Title:

Genetic prediction of myopia: Prospects and challenges

Authors:

Jeremy A. Guggenheim¹, Neema Ghorbani Mojarrad¹, Cathy Williams², D. Ian Flitcroft³

Affiliations:

¹ School of Optometry & Vision Sciences, Cardiff University, Cardiff, UK.
² School of Social and Community Medicine, University of Bristol, Bristol, UK.
³ Department of Ophthalmology, Children’s University Hospital, Dublin, Ireland.

Corresponding author:

Jeremy A. Guggenheim, PhD
School of Optometry & Vision Sciences, Cardiff University
Maindy Road, Cardiff, CF24 4HQ, UK
Tel +44 (0) 29 2087 4904
Fax +44 (0) 29 2087 4859
Email. GuggenheimJ1@cardiff.ac.uk

Running title:

Genetic prediction of myopia

Keywords:

Refractive error, myopia, genetic prediction, genome-wide association study, personalised medicine

Disclosure:

The authors report no conflicts of interest and have no proprietary interest in any of the materials mentioned in this article

Acknowledgements:

NGM is funded by a PhD Studentship grant from the College of Optometrists entitled “Genetic prediction of individuals at-risk for myopia development”. CW is funded by an NIHR Senior Research Fellowship award SRF-2015-08-005.
**Background**

Appeals have been made for eye care professionals to start prescribing anti-myopia therapies as part of their routine management of myopic children\(^1\)\(^\text{-}^3\). These calls are fuelled by two key considerations. Firstly, that interventions to slow myopia progression have shown success in randomized controlled trials (RCTs)\(^4\)\(^\text{-}^7\), and secondly, appreciation that the risk of sight-threatening complications rises dose-dependently with the level of myopia\(^8\)\(^,^9\). Notwithstanding existing gaps in knowledge regarding the efficacy of current treatments (see below), these considerations argue that myopia control interventions should be widely adopted, and that they should be instigated at an early age – especially in children most at risk – in order to reduce the final level of myopia. Therefore in managing a child with myopia, an eye care professional would have to decide not only which therapy to recommend, but at what age to start treatment. In this review we discuss the future role of genetic prediction in helping clinicians treat myopia.

**When might genetic prediction be useful?**

We envisage three situations when genetic prediction might be helpful to clinicians; the “personalized medicine” scenario, the “prophylactic intervention” scenario, and the “prevention of high myopia” scenario.

While the average response to the best anti-myopia therapy reported is approximately 40-50%, there is typically a wide variability in response between patients\(^5\)\(^\text{-}^7\). If a child’s genetic profile makes them a better responder to one type of therapy compared to another, e.g. orthokeratology vs. atropine, then clinicians could tailor treatments to individual patients who would benefit the most. This concept of personalized medicine has been around for a long time outside the field of myopia yet has only recently started to make a difference to clinical practice – most notably in cancer care\(^10\). To date there is no research evidence supporting the use of genetic testing in choosing the most suitable anti-myopia therapy for an individual child, but one study has shown how a specific myopia risk allele and a specific environmental risk factor (time spent reading) can interact to markedly elevate susceptibility\(^11\), hinting that the personalized medicine approach may have merit. A comprehensive evidence base for personalizing anti-myopia therapy is unlikely to be forthcoming in the short to medium term because to achieve the necessary statistical power such studies would need to enrol thousands of patients and, ideally, follow highly standardized treatment protocols. If future intervention trials routinely genotyped subjects, then such data could be derived through individual patient meta-analysis.
What is the case for prophylactic intervention? Here again there is a lack of research evidence showing that the currently-available treatment options work when initiated prior to myopia onset – with the single exception of increased time outdoors, which has been shown in 2 rigorous, RCTs to reduce incident myopia by a clinically significant degree\textsuperscript{12, 13}. So should eye care professionals be “prescribing” more time outdoors to children at high risk of developing myopia? Any benefit of genetic prediction for such a simple intervention appears minimal; after all, spending more time outdoors provides health benefits for a child beyond refractive development and poses minimal adverse effects (if the dangers of excessive sun exposure are mitigated by recommending that children wear appropriate sunglasses, sun cream, and a hat). This argues that prescribing extra time outdoors should be done routinely, though in the real world, prescribing extra time outdoors may not work as effectively as other possible interventions, since spending the (currently unknown) necessary number of hours outdoors every day may not fit with many families’ lifestyles. However, knowing that a child has a high genetic predisposition to myopia might give parents an extra incentive to increase their child’s time outdoors. Ultimately such interventions are likely to be worthwhile if they result in a lower degree of myopia, but final refraction was not significantly altered in the largest randomized trial to date of additional outdoor time\textsuperscript{12}.

