Article

Predicting Future Self-Reported Motor Vehicle Collisions in Subjects with Primary Open-Angle Glaucoma Using the Penalized Support Vector Machine Method

Kenya Yuki[1], Ryo Asaoka[2], Sachiko Awano-Tanabe[1], Takeshi Ono[1], Daisuke Shiba[1], Hiroshi Murata[2], and Kazuo Tsubota[1]

[1] Department of Ophthalmology, Keio University School of Medicine, Shinanomachi 35, Shinjuku-ku, Tokyo, Japan
[2] Department of Ophthalmology, the University of Tokyo, Graduate School of Medicine, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Purpose: We predict the likelihood of a future motor vehicle collision (MVC) from visual function data, attitudes to driving, and past MVC history using the penalized support vector machine (pSVM) in subjects with primary open-angle glaucoma (POAG).

Methods: Patients with POAG were screened prospectively for eligibility and 185 were analyzed in this study. Self-reported MVCs of all participants were recorded for 3 years from the baseline using a survey questionnaire every 12 months. A binocular integrated visual field (IVF) was calculated for each patient by merging a patient's monocular Humphrey Field Analyzer (HFA) visual fields (VFs). The IVF was divided into six regions, based on eccentricity and the right or left hemifield, and the average of the total deviation (TD) values in each of these six areas was calculated. Then, the future MVCs were predicted using various variables, including age, sex, 63 variables of 52 TD values, mean of the TD values, visual acuities (VAs), six sector average TDs with (pred$_{penSVM\_all}$) and without (pred$_{penSVM\_basic}$) the attitudes in driving, and also past MVC history, using the pSVM method, applying the leave-one-out cross validation.

Results: The relationship between pred$_{penSVM\_basic}$ and the future MVC approached significance (odds ratio = 1.15, [0.99–1.29], $P = 0.064$, logistic regression). A significant relationship was observed between pred$_{penSVM\_all}$ and the future MVC (odds ratio = 1.21, $P = 0.0015$).

Conclusions: It was useful to predict future MVCs in patients with POAG using visual function metrics, patients' attitudes to driving, and past MVC history, using the pSVM.

Translational Relevance: Careful consideration is needed when predicting future MVCs in POAG patients using visual function, and without driving attitude and MVC history.

Motor vehicle collisions (MVCs) are an incredibly serious public health problem. More than 30,000 people were killed in MVCs in the United States in 2013 alone (data available in the public domain at http://www-fars.nhtsa.dot.gov/Main/index.aspx), while 4113 people were killed in MVCs in Japan last year (data available in the public domain at http://www.jtsa.or.jp/topics/T-254.html). Driving clearly is a vision-related task, so most developed countries have implemented vision screening systems to determine drivers' fitness to drive.[1]

Glaucoma is a disease characterized by progressive retinal ganglion cell loss with concomitant peripheral visual field (VF) damage, which usually proceeds to central VF impairment and visual acuity (VA) loss.[2] It is the second leading cause of blindness in the world.[2] Aging is a significant risk factor for progression and onset of the disease.[3] Thus, in an aging society, the number of glaucoma patients worldwide is growing and estimated to be 111.8 million in 2040;[4] consequently, the number of drivers with glaucomatous VF

translational vision science & technology

defects will increase. However, to date, the association between glaucomatous VF defects and MVCs remains unclear, and there is no agreed method to predict glaucoma patients' risk of MVC.

We have reported previously that the relationship between VF damage and MVCs is weak and not straightforward to measure,[5] because patients with VF defects may try to drive more carefully. As a result, any impairment to the patient's visual function could be compensated by their driving behavior.[5] In addition, previous studies have shown that other factors, such as history of MVC, age,[6,7] sex,[8] diabetes,[9] body mass index (BMI),[8] and the use of sleeping drugs,[10,11] can influence the risk of a MVC.

In the current study, we developed a model to predict the risk of MVC using patients' VFs, VAs, self-reported assessment of driving behavior, and previous MVC history. As these variables are closely intercorrelated, we used a machine learning method to predict risk, namely the support vector machine (SVM), which has been applied to many research fields, including ophthalmology.[12] In particular, we presented the penalized SVM (pSVM), which is a hybrid of the SVM and a variable selection method,[13] and can improve prediction accuracy. Thus, obtaining self-reported MVC records longitudinally (annually), the purpose of the current study was to investigate the usefulness of the pSVM method to predict future MVC from visual function data, driving behavior, and past MVC history.

