Inheritance Pattern Prediction of Retinal Dystrophies: A Machine-Learning Model

Dana Schlegel, Edmond Cunningham, Xinghai Zhang, Yaman Abdulhak, Andrew DeOrio, and Thiran Jayasundera

Purpose: Knowing the pattern of inheritance of a retinal dystrophy can suggest causal mutated genes, inform appropriate genetic testing, and determine likely diagnoses. At this time, there is no known algorithm to predict the pattern of inheritance for a patient, and not all patients have access to genetic counseling resources. Therefore, using a retrospective chart review to collect data on patients with genetically-proven retinal dystrophies, a machine-learning algorithm has been created whose input is family history phenotype and whose output is the most likely pattern of inheritance.

Methods: The pedigrees of patients from the Kellogg Eye Center with genetically-proven retinal dystrophies were analyzed with 3 methods: 1) prediction of pattern of inheritance by students who were trained by genetic counselors (277 patients), 2) clinician-reported answers to 12 questions about family history (100 patients), and 3) computer- calculated answers to the same questions using feature extraction, with tolerance for user input errors, on a digitalized patient pedigree (90 patients). The 3 cohorts have an overlap of 70 patients. Questions were based on pedigree structure, severity and age of onset of disease in males vs females, symmetric vs asymmetric disease in females, and variability of retinal diagnoses in the family. A machine learning model using a Gradient Boosted Tree (80/20 training/testing split) was used on the second two methods to predict the pattern of inheritance. Results were compared with the pattern of inheritance associated with the causal mutated gene for that patient.

Results: Student-predicted pattern of inheritance had 84% accuracy, machine learning on clinician-reported answers to questions had 77% accuracy (SD 2.6%), and machine learning on computer-calculated answers to the questions had 70% accuracy (SD 3.1%). Our machine-learning algorithms provided results with similar accuracy to that of humans, and this algorithm will likely improve over time and with the addition of more data.

Conclusions: Machine learning methodology is becoming more common in medicine, and it is a useful tool for examining the pattern of inheritance for patients with retinal dystrophies. This may be helpful for predictive purposes, counseling about risk calculation for relatives, determining appropriate genetic testing to order, and refining the differential genetic and clinical diagnoses for a given patient.