The two types of intervention that have shown the most impact on myopia progression so far are pharmacological (atropine at various doses) and optical (multifocal contact lenses and orthokeratology) approaches\textsuperscript{4-7}. Therefore, if new research emerges showing that one or more of the known pharmacological or optical treatment interventions for myopia is effective when used prophylactically, then identifying children at high risk of myopia would become useful to clinicians and parents in making decisions about who to treat. Since pre-myopic subjects would not require optical correction, pharmacological treatments might appear more suitable for this application. Such an approach currently lacks appropriate clinical evidence, but securing research evidence about the efficacy or otherwise of prescribing existing anti-myopia therapies prophylactically could be readily achieved by means of small-scale clinical trials designed to address this specific issue.

The other plausible scenario where genetic prediction may play a useful role is in the avoidance of high myopia. In later life, eyes with longer axial lengths (>26 mm) have a significantly higher risk of uncorrectable visual loss\textsuperscript{14}. The strongest case for treatment of myopia progression therefore exists in
patients destined to have a final refraction more myopic than -6.00 D or an axial length > 26 mm. Enhanced genetic prediction of final refraction in a progressing myope may be very valuable in deciding whether treatment over many years with pharmacological or optical means is worthwhile in terms of both potential risks and costs.

**Is genetic prediction any better than prediction based on existing known risk factors?**

Researchers from the CLEERE study group\(^{15}\) recently evaluated which risk factors were most useful in predicting children at-risk of incident myopia, using data from a large, multi-centre longitudinal study of U.S. children. They considered 13 potential predictive factors, including spherical equivalent refractive error at baseline and parentally-reported number of myopic parents. Spherical equivalent at baseline was the single best predictive factor, and indeed this predictive factor alone performed as well, essentially, as any other combination of predictive factors. Notably, parental myopia (i.e. having 0, 1 or 2 myopic parents) was not a useful predictor once spherical equivalent refractive error at baseline was known. A predictive model for incident myopia using current refractive error and parental myopia as risk factors performed with a sensitivity and specificity of 62.5% and 81.9% in CLEERE participants\(^{16}\).

At first glance the CLEERE study’s findings\(^{15}\) would seem to imply that genetic prediction is probably not worth the trouble, as it would offer no benefit over-and-above measuring cycloplegic refractive error in early childhood. However, there are 3 reasons to expect that – at least in theory – genetic prediction will be a better predictive factor for myopia than number of myopic parents. Firstly, for a highly polygenic trait such as refractive error, parental myopia is only a good predictor of the refractive error of the parents’ children on average, not the refractive error of any individual child. Secondly, parental myopia uses a binary scale, which means that parents’ actual levels of myopia are ignored in the predictive modelling process. Thirdly, parental myopia is not well suited to predicting the future severity of refractive error in an individual, which means that the risk of future pathological complications will remain uncertain. While this limitation could be mitigated by predicting the risk of high myopia rather than myopia, there would still be benefits to knowing whether an individual was likely to attain a refractive error in adulthood of, for example, -6.50 D or -20.00 D.

To illustrate the first point mentioned above, Figure 1 shows the relationship between the refractive error of parents and their children using computer simulations performed under the assumption that refractive error is a purely polygenic trait (details of the simulation parameters are available from the...
authors on request). For heritability values across the range 25% to 75%, which is typical of the range reported for refractive error and myopia\textsuperscript{17-19}, it is evident that while the mean refractive error of parents is a good predictor of the average refractive error of their children (blue line of best fit in the graphs) it is only a crude guide to the refractive error of individual children. To illustrate the second point mentioned above, we generated simulations of a larger sample of 2-parent-plus-1-child families and evaluated the efficacy of 3 different predictive factors: (a) number of myopic parents, (b) the average refractive error of the parents (called the “midparent” value in the genetics literature), and (c) the precise genetic risk of the child. The results are shown in Figures 2 and 3. It is apparent that the loss of information inherent when simply counting the number of myopic parents adversely impacts the ability to predict the refractive error of children – at least under the conditions used in the computer simulation. However, it is also clear that “perfect” genetic prediction is a very much better predictive factor than the midparent value, capturing essentially all of the available variability (as governed in the simulation by the level of heritability). In reality it is not possible to predict the genetic risk of an individual with 100% precision, and in this sense the results of genetic prediction shown in Figure 2 are misleading. However, as we discuss further below, they do offer a glimpse of what genetic prediction has to offer.

