Machine Learning to Determine Risk Factors for Myopia Progression in Primary School Children: The Anyang Childhood Eye Study

Shi-Ming Li · Ming-Yang Ren · Jiahe Gan · San-Guo Zhang ·

Meng-Tian Kang · He Li · David A. Atchison · Jos Rozema ·

Andrzej Grzybowski · Ningli Wang [iD]

ABSTRACT

Introduction: To investigate the risk factors for myopia progression in primary school children and build prediction models by applying machine learning to longitudinal, cycloplegic autorefraction data.

Methods: A total of 2740 children from grade 1 to grade 6 were examined annually over a period of 5 years. Myopia progression was determined as change in cycloplegic autorefraction. Questionnaires were administered to gauge environmental factors. Each year, risk factors were evaluated and prediction models were built in a training group and then tested in an independent hold-out group using the random forest algorithm.

Results: Six variables appeared in prediction models on myopia progression for all 5 years, with combined weight of 77% and prediction accuracy over 80%. Uncorrected distance visual acuity (UDVA) had the greatest weight (mean 28%, range 22–39%), followed by spherical equivalent (20%, 7–28%), axial length (13%, 10–14%), flat

Shi-Ming Li, Ming-Yang Ren and Jiahe Gan contributed equally to this work.

S.-M. Li · J. Gan · M.-T. Kang · N. Wang (✉)
Beijing Tongren Eye Center, Beijing Tongren Hospital, Beijing Ophthalmology and Visual Science Key Lab, Beijing Institute of Ophthalmology, Capital Medical University, Beijing 100730, China
e-mail: wningli@vip.163.com

M.-Y. Ren · S.-G. Zhang (✉)
School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
e-mail: sgzhang@ucas.ac.cn

M.-Y. Ren · S.-G. Zhang
Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing, China

H. Li
Anyang Eye Hospital, Anyang, Henan Province, China

D. A. Atchison
Centre for Vision and Eye Research, Queensland University of Technology, Kelvin Grove, QLD, Australia

J. Rozema
Visual Optics Lab, Department of Ophthalmology, Antwerp University Hospital, Edegem, Belgium

J. Rozema
Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium

A. Grzybowski
Department of Ophthalmology, University of Warmia and Mazury, Olsztyn, Poland

A. Grzybowski
Institute for Research in Ophthalmology, Foundation for Ophthalmology Development, Poznan, Poland

keratometry reading (K1) (7%, 4–11%), gender (6%, 2–9%), and parental myopia (3%, 1–10%). UDVA and spherical equivalent had peak weight at the second and third study years, respectively. The weight of myopic parents decreased steadily over the 5 years (9.5%, 1.9%, 1.8%, 1%, and 1.3%). Weekly time spent reading, reading distance, reading in bed, and frequency of eating meat were included as variables in different study years.

Conclusions: Myopia progression in children was predicted well by machine learning models. UDVA and spherical equivalents were good predictive factors for myopia progression in children through primary school. Parental myopia was found to play a substantial role in the early stage of myopia progression but waned as children grew older.

Keywords: Risk factors; Myopia progression; Children; Machine learning

Key Summary Points

Why carry out this study?

Myopia prevalence is increasing worldwide, with half of the global population expected to have myopia by 2050.

Although the etiology of myopia remains unclear, it is important to control myopia early in children to avoid sight-threatening complications due to high myopia in the future.

The study asked: What are the main risk factors for myopia in children during primary school, and how can the change in these risk factors be predicted well?

What was learned from the study?

Myopia progression in primary school children could be predicted with good accuracy using machine learning models.

Ocular factors, such as spherical equivalent, had greater weight than environmental and genetic factors, and should be monitored annually to achieve early prediction and intervention in children with myopia.

INTRODUCTION

Myopia has become one of the most prevalent ocular disorders worldwide [1, 2] and has achieved an epidemic level of 90% prevalence in university students in China [3, 4]. Myopia-related complications and vision loss may become severe social concerns by 2050 [2]. The etiology of myopia remains unclear, although in recent years, some studies suggest that environmental factors may play as important a role as genetic factors [1]. Time outdoors is one promising environmental factor for myopia prevention in children, and can be increased by adding an outdoor physical class per day [5] or encouraging playing outside during recess [6]. In addition, education [7] and certain reading habits such as continuous reading, a close reading distance, and a low-light environment are reported to be associated with childhood myopia [8, 9].

In 2015, Zadnik et al. [10] found that cycloplegic spherical equivalent refraction was the single best predictive factor for myopia onset in school-age children and performed as well as all eight factors together, with the area under the curve (AUC) ranging from 0.87 to 0.93 depending on the factors included. Zhang et al. [11], using ocular biometry, height, weight, and presenting visual acuity in a model, reported AUCs of 0.82–0.97 in Chinese children. In 2010, Lim et al. [12] reported that higher intake of saturated fat and cholesterol were associated with longer axial length in schoolchildren, suggesting a possible relation between dietary factors and refractive errors. In 1996, Edwards et al. [13] reported that children who developed myopia had significantly lower intakes of fat, protein, vitamins B1, B2, and C, cholesterol, phosphorus and iron than children who did not become myopic. However, more studies are still needed to confirm the relation between dietary factors and myopia. Recently, Tideman et al. [14] reported that risk score combining environmental risk factors and ocular parameters can help to identify children at high risk of myopia. Nevertheless, environmental factors are various and changing every year for students during primary and middle school. Although

many researcher have built various myopia prediction models based on different datasets on risk factors for myopia [15], it remains unclear what the main risk factors for myopia are during primary school and how these risk factors change over time [16].

