Real-world big data demonstrates prevalence trends and developmental patterns of myopia in China: a retrospective, multicenter study

Erping Long[1#], Xiaohang Wu[1#], Xiaohu Ding[1], Yahan Yang[1], Xun Wang[1], Chong Guo[1], Xiayin Zhang[1], Kexin Chen[3], Tongyong Yu[3], Dongxuan Wu[3], Xutu Zhao[3], Zhenzhen Liu[1], Yizhi Liu[1], Haotian Lin[1,2]

[1]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China; [2]Center for Precision Medicine, Sun Yat-sen University, Guangzhou, China; [3]Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China

Contributions: (I) Concept and design: E Long, X Wu, H Lin; (II) Administrative support: H Lin, Y Liu; (III) Provision of study materials or patients: H Lin; (IV) Collection and assembly of data: X Ding, Y Yang, X Wu, C Guo, X Zhang, K Chen, T Yu, D Wu, X Zhao, Z Liu; (V) Data analysis and interpretation: E Long, X Wu, C Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Haotian Lin; Xiaohang Wu; Zhongshan Ophthalmic Center, Sun Yat-sen University, 54# Xianlie S. Road, Guangzhou 510060, China. Email: linht5@mail.sysu.edu.cn; wxhang@mail2.sysu.edu.cn.

Background: Myopia is a complex disease caused by a combination of multiple pathogenic factors. Prevalence trends and developmental patterns of myopia exhibit substantial variability that cannot be clearly assessed using limited sample sizes. This study aims to determine the myopia prevalence over the past 60 years and trace the myopia development in a school-aged population using medical big data.

Methods: The refraction data from electronic medical records in eight hospitals in South China were collected from January 2005 to October 2018; including patients' year of birth, refraction status, and age at the exam. All optometry tests were performed in accordance with standard procedures by qualified senior optometrists. The cross-sectional datasets (individuals with a single examination) and longitudinal datasets (individuals with multiple examinations) were analyzed respectively. SAS statistical software was used to extract and statistically analyse all target data and to identify prevalence trends and developmental patterns related to myopia.

Results: In total, 1,112,054 cross-sectional individual refraction records and 774,645 longitudinal records of 273,006 individuals were collected. The myopia prevalence significantly increased among individuals who were born after the 1960s and showed a steep rise until reaching a peak of 80% at the 1980s. Regarding developmental patterns, the cross-sectional data demonstrated that the myopia prevalence increased dramatically from 23.13% to 82.83% aging from 5 to 11, and the prevalence stabilized at the age of 20. The longitudinal data confirmed the results that the age of myopic onset was 7.47±1.67 years, the age of myopia stabilized at 17.14±2.61 years, and the degree of myopia stabilized at −4.35±3.81 D.

Conclusions: The medical big data used in this study demonstrated prevalence trends of myopia over the past 60 years and revealed developmental patterns in the onset, progression and stability of myopia in China.

Keywords: Myopia; prevalence; real-world analysis; developmental pattern

Submitted Sep 28, 2020. Accepted for publication Jan 03, 2021.

doi: 10.21037/atm-20-6663

View this article at: http://dx.doi.org/10.21037/atm-20-6663

Introduction

Myopia, the leading cause of visual impairment worldwide, is induced by a combination of multiple pathogenic factors. Social-environmental factors are well known as the main risk factors of school-aged myopia (1,2). In the past sixty years, social-environmental factors, such as sunlight exposure, reading habits, visual stress, and outdoor activities, have triggered a marked change in school myopia prevalence in China (3). In large Chinese cities (such as Guangzhou), the school-aged myopia prevalence among primary school students has reached 50% (4), and the rate is as high as 80% among university students (5), affecting most teenagers' study performance and daily life. Therefore, the increasing myopia prevalence has become an international focus of public health. Moreover, the myopia development among school-aged children exhibited substantial heterogeneity because most of the social-environmental factors were highly variable from person to person (6-9). Our previous work developed an algorithm to predict the myopia development (10); however, the heterogenous predictive outcomes between groups suggest that the myopia developmental patterns are non-linear and intrinsically complex.

