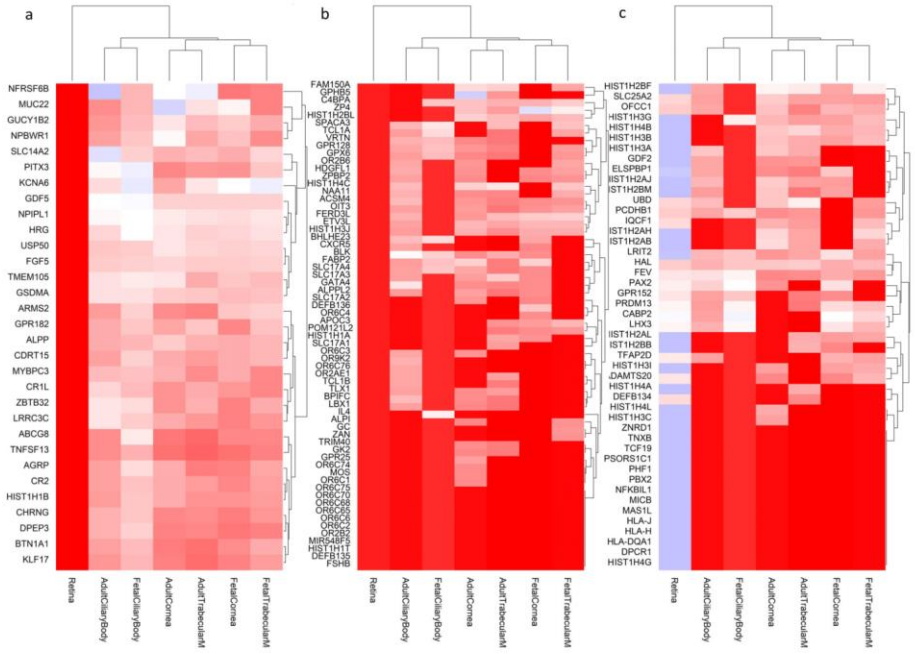
Extended Data Fig. 3:
Distribution of the base-pair length (red) of the 449 regions associated in the meta-analysis of all available cohorts (from Supplementary Table 3), alongside the distribution of number of SNPs (blue) for each region. Numbers in each of the axes in the figure are differentially colored to match the density curve they correspond to: red for the length of the region and blue for the number of SNPs.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Extended Data Fig. 4:
Expression of genes located in the associated loci (from Supplementary Table 3) along the x-axis, across several human body tissues (y-axis). The colors represent the centile ranking of the expression level of the gene in the tissue of interest. The hotter colors represent higher ranking of the gene expression and the colder colors low expression. Both genes and tissues are clustered in accordance with their pattern similarity. The symbol of all the genes could not be visualized and therefore are removed for the sake of clarity. Eye tissues, whether fetal or adult, appear to have similar patterns of gene expressions (clustered together at the bottom of the figure). Genes that are highly expressed in eye tissues fall in three clusters, shown with a black box. These clusters are displayed in more detail in Figure 4A, B and C.

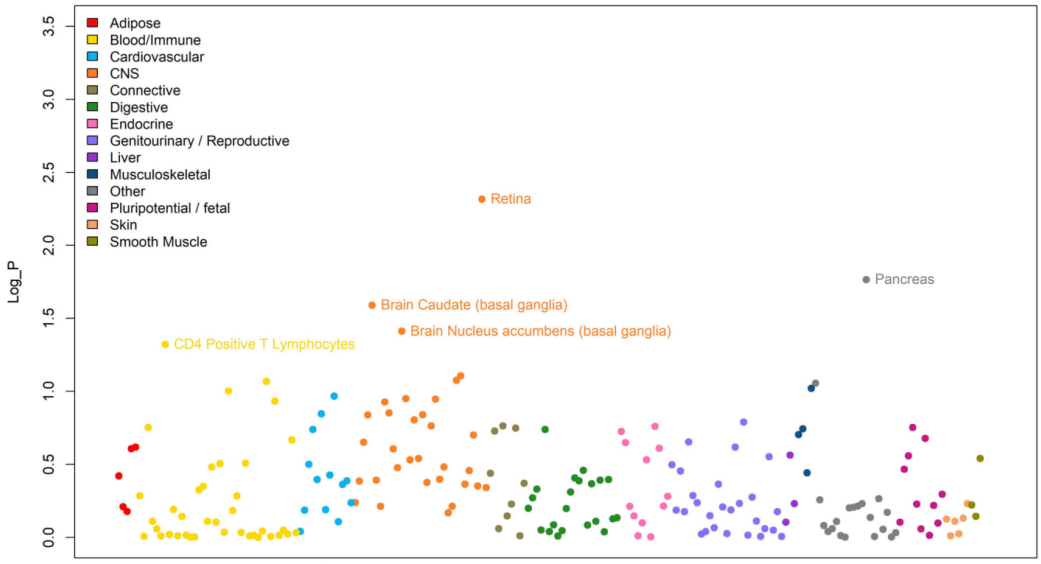
Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Extended Data Fig. 5:

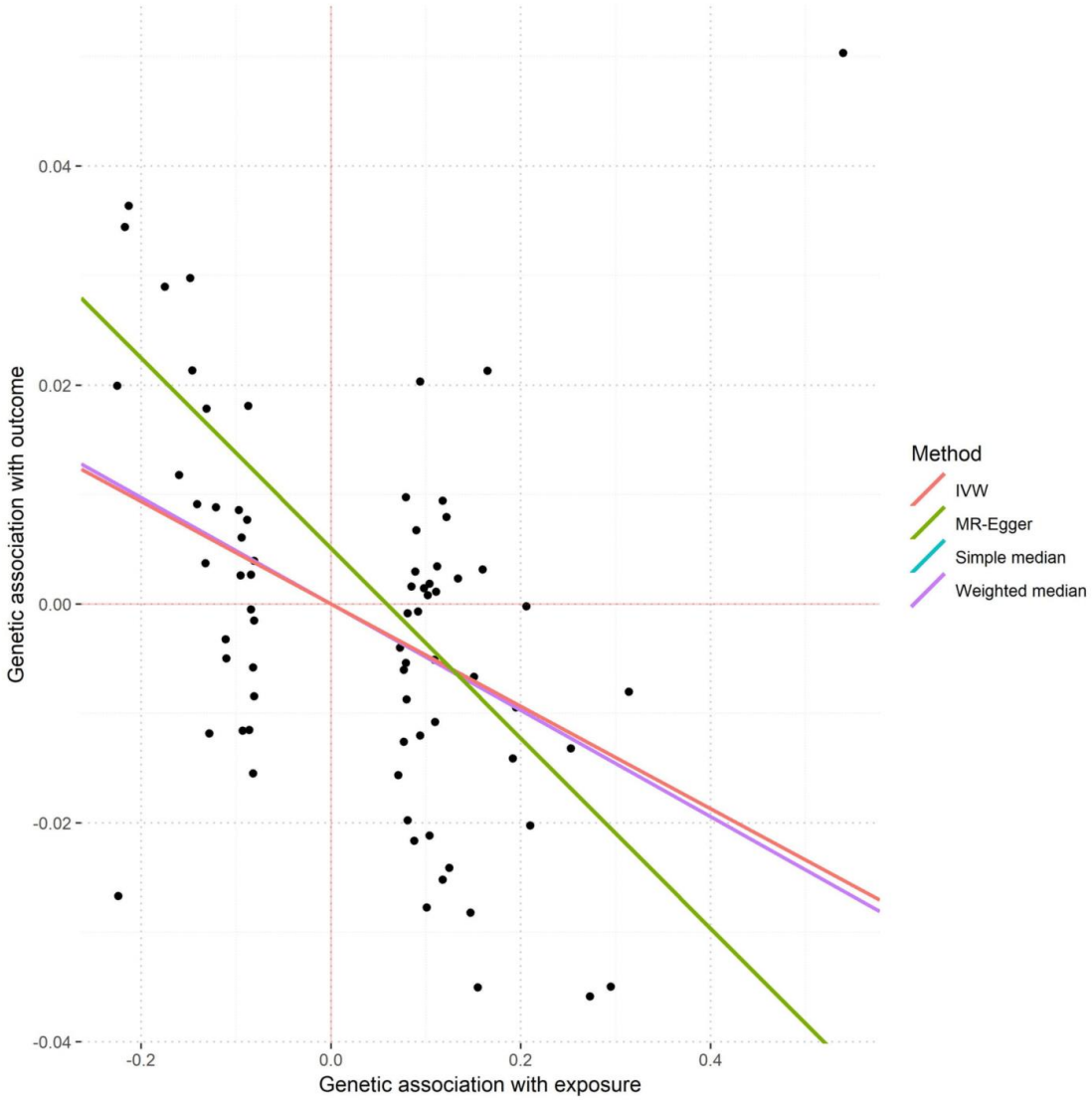
Genes from the regions associated with RE (from Supplementary Table 3) that are particularly expressed in eye tissues, compared to non-ocular tissues. These clusters are those highlighted in Supplementary Figure 3, but for the sake of clarity they are shown in transposed orientation compared to the previous figure (here genes in the y-axis and eye tissues in the x-axis), but same color codes as before. The dendrograms represent the degree of similarity observed for both tissues and gene expressions. The clusters are given in the order in which they were clustered together, from left to right: A) genes that are expressed more in other ocular tissues (fetal and adult) but much less in the adult retina. B) genes that are highly expressed in the retina and other ocular tissues, and C) genes that are expressed in the retina, but less in the other ocular tissues tested.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Extended Data Fig. 6:
Results of the LD score regression analysis applied to specifically expressed genes (LDSC-SEG) on multiple tissue for the meta-analysis results. Each point represents one tissue or cell line (along the x-axis) and the log10 value of the p-value for the enrichment of the meta-analysis results among genes expressed in these tissues. There were 205 tests carried out, one in each tissue and cell line, therefore only tissues with a correlation p-value< 0.00025 (Log_P> 3.6 in this figure), would have been significant after multiple testing. This condition was not fulfilled for any of the available tissues.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



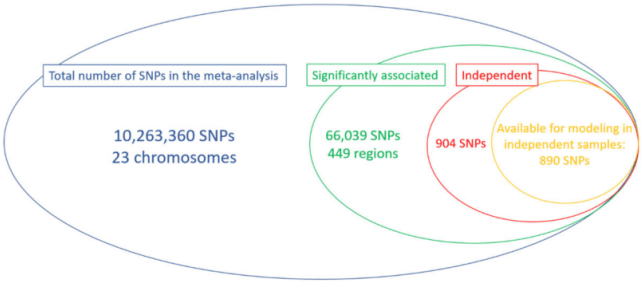
Extended Data Fig. 7:

Mendelian randomization results on causality of IOP over refractive error. Single points in the graph represent coordinates determined by the effect of each specific SNP over IOP (x-axis, mmHg) and spherical equivalent (y-axis, Diopter units). A total of 73 SNPs associated with IOP, but not directly associated with refractive error (i.e. p> 0.05) were selected as instruments. Values of associations with IOP were obtained from a meta-analysis of 139,555 European participants (Reference 50 in the manuscript) and the refractive error associations from 102,117 UK Biobank subjects. The lines represent the regression lines from each

model, as specified in the figure legend. In some cases, these lines may not visible because they overlap (please refer to the values underneath the figure).

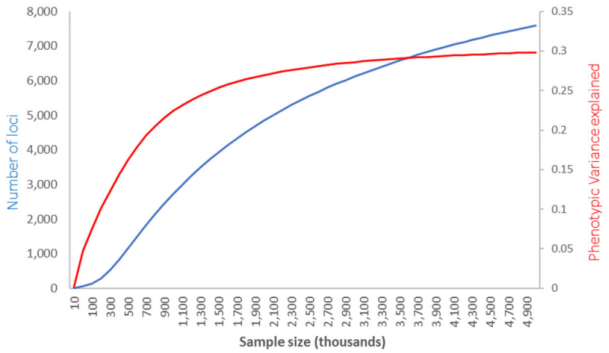
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



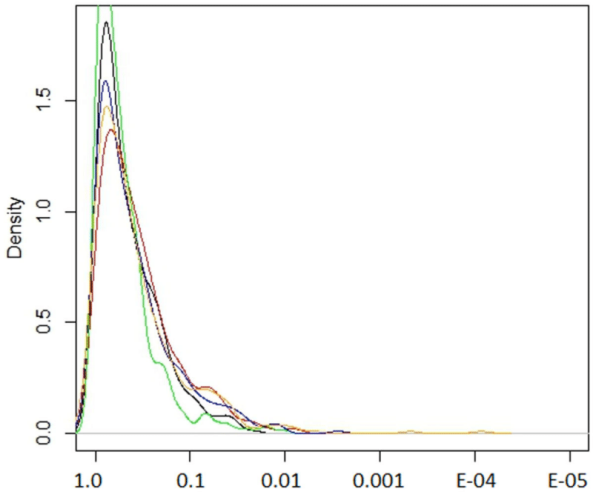
Extended Data Fig. 8:
Venn's Diagram of the number of SNPs considered in each of the stages of this study. The different circles represent various stages, inclusion in the meta-analysis (blue), identification of significant loci (green), conditional analysis results identifying independent effects (red) and the total number of SNPs available for inclusion in prediction and heritability estimation in the independent (i.e. not part of the original meta-analysis) EPIC-Norfolk cohort (orange).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Extended Data Fig. 9:
Prediction for the total number of SNPs and phenotypic variance explained as a function of GWAS sample size in future studies, based on the distribution of effects observed in the current meta-analysis. The plot lines show the predicted relationship between the number of loci associated with refractive error (left vertical axis, blue line) and the variance they help explain (red line, right vertical axis), as a function of the sample size (x-axis) used in future GWAS or meta-analyses. These projections are consistent with the observed results, where an effective sample of 379,227 identified 904 independent signals after a conditional analysis, explaining 12–16% of refractive error variability.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Extended Data Fig. 10:
The distribution of various natural selection test scores for SNPs associated with refractive error. The values on the x-axis represent the ranking in terms of natural selection observed and the y-axis the density of that rank. The different tests shown are iHS, XP-EHH (CEU vs YRI), XP-EHH average score, XP-EHH maximum score and Tajima scores (black, green, red, blue and yellow respectively).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

P.T.K. and P.J.F oversaw the UK Biobank eye data acquisition with support from The National Institute for Health Research (NIHR); Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology. The UK Biobank Eye and Vision Consortium was supported by grants from UK NIHR (BRC3_026), Moorfields Eye Charity (ST 15 11 E), Fight for Sight (1507/1508), The Macular Society, The International Glaucoma Association (IGA, Ashford UK) and Alcon Research Institute. V.V. is supported by a core UK Medical Research Council grant MC_UU_00007/10.

23andMe thanks research participants and employees of 23andMe for making this work possible (list of contributing staff in the Supplementary Note).

Genotyping of the GERA cohort was funded by the US National Institute on Aging; National Institute of Mental Health and National Institute of Health Common Fund (RC2 AG036607); data analyses by the National Eye Institute (NEI R01 EY027004, E.J.), National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK116738, E.J.).

The CREAM GWAS meta-analysis was supported by European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (grant 648268 to C.C.W.K), the Netherlands Organisation for Scientific Research (NWO, 91815655 to C.C.W.K) and the National Eye Institute (R01EY020483). V.J.M.V. acknowledges funding from the Netherlands Organisation for Scientific Research (NWO, grant 91617076).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

S.M. acknowledges support from the National Health and Medical Research Council (NHMRC) of Australia (grants 1150144, 1116360, 1154543, 1121979).

EPIC-Norfolk infrastructure and core functions are supported by the Medical Research Council (G1000143) and Cancer Research UK (C864/A14136). Genotyping funded by the UK MRC (MC_PC_13048). A.K.P. is supported by a Moorfields Eye Charity grant. P.J.F. received support from the Richard Desmond Charitable Trust, the National Institute for Health Research to Moorfields Eye Hospital and the Biomedical Research Centre for Ophthalmology.

RW and PGH were supported by the National Eye Institute of the National institutes of Health under award number R21EY029309. M.J.S. is a recipient of a Fight for Sight PhD studentship. K.P. is a recipient of a Fight for Sight PhD studentship. P.G.H. the recipient of a FfS ECI fellowship. P.G.H. and C.J.H. acknowledge the TFC Frost Charitable Trust Support for the KCL Department of Ophthalmology. Statistical analyses were run in King's College London Rosalind HPC LINUX Clusters and cloud servers. The UK Biobank data was accessed as part of the UK Biobank projects 669 and 17615.

J.S.R. is supported in part by the NIHR Biomedical Research Centres at Moorfields Eye Hospital/UCL Institute of Ophthalmology and at the UCL Institute of Child Health/Great Ormond Street Hospital and is an NIHR Senior Investigator. P.M.C. was funded by the Ulverscroft Foundation. O.A.M is supported by Wellcome Trust grant 206619_Z_17_Z and the NIHR Biomedical Research Centre at Moorfields Eye Hospital and the UCL Institute of Ophthalmology.

References:

1. Vos T et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet 390, 1211–1259 (2017).
2. WHO. The Global Burden of Disease. 2004 Update ISBN-13: 9789241563710. ISBN-10 651629118(2008).
3. Williams KM et al. Increasing Prevalence of Myopia in Europe and the Impact of Education. Ophthalmology 122, 1489–97 (2015). [PubMed: 25983215]
4. Sanfilippo PG, Hewitt AW, Hammond CJ & Mackey DA The heritability of ocular traits. Surv Ophthalmol 55, 561–83 (2010). [PubMed: 20851442]
5. Kiefer AK et al. Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. PLoS Genet 9, e1003299 (2013).
6. Verhoeven VJ et al. Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. Nat Genet 45, 314–8 (2013). [PubMed: 23396134]
7. Tedja MS et al. Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. Nat Genet 50, 834–848 (2018). [PubMed: 29808027]
8. Cumberland PM et al. Frequency and Distribution of Refractive Error in Adult Life: Methodology and Findings of the UK Biobank Study. PLoS One 10, e0139780 (2015).
9. Kvale MN et al. Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics 200, 1051–60 (2015). [PubMed: 26092718]
10. Pickrell JK et al. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet 48, 709–17 (2016). [PubMed: 27182965]
11. Bulik-Sullivan BK et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47, 291–5 (2015). [PubMed: 25642630]
12. Dudbridge F & Gusnanto A Estimation of significance thresholds for genomewide association scans. Genet Epidemiol 32, 227–34 (2008). [PubMed: 18300295]
13. Pe'er I, Yelensky R, Altshuler D & Daly MJ Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 32, 381–5 (2008). [PubMed: 18348202]
14. Wood AR et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46, 1173–86 (2014). [PubMed: 25282103]
15. Oliveira J et al. LAMA2 gene analysis in a cohort of 26 congenital muscular dystrophy patients. Clin Genet 74, 502–12 (2008). [PubMed: 18700894]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

16. Colognato H et al. Identification of dystroglycan as a second laminin receptor in oligodendrocytes, with a role in myelination. Development 134, 1723–36 (2007). [PubMed: 17395644]
17. Burkin DJ & Kaufman SJ The alpha7beta1 integrin in muscle development and disease. Cell Tissue Res 296, 183–90 (1999). [PubMed: 10199978]
18. Ervasti JM & Campbell KP Dystrophin-associated glycoproteins: their possible roles in the pathogenesis of Duchenne muscular dystrophy. Mol Cell Biol Hum Dis Ser 3, 139–66 (1993). [PubMed: 8111538]
19. Mayer U et al. Absence of integrin alpha 7 causes a novel form of muscular dystrophy. Nat Genet 17, 318–23 (1997). [PubMed: 9354797]
20. Jean D, Ewan K & Gruss P Molecular regulators involved in vertebrate eye development. Mech Dev 76, 3–18 (1998). [PubMed: 9767078]
21. Hammond CJ, Andrew T, Mak YT & Spector TD 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 (2004). [PubMed: 15307048]
22. Ali M et al. Null mutations in LTBP2 cause primary congenital glaucoma. Am J Hum Genet 84, 664–71 (2009). [PubMed: 19361779]
23. Clark AM et al. Negative regulation of Vsx1 by its paralog Chx10/Vsx2 is conserved in the vertebrate retina. Brain Res 1192, 99–113 (2008). [PubMed: 17919464]
24. Heon E et al. VSX1: a gene for posterior polymorphous dystrophy and keratoconus. Hum Mol Genet 11, 1029–36 (2002). [PubMed: 11978762]
25. Hysi PG et al. Common mechanisms underlying refractive error identified in functional analysis of gene lists from genome-wide association study results in 2 European British cohorts. JAMA Ophthalmol 132, 50–6 (2014). [PubMed: 24264139]
26. Ma M et al. Wnt signaling in form deprivation myopia of the mice retina. PLoS One 9, e91086 (2014).
27. Miyake M et al. Identification of myopia-associated WNT7B polymorphisms provides insights into the mechanism underlying the development of myopia. Nat Commun 6, 6689 (2015). [PubMed: 25823570]
28. Cuellar-Partida G et al. WNT10A exonic variant increases the risk of keratoconus by decreasing corneal thickness. Hum Mol Genet 24, 5060–8 (2015). [PubMed: 26049155]
29. Stone RA et al. Image defocus and altered retinal gene expression in chick: clues to the pathogenesis of ametropia. Invest Ophthalmol Vis Sci 52, 5765–77 (2011). [PubMed: 21642623]
30. Zhou H, Yoshioka T & Nathans J Retina-derived POU-domain factor-1: a complex POU-domain gene implicated in the development of retinal ganglion and amacrine cells. J Neurosci 16, 2261–74 (1996). [PubMed: 8601806]
31. Hysi PG et al. Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability. Nat Genet 50, 652–656 (2018). [PubMed: 29662168]
32. Fabian-Jessing BK et al. Ocular albinism with infertility and late-onset sensorineural hearing loss. Am J Med Genet A 176, 1587–1593 (2018). [PubMed: 30160833]
33. Thorleifsson G et al. Common sequence variants in the LOXL1 gene confer susceptibility to exfoliation glaucoma. Science 317, 1397–400 (2007). [PubMed: 17690259]
34. Rivera A et al. Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. Hum Mol Genet 14, 3227–36 (2005). [PubMed: 16174643]
35. Morimura H, Saindelle-Ribeaudeau F, Berson EL & Dryja TP Mutations in RGR, encoding a light-sensitive opsin homologue, in patients with retinitis pigmentosa. Nat Genet 23, 393–4 (1999). [PubMed: 10581022]
36. Robinson PN et al. Mutations of FBN1 and genotype-phenotype correlations in Marfan syndrome and related fibrillinopathies. Hum Mutat 20, 153–61 (2002). [PubMed: 12203987]
37. Rohde K, Moller M & Rath MF Homeobox genes and melatonin synthesis: regulatory roles of the cone-rod homeobox transcription factor in the rodent pineal gland. Biomed Res Int 2014, 946075 (2014).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

38. Chakraborty R et al. Circadian rhythms, refractive development, and myopia. Ophthalmic Physiol Opt 38, 217–245 (2018). [PubMed: 29691928]
39. Carnes MU, Allingham RR, Ashley-Koch A & Hauser MA Transcriptome analysis of adult and fetal trabecular meshwork, cornea, and ciliary body tissues by RNA sequencing. Exp Eye Res 167, 91–99 (2018). [PubMed: 27914989]
40. Ratnapriya R et al. Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. Nat Genet 51, 606–610 (2019). [PubMed: 30742112]
41. Finucane HK et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat Genet 50, 621–629 (2018). [PubMed: 29632380]
42. Zhu Z et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48, 481–7 (2016). [PubMed: 27019110]
43. Westra HJ et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45, 1238–1243 (2013). [PubMed: 24013639]
44. Qi T et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat Commun 9, 2282 (2018). [PubMed: 29891976]
45. Buniello A et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 47, D1005–D1012 (2019). [PubMed: 30445434]
46. Benyamin B et al. Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. Mol Psychiatry 19, 253–8 (2014). [PubMed: 23358156]
47. Mountjoy E et al. Education and myopia: assessing the direction of causality by mendelian randomisation. BMJ 361, k2022 (2018).
48. Pozarickij A et al. Quantile regression analysis reveals widespread evidence for gene-environment or gene-gene interactions in myopia development. Commun Biol 2, 167 (2019). [PubMed: 31069276]
49. Audo I et al. TRPM1 is mutated in patients with autosomal-recessive complete congenital stationary night blindness. Am J Hum Genet 85, 720–9 (2009). [PubMed: 19896113]
50. Khawaja AP et al. Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. Nat Genet 50, 778–782 (2018). [PubMed: 29785010]
51. Yang J et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 44, 369–75, S1–3 (2012). [PubMed: 22426310]
52. Zhang Y, Qi G, Park JH & Chatterjee N Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. Nat Genet 50, 1318–1326 (2018). [PubMed: 30104760]
53. Zadnik K et al. Normal eye growth in emmetropic schoolchildren. Optom Vis Sci 81, 819–28 (2004). [PubMed: 15545807]
54. Li Z et al. Recessive mutations of the gene TRPM1 abrogate ON bipolar cell function and cause complete congenital stationary night blindness in humans. Am J Hum Genet 85, 711–9 (2009). [PubMed: 19878917]
55. Chow RL et al. Vsx1, a rapidly evolving paired-like homeobox gene expressed in cone bipolar cells. Mech Dev 109, 315–22 (2001). [PubMed: 11731243]
56. Struck MC Albinism: Update on ocular features. Current Ophthalmology Reports 3, 232–237 (2015).
57. Mohammad S et al. Characterization of Abnormal Optic Nerve Head Morphology in Albinism Using Optical Coherence Tomography. Invest Ophthalmol Vis Sci 56, 4611–8 (2015). [PubMed: 26200501]
58. Yahalom C et al. Refractive profile in oculocutaneous albinism and its correlation with final visual outcome. Br J Ophthalmol 96, 537–9 (2012). [PubMed: 22133989]
59. Lopez VM, Decatur CL, Stamer WD, Lynch RM & McKay BS L-DOPA is an endogenous ligand for OA1. PLoS Biol 6, e236 (2008).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

60. Karouta C & Ashby RS Correlation between light levels and the development of deprivation myopia. Invest Ophthalmol Vis Sci 56, 299–309 (2014). [PubMed: 25491298]
61. Wu PC, Tsai CL, Wu HL, Yang YH & Kuo HK Outdoor activity during class recess reduces myopia onset and progression in school children. Ophthalmology 120, 1080–5 (2013). [PubMed: 23462271]
62. Troilo D, Gottlieb MD & Wallman J Visual deprivation causes myopia in chicks with optic nerve section. Curr Eye Res 6, 993–9 (1987). [PubMed: 3665562]
63. de Leeuw CA, Neale BM, Heskes T & Posthuma D The statistical properties of gene-set analysis. Nat Rev Genet 17, 353–64 (2016). [PubMed: 27070863]

References

64. Cumberland PM et al. Frequency and Distribution of Refractive Error in Adult Life: Methodology and Findings of the UK Biobank Study. PLoS One 10, e0139780 (2015).
65. Bycroft C et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 562, 203–209 (2018). [PubMed: 30305743]
66. Delaneau O, Marchini J & Zagury JF A linear complexity phasing method for thousands of genomes. Nat Methods 9, 179–81 (2011). [PubMed: 22138821]
67. Loh P-R, Kichaev G, Gazal S, Schoech AP & Price AL Mixed-model association for biobank-scale datasets. Nature genetics, 1 (2018).
68. Fuchsberger C, Abecasis GR & Hinds DA minimac2: faster genotype imputation. Bioinformatics 31, 782–4 (2015). [PubMed: 25338720]
69. Banda Y et al. Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics 200, 1285–95 (2015). [PubMed: 26092716]
70. Howie BN, Donnelly P & Marchini J A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5, e1000529 (2009).
71. Tedja MS et al. Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. Nat Genet 50, 834–848 (2018). [PubMed: 29808027]
72. Riboli E & Kaaks R The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 26 Suppl 1, S6–14 (1997). [PubMed: 9126529]
73. Hayat SA et al. Cohort profile: A prospective cohort study of objective physical and cognitive capability and visual health in an ageing population of men and women in Norfolk (EPIC-Norfolk 3). Int J Epidemiol 43, 1063–72 (2014). [PubMed: 23771720]
74. Delaneau O, Marchini J, Genomes Project, C. & Genomes Project, C. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat Commun 5, 3934 (2014). [PubMed: 25653097]
75. Willer CJ, Li Y & Abecasis GR METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–1 (2010). [PubMed: 20616382]
76. Winkler TW et al. Quality control and conduct of genome-wide association meta-analyses. Nat Protoc 9, 1192–212 (2014). [PubMed: 24762786]
77. Yang J, Lee SH, Goddard ME & Visscher PM GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88, 76–82 (2011). [PubMed: 21167468]
78. Yang J et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42, 565–9 (2010). [PubMed: 20562875]
79. Bulik-Sullivan B et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 47, 1236–41 (2015). [PubMed: 26414676]
80. Buniello A et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 47, D1005–D1012 (2019). [PubMed: 30445434]
81. Finucane HK et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat Genet 50, 621–629 (2018). [PubMed: 29632380]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

82. Carnes MU, Allingham RR, Ashley-Koch A & Hauser MA Transcriptome analysis of adult and fetal trabecular meshwork, cornea, and ciliary body tissues by RNA sequencing. Exp Eye Res 167, 91–99 (2018). [PubMed: 27914989]
83. Ratnapriya R et al. Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. Nat Genet 51, 606–610 (2019). [PubMed: 30742112]
84. Zhu Z et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48, 481–7 (2016). [PubMed: 27019110]
85. Westra HJ et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45, 1238–1243 (2013). [PubMed: 24013639]
86. Qi T et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat Commun 9, 2282 (2018). [PubMed: 29891976]
87. Segre AV et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet 6(2010).
88. Pybus M et al. 1000 Genomes Selection Browser 1.0: a genome browser dedicated to signatures of natural selection in modern humans. Nucleic Acids Res 42, D903–9 (2014). [PubMed: 24275494]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

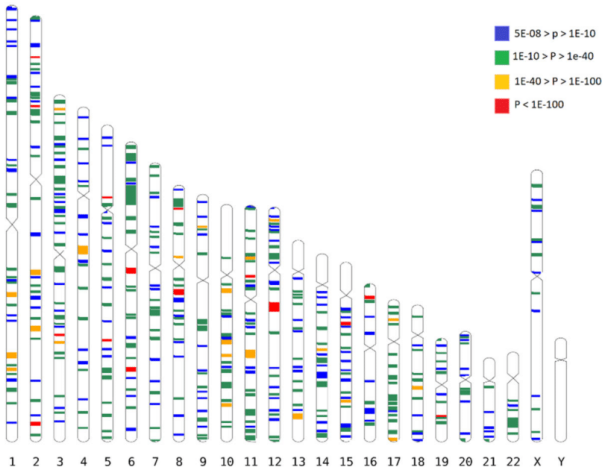


Figure 1.

All GWAS-associated regions from the main meta-analysis. Each band is a true scale of genomic regions associated with refractive error listed in Supplementary Table 1 (+250kbp on each side to make smaller regions more visible). The different color codes represent the significance (p-value) for the genetic variant within that region that displays the strongest evidence for association.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

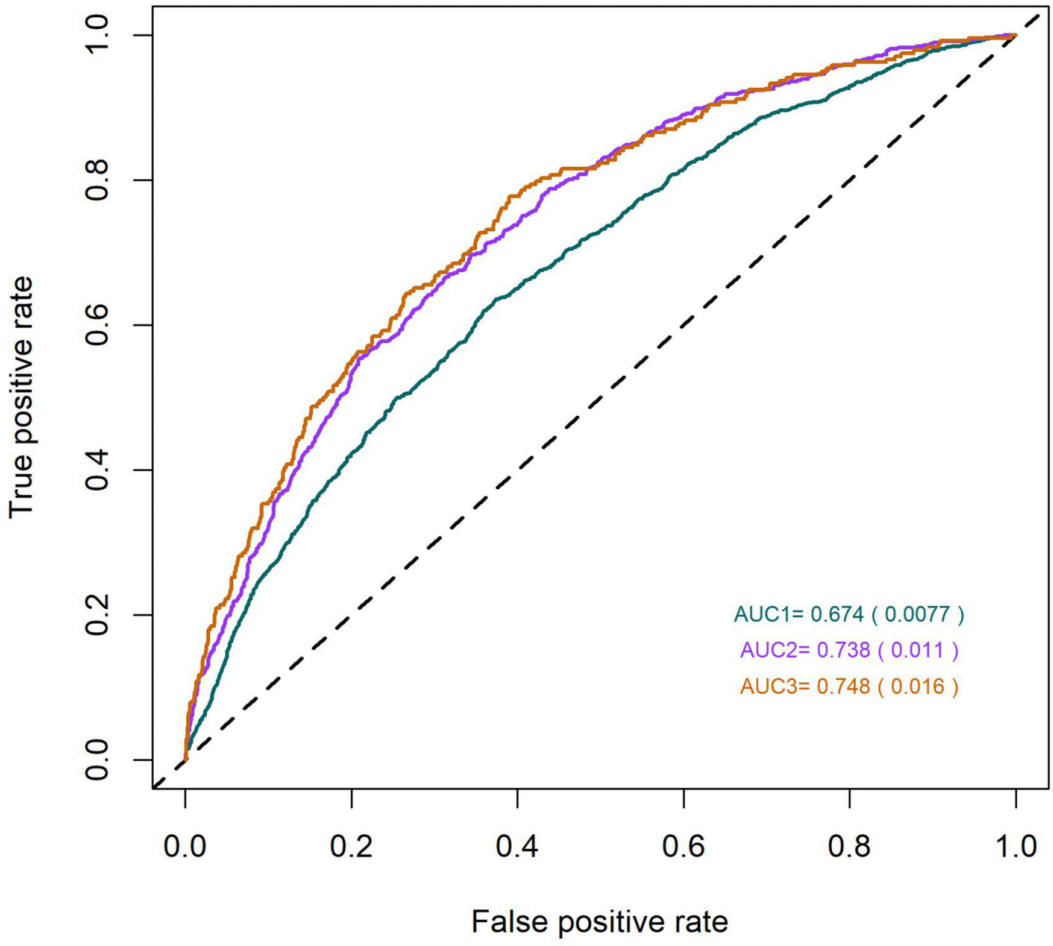


Figure 2.
Receiver Operating Characteristic (ROC) curves for myopia predictions, using information from 890 SNP markers identified in the meta-analysis. The three different colors represent three different curves for each of the different definition of myopia: green – all myopia (< −0.75D), magenta – moderate myopia (< −3.00 D) and brown - severe myopia (defined as < −5.00 D).

Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia

Supplementary Notes

Table of Contents

1 UK Biobank Eye and Vision Consortium Membership 3
2 The CREAM Consortium 6
3 23andMe 9
4 Cohort Description 10
4.1 UK Biobank 10
4.1.1 Phenotyping 10
4.1.2 Genotyping 14
4.1.3 Association analyses 15
4.2 23andMe 16
4.2.1 Phenotyping 16
4.2.2 Genotyping 17
4.2.3 Association analyses 18
4.3 GERA 18
4.3.1 Phenotyping 18
4.3.2 Genotyping 18
4.3.3 Association analyses 19
4.4 Cream Consortium 19
4.4.1 Phenotyping 19
4.4.2 Genotyping 19
4.4.3 Association analyses 19
4.5 EPIC 20
4.5.1 Phenotyping 20
4.5.2 Genotyping and imputation 20
4.5.3 Association analysis 21
5 STATISTICAL ANALYSES 21
5.1 Meta-analyses 21
5.2 Conditional analyses 21
5.3 Multiple testing correction 21
5.4 Genomic inflation 21
5.5 LDscore regression-based methods 22
5.5.1 Polygenicity vs. inflation 22
5.5.2 Calculation of genetic correlation 22
5.6 Associated SNPs and gene annotations 22
5.6.1 OMIM 22
5.6.2 The GWAS Catalog. 22

5.7 Graphical illustration of association 22
5.8 Mendelian randomization 22
5.9 Gene expression, GTEx and other transcription data 22
5.10 LD score regression applied to specifically expressed genes (LDSC-SEG) 23
5.11 SMR 23
5.11.1 Test description 23
5.11.2 Datasets for the SMR analyses: eQTL, *cis*-mQTL 23
5.12 Gene-set enrichment 24
5.12.1 GSEA definitions 24
5.13 Analyses of signals of natural selection 24
5.14 Estimation of effect size distributions for spherical equivalent 24
6 FULL ACKNOWLEDGEMENTS 25
6.1 UK Biobank 25
6.2 23andMe 25
6.3 GERA 25
6.4 Consortium for Refractive Error and Myopia (CREAM) 25
6.5 EPIC-Norfolk 25
6.6 King's College London authors 26
6.7 UK Biobank Eye Consortium members 26
6.8 University College London authors 26

1 UK Biobank Eye and Vision Consortium Membership

- Prof **Tariq ASLAM** - Manchester University, Manchester, United Kingdom
- Prof **Sarah BARMAN** - Kingston University, London, United Kingdom
- Prof **Jenny BARRETT** - University of Leeds, Yorkshire, United Kingdom
- Prof **Paul BISHOP** - Manchester University, Manchester, United Kingdom
- Mr **Peter BLOWS** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Catey BUNCE** - King's College London, London, United Kingdom
- Dr **Roxana CARARE** - University of Southampton, Southampton, United Kingdom
- Prof **Usha CHAKRAVARTHY** - Queens University Belfast, Belfast, Ireland
- Miss **Michelle CHAN** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Sharon CHUA** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **David CRABB** – City, University of London, London, United Kingdom
- Mrs **Philippa CUMBERLAND** - UCL Great Ormond Street Institute of Child Health, London, United Kingdom
- Dr **Alexander DAY** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Miss **Parul DESAI** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Bal DHILLON** - University of Edinburgh, Scotland, United Kingdom
- Prof Andrew DICK - University of Bristol, Bristol, United Kingdom
- Dr **Cathy EGAN** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Sarah ENNIS** - University of Southampton, Southampton, United Kingdom
- Prof **Paul FOSTER** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Marcus FRUTTIGER** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **John GALLACHER** - University of Oxford, Oxford, United Kingdom
- Prof **David (Ted) GARWAY-HEATH** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Jane GIBSON** - University of Southampton, Southampton, United Kingdom
- Mr **Dan GORE** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Jeremy GUGGENHEIM** - Cardiff University, Wales, United Kingdom
- Prof **Chris HAMMOND** - King's College London, London, United Kingdom
- Prof **Alison HARDCASTLE** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Simon HARDING** - University of Liverpool, London, United Kingdom

- Dr **Ruth HOGG** - Queens University Belfast, Belfast, Ireland
- Dr **Pirro HYS**I - King's College London, London, United Kingdom
- Mr **Pearse A KEANE** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Sir Peng Tee KHAW** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Mr **Anthony KHAWAJA** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Mr **Gerassimos LASCARATOS** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Andrew LOTERY**- University of Southampton, Southampton, United Kingdom
- Prof Phil LUTHERT - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Tom MACGILLIVRAY** - University of Edinburgh, Scotland, United Kingdom
- Dr **Sarah MACKIE** - University of Leeds, Yorkshire, United Kingdom
- Prof **Keith MARTIN** - University of Cambridge, Cambridge, United Kingdom
- Ms **Michelle MCGAUGHEY** - Queen's University Belfast, Belfast, Ireland
- Dr **Bernadette MCGUINNESS** - Queen's University Belfast, Belfast, Ireland
- Dr **Gareth MCKAY** - Queen's University Belfast, Belfast, Ireland
- Mr **Martin MCKIBBIN** - Leeds Teaching Hospitals NHS Trust, Yorkshire, United Kingdom
- Dr **Danny MITRY** – NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom & Royal Free Hospital, London, United Kingdom
- Prof **Tony MOORE** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **James MORGAN** - Cardiff University, Wales, United Kingdom
- Ms **Zaynah MUTHY** – NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Mr **Eoin O'SULLIVAN** - King's College Hospital NHS Foundation Trust, London, United Kingdom
- Dr **Chris OWEN** - St George's, University of London, London, United Kingdom
- Mr Praveen PATEL - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Mr **Euan PATERSON** - Queens University Belfast, Belfast, Ireland
- Dr **Tunde PETO** - Queen's University Belfast, Belfast, Ireland
- Dr **Axel PETZOLD** - UCL Institute of Neurology, London, United Kingdom
- Prof **Jugnoo RAHI** - UCL Great Ormond Street Institute of Child Health, London, United Kingdom
- Dr **Alicja RUDNICKA** - St George's, University of London, London, United Kingdom
- Mr **Jay SELF** - University of Southampton, Southampton, United Kingdom
- Prof **Sobha SIVAPRASAD** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Mr **David STEEL** - Newcastle University, Newcastle, United Kingdom
- Mrs **Irene STRATTON** - Gloucestershire Hospitals NHS Foundation Trust

- Mr **Nicholas STROUTHIDIS** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Cathie SUDLOW** - University of Edinburgh, Scotland, United Kingdom
- Dr **Caroline THAUNG** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Miss **Dhanes THOMAS** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Prof **Emanuele TRUCCO** - University of Dundee, Scotland, United Kingdom
- Prof **Adnan TUFAIL** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Veronique VITART** - University of Edinburgh, Scotland, United Kingdom
- Prof **Stephen VERNON** – Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom
- Mr **Ananth VISWANATHAN** - NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
- Dr **Cathy WILLIAMS** - University of Bristol, Bristol, United Kingdom
- Dr **Katie WILLIAMS** - King's College London, London, United Kingdom
- Prof **Jayne WOODSIDE** - Queen's University Belfast, Belfast, Ireland
- Dr **Max YATES** - University of East Anglia, Norwich, United Kingdom
- Ms **Jennifer YIP** - University of Cambridge, Cambridge, United Kingdom
- Dr **Yalin ZHENG** - University of Liverpool, London, United Kingdom

2 The CREAM Consortium

Joan E. Bailey-Wilson[1], Paul Nigel Baird[2], Amutha Barathi Veluchamy[3-5], Ginevra Biino[6], Kathryn P. Burdon[7], Harry Campbell[8], Li Jia Chen[9], Ching-Yu Cheng[10-12], Emily Y. Chew[13], Jamie E. Craig[14], Phillippa M. Cumberland[15], Margaret M. Deangelis[16], Cécile Delcourt[17], Xiaohu Ding[18], Cornelia M. van Duijn[19], David M. Evans[20-22], Qiao Fan[23], Maurizio Fossarello[24], Paul J. Foster[25], Puya Gharahkhani[26], Adriana I. Iglesias[19,27,28], Jeremy A. Guggenheim[29], Xiaobo Guo[18,30], Annechien E.G. Haarman[19,28], Toomas Haller[31], Christopher J. Hammond[32], Xikun Han[26], Caroline Hayward[33], Mingguang He[2,18], Alex W. Hewitt[2,7,34], Quan Hoang[3,35] , Pirro G. Hysi[32], Robert P. Igo Jr.[36], Sudha K. Iyengar[36-38], , Jost B. Jonas[39,40], Mika Kähönen[41,42], Jaakko Kaprio[43,44], , Anthony P. Khawaja[25,45], Caroline C. W. Klaver[19,28,46], Barbara E. Klein[47], Ronald Klein[47], Jonathan H. Lass[36,37], Kris Lee[47], Terho Lehtimäki[48,49], Deyana Lewis[1], Qing Li[50], Shi-Ming Li[40], Leo-Pekka Lyytikäinen[48,49], Stuart MacGregor[26], David A. Mackey[2,7,34], Nicholas G. Martin[51], Akira Meguro[52], Andres Metspalu[31], Candace Middlebrooks[1], Masahiro Miyake[53], Nobuhisa Mizuki[52], Anthony Musolf[1], Stefan Nickels[54], Konrad Oexle[55], Chi Pui Pang[9], Olavi Pärssinen[56,57], Andrew D. Paterson[58], Norbert Pfeiffer[54], Ozren Polasek[59,60], Jugnoo S. Rahi[15,25,61], Olli Raitakari[62,63], Igor Rudan[8], Srujana Sahebjada[2], Seang-Mei Saw[64,65], Dwight Stambolian[66], Claire L. Simpson[1,67], E-Shyong Tai[65], Milly S. Tedja[19,28], J. Willem L. Tideman[19,28], Akitaka Tsujikawa[53], Virginie J.M. Verhoeven[19,27,28], Veronique Vitart[33], Ningli Wang[40], Juho Wedenoja[43,68], Wen Bin Wei[69], Cathy Williams[22], Katie M. Williams[32], James F. Wilson[8,33], Robert Wojciechowski[1,70,71], Ya Xing Wang[40], Kenji Yamashiro[72], Jason C. S. Yam[9], Maurice K.H. Yap[73], Seyhan Yazar[34], Shea Ping Yip[74], Terri L. Young[47], Xiangtian Zhou[75]

Affiliations

1. Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.
2. Centre for Eye Research Australia, Ophthalmology, Department of Surgery, University of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Australia.
3. Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
4. Duke-NUS Medical School, Singapore, Singapore.
5. Department of Ophthalmology, National University Health Systems, National University of Singapore, Singapore.
6. Institute of Molecular Genetics, National Research Council of Italy, Pavia, Italy.
7. Department of Ophthalmology, Menzies Institute of Medical Research, University of Tasmania, Hobart, Australia.
8. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK.
9. Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong Eye Hospital, Kowloon, Hong Kong.
10. Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.
11. Ocular Epidemiology Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
12. Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore.
13. Division of Epidemiology and Clinical Applications, National Eye Institute/National Institutes of Health, Bethesda, USA.
14. Department of Ophthalmology, Flinders University, Adelaide, Australia.

15. Great Ormond Street Institute of Child Health, University College London, London, UK.
16. Department of Ophthalmology and Visual Sciences, John Moran Eye Center, University of Utah, Salt Lake City, Utah, USA.
17. Université de Bordeaux, Inserm, Bordeaux Population Health Research Center, team LEHA, UMR 1219, F-33000 Bordeaux, France.
18. State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China.
19. Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands.
20. Translational Research Institute, University of Queensland Diamantina Institute, Brisbane, Queensland, Australia.
21. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.
22. Department of Population Health Sciences, Bristol Medical School, Bristol, UK.
23. Centre for Quantitative Medicine, DUKE-National University of Singapore, Singapore.
24. University Hospital 'San Giovanni di Dio', Cagliari, Italy.
25. NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK.
26. Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
27. Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.
28. Department of Ophthalmology, Erasmus Medical Center, Rotterdam, The Netherlands.
29. School of Optometry & Vision Sciences, Cardiff University, Cardiff, UK.
30. Department of Statistical Science, School of Mathematics, Sun Yat-Sen University, Guangzhou, China.
31. Institute of Genomics, University of Tartu, Tartu, Estonia.
32. Section of Academic Ophthalmology, School of Life Course Sciences, King's College London, London, UK.
33. MRC Human Genetics Unit, MRC Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh, UK.
34. Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western Australia, Perth, Australia.
35. Department of Ophthalmology, Columbia University, New York, USA.
36. Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, USA.
37. Department of Ophthalmology and Visual Sciences, Case Western Reserve University and University Hospitals Eye Institute, Cleveland, Ohio, USA.
38. Department of Genetics, Case Western Reserve University, Cleveland, Ohio, USA.
39. Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University of Heidelberg, Mannheim, Germany.
40. Beijing Tongren Eye Center, Beijing Tongren Hospital, Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Capital Medical University, Beijing, China.
41. Department of Clinical Physiology, Tampere University Hospital and School of Medicine, University of Tampere, Tampere, Finland.
42. Finnish Cardiovascular Research Center, Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland.
43. Department of Public Health, University of Helsinki, Helsinki, Finland.
44. Institute for Molecular Medicine Finland FIMM, HiLIFE Unit, University of Helsinki, Helsinki, Finland.
45. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
46. Department of Ophthalmology, Radboud University Medical Center, Nijmegen, The Netherlands.
47. Department of Ophthalmology and Visual Sciences, University of Wisconsin–Madison, Madison, Wisconsin, USA.

48. Department of Clinical Chemistry, Finnish Cardiovascular Research Center-Tampere, Faculty of Medicine and Life Sciences, University of Tampere.
49. Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere, Tampere, Finland.
50. National Human Genome Research Institute, National Institutes of Health, Baltimore, USA.
51. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
52. Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama, Kanagawa, Japan.
53. Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan.
54. Department of Ophthalmology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany.
55. Institute of Neurogenomics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany.
56. Department of Ophthalmology, Central Hospital of Central Finland, Jyväskylä, Finland.
57. Gerontology Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland.
58. Program in Genetics and Genome Biology, Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada.
59. Gen-info Ltd, Zagreb, Croatia. .
60. University of Split School of Medicine, Soltanska 2, Split, Croatia.
61. Ulverscroft Vision Research Group, University College London, London, UK.
62. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland.
63. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland.
64. Myopia Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
65. Saw Swee Hock School of Public Health, National University Health Systems, National University of Singapore, Singapore.
66. Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
67. Department of Genetics, Genomics and Informatics, University of Tennessee Health Sciences Center, Memphis, Tenessee.
68. Department of Ophthalmology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.
69. Beijing Tongren Eye Center, Beijing Key Laboratory of Intraocular Tumor Diagnosis and Treatment, Beijing Ophthalmology & Visual Sciences Key Lab, Beijing Tongren Hospital, Capital Medical University, Beijing, China.
70. Department of Epidemiology and Medicine, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.
71. Wilmer Eye Institute, Johns Hopkins Medical Institutions, Baltimore, Maryland, USA.
72. Department of Ophthalmology, Otsu Red Cross Hospital, Nagara, Japan. .
73. Centre for Myopia Research, School of Optometry, The Hong Kong Polytechnic University, Hong Kong, Hong Kong.
74. Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, Hong Kong.
75. School of Ophthalmology and Optometry, Eye Hospital, Wenzhou Medical University, China.

3 23andMe

The following members of the 23andMe Research Team contributed to this study:

Michelle Agee

Babak Alipanahi

Adam Auton

Robert K. Bell

Katarzyna Bryc

Sarah L. Elson

Pierre Fontanillas

Nicholas A. Furlotte

David A. Hinds

Karen E. Huber

Aaron Kleinman

Nadia K. Litterman

Jennifer C. McCreight

Matthew H. McIntyre

Joanna L. Mountain

Elizabeth S. Noblin

Carrie A.M. Northover

Steven J. Pitts

J. Fah Sathirapongsasuti

Olga V. Sazonova

Janie F. Shelton

Suyash Shringarpure

Chao Tian

Joyce Y. Tung

Vladimir Vacic

Catherine H. Wilson.

4 Cohort Description

4.1 UK Biobank

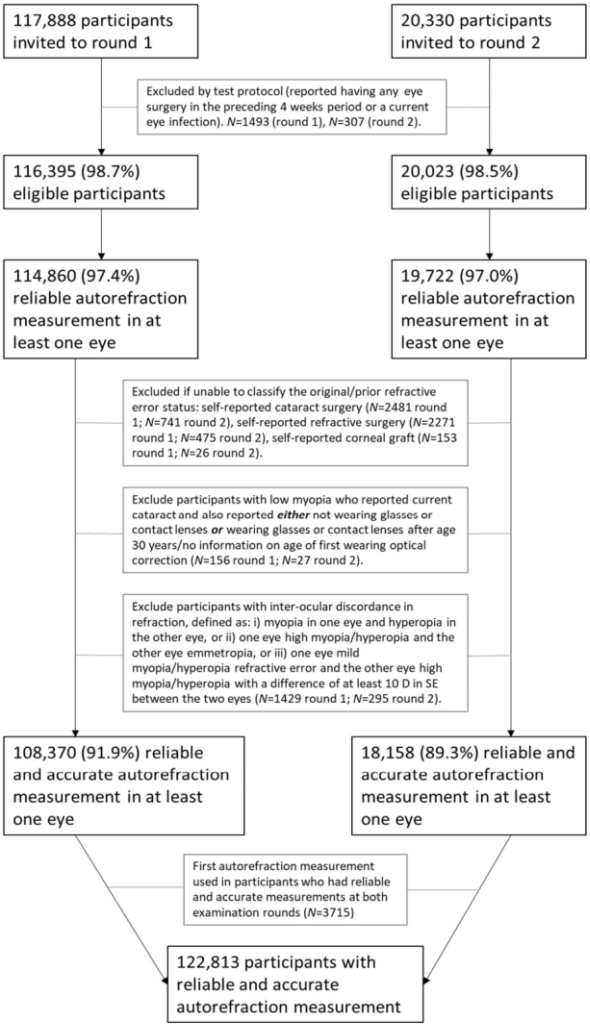
The UK Biobank is a very large multisite cohort study established by the Medical Research Council, Department of Health, Wellcome Trust medical charity, Scottish Government and Northwest Regional Development Agency. A baseline questionnaire, measurements, and biological samples were undertaken in 22 assessment centers across the UK between 2006 and 2010.

4.1.1 Phenotyping

4.1.1.1 Spherical equivalent (UK Biobank – 1)

Ophthalmic assessment was not part of the original baseline assessment and was introduced as an enhancement in 2009 for 6 assessment centers which are spread across the UK (Liverpool and Sheffield in North England, Birmingham in the Midlands, Swansea in Wales, and Croydon and Hounslow in Greater London). Participants completed a touch-screen self-administered questionnaire. The response options for ethnicity included White (English/Irish or other white background), Asian or British Asian (Indian/Pakistani/Bangladeshi or other Asian background), Black or Black British (Caribbean, African, or other black background), Chinese, mixed (White and Black Caribbean or African, White and Asian, or other mixed background), or other, non-defined, ethnic group.

Refractive error (RE) was measured by non-cycloplegic autorefraction in both eyes using the Tomey RC 5000 Auto Refkeratometer (Tomey Corp., Nagoya, Japan). The right eye was measured first and up to 10 measurements were taken per eye. The most representative result was automatically recorded. To ensure reliable and accurate RE data, we excluded participants based on previously published criteria[1] (Supplementary Note Figure 1). Spherical equivalent was calculated as spherical refractive error (UK Biobank codes 5084 and 5085) plus half the cylindrical error (UK Biobank 5086 and 5087) for each eye. If reliable data were only available for one eye, the RE of that eye was considered as the participant's RE. If reliable data were available for both eyes, we calculated the mean of right and left RE as the participant's RE.



Supplementary Note Figure 1. Flowchart explaining the exclusion of UK Biobank participants, as described elsewhere[1], for whom spherical equivalent measurements were available, based on existing clinical data.

A summary of the basic demographic characteristics of the subset of the UK Biobank for which the spherical equivalent was directly measured and which was used for our analysis is given below (Supplementary Note Table 1).

		Participants	Spherical Equivalent (diopters)	Age (years)
		N (%)	Mean ± SD	Mean ± SD
All		102,117	-0.28D (2.74)	57.26 (7.86)
Sex	**Male**	47,774	-0.27D (2.65)	57.59 (7.92)
	Female	54,343	-0.29D (2.82)	56.97(7.79)

Supplementary Note Table 1**.** Characteristics of UK Biobank participants for whom spherical equivalent measurements were available and included in the quantitative spherical equivalent analysis.
Abbreviations: N, number; SD, standard deviation; Age, age at first documented spherical equivalent assessment

4.1.1.2 Inferring myopia case-control status from self-reported age of spectacle wear (UKB2 sample)

Direct measurements of RE were only available for just 22.7% of the entire UK Biobank sample. However refractive error is strongly correlated with several clinical parameters and demographic factors[2], which can be used to predict refractive error and myopia. Some of that indirect information was present for significant numbers among the UK Biobank participants.

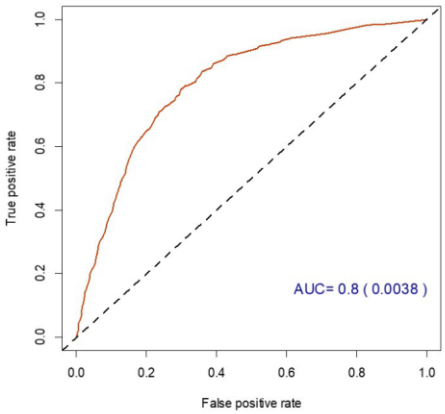
For example, age when the first lens correction is prescribed is strongly correlated to spherical equivalent both at a genetic[3] and phenotypic level[4]. It has used before as a proxy for RE previously[5] with reasonably low levels of genetic effect heterogeneity[6].A total of 87% of the participants, responded to the question on whether their vision needed correction when they were asked "Do you wear glasses or contact lenses to correct your vision?" (potential answers: Yes/No/Prefer not to answer) in a touch-screen self-administered questionnaire. If a participant answered "Yes" to this question, they were further asked "What age did you first start to wear glasses or contact lenses?", to which 67% of the participants responded. In addition, spherical equivalent and myopia affection status is highly correlated with age, sex, birth year[4,7] , all of which were available for UK Biobank participants.

We therefore aimed at harnessing the demographic and clinical information to obtain an estimate about the individual's likely myopia status. This general approach has been used successfully before[6], and to better classify the non-refracted subjects into myopia cases and non-myopia controls.

We proceeded in three steps: 1) training a Support Vector Machine model in 80% randomly selected UK Biobank participants of European descent for whom direct spherical equivalent and refractive error

status were available, 2) validate the prediction in the remaining 20% of UK Biobank participants of European descent for whom direct spherical equivalent and refractive error status were available, 3) make the SVM predictions in the remaining individuals with no direct spherical error measurements available using the model developed for the training data.

We initially fine-tuned the prediction model, in order to optimize the γ (gamma) and "cost" parameters in the 80% training data sample. The performance of the model was subsequently tested in the remaining 20% of the UK Biobank participants with spherical equivalent measurements. Receiver Operating Characteristic (ROC) curves were drawn and the area under the curve (AUC) calculated (Supplementary Note Figure 2). For the optimal SVM model, an Area Under the Curve (AUC) of 0.8 was obtained, providing strong evidence that myopia status could be inferred from 'age of first spectacle wear' with sufficient reliability to serve as a proxy for myopia status in a GWAS analysis. When we compared predicted vs observed myopia case-control status in the 20% validation subset of the UK Biobank for which spherical equivalent were available (cases defined as < -0.75 Diopters and controls > -0.5Diopters), the inferred myopia status was a strong predictor for actual (observed) myopia case-control status (OR=16.79, 95%CI 15.99-17.62).



Supplementary Note Figure 2. ROC curve and AUCs for the SVM model trained in 80% of fully phenotyped (i.e. spherical equivalent available) UK Biobank participants. For both ROC and AUC, the remaining 20% (i.e. not part of the initial training set) of the UK Biobank were used. Cases defined as < -0.75 Diopters and controls > -0.5Diopters. The standard error for the AUC is shown within brackets.

The optimal SVM model was applied to the UK Biobank sample with known 'age of first spectacle wear' but for whom no spherical equivalent measurements were available. That is, the GWAS for SVM-inferred myopia case-control status did not include participants with known (measured) refractive error. This inferred phenotypic status was used for a GWAS analysis, for which we followed the previously

described workflow (sections 4.1.3). In further support of SVM-inferred myopia case-control status as a valid proxy measure, the genetic effect sizes for variants assessed in the SVM-inferred phenotype were found to be highly correlated to those from the GWAS analyses for autorefractor-measured refractive error (Supplementary Note Table 2, Supplementary File 1**Error! Reference source not found.**).

	UKB1[1]	GERA[1]	23andMe[2,3]	UKB2[2,4]
Beta - UKB1	1.000	0.926	0.953	0.935
Beta - GERA	0.926	1.000	0.927	0.901
23andMe - log(OR)	0.953	0.927	1.000	0.941
UKB-2 log(OR)	0.935	0.901	0.941	1.000

Supplementary Note Table 2. Correlation of effects sizes between the UKB-2 subset (for which spherical equivalent was not available and myopia status was imputed using the SVM) with other cohorts in which refraction was directly measured or self-reported. "Beta" linear regression coefficients; log(OR) the logarithm of the logistic regression Odds Ratios; UKB-1 denotes the first subset of the UK Biobank participants (spherical equivalent available), UKB -2 the second subset (myopia case-control status inferred). For a description of the 23andMe and GERA cohorts, please refer to subsequent sections for further cohort descriptions.

The basic characteristics of the participants that were selected for the second UK Biobank subset (UKB-2) of our meta-GWAS are shown in the Supplementary Note Table 3.

		inferred cases/controls	**Age first spectacle wear (years)**	**Age (years)**
		N cases (%)	Mean ± SD	Mean ± SD
All		108,956/70,941 (60.56)	26.1 ± 13.58	56.61 ± 7.79
Sex	**Male**	45,994/30,388 (60.25)	26.38 ± 13.48	56.78 ± 7.88
	Female	62,962 / 40,553 (60.82)	25.74 ± 13.65	56.48 ± 7.73

Supplementary Note Table 3**.** Characteristics of UK Biobank participants included in the inferred myopia qualitative analysis (UKB-2). Abbreviations: N, number; SD, standard deviation; Age, age at first documented spherical equivalent assessment.

4.1.2 Genotyping

DNA extraction, genotyping and imputation of UK Biobank participants has been reported elsewhere[8]. DNA extraction begun on buffy coat samples. DNA was extracted from 850 µl buffy coat (recovered from 9 ml of whole blood) on customized TECAN Freedom EVO® 200 platform[9]. The samples were then processed in the approximate order received to produce genotype data. Genotyping was done using two arrays. The first array was the Affymetrix Axiom® platform with a custom-designed array described in the UK Biobank Axiom® Array Content Summary[10]. Processing was done using a LIMS system to track instrumentation, Axiom consumables arrays and reagents and operators. The process is described

elsewhere[11]. Details on genotyping procedure and quality control can be found elsewhere[12]. The second array is what was the UK BiLEVE, described elsewhere[13].

Phasing on the autosomes was carried out using a modified version of the SHAPEIT2[14] program modified to allow for very large sample sizes. This new method (which we refer to as SHAPEIT3) modifies the SHAPEIT2 surrogate family approach to remove a quadratic complexity component of the algorithm[15]. In small sample sizes of a few thousand samples, this part of the algorithm, which involves calculating Hamming distances between current haplotypes estimates, contributes only a relatively small part to the computational cost. As sample sizes increase over 10,000 samples then this component becomes significant. The new algorithm uses a divisive clustering algorithm to identify clusters of haplotypes, and then calculates Hamming distances only between pairs of haplotypes within each cluster. Only haplotypes within each cluster are used as candidates for the surrogate family copying states in the HMM model.

A total of 806,466 directly genotyped DNA sequence variants were available after variant quality control. The UK Biobank team then performed imputation from a combined Haplotype Reference Consortium (HRC) and UK10K reference panel; phasing was performed using SHAPEIT3 and imputation was carried out via the IMPUTE4 program[16]. Only HRC-imputed variants were used for the purpose our analyses of the UK Biobank participants. The variant-level quality control exclusion metrics applied to imputed data for GWAS included the following: call rate < 95%, Hardy–Weinberg equilibrium $P < 1 \times 10^{-6}$, posterior call probability < 0.9, imputation quality < 0.4, and MAF < 0.005. The Y chromosome and mitochondrial genetic data were excluded from this analysis. In total, 10,263,360 imputed DNA sequence variants were included in our analysis.

For sample quality control, we removed individuals of non-European ancestry and participants with relatedness corresponding to third-degree relatives or closer, and an additional 480 samples with an excess of missing genotype calls or more heterozygosity than expected were excluded. In total, genotypes were available for 102,117 participants of European ancestry with spherical equivalent data.

4.1.3 Association analyses

The basic model tested was the average of spherical equivalent measured in the left and right eye as an outcome of a regression model whose predictor is the allele dosage at a given polymorphic locus, adjusted for the effect of relevant covariables (see table below). The empirical association between spherical equivalent and other covariables is shown in Supplementary Note Table 4.

Effect size estimates (Beta) and P-values are from the multivariate regression model. Since demographic factors and principal components had a small yet real effect over Spherical Equivalent, the above variables were included in the model.
Therefore, models of mixed linear regressions, as described before[17], where the spherical equivalent was the outcome, the allele dosage the predictor, adjusted for age, sex and the first 10 principal components. Since there was, there is evidence of cryptic relatedness among the UK Biobank participants, a linear mixed model that controls for population structure was used as implemented in the Bolt-LMM software [18].

Variable	Beta	SE	Z	P-value
age	0.05247	0.001122	46.773	<E-308
sex	-0.08013	0.017988	-4.455	8.42E-06
PC1	0.000297	0.000125	2.367	0.017935
PC2	0.001096	0.000272	4.032	5.54E-05
PC3	0.001369	0.000516	2.655	0.007935
PC4	0.005147	0.000799	6.439	1.21E-10
PC5	0.003194	0.001253	2.55	0.010786
PC6	0.004701	0.001314	3.577	0.000348
PC7	0.008073	0.001614	5.002	5.68E-07
PC8	0.000269	0.001605	0.168	0.866871
PC9	0.008165	0.002385	3.424	0.000618
PC10	-0.00039	0.001885	-0.207	0.836117

Online Note Table 4. The association between spherical equivalent, age, sex and the first 10 Principal Components.

For the second UK Biobank subset, for which no spherical equivalent information was available, the mixed linear model was built with the predicted myopia status as outcome and using the same covariates as for the previously described linear regression analysis on spherical equivalent (paragraph 4.1.3). Odds Ratios were obtained from the beta regression coefficient using the equation:

$$\ln(\mathrm{OR}) = \frac{\beta}{\mu(1-\mu)}$$

where μ is the fraction of the cases in the sample (μ=0.606). Although the case-control analysis was quite balanced, we opted to remove genotypes with MAF <0.01 and MAC< 400 recommended elsewhere[19] (which in our samples, most often would be correspond to MAF < 0.001).

4.2 23andMe

4.2.1 Phenotyping

The subjects were all volunteers from the 23andMe (Mountain View, CA, USA) personal genomics company. All participants included in the analyses provided informed consent and answered surveys online according to the approved 23andMe human subjects protocol, which was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (http://www.eandireview.com). The participants were identified as myopia cases if they responded positively to any of the following questions:

1. "Have you ever been diagnosed by a doctor with nearsightedness (near objects are clear, far objects are blurry)?"
2. "Are you nearsighted (near objects are clear, far objects are blurry)?"
3. "What vision problems do you have? Please check all that apply." - Nearsightedness (near objects are clear, far objects are blurry.
4. "Prior to your LASIK eye surgery, what vision problems did you have? Please check all that apply." - Nearsightedness (near objects are clear, far objects are blurry.

Controls were defined as having said "No" or not checking nearsightedness to at least one of the questions above. Subjects who gave discordant answers were removed.

4.2.2 Genotyping

DNA extraction and genotyping were performed on saliva samples by CLIA-certified and CAP-accredited clinical laboratories of Laboratory Corporation of America. Samples were genotyped on one of four genotyping platforms. The V1 and V2 platforms were variants of the Illumina HumanHap550+ BeadChip, including about 25,000 custom SNPs selected by 23andMe, with a total of about 560,000 SNPs. The V3 platform was based on the Illumina OmniExpress+ BeadChip., with custom content to improve the overlap with our V2 array, with a total of about 950,000 SNPs. The V4 platform in current use is a fully custom array, including a lower redundancy subset of V2 and V3 SNPs, with additional coverage of lower-frequency coding variation, and about 570,000 SNPs. Samples that failed to reach 98.5% call rate were re-analyzed. For the GWAS only participants who have >97% European ancestry, as determined through an analysis of local ancestry, were included. For the purposes of ethnic categorization, an algorithm first partitioned phased genomic data into short windows of about 100 SNPs and used a support vector machine (SVM) to classify individual haplotypes into one of 31 reference populations. The SVM classifications then fed into a hidden Markov model (HMM) that accounts for switch errors and incorrect assignments and gives probabilities for each reference population in each window. The reference population data are derived from public datasets (the Human Genome Diversity Project, HapMap, and 1000 Genomes), as well as 23andMe customers who have reported having four grandparents from the same country. A maximal set of unrelated individuals was chosen for each analysis using a segmental identity-by-descent (IBD) estimation algorithm[20]. Individuals were defined as related if they shared more than 700 cM IBD, including regions where the two individuals share either one or both genomic segments identical-by-descent. This level of relatedness corresponds approximately to the minimal expected sharing between first cousins in an outbred population.

Participant genotype data were imputed against the September 2013 release of 1000 Genomes Phase1 reference haplotypes, phased with ShapeIt2[21]. We phased and imputed data for each genotyping platform separately. We phased using an internally developed phasing tool which implements the Beagle haplotype graph-based phasing algorithm[22].