Machine learning is a method of data analysis that automates analytical model building and has been successfully used in imaging recognition and classification [17]. In the field of ophthalmology, machine learning has been used in diagnosis of diabetic retinopathy [18], predictions of myopia development [19, 20], orthokeratology lens prescription [21, 22], and visual acuity in patients treated for neovascular age-related macular degeneration [23]. There are many machine learning algorithms, each with its own strengths and weaknesses. Random forest is a supervised learning algorithm with the major advantage that it can be used for both classification and regression tasks, which form the majority of current machine learning systems. Although the algorithm is at risk of overfitting the data, this can be avoided through careful system design. In addition, it can handle missing values and can be modeled for categorical values.

In this study, we developed a system that assigns probabilities for myopia progression of children using machine learning with random forest. We applied this to investigate the risk factors for myopia progression using a large sample of Chinese children that were followed for a period of 5 years.

METHODS

Study Population

The Anyang Childhood Eye Study (ACES) was a school-based cohort study, which was approved by the Ethics Committee of Beijing Tongren Hospital, Capital Medical University, and adhered to the tenets of the Declaration of Helsinki. Informed written consent was obtained from at least one parent, while verbal assent was obtained from each child. Details of the methodology have been reported previously [24]. A total of 2740 grade 1 students aged

7.1 $\pm$ 0.4 years (range 6–9 years) were measured annually with ocular biometry and cycloplegic autorefraction [25, 26]. We examined the students during a fixed period of 2 months every year. We defined grade 1 as the first study year, grade 2 as the second year, and so on.

Procedures

All students had distance visual acuity measured with and without spectacles, if worn, using a logarithmic visual acuity chart (Precision Vision, La Salle, IL, USA) at a distance of 4 m [27]. A Lenstar LS900 (Haag-Streit, Koeniz, Switzerland) was used to measure axial length before cycloplegia [24]. Five repeated measurements were taken and averaged. The cornea powers were measured in the principal meridians to give the lesser (the flat) and the greater (the steep) corneal powers. Mean corneal power was calculated as the average of these powers [26]. Cycloplegic autorefraction was performed 30 min after one drop of topical anesthetic agent (Alcaine; Alcon, Fort Worth, TX, USA), two drops of 1% cyclopentolate (Alcon), and one drop of 0.5% tropicamide (Mydrin P; Santen, Osaka, Japan) at 5-min intervals. Three measurements were averaged (HRK-7000A, Huvitz, Gunpo, Korea), and the spherical equivalent was calculated (sphere power + cylinder power/2). Myopia progression was considered to be any increase in the myopic spherical equivalent in myopic children, while in the full cohort such a shift towards a more negative or less positive refractive error was termed a myopic shift. Information including children's near work load, time outdoors, living habits, reading habits, food habits, and parental myopia were gathered using questionnaires administered to the parents, with five choices from which to select offered for each question, as described in previous studies [9, 28].

Data Analysis

Data analysis was performed using the R programming language (http://www.r-project.org/) [29] on right eyes only. The children were

divided into two groups: (1) a randomly chosen subset of 10% of children as an independent dataset or a "hold-out" group to test the performance of the prediction model, and (2) the remaining children (90%) as the training group to identify risk factors and establish prediction models. A regression model was first used to screen the factors that will be included to model using random forest.

Univariate analysis was performed with the dependent variable being myopia progression in each study year, and independent variables being ocular axial length, near work, time outdoors, living habits, nutritional habits, reading habits, habits of wearing spectacles, and parental myopia (see Supplement A for the meaning of each variable). In multivariate regression analysis, categorical variables were changed into dummy variables, and the best subset of detected variables in univariate analysis was determined based on the Akaike information criterion. The relative weights of the predictive variables in the multivariate regression model were calculated.

The prediction models of each study year were determined by the random forest method (*randomForest* package for R, http://www.r-project.org/) [30], a mature ensemble learning method in machine learning that can be applied for classification and regression. Among the 90% of children used for training, five-fold cross-validation (80% subjects for training and 20% subjects for validation) was used to tune parameters and train an optimal random forest model. There were two parameters, mTry and nTree, which represent the number of randomly chosen features at each split of decision trees and the number of trees in the random forest, respectively. The rates of samples with absolute error between actual and predicted myopia progression of less than a certain error threshold, set to various values, and coefficients of determination between actual and predicted myopia progression, were used as prediction indexes. Finally, the model's performance was validated by applying it to the hold-out group.

RESULTS

At baseline, 2740 grade 1 students aged 7.1 ± 0.4 years were included. Boys accounted for 57.4%. From the first to the fifth study years, 2559 (93.4%), 2611 (95.3%), 2531 (92.4%), 2342 (85.5%), and 2199 (80.3%) children were re-examined, respectively (Table 1). There were no significant differences in baseline characteristic between the students included and those excluded due to incomplete follow-up, nor between the training group and hold-out group (Table 1). Table 2 shows the descriptive statistics of spherical equivalent, axial length, uncorrected distance visual acuity, and flat keratometry reading for children in each grade.

In this study, 68 variables were screened, including 23 continuous variables, 16 nominally categorical variables, and 29 orderly categorical variables (Supplement variables). From the first to the fifth study year, 19, 23, 26, 20, and 25 variables (Fig. 1 and Supplement B) were screened out, respectively.

