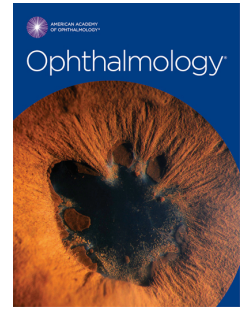


Journal Pre-proof

Distinguishing “Predictive Factors” from “Associated Factors” in the Establishment of Myopia Prediction Models

Shengsong Xu, Xiao Yang



PII: S0161-6420(26)00046-1

DOI: <https://doi.org/10.1016/j.ophtha.2026.01.016>

Reference: OPTHHA 13394

To appear in: *Ophthalmology*

Received Date: 14 January 2026

Accepted Date: 16 January 2026

Please cite this article as: Xu S, Yang X, Distinguishing “Predictive Factors” from “Associated Factors” in the Establishment of Myopia Prediction Models, *Ophthalmology* (2026), doi: <https://doi.org/10.1016/j.ophtha.2026.01.016>.

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier’s sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 Published by Elsevier Inc. on behalf of the American Academy of Ophthalmology

Distinguishing “Predictive Factors” from “Associated Factors” in the Establishment of Myopia Prediction Models

Shengsong Xu¹ and Xiao Yang²

Affiliations:

¹Shenzhen Eye Hospital, Shenzhen Eye Medical Center, Southern Medical University, Shenzhen, China

²State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, Guangzhou, China

Corresponding author

Xiao Yang, E-mail: yangx_zoc@163.com; Tel: +86-20-87330348

In the development of myopia prediction models, a wide range of variables have been reported to be statistically associated with myopia, including demographic characteristics (age, sex, etc.), ocular parameters (refractive error, axial length, lens parameters, etc.), environmental exposures (time outdoors, near-work activities, etc.), and genetic factors. These variables have been repeatedly demonstrated to be *associated factors* for myopia in epidemiologic studies.

Accordingly, we questioned the absence of certain well-established *associated factors*, such as age and corneal curvature, in our prediction model. This concern is highly representative. When reviewing the myopia prediction literature, we found that many researchers encounter similar confusion when interpreting prediction models. We therefore take this opportunity to clarify several key conceptual and methodological points:

1) First, by definition, *predictive factors* and *associated factors* represent distinct concepts. Not all associated factors necessarily qualify as *predictive factors*. This distinction is intuitive—as *predictive factors* must demonstrate not only statistical association but also plausible causal relevance to the outcome.

2) Second, from a statistical perspective, prediction models are inherently multivariable. Even if all *associated factors* are formally tested, deemed eligible *predictive factors*, and entered into the multivariable prediction model, multivariable model selection procedures—such as Akaike Information Criterion-based selection, penalization, or assessment of multicollinearity—may exclude variables that do not provide additional predictive value, even when they have demonstrated significance in

univariable analyses. In our study, age was carefully evaluated but did not significantly improve model performance after adjustment for other predictors and therefore was excluded from the final model. This phenomenon is not unique to our work (**Table**).¹⁻⁴ Notably, the seminal myopia prediction model developed by Zadnik and Mutti using the CLEERE cohort similarly excluded several well-established *associated factors* and ultimately retained only a single predictive factor, the baseline refractive error (**Table**).¹ Other well-established cohorts, such as the Anyang cohort² and the WEPrOM cohort,⁴ also did not include age in their final prediction models, suggesting that these cohorts may have identified similar patterns. Nevertheless, some studies did include age as a predictor; while, these analyses were partly based on cross-sectional data (**Table**).⁵

3) Third, from a mechanistic perspective, the inclusion of additional *associated factors* does not necessarily enhance predictive performance when existing predictors already encapsulate the relevant biological information. For example, corneal curvature, as mentioned by Ao, is indeed a determinant of refractive status. However, refractive error represents the integrated optical outcome of corneal curvature, lens power, and axial length. Consequently, refractive error already subsumes corneal information, rendering the additional inclusion of corneal curvature redundant. Similarly, adding absolute axial length does not meaningfully improve prediction. These observations have been consistently confirmed in prior studies.¹

With respect to the performance of the non-cycloplegic models questioned by Ao, we would like to clarify that these models were included primarily to illustrate the contrast between cycloplegic and non-cycloplegic conditions, and to highlight the necessity of cycloplegia for accurate prediction. The non-cycloplegic models are intended for use only in rare circumstances where cycloplegia is not feasible. On the other hand, regarding the concern that the cycloplegic model was not usable by Ao, we carefully re-examined the model inputs and outputs and confirmed that the model itself functions as intended. We suspect that the issue arose from incorrect data entry, particularly the omission or mis-specification of refractive error values in diopters (**Figure**). We received a small number of similar inquiries via the correspondence email and have been able to promptly resolve these issues by clarifying the correct input format (**Figure**). We encourage users encountering any technical difficulties to contact the corresponding author directly, as this typically allows for more efficient resolution.