Large-scale medical records from optometric clinics provide a unique opportunity to reveal the developmental tendency of the disease because these routinely collected records can comprehensively reflect the clinical reality in an evidence-based manner (11). Previous studies have leveraged real-world data to transform the health-care situation, such as management of hypertension (12) and cancer (13). However, whether and to what extent the real-world big data can contribute to the myopia management remains unclear.

This study utilized millions of refraction records from eight ophthalmic medical institutions in China. The goal was to reveal prevalence trends and individual developmental patterns of myopia. The results will provide evidence useful for health policy decision-making regarding the prevention and control of myopia and determining the appropriate treatment timing.

We present the following study in accordance with the MDAR reporting checklist (available at http://dx.doi.org/10.21037/atm-20-6663).

Methods

Study design

This is a retrospective, multicenter study based on the refraction data derived from electronic health records. Eight ophthalmic centres were included in the study, including 4 in Guangzhou (Zhongshan Ophthalmic Center, Haizhu Optometry Department, Huangpu Optometry Department and Panyu Optometry Department) and four outside Guangzhou (Guangming Eye Hospital in Dongguan City, Optometry Centre in Huizhou City, Haikou Longhua Optometry Department and Xiuying Optometry Department in Haikou City). These eight ophthalmic centres are in South China and collectively provided a representative medical big data sample of Southeast Asia, an area with one of the highest myopia prevalence worldwide.

Data pre-processing

Refraction data from January 2005 to October 2018 at all 8 ophthalmic centers from the electronic medical record system were extracted, which contained individuals with a single examination and multiple examinations. Then the extracted data were separated into the cross-sectional dataset (a single examination) and longitudinal datasets (multiple examinations) based on unique identifiers for each individual. The records for eyes with iatrogenic or complicated conditions indicating a pathological refraction status or invalid, repetitive and aberrant records were excluded (e.g., diagnosed pathological myopia, previous intraocular surgery, refractive media opacities, glaucoma, notable retinopathy unrelated to myopia, or optic atrophy).

Standardized protocol for refraction evaluations

All examinations of the refractive status spherical equivalent (SE) were conducted using an autorefractor and were confirmed with objective retinoscopy (under cycloplegia for first-visit patients <18 years old). The cycloplegic routine varied depending on the centers. For autorefractor measurements, the right eye was tested first, followed by the left eye. Three refractive error measurements were obtained from each eye, and the mean result was recorded. Objective retinoscopy was conducted by experienced optometrists using standardized optometry processes, and the objective retinoscopy results were recorded in the final database.

Outcome definitions

The SE [algebraic sum in diopters [D], sphere + 1/2 cylinder] was used to categorize refractive error using conventional cut-offs (14); emmetropia (SE −0.49 to +0.49 D), low primary