Figure 1 shows the weights of variables for the first study year with ocular parameters in green, environmental factors in yellow, nutrition factors in red, and genetic factors and gender in gray. The weights of variables from the second to the fifth study years are shown in Supplement B. Table 3 shows the regression coefficients of predictive variables in each study year. Six variables were significant risk factors for myopia progression in all study years: more myopia ($P < 0.01$ to 0.045), poorer uncorrected distance visual acuity (UDVA, $P < 0.0001$), longer axial length ($P < 0.0001$), being female ($P < 0.0001$), higher flat keratometry reading (K1, $P < 0.0001$), and having two myopic parents ($P < 0.0001$–0.027).

Figure 2 shows the weights of six prominent variables and two additional variables. During the five study years, these combined variables had a mean weight of 76.7% (range 69.1–86.1%). UDVA had the greatest weight (28.3%, 21.6–38.9%), followed by spherical equivalent (20.4%, 7–28.1%), axial length (12.6%, 10.2–14.4%), the flat keratometry reading (K1) (6.7%, 3.7–10.8%), gender (5.7%, 1.9–8.5%), and myopic parents (3.1%, 1–9.5%).

Table 1 Baseline characteristics of children between training group and hold-out group in each study year (mean ± SD)

Variables	Baseline (grade 1)		First study year[a]		Second study year[a]		Third study year[a]		Fourth study year[a]		Fifth study year[a]	
	Training	Hold-out	Training	Hold-out	Training	Hold-out	Training	Hold-out	Training	Hold-out	Training	Hold-out
Participants (%)	2740 (100%)		2559 (93.4%)		2611 (95.3%)		2531 (92.4%)		2342 (85.5%)		2199 (80.3%)	
SE (diopters, D)	0.93 ± 1.03	1.02 ± 1.09	−0.33 ± 0.60	−0.28 ± 0.50	−0.45 ± 0.60	−0.41 ± 0.44	−0.63 ± 1.46	−0.53 ± 0.54	−0.52 ± 0.55	−0.50 ± 0.48	−0.57 ± 0.59	−0.55 ± 0.53
Axial length (mm)	22.72 ± 0.76	22.63 ± 0.76	0.38 ± 2.06	0.29 ± 0.19	0.24 ± 2.04	0.27 ± 0.25	0.31 ± 0.26	0.29 ± 0.25	0.29 ± 0.30	0.31 ± 0.41	0.32 ± 0.33	0.28 ± 0.35

Participants (%): number of children and response rate

SE spherical equivalent

[a]Except for the baseline (grade 1), the values for each study year were the changes from the previous year

Other variables were found to be significant in different study years (Figs. 1, 2 and Supplement figures). Wearing spectacles was significant at the fourth study year, with weight of 12.8%. Undergoing other myopia treatments (OMT) was significant at the second and third study years, with weights of 4% and 2.7%, respectively. Weekly time spent reading was significant in the first, third, fourth, and fifth study years (more reading, more myopia). Distance between the child's eye and book when reading was significant in the third, fourth, and fifth study years (farther distance, less myopia).

Figure 3 shows the curves of prediction accuracy with different absolute errors in each study year. When the absolute error between predicted and actual myopia progression was set at 0.50 D, the prediction accuracy was 80%. The accuracy increased to 90% for an absolute error at 0.75 D. The differences in mean myopia progression between predicted and actual values of each study year were less than 0.05 D.

DISCUSSION

Machine learning often uses many more variables in its prediction models because the emphasis is not on significance of individual variables, but rather the ability of the machine learning model to predict the independent variable from a combination of factors. In this study on risk factors for myopia progression in primary school children, 68 variables, not including variables of binocular vision and accommodation such as phoric state or accommodative lag, were first screened using multivariate regression analysis. Among these, six variables comprising uncorrected distance visual acuity (UDVA), spherical equivalent, axial length, the flat keratometry reading (K1), gender, and myopic parents were included in the models for all study years, with a mean combined weight of 76.7%. The prediction accuracy based on these variables was greater than 80%.

During the five study years, UDVA always had the greatest weight (28.3%) with a peak at the second study year, indicating that UDVA was the best predictor for myopia progression

Table 2 Distribution of SE, AL, AR, UDVA, and K1 of children in each grade (mean ± SD)

Variables	Baseline (grade 1)	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
SE (D)	+0.94 ± 0.84	+0.64 ± 1.01	+0.20 ± 1.27	−0.38 ± 1.58	−0.84 ± 1.77	−1.34 ± 1.97
AL (mm)	22.71 ± 0.73	23.01 ± 0.81	23.31 ± 0.86	23.60 ± 0.93	23.90 ± 1.00	24.19 ± 1.05
UDVA (logMAR)	0.08 ± 0.10	0.09 ± 0.13	0.14 ± 0.21	0.19 ± 0.28	0.27 ± 0.31	0.33 ± 0.30
K1 (D)	42.82 ± 1.35	42.83 ± 1.40	42.82 ± 1.35	42.78 ± 1.31	42.77 ± 1.33	42.74 ± 1.34