Method

This study's procedures conformed to the tenets of the Declaration of Helsinki and to national (Japanese) and institutional (Keio University School of Medicine) regulations. The study was approved by the Ethics Committee of the Keio University School of Medicine (#2010293). All study subjects gave informed, written consent before being enrolled.

Subjects

This was a prospective observational study. Japanese patients between 40 and 85 years of age who visited Keio University Hospital (Tokyo, Japan), the Iidabashi Eye Clinic (Tokyo, Japan), or the Tanabe Eye Clinic (Yamanashi, Japan) between May 1, 2011 and November 30, 2011 were screened for eligibility.

Baseline Evaluation of Subjects with Glaucoma

Patients with glaucoma were screened consecutively for eligibility using a battery of ophthalmic examinations, including slit-lamp biomicroscopy, funduscopy, gonioscopy, intraocular pressure measurements by Goldmann applanation tonometry, and VF examination with a Humphrey VF analyzer (HFA) and the 24-2 Swedish Interactive Threshold Algorithm Standard Strategy (Carl Zeiss Meditec, Dublin, CA). The findings were analyzed by two of us (TS and KY), both of whom subspecialize in glaucoma. The reliability of VF measurements was confirmed to be high with tests required to have less than a 20% fixation loss rate and less than a 15% false-positive rate.[14]

Diagnostic Criteria for Primary Open-Angle Glaucoma (POAG)

POAG was diagnosed when three findings were present: (1) glaucomatous optic cupping, represented by notch formation, generalized cup enlargement, a senile sclerotic or myopic disc, or nerve fiber layer defects; (2) glaucomatous VF defects, defined according to the criteria of Anderson and Patella[15] (a cluster of 3 or more points in the pattern deviation plot within a single hemifield [superior or inferior] with a P value $<5\%$, one of which must have a P value $<1\%$) ; and (3) an open angle observed on gonioscopy.

Exclusion Criteria

Subjects were excluded if they had an ophthalmologic disease other than POAG that could potentially compromise VA or contribute to VF loss. Subjects also were excluded if they had a decimal best-corrected VA (BCVA) of less than 0.7 in either eye, if they did not have a driver's license or drove 1 km or less per week, or had a mental disorder that prevented them from understanding the questionnaire. Of the 431 consecutive POAG patients screened, 204 were excluded. The reasons for excluding subjects were refusal to participate (3 patients), dementia (2), posterior retinal detachment (20), diabetic retinopathy (21), bullous keratopathy (2), age-related macular degeneration (5), other ocular disease (7), never had a driver's license (73), had already returned their driver's license (15), and drive less than 1 km per week (56). Consequently, 227 POAG patients were included in this prospective study.

Baseline Questionnaire of Motor Vehicle Collisions and Driving Behavior

All study participants answered the following questionnaire in Japanese (translated) at baseline ophthalmic examination:

1. Do you have a driver's license? (Yes/No/ Previously)
2. How long have you driven/did you drive a car? (__years)
3. How many kilometers per week do you normally drive? (__km)
4. Have you been involved in one or more traffic accidents in the past 5 years, including a single-car or minor accident including a slight scratch or dent, in which you were driving the car? (Yes/No)
5. How many traffic accidents have you ever been involved in, in the past 5 years? (__)
6. Circle any of the following driving situations that you try to avoid: at night, in rain, in fog, on freeways, lane changing, high speed driving, close to the car in front.

These questions were modified from the Driving Habits Questionnaire (DHQ).[16,17]

7. Have you ever failed to see the stop signal? (Yes/ No)
8. Are you aware of your VF defect when you drive? (Yes/No)
9. Do you have difficulty driving because of glaucoma? (Yes/No)

Demographic information recorded for all subjects included age, sex, height, weight, alcohol intake (yes/ no), smoking history (yes/no/previous), current and previous illnesses (e.g., systemic hypertension [HT], diabetes mellitus [DM]), and medical history, including oral medications, such as sleeping aids or sedatives, and anti-HT drugs.