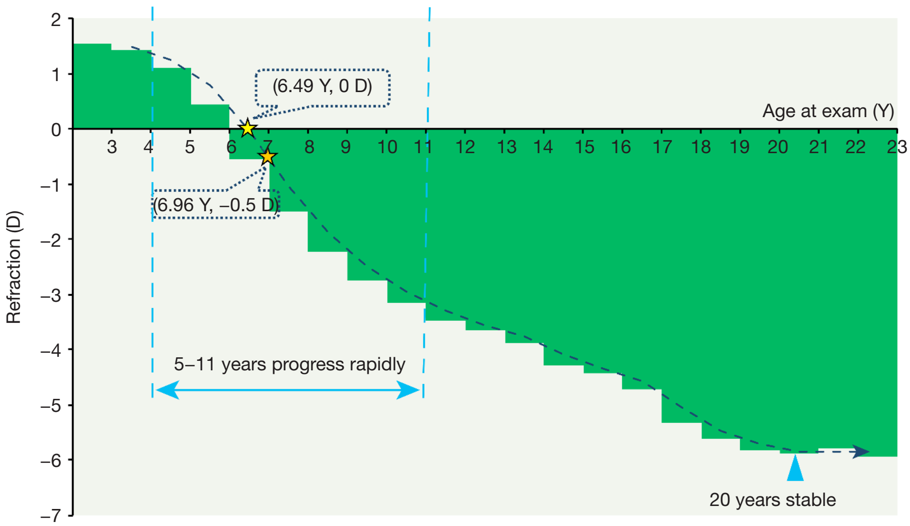

Figure 1 Distribution of mean diopters of refraction from ages 3 to 23 years in the cross-sectional data analysis. The cross-sectional data exhibited a significant downward trend (dark-blue dotted line) in the mean value of refraction from ages 3 (1.53 D) to 21 years (−5.88 D), especially between the ages of 5 and 11 years (blue interval). The rapid growth period is determined by the absolute value of the slope of the tendency line (dark blue dotted line). In particular, the refraction mean value passed 0 at age 6.49 years (yellow star) and passed −0.5 D at age 6.96 years (orange star), which was calculated with the linear regression equation: $y = -1.04x + 6.75$ ($6< x <7$). The stabilization age was approximately 20 years (blue triangle). D, diopters; Y, years.

myopia (SE −0.50 to −2.99 D), moderate primary myopia (SE −3.00 to −5.99 D), high primary myopia (SE −6.00 D or greater), and hyperopia (SE +0.50 D or greater).

Onset age was defined as the age at which the mean refractive status first reached −0.5 D (14) in the cross-sectional data. For the longitudinal cohort, the onset age was defined as the age at which the individual's refractive status was first measured from 0 to −1 D with a change of more than 0.5 D in the previous year (15). The additional criteria in the previous year aimed to detect the true onset timepoint.

Stabilization age (16) was defined as the age after which refractive status' change was no more than 0.5 D and persisted for more than two years in the longitudinal cohort. For cross-sectional data, the stabilization age was defined as the age at which the absolute value slope of the mean refractive status tendency (presented in *Figure 1*) line first dropped below 0.2.

Statistical analyses

The right eye was arbitrarily chosen to represent a specific individual, considering the refractive correlation between bilateral eyes (R^2=0.96, P=0.99, calculated by Pearson correlation). The individuals with missing records were excluded when calculating the onset and stabilization ages. Linear regression was used to calculate the absolute value slope and the age of the refraction mean value passing 0 D and −0.5 D in cross-sectional data. P value below 0.05 was considered statistically significant. The results were presented as mean ± standard deviation (SD). All analyses were performed using SAS statistical package version 9.2 (SAS Institute, Inc., Cary, NC, USA).

Ethical statement

The study adhered to the tenets of the Declaration of Helsinki (as revised in 2013), and approval for the study protocol was obtained from the Institutional Review Board/Ethics Committee of Sun Yat-Sen University, Guangzhou, China (No. 2010175). Consents were waived as all datasets used throughout the research were deidentified prior to transfer to the study investigators. The study was registered with ClinicalTrials.gov (identifier: NCT02667509).