SNPs with Hardy-Weinberg equilibrium $P<10^{-20}$, call rate < 95%, or with large allele frequency discrepancies compared to European 1000 Genomes reference data were excluded from imputation. Imputation was done against all-ethnicity 1000 Genomes haplotypes (excluding monomorphic and singleton sites) using Minimac[23]. For the X chromosome, separate haplotype graphs were built for the non-pseudoautosomal region and each pseudoautosomal region, and these regions were phased separately. Males and females were imputed together using Minimac2[23], as with the autosomes, treating males as homozygous pseudo-diploids for the non-pseudoautosomal region.

HLA allele dosages were imputed from SNP genotype data using HIBAG[24]. We imputed alleles for HLA-A, B, C, DPB1, DQA1, DQB1, and DRB1 loci at four-digit resolution. To test associations between HLA allele dosages and phenotypes, we performed logistic or linear regression using the same set of covariates used in the SNP-based GWAS for that phenotype. We performed separate association tests for each imputed allele.

4.2.3 Association analyses

Association test results were computed by linear regression assuming additive allelic effects. For tests imputed dosages rather than best-guess genotypes were used. Covariates for age, gender, the first ten principal components to account for residual population structure were also included into the model. Results for the X chromosome are computed similarly, with male genotypes coded as if they were homozygous diploid for the observed allele.

4.3 GERA

The Genetic Epidemiology Research in Adult Health and Aging (GERA) cohort is part of the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH) and has been described in detail elsewhere[25,26]. The GERA cohort comprises 110,266 adult men and women who are consented participants in the RPGEH, an unselected cohort of adult participants who are members of Kaiser Permanente Northern California (KPNC), an integrated health care delivery system, with ongoing longitudinal records from vision examinations. For this analysis, 34,998 adults (25 years and older), who self-reported as non-Hispanic white, and who had at least one assessment of spherical equivalent obtained between 2008 and 2014 were included (Supplementary Note Table 5). All study procedures were approved by the Institutional Review Board of the Kaiser Foundation Research Institute.

		Participants	**Spherical Equivalent (diopters)**	**Age (years)**
		N (%)	Mean ± SD	Mean ± SD
All		34,998 (100)	-0.35 ± 2.56	66.54 ± 11.55
Sex	**Male**	14,431 (41.23)	-0.32 ± 2.46	68.84 ± 10.70
	Female	20,567 (58.77)	-0.38 ± 2.64	64.93 ± 11.84

Supplementary Note Table 5**.** Characteristics of GERA non-Hispanic white subjects included in the GWAS of spherical equivalent by sex. Abbreviations: N, number; SD, standard deviation; Age, age at first documented spherical equivalent assessment

4.3.1 Phenotyping

All participants underwent vision examinations, and most subjects had multiple measures for both eyes. Spherical equivalent was assessed as the sphere + cylinder/2. For this analysis, spherical equivalent was selected from the first documented assessment, and the mean of both eyes was used. As previously described[27], individuals with histories of cataract surgery (in either eye), refractive surgery, keratitis, or corneal diseases were excluded.

4.3.2 Genotyping

DNA samples from GERA individuals were extracted from Oragene kits (DNA Genotek Inc., Ottawa, ON, Canada) at KPNC and genotyped at the Genomics Core Facility of the University of California, San Francisco (UCSF). DNA samples were genotyped at over 665,000 single nucleotide polymorphisms (SNPs) on Affymetrix Axiom arrays (Affymetrix, Santa Clara, CA, USA)[28,29]. SNPs with initial genotyping call rate ≥97%, allele frequency difference ≤0.15 between males and females for autosomal markers, and genotype concordance rate >0.75 across duplicate samples were included[26]. Around 94% of samples and more than 98% of genetic markers assayed passed quality control (QC) procedures. In addition to those QC criteria,

SNPs with genotype call rates <90% were removed, as well as SNPs with a minor allele frequency (MAF) < 1%.
Following genotyping QC, we conducted statistical imputation of additional genetic variants. Following the pre-phasing of genotypes with Shape-IT v2.r72719[30], variants were imputed from the cosmopolitan 1000 Genomes Project reference panel (phase I integrated release; http://1000genomes.org) using IMPUTE2 v2.3.0.[31-33] As a QC metric, we used the info r^2 from IMPUTE2, which is an estimate of the correlation of the imputed genotype to the true genotype[34]. Variants with an imputation $r^2 < 0.3$ were excluded, and we restricted to SNPs that had a minor allele count (MAC) ≥ 20.

4.3.3 Association analyses

A linear regression of each individual's spherical equivalent was performed with the following covariates: age at first documented spherical equivalent assessment, sex, and genetic principal components. A linear regression of the residuals on each SNP was then performed using PLINK[35] v1.9 (www.cog-genomics.org/plink/1.9/) to assess genetic associations. Data from each SNP were modeled using additive dosages to account for the uncertainty of imputation[36]. Eigenstrat[37] v4.2 was used to calculate the PCs[25]. The top 10 ancestry PCs were included as covariates, as well as the percentage of Ashkenazi ancestry to adjust for genetic ancestry, as described previously[25].

4.4 Cream Consortium

4.4.1 Phenotyping

All participants included in this analysis from CREAM were 25 years of age or older. RE was represented by measurements of refraction and spherical equivalent (SphE = spherical refractive error +1/2 cylinder refractive error) was the outcome variable for CREAM. Participants with conditions that might alter refraction, such as cataract surgery, laser refractive procedures, retinal detachment surgery, keratoconus, or ocular or systemic syndromes were excluded from the analyses. Recruitment and ascertainment strategies varied by study and were previously published elsewhere[6].

4.4.2 Genotyping

The genotyping process has been described elsewhere[6]. Samples were genotyped on different platforms, and study-specific QC measures of the genotyped variants were implemented before association analysis. Genotypes were imputed with the appropriate ancestry-matched reference panel for all cohorts from the 1000 Genomes Project (Phase I version 3, March 2012 release) with either minimac[23] or IMPUTE[16]. The metrics for preimputation QC varied among studies, but genotype call-rate thresholds were set at a high level (≥0.95). These metrics were similar to those described in a previous GWAS analyses[38]; detailed information for each cohort is described elsewhere[6].

4.4.3 Association analyses

To prevent overlap of samples, cohorts from the United Kingdom (1985BBC, ALSPAC-Mothers, EPIC-Norfolk, ORCADES and Twins UK) were excluded from the GWAS meta-analysis. Association analyses were performed following the workflow elsewhere[6]: All samples analyzed were of European descent, for each CREAM cohort, a single-marker analysis for the phenotype of Spherical equivalent (in diopters) was carried out with linear regression with adjustment for age, sex and up to the first five principal components. For all non-family-based cohorts, one of each pair of relatives was removed (after detection through either GCTA or identity by sequence (IBS)/identity by descent (IBD) analysis). In

family-based cohorts, a score test-based association was used to adjust for within-family relatedness. We used an additive SNP allelic-effect model.

4.5 EPIC

The European Prospective Investigation into Cancer (EPIC) study is a pan-European prospective cohort study designed to investigate the etiology of major chronic diseases[39]. EPIC-Norfolk , one of the UK arms of EPIC, recruited and examined 25,639 participants between 1993 and 1997 for the baseline examination[40]. Recruitment was via general practices in the city of Norwich and the surrounding small towns and rural areas, and methods have been described in detail previously[41]. Since virtually all residents in the UK are registered with a general practitioner through the National Health Service, general practice lists serve as population registers. Ophthalmic assessment formed part of the third health examination and this has been termed the EPIC-Norfolk Eye Study[42]. In total, 8,623 participants were seen for the Eye Study, between 2004 and 2011. The EPIC-Norfolk Eye Study was carried out following the principles of the Declaration of Helsinki and the Research Governance Framework for Health and Social Care. The study was approved by the Norfolk Local Research Ethics Committee (05/Q0101/191) and East Norfolk & Waveney NHS Research Governance Committee (2005EC07L). All participants gave written, informed consent.

4.5.1 Phenotyping

Refractive error was measured in both eyes using a Humphrey Auto-Refractor 500 (Humphrey Instruments, San Leandro, California, USA). Spherical equivalent was calculated as spherical refractive error plus half the cylindrical error for each eye.
Some basic demographic and clinical information about the samples used for the validation analyses is given below (Supplementary Note Table 6).

		Participants	**Spherical Equivalent (diopters)**	**Age (years)**
		N (%)	Mean ± SD	Mean ± SD
All		7,117 (100)	+0.16 ± 2.25	68.80 ± 8.18
Sex	**Male**	3,253 (45.71)	+0.15 ± 2.23	69.60 ± 8.20
	Female	3,864 (54.29)	+0.18 ± 2.27	68.13 ± 8.11

Supplementary Note Table 6**.** Characteristics of the participants in the EPIC-Norfolk cohort, included in the heritability and risk prediction analyses. Abbreviations: N, number; SD, standard deviation; Age, age at first documented spherical equivalent assessment

4.5.2 Genotyping and imputation

Genotypes obtained using the Affymetrix UK Biobank Axiom Array on 7,117 subjects contributed to the current study were excluded if they had low call rates, poor clustering, batch effects across genotyping plates and/or Hardy-Weinberg equilibrium $P < 10^{-7}$. Samples were excluded on grounds of poor genotyping across all SNPs, sex discordance , exces or low heterozygosity and unexplainable identity-by-descent values. Third-degree relatives or closer participants were also removed. Data were pre-phased using SHAPEIT[14] version 2 and imputed to the Phase 3 build of the 1000 Genomes project[43] (October 2014) using IMPUTE[16] version 2.3.2.

4.5.3 Association analysis

We examined the relationship between allele dosage and mean spherical equivalent using linear regression adjusted for age, sex and the first 5 principal components. Analyses were carried out using SNPTEST version 2.5.1.

5 STATISTICAL ANALYSES

5.1 Meta-analyses

For all meta-analyses we applied a Z-score method, weighted by the effective population sample size, as implemented in METAL[44]. No genomic control adjustment was applied during the meta-analysis.

5.2 Conditional analyses

The conditional and joint analysis on summary data (COJO) [45] as implemented in the GCTA program [46] was used to identify independent effects within associated loci as well as the calculation of the phenotypic variance explained[47] by all polymorphisms associated with the trait after the conditional analyses. Default parameters were used for the analysis. The LD estimates were derived from a randomly selected sample of 10,000 unrelated subjects the UK Biobank cohort.

5.3 Multiple testing correction

Two methods of correcting for multiple testing were used. The first was a classic Bonferroni correction, in which the threshold of significance (0.05) was divided by the number of tests (n):

$$\alpha = \frac{0.05}{n}$$

Given the large number of loci for which replication was needed, we additionally calculated the False Discovery Rates, using the Benjamini-Hochberg method[48].

5.4 Genomic inflation

To assess the potential inflation of association probabilities, genomic inflation factors[49] were calculated and Q-Q plots were drawn using the package 'gap' in R (https://cran.r-project.org/).

5.5 LDscore regression-based methods

5.5.1 Polygenicity vs. inflation

To distinguish between the effect of polygenicity and those arising from sample stratification or uncontrolled population admixture, we followed previously suggested approaches[50] to calculate the LD score regression intercepts using the program LD Score (https://github.com/bulik/ldsc).

5.5.2 Calculation of genetic correlation

Bivariate genetic correlations between refractive error and other complex traits whose summary statistics are publicly available were assessed following previously described methodologies[51], using the program LD Score (https://github.com/bulik/ldsc).

5.6 Associated SNPs and gene annotations

Polymorphisms associated at a GWAS level ($P<5x10^{-08}$) were clustered within an "associated genomic region", defined as a contiguous genomic region where GWAS-significant markers were within 1 million base pairs from each other, as suggested elsewhere[52]. Significant polymorphisms were annotated with the gene inside whose transcript-coding region they are located, or alternatively, if located between two genes, with the gene nearest to it. The associated genomic regions were collectively annotated with the gene overlapping, or nearest the most significantly associated variant within that region. In addition, the polymorphic sites were functionally annotated using SNPnexus[53].

5.6.1 OMIM

The Online Mendelian Inheritance In Man (OMIM) is a continuously curated catalog of human genes and phenotypic changes their polymorphic forms cause in humans[54]. This catalogue contains a still partial, but highly reliable list of gene-phenotype pairs and was used retrieve data that could inform about the functionality of specific genes with particular focus on phenotypic expressions of extremely penetrant mutations.

5.6.2 The GWAS Catalog.

Previous GWAS association of SNPs or genes with other phenotypic traits was conducted through queries of the GWAS Catalog[55]. Results were downloaded from the official site hosted at the European Bioinformatics Institute: https://www.ebi.ac.uk/gwas/downloads.

5.7 Graphical illustration of association

LocusZoom[56] was used to generate plot that visualize regional association and its genomic context. Data from the European participants in the 1000 Genome Project, November 2014 was used, and the graphs were generated using the online LocusZoomserver (http://locuszoom.org/).

5.8 Mendelian randomization

The R (https://cran.r-project.org) package MendelianRandomization v3.4.4 was used for Mendelian randomization analyses.

5.9 Gene expression, GTEx and other transcription data

We obtained data on tissue expression from several sources for genes that map within RE associated loci defined as described before (section **Error! Reference source not found.**). Information about the expression of the genes of interest in systemic (i.e. non-ocular) tissues was obtained from the GTEx Portal for GTEx release v7 (https://gtexportal.org/home/datasets). RNA sequencing data was obtained

for both fetal and adult corneal, trabecular meshwork and ciliary body, as described elsewhere[57], which we downloaded from the authors' supplementary information. In addition, we extracted data from the subset of subjects with presumed healthy adult retinas (AMD=1), described elsewhere[58] that obtained from the GTEx Portal (https://gtexportal.org/home/datasets).

Transcription data was processed using different platforms and were available in different units (Transcripts per Million bases, TPM, for the retina and GTEx tissues, and Fragments per Kilobase, FPKM for the other tissues). For purposes of comparing expression across different tissues for which different methodologies may have been used, expression levels for all tissues were rank-transformed. Hierarchical clustering was used to help visualize similarities and differences of patterns of transcript expression across different tissues ('hclust' package in R).

5.10 LD score regression applied to specifically expressed genes (LDSC-SEG)

Disease-relevant tissues and cell types were identified by analyzing gene expression data together with summary statistics from the meta-analysis of refractive error in all five cohorts, as described elsewhere[59]. Briefly, genes were ranked based on the t-statistic of their expression in each tisue and the 10% most expressed genes for each tissue were considered "specifically expressed genes". A stratified LD score regression was applied to the meta-analysis summary statistics to evaluate the contribution of the focal genome annotation to trait heritability.

5.11 SMR

SMR (Summary data–based Mendelian randomization) assesses the relationship between genetic variant, intermediate variables such as gene expression levels or methylation levels as mediating traits, to test causality on a specific phenotype[60].

5.11.1 Test description

The SMR package helps perform two tests. The first is an SMR test, which correlates GWAS effects with eQTL or methylation effects (or any other intermediate trait)[60]. This test suggests causation, although it is unable to fully differentiate between it and pleiotropy. The second test is that of Heterogeneity in Dependent Instrument (HEIDI). This test against the null hypothesis that changes in both eQTL (or other intermediary traits) and the phenotype of interest are caused by one single SNP, which is therefore considered as the candidate for the putative causal effect.

5.11.2 Datasets for the SMR analyses: eQTL, *cis*-mQTL

To perform the above-mentioned tests of causation/pleiotropy, we used three different datasets of association between genetic variants and intermediate traits. The first was the summary statistics of eQTL associations in the untransformed peripheral blood samples of 5,311 subjects[61]. There were two advantages in using these data: 1) this was the largest eQTL dataset available and 2) the use of a highly heterogeneous tissue such as peripheral blood would be more likely than any other single more homogeneous tissue to overcome any heterogeneity of eQTL effects with eye and retinal tissues that were unavailable at the time of the analysis and manuscript writing.

Assuming that tissues relevant to the development of refractive error are similar to the brain, we also used two datasets, one with eQTL effects and the other with results of a *cis*- methylation analysis (*cis*-mQTL), both in brain tissues[62].

5.12 Gene-set enrichment

To identify pathways or other gene sets that were over-represented among our results, we used a Gene-Set Enrichment Analysis (GSEA) as implemented in the Meta-Analysis Gene Set Enrichment of Variant (MAGENTA) software[63]. This program assigns scores to each gene based on the strength of association with refractive error, adjusting for potential confounders such as gene length and linkage disequilibrium. Enrichment for any gene set was assessed within genes above the cut-off of the highest 75th centile of significant gene scores. For the current study, the most recent versions of Gene Ontology (GO), Panther, KGG, Biocarta and MSigDB databases were used. We also carried out a similar enrichment analysis for the presence of transcription factor binding sites. A permutational procedure and false-discovery rates were used to calculate significance of enrichment and control for multiple testing.

5.12.1 GSEA definitions

For the enrichment analyses we used updated versions of the GSEA gene sets as described before[64]. We used the versions from September 2017 which were downloaded from:
http://software.broadinstitute.org/gsea/login.jsp

5.13 Analyses of signals of natural selection

Results of three statistical tests for natural selection were imported from the 1000 Genomes Selection Browser[65]. We downloaded and reported results from several tests such as iHS[66] and a cross-population comparison, XP-EHH, based on extended haplotype homozygosity test (average and maximum CEU, CEU vs YRI) [67] and the Tajima's D test. The absolute test scores and the rank scores (-log10 of the centile of the absolute test score across the genome) were reported.

5.14 Estimation of effect size distributions for spherical equivalent

We used a maximum-likelihood model to estimate the distribution of effect sizes, based on summary statistics of observations and linkage disequilibrium patterns to predict the likely number of SNPs that explain spherical equivalent heritability as well as explore the relationship between future sample sizes and the number of SNPs identified and variance or heritability explained as described elsewhere[68] and implemented in the GENESIS R package (https://github.com/yandorazhang/GENESIS).

6 FULL ACKNOWLEDGEMENTS

6.1 UK Biobank

UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government, and Northwest Regional Development Agency. It also had funding from the Welsh Assembly Government, British Heart Foundation, and Diabetes UK.

Development of the eye and vison dataset was led by Prof Sir Peng T Khaw and Prof Paul Foster, funded by a grant from The National Institute for Health Research (NIHR) to Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology for a biomedical research centre.

6.2 23andMe

23andMe would like to thank the research participants and employees of 23andMe for making this work possible.

6.3 GERA

GERA are grateful to the Kaiser Permanente Northern California members who have generously agreed to participate in the Kaiser Permanente Research Program on Genes, Environment, and Health. Support for participant enrollment, survey completion, and biospecimen collection for the RPGEH was provided by the Robert Wood Johnson Foundation, the Wayne and Gladys Valley Foundation, the Ellison Medical Foundation, and Kaiser Permanente Community Benefit Programs. Genotyping of the GERA cohort was funded by a grant from the National Institute on Aging, National Institute of Mental Health, and National Institute of Health Common Fund (RC2 AG036607). Data analyses were facilitated by National Eye Institute (NEI) grant R01 EY027004 (E.J.), National Institute of Diabetes and Digestive and Kidney Diseases grant R01 DK116738 (E.J.).

6.4 Consortium for Refractive Error and Myopia (CREAM)

CREAM investigators are gratefully thank all study participants, their relatives and the staff at the recruitment centers for their invaluable contributions. Funding for this particular GWAS mega-analysis was provided by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (grant 648268), the Netherlands Organisation for Scientific Research (NWO, grant 91815655) and the National Eye Institute (grant R01EY020483). VJMV acknowledges funding from the Netherlands Organisation for Scientific Research (NWO, grant 91815655 & 91617076).

6.5 EPIC-Norfolk

EPIC-Norfolk infrastructure and core functions are supported by grants from the Medical Research Council (G1000143) and Cancer Research UK (C864/A14136). The clinic for the third health examination was funded by Research into Ageing (262). Genotyping was funded by the Medical Research Council (MC_PC_13048). We thank all staff from the MRC Epidemiology laboratory team for the preparation and quality control of DNA samples. Mr. Khawaja is supported by a Moorfields Eye Charity grant. Professor Foster has received additional support from the Richard Desmond Charitable Trust (via Fight for Sight) and the Department for Health through the award made by the National Institute for Health Research to

Moorfields Eye Hospital and the UCL Institute of Ophthalmology for a specialist Biomedical Research Centre for Ophthalmology.

6.6 King's College London authors

M.J. Simcoe is a recipient of a Fight for Sight PhD studentship. K. Patasova is a recipient of a Fight for Sight PhD studentship. P.G. Hysi the recipient of a FfS ECI fellowship. P.G. Hysi and C.J. Hammond acknowledge the TFC Frost Charitable Trust Support for the KCL Department of Ophthalmology. The statistical analyses were run in King's College London Rosalind HPC LINUX Clusters and cloud server. The UK Biobank data was accessed as part of the UK Biobank projects 669 and 17615.

6.7 UK Biobank Eye Consortium members

Data analyses were carried out using the RAVEN computing cluster, maintained by the ARCCA group. UK Biobank data was accessed as part of the UK Biobank Project 17615. Veronique Vitart is supported by core grant MC_UU_00007/10 from the UK Medical Research Council. The UK Biobank Eye and Vision Consortium has been supported by grants from NIHR (BRC3_026), Moorfields Eye Charity (ST 15 11 E), Fight for Sight (1507/1508), The Macular Society, The International Glaucoma Association (IGA, Ashford UK) and Alcon Research Institute.

6.8 University College London authors

Jugnoo Rahi is a National Institute for Health Research (NIHR) Senior Investigator, supported by the NIHR Biomedical Research Centres at Moorfields Eye Hospital/UCL Institute of Ophthalmology, and at the UCL Institute of Child Health/Great Ormond Street Hospital. Views expressed are those of the authors and not necessarily those of the NHS/NIHR. Phillippa Cumberland was funded by the Ulverscroft Foundation. Omar Mahroo is supported by Wellcome Trust grant 206619_Z_17_Z and the NIHR Biomedical Research Centre at Moorfields Eye Hospital and the UCL Institute of Ophthalmology. The UK Biobank data was accessed as part of the UK Biobank projects 669 and 17615.

References:

1. Cumberland, P.M. *et al.* Frequency and Distribution of Refractive Error in Adult Life: Methodology and Findings of the UK Biobank Study. *PLoS One* **10**, e0139780 (2015).
2. Williams, K.M. *et al.* Early life factors for myopia in the British Twins Early Development Study. *British Journal of Ophthalmology*, bjophthalmol-2018-312439 (2018).
3. Wojciechowski, R. & Hysi, P.G. Focusing in on the complex genetics of myopia. *PLoS Genet* **9**, e1003442 (2013).
4. Williams, K.M. *et al.* Age of myopia onset in a British population-based twin cohort. *Ophthalmic Physiol Opt* **33**, 339-45 (2013).
5. Kiefer, A.K. *et al.* Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet* **9**, e1003299 (2013).
6. Tedja, M.S. *et al.* Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nat Genet* **50**, 834-848 (2018).
7. Williams, K.M. *et al.* Increasing Prevalence of Myopia in Europe and the Impact of Education. *Ophthalmology* **122**, 1489-97 (2015).
8. Bycroft, C. *et al.* Genome-wide genetic data on~ 500,000 UK Biobank participants. *bioRxiv*, 166298 (2017).
9. http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/DNA-Extraction-at-UK-Biobank-October-2014.pdf.
10. http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UK-Biobank-Axiom-Array-Content-Summary-2014.pdf.
11. http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=155583.
12. https://biobank.ctsu.ox.ac.uk/crystal/docs/impute_ukb_v1.pdf.
13. Wain, L.V. *et al.* Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med* **3**, 769-81 (2015).
14. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat Methods* **9**, 179-81 (2011).
15. O'Connell, J. *et al.* Haplotype estimation for biobank-scale data sets. *Nat Genet* **48**, 817-20 (2016).
16. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
17. Loh, P.R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* **47**, 284-90 (2015).
18. Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A.P. & Price, A.L. Mixed-model association for biobank-scale datasets. *Nature genetics*, 1 (2018).
19. Ma, C., Blackwell, T., Boehnke, M., Scott, L.J. & Go, T.D.i. Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. *Genet Epidemiol* **37**, 539-50 (2013).
20. Henn, B.M. *et al.* Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS One* **7**, e34267 (2012).
21. Delaneau, O., Zagury, J.F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* **10**, 5-6 (2013).
22. Browning, S.R. & Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* **81**, 1084-97 (2007).

23. Fuchsberger, C., Abecasis, G.R. & Hinds, D.A. minimac2: faster genotype imputation. *Bioinformatics* **31**, 782-4 (2015).
24. Zheng, X. *et al.* HIBAG--HLA genotype imputation with attribute bagging. *Pharmacogenomics J* **14**, 192-200 (2014).
25. Banda, Y. *et al.* Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* **200**, 1285-95 (2015).
26. Kvale, M.N. *et al.* Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* **200**, 1051-60 (2015).
27. Shen, L. *et al.* The Association of Refractive Error with Glaucoma in a Multiethnic Population. *Ophthalmology* **123**, 92-101 (2016).
28. Hoffmann, T.J. *et al.* Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. *Genomics* **98**, 79-89 (2011).
29. Hoffmann, T.J. *et al.* Design and coverage of high throughput genotyping arrays optimized for individuals of East Asian, African American, and Latino race/ethnicity using imputation and a novel hybrid SNP selection algorithm. *Genomics* **98**, 422-30 (2011).
30. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nature methods* **9**, 179-81 (2012).
31. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature genetics* **44**, 955-9 (2012).
32. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3* **1**, 457-70 (2011).
33. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics* **5**, e1000529 (2009).
34. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nature reviews. Genetics* **11**, 499-511 (2010).
35. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
36. Huang, L., Wang, C. & Rosenberg, N.A. The relationship between imputation error and statistical power in genetic association studies in diverse populations. *American journal of human genetics* **85**, 692-8 (2009).
37. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38**, 904-9 (2006).
38. Verhoeven, V.J. *et al.* Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat Genet* **45**, 314-8 (2013).
39. Riboli, E. & Kaaks, R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26 Suppl 1**, S6-14 (1997).
40. Day, N. *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* **80 Suppl 1**, 95-103 (1999).
41. Hayat, S.A. *et al.* Cohort profile: A prospective cohort study of objective physical and cognitive capability and visual health in an ageing population of men and women in Norfolk (EPIC-Norfolk 3). *Int J Epidemiol* **43**, 1063-72 (2014).
42. Khawaja, A.P. *et al.* The EPIC-Norfolk Eye Study: rationale, methods and a cross-sectional analysis of visual impairment in a population-based cohort. *BMJ Open* **3**(2013).

43. Delaneau, O., Marchini, J., Genomes Project, C. & Genomes Project, C. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat Commun* **5**, 3934 (2014).
44. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
45. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369-75, S1-3 (2012).
46. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
47. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-9 (2010).
48. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289-300 (1995).
49. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
50. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
51. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
52. Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173-86 (2014).
53. Dayem Ullah, A.Z., Lemoine, N.R. & Chelala, C. SNPnexus: a web server for functional annotation of novel and publicly known genetic variants (2012 update). *Nucleic Acids Res* **40**, W65-70 (2012).
54. McKusick, V.A. Mendelian Inheritance in Man and its online version, OMIM. *Am J Hum Genet* **80**, 588-604 (2007).
55. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005-D1012 (2019).
56. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
57. Carnes, M.U., Allingham, R.R., Ashley-Koch, A. & Hauser, M.A. Transcriptome analysis of adult and fetal trabecular meshwork, cornea, and ciliary body tissues by RNA sequencing. *Exp Eye Res* **167**, 91-99 (2018).
58. Ratnapriya, R. *et al.* Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. *Nat Genet* **51**, 606-610 (2019).
59. Finucane, H.K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet* **50**, 621-629 (2018).
60. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* **48**, 481-7 (2016).
61. Westra, H.J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238-1243 (2013).
62. Qi, T. *et al.* Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat Commun* **9**, 2282 (2018).
63. Segre, A.V. *et al.* Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet* **6**(2010).
64. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-50 (2005).

65. Pybus, M. *et al.* 1000 Genomes Selection Browser 1.0: a genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res* **42**, D903-9 (2014).
66. Voight, B.F., Kudaravalli, S., Wen, X. & Pritchard, J.K. A map of recent positive selection in the human genome. *PLoS Biol* **4**, e72 (2006).
67. Sabeti, P.C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-8 (2007).
68. Zhang, Y., Qi, G., Park, J.H. & Chatterjee, N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nat Genet* **50**, 1318-1326 (2018).

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_1773395	*RDH5*	rs3138142	T	C	0.17	0.08	0.003	$5.7x10^{-174}$	-0.75	0.02	$8.3x10^{-253}$	-0.11	0.005	$5.0x10^{-104}$	$1.83\ x10^{-22}$
ILMN_1671237	*GNGT2*	rs7222840	T	C	0.31	0.04	0.003	$1.3x10^{-44}$	0.44	0.02	$7.4x10^{-109}$	0.08	0.007	$2.3x10^{-32}$	0.0003
ILMN_1779182	*TMEM98*	rs29009	G	A	0.14	-0.04	0.003	$2.7x10^{-26}$	0.52	0.03	$2.76x10^{-84}$	-0.07	0.007	$1.2x10^{-20}$	0.10
ILMN_1673682	*GATAD2A*	rs6909	G	A	0.35	-0.02	0.003	$3.1x10^{-20}$	0.58	0.02	$5.1x10^{-210}$	-0.04	0.005	$1.0x10^{-18}$	$1.5\ x10^{-07}$
ILMN_1757350	*CTNNB1*	rs11711946	C	T	0.48	-0.03	0.002	$1.0x10^{-26}$	-0.28	0.02	$2.4x10^{-47}$	0.09	0.011	$8.0x10^{-18}$	0.04
ILMN_1679917	*ARL16*	rs11150805	T	C	0.17	0.03	0.003	$3.0x10^{-19}$	-0.73	0.02	$7.4x10^{-201}$	-0.04	0.005	$8.1x10^{-18}$	$7.2\ x10^{-15}$
ILMN_1810474	*UBE2I*	rs710902	T	C	0.11	0.03	0.004	$6.4x10^{-18}$	0.84	0.03	$5.4x10^{-200}$	0.04	0.005	$1.1x10^{-16}$	0.01
ILMN_2367165	*ABTB1*	rs9859117	C	G	0.20	0.03	0.003	$3.4x10^{-19}$	0.50	0.02	$3.2x10^{-101}$	0.05	0.006	$1.5x10^{-16}$	0.002
ILMN_1802096	*ABTB1*	rs9859117	C	G	0.20	0.03	0.003	$3.4x10^{-19}$	0.43	0.02	$1.3x10^{-72}$	0.06	0.008	$1.1x10^{-15}$	0.01
ILMN_1754489	*FBXL20*	rs8076462	G	C	0.26	-0.03	0.003	$5.2x10^{-23}$	0.30	0.02	$3.14x10^{-42}$	-0.09	0.011	$1.3\ x10^{-15}$	0.5
ILMN_1776723	*PHF11*	rs7322077	A	G	0.46	0.02	0.002	$3.9x10^{-16}$	-0.68	0.02	0.00E+00	-0.03	0.004	$1.5\ x10^{-15}$	$2.9\ x10^{-11}$
ILMN_2336109	*L3MBTL2*	rs139486	G	C	0.37	-0.02	0.002	$5.5x10^{-17}$	-0.49	0.02	$1.4x10^{-149}$	0.04	0.005	$1.5\ x10^{-15}$	$1.5\ x10^{-09}$
ILMN_1823231	*HS.571502*	rs13438889	A	G	0.18	0.05	0.003	$4.7x10^{-60}$	0.23	0.03	$4.61x10^{-19}$	0.22	0.029	$4.8\ x10^{-15}$	$1.4x10^{-07}$
ILMN_2300695	*IKZF3*	rs2517955	C	T	0.32	-0.02	0.003	$1.0x10^{-17}$	-0.38	0.02	$5.17x10^{-82}$	0.06	0.007	$5.1\ x10^{-15}$	$1.7x10^{-06}$
ILMN_1727495	*L3MBTL3*	rs7740107	T	A	0.26	0.02	0.003	$4.6x10^{-15}$	0.59	0.02	$2.4x10^{-180}$	0.04	0.005	$4.1x10^{-14}$	0.005
ILMN_1662741	*EDG4*	rs880090	C	G	0.25	-0.02	0.003	$3.2x10^{-15}$	-0.53	0.02	$2.1x10^{-144}$	0.04	0.005	$5.0x10^{-14}$	$3.9\ x10^{-09}$
ILMN_1680673	*NT5DC1*	rs9387383	G	A	0.46	-0.02	0.002	$9.8x10^{-20}$	-0.25	0.02	$4.21x10^{-40}$	0.09	0.011	$6.5\ x10^{-14}$	$2.2\ x10^{-07}$
ILMN_1784428	*MGC57346*	rs4074462	T	G	0.23	0.02	0.003	$3.1x10^{-14}$	0.86	0.02	0.00E+00	0.03	0.003	$8.1\ x10^{-14}$	0.36
ILMN_1683127	*ZNF281*	rs2808519	G	T	0.28	0.02	0.003	$2.0x10^{-16}$	0.37	0.02	$2.53x10^{-67}$	0.06	0.008	$1.1\ x10^{-13}$	$1.9\ x10^{-06}$
ILMN_1696463	*SPI1*	rs1057233	G	A	0.32	0.02	0.003	$8.7x10^{-14}$	-0.82	0.02	0.00E+00	-0.02	0.003	$1.7\ x10^{-13}$	0.0002
ILMN_1699631	*GATS*	rs7811662	G	A	0.28	-0.02	0.003	$6.9x10^{-14}$	0.77	0.02	0.00E+00	-0.03	0.004	$1.8\ x10^{-13}$	0.05
ILMN_1781184	*MYBPC3*	rs7105851	T	C	0.32	0.02	0.003	$9.5x10^{-14}$	-0.64	0.02	$4.4x10^{-262}$	-0.03	0.004	$3.3\ x10^{-13}$	0.0004
ILMN_1724367	*NDUFB1*	rs3814833	G	A	0.31	0.02	0.003	$1.7x10^{-17}$	-0.27	0.02	$1.04x10^{-40}$	-0.08	0.011	$7.1\ x10^{-13}$	$1.3\ x10^{-07}$
ILMN_1743621	*C17ORF69*	rs17426195	A	G	0.23	0.02	0.003	$3.2x10^{-14}$	0.50	0.02	$5.7x10^{-105}$	0.05	0.006	$7.6\ x10^{-13}$	0.3
ILMN_2043615	*C17ORF90*	rs7219915	T	C	0.49	0.02	0.002	$8.8x10^{-14}$	-0.45	0.02	$2.7x10^{-130}$	-0.04	0.006	$1.0\ x10^{-12}$	$4.3x10^{-11}$
ILMN_1805344	*DDX5*	rs1991401	G	A	0.32	-0.02	0.003	$1.8x10^{-13}$	-0.54	0.02	$4.9x10^{-176}$	0.03	0.005	$1.0\ x10^{-12}$	NA