SE spherical equivalent; *AL* axial length; *UDVA* uncorrected distance visual acuity; *K1* flat keratometry reading
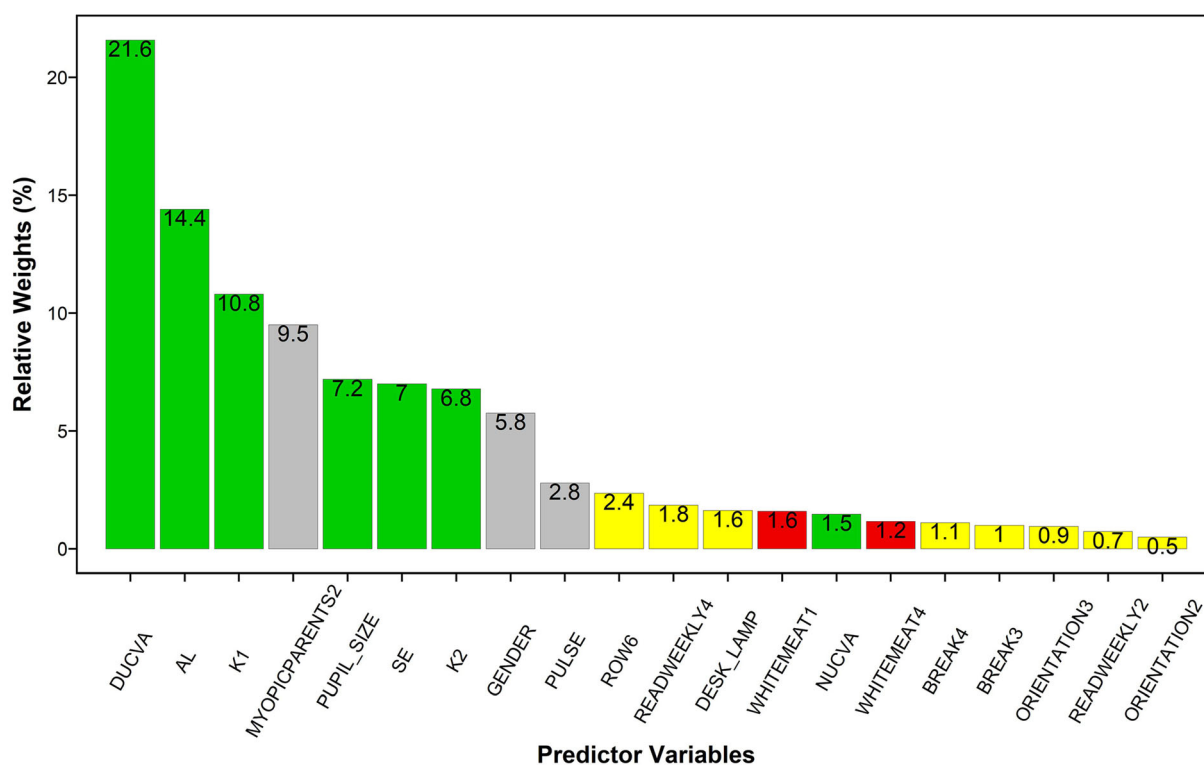
Fig. 1 Weights of predictor variables in the first study year using a random forest model. Ocular parameters are shown in green, environmental factors in yellow, nutrition factors in red, and genetic factors and gender in gray. *UDVA* uncorrected distance visual acuity; *AL* axial length; *K1* the flat keratometry reading; *MYOPICPARENTS2* two myopic parents; *PUPIL_SIZE* pupil diameter; *SE* spherical equivalent after cycloplegia; *K2* the steep keratometry reading; *GENDER* male or female; *PULSE* heart rate; *ROW* quantiles of rows children sit in the classroom from 1 least to 6 most; *READWEEKLY* quartiles of weekly reading from 1 (lowest) to 4 (highest); *DESK_LAMP* the type of lamp (bulb); *WHITEMEATS* quartiles of frequency of eating white meat, such as fish and chicken, in the last 4 weeks from 1 lower to 4 upper; *NUCVA* near uncorrected visual acuity; *BREAK* quartiles of time keeping reading or doing close work before a break from 1 least to 4 most; *ORIENTATION1,2,3* bedroom window orientated to south, west, and north, respectively (east as reference);

Table 3 Regression coefficients of predictor variables in the multivariate regression models of each study year

	First study year		Second study year		Third study year		Fourth study year		Fifth study year	
	β (95% CI)	P	β (95% CI)	P	β (95% CI)	P	β (95% CI)	P	β (95% CI)	P
Axial length	−0.19 (−0.24, −0.15)	< 0.0001	−0.14 (−0.18, −0.10)	< 0.0001	−0.10 (−0.14, −0.06)	< 0.0001	−0.12 (−0.17, −0.07)	< 0.0001	−0.14 (−0.19, −0.09)	< 0.0001
Spherical equivalent	−0.04 (−0.06, −0.01)	0.0087	0.04 (0.01, 0.07)	0.0024	0.05 (0.03, 0.08)	0.0001	0.04 (0.02, 0.07)	0.0003	0.03 (0.0005, 0.05)	0.0455
UDVA	−0.77 (−0.93, −0.62)	< 0.0001	−0.94 (−1.07, −0.81)	< 0.0001	−0.58 (−0.69, −0.47)	< 0.0001	−0.37 (−0.48, −0.26)	< 0.0001	−0.38 (−0.48, −0.27)	< 0.0001
K1	−0.14 (−0.18, −0.11)	< 0.0001	−0.09 (−0.12, −0.05)	< 0.0001	−0.04 (−0.06, −0.02)	< 0.0001	−0.07 (−0.10, −0.05)	< 0.0001	−0.08 (−0.10, −0.05)	< 0.0001
Gender	−0.09 (−0.12, −0.05)	< 0.0001	−0.07 (−0.10, −0.04)	< 0.0001	−0.12 (−0.16, −0.09)	< 0.0001	−0.13 (−0.16, −0.09)	< 0.0001	−0.10 (−0.13, −0.06)	< 0.0001
Two myopic parents	−0.17 (−0.23, −0.11)	< 0.0001	−0.10 (−0.16, −0.04)	0.0012	−0.09 (−0.16, −0.03)	0.0054	−0.07 (−0.15, 0.01)	0.079	−0.09 (−0.17, −0.01)	0.027