**How effective is the genetic prediction of refractive error currently?**

Very few studies have tested the current efficacy of genetic prediction for refractive error. In those that have, the *variance in refractive error explained* – i.e. the coefficient of determination ($r^2$ value) from a linear regression of “true” refractive error (y-axis) on “predicted” refractive error (x-axis) – was used as the measure of predictive performance. $R^2$ values are well-suited to quantifying the efficacy of continuous traits compared to statistics such as the positive/negative predictive value or area under the receiver-operating curve (AUC) commonly used for classifying individuals as either affected or unaffected by a disorder\textsuperscript{20}. The CREAM consortium\textsuperscript{21} identified 26 genetic variants associated with refractive error in a genome-wide association study (GWAS) and reported that a “polygenic risk score” (PGRS) comprised of these variants explained 3.4% of the phenotypic variation in refractive error. Fan *et al.*\textsuperscript{22} used a simpler approach – an “allele score” calculated by counting the number of “high risk” alleles carried by individuals for 39 genetic variants identified from the GWAS studies of the CREAM consortium\textsuperscript{21} and Kiefer *et al.*\textsuperscript{23} – and were able to explain 2.3% of the variation in refractive error for a sample of 15-year old children. Thus, using a PGRS or allele score comprised of the small numbers of genetic variants that meet the extremely stringent threshold of “genome-wide significance” in a current GWAS provides poor predictive performance ($r^2<0.05$).
As shown in Figure 4, the performance of a genetic prediction algorithm varies primarily depending on (1) a property of the trait known as the SNP-heritability ($h^2_{SNP}$), which is closely related to other measures of heritability, and (2) the sample size of the genome-wide association study (GWAS) used to build the predictive model\textsuperscript{20}. For refractive error, $h^2_{SNP}$ has been estimated\textsuperscript{24} at 0.35 and the largest GWAS for refractive error to date\textsuperscript{21, 23} had a sample size of approximately 45,000 participants. This suggests that including a greater number of genetic variants in a PGRS than the 26-39 that have hitherto been used should yield a PGRS model with a predictive accuracy of approximately 10% (if the assumptions relating to Figure 4 hold).

Refraction is the end result of a complex growth process\textsuperscript{25}, and the value of GWAS data may be enhanced by asking more specific questions related to different aspects of refractive development. Evaluating genetic associations with factors known to be associated with higher rates of final myopia (i.e. age of onset, initial rate of progression and duration of progression), or looking for genetic associations with a final axial length in excess of 26 mm may further enhance the genetic predictability of final refraction.

**How effective will genetic prediction of refractive error be in the future?**

Figure 4 gives an indication of the prospects for genetic prediction of refractive error in the future. As GWAS meta-analysis sample sizes approach 500,000 – 1,000,000 it should be possible to estimate the parameters in a PGRS or an equivalent Bayesian model with sufficient precision that prediction accuracy ($r^2$) in independent samples approaches $h^2_{SNP}$. Thus, for refractive error, prediction accuracy should approach 35% ($r^2=0.35$). (Note from Figure 3 that, by contrast, number of myopic parents has a maximum predictive accuracy of approximately $r^2=0.05$ if the true heritability of refractive error is $h^2=0.5$). An important caveat here, however, is that prediction accuracy is likely to be population-specific, like heritability estimates themselves. Thus, prediction accuracy would be expected to be worse when a prediction model is developed and tested in individuals of differing ancestry\textsuperscript{26}. For refractive error prediction this could be highly disadvantageous, since most large GWAS analyses are carried out in individuals of European ancestry (due to high investment in population-based genetics research in Europe and the United States) whereas the most urgent need for genetic prediction of myopia is in South-East Asia\textsuperscript{27}. It may also be the case that genetic prediction may be suboptimal when the prediction model is developed and tested in individuals born in differing generations. Suppose one
generation grew up in an environment with limited exposure to environmental risk factors for myopia, whereas for a younger generation these risk factors had become pervasive, then major influences from gene-environment interactions could impair genetic prediction models\textsuperscript{28}.