Table. Predictive factors ultimately included in prediction models in published cohorts

Study (Author, Journal, Year)	Cohort	Final Predictors Included
Zadnik et al., <i>JAMA Ophthalmol</i> 2015 ¹	CLEERE	SE
Li et al., <i>Ophthalmol Ther</i> 2022 ²	ACES	UDVA; SE; AL; K1; sex; parental myopia
Tideman et al., <i>Ophthalmology</i> 2019 ³	Generation R Study	Parental myopia, near work, outdoor activity, ethnicity, AL/CR ratio
Guo et al., <i>American Journal of Ophthalmology</i> 2022 ⁴	WEPrOM	Sex; SE; AL; CR; PRA
Chen et al., <i>Asia Pac J Ophthalmol (Phila)</i> 2023 ⁵	GTES	Age; sex; SE; ocular biometry; UDVA; parental myopia

Abbr: CLEERE (Collaborative Longitudinal Evaluation of Ethnicity and Refractive Error Study); WEPrOM (Wenzhou Prediction of Myopia Study); ACES (Anyang Childhood Eye Study); GTES (Guangzhou Twin Eye Study); Spherical equivalent (SE); axial length (AL); corneal refractive power (CR); positive relative accommodation (PRA); Visual acuity (VA); Uncorrected distance visual acuity (UDVA); flat keratometry (K1)

School Myopia-Onset Risk Calculator

Choose Model Group:
Cyc Data Models

Gender:
Male

Baseline SER:
1

AL Change in Past 1 Year:
0.2

Number of Myopic Parents:
0

Predict Risk

School Myopia-Onset Risk Calculator

Choose Model Group:
Cyc Data Models

Gender:
Male

Baseline SER:
1


AL Change in Past 1 Year:
0.2

Number of Myopic Parents:
0

Predict Risk

Click →

Baseline SER:
100

 **Incorrect input format:**
entering myopia in degrees rather than in diopters

Prediction Results:

Cyc Data Model for 1 year
Myopia-onset risk probability: **0.010**
95% Confidence Interval: [0.003, 0.037]

Cyc Data Model for 2 years
Myopia-onset risk probability: **0.084**
95% Confidence Interval: [0.030, 0.215]

Cyc Data Model for 4 years
Myopia-onset risk probability: **0.319**
95% Confidence Interval: [0.146, 0.561]

Figure. Proper use of the online calculator (<https://zoc-qgzk.github.io/pre-myopia-calculator/>). Incorrect input formats, particularly entering refractive error in degrees rather than in diopters, will result in output values consistently equal to zero.

Reference

1. Zadnik K, Sinnott LT, Cotter SA, et al. Prediction of Juvenile-Onset Myopia. *JAMA Ophthalmology*. 2015;133(6):683-689. doi:10.1001/jamaophthalmol.2015.0471
2. Li SM, Ren MY, Gan J, et al. Machine Learning to Determine Risk Factors for Myopia Progression in Primary School Children: The Anyang Childhood Eye Study. *Ophthalmol Ther*. 2022;11(2):573-585. doi:10.1007/s40123-021-00450-2
3. Tideman JW, Polling JR, Jaddoe VV, Vingerling JR, Klaver CC. Environmental Risk Factors Can Reduce Axial Length Elongation and Myopia Incidence in 6- to 9-Year-Old Children. *Ophthalmology*. 2019;126(1):127-136. doi:10.1016/j.ophtha.2018.06.029
4. Guo C, Ye Y, Yuan Y, et al. Development and validation of a novel nomogram for predicting the occurrence of myopia in schoolchildren: A prospective cohort study. *American Journal of Ophthalmology*. 2022;242:96-106. doi:10.1016/j.ajo.2022.05.027
5. Chen Y, Tan C, Foo LL, et al. Development and Validation of a Model to Predict Who Will Develop Myopia in the Following Year as a Criterion to Define Premyopia. *Asia Pac J Ophthalmol (Phila)*. 2023;12(1):38-43. doi:10.1097/APO.0000000000000591