UDVA uncorrected distance visual acuity; *K1* flat keratometry reading
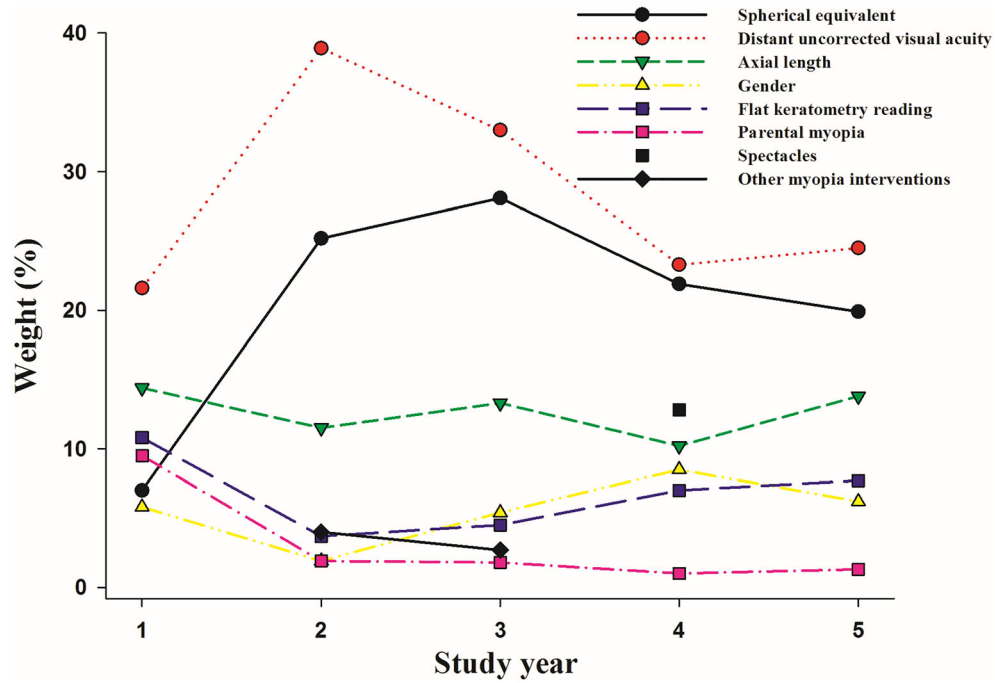
Fig. 2 Weights of important predictor variables in the random forest model during the five study years
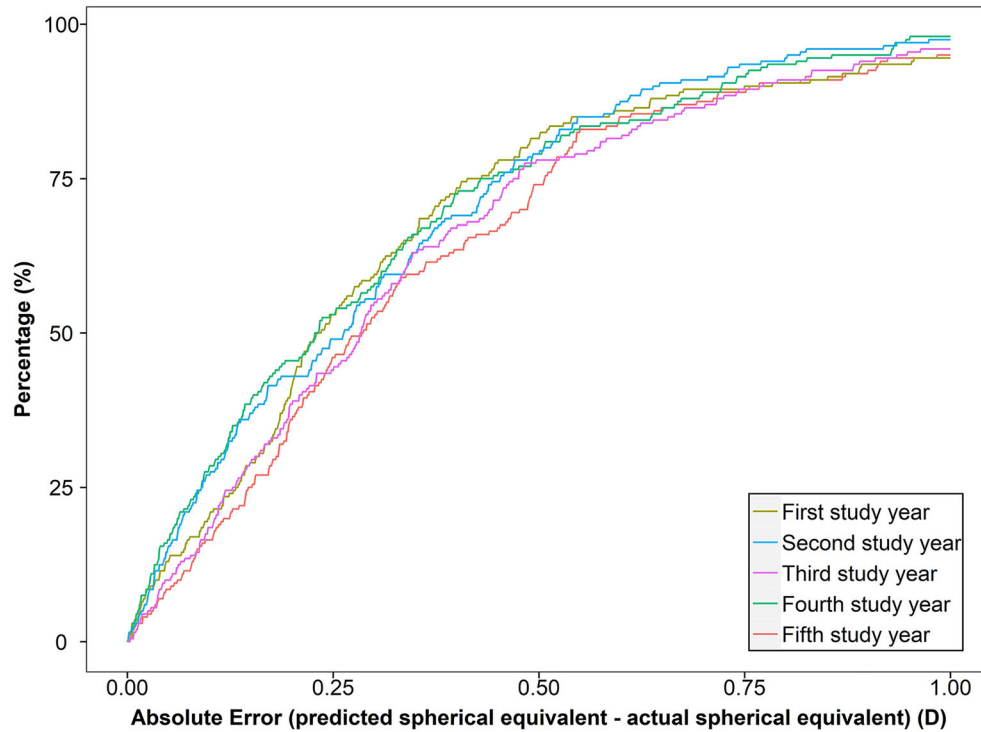

Fig. 3 Prediction accuracy curves of random forest models in five study years, that is, the accumulated percentage of samples as a function of the absolute difference between predicted and actual spherical equivalent refractions

and myopia screening [31]. This implies that UDVA of primary school children should be monitored frequently to identify children at risk of myopia. We also found that spherical equivalent was a significant predictive variable with a peak weight at the third study year. The successive weight peaks for UDVA and spherical equivalent might be explained by the peaks in myopia onset in grade 2 to grade 4 reported by previous studies [5, 32]. In other words, although more myopia was closely related to a lower UDVA, reduced UDVA occurred before myopia and thus acted as a more sensitive predictor of myopia progression.