**What about the environment?**

The extent to which genetic risk factors for myopia are deterministic, i.e. act independently of environmental exposures such as near work, education or time outdoors, is largely unknown. To date, just a handful of genetic variants have been reported to confer a greater risk in children who spend relatively more time than average engaged in near work or in full-time education (and of these, only one such gene-environment interaction has been replicated)\textsuperscript{11, 22, 29, 30}. No gene-environment (G x E) interactions involving time outdoors have been discovered. If G x E interactions turn out to be commonplace, then genetic predictions might benefit from taking lifestyle factors into account, e.g. given a particular child’s genetic profile, their risk may be different if they leave full-time education at age 18 compared to if they go on to University. This could improve the accuracy of genetic prediction above the likely $h^2_{SNP}$-based limit of approximately 35%, at the expense of a parents being given a more complicated message about their child’s risk of myopia.

Some of the effects of the environment and some G x E interactions affecting refractive development are likely to operate via epigenetics, i.e. modifications to the chemical nature of specific regions of the genome that alter gene expression patterns without changing the genome sequence. No such epigenetic changes with effects on myopia are known currently, and, at least according to theory, the utility of epigenetics for genetic prediction is likely to be low, since the changes in chemistry and gene expression may be restricted to particular ocular tissues, making them unavailable for analysis (in marked contrast to the readily-captured and decoded genome sequence).

The research literature currently suggests that for levels of time reading and time outdoors within the normal range, spending more time outdoors or less time on near work than average has limited impact on refractive error (e.g. time reading and time outdoors ascertained at age 8 years-old together explained <2% of the variation in refractive error at age 15 years-old\textsuperscript{24}). Therefore unless future research finds that spending much greater time outdoors than is commonly the case has a far larger benefit than this, then augmenting genetic predictions with pure (non-GxE) lifestyle-related risks will not be worthwhile\textsuperscript{15}. In contrast, there may be benefits in combining genetic risk scores with information about
parental myopia, under the assumption that myopic parents pass on not only their genes, but also their relatively myopiagenic environment to their children. As a precedent for this approach, a recent study\textsuperscript{31} suggested the genetic risk for heart disease was independent of self-reported family history, making a strong case for combining these sources of information.

**Questions for society**

Is it better to know or not to know one’s risk of Alzheimer’s disease, coronary artery disease, major depression, and myopia? Routine genetic testing/genome sequencing of every person in the population is now within reach for rich societies, potentially providing every individual with a personalized assessment of their risk of myopia and a multitude of other diseases and disorders. However, the American Society for Human Genetics’ 2015 position statement\textsuperscript{32} on “Genetic Testing in Children and Adolescents” argued against routine genetic sequencing of newborn babies and children (although the position statement did accept the legitimate case for testing specific candidate genes in children suffering from a likely genetic disease). The main obstacles arguing against routine genetic testing were deemed to be (i) the limited precision of current tests for the majority of diseases and disorders, (ii) a lack of evidence about how probabilistic information is best communicated to parents and children, and (iii) issues surrounding stigma and discrimination. Research aiming to address these issues is currently on-going, for instance, evaluating the pros and cons of using genome sequencing to diagnose the cause of intellectual disability, and how best to communicate this information to parents and those affected\textsuperscript{33}. As regards common eye disorders such as myopia, the American Academy of Ophthalmology’s 2014 “Recommendations for Genetic Testing of Inherited Eye Diseases” advised against routine testing until treatment strategies have been shown in published prospective clinical trials to benefit individuals with “specific disease-associated genotypes.”

Some individuals have taken steps to investigate their risk of genetic disorders themselves. For instance, the consumer genetics testing company 23andMe provided its one millionth customer with genetic predictions for a range of disorders, in 2015. (Most of these customers are from the United States, and paid around US$100 for the test)\textsuperscript{34}. Questions surrounding the use, and possible misuse, of genetic testing are important topics for debate that will become ever more important as prediction accuracy continues to improve.
**Unanswered Clinical Questions**

Although good evidence exists for the short term effectiveness of interventions to slow myopia progression, such approaches are not in routine use. Alongside developing predictive tests to help decide who might benefit from treatment, it is important to recognise that many treatment-related questions have yet to be addressed. These include the long term safety of pharmacological treatments, risks of keratitis in contact lens treatment modalities in young children, appropriate duration of treatment, decreased efficacy of treatments when used over several years, optimal age of treatment, and the potential for later rebound effects. Due to the time delay of many years between myopia onset and the manifestation of myopia-related ocular pathologies (cataract, glaucoma, retinal detachment and myopic maculopathy), the most challenging evidence to collect will be the impact of reduction in myopia progression on such ocular pathology in later life.