Results

Overall participants demographics

For our cross-sectional dataset, 1,112,054 valid refraction

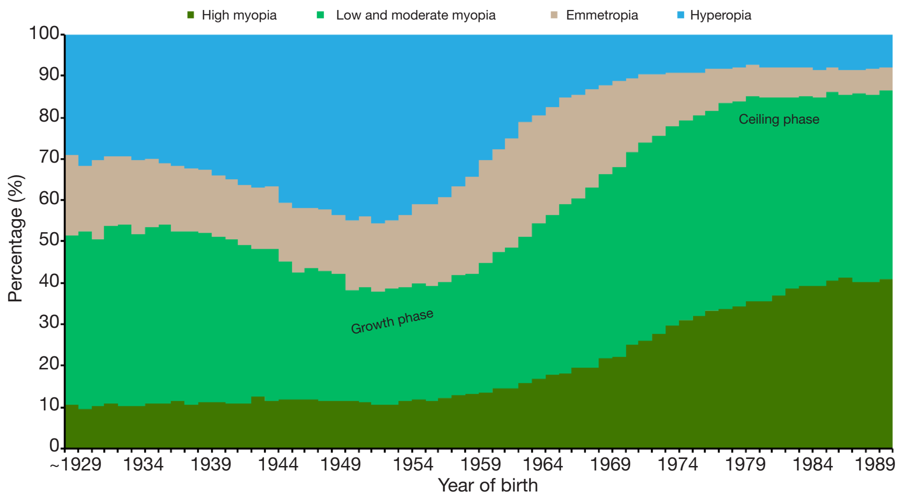

Figure 2 Dynamic changes in the refraction status of individuals born from 1929 to 1989 according to cross-sectional data. The cross-sectional data exhibited a downward trend in the percentage of myopia patients who were born in 1941 (50.63%) versus 1952 (37.98%). This low myopia proportion remained stable from the birth years of 1952 to 1959 (growth phase). From the birth years of 1960 to 1978, the cross-sectional data exhibited a significant upward trend in the percentage of myopia patients from 44.94% to 83.37%, for an average growth of 2.14% per year. This high proportion of myopia remained stable from the birth years of 1979 to 1990 (ceiling phase).

cases (male: 564,165; female: 547,889) were included; 891,421 (80.16%) participants were from Guangzhou, and 220,633 (19.84%) participants were outside Guangzhou. The longitudinal data were obtained from 273,006 participants (male: 139,176; female: 133,830; 774,645 re-examinations; follow-up length 0.04 to 11 years, mean 2.41±1.82 years); 228,567 (83.72%) participants were from Guangzhou, and 44,439 (16.28%) participants were from outside Guangzhou.

Population trend of myopia

The age range of the cross-sectional data is from 1 to 95 years old, which allows to describe the dynamic trend in the refraction status following the year of birth, including a “growth phase” and a “ceiling phase” (*Figure 2*). Specifically, the “growth phase” refers to the situation that the school-aged myopia prevalence significantly increased among individuals who were born after the 1960s. The “ceiling phase refers to the situation that myopia prevalence continued to increase until the 1980s, reaching a peak at approximately 80% of the current population.

Developmental pattern of myopia

As shown in *Figure 1*, in the cross-sectional dataset, the refractive tendency revealed that the onset age for school-aged myopia was approximately 7, and the stabilization age was approximately 20, with a stable refractive status at −5.81±3.74 D. There was a significant upward trend in myopia prevalence from ages of 5 (23.13%) to 11 (82.83%) years, with an average growth of 9.95% per year (*Figure 3*).

The longitudinal data indicated that the myopic onset age was 7.47±1.67 years and became stable at a refractive degree of −4.35±3.81 D at 17.14±2.61 years of age; the developmental pattern was similar, and statistically covered the cross-sectional values with the range between the first and third quartile (*Figure 4*). The onset age was calculated by 4,915 eligible individuals and the stabilization age was calculated by 3,072 eligible individuals (eligibility based on definition in Methods). In each cluster, the age of onset and stabilization was presented as mean ± SD.

Discussion

Risk factors, such as the social environment, are essential to the occurrence and development of many common diseases, of which myopia is one of the best representatives. In the past sixty years, the combination of dramatic changes has triggered a considerable increase in myopia (3,17,18), including education and reading habits, decreasing sunlight exposure and outdoor activities, and the increasing short-distance eye usage. However, previous epidemiological