Axial length was also found to be significant in our models, with moderate weight (12.6%), followed by the flat keratometry reading (K1) with lower weight (6.7%). Although the ratio of axial length to the corneal radius of curvature (AL/CR) is relatively good at classifying myopia grades [33], a recent study found that AL/CR was not useful in monitoring myopia progression in children due to a nonlinear relation between axial length and the corneal radius of curvature [34]. The successive lower weight of K1 and higher weight of axial length might reflect an active emmetropization process and final match between them [35].

A study on children in Singapore aged 6 months to 6 years reported that genetic factors (number of myopic parents) may play a more substantial role in early-onset myopia than environmental factors, of which neither near work nor outdoor activity was associated with early myopia [36]. Our findings confirmed this, as the 7-year-old children at grade 1 were affected by 7% by having myopic parents, but not by the total time spent in near work or outdoors. Furthermore, the influence of myopic parents decreased steadily over the five study years (9.5%, 1.9%, 1.8%, 1.0%, 1.3%, respectively), possibly reflecting a decreased importance of genetic factors on myopia with age. It should be noted that myopic parents not only constitute a genetic factor, but are likely associated with myopigenic environments, such as more time spent at near work and less time outdoors [37].

Interestingly, continuous reading (i.e., the break variable) was significant in all five study years during primary school, indicating its association with myopia, which was consistent with our previous reports in grade 7 children [9] and a study of Australian children [8]. In the Australian children study, myopia was also not associated with time spent doing near work [8]. The reasons for the lack of relation include inaccurate measurement of time on near work by questionnaire and dynamic changes in near work through different school years, as well as the effect of continuous reading (break) masked in the time on near work. In clinical trials, it has been demonstrated that making full use of recess time for outdoor activities can significantly control the development of myopia in children [6, 38]. Therefore, children's myopia may be more affected by the breaks between periods of continuous reading than total time spent on near work. In addition, use of smart phones or digital devices, which might be associated with myopia [39], were not measured alone in this study but included in the total time on near work.

In this study, time spent outdoors was not significant in the model, even though it is often regarded in the literature as the most promising environmental factor for controlling myopia onset. In our previous study, time spent outdoors was associated with a change in axial length but not with a change in spherical equivalent, possibly due to insufficient statistical power [28]. Moreover, the longitudinal results in the present study suggest that environmental factors are dynamic, making it more difficult to accurately evaluate their effects on myopia progression in children. This is difficult to prove, however, because time spent outdoors is usually determined through generally rather inaccurate questionnaires. The use of wearable devices to monitor children's studying and living environment, including time outdoors, might resolve the problem. This leaves the challenge of coping with the large amounts of data, which can be handled by machine learning.

K1 was probably a risk factor because it helps to form hyperopic defocus which has been shown to be the cause of prolonged axial myopia in animal studies. Although corneal power is basically stable after the age of 3 years, in some

cases, it might produce a certain compensatory reaction in the case of excessive near work, especially when the myopic refractive error is not corrected by glasses for a long time.

The strengths of this study include a large sample size and a high follow-up rate. This study also has some limitations. First, the environmental variables were determined using a questionnaire, which might lead to recall bias. Second, our study built the prediction model in the same cohort, although we divided the children into two groups; testing the model in other cohorts and populations is necessary to evaluate its generalizability. Third, the present work was designed to predict risk factors for myopia progression only 12 months ahead. It is very challenging to accurately forecast the occurrence of any medical condition 10 years in the future, since during that period there may be many important environmental or behavioral changes in the patient's life that such long-term predictions cannot take into account. Also, while it is possible to perform a retrospective analysis of the best predictors, it will be more difficult to prospectively validate these results in a 10-year follow-up study. For this reason. it is more realistic to use a series of short-term forecasts, allowing for treatment that is more relevant to the patient's current situation.

Random forest, a machine learning model widely used in many data analysis studies, performed well in predicting myopia progression. Methods for training the random forest model included bootstrap samples and random feature selection, thus reducing model variance, improving generalizability, and avoiding overfitting. This stabilizes the proposed random forest model and makes it more suitable for clinical practice. However, machine learning methods such as random forest act like a black box and are not easy to interpret. These techniques often use many variables for their prediction models because the emphasis is not on the significance of individual variables, but rather the ability of the machine learning-produced model to predict the independent variable from a combination of factors. This may occasionally lead to assigned values that seem irrelevant to a clinician.

Multivariate linear regression can compensate for some of the shortcomings of random forest, as they can be easily interpreted, reflecting the effect of predictive variables on response variables, and they can include important predictive variables and can analyze the relative weights of these variables. It is useful to combine classic regression models to explore risk factors with random forest to build highly accurate prediction models. Moreover, the random forest model can evaluate the importance of each predictive variable based on node purity from a decision tree, which might be different from linear models.