**Concluding remarks**

In the longer term, genetic prediction has the potential to enhance the detection of children who would benefit from treatment to slow, or ideally prevent, myopia progression. Based on current knowledge of the aetiology of refractive error, genetic prediction has the potential to perform better than existing methods (the best of which – using baseline refractive error and number of myopic parents as predictors – has a sensitivity and specificity of 62.5% and 81.9%). Compared to the use of baseline refractive error as a predictor, genetic prediction has the advantages that it can be employed at a very young age and does not rely on cooperation by the child being tested. Therefore, it is likely to be most effective when employed prior to the development of myopia – although, at present there is no research evidence supporting the prophylactic use of anti-myopia interventions other than time outdoors. Current genetic prediction models perform very poorly (explaining <5% of the variance in refractive error). To reach their full potential – explaining up to approximately 35% of the variance in refractive error – will require genetic prediction models to leverage information from GWAS involving 10-20 times larger sample sizes than those performed to date, carried out in participants with a range of ancestries/ethnicities. Incorporating the collection of information on genetic and other identifiable risk factors into the design of future myopia interventions trials will maximise the prospects for predicting myopia.
Glossary

**Additive effects** A polymorphism has an additive effect if the phenotype in heterozygotes is midway between that in the two homozygous genotype classes. This is not the case with dominant or recessive effects.

**Allele score** A simpler alternative to a polygenic risk score, in which equal weight is assigned to each of the predictors (polymorphisms) in the regression model. See polygenic risk score.

**Allele** The genome sequence of an individual chromosome at a particular polymorphic site. See also genotype.

**Genetic variant** A polymorphism.

**Genome-wide association study (GWAS)** A systematic search through the genome for polymorphisms associated with a phenotype.

**Genotype** The genome sequence of each of the chromosomes carried by an individual at a particular polymorphic site. Apart from polymorphisms on the sex chromosomes, genotypes comprise of 2 alleles in humans.

**Heritability** The genetic contribution to variation in a phenotype between individuals in a specific population.

**Negative predictive value** The probability that a patient with a negative screening result truly does not have the disorder.

**Phenotype** An observed trait or disease, e.g. height or refractive error.

**Polygenic** A phenotype caused by a large number of separate polymorphisms, each typically having only a subtle effect.

**Polygenic risk score** A metric used to predict an individual’s future phenotype. It is derived using a linear regression model that includes the average phenotype in the sample as the intercept, and sums the number of risk alleles carried by an individual multiplied by their corresponding phenotypic effect. The regression coefficients needed for the above linear regression model are found using a genome-wide association study in a reference sample.

**Polymorphism** A difference in the genome sequence between individuals at a specified position. The most abundant class comprise single nucleotide polymorphisms (SNPs).

**Positive predictive value** The probability that a patient with a positive screening result really does have the disorder.

**SNP heritability** The variation in a phenotype between individuals in a specific population that can be explained (i.e. accounted for in a regression model) by the genotype of commonly-occurring polymorphisms in the genome.
**Figure 1. Refractive error in parents and the diversity of refractive error in their children.** Computer simulations were carried out assuming a purely polygenic mode of inheritance for refractive error, at three different levels of heritability (25%, 50% and 75%). Results are shown for 8 separate families, each consisting of 2 parents and 3 children, with each family labelled using a different colour. The line of best fit is shown in blue. Note the wide range of refractive errors for the children in each family, even when the heritability is high.
Figure 2. Predicting refractive error in children using parental information or genetic prediction.

Computer simulations were carried out assuming a purely polygenic mode of inheritance for refractive error, at three different levels of heritability (25%, 50% and 75%). The refractive error of children is plotted as a function of either: (a) the number of myopic parents, (b) the average refractive error of the parents, or (c) the genetic risk of the child evaluated assuming perfect precision. For each graph the line of best fit is shown in blue and the r-squared value quantifies predictive efficacy (how much of the variation in children’s refractive error can be explained by the predictive factor).
Figure 3. Predicting refractive error in children using parental information or genetic prediction.
Computer simulations were carried out as in Figure 2 assuming a purely polygenic mode of inheritance for refractive error with heritability varying from 0–100%. The accuracy ($r^2$) of predicting refractive error in children is plotted as a function of either: (a) the number of myopic parents, (b) the average refractive error of the parents, or (c) the genetic risk of the child evaluated assuming perfect precision. The error bars show the 95% confidence intervals from 100 replicate simulations.
Figure 4. Accuracy ($r^2$) of genetic prediction for polygenic traits varies with SNP-heritability ($h^2_{SNP}$) and the sample size of the GWAS used to select predictor regression coefficients. The curves show fits to equation 1 of Wray et al. at different heritability levels.
References