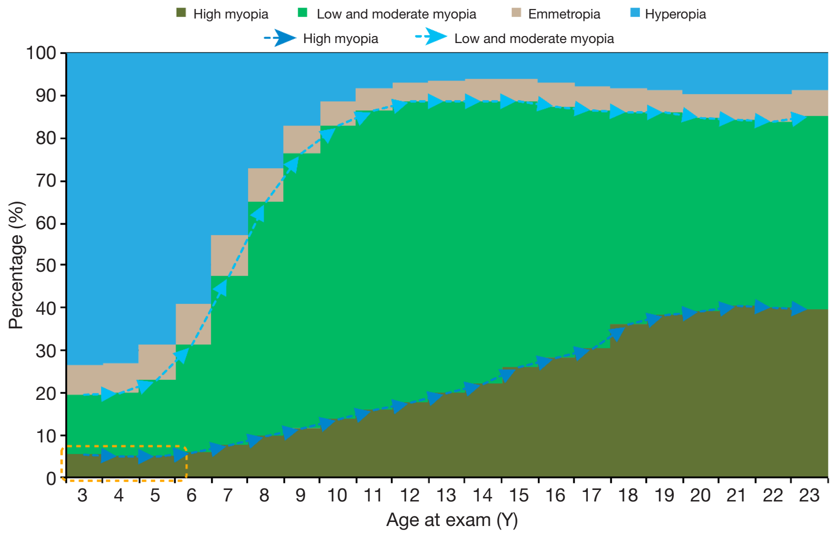

Figure 3 Dynamic changes in the myopia proportion from ages 3 to 23 years. Among all of the included patients, 5% of subjects developed high myopia by 5 years of age, but the increased rate of high myopia (dark blue arrow) was significantly more hysteretic than that of school myopia (light blue arrow). The cross-sectional data exhibited a significant upward trend in the percentage of myopia patients from ages 5 years (23.13%) to 11 years (82.83%), with an average growth of 9.95% per year. In particular, this high proportion of myopia remained stable until age 23. Y, years.

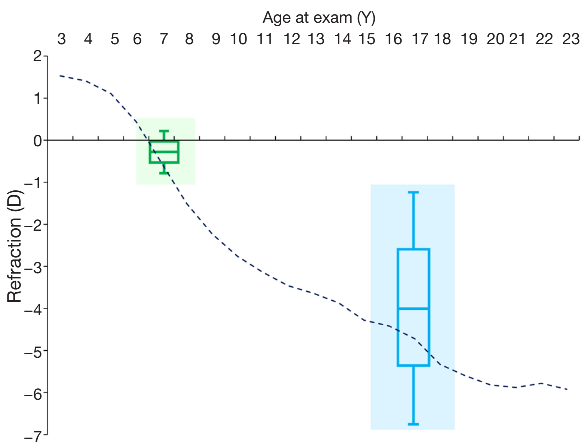

Figure 4 Trend of myopia development from ages 3 to 23 years as indicated by longitudinal data. The longitudinal data indicated that the age of onset of school myopia was 7.47±1.67 years (green box chart) and became stable at a refractive degree of –4.35±3.81 D at 17.14±2.61 years of age (blue box chart), which was consistent with the trend from the cross-sectional data (dark-blue dotted line generated from the same data shown in *Figure 1*). The onset age was calculated in clusters by valid records (4,915 patients) according to the definition. The stable age was calculated in clusters by valid records (3,072 patients) according to the definition. The box-whisker plot consists of the maximum, first quartile, median, third quartile, and minimum. D, diopters; Y, years.

studies have been based on short study durations and smaller samples that often lack long-term, panoramic, large data to represent the changes over the entire course of the disease (19-23). The present analysis consisting of millions of optometry records from eight Chinese ophthalmic institutes, is the first to reveal the pattern of myopia, regarding the onset, progression and stabilization, indicating the comprehensive and long-term impact of risk factor on the pathogenesis.

This study identified the critical period of time for the myopic onset and progression using this refractive big data and found that the rapid development of school-aged myopia occurred from 5 to 11 years of age, which may be the most effective period for the myopic prevention and control. A recent study showed that outdoor activities are effectual for myopia protection, but this protective effect is limited to emmetropic children (24). Based on these facts and the current findings, it is particularly essential to strengthen the prevention and control measures for juvenile myopia among children aged from 5 to 11 years to maximize their prophylactic effect in reducing the myopic onset.