CONCLUSIONS

We have built machine learning-based prediction models for myopia progression in primary school children for a range of study years. The models demonstrated good accuracy for predicting myopia progression and showed the interaction among different factors. Ocular factors had greater weight than environmental and genetic factors. The environmental factors are modifiable to control myopia in children, which deserves further study to evaluate their interaction effect and feasibility in different populations and individuals.

ACKNOWLEDGEMENTS

REFERENCES

1. Morgan IG, Ohno-Matsui K, Saw SM. Myopia. Lancet. 2012;379(9827):1739–48.

2. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, Wong TY, Naduvilath TJ, Resnikoff S. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology. 2016;123(5):1036–42.

3. Sun YY, Wei SF, Li SM, Hu JP, Yang XH, Cao K, Lin CX, Du JL, Guo JY, Li H, Liu LR, Morgan IG, Wang NL. Cycloplegic refraction by 1% cyclopentolate in young adults: is it the gold standard? The Anyang University Students Eye Study (AUSES). Br J Ophthalmol. 2018. https://bjo.bmj.com/content/103/5/654.long.

4. Sun J, Zhou J, Zhao P, Lian J, Zhu H, Zhou Y, Sun Y, Wang Y, Zhao L, Wei Y, Wang L, Cun B, Ge S, Fan X. High prevalence of myopia and high myopia in 5060 Chinese university students in Shanghai. Invest Ophthalmol Vis Sci. 2012;53(12):7504–9.

5. He M, Xiang F, Zeng Y, Mai J, Chen Q, Zhang J, Smith W, Rose K, Morgan IG. Effect of time spent outdoors at school on the development of myopia among children in China: a randomized clinical trial. JAMA. 2015;314(11):1142–8.

6. Wu PC, Chen CT, Lin KK, Sun CC, Kuo CN, Huang HM, Poon YC, Yang ML, Chen CY, Huang JC, Wu PC, Yang IH, Yu HJ, Fang PC, Tsai CL, Chiou ST, Yang YH. Myopia prevention and outdoor light intensity in a school-based cluster randomized trial. Ophthalmology. 2018;125:1239–50.

7. Mountjoy E, Davies NM, Plotnikov D, Smith GD, Rodriguez S, Williams CE, Guggenheim JA, Atan D. Education and myopia: assessing the direction of causality by mendelian randomisation. BMJ. 2018;361:k2022.

8. Ip JM, Saw SM, Rose KA, Morgan IG, Kifley A, Wang JJ, Mitchell P. Role of near work in myopia: findings

in a sample of Australian school children. Invest Ophthalmol Vis Sci. 2008;49(7):2903–10.

9. Li SM, Li SY, Kang MT, Zhou Y, Liu LR, Li H, Wang YP, Zhan SY, Gopinath B, Mitchell P, Wang N, Anyang Childhood Eye Study G. Near work related parameters and myopia in Chinese children: the Anyang Childhood Eye Study. PLoS ONE. 2015;10(8): e0134514.

10. Zadnik K, Sinnott LT, Cotter SA, Jones-Jordan LA, Kleinstein RN, Manny RE, Twelker JD, Mutti DO. Prediction of juvenile-onset myopia. JAMA Ophthalmol. 2015;133(6):683–9.

11. Zhang M, Gazzard G, Fu Z, Li L, Chen B, Saw SM, Congdon N. Validating the accuracy of a model to predict the onset of myopia in children. Invest Ophthalmol Vis Sci. 2011;52(8):5836–41.

12. Lim LS, Gazzard G, Low YL, Choo R, Tan DT, Tong L, Yin Wong T, Saw SM. Dietary factors, myopia, and axial dimensions in children. Ophthalmology. 2010;117(5):993–7.

13. Edwards MH. Do variations in normal nutrition play a role in the development of myopia? Optom Vis Sci. 1996;73(10):638–43.

14. Tideman JWL, Polling JR, Jaddoe VWV, Vingerling JR, Klaver CCW. Environmental risk factors can reduce axial length elongation and myopia incidence in 6- to 9-year-old children. Ophthalmology. 2019;126(1):127–36.

15. Han X, Liu C, Chen Y, He M. Myopia prediction: a systematic review. Eye. 2021. https://doi.org/10.1038/s41433-021-01805-6.

16. Morgan IG, Wu PC, Ostrin LA, Tideman JWL, Yam JC, Lan W, Baraas RC, He X, Sankaridurg P, Saw SM, French AN, Rose KA, Guggenheim JA. IMI risk factors for myopia. Invest Ophthalmol Vis Sci. 2021;62(5):3.

17. Consejo A, Melcer T, Rozema JJ. Introduction to machine learning for ophthalmologists. Semin Ophthalmol. 2018. https://doi.org/10.1080/08820538.2018.1551496.

18. Grzybowski A, Brona P, Lim G, Ruamviboonsuk P, Tan GSW, Abramoff M, Ting DSW. Artificial intelligence for diabetic retinopathy screening: a review. Eye (Lond). 2020;34(3):451–60.

19. Lin H, Long E, Ding X, Diao H, Chen Z, Liu R, Huang J, Cai J, Xu S, Zhang X, Wang D, Chen K, Yu T, Wu D, Zhao X, Liu Z, Wu X, Jiang Y, Yang X, Cui D, Liu W, Zheng Y, Luo L, Wang H, Chan CC, Morgan IG, He M, Liu Y. Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical

records: a retrospective, multicentre machine learning study. PLoS Med. 2018;15(11): e1002674.