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_1652525	*FAM125B*	rs10122788	G	A	0.43	-0.02	0.002	$1.1x10^{-13}$	0.46	0.02	$9.9 x10^{-139}$	-0.04	0.005	$1.1 x10^{-12}$	0.02
ILMN_2390162	*PHF11*	rs7322886	T	C	0.37	0.03	0.002	$5.8x10^{-26}$	0.19	0.02	$6.66 x10^{-22}$	0.14	0.019	$1.2 x10^{-12}$	0.02
ILMN_2344956	*ACP1*	rs4455191	C	T	0.33	-0.02	0.003	$1.1 x10^{-17}$	0.25	0.02	$1.81 x10^{-35}$	-0.09	0.012	$1.8 x10^{-12}$	0.45
ILMN_1812926	*ANTXR2*	rs7699282	T	C	0.39	0.03	0.002	$9.0x10^{-27}$	-0.18	0.02	$2.99 x10^{-20}$	-0.14	0.021	$2.8 x10^{-12}$	$4.1 x10^{-05}$
ILMN_2226519	*CYP2D7P1*	rs2413667	C	A	0.21	-0.02	0.003	$1.3x10^{-15}$	-0.34	0.02	$4.56 x10^{-46}$	0.07	0.010	$3.1 x10^{-12}$	0.3
ILMN_1823112	*HS.377257*	rs12618537	G	C	0.08	-0.05	0.004	$8.8x10^{-26}$	0.33	0.04	$2.65 x10^{-20}$	-0.14	0.020	$4.1 x10^{-12}$	0.00
ILMN_2347193	*GSDMB*	rs9901146	A	G	0.51	0.02	0.002	$3.5x10^{-12}$	-0.63	0.02	$2.4 x10^{-292}$	-0.03	0.004	$8.4 x10^{-12}$	$7.1 x10^{-11}$
ILMN_1809883	*CCDC134*	rs9607821	T	C	0.45	-0.02	0.002	$3.3x10^{-18}$	-0.21	0.02	$1.48 x10^{-27}$	0.10	0.015	$1.1 x10^{-11}$	0.001
ILMN_2228180	*MSRA*	rs9329221	T	G	0.47	0.02	0.002	$8.3 x10^{-13}$	-0.39	0.02	$6.18 x10^{-94}$	-0.04	0.007	$1.4 x10^{-11}$	0.001
ILMN_2315979	*LBH*	rs7562727	G	A	0.48	-0.02	0.002	$2.8 x10^{-12}$	0.45	0.02	$2.3 x10^{-130}$	-0.04	0.006	$1.9 x10^{-11}$	$3.4 x10^{-08}$
ILMN_1794612	*UBA7*	rs1128535	T	C	0.50	0.02	0.002	$8.2 x10^{-12}$	-0.58	0.02	$4.2 x10^{-237}$	-0.03	0.004	$2.2 x10^{-11}$	0.02
ILMN_1784264	*TNFSF13*	rs6608	T	C	0.18	0.02	0.003	$2.6 x10^{-13}$	-0.40	0.03	$7.82 x10^{-58}$	-0.06	0.009	$2.8 x10^{-11}$	0.0003
ILMN_2259223	*TMTC4*	rs9582406	C	G	0.33	0.02	0.003	$3.6 x10^{-13}$	0.32	0.02	$7.13 x10^{-57}$	0.06	0.009	$3.8 x10^{-11}$	0.003
ILMN_1686135	*CCDC45*	rs1991401	G	A	0.32	-0.02	0.003	$1.8 x10^{-13}$	0.30	0.02	$7.8 x10^{-51}$	-0.06	0.009	$3.9 x10^{-11}$	NA
ILMN_1811650	*DUS2L*	rs6499163	T	G	0.19	-0.02	0.003	$1.4 x10^{-11}$	0.71	0.02	$1.9 x10^{-201}$	-0.03	0.005	$4.3 x10^{-11}$	0.002
ILMN_1738239	*RBM6*	rs2245365	A	C	0.47	-0.02	0.002	$3.4 x10^{-11}$	0.66	0.02	0.00E+00	-0.02	0.004	$6.5 x10^{-11}$	0.02
ILMN_1682572	*KIAA0528*	rs10842030	C	T	0.05	-0.04	0.005	$2.8 x10^{-13}$	0.61	0.04	$2.51 x10^{-47}$	-0.07	0.010	$7.1 x10^{-11}$	0.00001
ILMN_2183687	*LIME1*	rs6011066	G	A	0.34	-0.02	0.003	$2.5 x10^{-11}$	0.46	0.02	$2.5 x10^{-118}$	-0.04	0.006	$1.4 x10^{-10}$	0.009
ILMN_2284706	*PHF11*	rs7317982	T	C	0.46	0.02	0.002	$4.9 x10^{-16}$	-0.20	0.02	$1.98 x10^{-25}$	-0.10	0.015	$1.5 x10^{-10}$	0.0002
ILMN_1712929	*DNAJB12*	rs9415064	C	T	0.36	-0.02	0.002	$1.9 x10^{-15}$	0.21	0.02	$4.73 x10^{-27}$	-0.09	0.014	$1.6 x10^{-10}$	0.31
ILMN_1726769	*CNDP2*	rs8084058	A	G	0.40	0.02	0.002	$4.0 x10^{-11}$	-0.46	0.02	$5.7 x10^{-129}$	-0.04	0.006	$1.9 x10^{-10}$	0.001
ILMN_1811110	*TDRD9*	rs6576012	T	C	0.29	-0.02	0.003	$2.9 x10^{-11}$	0.45	0.02	$2.1 x10^{-107}$	-0.04	0.006	$1.9 x10^{-10}$	0.003
ILMN_2336609	*SYTL2*	rs290200	A	G	0.33	-0.02	0.003	$1.0 x10^{-10}$	-0.67	0.02	$8 x10^{-293}$	0.02	0.004	$2.0 x10^{-10}$	$5.9 x10^{-07}$
ILMN_1688240	*PHOSPHO1*	rs11650975	T	C	0.24	-0.02	0.003	$6.3 x10^{-12}$	0.36	0.02	$1.54 x10^{-59}$	-0.05	0.008	$2.4 x10^{-10}$	$2.5 x10^{-12}$
ILMN_1682929	*SYTL2*	rs290200	A	G	0.33	-0.02	0.003	$1.0 x10^{-10}$	-0.55	0.02	$1.2 x10^{-184}$	0.03	0.005	2.9E-10	$1.1 x10^{-06}$
ILMN_1653708	*CORO1B*	rs1476792	C	T	0.44	-0.02	0.002	$1.1 x10^{-11}$	-0.30	0.02	$1.14 x10^{-56}$	0.05	0.009	4.2E-10	0.6

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_2333865	*DNAJB12*	rs9415066	T	C	0.30	-0.02	0.002	4.6×10^{-17}	0.19	0.02	1.07×10^{-20}	-0.11	0.017	4.4E-10	0.6
ILMN_1678235	*KIAA1267*	rs4630591	T	C	0.22	0.02	0.003	4.1×10^{-13}	-0.29	0.02	3.0×10^{-34}	-0.08	0.012	4.5×10^{-10}	0.01
ILMN_2094952	*NUAK2*	rs7526603	T	G	0.39	0.02	0.002	7.4×10^{-12}	0.28	0.02	3.8×10^{-47}	0.06	0.010	6.1×10^{-10}	0.02
ILMN_1657347	*PODXL2*	rs9859117	C	G	0.20	0.03	0.003	3.4×10^{-19}	0.21	0.02	1.3×10^{-17}	0.13	0.021	6.4×10^{-10}	$2.3\text{x} \times 10^{-05}$
ILMN_2377019	*CORO1B*	rs1476792	C	T	0.44	-0.02	0.002	1.1×10^{-11}	-0.28	0.02	1.9×10^{-49}	0.06	0.009	6.7×10^{-10}	0.1
ILMN_1786759	*C11ORF10*	rs174538	A	G	0.32	-0.02	0.003	5.4×10^{-11}	0.36	0.02	$2.5 \times 10^{-6}7$	-0.05	0.008	8.5×10^{-10}	0.01
ILMN_1698218	*TRAF1*	rs2416804	G	C	0.44	-0.02	0.002	7.1×10^{-11}	-0.34	0.02	3.83×10^{-73}	0.05	0.007	8.7×10^{-10}	0.03
ILMN_1765851	*TRADD*	rs1053612	G	A	0.08	-0.03	0.005	2.7×10^{-10}	-0.87	0.03	2.3×10^{-137}	0.03	0.005	9.4×10^{-10}	0.3
ILMN_1741994	*L3MBTL3*	rs7740107	T	A	0.26	0.02	0.003	4.6×10^{-15}	0.21	0.02	6.21×10^{-22}	0.10	0.017	1.2×10^{-09}	0.04
ILMN_1862217	*HS.532698*	rs8040632	G	C	0.06	0.03	0.004	9.2×10^{-11}	-0.63	0.04	7.84×10^{-67}	-0.05	0.008	1.3×10^{-09}	0.0004
ILMN_2234229	*PRMT6*	rs1623927	T	C	0.34	0.02	0.003	6.0×10^{-10}	-0.57	0.02	5.3×10^{-202}	-0.03	0.004	1.3×10^{-09}	0.2
ILMN_2061435	*MEG3*	rs7141210	T	C	0.31	0.02	0.003	3.8×10^{-10}	-0.42	0.02	1.14×10^{-96}	-0.04	0.006	2.0×10^{-09}	0.2
ILMN_1798014	*EIF2S2*	rs2268079	A	G	0.41	0.02	0.002	1.6×10^{-11}	-0.25	0.02	2.34×10^{-37}	-0.07	0.011	2.5×10^{-09}	2.1E-06
ILMN_1668924	*BEGAIN*	rs4074037	G	A	0.40	0.01	0.002	1.7×10^{-09}	-0.51	0.02	4.3×10^{-163}	-0.03	0.005	4.0×10^{-09}	0.2
ILMN_2364022	*SLC16A3*	rs9747201	A	C	0.32	-0.02	0.003	4.9×10^{-10}	0.36	0.02	$5.42\text{x}10^{-72}$	-0.04	0.007	$4.1\text{x}10^{-09}$	$4.5\text{x}10^{-07}$
ILMN_1807540	*CBARA1*	rs12357282	A	G	0.34	-0.02	0.002	3.3×10^{-14}	0.19	0.02	1.84×10^{-20}	-0.10	0.017	4.3×10^{-09}	0.02
ILMN_1778668	*TAGLN*	rs10790177	A	G	0.23	-0.02	0.003	3.4×10^{-09}	-0.93	0.02	0.00E+00	0.02	0.003	4.5×10^{-09}	0.1
ILMN_1813834	*PRMT6*	rs1730858	T	C	0.34	0.02	0.002	6.6×10^{-10}	-0.36	0.02	1.02×10^{-74}	-0.04	0.007	4.9×10^{-09}	0.4
ILMN_1791912	*SIDT2*	rs4938353	G	A	0.18	-0.02	0.003	5.8×10^{-09}	-1.09	0.02	0.00E+00	0.02	0.003	7.1×10^{-09}	0.02
ILMN_1683096	*ASB1*	rs483166	A	G	0.43	-0.01	0.002	4.9×10^{-10}	-0.30	0.02	$1.05\text{x}10^{-54}$	0.05	0.009	7.5×10^{-09}	0.0003
ILMN_1814022	*NR1H3*	rs3758673	T	C	0.30	-0.02	0.003	8.2×10^{-10}	-0.34	0.02	2.06×10^{-59}	0.05	0.008	9.3×10^{-09}	0.02
ILMN_2120965	*NPAT*	rs623860	C	T	0.44	-0.02	0.002	6.7×10^{-18}	-0.15	0.02	1.65×10^{-14}	0.14	0.024	9.9×10^{-09}	0.1
ILMN_1742544	*MEF2C*	rs661311	T	C	0.22	-0.02	0.003	4.2×10^{-10}	-0.34	0.02	1.03×10^{-46}	0.05	0.009	1.0×10^{-08}	0.01
ILMN_1664978	*TJP2*	rs2309424	G	A	0.41	0.01	0.002	1.9×10^{-09}	-0.36	0.02	1.16×10^{-79}	-0.04	0.007	1.1×10^{-08}	$6.2\text{x}10^{-07}$
ILMN_1829845	*HS.553301*	rs2148568	T	C	0.40	-0.01	0.002	7.2×10^{-09}	0.67	0.02	0.00E+00	-0.02	0.004	1.1×10^{-08}	0.01
ILMN_1799011	*CEP250*	rs10359	G	A	0.23	-0.02	0.003	1.1×10^{-14}	-0.19	0.02	1.91×10^{-17}	0.11	0.020	1.1×10^{-08}	0.01

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_1699545	*PCSK7*	rs7107152	A	G	0.21	-0.02	0.003	2.4×10^{-09}	0.44	0.02	4.04×10^{-85}	-0.04	0.007	1.1×10^{-08}	0.1
ILMN_1763972	*TJP2*	rs2993820	C	A	0.50	-0.02	0.002	3.2×10^{-12}	0.19	0.02	6.96×10^{-23}	-0.09	0.015	1.3×10^{-08}	0.0003
ILMN_1690921	*STAT2*	rs808919	G	C	0.07	0.03	0.005	2.6×10^{-10}	-0.48	0.04	2.61×10^{-38}	-0.06	0.011	1.4×10^{-08}	0.04
ILMN_1711823	*C17ORF70*	rs7224579	G	A	0.34	0.03	0.003	4.8×10^{-26}	-0.14	0.02	1.68×10^{-11}	-0.20	0.036	1.4×10^{-08}	0.0004
ILMN_1714384	*PCCA*	rs2254579	C	G	0.42	0.03	0.002	1.7×10^{-25}	0.13	0.02	1.37×10^{-11}	0.19	0.034	1.4×10^{-08}	0.2
ILMN_2048607	*ANKRD9*	rs1044502	A	G	0.31	0.01	0.003	9.2×10^{-09}	-0.54	0.02	1.7×10^{-169}	-0.03	0.005	1.9×10^{-08}	0.02
ILMN_2134224	*ATP13A1*	rs2304130	G	A	0.09	-0.02	0.004	1.5×10^{-08}	1.17	0.03	0.00E+00	-0.02	0.004	2.1×10^{-08}	5.1×10^{-11}
ILMN_1715994	*HGS*	rs11150805	T	C	0.17	0.03	0.003	3.0×10^{-19}	0.18	0.03	1.51×10^{-12}	0.16	0.028	2.8×10^{-08}	0.04
ILMN_2413318	*C15ORF57*	rs4923876	A	G	0.41	-0.01	0.002	4.9×10^{-09}	0.34	0.02	1.4×10^{-69}	-0.04	0.008	2.8×10^{-08}	0.5
ILMN_1807600	*NPLOC4*	rs9913546	A	C	0.22	0.02	0.003	1.1×10^{-11}	0.23	0.02	6.15×10^{-22}	0.09	0.015	2.9×10^{-08}	9.4×10^{-07}
ILMN_1725071	*CCDC12*	rs11130111	C	T	0.43	-0.02	0.002	4.6×10^{-12}	-0.18	0.02	2.53×10^{-20}	0.09	0.017	3.1×10^{-08}	0.8
ILMN_2359742	*CTSB*	rs7013950	C	T	0.16	-0.02	0.003	1.8×10^{-08}	0.75	0.02	1.2×10^{-210}	-0.02	0.004	3.1×10^{-08}	1.9×10^{-07}
ILMN_1764320	*C8ORF58*	rs11782130	T	G	0.33	-0.02	0.003	6.8×10^{-17}	0.15	0.02	1.46×10^{-13}	-0.14	0.025	3.1×10^{-08}	0.10
ILMN_1680353	*NSF*	rs183211	A	G	0.24	0.02	0.003	2.4×10^{-11}	0.22	0.02	4.45×10^{-22}	0.08	0.015	4.0×10^{-08}	0.02
ILMN_1677376	*CHD7*	rs4237038	G	A	0.25	0.02	0.003	2.2×10^{-11}	-0.21	0.02	2.09×10^{-21}	-0.09	0.016	4.4×10^{-08}	0.01
ILMN_2304495	*PPP1R1B*	rs879606	A	G	0.16	-0.02	0.003	2.0×10^{-12}	0.22	0.03	8×10^{-18}	-0.10	0.019	5.2×10^{-08}	0.001
ILMN_1694731	*CLCN7*	rs12923519	T	C	0.11	0.02	0.004	2.4×10^{-08}	-0.72	0.03	3.2×10^{-130}	-0.03	0.006	5.3×10^{-08}	0.03
ILMN_1658399	*KLRG1*	rs7311267	C	T	0.34	0.01	0.002	2.1×10^{-08}	-0.41	0.02	8.47×10^{-98}	-0.03	0.006	6.2×10^{-08}	1.1×10^{-08}
ILMN_1812474	*TFG*	rs591728	C	T	0.36	-0.01	0.003	4.8×10^{-08}	0.59	0.02	2.9×10^{-219}	-0.02	0.004	7.5×10^{-08}	0.001
ILMN_2073202	*MEI1*	rs6002406	T	A	0.19	0.02	0.003	4.1×10^{-09}	0.32	0.02	1.85×10^{-40}	0.06	0.010	7.5×10^{-08}	0.0003
ILMN_1707448	*CRKRS*	rs7219814	T	C	0.10	-0.02	0.004	3.6×10^{-08}	-0.69	0.03	2.1×10^{-111}	0.03	0.006	8.8×10^{-08}	4.1×10^{-09}
ILMN_2400935	*TAGLN*	rs1871757	T	A	0.18	-0.02	0.003	1.0×10^{-08}	-0.38	0.03	1.41×10^{-50}	0.05	0.009	9.0×10^{-08}	0.1
ILMN_2384561	*TJP2*	rs10869976	T	C	0.41	0.01	0.002	1.8×10^{-09}	-0.22	0.02	1.74×10^{-30}	-0.07	0.012	1.0×10^{-07}	0.0003
ILMN_1663195	*MCM7*	rs11771331	G	A	0.34	-0.02	0.003	8.1×10^{-11}	0.19	0.02	1.54×10^{-20}	-0.09	0.016	1.0×10^{-10}	0.0002
ILMN_2394250	*PLEKHA1*	rs11200629	G	A	0.48	-0.01	0.002	1.9×10^{-08}	-0.31	0.02	1.7×10^{-59}	0.04	0.008	1.1×10^{-07}	0.01
ILMN_2042343	*MRPL42P5*	rs896795	T	G	0.41	-0.01	0.002	4.3×10^{-09}	0.23	0.02	8.94×10^{-32}	-0.06	0.012	1.5×10^{-07}	0.02

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_1789793	*NUAK2*	rs7526603	T	G	0.39	0.02	0.002	7.4×10^{-12}	0.16	0.02	5.85×10^{-16}	0.10	0.020	1.7×10^{-07}	0.06
ILMN_1721225	*C20ORF4*	rs4142032	C	T	0.06	-0.03	0.005	5.1×10^{-09}	-0.43	0.04	5.22×10^{-31}	0.06	0.012	1.8^{-07}	0.06
ILMN_1674038	*CTSD*	rs7943617	A	G	0.23	0.02	0.003	1.0 x10-10	-0.20	0.02	1.15×10^{-18}	-0.09	0.017	1.8×10^{-07}	0.02
ILMN_1675048	*IGF1R*	rs2654980	T	C	0.23	0.02	0.003	3.7×10^{-08}	-0.36	0.02	2.24×10^{-58}	-0.04	0.008	1.9×10^{-07}	0.07
ILMN_2186877	*FLJ10213*	rs2231928	A	G	0.49	0.01	0.002	9.2×10^{-09}	-0.24	0.02	5.05×10^{-35}	-0.06	0.011	1.9×10^{-07}	0.004
ILMN_1725244	*HAT1*	rs6433321	G	T	0.44	-0.02	0.002	6.8×10^{-12}	0.16	0.02	1.53×10^{-15}	-0.11	0.020	2.0×10^{-07}	2.7×10^{-10}
ILMN_1804789	*KIAA1967*	rs4592028	C	T	0.32	-0.02	0.003	4.0×10^{-17}	-0.13	0.02	3.91×10^{-11}	0.16	0.030	2.0×10^{-07}	0.4
ILMN_2392043	*SPI1*	rs326217	C	T	0.29	0.01	0.003	2.1×10^{-08}	0.28	0.02	1.73×10^{-43}	0.05	0.010	2.1×10^{-07}	2.8×10^{-06}
ILMN_2311278	*ADD3*	rs7071509	A	G	0.14	-0.02	0.003	9.8×10^{-09}	0.33	0.03	3.19×10^{-33}	-0.06	0.012	2.3×10^{-07}	0.1
ILMN_1690454	*C3ORF54*	rs695238	C	A	0.50	0.02	0.002	1.1×10^{-11}	-0.15	0.02	1.56×10^{-15}	-0.11	0.020	2.3×10^{-07}	0.3
ILMN_1662839	*PLEKHA1*	rs11200594	T	C	0.48	-0.01	0.002	2.3×10^{-08}	-0.26	0.02	1.3×10^{-40}	0.05	0.010	2.5×10^{-07}	0.06
ILMN_1741131	*CHRNB1*	rs2302761	T	C	0.21	0.02	0.003	8.33×10^{-09}	0.26	0.02	1.513×10^{-28}	0.06	0.013	3.2×10^{-07}	0.0002
ILMN_2244009	*LBH*	rs6755863	A	G	0.35	-0.02	0.003	6.23×10^{-10}	0.18	0.02	1.723×10^{-19}	-0.09	0.017	3.3×10^{-07}	0.001
ILMN_1781098	*DNAJB12*	rs9415064	C	T	0.36	-0.02	0.002	1.93×10^{-15}	0.13	0.02	3.93×10^{-11}	-0.15	0.029	3.8×10^{-07}	0.3
ILMN_2139052	*ZKSCAN1*	rs4729561	T	C	0.34	-0.02	0.003	1.53×10^{-11}	-0.16	0.02	1.61×10^{-14}	0.11	0.022	4.0×10^{-07}	0.002
ILMN_2200503	*NIT2*	rs13071905	T	G	0.31	0.02	0.003	1.93×10^{-09}	-0.19	0.02	1.17×10^{-20}	-0.08	0.016	4.4×10^{-07}	0.02
ILMN_1717594	*DKFZP761E198*	rs537786	C	T	0.46	-0.02	0.002	1.03×10^{-10}	-0.15	0.02	2.87×10^{-15}	0.10	0.020	5.7×10^{-07}	0.003
ILMN_1723522	*APOLD1*	rs34322	T	C	0.48	0.01	0.002	4.03×10^{-09}	-0.18	0.02	9.54×10^{-21}	-0.08	0.016	6.4×10^{-07}	0.03
ILMN_1732151	*COL6A1*	rs2839029	G	A	0.44	-0.02	0.002	1.43×10^{-14}	-0.12	0.02	3.04×10^{-10}	0.15	0.031	1.1×10^{-06}	3.4×10^{-06}
ILMN_2067370	*SNRPF*	rs17289597	T	C	0.06	0.02	0.004	3.53×10^{-08}	0.36	0.03	2.22×10^{-25}	0.06	0.013	1.1×10^{-06}	0.02
ILMN_1709809	*NHP2L1*	rs132765	A	G	0.19	0.02	0.003	5.63×10^{-09}	0.22	0.02	7.5×10^{-19}	0.08	0.017	1.1×10^{-06}	0.001
ILMN_1794074	*MXI1*	rs1318388	T	C	0.17	0.02	0.003	1.63×10^{-09}	-0.20	0.02	2.61×10^{-16}	-0.09	0.019	1.2×10^{-06}	0.1
ILMN_1787378	*ADD3*	rs7069128	G	A	0.14	-0.02	0.003	3.03×10^{-08}	0.27	0.03	8.52×10^{-23}	-0.07	0.014	1.4×10^{-06}	0.03
ILMN_1706511	*TEF*	rs9611565	C	T	0.25	0.02	0.003	4.53×10^{-08}	0.23	0.02	1.31×10^{-23}	0.07	0.014	1.6×10^{-06}	0.001
ILMN_1810550	*BAIAP3*	rs8063	A	G	0.29	0.02	0.003	2.9×10^{-13}	0.13	0.02	4.84×10^{-10}	0.16	0.033	2.2×10^{-06}	0.1
ILMN_1684549	*RNPC2*	rs6120993	T	G	0.13	-0.03	0.003	1.6×10^{-20}	0.16	0.03	3.75×10^{-08}	-0.21	0.044	2.2×10^{-06}	0.5

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_1806456	*C14ORF45*	rs1005888	T	C	0.13	0.02	0.003	$4.3x10^{-08}$	0.24	0.03	$5 x10^{-18}$	0.08	0.017	$3.7 x10^{-06}$	0.01
ILMN_1815102	*LCAT*	rs255049	C	T	0.20	-0.02	0.003	$2.5 x10^{-14}$	0.14	0.02	$6.24 x10^{-09}$	-0.16	0.035	$3.8 x10^{-06}$	0.003
ILMN_2286783	*LRRC37A4*	rs1635298	T	A	0.24	0.02	0.003	$4.0x10^{-12}$	-0.14	0.02	$6.1 x10^{-10}$	-0.14	0.030	$3.9 x10^{-06}$	0.1
ILMN_1797519	*DLK1*	rs12882815	C	G	0.31	0.02	0.003	$4.4 x10^{-10}$	-0.14	0.02	$6.37 x10^{-12}$	-0.11	0.024	$3.9 x10^{-06}$	0.7
ILMN_1675937	*ANKRD9*	rs1044502	A	G	0.31	0.01	0.003	$9.2 x10^{-09}$	-0.16	0.02	$8.62 x10^{-15}$	-0.09	0.020	$3.9 x10^{-06}$	0.0
ILMN_1729915	*PILRA*	rs2405442	T	C	0.31	-0.02	0.003	$1.0 x10^{-12}$	-0.13	0.02	$1.65 x10^{-09}$	0.14	0.031	$4.2 x10^{-06}$	0.5
ILMN_1713249	*PHF19*	rs10760129	T	C	0.44	-0.02	0.002	$4.3 x10^{-11}$	-0.12	0.02	$2.21 x10^{-10}$	0.13	0.028	$4.8 x10^{-06}$	0.01
ILMN_1724762	*XKR6*	rs4618656	A	T	0.38	-0.01	0.002	$1.5 x10^{-08}$	0.15	0.02	$2.17 x10^{-14}$	-0.09	0.020	$5.4 x10^{-06}$	0.004
ILMN_1814526	*ADD3*	rs17126936	T	C	0.13	-0.02	0.003	$1.6 x10^{-08}$	0.20	0.03	$6.54 x10^{-14}$	-0.10	0.021	$6.3 x10^{-06}$	0.1
ILMN_1790114	*LOC474170*	rs9915547	C	T	0.23	0.02	0.003	$2.4 x10^{-14}$	0.13	0.02	$2.16 x10^{-08}$	0.17	0.038	$6.4 x10^{-06}$	0.01
ILMN_1676254	*MTMR3*	rs2074204	T	C	0.25	-0.02	0.003	$7.4 x10^{-09}$	0.16	0.02	$1.29 x10^{-12}$	-0.10	0.022	$7.4 x10_{-06}$	0.1
ILMN_1745256	*CXXC5*	rs682164	G	A	0.32	-0.02	0.003	$4.9 x10^{-10}$	-0.13	0.02	$1.22 x10^{-10}$	0.13	0.028	$7.7 x10^{-06}$	0.2
ILMN_2103841	*AIP*	rs917570	G	C	0.43	-0.02	0.003	$1.8 x10^{-10}$	-0.12	0.02	$3.61 x10^{-10}$	0.13	0.030	$7.8 x10^{-06}$	0.5
ILMN_1686516	*CUGBP1*	rs12419692	A	C	0.35	0.02	0.003	$2.1 x10^{-12}$	0.12	0.02	$7.03 x10^{-09}$	0.15	0.034	$7.8 x10^{-06}$	0.0003
ILMN_1662198	*RANGAP1*	rs139519	G	C	0.43	-0.02	0.002	$5.3 x10^{-13}$	0.11	0.02	$1.4 x10^{-08}$	-0.16	0.035	$8.2 x10^{-06}$	0.0017
ILMN_2344079	*ZGPAT*	rs6062509	G	T	0.34	-0.02	0.003	$2.2 x10^{-11}$	-0.12	0.02	$2.23 x10^{-09}$	0.14	0.031	$8.2 x10^{-06}$	0.00001
ILMN_1665235	*CRTAP*	rs4075293	G	T	0.24	0.02	0.003	$2.0 x10^{-10}$	-0.14	0.02	$4.57 x10^{-10}$	-0.13	0.028	$8.5 x10^{-06}$	0.0
ILMN_2123402	*TMEM4*	rs2306693	C	T	0.07	0.03	0.005	$3.9 x10^{-10}$	0.23	0.04	$3.65 x10^{-10}$	0.13	0.029	$9.5 x10^{-06}$	0.7
ILMN_1652521	*MTMR9*	rs4410870	G	C	0.39	-0.01	0.002	$9.1 x10^{-10}$	-0.12	0.02	$2.19 x10^{-10}$	0.12	0.027	$1.0 x10^{-05}$	0.1
ILMN_2089752	*LOC285016*	rs366330	C	A	0.48	0.02	0.002	$1.2 x10^{-11}$	0.11	0.02	$6.74 x10^{-09}$	0.14	0.033	$1.1 x10^{-05}$	0.0
ILMN_1707342	*LRIG1*	rs2306272	C	T	0.28	-0.01	0.003	$1.8 x10^{-08}$	0.15	0.02	$8.96 x10^{-12}$	-0.10	0.023	$1.4 x10^{-05}$	0.6
ILMN_2344182	*TMTC4*	rs1625099	G	T	0.30	0.02	0.003	$2.5 x10^{-10}$	0.13	0.02	$2.77 x10^{-09}$	0.13	0.030	$1.5 x10^{-05}$	0.3
ILMN_2141941	*TOR1AIP1*	rs480377	T	A	0.37	0.01	0.002	$3.2 x10^{-09}$	-0.13	0.02	$2.61 x10^{-10}$	-0.12	0.027	$1.5 x10^{-05}$	0.01
ILMN_1792538	*CD7*	rs11658061	T	C	0.24	0.02	0.003	1.9 x10-10	0.13	0.02	$1.03 x10^{-08}$	0.14	0.032	$2.1 x10^{-05}$	0.01
ILMN_2388272	*MED24*	rs6503525	C	G	0.45	-0.01	0.002	$3.6 x10^{-09}$	-0.12	0.02	$1.26 x10^{-09}$	0.12	0.028	$2.3 x10^{-05}$	0.01
ILMN_1703412	*LATS2*	rs8001991	T	C	0.27	0.02	0.003	$9.1 x10^{-09}$	0.14	0.02	$4.59 x10^{-10}$	0.11	0.027	$2.4 x10^{-05}$	0.01

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ILMN_1715968	*MLL4*	rs2293688	G	C	0.33	-0.02	0.003	2.3×10^{-09}	0.11	0.02	1.77×10^{-08}	-0.14	0.034	4.2×10^{-05}	0.1
ILMN_1657405	*SPPL3*	rs3213566	T	C	0.47	-0.01	0.002	6.1×10^{-09}	0.11	0.02	1.52×10^{-08}	-0.13	0.031	5.0×10^{-05}	0.1
ILMN_1782385	*POLR2A*	rs9889368	A	G	0.25	0.02	0.003	4.8×10^{-08}	-0.13	0.02	1.62×10^{-08}	-0.12	0.030	8.7×10^{-05}	0.01

Table 21 SMR results using the GTEx-brain summary eQTL data. The field "Probe ID" lists the ID for the eQTL probe. The column "Gene" is the gene that the probe eQTL effects act on, "topSNP" contains the SNP associated with the largest effects on the trait (AFSW); "A1", "A2" and "Freq" are the position (Build 37), effect allele, another allele, and effect allele frequency for the "topSNP", respectively. The fields "b_GWAS", "se_GWAS " and "p_GWAS " display the effect size, standard error, and p-value for the "topSNP". Columns "b_eQTL", "se_eQTL" and " p_eQTL" include the effect size, standard error, and p-value for the "Probe ID", "b_SMR", "se_SMR", and "p_SMR" are the effect size, standard error, and p-value results from the SMR analysis. The fields " p_HEIDI " include the p-value from the HEIDI (Heterogeneity In Dependent Instruments) test.

