AMERICAN ACADEMY
OF OPHTHALMOLOGY®

Clinical Prediction Performance of Glaucoma Progression Using a 2-Dimensional Continuous-Time Hidden Markov Model with Structural and Functional Measurements

Youngseok Song, MD, PhD,[1,] Hiroshi Ishikawa, MD,[1,*] Mengfei Wu, MS,[1,2] Yu-Ying Liu, PhD,[3] Katie A. Lucy, BS,[1] Fabio Lavinsky, MD, MBA,[1] Mengling Liu, PhD,[2] Gadi Wollstein, MD,[1] Joel S. Schuman, MD[1]*

Purpose: Previously, we introduced a state-based 2-dimensional continuous-time hidden Markov model (2D CT HMM) to model the pattern of detected glaucoma changes using structural and functional information simultaneously. The purpose of this study was to evaluate the detected glaucoma change prediction performance of the model in a real clinical setting using a retrospective longitudinal dataset.

Design: Longitudinal, retrospective study.

Participants: One hundred thirty-four eyes from 134 participants diagnosed with glaucoma or as glaucoma suspects (average follow-up, 4.4±1.2 years; average number of visits, 7.1±1.8).

Methods: A 2D CT HMM model was trained using OCT (Cirrus HD-OCT; Zeiss, Dublin, CA) average circumpapillary retinal nerve fiber layer (cRNFL) thickness and visual field index (VFI) or mean deviation (MD; Humphrey Field Analyzer; Zeiss). The model was trained using a subset of the data (107 of 134 eyes [80%]) including all visits except for the last visit, which was used to test the prediction performance (training set). Additionally, the remaining 27 eyes were used for secondary performance testing as an independent group (validation set). The 2D CT HMM predicts 1 of 4 possible detected state changes based on 1 input state.

Main Outcome Measures: Prediction accuracy was assessed as the percentage of correct prediction against the patient's actual recorded state. In addition, deviations of the predicted long-term detected change paths from the actual detected change paths were measured.

Results: Baseline mean ± standard deviation age was 61.9±11.4 years, VFI was 90.7±17.4, MD was −3.50±6.04 dB, and cRNFL thickness was 74.9±12.2 μm. The accuracy of detected glaucoma change prediction using the training set was comparable with the validation set (57.0% and 68.0%, respectively). Prediction deviation from the actual detected change path showed stability throughout patient follow-up.

Conclusions: The 2D CT HMM demonstrated promising prediction performance in detecting glaucoma change performance in a simulated clinical setting using an independent cohort. The 2D CT HMM allows information from just 1 visit to predict at least 5 subsequent visits with similar performance. *Ophthalmology 2018;■:1–8 © 2018 by the American Academy of Ophthalmology*

For clinical management of a slowly progressing, irreversible disease like glaucoma, longitudinal progression assessment is essential to early detection and prediction of the disease.[1–3] However, because of the test-to-test measurement variability of established glaucoma biomarkers such as the retinal nerve fiber layer (RNFL) thickness and visual field (VF) assessments, it takes at least a few years to assess longitudinal changes reliably.[2,4] Some longitudinal studies of glaucoma progression reported heterogeneous observations of the relationships between structural and functional measurements.[2,5–8] Although most cases showed that structural progression preceded functional progression, a sizable number of cases from various reports showed that functional progression preceded structural progression. Attempts have been made to combine structural and functional measurements in the same statistical model to investigate heterogeneous observations further.[9–12] Although these combination approaches marginally improved glaucoma progression prediction performance, there is still no widely accepted model to predict glaucoma progression.

Previously, we reported a statistical approach called the 2-dimensional continuous-time hidden Markov model (2D CT HMM), which can predict glaucoma change by taking both structural and functional measurements into account.[13,14] The mean absolute error (MAE) of the 2D CT HMM outperformed conventional ordinary least squares regression (OLSR) and Bayesian joint linear regression (BJLR) models, which are the basis of conventional glaucoma progression analysis.[13,14] However, previously we used time-domain OCT to measure RNFL thickness, which has become practically obsolete and replaced by spectral-domain OCT,

https://doi.org/10.1016/j.ophtha.2018.02.010
ISSN 0161-6420/18

which has better resolution and measurement reproducibility.[15] In this study, we revisited and assessed the potential of the 2D CT HMM by using an updated imaging device and an individual cohort to gauge the practical clinical performance of the 2D CT HMM. The purpose of this study was to assess the prediction performance in detecting glaucoma change of the 2D CT HMM using a retrospective, longitudinal clinical dataset.

Methods

Patient Population

This was an observational study that was conducted in accordance with the tenets of the Declaration of Helsinki and the Health Insurance Portability and Accountability Act. The institutional review boards of New York University and the University of Pittsburgh approved the study, and all participants gave written consent before enrollment.