20. Tang T, Yu Z, Xu Q, Peng Z, Fan Y, Wang K, Ren Q, Qu J, Zhao M. A machine learning-based algorithm used to estimate the physiological elongation of ocular axial length in myopic children. Eye Vis (Lond). 2020;7:50.

21. Fan Y, Yu Z, Tang T, Liu X, Xu Q, Peng Z, Li Y, Wang K, Qu J, Zhao M. Machine learning algorithm improves accuracy of ortho-K lens fitting in vision shaping treatment. Cont Lens Anterior Eye. 2021. https://doi.org/10.1016/j.clae.2021.101474.

22. Fan Y, Yu Z, Peng Z, Xu Q, Tang T, Wang K, Ren Q, Zhao M, Qu J. Machine learning based strategy surpasses the traditional method for selecting the first trial Lens parameters for corneal refractive therapy in Chinese adolescents with myopia. Cont Lens Anterior Eye. 2021;44(3):101330.

23. Rohm M, Tresp V, Muller M, Kern C, Manakov I, Weiss M, Sim DA, Priglinger S, Keane PA, Kortuem K. Predicting visual acuity by using machine learning in patients treated for neovascular age-related macular degeneration. Ophthalmology. 2018;125(7):1028–36.

24. Li SM, Liu LR, Li SY, Ji YZ, Fu J, Wang Y, Li H, Zhu BD, Yang Z, Li L, Chen W, Kang MT, Zhang FJ, Zhan SY, Wang NL, Mitchell P, Anyang Childhood Eye Study G. Design, methodology and baseline data of a school-based cohort study in central china: the Anyang Childhood Eye Study. Ophthalmic Epidemiol. 2013;20(6):348–59.

25. Li S, Li SM, Wang XL, Kang MT, Liu LR, Li H, Wei SF, Ran AR, Zhan S, Thomas R, Wang N, Anyang Childhood Eye Study G. Distribution and associations of intraocular pressure in 7- and 12-year-old Chinese children: the Anyang Childhood Eye Study. PLoS ONE. 2017;12(8): e0181922.

26. Li SM, Li SY, Kang MT, Zhou YH, Li H, Liu LR, Yang XY, Wang YP, Yang Z, Zhan SY, Gopinath B, Mitchell P, Atchison DA, Wang N. Distribution of ocular biometry in 7- and 14-year-old Chinese children. Optom Vis Sci. 2015;92(5):566–72.

27. Li SM, Liu LR, Li SY, Ji YZ, Fu J, Wang Y, Li H, Zhu BD, Yang Z, Li L, Chen W, Kang MT, Zhang FJ, Zhan SY, Wang NL, Mitchell P, The Anyang Childhood Eye Study G. Design, methodology and baseline data of a school-based cohort study in central China: the Anyang Childhood Eye Study. Ophthalmic Epidemiol. 2013;20(6):348–59.

28. Li SM, Li H, Li SY, Liu LR, Kang MT, Wang YP, Zhang F, Zhan SY, Gopinath B, Mitchell P, Wang N, The Anyang Childhood Eye Study G. Time outdoors and myopia progression over 2 years in Chinese

children: the Anyang Childhood Eye Study. Invest Ophthalmol Vis Sci. 2015;56(8):4734–40.

29. Team RC. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. https://www.R-project.org/. Accessed Dec 2018.

30. Lantz B. Machine learning with R. Birmingham: Packt Publishing; 2013.

31. Lin S, Ma Y, He X, Zhu J, Zou H. Using decision curve analysis to evaluate common strategies for myopia screening in school-aged children. Ophthalmic Epidemiol. 2019;26(4):286–94.

32. Saw SM, Tong L, Chua WH, Chia KS, Koh D, Tan DT, Katz J. Incidence and progression of myopia in Singaporean school children. Invest Ophthalmol Vis Sci. 2005;46(1):51–7.

33. Scheiman M, Gwiazda J, Zhang Q, Deng L, Fern K, Manny RE, Weissberg E, Hyman L, Group C. Longitudinal changes in corneal curvature and its relationship to axial length in the Correction of Myopia Evaluation Trial (COMET) cohort. J Optom. 2016;9(1):13–21.

34. Jong M, Sankaridurg P, Naduvilath TJ, Li W, He M. The relationship between progression in axial length/corneal radius of curvature ratio and spherical equivalent refractive error in myopia. Optom Vis Sci. 2018;95(10):921–9.

35. Xiang F, He M, Morgan IG. Annual changes in refractive errors and ocular components before and after the onset of myopia in chinese children. Ophthalmology. 2012;119(7):1478–84.

36. Low W, Dirani M, Gazzard G, Chan YH, Zhou HJ, Selvaraj P, Au Eong KG, Young TL, Mitchell P, Wong TY, Saw SM. Family history, near work, outdoor activity, and myopia in Singapore Chinese preschool children. Br J Ophthalmol. 2010;94(8): 1012–6.

37. Xiang F, He M, Morgan IG. The impact of severity of parental myopia on myopia in chinese children. Optom Vis Sci. 2012;89(6):884–91.

38. Wu PC, Tsai CL, Wu HL, Yang YH, Kuo HK. Outdoor activity during class recess reduces myopia onset and progression in school children. Ophthalmology. 2013;120(5):1080–5.

39. Foreman J, Salim AT, Praveen A, Fonseka D, Ting DSW, Guang He M, Bourne RRA, Crowston J, Wong TY, Dirani M. Association between digital smart device use and myopia: a systematic review and meta-analysis. Lancet Digit Health. 2021;S2589–7500(2521):00135–7.