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ENSG00000163634.7	*THOC7*	rs832187	T	C	0.64	0.01	0.002	$8.38x10^{-09}$	0.50	0.08	$5.4x10^{-10}$	0.03	0.007	$2.42x10^{-05}$	0.99
ENSG00000260051.1	*LA16c-390E6.4*	rs17135338	C	T	0.11	0.02	0.004	$1.62x10^{-08}$	-0.58	0.10	$3.7x10^{-08}$	-0.04	0.010	$8.08x10^{-05}$	0.98
ENSG00000121454.4	*LHX4*	rs7536561	G	A	0.52	0.01	0.002	$2.21x10^{-08}$	-0.52	0.07	$3.6x10^{-14}$	-0.03	0.006	$6.78x10^{-06}$	0.92
ENSG00000148700.9	*ADD3*	rs6584962	A	G	0.15	-0.02	0.003	$1.73x10^{-10}$	0.46	0.08	$8.9x10^{-09}$	-0.05	0.011	$1.93x10^{-05}$	0.85
ENSG00000227888.3	*FAM66A*	rs6601649	T	C	0.57	0.01	0.002	$4.88x10^{-10}$	0.80	0.10	$8.1x10^{-16}$	0.02	0.004	$8.48x10^{-07}$	0.84
ENSG00000213402.2	*PTPRCAP*	rs1638556	C	T	0.44	-0.02	0.002	$9.14x10^{-12}$	-0.88	0.08	$7.3x10^{-30}$	0.02	0.003	$5.05x10^{-09}$	0.84
ENSG00000198890.6	*PRMT6*	rs1623927	C	T	0.66	-0.02	0.003	$5.98x10^{-10}$	0.40	0.07	$1.5x10^{-08}$	-0.04	0.009	$2.93x10^{-05}$	0.81
ENSG00000218016.2	*ZNF192P2*	rs1225618	C	A	0.39	-0.03	0.002	$1.68x10^{-33}$	-0.69	0.09	$1.1x10^{-13}$	0.04	0.007	$2.56x10^{-10}$	0.81
ENSG00000204961.5	*PCDHA9*	rs155814	C	T	0.34	-0.02	0.003	$1.98x10^{-12}$	0.79	0.11	$4.4x10^{-12}$	-0.02	0.005	$8.00x10^{-07}$	0.80
ENSG00000120322.2	*PCDHB8*	rs10516156	A	C	0.24	-0.02	0.003	$7.38x10^{-10}$	0.62	0.10	$7.0x10^{-10}$	-0.03	0.007	$1.32x10^{-05}$	0.76
ENSG00000264070.1	*DND1P1*	rs17689882	A	G	0.23	0.02	0.003	$2.86x10^{-14}$	0.66	0.11	$1.1x10^{-09}$	0.03	0.007	$1.95x10^{-06}$	0.75
ENSG00000262539.1	*RP11-259G18.3*	rs199533	A	G	0.22	0.02	0.003	$2.05x10^{-14}$	1.34	0.11	$2.2x10^{-34}$	0.02	0.003	$8.95x10^{-11}$	0.73
ENSG00000148143.8	*ZNF462*	rs7049113	G	A	0.14	-0.03	0.003	$9.85x10^{-14}$	0.72	0.11	$1.4x10^{-10}$	-0.04	0.007	$1.16x10^{-06}$	0.69
ENSG00000086504.11	*MRPL28*	rs3743888	G	C	0.61	0.02	0.003	$3.16x10^{-09}$	-0.46	0.07	$4.5x10^{-10}$	-0.03	0.008	$1.75x10^{-05}$	0.69
ENSG00000214826.4	*DDX12P*	rs1549429	C	T	0.48	-0.02	0.002	$1.32x10^{-20}$	-0.94	0.08	$3.4x10^{-29}$	0.02	0.003	$7.92x10^{-13}$	0.66
ENSG00000197062.7	*RP5-874C20.3*	rs1531681	A	G	0.58	0.03	0.002	$1.23x10^{-31}$	0.70	0.08	$7.0x10^{-17}$	0.04	0.006	$1.08x10^{-11}$	0.66
ENSG00000253893.2	*FAM85B*	rs2980439	A	G	0.49	0.01	0.002	$1.90x10^{-09}$	-0.65	0.10	$4.9x10^{-10}$	-0.02	0.005	$1.55x10^{-05}$	0.65
ENSG00000241186.3	*TDGF1*	rs3806702	C	T	0.25	-0.02	0.003	$1.43x10^{-08}$	0.97	0.10	$1.2x10^{-22}$	-0.02	0.003	$9.29x10^{-07}$	0.63
ENSG00000165934.8	*CPSF2*	rs2025070	C	T	0.15	0.03	0.003	$8.46x10^{-24}$	-0.77	0.10	$7.2x10^{-15}$	-0.04	0.007	$7.54x10^{-10}$	0.60
ENSG00000186868.11	*MAPT*	rs1876831	T	C	0.23	0.02	0.003	$1.26x10^{-14}$	-0.68	0.10	$2.4x10^{-11}$	-0.03	0.006	$4.44x10^{-07}$	0.53
ENSG00000131484.3	*RP11-798G7.5*	rs1876831	T	C	0.23	0.02	0.003	$1.26x10^{-14}$	-0.66	0.11	$2.0x10^{-09}$	-0.03	0.007	$2.20x10^{-06}$	0.50
ENSG00000214425.2	*LRRC37A4P*	rs17426174	C	G	0.23	0.02	0.003	$2.63x10^{-14}$	-1.27	0.11	$2.8x10^{-33}$	-0.02	0.003	$1.26x10^{-10}$	0.47
ENSG00000169738.3	*DCXR*	rs9903371	G	C	0.11	-0.02	0.004	$3.16x10^{-08}$	0.74	0.10	$2.4x10^{-12}$	-0.03	0.007	$1.41x10^{-05}$	0.43
ENSG00000111788.9	*RP11-22B23.1*	rs1549429	C	T	0.48	-0.02	0.002	$1.32x10^{-20}$	-1.04	0.08	$3.5x10^{-38}$	0.02	0.003	$4.29x10^{-14}$	0.43
ENSG00000249244.1	*RP11-548H18.2*	rs1546505	G	A	0.31	0.02	0.003	$4.00x10^{-09}$	0.68	0.09	$4.1x10^{-14}$	0.02	0.005	$3.44x10^{-06}$	0.41
ENSG00000255556.2	*RP11-351I21.6*	rs11250186	G	A	0.57	0.01	0.002	$6.51x10^{-09}$	0.68	0.10	$2.7x10^{-11}$	0.02	0.005	$1.21x10^{-05}$	0.40

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ENSG00000255098.1	*RP11-481A20.11*	rs11250186	G	A	0.57	0.01	0.002	$6.51x10^{-09}$	0.59	0.10	$6.4x10^{-09}$	0.02	0.006	$4.05x10^{-05}$	0.40
ENSG00000214401.4	*KANSL1-AS1*	rs199533	A	G	0.22	0.02	0.003	$2.05x10^{-14}$	1.24	0.12	$2.3x10^{-26}$	0.02	0.003	$5.43x10^{-10}$	0.38
ENSG00000182477.5	*OR2B8P*	rs4713135	A	G	0.24	-0.03	0.003	$9.75x10^{-28}$	0.78	0.11	$5.1x10^{-12}$	-0.04	0.007	$5.39x10^{-09}$	0.37
ENSG00000172493.16	*AFF1*	rs10489044	G	A	0.20	0.02	0.003	$1.25x10^{-10}$	-0.56	0.08	$2.3x10^{-11}$	-0.03	0.007	$3.57x10^{-06}$	0.34
ENSG00000262500.1	*RP11-259G18.2*	rs199528	T	C	0.21	0.02	0.003	$5.28x10^{-15}$	1.22	0.11	$2.3x10^{-26}$	0.02	0.003	$3.02x10^{-10}$	0.33
ENSG00000261575.2	*RP11-259G18.1*	rs17689882	A	G	0.23	0.02	0.003	$2.86x10^{-14}$	0.83	0.11	$7.8x10^{-15}$	0.03	0.005	$5.47x10^{-08}$	0.29
ENSG00000173295.3	*FAM86B3P*	rs2945232	C	T	0.49	0.01	0.002	$4.45x10^{-10}$	-0.49	0.07	$5.1x10^{-14}$	-0.03	0.006	$1.56x10^{-06}$	0.27
ENSG00000185298.8	*CCDC137*	rs11150805	T	C	0.17	0.03	0.003	$3.02x10^{-19}$	0.56	0.10	$4.5x10^{-08}$	0.05	0.011	$3.00x10^{-06}$	0.25
ENSG00000263503.1	*RP11-707O23.5*	rs1876831	T	C	0.23	0.02	0.003	$1.26x10^{-14}$	1.17	0.11	$1.1x10^{-26}$	0.02	0.003	$4.00x10^{-10}$	0.21
ENSG00000254507.2	*RP11-481A20.10*	rs11250186	G	A	0.57	0.01	0.002	$6.51x10^{-09}$	0.71	0.10	$2.6x10^{-13}$	0.02	0.004	$5.47x10^{-06}$	0.21
ENSG00000239389.3	*PCDHA13*	rs251379	A	G	0.65	0.02	0.003	$8.76x10^{-12}$	-0.86	0.09	$6.8x10^{-20}$	-0.02	0.004	$4.57x10^{-08}$	0.20
ENSG00000226314.3	*ZNF192P1*	rs4713145	C	T	0.24	-0.03	0.003	$1.56x10^{-27}$	0.53	0.09	$4.6x10^{-10}$	-0.06	0.010	$6.39x10^{-08}$	0.19
ENSG00000056558.6	*TRAF1*	rs10733648	C	T	0.63	0.01	0.002	$9.06x10^{-09}$	0.76	0.10	$1.4x10^{-15}$	0.02	0.004	$3.10x10^{-06}$	0.19
ENSG00000254532.1	*RP11-624D11.2*	rs1222210	T	C	0.22	0.03	0.003	$1.82x10^{-20}$	0.78	0.12	$5.7x10^{-11}$	0.03	0.006	$8.74x10^{-08}$	0.19
ENSG00000204650.9	*CRHR1-IT1*	rs1876831	T	C	0.23	0.02	0.003	$1.26x10^{-14}$	0.91	0.12	$2.8x10^{-14}$	0.02	0.004	$6.13x10^{-08}$	0.18
ENSG00000066923.13	*STAG3*	rs2272345	G	C	0.28	-0.02	0.003	$1.06x10^{-13}$	0.51	0.09	$3.6x10^{-09}$	-0.04	0.009	$3.79x10^{-06}$	0.18
ENSG00000147439.7	*BIN3*	rs11776549	T	C	0.33	-0.02	0.003	$6.40x10^{-17}$	0.50	0.09	$1.2x10^{-08}$	-0.04	0.009	$2.47x10^{-06}$	0.13
ENSG00000164078.8	*MST1R*	rs2280405	C	G	0.46	-0.01	0.002	$2.45x10^{-08}$	-0.62	0.10	$2.4x10^{-10}$	0.02	0.005	$2.84x10^{-05}$	0.13
ENSG00000128891.11	*C15orf57*	rs17673703	A	G	0.44	0.02	0.002	$4.95x10^{-11}$	-0.76	0.08	$1.0x10^{-23}$	-0.02	0.004	$3.82x10^{-08}$	0.12
ENSG00000004534.10	*RBM6*	rs2352974	T	C	0.50	0.02	0.002	$4.38x10^{-13}$	-0.93	0.08	$3.1x10^{-32}$	-0.02	0.003	$6.58x10^{-10}$	0.11
ENSG00000114021.7	*NIT2*	rs2289506	T	C	0.33	0.02	0.003	$1.11x10^{-10}$	-0.75	0.09	$2.2x10^{-16}$	-0.02	0.004	$3.91x10^{-07}$	0.10
ENSG00000267280.1	*RP11-332H18.4*	rs2240736	T	C	0.75	-0.02	0.003	$1.92x10^{-12}$	-0.46	0.07	$3.0x10^{-10}$	0.04	0.009	$2.69x10^{-06}$	0.09
ENSG00000242353.1	*RP4-710M3.1*	rs594937	G	A	0.78	-0.02	0.003	$2.50x10^{-15}$	0.55	0.10	$1.6x10^{-08}$	-0.04	0.009	$4.22x10^{-06}$	0.08
ENSG00000206535.3	*LNP1*	rs4928050	A	G	0.41	0.01	0.002	$2.16x10^{-08}$	-0.68	0.09	$2.1x10^{-14}$	-0.02	0.004	$6.29x10^{-06}$	0.07
ENSG00000238083.3	*LRRC37A2*	rs17763086	G	T	0.23	0.02	0.003	$1.08x10^{-14}$	1.23	0.09	$3.8x10^{-40}$	0.02	0.003	$2.41x10^{-11}$	0.07
ENSG00000011590.9	*ZBTB32*	rs231592	G	A	0.46	-0.01	0.003	$3.17x10^{-09}$	0.55	0.07	$2.7x10^{-15}$	-0.03	0.006	$2.14x10^{-06}$	0.06

Probe ID	Gene	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ENSG00000225190.4	*PLEKHM1*	rs1876831	T	C	0.23	0.02	0.003	$1.26x10^{-14}$	-0.41	0.07	$2.0x10^{-08}$	-0.05	0.012	$5.71x10^{-06}$	0.05
ENSG00000219392.1	*RP1-265C24.5*	rs2281588	A	G	0.24	-0.03	0.003	$8.58x10^{-27}$	-0.77	0.10	$1.9x10^{-14}$	0.04	0.006	$4.70x10^{-10}$	0.05
ENSG00000245958.2	*RP11-33B1.1*	rs3733520	G	C	0.30	0.02	0.003	$7.29x10^{-10}$	-0.91	0.10	$1.1x10^{-20}$	-0.02	0.003	$2.75x10^{-07}$	0.04
ENSG00000268340.1	*RP3-477O4.16*	rs2236160	G	A	0.23	-0.02	0.003	$2.67x10^{-14}$	0.71	0.10	$3.6x10^{-13}$	-0.03	0.006	$1.45x10^{-07}$	0.04
ENSG00000256427.1	*RP11-118B22.4*	rs34371	G	A	0.63	-0.01	0.002	$4.52x10^{-08}$	-0.60	0.09	$1.3x10^{-11}$	0.02	0.005	$2.09x10^{-05}$	0.04
ENSG00000100413.12	*POLR3H*	rs17367716	G	A	0.20	0.02	0.003	$5.34x10^{-09}$	0.72	0.12	$6.4x10^{-10}$	0.02	0.006	$2.21x10^{-05}$	0.03
ENSG00000215347.3	*SLC25A5P1*	rs1062753	A	G	0.29	-0.02	0.003	$8.57x10^{-11}$	-0.46	0.08	$1.6x10^{-09}$	0.04	0.008	$9.86x10^{-06}$	0.03
ENSG00000255753.1	*RP11-22B23.2*	rs6487747	C	T	0.48	-0.02	0.002	$9.55x10^{-21}$	-0.84	0.11	$7.8x10^{-14}$	0.03	0.005	$5.37x10^{-09}$	0.03
ENSG00000106392.6	*C1GALT1*	rs11771259	G	C	0.11	-0.02	0.004	$3.26x10^{-10}$	1.12	0.13	$3.2x10^{-17}$	-0.02	0.004	$4.63x10^{-07}$	0.03
ENSG00000187987.5	*ZSCAN23*	rs916403	G	A	0.48	-0.02	0.002	$2.57x10^{-22}$	0.51	0.07	$1.5x10^{-12}$	-0.05	0.008	$1.06x10^{-08}$	0.03
ENSG00000156414.14	*TDRD9*	rs1187448	T	C	0.29	-0.02	0.003	$5.73x10^{-11}$	0.57	0.10	$4.3x10^{-09}$	-0.03	0.007	$1.23x10^{-05}$	0.03
ENSG00000237037.5	*NDUFA6-AS1*	rs5758610	T	C	0.21	-0.02	0.003	$1.06x10^{-13}$	0.39	0.06	$1.2x10^{-09}$	-0.05	0.012	$2.53x10^{-06}$	0.03
ENSG00000016391.6	*CHDH*	rs17641133	A	T	0.31	0.02	0.003	$1.01x10^{-12}$	-0.48	0.08	$1.4x10^{-08}$	-0.04	0.009	$8.90x10^{-06}$	0.03
ENSG00000237412.2	*PRSS56*	rs733603	G	A	0.81	0.06	0.003	$3.76x10^{-81}$	-0.57	0.09	$1.6x10^{-10}$	-0.11	0.018	$1.31x10^{-09}$	0.02
ENSG00000259536.1	*RP11-111A22.1*	rs4924472	G	T	0.57	0.01	0.002	$1.79x10^{-08}$	-0.41	0.06	$5.7x10^{-11}$	-0.03	0.008	$1.95x10^{-05}$	0.02
ENSG00000164068.11	*RNF123*	rs2271960	C	T	0.50	0.02	0.002	$9.67x10^{-12}$	0.65	0.08	$1.0x10^{-16}$	0.02	0.005	$1.40x10^{-07}$	0.02
ENSG00000162873.10	*KLHDC8A*	rs7516786	A	C	0.40	-0.02	0.002	$6.16x10^{-10}$	-0.41	0.06	$1.7x10^{-11}$	0.04	0.008	$5.28x10^{-06}$	0.01
ENSG00000260423.1	*RP13-735L24.1*	rs2113899	C	T	0.48	-0.02	0.002	$2.15x10^{-20}$	-0.52	0.09	$7.0x10^{-09}$	0.04	0.009	$9.12x10^{-07}$	0.01
ENSG00000197599.8	*CCDC154*	rs4984834	G	A	0.86	-0.02	0.004	$1.44x10^{-08}$	-1.17	0.12	$4.7x10^{-21}$	0.02	0.004	$1.19x10^{-06}$	0.01
ENSG00000135930.9	*EIF4E2*	rs7571154	T	C	0.29	0.02	0.003	$1.64x10^{-16}$	-0.49	0.09	$4.3x10^{-08}$	-0.04	0.010	$5.05x10^{-06}$	0.01
ENSG00000176681.10	*LRRC37A*	rs17763086	G	T	0.23	0.02	0.003	$1.08x10^{-14}$	1.02	0.13	$1.2x10^{-14}$	0.02	0.004	$4.72x10^{-08}$	0.01
ENSG00000183066.10	*WBP2NL*	rs1062753	A	G	0.29	-0.02	0.003	$8.57x10^{-11}$	-0.67	0.08	$2.3x10^{-16}$	0.03	0.005	$3.57x10^{-07}$	0.01
ENSG00000156026.10	*MCU*	rs7075993	G	C	0.36	-0.02	0.002	$8.09x10^{-13}$	0.25	0.04	$2.1x10^{-08}$	-0.07	0.016	$1.03x10^{-05}$	0.01
ENSG00000214087.4	*ARL16*	rs11150805	T	C	0.17	0.03	0.003	$3.02x10^{-19}$	-1.18	0.11	$1.2x10^{-27}$	-0.02	0.004	$4.37x10^{-12}$	0.01
ENSG00000216901.1	*AL022393.7*	rs2531806	A	C	0.49	-0.02	0.002	$9.26x10^{-21}$	1.03	0.08	$1.6x10^{-35}$	-0.02	0.003	$7.93x10^{-14}$	0.004
ENSG00000271904.1	*CTC-498M16.4*	rs6884773	G	A	0.24	-0.03	0.003	$6.47x10^{-23}$	0.86	0.08	$2.7x10^{-28}$	-0.03	0.004	$1.99x10^{-13}$	0.004

Probe ID	*Gene*	topSNP	A1	A2	Freq	b_GWAS	se_GWAS	p_GWAS	b_eQTL	se_eQTL	p_eQTL	b_SMR	se_SMR	p_SMR	p_HEIDI
ENSG00000179296.9	*CTGLF12P*	rs769009	T	G	0.65	-0.02	0.002	$2.86x10^{-10}$	0.58	0.10	$6.8x10^{-09}$	-0.03	0.006	$1.98x10^{-05}$	0.004
ENSG00000170324.15	*FRMPD2*	rs769009	T	G	0.65	-0.02	0.002	$2.86x10^{-10}$	-0.38	0.06	$3.2x10^{-09}$	0.04	0.009	$1.59x10^{-05}$	0.003
ENSG00000126267.4	*COX6B1*	rs231592	G	A	0.46	-0.01	0.003	$3.17x10^{-09}$	-0.56	0.07	$7.2x10^{-14}$	0.03	0.006	$3.41x10^{-06}$	0.002
ENSG00000204622.6	*HLA-J*	rs2524005	A	G	0.22	-0.02	0.003	$5.16x10^{-09}$	-0.77	0.11	$1.9x10^{-11}$	0.02	0.005	$1.05x10^{-05}$	0.002
ENSG00000186470.9	*BTN3A2*	rs9358932	T	C	0.14	-0.03	0.004	$5.87x10^{-20}$	-1.47	0.17	$1.0x10^{-18}$	0.02	0.003	$2.10x10^{-10}$	0.001
ENSG00000100395.10	*L3MBTL2*	rs2038209	C	T	0.63	0.02	0.002	$7.37x10^{-17}$	0.42	0.07	$5.3x10^{-10}$	0.05	0.010	$6.34x10^{-07}$	0.001
ENSG00000126001.11	*CEP250*	rs10359	G	A	0.23	-0.02	0.003	$1.13x10^{-14}$	-1.02	0.08	$4.4x10^{-34}$	0.02	0.003	$6.98x10^{-11}$	0.001
ENSG00000172346.10	*CSDC2*	rs1006407	C	T	0.19	0.02	0.003	$1.46x10^{-08}$	0.57	0.08	$4.5x10^{-13}$	0.03	0.007	$8.11x10^{-06}$	0.001
ENSG00000224497.1	*RPL36P4*	rs2277862	T	C	0.15	-0.02	0.003	$8.15x10^{-09}$	1.09	0.12	$1.4x10^{-18}$	-0.02	0.004	$1.42x10^{-06}$	0.001
ENSG00000216915.2	*RP1-97D16.1*	rs2179096	T	C	0.34	-0.02	0.003	$4.99x10^{-12}$	-0.45	0.08	$4.1x10^{-09}$	0.04	0.009	$7.57x10^{-06}$	0.0004
ENSG00000257303.1	*RP11-977G19.11*	rs2066818	A	C	0.07	0.03	0.005	$8.89x10^{-10}$	-1.30	0.16	$1.2x10^{-16}$	-0.02	0.005	$8.36x10^{-07}$	0.0002
ENSG00000109919.5	*MTCH2*	rs10838699	C	T	0.68	-0.02	0.003	$1.39x10^{-13}$	-0.44	0.06	$6.1x10^{-12}$	0.04	0.008	$4.72x10^{-07}$	0.0001
ENSG00000181315.6	*ZNF322*	rs3800303	G	A	0.42	0.02	0.002	$2.24x10^{-10}$	-0.52	0.08	$5.1x10^{-11}$	-0.03	0.006	$5.05x10^{-06}$	$1.7x10^{-05}$
ENSG00000057019.11	*DCBLD2*	rs828617	T	G	0.39	-0.01	0.002	$1.92x10^{-08}$	-0.44	0.07	$3.6x10^{-10}$	0.03	0.007	$2.86x10^{-05}$	$1.3x10^{-05}$
ENSG00000148604.9	*RGR*	rs10736319	C	T	0.48	-0.04	0.002	$4.31x10^{-59}$	0.76	0.08	$1.6x10^{-22}$	-0.05	0.006	$6.07x10^{-17}$	$1.5x10^{-06}$
ENSG00000163870.10	*TPRA1*	rs4974443	A	G	0.11	-0.03	0.004	$6.86x10^{-13}$	-0.64	0.10	$1.5x10^{-10}$	0.04	0.009	$1.75x10^{-06}$	$5.1x10^{-08}$
ENSG00000176387.6	*HSD11B2*	rs12449157	G	A	0.16	-0.02	0.003	$1.90x10^{-12}$	-0.44	0.05	$1.5x10^{-20}$	0.05	0.010	$1.98x10^{-08}$	$4.4x10^{-08}$
ENSG00000137338.4	*PGBD1*	rs2531804	G	A	0.44	-0.02	0.002	$3.32x10^{-17}$	0.78	0.01	0.0E+00	-0.03	0.003	$4.44x10^{-17}$	$3.8x10^{-08}$

Table 22 SMR results using summary eQTL data from the peripheral blood (Westra *et al.*). The field "Probe ID" lists the ID for the eQTL probe. The column "Gene" is the gene that the probe eQTL effects act on, "topSNP" contains the SNP associated with the largest effects on the trait (AFSW); "A1", "A2" and "Freq" are the position (Build 37), effect allele, another allele, and effect allele frequency for the "topSNP " respectively. The fields "b_GWAS", "se_GWAS " and "p_GWAS " display the effect size, standard error, and p-value for the "topSNP". Columns "b_eQTL", "se_eQTL" and " p_eQTL" include the effect size, standard error, and p-value for the "Probe ID", "b_SMR", "se_SMR", and "p_SMR" are the effect size, standard error, and p-value results from the SMR analysis. The fields " p_HEIDI " include the p-value from the HEIDI (Heterogeneity In Dependent Instruments) test.

Chapter 10 | References

References

1. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, et al. Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. Ophthalmology. 2016 May;123(5):1036–42.

2. Mohlin C, Sandholm K, Ekdahl KN, Nilsson B. The link between morphology and complement in ocular disease. Mol Immunol. 2017 Sep 1;89:84–99.

3. Kolb H. Gross Anatomy of the Eye. In: Kolb H, Fernandez E, Nelson R, editors. Webvision: The Organization of the Retina and Visual System [Internet]. Salt Lake City (UT): University of Utah Health Sciences Center; 1995 [cited 2021 Nov 5]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK11534/

4. Kolb H. Simple Anatomy of the Retina. In: Kolb H, Fernandez E, Nelson R, editors. Webvision: The Organization of the Retina and Visual System [Internet]. Salt Lake City (UT): University of Utah Health Sciences Center; 1995 [cited 2021 Nov 5]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK11533/

5. Wojciechowski R. Nature and nurture: the complex genetics of myopia and refractive error. Clin Genet. 2011 Apr;79(4):301–20.

6. Flitcroft DI, He M, Jonas JB, Jong M, Naidoo K, Ohno-Matsui K, et al. IMI – Defining and Classifying Myopia: A Proposed Set of Standards for Clinical and Epidemiologic Studies. Invest Ophthalmol Vis Sci. 2019 Feb 28;60(3):M20–30.

7. Morgan IG, Ohno-Matsui K, Saw SM. Myopia. The Lancet. 2012 May 5;379(9827):1739–48.

8. Espandar L, Meyer J. Keratoconus: Overview and Update on Treatment. Middle East Afr J Ophthalmol. 2010;17(1):15–20.

9. Shiels A, Hejtmancik JF. Mutations and mechanisms in congenital and age-related cataracts. Exp Eye Res. 2017 Mar 1;156:95–102.

10. Baird PN, Schäche M, Dirani M. The GEnes in Myopia (GEM) study in understanding the aetiology of refractive errors. Prog Retin Eye Res. 2010 Nov 1;29(6):520–42.

11. The Prevalence of Refractive Errors Among Adults in the United States,Western Europe, and Australia. Arch Ophthalmol. 2004 Apr 1;122(4):495.

12. Sanfilippo PG, Chu BS, Bigault O, Kearns LS, Boon MY, Young TL, et al. What is the appropriate age cut-off for cycloplegia in refraction? Acta Ophthalmol (Copenh). 2014;92(6):e458–62.

13. Flitcroft DI. Is myopia a failure of homeostasis? Exp Eye Res. 2013 Sep 1;114:16–24.

14. Wojciechowski R. Nature and Nurture: the complex genetics of myopia and refractive error. Clin Genet. 2011 Apr;79(4):301–20.

15. Siegwart JT, Norton TT. Binocular lens treatment in tree shrews: Effect of age and comparison of plus lens wear with recovery from minus lens-induced myopia. Exp Eye Res. 2010 Nov;91(5):660–9.

16. Cook RC, Glasscock RE. Refractive and Ocular Findings in the Newborn*. Am J Ophthalmol. 1951 Oct 1;34(10):1407–13.

17. Larsen JS. The sagittal growth of the eye. IV. Ultrasonic measurement of the axial length of the eye from birth to puberty. Acta Ophthalmol (Copenh). 1971;49(6):873–86.

18. Barishak YR. Embryology of the eye and its adnexae. Dev Ophthalmol. 1992;24:1–142.

19. Refractive and ocular findings in the newborn. - PubMed - NCBI [Internet]. [cited 2020 Mar 31]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/14877971

20. Zadnik K, Mutti DO, Friedman NE, Adams AJ. Initial Cross-Sectional Results from the Orinda Longitudinal Study of Myopia. Optom Vis Sci. 1993 Sep;70(9):750–8.

21. Hung GK, Mahadas K, Mohammad F. Eye growth and myopia development: Unifying theory and Matlab model. Comput Biol Med. 2016 Mar 1;70:106–18.

22. Iribarren R, Hashemi H, Khabazkhoob M, Morgan IG, Emamian MH, Shariati M, et al. Hyperopia and Lens Power in an Adult Population: The Shahroud Eye Study. J Ophthalmic Vis Res. 2015;10(4):400–7.

23. Tay SA, Farzavandi S, Tan D. Interventions to Reduce Myopia Progression in Children. Strabismus. 2017 Jan 2;25(1):23–32.

24. Ramamurthy D, Chua SYL, Saw SM. A review of environmental risk factors for myopia during early life, childhood and adolescence. Clin Exp Optom. 2015;98(6):497–506.

25. Samarawickrama C, Mitchell P, Tong L, Gazzard G, Lim L, Wong TY, et al. Myopia-related optic disc and retinal changes in adolescent children from singapore. Ophthalmology. 2011 Oct;118(10):2050–7.

26. Mäntyjärvi MI. Predicting of myopia progression in school children. J Pediatr Ophthalmol Strabismus. 1985 Apr;22(2):71–5.

27. Zadnik K, Sinnott LT, Cotter SA, Jones-Jordan LA, Kleinstein RN, Manny RE, et al. Prediction of Juvenile-Onset Myopia. JAMA Ophthalmol. 2015 Jun 1;133(6):683–9.

28. Williams KM, Hysi PG, Nag A, Yonova-Doing E, Venturini C, Hammond CJ. Age of myopia onset in a British population-based twin cohort. Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom. 2013 May;33(3):339–45.

29. Kiefer AK, Tung JY, Do CB, Hinds DA, Mountain JL, Francke U, et al. Genome-Wide Analysis Points to Roles for Extracellular Matrix Remodeling, the Visual Cycle, and Neuronal Development in Myopia. PLOS Genet. 2013 Feb 28;9(2):e1003299.

30. Hysi PG, Young TL, Mackey DA, Andrew T, Fernández-Medarde A, Solouki AM, et al. A genome-wide association study for myopia and refractive error identifies a susceptibility locus at 15q25. Nat Genet. 2010 Oct;42(10):902–5.

31. Verhoeven VJM, Hysi PG, Wojciechowski R, Fan Q, Guggenheim JA, Höhn R, et al. Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. Nat Genet. 2013 Mar;45(3):314–8.

32. Grzybowski A, Kanclerz P, Tsubota K, Lanca C, Saw SM. A review on the epidemiology of myopia in school children worldwide. BMC Ophthalmol. 2020 Jan 14;20(1):27.

33. Haarman AEG, Enthoven CA, Tideman JWL, Tedja MS, Verhoeven VJM, Klaver CCW. The Complications of Myopia: A Review and Meta-Analysis. Invest Ophthalmol Vis Sci. 2020 Apr 9;61(4):49.

34. Morgan IG, French AN, Ashby RS, Guo X, Ding X, He M, et al. The epidemics of myopia: Aetiology and prevention. Prog Retin Eye Res. 2018;62:134–49.

35. Rudnicka AR, Kapetanakis VV, Wathern AK, Logan NS, Gilmartin B, Whincup PH, et al. Global variations and time trends in the prevalence of childhood myopia, a systematic review and quantitative meta-analysis: implications for aetiology and early prevention. Br J Ophthalmol. 2016 Jul 1;100(7):882–90.

36. Ding BY, Shih YF, Lin LLK, Hsiao CK, Wang IJ. Myopia among schoolchildren in East Asia and Singapore. Surv Ophthalmol. 2017 Sep 1;62(5):677–97.

37. Ojaimi E, Rose KA, Morgan IG, Smith W, Martin FJ, Kifley A, et al. Distribution of Ocular Biometric Parameters and Refraction in a Population-Based Study of Australian Children. Invest Ophthalmol Vis Sci. 2005 Aug 1;46(8):2748–54.

38. Ip JM, Huynh SC, Robaei D, Kifley A, Rose KA, Morgan IG, et al. Ethnic differences in refraction and ocular biometry in a population-based sample of 11–15-year-old Australian children. Eye. 2008 May;22(5):649–56.

39. Wu PC, Huang HM, Yu HJ, Fang PC, Chen CT. Epidemiology of Myopia. Asia-Pac J Ophthalmol. 2016 Dec;5(6):386–93.

40. Kleinstein RN. Refractive Error and Ethnicity in Children. Arch Ophthalmol. 2003 Aug 1;121(8):1141.

41. Carter MJ, Lansingh VC, Schacht G, Río del Amo M, Scalamogna M, France TD. Visual acuity and refraction by age for children of three different ethnic groups in Paraguay. Arq Bras Oftalmol. 2013 Apr;76(2):94–7.

42. O'Donoghue L, McClelland JF, Logan NS, Rudnicka AR, Owen CG, Saunders KJ. Refractive error and visual impairment in school children in Northern Ireland. Br J Ophthalmol. 2010 Sep;94(9):1155–9.

43. Villarreal MG, Ohlsson J, Abrahamsson M, Sjöström A, Sjöstrand J. Myopisation: The refractive tendency in teenagers. Prevalence of myopia among young teenagers in Sweden. Acta Ophthalmol Scand. 2000;78(2):177–81.

44. Matamoros E, Ingrand P, Pelen F, Bentaleb Y, Weber M, Korobelnik JF, et al. Prevalence of Myopia in France: A Cross-Sectional Analysis. Medicine (Baltimore). 2015 Nov;94(45):e1976.

45. Plainis S, Moschandreas J, Nikolitsa P, Plevridi E, Giannakopoulou T, Vitanova V, et al. Myopia and visual acuity impairment: a comparative study of Greek and Bulgarian school children. Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom. 2009 May;29(3):312–20.

46. Ding BY, Shih YF, Lin LLK, Hsiao CK, Wang IJ. Myopia among schoolchildren in East Asia and Singapore. Surv Ophthalmol. 2017 Sep 1;62(5):677–97.

47. He M, Zeng J, Liu Y, Xu J, Pokharel GP, Ellwein LB. Refractive error and visual impairment in urban children in southern china. Invest Ophthalmol Vis Sci. 2004 Mar;45(3):793–9.

48. Zhao J, Pan X, Sui R, Munoz SR, Sperduto RD, Ellwein LB. Refractive Error Study in Children: results from Shunyi District, China. Am J Ophthalmol. 2000 Apr;129(4):427–35.

49. Murthy GVS, Gupta SK, Ellwein LB, Muñoz SR, Pokharel GP, Sanga L, et al. Refractive error in children in an urban population in New Delhi. Invest Ophthalmol Vis Sci. 2002 Mar;43(3):623–31.

50. Pokharel GP, Negrel AD, Munoz SR, Ellwein LB. Refractive Error Study in Children: results from Mechi Zone, Nepal. Am J Ophthalmol. 2000 Apr;129(4):436–44.

51. Donovan L, Sankaridurg P, Ho A, Naduvilath T, Smith EL, Holden BA. Myopia progression rates in urban children wearing single-vision spectacles. Optom Vis Sci Off Publ Am Acad Optom. 2012 Jan;89(1):27–32.

52. Lin LLK, Shih YF, Hsiao CK, Chen CJ. Prevalence of myopia in Taiwanese schoolchildren: 1983 to 2000. Ann Acad Med Singapore. 2004 Jan;33(1):27–33.