The current study is a pioneering, big data-driven research project using large-scale electronic medical records and taking myopia, a prevalent disease in the Chinese population, as an example. Our data not only clearly demonstrated the developmental pattern of school-aged myopia but also recorded the prevalence trends during the past sixty years. The significant dynamic trend of the "growth phase" is consistent with the timing of the recovery of China's economy and educational system in the mid-1970s, and the "ceiling phase" may possibly be due to increased schooling intensity after social and economic developments regarding reforms and globalization of the country, as well as the implementation of a compulsory education policy in China. These policies have resulted in an explosive increase in myopia-related risk factors and an abrupt reduction in protective factors. However, the remaining 20% of the population remaining non-myopic may be due to certain gene(s) with protective effects, even in a highly myogenic environment (25).

It should be also noted that our study has several limitations. First, our records include missing information, as many of the individuals in our study did not attend multiple regular visits. It resulted in the limited sample size for calculating the onset and stabilization ages in the longitudinal analysis. Second, this study did not contain records on behaviors and environmental factors that may influence myopia (e.g., near-work habits). However, the large sample size and a multicenter design may minimize the potential bias caused by this latent structure. Third, the eight centers were chosen because their refraction data are available. It should be noted that there is no randomization and the main (80.16%) participants in this study were from Guangzhou, which will inevitably induce cluster effects. In future, it would be interesting to investigate the myopia pattern in diverse population by randomly choosing from a worldwide dataset to minimize these effects.

In summary, our study used large-scale data collected from the electronic health records of the eight ophthalmic institutions in China to demonstrate the contribution of big data to the better understanding of a disease. In the context of school-aged myopia, the most prevalent disease in the Chinese population, our results provided robust evidence on the long-term tendency of the proportion of myopia that may indicate the impact of social-environmental factors. Broadly, this work proposes a novel direction for the use of medical big data mining and its impact on transforming clinical practice and guiding health policy making.

Acknowledgments

We would like to thank TopEdit LLC for helping polish the language of our paper.

Funding: The study was supported by the National Key R&D Program of China (2018YFC0116500); Guangzhou Key Laboratory Project (202002010006); Guangdong Science and Technology Innovation Leading Talents (2017TX04R031); the National Natural Science Foundation of China-General Programs (81770967, 81873675); National Natural Science Fund for Distinguished Young Scholars (81822010); the National Natural Science Foundation-Young Scientists Fund (81800810) and the Science and Technology Planning Projects of Guangdong Province (2018B010109008).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-6663

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-20-6663

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-6663). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study adhered to the tenets of the Declaration of Helsinki (as revised in 2013), and approval for the study protocol was obtained from the Institutional Review Board/Ethics Committee of Sun Yat-Sen University, Guangzhou, China (No. 2010175). Consents were waived as all datasets used throughout the research were deidentified prior to transfer to the study investigators. The study was registered with ClinicalTrials.gov (identifier: NCT02667509).

