tvst

Letter to the Editor

Predicting the Real-World Future of Glaucoma Patients? Cautions Are Required for Machine Learning

Erping Long, Peixing Wan, and Yehong Zhuo

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

Correspondence: Peixing Wan, State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, Xian Lie South Road 54#, Guangzhou, 510060, China. e-mail: peixing_wan@163.com

We read with great interest the recent article by Yuki et al.[1] It is very thought-provoking that the authors predicted the likelihood of a future motor vehicle collision (MVC) among patients with primary open-angle glaucoma (POAG) based on multiple attributes using the penalized support vector machine (pSVM). However, several points remain to be discussed in terms of model design, data interpretation, as well as clinical application.

On one hand, the design of the model is somewhat controversial (recapitulation in Fig. 1A). The high-dimensional dataset (62 variables in the predpenSVM_basic model and 84 variables in the predSVM and predpenSVM_all models) is inclined to overfit when the sample size is small, because the cover range is not sufficient for pattern recognition on each attribute.[2] Elaboration of attributes' contribution is a promising alternative to address this issue. Model simplification (excluding the less contributed attributes) will broaden dramatically the application merit and be more cost-effective. Moreover, the progress of POAG and corresponding controlling efficacy have been reported to exhibit great heterogeneity.[3] Thus, potential indictors regarding intraocular pressure, visual field, and drug intake should be included. Given that best-corrected visual acuity has a significant effect on the likelihood of MVC, other variables of visual function, including refractive conditions, stereoscopic vision, contrast sensitivity, and color vision, can improve the model performance as well.[4,5]

On the other hand, the model validation and its clinical merit should be interpreted cautiously (recapitulation in Fig. 1B). First concern lies in the consistency of predicted results among patients at the three involved centers. Differences should be evaluated carefully to investigate whether this predictive model could be extended to patients at other
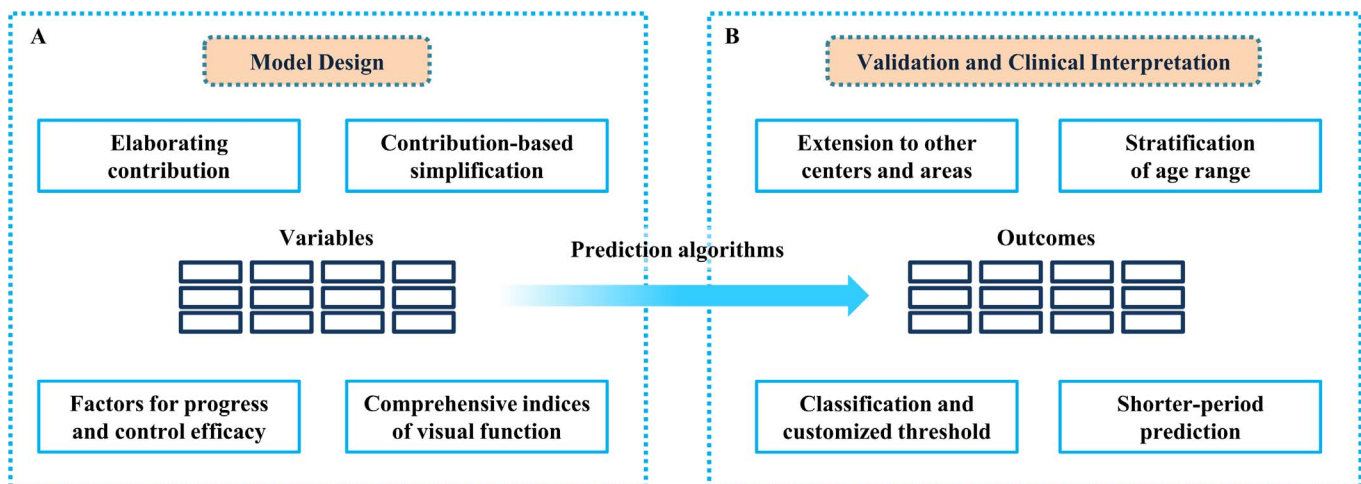

Figure 1. Overall suggestions on the design, verification, and interpretation of study. (A) For the model design, elaboration of the variables' contribution and model simplification will broaden the application merit dramatically. Moreover, potential indictors regarding the progress of POAG, controlling efficacy, and comprehensive indices of visual function can improve the model performance further. (B) For the model validation and clinical merit, the variety among areas and generations should be investigated and stratified. Besides, customized classifiers and a shorter-period prediction could provide direct and practical guideline for real-world decision making.

translational vision science & technology

centers or areas. Second, the age range (with a span of above 40 years) should be stratified before modeling due to the distinct behavioral pattern and life-styles among different generations.[6] Third, the predicted outcome is equivocal; it is unclear whether intervention should be conducted based on a probability. Alternatively, the classification outcome could provide direct and reliable guidelines.[7] Classifiers also could be customized to fit in various real-world situations: loose thresholds for better sensitivity in wide-range primary screening and more rigorous thresholds when making vital decision.[8] Fourth, the predictive power varies within different time frames: it is easier to achieve strong predictive performance within the first year than within 3 years. It also is a dilemma for decision-making when facing a likely incident within a 3-year period: when will the incident happen within this span of 3 years? Should all these high-risk patients be disqualified? Therefore, a shorter-period prediction will be more practical and cost-efficient.

Machine lea rning for prediction promises to transform health care.[9] However, in the "hype cycle" of emerging technologies, machine learning currently rides atop the "peak of inflated expectations"[10] Therefore, the capabilities and limitations of this technology should be validated and acknowledged more cautiously. Although we are conservative about the effect of the model, we appreciate the remarkable contribution made by Yuki and hope our suggestions on the design, verification and interpretation will help this work move toward real-world application.

Acknowledgments

References

1. Yuki K, Asaoka R, Awano-Tanabe S, et al. Predicting future self-reported motor vehicle collisions in subjects with primary open-angle glaucoma using the penalized support vector machine method. *Trans Vis Sci Tech.* 2017;6:1–10.
2. Hawkins DM. The Problem of Overfitting. *J Chem Info Computer Sci.* 2011;44:1.
3. Kymes SM, Kass MA, Anderson DR, Miller JP, Gordon MO. Management of ocular hypertension: a cost-effectiveness approach from the Ocular Hypertension Treatment Study. *Am J Ophthalmol.* 2006;141:997–1008.
4. Shen L, Melles RB, Metlapally R, et al. The association of refractive error with glaucoma in a multiethnic population. *Ophthalmology.* 2016;123:92–101.
5. Ugurlu SK, Altundal AEK, Ekin MA. Comparison of vision-related quality of life in primary open-angle glaucoma and dry-type age-related macular degeneration. *Eye,* 2017;31:395–405.
6. Cedrone C, Mancino R, Cerulli A, et al. Epidemiology of primary glaucoma: prevalence, incidence, and blinding effects. *Prog Brain Res.* 2008;173:3–14.
7. Guo G, Fu Y, Dyer CR, et al. Head pose estimation: classification or regression? *Int Conf Pattern Recog IEEE.* 2012:1–4.
8. Hong CS. Optimal Threshold from ROC and CAP Curves. *Comm Stat Simul Comput.* 2009;38:2060–2072.
9. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med.* 2016;375:1216–1219.
10. Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. *N Engl J Med.* 2017;376:2507–2509.