53. Dirani M, Chan YH, Gazzard G, Hornbeak DM, Leo SW, Selvaraj P, et al. Prevalence of refractive error in Singaporean Chinese children: the strabismus, amblyopia, and refractive error in young Singaporean Children (STARS) study. Invest Ophthalmol Vis Sci. 2010 Mar;51(3):1348–55.

54. Saw SM, Carkeet A, Chia KS, Stone RA, Tan DTH. Component dependent risk factors for ocular parameters in Singapore Chinese children. Ophthalmology. 2002 Nov;109(11):2065–71.

55. Yoon KC, Mun GH, Kim SD, Kim SH, Kim CY, Park KH, et al. Prevalence of eye diseases in South Korea: data from the Korea National Health and Nutrition Examination Survey 2008-2009. Korean J Ophthalmol KJO. 2011 Dec;25(6):421–33.

56. Pan CW, Dirani M, Cheng CY, Wong TY, Saw SM. The Age-Specific Prevalence of Myopia in Asia: A Meta-analysis. Optom Vis Sci. 2015 Mar;92(3):258–66.

57. McCullough SJ, O'Donoghue L, Saunders KJ. Six Year Refractive Change among White Children and Young Adults: Evidence for Significant Increase in Myopia among White UK Children. PLOS ONE. 2016 Jan 19;11(1):e0146332.

58. Williams KM, Bertelsen G, Cumberland P, Wolfram C, Verhoeven VJM, Anastasopoulos E, et al. Increasing Prevalence of Myopia in Europe and the Impact of Education. Ophthalmology. 2015 Jul;122(7):1489–97.

59. Vitale S. Increased Prevalence of Myopia in the United States Between 1971-1972 and 1999-2004. Arch Ophthalmol. 2009 Dec 14;127(12):1632.

60. Zhang X, Qu X, Zhou X. Association between parental myopia and the risk of myopia in a child. Exp Ther Med. 2015 Jun 1;9(6):2420–8.

61. Jones LA, Sinnott LT, Mutti DO, Mitchell GL, Moeschberger ML, Zadnik K. Parental History of Myopia, Sports and Outdoor Activities, and Future Myopia. Invest Ophthalmol Vis Sci. 2007 Aug 1;48(8):3524–32.

62. Jones-Jordan LA, Sinnott LT, Manny RE, Cotter SA, Kleinstein RN, Mutti DO, et al. Early childhood refractive error and parental history of myopia as predictors of myopia. Invest Ophthalmol Vis Sci. 2010 Jan;51(1):115–21.

63. Xiang F, He M, Morgan IG. The impact of parental myopia on myopia in Chinese children: population-based evidence. Optom Vis Sci Off Publ Am Acad Optom. 2012 Oct;89(10):1487–96.

64. Hammond CJ, Snieder H, Gilbert CE, Spector TD. Genes and Environment in Refractive Error: The Twin Eye Study. Invest Ophthalmol Vis Sci. 2001 May 1;42(6):1232–6.

65. Lyhne N, Sjølie AK, Kyvik KO, Green A. 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. 2001 Dec 1;85(12):1470–6.

66. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era — concepts and misconceptions. Nat Rev Genet. 2008 Apr;9(4):255–66.

67. Young FA, Leary GA, Baldwin WR, West DC, Box RA, Harris E, et al. The transmission of refractive errors within eskimo families. Am J Optom Arch Am Acad Optom. 1969 Sep;46(9):676–85.

68. Angi MR, Clementi M, Sardei C, Piattelli E, Bisantis C. Heritability of myopic refractive errors in identical and fraternal twins. Graefes Arch Clin Exp Ophthalmol Albrecht Von Graefes Arch Klin Exp Ophthalmol. 1993 Oct;231(10):580–5.

69. Teikari JM, Kaprio J, Koskenvuo MK, Vannas A. Heritability estimate for refractive errors--a population-based sample of adult twins. Genet Epidemiol. 1988;5(3):171–81.

70. Sanfilippo PG, Hewitt AW, Hammond CJ, Mackey DA. The heritability of ocular traits. Surv Ophthalmol. 2010 Dec;55(6):561–83.

71. Tedja MS, Haarman AEG, Meester-Smoor MA, Kaprio J, Mackey DA, Guggenheim JA, et al. IMI - Myopia Genetics Report. Invest Ophthalmol Vis Sci. 2019 Feb 28;60(3):M89–105.

72. Chen CYC, Scurrah KJ, Stankovich J, Garoufalis P, Dirani M, Pertile KK, et al. Heritability and shared environment estimates for myopia and associated ocular biometric traits: the Genes in Myopia (GEM) family study. Hum Genet. 2007 May;121(3–4):511–20.

73. Wojciechowski R, Congdon N, Bowie H, Munoz B, Gilbert D, West SK. Heritability of refractive error and familial aggregation of myopia in an elderly American population. Invest Ophthalmol Vis Sci. 2005 May;46(5):1588–92.

74. Sorsby A, Fraser GR. STATISTICAL NOTE ON THE COMPONENTS OF OCULAR REFRACTION IN TWINS. J Med Genet. 1964 Sep;1(1):47–9.

75. Hammond CJ, Snieder H, Gilbert CE, Spector TD. Genes and Environment in Refractive Error: The Twin Eye Study. Invest Ophthalmol Vis Sci. 2001 May 1;42(6):1232–6.

76. Dirani M, Chamberlain M, Shekar SN, Islam AFM, Garoufalis P, Chen CY, et al. Heritability of refractive error and ocular biometrics: the Genes in Myopia (GEM) twin study. Invest Ophthalmol Vis Sci. 2006 Nov;47(11):4756–61.

77. Baird PN, Schäche M, Dirani M. The GEnes in Myopia (GEM) study in understanding the aetiology of refractive errors. Prog Retin Eye Res. 2010 Nov;29(6):520–42.

78. Hornbeak DM, Young TL. Myopia genetics: a review of current research and emerging trends. Curr Opin Ophthalmol. 2009 Sep;20(5):356–62.

79. Jacobi FK, Pusch CM. A decade in search of myopia genes. Front Biosci Landmark Ed. 2010 Jan 1;15:359–72.

80. Zhang Q, Guo X, Xiao X, Jia X, Li S, Hejtmancik JF. A new locus for autosomal dominant high myopia maps to 4q22-q27 between D4S1578 and D4S1612. Mol Vis. 2005 Jul 22;11:554–60.

81. Young TL, Ronan SM, Alvear AB, Wildenberg SC, Oetting WS, Atwood LD, et al. A second locus for familial high myopia maps to chromosome 12q. Am J Hum Genet. 1998 Nov;63(5):1419–24.

82. Naiglin L, Gazagne C, Dallongeville F, Thalamas C, Idder A, Rascol O, et al. A genome wide scan for familial high myopia suggests a novel locus on chromosome 7q36. J Med Genet. 2002 Feb;39(2):118–24.

83. Paluru P, Ronan SM, Heon E, Devoto M, Wildenberg SC, Scavello G, et al. New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. Invest Ophthalmol Vis Sci. 2003 May;44(5):1830–6.

84. Nallasamy S, Paluru PC, Devoto M, Wasserman NF, Zhou J, Young TL. Genetic linkage study of high-grade myopia in a Hutterite population from South Dakota. Mol Vis. 2007 Feb 15;13:229–36.

85. Lam CY, Tam POS, Fan DSP, Fan BJ, Wang DY, Lee CWS, et al. A Genome-wide Scan Maps a Novel High Myopia Locus to 5p15. Invest Ophthalmol Vis Sci. 2008 Sep 1;49(9):3768–78.

86. Stambolian D. Genetic susceptibility and mechanisms for refractive error. Clin Genet. 2013 Aug;84(2):102–8.

87. Hawthorne FA, Young TL. Genetic contributions to myopic refractive error: Insights from human studies and supporting evidence from animal models. Exp Eye Res. 2013 Sep 1;114:141–9.

88. Hart AB, de Wit H, Palmer AA. Candidate Gene Studies of a Promising Intermediate Phenotype: Failure to Replicate. Neuropsychopharmacology. 2013 Apr;38(5):802–16.

89. Cordell HJ, Clayton DG. Genetic association studies. The Lancet. 2005 Sep 24;366(9491):1121–31.

90. Mutti DO, Cooper ME, O'Brien S, Jones LA, Marazita ML, Murray JC, et al. Candidate gene and locus analysis of myopia. Mol Vis. 2007 Jun 28;13:1012–9.

91. Metlapally R, Li YJ, Tran-Viet KN, Abbott D, Czaja GR, Malecaze F, et al. COL1A1 and COL2A1 genes and myopia susceptibility: evidence of association and suggestive linkage to the COL2A1 locus. Invest Ophthalmol Vis Sci. 2009 Sep;50(9):4080–6.

92. Lin HJ, Wan L, Tsai Y, Tsai YY, Fan SS, Tsai CH, et al. The TGFbeta1 gene codon 10 polymorphism contributes to the genetic predisposition to high myopia. Mol Vis. 2006 Jun 21;12:698–703.

93. Veerappan S, Pertile KK, Islam AFM, Schäche M, Chen CY, Mitchell P, et al. Role of the hepatocyte growth factor gene in refractive error. Ophthalmology. 2010 Feb;117(2):239-245.e1-2.

94. Khor CC, Fan Q, Goh L, Tan D, Young TL, Li YJ, et al. Support for TGFB1 as a susceptibility gene for high myopia in individuals of Chinese descent. Arch Ophthalmol Chic Ill 1960. 2010 Aug;128(8):1081–4.

95. Hall NF, Gale CR, Ye S, Martyn CN. Myopia and Polymorphisms in Genes for Matrix Metalloproteinases. Invest Ophthalmol Vis Sci. 2009 Jun 1;50(6):2632–6.

96. Wojciechowski R, Yee SS, Simpson CL, Bailey-Wilson JE, Stambolian D. Matrix Metalloproteinases and Educational Attainment in Refractive Error: Evidence of Gene–Environment Interactions in the Age-related Eye Disease Study. Ophthalmology. 2013 Feb 1;120(2):298–305.

97. Chen KC, Hsi E, Hu CY, Chou WW, Liang CL, Juo SHH. MicroRNA-328 May Influence Myopia Development by Mediating the PAX6 Gene. Invest Ophthalmol Vis Sci. 2012 May 1;53(6):2732–9.

98. Tang SM, Rong SS, Young AL, Tam POS, Pang CP, Chen LJ. PAX6 gene associated with high myopia: a meta-analysis. Optom Vis Sci Off Publ Am Acad Optom. 2014 Apr;91(4):419–29.

99. Hysi PG, Choquet H, Khawaja AP, Wojciechowski R, Tedja MS, Yin J, et al. Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. Nat Genet. 2020 Apr;52(4):401–7.

100. Williams KM, Hammond CJ. GWAS in myopia: insights into disease and implications for the clinic. Expert Rev Ophthalmol. 2016 Mar 3;11(2):101–10.

101. Cai XB, Shen SR, Chen DF, Zhang Q, Jin ZB. An overview of myopia genetics. Exp Eye Res. 2019 Nov 1;188:107778.

102. Nakanishi H, Yamada R, Gotoh N, Hayashi H, Yamashiro K, Shimada N, et al. A Genome-Wide Association Analysis Identified a Novel Susceptible Locus for Pathological Myopia at 11q24.1. PLOS Genet. 2009 Sep 25;5(9):e1000660.

103. Liu J, Zhang H xin. Polymorphism in the 11q24.1 genomic region is associated with myopia: A comprehensive genetic study in Chinese and Japanese populations. Mol Vis. 2014 Mar 21;20:352–8.

104. Lu B, Jiang D, Wang P, Gao Y, Sun W, Xiao X, et al. Replication Study Supports CTNND2 as a Susceptibility Gene for High Myopia. Invest Ophthalmol Vis Sci. 2011 Oct 1;52(11):8258–61.

105. Yu Z, Zhou J, Chen X, Zhou X, Sun X, Chu R. Polymorphisms in the CTNND2 gene and 11q24.1 genomic region are associated with pathological myopia in a Chinese

population. Ophthalmol J Int Ophtalmol Int J Ophthalmol Z Augenheilkd. 2012;228(2):123–9.

106. Solouki AM, Verhoeven VJM, van Duijn CM, Verkerk AJMH, Ikram MK, Hysi PG, et al. A genome-wide association study identifies a susceptibility locus for refractive errors and myopia at 15q14. Nat Genet. 2010 Oct;42(10):897–901.

107. Stambolian D, Wojciechowski R, Oexle K, Pirastu M, Li X, Raffel LJ, et al. Meta-analysis of genome-wide association studies in five cohorts reveals common variants in RBFOX1, a regulator of tissue-specific splicing, associated with refractive error. Hum Mol Genet. 2013 Jul 1;22(13):2754–64.

108. Kiefer AK, Tung JY, Do CB, Hinds DA, Mountain JL, Francke U, et al. Genome-Wide Analysis Points to Roles for Extracellular Matrix Remodeling, the Visual Cycle, and Neuronal Development in Myopia. PLOS Genet. 2013 Feb 28;9(2):e1003299.

109. Wojciechowski R, Hysi PG. Focusing in on the complex genetics of myopia. PLoS Genet. 2013 Apr;9(4):e1003442.

110. Tedja MS, Wojciechowski R, Hysi PG, Eriksson N, Furlotte NA, Verhoeven VJM, et al. Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. Nat Genet. 2018 Jun;50(6):834–48.

111. Troilo D, Smith EL, Nickla DL, Ashby R, Tkatchenko AV, Ostrin LA, et al. IMI – Report on Experimental Models of Emmetropization and Myopia. Invest Ophthalmol Vis Sci. 2019 Feb 28;60(3):M31–88.

112. Hysi PG, Mahroo OA, Cumberland P, Wojciechowski R, Williams KM, Young TL, et al. Common Mechanisms Underlying Refractive Error Identified in Functional Analysis of Gene Lists From Genome-Wide Association Study Results in 2 European British Cohorts. JAMA Ophthalmol. 2014 Jan 1;132(1):50–6.

113. Hysi PG, Wojciechowski R, Rahi JS, Hammond CJ. Genome-Wide Association Studies of Refractive Error and Myopia, Lessons Learned, and Implications for the Future. Invest Ophthalmol Vis Sci. 2014 May;55(5):3344–51.

114. Stone RA, Lin T, Luvone PM, Laties AM. Postnatal Control of Ocular Growth: Dopaminergic Mechanisms. In: Ciba Foundation Symposium 155 - Myopia and the Control of Eye Growth [Internet]. John Wiley & Sons, Ltd; [cited 2021 Aug 18]. p. 45–62. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470514023.ch4

115. Brzyski D, Peterson CB, Sobczyk P, Candès EJ, Bogdan M, Sabatti C. Controlling the Rate of GWAS False Discoveries. Genetics. 2017 Jan 1;205(1):61–75.

116. Guggenheim JA, St Pourcain B, McMahon G, Timpson NJ, Evans DM, Williams C. Assumption-free estimation of the genetic contribution to refractive error across childhood. Mol Vis. 2015 May 26;21:621–32.

117. Cirulli ET, White S, Read RW, Elhanan G, Metcalf WJ, Tanudjaja F, et al. Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. Nat Commun. 2020 Jan 28;11(1):542.

118. Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, et al. Rare and low-frequency coding variants alter human adult height. Nature. 2017 Feb;542(7640):186–90.

119. Liu C, Kraja AT, Smith JA, Brody JA, Franceschini N, Bis JC, et al. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. Nat Genet. 2016 Oct;48(10):1162–70.

120. Flannick J, Mercader JM, Fuchsberger C, Udler MS, Mahajan A, Wessel J, et al. Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. Nature. 2019 Jun;570(7759):71–6.

121. Do R, Stitziel NO, Won HH, Jørgensen AB, Duga S, Angelica Merlini P, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature. 2015 Feb;518(7537):102–6.

122. Napolitano F, Iorio VD, Testa F, Tirozzi A, Reccia MG, Lombardi L, et al. Autosomal-dominant myopia associated to a novel P4HA2 missense variant and defective collagen hydroxylation. Clin Genet. 2018;93(5):982–91.

123. Feng L, Zhou D, Zhang Z, He L, Liu Y, Yang Y. Exome sequencing identifies a novel UNC5D mutation in a severe myopic anisometropia family: A case report. Medicine (Baltimore). 2017 Jun;96(24):e7138.

124. Zhao F, Wu J, Xue A, Su Y, Wang X, Lu X, et al. Exome sequencing reveals CCDC111 mutation associated with high myopia. Hum Genet. 2013 Aug;132(8):913–21.

125. Sun W, Huang L, Xu Y, Xiao X, Li S, Jia X, et al. Exome Sequencing on 298 Probands With Early-Onset High Myopia: Approximately One-Fourth Show Potential Pathogenic Mutations in RetNet Genes. Invest Ophthalmol Vis Sci. 2015 Dec 1;56(13):8365–72.

126. Goldschmidt E, Jacobsen N. Genetic and environmental effects on myopia development and progression. Eye. 2014 Feb;28(2):126–33.

127. Morgan I, Rose K. How genetic is school myopia? Prog Retin Eye Res. 2005 Jan 1;24(1):1–38.

128. Kim EC, Morgan IG, Kakizaki H, Kang S, Jee D. Prevalence and Risk Factors for Refractive Errors: Korean National Health and Nutrition Examination Survey 2008-2011. PLOS ONE. 2013 Nov 5;8(11):e80361.

129. Morgan IG, French AN, Rose KA. Intense schooling linked to myopia. BMJ. 2018 Jun 6;361:k2248.

130. Guggenheim JA, Northstone K, McMahon G, Ness AR, Deere K, Mattocks C, et al. Time Outdoors and Physical Activity as Predictors of Incident Myopia in Childhood: A Prospective Cohort Study. Invest Ophthalmol Vis Sci. 2012 May 1;53(6):2856–65.

131. Plotnikov D, Williams C, Atan D, Davies NM, Mojarrad NG, Guggenheim JA, et al. Effect of Education on Myopia: Evidence from the United Kingdom ROSLA 1972 Reform. Invest Ophthalmol Vis Sci. 2020 Sep 1;61(11):7–7.

132. Mountjoy E, Davies NM, Plotnikov D, Smith GD, Rodriguez S, Williams CE, et al. Education and myopia: assessing the direction of causality by mendelian randomisation. BMJ. 2018 Jun 6;361:k2022.

133. D C, E L, M C. Are children with myopia more intelligent? A literature review. Ann Acad Med Stetin. 2008 Jan 1;54(1):13–6; discussion 16.

134. Saw SM, Tan SB, Fung D, Chia KS, Koh D, Tan DTH, et al. IQ and the Association with Myopia in Children. Invest Ophthalmol Vis Sci. 2004 Sep 1;45(9):2943–8.

135. Rosner M, Belkin M. Intelligence, Education, and Myopia in Males. Arch Ophthalmol. 1987 Nov 1;105(11):1508–11.

136. Tw T, J F, E G. Degree of myopia in relation to intelligence and educational level. Lancet Lond Engl. 1988 Dec 1;2(8624):1351–4.

137. Megreli J, Barak A, Bez M, Bez D, Levine H. Association of Myopia with cognitive function among one million adolescents. BMC Public Health. 2020 May 8;20(1):647.

138. Mirshahi A, Ponto KA, Laubert-Reh D, Rahm B, Lackner KJ, Binder H, et al. Myopia and Cognitive Performance: Results From the Gutenberg Health Study. Invest Ophthalmol Vis Sci. 2016 Oct 1;57(13):5230–6.

139. Spierer O, Fischer N, Barak A, Belkin M. Correlation Between Vision and Cognitive Function in the Elderly. Medicine (Baltimore) [Internet]. 2016 Jan 22 [cited 2021 Apr 22];95(3). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4998246/

140. Mutti DO, Mitchell GL, Moeschberger ML, Jones LA, Zadnik K. Parental Myopia, Near Work, School Achievement, and Children's Refractive Error. Invest Ophthalmol Vis Sci. 2002 Dec 1;43(12):3633–40.

141. Ramsden S, Richardson FM, Josse G, Shakeshaft C, Seghier ML, Price CJ. The influence of reading ability on subsequent changes in verbal IQ in the teenage years. Dev Cogn Neurosci. 2013 Oct 1;6:30–9.

142. Li SM, Li SY, Kang MT, Zhou Y, Liu LR, Li H, et al. Near Work Related Parameters and Myopia in Chinese Children: the Anyang Childhood Eye Study. PLOS ONE. 2015 Aug 5;10(8):e0134514.

143. Young FA. READING, MEASURES OF INTELLIGENCE AND REFRACTIVE ERRORS*. Optom Vis Sci. 1963 May;40(5):257–64.

144. Mak W, Kwan MWM, Cheng TS, Chan KH, Cheung RTF, Ho SL. Myopia as a latent phenotype of a pleiotropic gene positively selected for facilitating neurocognitive development, and the effects of environmental factors in its expression. Med Hypotheses. 2006 Jan 1;66(6):1209–15.

145. Williams KM, Hysi PG, Yonova-Doing E, Mahroo OA, Snieder H, Hammond CJ. Phenotypic and genotypic correlation between myopia and intelligence. Sci Rep [Internet]. 2017 Apr 6 [cited 2018 Aug 21];7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5382686/

146. Mutti DO, Mitchell GL, Moeschberger ML, Jones LA, Zadnik K. Parental myopia, near work, school achievement, and children's refractive error. Invest Ophthalmol Vis Sci. 2002 Dec;43(12):3633–40.

147. Saw SM, Chua WH, Hong CY, Wu HM, Chan WY, Chia KS, et al. Nearwork in early-onset myopia. Invest Ophthalmol Vis Sci. 2002 Feb;43(2):332–9.

148. Wallman J, Winawer J. Homeostasis of eye growth and the question of myopia. Neuron. 2004 Aug 19;43(4):447–68.

149. Huang HM, Chang DST, Wu PC. The Association between Near Work Activities and Myopia in Children—A Systematic Review and Meta-Analysis. PLOS ONE. 2015 Oct 20;10(10):e0140419.

150. Jones-Jordan LA, Mitchell GL, Cotter SA, Kleinstein RN, Manny RE, Mutti DO, et al. Visual Activity before and after the Onset of Juvenile Myopia. Invest Ophthalmol Vis Sci. 2011 Mar 1;52(3):1841–50.

151. French AN, Morgan IG, Mitchell P, Rose KA. Risk Factors for Incident Myopia in Australian Schoolchildren: The Sydney Adolescent Vascular and Eye Study. Ophthalmology. 2013 Oct 1;120(10):2100–8.

152. Rose KA, Morgan IG, Ip J, Kifley A, Huynh S, Smith W, et al. Outdoor Activity Reduces the Prevalence of Myopia in Children. Ophthalmology. 2008 Aug 1;115(8):1279–85.

153. He M, Xiang F, Zeng Y, Mai J, Chen Q, Zhang J, et al. Effect of Time Spent Outdoors at School on the Development of Myopia Among Children in China: A Randomized Clinical Trial. JAMA. 2015 Sep 15;314(11):1142–8.

154. Wu PC, Chen CT, Lin KK, Sun CC, Kuo CN, Huang HM, et al. Myopia Prevention and Outdoor Light Intensity in a School-Based Cluster Randomized Trial. Ophthalmology. 2018 Jan 19;

155. Sherwin JC, Reacher MH, Keogh RH, Khawaja AP, Mackey DA, Foster PJ. The Association between Time Spent Outdoors and Myopia in Children and Adolescents:

A Systematic Review and Meta-analysis. Ophthalmology. 2012 Oct 1;119(10):2141–51.

156. Xiong S, Sankaridurg P, Naduvilath T, Zang J, Zou H, Zhu J, et al. Time spent in outdoor activities in relation to myopia prevention and control: a meta-analysis and systematic review. Acta Ophthalmol (Copenh). 2017 Sep;95(6):551–66.

157. Lingham G, Mackey DA, Lucas R, Yazar S. How does spending time outdoors protect against myopia? A review. Br J Ophthalmol. 2020 May 1;104(5):593–9.

158. Cuellar-Partida G, Williams KM, Yazar S, Guggenheim JA, Hewitt AW, Williams C, et al. Genetically low vitamin D concentrations and myopic refractive error: a Mendelian randomization study. Int J Epidemiol. 2017 Dec 1;46(6):1882–90.

159. Harrington SC, Stack J, O'Dwyer V. Risk factors associated with myopia in schoolchildren in Ireland. Br J Ophthalmol. 2019 Dec 1;103(12):1803–9.

160. Mandel Y, Grotto I, El-Yaniv R, Belkin M, Israeli E, Polat U, et al. Season of Birth, Natural Light, and Myopia. Ophthalmology. 2008 Apr 1;115(4):686–92.

161. McMahon G, Zayats T, Chen YP, Prashar A, Williams C, Guggenheim JA. Season of Birth, Daylight Hours at Birth, and High Myopia. Ophthalmology. 2009 Mar 1;116(3):468–73.

162. Williams KM, Kraphol E, Yonova-Doing E, Hysi PG, Plomin R, Hammond CJ. Early life factors for myopia in the British Twins Early Development Study. Br J Ophthalmol. 2019 Aug 1;103(8):1078–84.

163. Rahi JS, Cumberland PM, Peckham CS. Myopia Over the Lifecourse: Prevalence and Early Life Influences in the 1958 British Birth Cohort. Ophthalmology. 2011 May 1;118(5):797–804.

164. Ip JM, Huynh SC, Robaei D, Rose KA, Morgan IG, Smith W, et al. Ethnic differences in the impact of parental myopia: findings from a population-based study of 12-year-old Australian children. Invest Ophthalmol Vis Sci. 2007 Jun;48(6):2520–8.

165. Wei CC, Lin HJ, Lim YP, Chen CS, Chang CY, Lin CJ, et al. PM2.5 and NOx exposure promote myopia: clinical evidence and experimental proof. Environ Pollut. 2019 Nov 1;254:113031.

166. Dadvand P, Nieuwenhuijsen MJ, Basagaña X, Alvarez-Pedrerol M, Dalmau-Bueno A, Cirach M, et al. Traffic-related air pollution and spectacles use in schoolchildren. PLOS ONE. 2017 Apr 3;12(4):e0167046.

167. Morgan IG, Wu PC, Ostrin LA, Tideman JWL, Yam JC, Lan W, et al. IMI Risk Factors for Myopia. Invest Ophthalmol Vis Sci. 2021 Apr 28;62(5):3–3.

168. Ip JM, Rose KA, Morgan IG, Burlutsky G, Mitchell P. Myopia and the Urban Environment: Findings in a Sample of 12-Year-Old Australian School Children. Invest Ophthalmol Vis Sci. 2008 Sep 1;49(9):3858–63.

169. Choi KY, Yu WY, Lam CHI, Li ZC, Chin MP, Lakshmanan Y, et al. Childhood exposure to constricted living space: a possible environmental threat for myopia development. Ophthalmic Physiol Opt. 2017;37(5):568–75.

170. Zhang M, Li L, Chen L, Lee J, Wu J, Yang A, et al. Population Density and Refractive Error among Chinese Children. Invest Ophthalmol Vis Sci. 2010 Oct 1;51(10):4969–76.

171. Morris TT, Guggenheim JA, Northstone K, Williams C. Geographical Variation in Likely Myopia and Environmental Risk Factors: A Multilevel Cross Classified Analysis of A UK Cohort. Ophthalmic Epidemiol. 2020 Jan 2;27(1):1–9.

172. Kong A, Thorleifsson G, Frigge ML, Vilhjalmsson BJ, Young AI, Thorgeirsson TE, et al. The nature of nurture: Effects of parental genotypes. Science. 2018 Jan 26;359(6374):424–8.

173. Williams KM, Kraphol E, Yonova-Doing E, Hysi PG, Plomin R, Hammond CJ. Early life factors for myopia in the British Twins Early Development Study. Br J Ophthalmol. 2019 Aug 1;103(8):1078–84.

174. Rahi JS, Cumberland PM, Peckham CS. Myopia over the lifecourse: prevalence and early life influences in the 1958 British birth cohort. Ophthalmology. 2011 May;118(5):797–804.

175. GARDINER PA. The diet of growing myopes. Trans Ophthalmol Soc U K. 1956 Jan 1;76:171–80.

176. Cordain L, Eaton SB, Miller JB, Lindeberg S, Jensen C. An evolutionary analysis of the aetiology and pathogenesis of juvenile-onset myopia. Acta Ophthalmol Scand. 2002;80(2):125–35.

177. Feldkaemper MP, Neacsu I, Schaeffel F. Insulin Acts as a Powerful Stimulator of Axial Myopia in Chicks. Invest Ophthalmol Vis Sci. 2009 Jan 1;50(1):13–23.

178. Zhu X, Wallman J. Opposite Effects of Glucagon and Insulin on Compensation for Spectacle Lenses in Chicks. Invest Ophthalmol Vis Sci. 2009 Jan 1;50(1):24–36.

179. Sheng C, Zhu X, Wallman J. In vitro effects of insulin and RPE on choroidal and scleral components of eye growth in chicks. Exp Eye Res. 2013 Nov 1;116:439–48.

180. Tang R hong, Tan J, Deng Z hong, Zhao S zhen, Miao Y bin, Zhang W juan. Insulin-like growth factor-2 antisense oligonucleotides inhibits myopia by expression blocking of retinal insulin-like growth factor-2 in guinea pig. Clin Experiment Ophthalmol. 2012;40(5):503–11.

181. Galvis V, López-Jaramillo P, Tello A, Castellanos-Castellanos YA, Camacho PA, Cohen DD, et al. Is myopia another clinical manifestation of insulin resistance? Med Hypotheses. 2016 May;90:32–40.

182. Liu X, Wang P, Qu C, Zheng H, Gong B, Ma S, et al. Genetic association study between INSULIN pathway related genes and high myopia in a Han Chinese population. Mol Biol Rep. 2015 Jan;42(1):303–10.

183. Zhuang W, Yang P, Li Z, Sheng X, Zhao J, Li S, et al. Association of insulin-like growth factor-1 polymorphisms with high myopia in the Chinese population. Mol Vis. 2012 Mar 13;18:634–44.

184. Metlapally R, Ki CS, Li YJ, Tran-Viet KN, Abbott D, Malecaze F, et al. Genetic Association of Insulin-like Growth Factor-1 Polymorphisms with High-Grade Myopia in an International Family Cohort. Invest Ophthalmol Vis Sci. 2010 Sep 1;51(9):4476–9.

185. Jacobsen N, Jensen H, Lund-Andersen H, Goldschmidt E. Is poor glycaemic control in diabetic patients a risk factor of myopia? Acta Ophthalmol (Copenh). 2008 Jul 22;86(5):510–4.

186. Chen SJ, Tung TH, Liu JH, Lee AF, Lee FL, Hsu WM, et al. Prevalence and Associated Factors of Refractive Errors Among Type 2 Diabetics in Kinmen, Taiwan. Ophthalmic Epidemiol. 2008 Jan 1;15(1):2–9.

187. Umezurike BC, Udeala O, Green UG, Okpechi-Agbo U, Ohaeri MU. The Pathogenesis of Index Myopia in Hyperglycemia in Type 2 Diabetes: A Review. Ophthalmol Res Int J. 2018 Jul 5;1–17.

188. Gardiner PA. The relation of myopia to growth. Lancet. 1954;266:476–9.

189. Harrington SC, Stack J, O'Dwyer V. Risk factors associated with myopia in schoolchildren in Ireland. Br J Ophthalmol. 2019 Dec 1;103(12):1803–9.

190. Terasaki H, Yamashita T, Yoshihara N, Kii Y, Sakamoto T. Association of lifestyle and body structure to ocular axial length in Japanese elementary school children. BMC Ophthalmol. 2017 Jul 12;17(1):123.

191. Tideman JWL, Polling JR, Hofman A, Jaddoe VW, Mackenbach JP, Klaver CC. Environmental factors explain socioeconomic prevalence differences in myopia in 6-year-old children. Br J Ophthalmol. 2018 Feb;102(2):243–7.

192. Pan CW, Saw SM, Wong TY. Epidemiology of Myopia. In: Spaide RF, Ohno-Matsui K, Yannuzzi LA, editors. Pathologic Myopia [Internet]. New York, NY: Springer; 2014 [cited 2020 May 27]. p. 25–38. Available from: https://doi.org/10.1007/978-1-4614-8338-0_3

193. Vutipongsatorn K, Yokoi T, Ohno-Matsui K. Current and emerging pharmaceutical interventions for myopia. Br J Ophthalmol. 2019 Nov 1;103(11):1539–48.

194. Wildsoet CF, Chia A, Cho P, Guggenheim JA, Polling JR, Read S, et al. IMI – Interventions for Controlling Myopia Onset and Progression Report. Invest Ophthalmol Vis Sci. 2019 Feb 28;60(3):M106–31.

195. Wu Y, Byrne EM, Zheng Z, Kemper KE, Yengo L, Mallett AJ, et al. Genome-wide association study of medication-use and associated disease in the UK Biobank. Nat Commun. 2019 Dec;10(1):1891.

196. Wu A, Khawaja AP, Pasquale LR, Stein JD. A review of systemic medications that may modulate the risk of glaucoma. Eye. 2020 Jan;34(1):12–28.

197. Ho H, Shi Y, Chua J, Tham YC, Lim SH, Aung T, et al. Association of Systemic Medication Use With Intraocular Pressure in a Multiethnic Asian Population: The Singapore Epidemiology of Eye Diseases Study. JAMA Ophthalmol. 2017 Mar 1;135(3):196–202.

198. Saw SM, Gazzard G, Shih-Yen EC, Chua WH. Myopia and associated pathological complications. Ophthalmic Physiol Opt. 2005;25(5):381–91.

199. Bellezza AJ, Hart RT, Burgoyne CF. The Optic Nerve Head as a Biomechanical Structure: Initial Finite Element Modeling. Invest Ophthalmol Vis Sci. 2000 Sep 1;41(10):2991–3000.

200. Kim TW, Kim M, Weinreb RN, Woo SJ, Park KH, Hwang JM. Optic Disc Change with Incipient Myopia of Childhood. Ophthalmology. 2012 Jan 1;119(1):21-26.e3.