References

1. Morgan IG, Ohno-Matsui K, Saw SM. Myopia. Lancet 2012;379:1739-48.
2. Goss DA. Nearwork and myopia. Lancet 2000;356:1456-7.
3. Dolgin E. The myopia boom. Nature 2015;519:276-8.
4. He M, Zeng J, Liu Y, et al. Refractive error and visual impairment in urban children in southern China. Invest Ophthalmol Vis Sci 2004;45:793-9.
5. Sun J, Zhou J, Zhao P, et al. High prevalence of myopia and high myopia in 5060 Chinese university students in Shanghai. Invest Ophthalmol Vis Sci 2012;53:7504-9.
6. Rusnak S, Salcman V, Hecova L, et al. Myopia Progression Risk: Seasonal and Lifestyle Variations in Axial Length Growth in Czech Children. J Ophthalmol 2018;2018:5076454.
7. Saw SM, Chua WH, Hong CY, et al. Nearwork in early-onset myopia. Invest Ophthalmol Vis Sci 2002;43:332-9.
8. Huang HM, Chang DS, Wu PC. The association between near work activities and myopia in children-a systematic review and meta-analysis. PLoS One 2015;10:e0140419.
9. Ip JM, Saw SM, Rose KA, et al. Role of near work in myopia: findings in a sample of Australian school children. Invest Ophthalmol Vis Sci 2008;49:2903-10.
10. Lin H, Long E, Ding X, et al. Prediction of myopia development among Chinese school-agedd children using refraction data from electronic medical records: A retrospective, multicentre machine learning study. PLoS Med 2018;15:e1002674.
11. Maissenhaelter BE, Woolmore AL, Schlag PM. Real-world evidence research based on big data: Motivation-challenges-success factors. Onkologe (Berl) 2018;24:91-8.
12. Okada M. Big data and real-world data-based medicine in the management of hypertension. Hypertens Res 2021;44:147-53.
13. Tsai CJ, Riaz N, Gomez SL. Big Data in Cancer Research: Real-World Resources for Precision Oncology to Improve Cancer Care Delivery. Semin Radiat Oncol 2019;29:306-10.
14. Flitcroft DI, He M, Jonas JB, et al. IMI - Defining and Classifying Myopia: A Proposed Set of Standards for Clinical and Epidemiologic Studies. Invest Ophthalmol Vis Sci 2019;60:M20-30.
15. Negrel AD, Maul E, Pokharel GP, et al. Refractive Error Study in Children: sampling and measurement methods for a multi-country survey. Am J Ophthalmol 2000;129:421-6.
16. COMET Group. Myopia stabilization and associated factors among participants in the Correction of Myopia Evaluation Trial (COMET). Invest Ophthalmol Vis Sci 2013;54:7871-84.
17. Sorsby A, Young FA. Transmission of refractive errors within Eskimo families. Am J Optom Arch Am Acad Optom 1970;47:244-9.
18. Rose KA, Morgan IG, Smith W, et al. High heritability of myopia does not preclude rapid changes in prevalence. Clin Exp Ophthalmol 2002;30:168-72.
19. Seet B, Wong TY, Tan DT, et al. Myopia in Singapore: taking a public health approach. Br J Ophthalmol 2001;85:521-6.
20. Lee KE, Klein BE, Klein R, et al. Changes in refraction over 10 years in an adult population: the Beaver Dam Eye study. Invest Ophthalmol Vis Sci 2002;43:2566-71.
21. Edwards MH, Lam CS. The epidemiology of myopia in Hong Kong. Ann Acad Med Singapore 2004;33:34-8.
22. Hrynchak PK, Mittelstaedt A, Machan CM, et al. Increase in myopia prevalence in clinic-based populations across a century. Optom Vis Sci 2013;90:1331-41.
23. Xiang F, He M, Zeng Y, et al. Increases in the prevalence of reduced visual acuity and myopia in Chinese children in Guangzhou over the past 20 years. Eye 2013;27:1353-8.
24. He M, Xiang F, Zeng Y, et al. Effect of time spent outdoors at school on the development of myopia among children in China: a randomized clinical trial. JAMA 2015;314:1142-8.
25. Wojciechowski R. Nature and nurture: the complex genetics of myopia and refractive error. Clin Genet 2011;79:301-20.

Cite this article as: Long E, Wu X, Ding X, Yang Y, Wang X, Guo C, Zhang X, Chen K, Yu T, Wu D, Zhao X, Liu Z, Liu Y, Lin H. Real-world big data demonstrates prevalence trends and developmental patterns of myopia in China: a retrospective, multicenter study. Ann Transl Med 2021;9(7):554. doi: 10.21037/atm-20-6663