Participants diagnosed with glaucoma or as glaucoma suspects were enrolled to the study. All participants underwent a comprehensive ophthalmic examination, including best-corrected visual acuity measurement, slit-lamp biomicroscopy, intraocular pressure measurement, gonioscopy, dilated funduscopic examination, stereoscopic optic disc photography, and VF testing. The participants were included in the study according to the following inclusion criteria: best-corrected visual acuity of 20/60 or better, refractive error between −6.00 and 3.00 diopters, a minimum of 6 reliable VF tests, and good-quality OCT scans (see the testing protocols and criteria below). Participants were excluded based on the following exclusion criteria: history of intraocular surgery without cataract, glaucoma, laser surgery (participants were allowed to have these surgeries even during the follow-up period), or a combination thereof; diagnosis of diabetes; or diagnosis of a posterior pole pathologic condition other than glaucoma.

Clinical Diagnosis

Our study included participants diagnosed with glaucoma or as glaucoma suspects at their baseline visit. Glaucoma suspect eyes were defined as eyes with normal VF testing results with any of the following criteria: intraocular pressure of 22 to 30 mmHg, asymmetric optic nerve head cupping, or both; abnormal optic nerve head appearance; or an eye that was the contralateral eye of unilateral glaucoma.[7] Glaucomatous eyes were those with glaucomatous VF defects (at least 2 consecutive abnormal VF test results).

Visual Field Testing

All patients underwent Swedish interactive thresholding algorithm standard 24-2 perimetry (Humphrey Field Analyzer; Zeiss, Dublin, CA). A reliable VF test was defined as one with less than 30% fixation losses, false-positive responses, or false-negative responses. Abnormal VF tests were defined as tests featuring a cluster of 3 or more adjacent points in the pattern deviation plot depressed more than 5 dB, or 2 adjacent points depressed more than 10 dB and pattern standard deviation or glaucoma hemifield test results outside normal limits. The VFI and mean deviation (MD) from 2 consecutive visits were used as functional measurements.

OCT

Eyes were imaged using a commercially available spectral-domain OCT device (Cirrus OCT; Zeiss), and the average thickness of the circumpapillary RNFL (cRNFL) was acquired. Only good-quality scans that were well centered, had no detectable eye motion within the scan area, and had a signal strength of 7 or more were included.

2-Dimensional Continuous-Time Hidden Markov Model

Details of the 2D CT HMM have been described in previous publications.[13,14] Briefly, the 2D CT HMM is a state-based model where detected disease changes within the defined ranges[13] can be predicted based on 2 different variables such as structural and functional measurements, depending on each state.[13,14] In glaucoma assessment, this model predicts the next transition, including no state change, solely by taking into account the current visit information, regardless of the paths by which a given participant reached a specific state. Additionally, the model can be used to generate an intuitive graphical representation of detected glaucoma change trajectories that could benefit clinical glaucoma management by providing visually intelligible cues of the detected change direction for structure, function, or both.[13,14] We used 2 hidden state dimensions, including OCT cRNFL thickness as the structural dimension and VFI as the functional dimension, to model the detected disease status changes for glaucoma (VFI-trained model). Additionally, another model was trained with cRNFL thickness and MD instead of VFI (MD-trained model).

Study Protocol

One hundred four eyes with glaucoma and 30 glaucoma suspect eyes were enrolled (total of 134 eyes from 134 participants). One hundred seven eyes of 134 participants (80%) were used to train the model (training set). All data except for the last visit were used for training, and only the last visit was used for prediction performance testing, so that we avoided overfitting by using overlapped data that was used for training the model. The prediction accuracy of the model was evaluated using the following method. The information of the penultimate visit was used as the only input to predict the last visit state, which was compared with the actual state at the last visit. For example, if there were 6 visits in total, the first 5 visits were used for training the model (4 transitions), then the fifth visit was used as an input and the most likely state transition was predicted. This predicted state was compared with the actual sixth visit (last visit). To calculate the prediction paths using more than 1 visit interval, the only input required was the baseline visit state. Then, using the dwelling time and the most likely transition connection information, we traced multiple transitions on the trained model from the baseline state until the aimed interval was reached. To maximize training efficiency, the selection of participants in the training set was prioritized to include those with a longer observation period. The remaining 27 eyes, which had fewer than 6 visits per observation period, formed an independent sample (validation set).

Comparison of Mean Absolute Errors between the 2-Dimensional Continuous-Time Hidden Markov Model and Conventional Statistical Approaches

To assess the performance of the hidden Markov model against conventional statistical approaches, MAEs of both cRNFL thickness and VFI were calculated. The MAEs then were compared between the hidden Markov model and OLSR as a conventional linear regression model and a state-of-the-art method of BJLR.[11] Because 2D CT HMM requires information from only 1 visit for prediction, we restricted the number of visits for acquiring the conventional