201. Spaide RF. Staphyloma: Part 1. In: Spaide RF, Ohno-Matsui K, Yannuzzi LA, editors. Pathologic Myopia [Internet]. New York, NY: Springer; 2014 [cited 2021 Oct 3]. p. 167–76. Available from: https://doi.org/10.1007/978-1-4614-8338-0_12

202. Ohno-Matsui K, Lai TYY, Lai CC, Cheung CMG. Updates of pathologic myopia. Prog Retin Eye Res. 2016;52:156–87.

203. Ohno-Matsui K, Jonas JB. Posterior staphyloma in pathologic myopia. Prog Retin Eye Res. 2019 May 1;70:99–109.

204. Ohno-Matsui K. Myopic Chorioretinal Atrophy. In: Spaide RF, Ohno-Matsui K, Yannuzzi LA, editors. Pathologic Myopia [Internet]. New York, NY: Springer; 2014 [cited 2021 Oct 3]. p. 187–209. Available from: https://doi.org/10.1007/978-1-4614-8338-0_14

205. Kobayashi K, Ohno-Matsui K, Kojima A, Shimada N, Yasuzumi K, Yoshida T, et al. Fundus Characteristics of High Myopia in Children. Jpn J Ophthalmol. 2005 Jul 1;49(4):306–11.

206. Ohno-Matsui K, Yoshida T, Futagami S, Yasuzumi K, Shimada N, Kojima A, et al. Patchy atrophy and lacquer cracks predispose to the development of choroidal neovascularisation in pathological myopia. Br J Ophthalmol. 2003 May 1;87(5):570–3.

207. Grossniklaus HE, Green WR. Choroidal neovascularization. Am J Ophthalmol. 2004 Mar 1;137(3):496–503.

208. Avila MP, Weiter JJ, Jalkh AE, Trempe CL, Pruett RC, Schepens CL. Natural History of Choroidal Neovascularization in Degenerative Myopia. Ophthalmology. 1984 Dec 1;91(12):1573–81.

209. Fried M, Siebert A, Meyer-Schwickerath G, Wessing A. Natural History of Fuchs' Spot: A Long-Term Follow-Up Study. In: Fledelius HC, Alsbirk PH, Goldschmidt E, editors. Third International Conference on Myopia Copenhagen, August 24–27, 1980 [Internet]. Dordrecht: Springer Netherlands; 1981 [cited 2021 Oct 2]. p. 215–21. (Documenta Ophthalmologica Proceedings Series). Available from: https://doi.org/10.1007/978-94-009-8662-6_31

210. Hampton GR, Kohen D, Bird AC. Visual Prognosis of Disciform Degeneration in Myopia. Ophthalmology. 1983 Aug 1;90(8):923–6.

211. Ruiz-Medrano J, Montero JA, Flores-Moreno I, Arias L, García-Layana A, Ruiz-Moreno JM. Myopic maculopathy: Current status and proposal for a new classification and grading system (ATN). Prog Retin Eye Res. 2019 Mar 1;69:80–115.

212. Chang L, Pan CW, Ohno-Matsui K, Lin X, Cheung GCM, Gazzard G, et al. Myopia-Related Fundus Changes in Singapore Adults With High Myopia. Am J Ophthalmol. 2013 Jun 1;155(6):991-999.e1.

213. Vongphanit J, Mitchell P, Wang JJ. Prevalence and progression of myopic retinopathy in an older population. Ophthalmology. 2002 Apr 1;109(4):704–11.

214. Fricke TR, Jong M, Naidoo KS, Sankaridurg P, Naduvilath TJ, Ho SM, et al. Global prevalence of visual impairment associated with myopic macular degeneration and temporal trends from 2000 through 2050: systematic review, meta-analysis and modelling. Br J Ophthalmol. 2018 Jul 1;102(7):855–62.

215. Klaver CCW, Wolfs RCW, Vingerling JR, Hofman A, Jong PTVM de. Age-Specific Prevalence and Causes of Blindness and Visual Impairment in an Older Population: The Rotterdam Study. Arch Ophthalmol. 1998 May 1;116(5):653–8.

216. Liu JH, Cheng CY, Chen SJ, Lee FL. Visual impairment in a Taiwanese population: Prevalence, causes, and socioeconomic factors. Ophthalmic Epidemiol. 2001 Dec;8(5):339–50.

217. Naidoo KS, Fricke TR, Frick KD, Jong M, Naduvilath TJ, Resnikoff S, et al. Potential Lost Productivity Resulting from the Global Burden of Myopia: Systematic Review, Meta-analysis, and Modeling. Ophthalmology. 2019 Mar 1;126(3):338–46.

218. Smith TST, Frick KD, Holden BA, Fricke TR, Naidoo KS. Potential lost productivity resulting from the global burden of uncorrected refractive error. Bull World Health Organ. 2009 Jun;87:431–7.

219. Naidoo KS, Leasher J, Bourne RR, Flaxman SR, Jonas JB, Keeffe J, et al. Global Vision Impairment and Blindness Due to Uncorrected Refractive Error, 1990–2010. Optom Vis Sci. 2016 Mar;93(3):227–34.

220. Holden BA. Blindness and poverty: a tragic combination. Clin Exp Optom. 2007;90(6):401–3.

221. Thiagalingam S, Cumming RG, Mitchell P. Factors associated with undercorrected refractive errors in an older population: the Blue Mountains Eye Study. Br J Ophthalmol. 2002 Sep 1;86(9):1041–5.

222. Varma R, Wang MY, Ying-Lai M, Donofrio J, Azen SP. The Prevalence and Risk Indicators of Uncorrected Refractive Error and Unmet Refractive Need in Latinos: The Los Angeles Latino Eye Study. Invest Ophthalmol Vis Sci. 2008 Dec 1;49(12):5264–73.

223. Saw SM, Foster PJ, Gazzard G, Friedman D, Hee J, Seah S. Undercorrected refractive error in Singaporean Chinese adults: The Tanjong Pagar survey. Ophthalmology. 2004 Dec 1;111(12):2168–74.

224. Rosman M, Wong TY, Tay WT, Tong L, Saw SM. Prevalence and Risk Factors of Undercorrected Refractive Errors among Singaporean Malay Adults: The Singapore Malay Eye Study. Invest Ophthalmol Vis Sci. 2009 Aug 1;50(8):3621–8.

225. Murray CJL, Barber RM, Foreman KJ, Ozgoren AA, Abd-Allah F, Abera SF, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. The Lancet. 2015 Nov 28;386(10009):2145–91.

226. Lou L, Liu X, Tang X, Wang L, Ye J. Gender Inequality in Global Burden of Uncorrected Refractive Error. Am J Ophthalmol. 2019 Feb 1;198:1–7.

227. Lou L, Yao C, Jin Y, Perez V, Ye J. Global Patterns in Health Burden of Uncorrected Refractive Error. Invest Ophthalmol Vis Sci. 2016 Nov 1;57(14):6271–7.

228. Ganesan P, Wildsoet CF. Pharmaceutical intervention for myopia control. Expert Rev Ophthalmol. 2010 Dec 1;5(6):759–87.

229. Lin HJ, Wei CC, Chang CY, Chen TH, Hsu YA, Hsieh YC, et al. Role of Chronic Inflammation in Myopia Progression: Clinical Evidence and Experimental Validation. EBioMedicine. 2016 Jul 19;10:269–81.

230. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018 Oct;562(7726):203–9.

231. Cumberland PM, Bao Y, Hysi PG, Foster PJ, Hammond CJ, Rahi JS, et al. Frequency and Distribution of Refractive Error in Adult Life: Methodology and Findings of the UK Biobank Study. PLOS ONE. 2015 Oct 2;10(10):e0139780.

232. Cheesman R, Coleman J, Rayner C, Purves KL, Morneau-Vaillancourt G, Glanville K, et al. Familial Influences on Neuroticism and Education in the UK Biobank. Behav Genet. 2020 Mar 1;50(2):84–93.

233. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North. London; New York: Croom Helm; 1988.

234. Ye J, Wen Y, Sun X, Chu X, Li P, Cheng B, et al. Socioeconomic Deprivation Index Is Associated With Psychiatric Disorders: An Observational and Genome-wide Gene-by-Environment Interaction Analysis in the UK Biobank Cohort. Biol Psychiatry. 2021 May 1;89(9):888–95.

235. Hout CVV, Tachmazidou I, Backman JD, Hoffman JX, Ye B, Pandey AK, et al. Whole exome sequencing and characterization of coding variation in 49,960 individuals in the UK Biobank. bioRxiv. 2019 Mar 9;572347.

236. Szustakowski JD, Balasubramanian S, Sasson A, Khalid S, Bronson PG, Kvikstad E, et al. Advancing Human Genetics Research and Drug Discovery through Exome Sequencing of the UK Biobank. medRxiv. 2020 Nov 4;2020.11.02.20222232.

237. Shen L, Melles RB, Metlapally R, Barcellos LF, Schaefer C, Risch NJ, et al. The Association of Refractive Error with Glaucoma in a Multiethnic Population. [Internet]. undefined. 2016 [cited 2018 Sep 4]. Available from: /paper/The-Association-of-Refractive-Error-with-Glaucoma-a-Shen-Melles/92904e1f310002cd9fde42b6e401478e0493d256

238. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods. 2013 Jan;10(1):5–6.

239. Roshyara NR, Horn K, Kirsten H, Ahnert P, Scholz M. Comparing performance of modern genotype imputation methods in different ethnicities. Sci Rep. 2016 Oct 4;6(1):34386.

240. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007 Jul;39(7):906–13.

241. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012 Aug;44(8):955–9.

242. Verdi S, Abbasian G, Bowyer RCE, Lachance G, Yarand D, Christofidou P, et al. TwinsUK: The UK Adult Twin Registry Update. Twin Res Hum Genet. 2019 Dec;22(6):523–9.

243. Ramessur R, Williams KM, Hammond CJ. Risk factors for myopia in a discordant monozygotic twin study. Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom. 2015 Nov;35(6):643–51.

244. Long T, Hicks M, Yu HC, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet. 2017 Apr;49(4):568–78.

245. Raczy C, Petrovski R, Saunders CT, Chorny I, Kruglyak S, Margulies EH, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. Bioinformatics. 2013 Aug 15;29(16):2041–3.

246. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009 Sep;19(9):1655–64.

247. Moltke I, Albrechtsen A. RelateAdmix: a software tool for estimating relatedness between admixed individuals. Bioinformatics. 2014 Apr 1;30(7):1027–8.

248. Klein R, Klein BEK, Lee KE. Changes in Visual Acuity in a Population: The Beaver Dam Eye Study. Ophthalmology. 1996 Aug 1;103(8):1169–78.

249. Vergara C, Bomotti SM, Valencia C, Klein BEK, Lee KE, Klein R, et al. Association analysis of exome variants and refraction, axial length, and corneal curvature in a European–American population. Hum Mutat. 2018;39(12):1973–9.

250. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016 Oct;48(10):1284–7.

251. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2012 Feb;9(2):179–81.

252. Kretzschmar W, Mahajan A, Sharp K, McCarthy MI, Consortium HR. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet [Internet]. 2016 [cited 2021 Sep 22];48. Available from: https://ora.ox.ac.uk/objects/uuid:f64e7ec1-b682-4ce8-bcf3-4f6c14d3478e

253. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol. 1997 Jan 1;26(suppl_1):S6–S6.

254. Hayat SA, Luben R, Keevil VL, Moore S, Dalzell N, Bhaniani A, et al. Cohort Profile: A prospective cohort study of objective physical and cognitive capability and visual health in an ageing population of men and women in Norfolk (EPIC-Norfolk 3). Int J Epidemiol. 2014 Aug 1;43(4):1063–72.

255. Delaneau O, Marchini J. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat Commun. 2014 Jun 13;5(1):3934.

256. Saw SM, Shankar A, Tan SB, Taylor H, Tan DTH, Stone RA, et al. A Cohort Study of Incident Myopia in Singaporean Children. Invest Ophthalmol Vis Sci. 2006 May 1;47(5):1839–44.

257. Saw SM, Chua WH, Hong CY, Wu HM, Chia KS, Stone RA, et al. Height and Its Relationship to Refraction and Biometry Parameters in Singapore Chinese Children. Invest Ophthalmol Vis Sci. 2002 May 1;43(5):1408–13.

258. Dirani M, Tong L, Gazzard G, Zhang X, Chia A, Young TL, et al. Outdoor activity and myopia in Singapore teenage children. Br J Ophthalmol. 2009 Aug 1;93(8):997–1000.

259. Li YJ, Goh L, Khor CC, Fan Q, Yu M, Han S, et al. Genome-Wide Association Studies Reveal Genetic Variants in CTNND2 for High Myopia in Singapore Chinese. Ophthalmology. 2011 Feb 1;118(2):368–75.

260. Lanca C, Kassam I, Patasova K, Foo LL, Li J, Ang M, et al. New Polygenic Risk Score to Predict High Myopia in Singapore Chinese Children. Transl Vis Sci Technol. 2021 Jul 1;10(8):26–26.

261. Fan Q, Wojciechowski R, Kamran Ikram M, Cheng CY, Chen P, Zhou X, et al. Education influences the association between genetic variants and refractive error: a meta-analysis of five Singapore studies. Hum Mol Genet. 2014 Jan 15;23(2):546–54.

262. Soh SE, Tint MT, Gluckman PD, Godfrey KM, Rifkin-Graboi A, Chan YH, et al. Cohort Profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study. Int J Epidemiol. 2014 Oct 1;43(5):1401–9.

263. Bůžková P. Linear Regression in Genetic Association Studies. PLOS ONE. 2013 Feb 21;8(2):e56976.

264. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019 Aug;20(8):467–84.

265. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559–75.

266. Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 2010 Apr;42(4):355–60.

267. Jiang L, Zheng Z, Qi T, Kemper KE, Wray NR, Visscher PM, et al. A resource-efficient tool for mixed model association analysis of large-scale data. Nat Genet. 2019 Dec;51(12):1749–55.

268. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet. 2015 Mar;47(3):284–90.

269. Pinheiro JC, Bates DM, editors. Linear Mixed-Effects Models: Basic Concepts and Examples. In: Mixed-Effects Models in Sand S-PLUS [Internet]. New York, NY: Springer; 2000 [cited 2021 Aug 24]. p. 3–56. (Statistics and Computing). Available from: https://doi.org/10.1007/978-1-4419-0318-1_1

270. Fox J, Monette G. An R and S-Plus Companion to Applied Regression. SAGE; 2002. 332 p.

271. O'Brien RM, Hudson K, Stockard J. A Mixed Model Estimation of Age, Period, and Cohort Effects. Sociol Methods Res. 2008 Feb 1;36(3):402–28.

272. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. Am J Hum Genet. 2014 Jul 3;95(1):5–23.

273. Li B, Leal SM. Methods for Detecting Associations with Rare Variants for Common Diseases: Application to Analysis of Sequence Data. Am J Hum Genet. 2008 Sep 12;83(3):311–21.

274. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. Optimal Unified Approach for Rare-Variant Association Testing with Application to Small-Sample Case-Control Whole-Exome Sequencing Studies. Am J Hum Genet. 2012 Aug 10;91(2):224–37.

275. Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, et al. Testing for an Unusual Distribution of Rare Variants. PLOS Genet. 2011 Mar 3;7(3):e1001322.

276. Basu S, Pan W. Comparison of statistical tests for disease association with rare variants. Genet Epidemiol. 2011;35(7):606–19.

277. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. Am J Hum Genet. 2011 Jul 15;89(1):82–93.

278. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics. 2016 May 1;32(9):1423–6.

279. Quick C, Wen X, Abecasis G, Boehnke M, Kang HM. Integrating comprehensive functional annotations to boost power and accuracy in gene-based association analysis. PLOS Genet. 2020 Dec 15;16(12):e1009060.

280. Schober P, Vetter TR. Survival Analysis and Interpretation of Time-to-Event Data: The Tortoise and the Hare. Anesth Analg. 2018 Sep;127(3):792–8.

281. Bi W, Fritsche LG, Mukherjee B, Kim S, Lee S. A Fast and Accurate Method for Genome-Wide Time-to-Event Data Analysis and Its Application to UK Biobank. Am J Hum Genet. 2020 Aug 6;107(2):222–33.

282. Aa R, E K, M M, L P, J W, M S, et al. gwasurvivr: an R package for genome-wide survival analysis. Bioinforma Oxf Engl. 2019 Jun 1;35(11):1968–70.

283. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015 Mar;47(3):291–5.

284. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019 Jan 8;47:D1005–12.

285. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015 Nov;47(11):1236–41.

286. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A Tool for Genome-wide Complex Trait Analysis. Am J Hum Genet. 2011 Jan 7;88(1):76–82.

287. Duncan LE, Holmans PA, Lee PH, O'Dushlaine CT, Kirby AW, Smoller JW, et al. Pathway Analyses Implicate Glial Cells in Schizophrenia. PLOS ONE. 2014 Feb 24;9(2):e89441.

288. Yoon S, Nguyen HCT, Yoo YJ, Kim J, Baik B, Kim S, et al. Efficient pathway enrichment and network analysis of GWAS summary data using GSA-SNP2. Nucleic Acids Res. 2018 Jun 1;46(10):e60.

289. Goeman JJ, Bühlmann P. Analyzing gene expression data in terms of gene sets: methodological issues. Bioinformatics. 2007 Apr 15;23(8):980–7.

290. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000 May;25(1):25–9.

291. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999 Jan 1;27(1):29–34.

292. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. JAMA. 2017 Nov 21;318(19):1925–6.

293. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nat Commun. 2020 Jan 17;11(1):376.

294. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. Res Synth Methods. 2019;10(4):486–96.

295. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. Genet Epidemiol. 2013;37(7):658–65.

296. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016 May;40(4):304–14.

297. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015 Apr 1;44(2):512–25.

298. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017 May;32(5):377–89.

299. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016 May;48(5):481–7.

300. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet. 2013 Oct;45(10):1238–43.

301. Qi T, Wu Y, Zeng J, Zhang F, Xue A, Jiang L, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat Commun. 2018 Jun 11;9(1):2282.

302. Ratnapriya R, Sosina OA, Starostik MR, Kwicklis M, Kapphahn RJ, Fritsche LG, et al. Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. Nat Genet. 2019 Apr;51(4):606–10.

303. Fricke T, Holden B, Wilson D, Schlenther G, Naidoo K, Resnikoff S, et al. Global cost of correcting vision impairment from uncorrected refractive error. Bull World Health Organ. 2012 Oct 1;90(10):728–38.

304. Mutti DO, Zadnik K. Age-related decreases in the prevalence of myopia: longitudinal change or cohort effect? Invest Ophthalmol Vis Sci. 2000 Jul;41(8):2103–7.

305. Wong TY, Foster PJ, Johnson GJ, Seah SKL. Education, socioeconomic status, and ocular dimensions in Chinese adults: the Tanjong Pagar Survey. Br J Ophthalmol. 2002 Sep;86(9):963–8.

306. Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet. 2001 Nov;29(3):306–9.

307. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010 Sep 1;26(17):2190–1.

308. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Mägi R, et al. Quality control and conduct of genome-wide association meta-analyses. Nat Protoc. 2014 May;9(5):1192–212.

309. Khawaja AP, Cooke Bailey JN, Wareham NJ, Scott RA, Simcoe M, Igo RP, et al. Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. Nat Genet. 2018 Jun;50(6):778–82.

310. Benyamin B, Pourcain Bs, Davis OS, Davies G, Hansell NK, Brion MJ, et al. Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. Mol Psychiatry. 2014 Feb;19(2):253–8.

311. Carnes MU, Allingham RR, Ashley-Koch A, Hauser MA. Transcriptome analysis of adult and fetal trabecular meshwork, cornea, and ciliary body tissues by RNA sequencing. Exp Eye Res. 2018 Feb 1;167:91–9.

312. Segrè AV, Consortium D, Investigators M, Groop L, Mootha VK, Daly MJ, et al. Common Inherited Variation in Mitochondrial Genes Is Not Enriched for Associations with Type 2 Diabetes or Related Glycemic Traits. PLOS Genet. 2010 Aug 12;6(8):e1001058.

313. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, et al. Improved polygenic prediction by Bayesian multiple regression on summary statistics. Nat Commun. 2019 Nov 8;10(1):5086.

314. Zeng J, Xue A, Jiang L, Lloyd-Jones LR, Wu Y, Wang H, et al. Widespread signatures of natural selection across human complex traits and functional genomic categories. Nat Commun. 2021 Feb 19;12(1):1164.

315. Oliveira J, Santos R, Soares-Silva I, Jorge P, Vieira E, Oliveira ME, et al. LAMA2 gene analysis in a cohort of 26 congenital muscular dystrophy patients. Clin Genet. 2008;74(6):502–12.

316. Ervasti JM, Campbell KP. Dystrophin-associated glycoproteins: their possible roles in the pathogenesis of Duchenne muscular dystrophy. In: Partridge T, editor. Molecular and Cell Biology of Muscular Dystrophy [Internet]. Dordrecht: Springer Netherlands; 1993 [cited 2021 Jun 25]. p. 139–66. (Molecular and Cell Biology of Human Diseases Series). Available from: https://doi.org/10.1007/978-94-011-1528-5_6

317. Mayer U, Saher G, Fässler R, Bornemann A, Echtermeyer F, Mark H von der, et al. Absence of integrin α7 causes a novel form of muscular dystrophy. Nat Genet. 1997 Nov;17(3):318–23.

318. Thorleifsson G, Magnusson KP, Sulem P, Walters GB, Gudbjartsson DF, Stefansson H, et al. Common sequence variants in the LOXL1 gene confer susceptibility to exfoliation glaucoma. Science. 2007 Sep 7;317(5843):1397–400.

319. Rivera A, Fisher SA, Fritsche LG, Keilhauer CN, Lichtner P, Meitinger T, et al. Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. Hum Mol Genet. 2005 Nov 1;14(21):3227–36.

320. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48(7):709–17.

321. Wang Q, Chen Q, Zhao K, Wang L, Wang L, Traboulsi EI. Update on the molecular genetics of retinitis pigmentosa. Ophthalmic Genet. 2001 Jan 1;22(3):133–54.

322. Robinson PN, Booms P, Katzke S, Ladewig M, Neumann L, Palz M, et al. Mutations of FBN1 and genotype–phenotype correlations in Marfan syndrome and related fibrillinopathies. Hum Mutat. 2002;20(3):153–61.

323. Rohde K, Møller M, Rath MF. Homeobox Genes and Melatonin Synthesis: Regulatory Roles of the Cone-Rod Homeobox Transcription Factor in the Rodent Pineal Gland. BioMed Res Int. 2014 Apr 30;2014:e946075.

324. Chakraborty R, Ostrin LA, Nickla DL, Iuvone PM, Pardue MT, Stone RA. Circadian rhythms, refractive development, and myopia. Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom. 2018 May;38(3):217–45.

325. Ratnapriya R, Sosina OA, Starostik MR, Kwicklis M, Kapphahn RJ, Fritsche LG, et al. Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. Nat Genet. 2019 Apr;51(4):606–10.

326. Audo I, Kohl S, Leroy BP, Munier FL, Guillonneau X, Mohand-Saïd S, et al. TRPM1 Is Mutated in Patients with Autosomal-Recessive Complete Congenital Stationary Night Blindness. Am J Hum Genet. 2009 Nov 13;85(5):720–9.

327. Pozarickij A, Williams C, Hysi PG, Guggenheim JA. Quantile regression analysis reveals widespread evidence for gene-environment or gene-gene interactions in myopia development. Commun Biol [Internet]. 2019 May 6 [cited 2019 May 14];2. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6502837/

328. Meguro A, Yamane T, Takeuchi M, Miyake M, Fan Q, Zhao W, et al. Genome-Wide Association Study in Asians Identifies Novel Loci for High Myopia and Highlights a Nervous System Role in Its Pathogenesis. Ophthalmology. 2020 Dec 1;127(12):1612–24.

329. Zadnik K, Mutti DO, Mitchell GL, Jones LA, Burr D, Moeschberger ML. Normal Eye Growth in Emmetropic Schoolchildren. Optom Vis Sci. 2004 Nov;81(11):819–28.

330. Karouta C, Kucharski R, Hardy K, Thomson K, Maleszka R, Morgan I, et al. Transcriptome-based insights into gene networks controlling myopia prevention. FASEB J. 2021;35(9):e21846.

331. Troilo D, Gottlieb MD, Wallman J. Visual deprivation causes myopia in chicks with optic nerve section. Curr Eye Res. 1987 Jan 1;6(8):993–9.

332. Stone RA, Wei W, Sarfare S, McGeehan B, Engelhart KC, Khurana TS, et al. Visual Image Quality Impacts Circadian Rhythm-Related Gene Expression in Retina and in Choroid: A Potential Mechanism for Ametropias. Invest Ophthalmol Vis Sci. 2020 May 12;61(5):13.

333. Nickla DL. The phase relationships between the diurnal rhythms in axial length and choroidal thickness and the association with ocular growth rate in chicks. J Comp Physiol A. 2006 Apr 1;192(4):399–407.

334. Ghorbani Mojarrad N, Williams C, Guggenheim JA. A genetic risk score and number of myopic parents independently predict myopia. Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom. 2018 Sep;38(5):492–502.

335. Wang Y, Guo J, Ni G, Yang J, Visscher PM, Yengo L. Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. Nat Commun. 2020 Jul 31;11(1):3865.

336. Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet. 2012 Feb;13(2):135–45.

337. Stefansson H, Rujescu D, Cichon S, Pietiläinen OPH, Ingason A, Steinberg S, et al. Large recurrent microdeletions associated with schizophrenia. Nature. 2008 Sep;455(7210):232–6.

338. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong Association of De Novo Copy Number Mutations with Autism. Science. 2007 Apr 20;316(5823):445–9.

339. Van Hout CV, Tachmazidou I, Backman JD, Hoffman JD, Liu D, Pandey AK, et al. Exome sequencing and characterization of 49,960 individuals in the UK Biobank. Nature. 2020 Oct;586(7831):749–56.

340. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018 Oct;562(7726):203–9.

341. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics. 2012 Sep 1;13(4):762–75.

342. Vergara C, Bomotti SM, Valencia C, Klein BEK, Lee KE, Klein R, et al. Association analysis of exome variants and refraction, axial length, and corneal curvature in a European-American population. Hum Mutat. 2018 Dec;39(12):1973–9.

343. Bomotti S, Lau B, Klein BEK, Lee KE, Klein R, Duggal P, et al. Refraction and Change in Refraction Over a 20-Year Period in the Beaver Dam Eye Study. Invest Ophthalmol Vis Sci. 2018 Sep 4;59(11):4518–24.

344. Slavotinek AM. Eye development genes and known syndromes. Mol Genet Metab. 2011 Dec 1;104(4):448–56.

345. Gallardo ME, Lopez-Rios J, Fernaud-Espinosa I, Granadino B, Sanz R, Ramos C, et al. Genomic Cloning and Characterization of the Human Homeobox Gene SIX6 Reveals a Cluster of SIX Genes in Chromosome 14 and Associates SIX6 Hemizygosity with Bilateral Anophthalmia and Pituitary Anomalies. Genomics. 1999 Oct 1;61(1):82–91.

346. Ledford KL, Martinez-De Luna RI, Theisen MA, Rawlins KD, Viczian AS, Zuber ME. Distinct cis-acting regions control six6 expression during eye field and optic cup stages of eye formation. Dev Biol. 2017 Jun 15;426(2):418–28.

347. Carnes MU, Liu YP, Allingham RR, Whigham BT, Havens S, Garrett ME, et al. Discovery and functional annotation of SIX6 variants in primary open-angle glaucoma. PLoS Genet. 2014;10(5):e1004372.

348. Iglesias AI, Springelkamp H, van der Linde H, Severijnen LA, Amin N, Oostra B, et al. Exome sequencing and functional analyses suggest that SIX6 is a gene involved in an altered proliferation–differentiation balance early in life and optic nerve degeneration at old age. Hum Mol Genet. 2014 Mar 1;23(5):1320–32.

349. Mastropasqua L, Lobefalo L, Mancini A, Ciancaglini M, Palma S. Prevalence of myopia in open angle glaucoma. Eur J Ophthalmol. 1992 Mar;2(1):33–5.

350. Wong TY, Klein BEK, Klein R, Knudtson M, Lee KE. Refractive errors, intraocular pressure, and glaucoma in a white population. Ophthalmology. 2003 Jan;110(1):211–7.

351. Griffith JF, DeBenedictis MJ, Traboulsi EI. A novel dominant CRX mutation causes adult-onset macular dystrophy. Ophthalmic Genet. 2018 Jan 2;39(1):120–4.

352. Furukawa T, Morrow EM, Cepko CL. Crx, a Novel otx-like Homeobox Gene, Shows Photoreceptor-Specific Expression and Regulates Photoreceptor Differentiation. Cell. 1997 Nov 14;91(4):531–41.

353. Andzelm MM, Cherry TJ, Harmin DA, Boeke AC, Lee C, Hemberg M, et al. MEF2D Drives Photoreceptor Development through a Genome-wide Competition for Tissue-Specific Enhancers. Neuron. 2015 Apr 8;86(1):247–63.

354. Sohocki MM, Sullivan LS, Mintz-Hittner HA, Birch D, Heckenlively JR, Freund CL, et al. A Range of Clinical Phenotypes Associated with Mutations in CRX, a Photoreceptor Transcription-Factor Gene. Am J Hum Genet. 1998 Nov 1;63(5):1307–15.

355. Wu Y, Qiu J, Chen S, Chen X, Zhang J, Zhuang J, et al. Crx Is Posttranscriptionally Regulated by Light Stimulation in Postnatal Rat Retina. Front Cell Dev Biol [Internet]. 2020 [cited 2021 Jul 17];0. Available from: https://www.frontiersin.org/articles/10.3389/fcell.2020.00174/full

356. Dryja TP, Berson EL, Rao VR, Oprian DD. Heterozygous missense mutation in the rhodopsin gene as a cause of congenital stationary night blindness. Nat Genet. 1993 Jul;4(3):280–3.

357. Craft CM, Deming JD. Cone Arrestin: Deciphering the Structure and Functions of Arrestin 4 in Vision. In: Gurevich VV, editor. Arrestins - Pharmacology and Therapeutic Potential [Internet]. Berlin, Heidelberg: Springer; 2014 [cited 2021 Jul 17]. p. 117–31. (Handbook of Experimental Pharmacology). Available from: https://doi.org/10.1007/978-3-642-41199-1_6

358. Albertsen M, Teperek M, Elholm G, Füchtbauer EM, Lykke-Hartmann K. Localization and Differential Expression of the Krüppel-Associated Box Zinc Finger Proteins 1 and 54 in Early Mouse Development. DNA Cell Biol. 2010 Oct;29(10):589–601.

359. Fritsche LG, Igl W, Bailey JNC, Grassmann F, Sengupta S, Bragg-Gresham JL, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat Genet. 2016 Feb;48(2):134–43.

360. Elsheikh SSM, Chimusa ER, Mulder NJ, Crimi A. Genome-Wide Association Study of Brain Connectivity Changes for Alzheimer's Disease. Sci Rep. 2020 Jan 29;10(1):1433.

361. Shi Y, Li Y, Zhang D, Zhang H, Li Y, Lu F, et al. Exome Sequencing Identifies ZNF644 Mutations in High Myopia. PLOS Genet. 2011 Jun 9;7(6):e1002084.

362. Argüeso P, Guzman-Aranguez A, Mantelli F, Cao Z, Ricciuto J, Panjwani N. Association of Cell Surface Mucins with Galectin-3 Contributes to the Ocular Surface Epithelial Barrier *. J Biol Chem. 2009 Aug 21;284(34):23037–45.

363. Ugarte M, Pérez-Cerdá C, Rodríguez-Pombo P, Desviat LR, Pérez B, Richard E, et al. Overview of mutations in the PCCA and PCCB genes causing propionic acidemia. Hum Mutat. 1999;14(4):275–82.

364. Li X, Zhang J, Wang Y, Ji J, Yang F, Wan C, et al. Association study on the NAPG gene and bipolar disorder in the Chinese Han population. Neurosci Lett. 2009 Jul 3;457(3):159–62.

365. Zhao B, Shan Y, Yang Y, Yu Z, Li T, Wang X, et al. Transcriptome-wide association analysis of brain structures yields insights into pleiotropy with complex neuropsychiatric traits. Nat Commun. 2021 May 17;12(1):2878.

366. Bulayeva K, Lesch KP, Bulayev O, Walsh C, Glatt S, Gurgenova F, et al. Genomic structural variants are linked with intellectual disability. J Neural Transm. 2015 Sep 1;122(9):1289–301.

367. Goes FS, McGrath J, Avramopoulos D, Wolyniec P, Pirooznia M, Ruczinski I, et al. Genome-wide association study of schizophrenia in Ashkenazi Jews. Am J Med Genet B Neuropsychiatr Genet. 2015;168(8):649–59.

368. Han MR, Schellenberg GD, Wang LS, Alzheimer's Disease Neuroimaging Initiative. Genome-wide association reveals genetic effects on human Aβ42 and τ protein levels in cerebrospinal fluids: a case control study. BMC Neurol. 2010 Oct 8;10:90.

369. Cui L, Tao H, Wang Y, Liu Z, Xu Z, Zhou H, et al. A functional polymorphism of the microRNA-146a gene is associated with susceptibility to drug-resistant epilepsy and seizures frequency. Seizure. 2015 Apr 1;27:60–5.

370. Luo M, Ma W, Sand Z, Finlayson J, Wang T, Brinton RD, et al. Von Willebrand factor A domain-containing protein 8 (VWA8) localizes to the matrix side of the inner mitochondrial membrane. Biochem Biophys Res Commun. 2020 Jan 1;521(1):158–63.

371. Anney R, Klei L, Pinto D, Regan R, Conroy J, Magalhaes TR, et al. A genome-wide scan for common alleles affecting risk for autism. Hum Mol Genet. 2010 Oct 15;19(20):4072–82.

372. Greenwood TA, Akiskal HS, Akiskal KK, Bipolar Genome Study, Kelsoe JR. Genome-wide association study of temperament in bipolar disorder reveals significant associations with three novel Loci. Biol Psychiatry. 2012 Aug 15;72(4):303–10.

373. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet. 2018 Aug;50:1112–21.

374. Pitts MW, Byrns CN, Ogawa-Wong AN, Kremer P, Berry MJ. Selenoproteins in Nervous System Development and Function. Biol Trace Elem Res. 2014 Dec 1;161(3):231–45.

375. Dekker AM, Diekstra FP, Pulit SL, Tazelaar GHP, van der Spek RA, van Rheenen W, et al. Exome array analysis of rare and low frequency variants in amyotrophic lateral sclerosis. Sci Rep [Internet]. 2019 Apr 11 [cited 2021 Mar 25];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6459905/

376. Verheijen J, Van den Bossche T, van der Zee J, Engelborghs S, Sanchez-Valle R, Lladó A, et al. A comprehensive study of the genetic impact of rare variants in SORL1 in European early-onset Alzheimer's disease. Acta Neuropathol (Berl). 2016;132:213–24.

377. Liskova P, Dudakova L, Evans CJ, Rojas Lopez KE, Pontikos N, Athanasiou D, et al. Ectopic GRHL2 Expression Due to Non-coding Mutations Promotes Cell State Transition and Causes Posterior Polymorphous Corneal Dystrophy 4. Am J Hum Genet. 2018 Mar 1;102(3):447–59.

378. Zhang Z, Xiao X, Zhou W, Zhu D, Amos CI. False positive findings during genome-wide association studies with imputation: influence of allele frequency and imputation accuracy. Hum Mol Genet. 2021 Dec 17;31(1):146–55.

379. Maróti Z, Boldogkői Z, Tombácz D, Snyder M, Kalmár T. Evaluation of whole exome sequencing as an alternative to BeadChip and whole genome sequencing in human population genetic analysis. BMC Genomics. 2018 Oct 29;19(1):778.

380. Rymer J, Wildsoet CF. The role of the retinal pigment epithelium in eye growth regulation and myopia: a review. Vis Neurosci. 2005 Jun;22(3):251–61.

381. Wildsoet C, Wallman J. Choroidal and scleral mechanisms of compensation for spectacle lenses in chicks. Vision Res. 1995 May 1;35(9):1175–94.

382. Rada JAS, Shelton S, Norton TT. The sclera and myopia. Exp Eye Res. 2006 Feb;82(2):185–200.

383. Mutti DO, Zadnik K, Adams AJ. Myopia. The nature versus nurture debate goes on. Invest Ophthalmol Vis Sci. 1996 May;37(6):952–7.

384. Fan DSP, Lam DSC, Lam RF, Lau JTF, Chong KS, Cheung EYY, et al. Prevalence, incidence, and progression of myopia of school children in Hong Kong. Invest Ophthalmol Vis Sci. 2004 Apr;45(4):1071–5.

385. Iribarren R, Cortinez MF, Chiappe JP. Age of First Distance Prescription and Final Myopic Refractive Error. Ophthalmic Epidemiol. 2009 Jan 1;16(2):84–9.

386. Hysi PG, Choquet H, Khawaja AP, Wojciechowski R, Tedja MS, Yin J, et al. Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. Nat Genet. 2020 Apr;52(4):401–7.

387. Wojciechowski R, Hysi PG. Focusing In on the Complex Genetics of Myopia. PLoS Genet [Internet]. 2013 Apr 4 [cited 2018 Aug 15];9(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3617090/

388. Wong YL, Hysi P, Cheung G, Tedja M, Hoang QV, Tompson SWJ, et al. Genetic variants linked to myopic macular degeneration in persons with high myopia: CREAM Consortium. PloS One. 2019;14(8):e0220143.

389. Tideman JWL, Fan Q, Polling JR, Guo X, Yazar S, Khawaja A, et al. When do myopia genes have their effect? Comparison of genetic risks between children and adults. Genet Epidemiol. 2016;40(8):756–66.

390. Shah RL, Guggenheim JA, UK Biobank Eye and Vision Consortium. Genome-wide association studies for corneal and refractive astigmatism in UK Biobank demonstrate a shared role for myopia susceptibility loci. Hum Genet. 2018 Dec;137(11–12):881–96.

391. Fan Q, Guo X, Tideman JWL, Williams KM, Yazar S, Hosseini SM, et al. Childhood gene-environment interactions and age-dependent effects of genetic variants associated with refractive error and myopia: The CREAM Consortium. Sci Rep. 2016 May 13;6(1):25853.

392. Kiefer AK, Tung JY, Do CB, Hinds DA, Mountain JL, Francke U, et al. Genome-Wide Analysis Points to Roles for Extracellular Matrix Remodeling, the Visual Cycle, and Neuronal Development in Myopia. PLoS Genet [Internet]. 2013 Feb 28 [cited 2018 Aug 20];9(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3585144/

393. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. Transcriptional landscape of the prenatal human brain. Nature. 2014 Apr;508(7495):199–206.

394. Strunz T, Kiel C, Grassmann F, Ratnapriya R, Kwicklis M, Karlstetter M, et al. A mega-analysis of expression quantitative trait loci in retinal tissue. PLOS Genet. 2020 Sep 1;16(9):e1008934.

395. Singh K, Loreth D, Pöttker B, Hefti K, Innos J, Schwald K, et al. Neuronal Growth and Behavioral Alterations in Mice Deficient for the Psychiatric Disease-Associated Negr1 Gene. Front Mol Neurosci [Internet]. 2018 [cited 2021 Jan 26];11. Available from: https://www.frontiersin.org/articles/10.3389/fnmol.2018.00030/full#B66

396. Brugger SW, Gardner MC, Beales JT, Briggs F, Davis MF. Depression in multiple sclerosis patients associated with risk variant near NEGR1. Mult Scler Relat Disord. 2020 Nov;46:102537.

397. Eyers PA, Keeshan K, Kannan N. Tribbles in the 21st Century: The Evolving Roles of Tribbles Pseudokinases in Biology and Disease. Trends Cell Biol. 2017 Apr 1;27(4):284–98.

398. Xie J, Heim EN, Crite M, DiMaio D. TBC1D5-Catalyzed Cycling of Rab7 Is Required for Retromer-Mediated Human Papillomavirus Trafficking during Virus Entry. Cell Rep. 2020 Jun 9;31(10):107750.

399. Yasui D, Miyano M, Cai S, Varga-Weisz P, Kohwi-Shigematsu T. SATB1 targets chromatin remodelling to regulate genes over long distances. Nature. 2002 Oct;419(6907):641–5.

400. Hsia HE, Tüshaus J, Brummer T, Zheng Y, Scilabra SD, Lichtenthaler SF. Functions of 'A disintegrin and metalloproteases (ADAMs)' in the mammalian nervous system. Cell Mol Life Sci. 2019 Aug 1;76(16):3055–81.

401. Mandal M, Hamel KM, Maienschein-Cline M, Tanaka A, Teng G, Tuteja JH, et al. Histone reader BRWD1 targets and restricts recombination to the Igk locus. Nat Immunol. 2015 Oct;16(10):1094–103.

402. Hever AM, Williamson KA, van Heyningen V. Developmental malformations of the eye: the role of PAX6, SOX2 and OTX2. Clin Genet. 2006 Jun;69(6):459–70.

403. Almoallem B, Arno G, De Zaeytijd J, Verdin H, Balikova I, Casteels I, et al. The majority of autosomal recessive nanophthalmos and posterior microphthalmia can be attributed to biallelic sequence and structural variants in MFRP and PRSS56. Sci Rep. 2020 Jan 28;10(1):1289.

404. Singh K, Jayaram M, Kaare M, Leidmaa E, Jagomäe T, Heinla I, et al. Neural cell adhesion molecule Negr1 deficiency in mouse results in structural brain endophenotypes and behavioral deviations related to psychiatric disorders. Sci Rep. 2019 Apr 1;9(1):5457.

405. Funatsu N, Miyata S, Kumanogoh H, Shigeta M, Hamada K, Endo Y, et al. Characterization of a Novel Rat Brain Glycosylphosphatidylinositol-anchored Protein (Kilon), a Member of the IgLON Cell Adhesion Molecule Family *. J Biol Chem. 1999 Mar 19;274(12):8224–30.

406. Miyata S, Matsumoto N, Taguchi K, Akagi A, Iino T, Funatsu N, et al. Biochemical and ultrastructural analyses of iglon cell adhesion molecules, kilon and obcam in the rat brain. Neuroscience. 2003 Mar 31;117(3):645–58.

407. Noh K, Lee H, Choi TY, Joo Y, Kim SJ, Kim H, et al. Negr1 controls adult hippocampal neurogenesis and affective behaviors. Mol Psychiatry. 2019 Aug;24(8):1189–205.

408. Ni H, Xu M, Zhan GL, Fan Y, Zhou H, Jiang HY, et al. The GWAS Risk Genes for Depression May Be Actively Involved in Alzheimer's Disease. J Alzheimers Dis. 2018 Jan 1;64(4):1149–61.

409. Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. Nat Genet. 2016 Sep;48(9):1031–6.

410. Ripke S, Neale BM, Corvin A, Walters JT, Farh KH, Holmans PA, et al. Biological Insights From 108 Schizophrenia-Associated Genetic Loci. Nature. 2014 Jul 24;511(7510):421–7.

411. Dennis EL, Jahanshad N, Braskie MN, Warstadt NM, Hibar DP, Kohannim O, et al. Obesity gene NEGR1 associated with white matter integrity in healthy young adults. NeuroImage. 2014 Nov 15;102:548–57.

412. Genovese A, Cox DM, Butler MG. Partial Deletion of Chromosome 1p31.1 Including only the Neuronal Growth Regulator 1 Gene in Two Siblings. J Pediatr Genet. 2015 Mar;04(1):23–8.

413. Tassano E, Gamucci A, Celle ME, Ronchetto P, Cuoco C, Gimelli G. Clinical and Molecular Cytogenetic Characterization of a de novo Interstitial 1p31.1p31.3 Deletion in a Boy with Moderate Intellectual Disability and Severe Language Impairment. Cytogenet Genome Res. 2015;146(1):39–43.

414. Sagane K, Yamazaki K, Mizui Y, Tanaka I. Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23. Gene. 1999 Aug 5;236(1):79–86.

415. Takahashi E, Sagane K, Nagasu T, Kuromitsu J. Altered nociceptive response in ADAM11-deficient mice. Brain Res. 2006 Jun 30;1097(1):39–42.

416. Takahashi E, Sagane K, Oki T, Yamazaki K, Nagasu T, Kuromitsu J. Deficits in spatial learning and motor coordination in ADAM11-deficient mice. BMC Neurosci. 2006 Feb 26;7:19.

417. Rybnikova E, Kärkkäinen I, Pelto-Huikko M, Huovila APJ. Developmental regulation and neuronal expression of the cellular disintegrin ADAM11 gene in mouse nervous system. Neuroscience. 2002 Jul 19;112(4):921–34.

418. Yamagata A, Fukai S. Insights into the mechanisms of epilepsy from structural biology of LGI1–ADAM22. Cell Mol Life Sci. 2020 Jan 1;77(2):267–74.

419. Verhaaren BFJ, Debette S, Bis JC, Smith JA, Ikram MK, Adams HH, et al. Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. Circ Cardiovasc Genet. 2015 Apr;8(2):398–409.

420. Narboux-Nême N, Goïame R, Mattéi MG, Cohen-Tannoudji M, Wassef M. Integration of H-2Z1, a Somatosensory Cortex-Expressed Transgene, Interferes with the Expression of the Satb1 and Tbc1d5 Flanking Genes and Affects the Differentiation of a Subset of Cortical Interneurons. J Neurosci. 2012 May 23;32(21):7287–300.

421. Balamotis MA, Tamberg N, Woo YJ, Li J, Davy B, Kohwi-Shigematsu T, et al. Satb1 Ablation Alters Temporal Expression of Immediate Early Genes and Reduces Dendritic Spine Density during Postnatal Brain Development. Mol Cell Biol. 2012 Jan 15;32(2):333–47.

422. van der Meer D, Frei O, Kaufmann T, Shadrin AA, Devor A, Smeland OB, et al. Understanding the genetic determinants of the brain with MOSTest. Nat Commun. 2020 Jul 14;11:3512.

423. Periyasamy S, John S, Padmavati R, Rajendren P, Thirunavukkarasu P, Gratten J, et al. Association of Schizophrenia Risk With Disordered Niacin Metabolism in an Indian Genome-wide Association Study. JAMA Psychiatry. 2019 Oct;76(10):1026–34.

424. Peyrot WJ, Price AL. Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. Nat Genet. 2021 Apr 1;53(4):445–54.

425. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020 Feb 6;180(3):568-584.e23.

426. Coleman JRI, Peyrot WJ, Purves KL, Davis KAS, Rayner C, Choi SW, et al. Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. Mol Psychiatry. 2020 Jul;25(7):1430–46.

427. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al. Multivariate genome-wide analyses of the well-being spectrum. Nat Genet. 2019 Mar;51(3):445–51.

428. Okbay A, Baselmans BML, De Neve JE, Turley P, Nivard MG, Fontana MA, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. Nat Genet. 2016 Jun;48(6):624–33.

429. Fajardo D, Vinasco K, Montoya JC, Satizabal JM, Sanchez A, GarcÃ-a-Vallejo F. Complex networks of interaction of genes located in the critical region of Down syndrome expressed in the normal human brain. Biomed Res [Internet]. 2018 [cited 2021 Jul 21];29(18). Available from: https://www.alliedacademies.org/abstract/complex-networks-of-interaction-of-genes-located-in-the-critical-region-of-down-syndrome-expressed-in-the-normal-human-brain-10859.html

430. Farrelly LA, Maze I. An emerging perspective on 'histone code' mediated regulation of neural plasticity and disease. Curr Opin Neurobiol. 2019 Dec 1;59:157–63.

431. Davies G, Armstrong N, Bis JC, Bressler J, Chouraki V, Giddaluru S, et al. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N =53 949). Mol Psychiatry. 2015 Feb;20(2):183–92.

432. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet. 2018 Jul;50(7):912–9.

433. Fang Y, Zekiy AO, Ghaedrahmati F, Timoshin A, Farzaneh M, Anbiyaiee A, et al. Tribbles homolog 2 (Trib2), a pseudo serine/threonine kinase in tumorigenesis and stem cell fate decisions. Cell Commun Signal. 2021 Apr 1;19(1):41.

434. Dobens LL, Bouyain S. Developmental roles of tribbles protein family members. Dev Dyn. 2012;241(8):1239–48.

435. Zhang Y, Davis JL, Li W. Identification of tribbles homolog 2 as an autoantigen in autoimmune uveitis by phage display. Mol Immunol. 2005 Jul 1;42(11):1275–81.

436. Deng Y, Qiao L, Du M, Qu C, Wan L, Li J, et al. Age-related macular degeneration: Epidemiology, genetics, pathophysiology, diagnosis, and targeted therapy. Genes Dis [Internet]. 2021 Feb 27 [cited 2021 Jul 21]; Available from: https://www.sciencedirect.com/science/article/pii/S2352304221000295

437. Choquet H, Melles RB, Anand D, Yin J, Cuellar-Partida G, Wang W, et al. A multiethnic GWAS meta-analysis of 585,243 individuals identifies new risk loci associated with cataract and reveals sex-specific effects. medRxiv. 2020 Sep 24;2020.09.23.20200428.

438. Liu Y, Allingham RR. Major review: Molecular genetics of primary open-angle glaucoma. Exp Eye Res. 2017 Jul 1;160:62–84.

439. Paylakhi SH, Yazdani S, April C, Fan JB, Moazzeni H, Ronaghi M, et al. Non-housekeeping genes expressed in human trabecular meshwork cell cultures. Mol Vis. 2012 Jan 28;18:241–54.

440. H S, A M, Pg H, P G, R H, Cc K, et al. Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. Genet Epidemiol. 2015 Jan 28;39(3):207–16.

441. Plotnikov D, Shah RL, Rodrigues JN, Cumberland PM, Rahi JS, Hysi PG, et al. A commonly occurring genetic variant within the NPLOC4–TSPAN10–PDE6G gene cluster is associated with the risk of strabismus. Hum Genet. 2019 Jul 1;138(7):723–37.

442. Tsai TH, Liu YL, Ma IH, Su CC, Lin CW, Lin LLK, et al. Evolution of the Prevalence of Myopia among Taiwanese Schoolchildren: A Review of Survey Data from 1983 through 2017. Ophthalmology. 2021 Feb 1;128(2):290–301.

443. Li Y, Liu J, Qi P. The increasing prevalence of myopia in junior high school students in the Haidian District of Beijing, China: a 10-year population-based survey. BMC Ophthalmol. 2017 Jun 12;17(1):88.

444. Dong L, Kang YK, Li Y, Wei WB, Jonas JB. PREVALENCE AND TIME TRENDS OF MYOPIA IN CHILDREN AND ADOLESCENTS IN CHINA: A Systemic Review and Meta-Analysis. Retina Phila Pa. 2020 Mar;40(3):399–411.

445. French AN, Morgan IG, Burlutsky G, Mitchell P, Rose KA. Prevalence and 5- to 6-year incidence and progression of myopia and hyperopia in Australian schoolchildren. Ophthalmology. 2013 Jul;120(7):1482–91.

446. Harrington SC, Stack J, Saunders K, O'Dwyer V. Refractive error and visual impairment in Ireland schoolchildren. Br J Ophthalmol. 2019 Aug;103(8):1112–8.

447. Dayan YB, Levin A, Morad Y, Grotto I, Ben-David R, Goldberg A, et al. The Changing Prevalence of Myopia in Young Adults: A 13-Year Series of Population-Based Prevalence Surveys. Invest Ophthalmol Vis Sci. 2005 Aug 1;46(8):2760–5.

448. Sankaridurg P, Tahhan N, Kandel H, Naduvilath T, Zou H, Frick KD, et al. IMI Impact of Myopia. Invest Ophthalmol Vis Sci. 2021 Apr 28;62(5):2–2.

449. Saw SM, Goh PP, Cheng A, Shankar A, Tan DTH, Ellwein LB. Ethnicity-specific prevalences of refractive errors vary in Asian children in neighbouring Malaysia and Singapore. Br J Ophthalmol. 2006 Oct 1;90(10):1230–5.

450. Kaštelan S, Gverović-Antunica A, Pelčić G, Gotovac M, Marković I, Kasun B. Refractive Changes Associated with Diabetes Mellitus. Semin Ophthalmol. 2018 Nov 17;33(7–8):838–45.

451. Zheng W, Dryja TP, Wei Z, Song D, Tian H, Kahler KH, et al. Systemic Medication Associations with Presumed Advanced or Uncontrolled Primary Open-Angle Glaucoma. Ophthalmology. 2018 Jul 1;125(7):984–93.

452. Khawaja AP, Chan MPY, Broadway DC, Garway-Heath DF, Luben R, Yip JLY, et al. Systemic Medication and Intraocular Pressure in a British Population: The EPIC-Norfolk Eye Study. Ophthalmology. 2014 Aug 1;121(8):1501–7.

453. Pappelis K, Loiselle AR, Visser S, Jansonius NM. Association of Systemic Medication Exposure With Glaucoma Progression and Glaucoma Suspect Conversion in the Groningen Longitudinal Glaucoma Study. Invest Ophthalmol Vis Sci. 2019 Nov 1;60(14):4548–55.

454. Bery A, Bagchi U, Bergen AA, Felder-Schmittbuhl MP. Circadian clocks, retinogenesis and ocular health in vertebrates: new molecular insights. Dev Biol. 2022 Apr 1;484:40–56.

455. Dyer MA. The Homeodomain Proteins Prox1, Six3 and Chx10 Regulate Proliferation, Cell Fate Specification and Differentiation in the Developing Retina. Cell Cycle. 2003 Jul 1;2(4):347–54.

456. Gaston-Massuet C, Andoniadou CL, Signore M, Sajedi E, Bird S, Turner JMA, et al. Genetic interaction between the homeobox transcription factors HESX1 and SIX3 is required for normal pituitary development. Dev Biol. 2008 Dec 15;324(2):322–33.

457. Kelberman D, Rizzoti K, Lovell-Badge R, Robinson ICAF, Dattani MT. Genetic Regulation of Pituitary Gland Development in Human and Mouse. Endocr Rev. 2009 Dec;30(7):790–829.

458. Hoffmann HM, Meadows JD, Breuer JA, Yaw AM, Nguyen D, Tonsfeldt KJ, et al. The transcription factors SIX3 and VAX1 are required for suprachiasmatic nucleus circadian output and fertility in female mice. J Neurosci Res. 2021;99(10):2625–45.

459. Rong SS, Lu SY, Matsushita K, Huang C, Leung CKS, Kawashima R, et al. Association of the SIX6 locus with primary open angle glaucoma in southern Chinese and Japanese. Exp Eye Res. 2019 Mar 1;180:129–36.

460. Huang L, Xiao X, Li S, Jia X, Wang P, Guo X, et al. CRX variants in cone–rod dystrophy and mutation overview. Biochem Biophys Res Commun. 2012 Oct 5;426(4):498–503.

461. Mansergh FC, Carrigan M, Hokamp K, Farrar GJ. Gene expression changes during retinal development and rod specification. Mol Vis. 2015 Jan 20;21:61–87.

462. Hu M, Zhong H, Liu Y, Liu R, Chen C, Chu R. [Expression of rhodopsin for experimental myopia of form-deprivation and defocus in guinea pig]. Yan Ke Xue Bao 2016. 2008 Mar 1;24(1):1–5.

463. Sakamoto K, Oishi K, Okada T, Onuma Y, Yokoyama K, Sugimoto K, et al. Molecular cloning of the cone-rod homeobox gene (Crx) from the rat and its temporal expression pattern in the retina under a daily light–dark cycle. Neurosci Lett. 1999 Feb 12;261(1):101–4.

464. Rohde K, Rovsing L, Ho AK, Møller M, Rath MF. Circadian Dynamics of the Cone-Rod Homeobox (CRX) Transcription Factor in the Rat Pineal Gland and Its Role in Regulation of Arylalkylamine N-Acetyltransferase (AANAT). Endocrinology. 2014 Aug 1;155(8):2966–75.

465. Wojciechowski R, Cheng CY. Involvement of multiple molecular pathways in the genetics of ocular refraction and myopia. Retina Phila Pa. 2018 Jan;38(1):91–101.

466. Yang Y, Wu J, Wu D, Wei Q, Zhong T, Yang J, et al. Intravitreal brimonidine inhibits form-deprivation myopia in guinea pigs. Eye Vis. 2021 Jul 14;8(1):27.

467. Liu Y, Wang Y, Lv H, Jiang X, Zhang M, Li X. α-adrenergic agonist brimonidine control of experimentally induced myopia in guinea pigs: A pilot study. Mol Vis. 2017;23:785–98.

468. Carr BJ, Nguyen CT, Stell WK. Alpha2-adrenoceptor agonists inhibit form-deprivation myopia in the chick. Clin Exp Optom. 2019 Jul 1;102(4):418–25.

469. Carr BJ, Mihara K, Ramachandran R, Saifeddine M, Nathanson NM, Stell WK, et al. Myopia-Inhibiting Concentrations of Muscarinic Receptor Antagonists Block Activation of Alpha2A-Adrenoceptors In Vitro. Invest Ophthalmol Vis Sci. 2018 Jun 1;59(7):2778–91.

470. Näreoja K, Kukkonen JP, Rondinelli S, Toivola DM, Meriluoto J, Näsman J. Adrenoceptor activity of muscarinic toxins identified from mamba venoms. Br J Pharmacol. 2011;164(2b):538–50.

471. El-Nimri NW, Wildsoet CF. Effects of Topical Latanoprost on Intraocular Pressure and Myopia Progression in Young Guinea Pigs. Invest Ophthalmol Vis Sci. 2018 May 1;59(6):2644–51.

472. Tokoro T, Funata M, Akazawa Y. Influence of intraocular pressure on axial elongation. J Ocul Pharmacol. 1990;6(4):285–91.

473. Leydolt C, Findl O, Drexler W. Effects of change in intraocular pressure on axial eye length and lens position. Eye. 2008 May;22(5):657–61.

474. Edwards MH, Brown B. IOP in myopic children: the relationship between increases in IOP and the development of myopia. Ophthalmic Physiol Opt J Br Coll Ophthalmic Opt Optom. 1996 May;16(3):243–6.

475. Shen M, Fan F, Xue A, Wang J, Zhou X, Lu F. Biomechanical properties of the cornea in high myopia. Vision Res. 2008 Sep;48(21):2167–71.

476. Genest R, Chandrashekar N, Irving E. The effect of intraocular pressure on chick eye geometry and its application to myopia. Acta Bioeng Biomech. 2012;14(2):3–8.

477. Saeedi O, Pillar A, Jefferys J, Arora K, Friedman D, Quigley H. Change in choroidal thickness and axial length with change in intraocular pressure after trabeculectomy. Br J Ophthalmol. 2014 Jul;98(7):976–9.

478. Yan L, Huibin L, Xuemin L. Accommodation-induced intraocular pressure changes in progressing myopes and emmetropes. Eye. 2014 Nov;28(11):1334–40.

479. Fillingim R, Wallace M, Herbstman D, Ribeiro-Dasilva M, Staud R. Genetic contributions to pain: a review of findings in humans. Oral Dis. 2008 Nov;14(8):673–82.

480. Zorina-Lichtenwalter K, Meloto CB, Khoury S, Diatchenko L. Genetic predictors of human chronic pain conditions. Neuroscience. 2016 Dec 3;338:36–62.

481. Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, et al. Genome-wide association study of multisite chronic pain in UK Biobank. PLOS Genet. 2019 Jun 13;15(6):e1008164.

482. Bountra C, Munglani R, Schmidt WK. Pain: Current Understanding, Emerging Therapies, and Novel Approaches to Drug Discovery. CRC Press; 2003. 1006 p.

483. Wu SY, Yoo YJ, Nemesure B, Hennis A, Leske MC. Nine-Year Refractive Changes in the Barbados Eye Studies. Invest Ophthalmol Vis Sci. 2005 Nov 1;46(11):4032–9.

484. Tarczy-Hornoch K, Ying-Lai M, Varma R, Group the LALES. Myopic Refractive Error in Adult Latinos: The Los Angeles Latino Eye Study. Invest Ophthalmol Vis Sci. 2006 May 1;47(5):1845–52.

485. Fledelius HC, Miyamoto K. Diabetic myopia - is it lens-induced?: An oculometric study comprising ultrasound measurements. Acta Ophthalmol (Copenh). 2009 May 27;65(4):469–73.

486. Pittasch D, Lobmann R, Behrens-Baumann W, Lehnert H. Pupil Signs of Sympathetic Autonomic Neuropathy in Patients With Type 1 Diabetes. Diabetes Care. 2002 Sep 1;25(9):1545–50.

487. Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. Nat Rev Immunol. 2011 Feb;11(2):98–107.

488. King PT. Inflammation in chronic obstructive pulmonary disease and its role in cardiovascular disease and lung cancer. Clin Transl Med. 2015 Jul 29;4:26.

489. Fledelius H, Zak M, Pedersen FK. Refraction in juvenile chronic arthritis: a long-term follow-up study, with emphasis on myopia. Acta Ophthalmol Scand. 2001 Jun;79(3):237–9.

490. Ayazi S, Armstrong WL, Weinstein A. Transient high myopia in systemic lupus erythematosus associated with anasarca. Ann Ophthalmol. 1982 Aug 1;14(8):785–7.

491. Kamath Y, Singh A, Bhat S, Sripathi H. Acute onset myopia as a presenting feature of systemic lupus erythematosus. J Postgrad Med. 2013 Jul;59(3):245–6.

492. Shu U, Takeuchi F, Tanimoto K, Moroi Y, Uchida K, Ito K. Transient myopia with severe chemosis associated with exacerbation of disease activity in systemic lupus erythematosus. J Rheumatol. 1992 Feb 1;19(2):297–301.

493. Boone MI, Moore TL, Cruz OA. Screening for Uveitis in Juvenile Rheumatoid Arthritis. J Pediatr Ophthalmol Strabismus. 1998 Jan;35(1):41–3.

494. Kesen MR, Setlur V, Goldstein DA. Juvenile Idiopathic Arthritis-related Uveitis. Int Ophthalmol Clin. 2008 Summer;48(3):21–38.

495. Malleson P. Prevalence and outcome of uveitis in a regional cohort of patients with juvenile rheumatoid arthritis. J Rheumatol. 1998 Jun 1;25(6):1242.

496. Chen W, Zhao B, Jiang R, Zhang R, Wang Y, Wu H, et al. Cytokine Expression Profile in Aqueous Humor and Sera of Patients with Acute Anterior Uveitis. Curr Mol Med. 2015 Jul 1;15(6):543–9.

497. Wang Y, Tang Z, Xue R, Singh GK, Lv Y, Shi K, et al. TGF-β1 promoted MMP-2 mediated wound healing of anterior cruciate ligament fibroblasts through NF-κB. Connect Tissue Res. 2011 Jun 1;52(3):218–25.

498. Lin Z, Chen X, Ge J, Cui D, Wu J, Tang F, et al. Effects of Direct Intravitreal Dopamine Injection on Sclera and Retina in Form-Deprived Myopic Rabbits. J Ocul Pharmacol Ther. 2008 Dec;24(6):543–50.

499. Wang Q, Zhao G, Xing S, Zhang L, Yang X. Role of bone morphogenetic proteins in form-deprivation myopia sclera. Mol Vis. 2011 Mar 8;17:647–57.

500. Aït-Khaled N, Pearce N, Anderson HR, Ellwood P, Montefort S, Shah J, et al. Global map of the prevalence of symptoms of rhinoconjunctivitis in children: The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three. Allergy. 2009 Jan;64(1):123–48.

501. Gordon-Shaag A, Millodot M, Shneor E, Liu Y. The genetic and environmental factors for keratoconus. BioMed Res Int. 2015;2015:795738.

502. Shafer BM, Qiu M, Rapuano CJ, Shields CL. Association Between Hay Fever and High Myopia in United States Adolescents and Adults. Eye Contact Lens. 2017 May;43(3):186–91.

503. Danescu LG, Levy S, Levy J. Vitamin D and diabetes mellitus. Endocrine. 2009 Feb 1;35(1):11–7.

504. Choi JA, Han K, Park YM, La TY. Low serum 25-hydroxyvitamin D is associated with myopia in Korean adolescents. Invest Ophthalmol Vis Sci. 2014 Apr 3;55(4):2041–7.

505. Yazar S, Hewitt AW, Black LJ, McKnight CM, Mountain JA, Sherwin JC, et al. Myopia is associated with lower vitamin D status in young adults. Invest Ophthalmol Vis Sci. 2014 Jun 26;55(7):4552–9.

506. Guggenheim JA, Williams C, Northstone K, Howe LD, Tilling K, St Pourcain B, et al. Does vitamin D mediate the protective effects of time outdoors on myopia? Findings from a prospective birth cohort. Invest Ophthalmol Vis Sci. 2014 Nov 18;55(12):8550–8.

507. Cuellar-Partida G, Williams KM, Yazar S, Guggenheim JA, Hewitt AW, Williams C, et al. Genetically low vitamin D concentrations and myopic refractive error: a Mendelian randomization study. Int J Epidemiol. 2017 Dec 1;46(6):1882–90.

508. Liu YQ, Liu Y, Chen ZY, Li H, Xiao T. Rheumatoid arthritis and osteoporosis: a bi-directional Mendelian randomization study. Aging. 2021 May 18;13(10):14109–30.

509. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci. 2009 Jun 9;106(23):9362–7.

510. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet. 2018 Apr;50(4):559–71.

511. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell. 2017 Jun 15;169(7):1177–86.

512. Helene C, David M. Genetics of Obesity: What have we Learned? Curr Genomics. 2011 May 1;12(3):169–79.

513. Yang J, Lee T, Kim J, Cho MC, Han BG, Lee JY, et al. Ubiquitous Polygenicity of Human Complex Traits: Genome-Wide Analysis of 49 Traits in Koreans. PLOS Genet. 2013 Mar 7;9(3):e1003355.

514. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009 Oct;461(7265):747–53.

515. Stephens M, Balding DJ. Bayesian statistical methods for genetic association studies. Nat Rev Genet. 2009 Oct;10(10):681–90.

516. Paré G, Asma S, Deng WQ. Contribution of Large Region Joint Associations to Complex Traits Genetics. PLOS Genet. 2015 Apr 9;11(4):e1005103.

517. Szymczak S, Biernacka JM, Cordell HJ, González-Recio O, König IR, Zhang H, et al. Machine learning in genome-wide association studies. Genet Epidemiol. 2009;33(S1):S51–7.

518. Shen H, Feldman MW. Genetic nurturing, missing heritability, and causal analysis in genetic statistics. Proc Natl Acad Sci. 2020 Oct 13;117(41):25646–54.

519. Wei WH, Hemani G, Haley CS. Detecting epistasis in human complex traits. Nat Rev Genet. 2014 Nov;15(11):722–33.

520. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J Hum Genet. 2017 Jul 6;101(1):5–22.

521. Wang H, Zhang F, Zeng J, Wu Y, Kemper KE, Xue A, et al. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. Sci Adv. 2019 Aug 1;5(8):eaaw3538.

522. Mat Ripen A, Ghani H, Chear CT, Chiow MY, Syed Yahya SNH, Kassim A, et al. Whole exome sequencing identifies compound heterozygous variants of CR2 gene in monozygotic twin patients with common variable immunodeficiency. SAGE Open Med. 2020 May 22;8:2050312120922652.

523. Das S, Abecasis GR, Browning BL. Genotype Imputation from Large Reference Panels. Annu Rev Genomics Hum Genet. 2018 Aug 31;19(1):73–96.

524. Lee S, Teslovich TM, Boehnke M, Lin X. General Framework for Meta-analysis of Rare Variants in Sequencing Association Studies. Am J Hum Genet. 2013 Jul 11;93(1):42–53.

525. Zhang M, Gazzard G, Fu Z, Li L, Chen B, Saw SM, et al. Validating the Accuracy of a Model to Predict the Onset of Myopia in Children. Invest Ophthalmol Vis Sci. 2011 Jul 1;52(8):5836–41.

526. Bachtiar M, Ooi BNS, Wang J, Jin Y, Tan TW, Chong SS, et al. Towards precision medicine: interrogating the human genome to identify drug pathways associated with potentially functional, population-differentiated polymorphisms. Pharmacogenomics J. 2019 Dec;19(6):516–27.

527. Parmeggiani F, Gemmati D, Costagliola C, Semeraro F, D'Angelo S, Perri P, et al. Impact of Coagulation-Balance Gene Predictors on Efficacy of Photodynamic Therapy for Choroidal Neovascularization in Pathologic Myopia. Ophthalmology. 2010 Mar 1;117(3):517–23.

528. Tkatchenko TV, Tkatchenko AV. Pharmacogenomic Approach to Antimyopia Drug Development: Pathways Lead the Way. Trends Pharmacol Sci. 2019 Nov 1;40(11):833–52.