Follow-up Questionnaire of MVCs

All study participants answered the following questionnaires in Japanese (translated) every 12 ± 1 months after the baseline interview:

1. Do you have a driver's license? (Yes/No/ Previously)
2. How many kilometers per week do you normally drive? (__km)
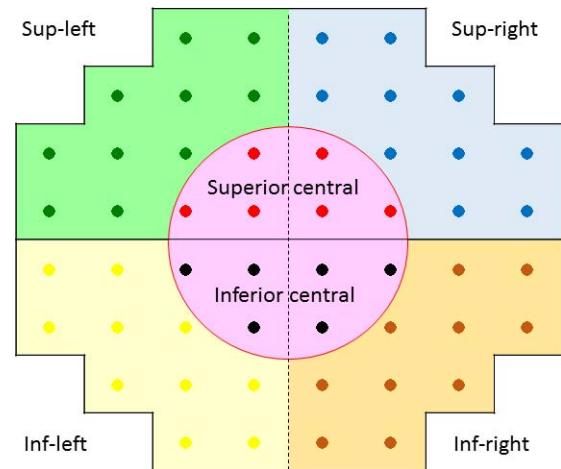3. Have you been involved in one or more traffic accidents in the past 1 year, including a single-

Figure 1. Six visual field sectors. The IVF was divided into six regions, using eccentricity and hemifield position; the average of the TD values in each area was calculated.

car or minor accident, in which you were driving the car? (Yes/No)

Subjects who answered "Yes" to Question 3 were defined as subjects with an incidence of MVCs.

Integrated Binocular Visual Field (IVF)

A binocular IVF was calculated for each patient by merging a patient's monocular HFA VFs, using the "best sensitivity" method, where the IVF total deviation (TD) at each point was calculated using the maximum TD (mTD; least negative) value from each of the two overlapping points, as if the subject was viewing the field binocularly.[18] The IVF mTD was calculated as the mean of 52 TD values across the VF. We were unable to obtain IVF data for 8 POAG subjects, and these were excluded from the following analyses. The IVF was divided into six regions, based on eccentricity and superior/inferior or right/left hemifields, as shown in Figure 1. Then, the average of the TD values in each of these six areas was calculated ($mTD_{sup-right}$, $mTD_{sup-left}$, $mTD_{sup-center}$, $mTD_{inf-right}$, $mTD_{inf-left}$, $mTD_{inf-center}$).

Statistical Analysis

For pSVM, an SVM constructs a "hyperplane" onto which each variable is projected. The hyperplane gives the largest separation margin between two classes, using the kernel function. This allows a linear separation to a nonlinear classification problem, which generally results in better discrimination performance.[19] Classification by SVM has been used widely and is one of the most powerful supervised

classification techniques, especially for high-dimension data. It has been shown that classification can be improved further by combining the SVM approach with variable selection methods.[13] Variable selection is the process of selecting a subset of relevant variables or features to be used in a model, thereby avoiding the over-fitting problem.[20] We recently reported the usefulness of this approach in analyzing clinical glaucoma VF data.[21,22] In the variable selection approach, the sum of the absolute values of the regression coefficients is constrained or penalized. The elastic net is one particular variable selection method that can be combined with the SVM approach.[23] Given P predictors $x_1,..., x_p$ and where $\hat{y}$ is the predicted value of the response y:

$$\hat{y} = \hat{\beta}_0 + x_1\hat{\beta}_1 + x_1\hat{\beta}_1 + ... + x_p\hat{\beta}_p$$

a model fitting procedure produces the vector of coefficients $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2...\hat{\beta}_p)$. In the elastic net, these estimates are defined as:[20]

$$\hat{\beta} = argmin_\beta(\| y - x\beta \|^2 + \lambda_2 \| \beta \|^2 + \lambda_1 \| \beta \|_1)$$

Thus, pSVM combines the merits of the SVM approach and the elastic net variable selection method.[24] Put simply, a penalty is applied to the independent variables when they are projected onto a hyperplane in the SVM model.[25,26]

The pSVM method, with a Gaussian RBF kernel, was used to predict future MVCs, using the 84 independent variables shown in Table 1, including better and worse VAs, the 52 TD values, the six sectorial average TD values, driving attitude, past history of MVCs, HT status, DM status, use of anti-HT drugs, smoking habits, alcohol intake, years with driving license, distance driven per week, BMI, and use of sleep aid/sedatives. This pSVM model is denoted, pred$_{penSVM\_all}$. We also fitted a standard SVM model (without variable selection) and denote this model, pred$_{SVM}$. As the dataset is unbalanced with respect to the measured outcome (merely 28 patients experienced a MVC on self-report ["MVC+" group], while 157 patients did not ["MVC–" group]), we applied the synthetic minority over-sampling technique (SMOTE) algorithm[25] to generate an "artificial" dataset with a balanced outcome. In addition, weights for the different classes were applied: a weight of 157/185 for the MVC+ group and a weight of 28/185 for the MVC– group. Future MVCs were predicted using the pSVM model and leave-one-out cross validation. In leave-one-out cross validation, data from a single patient are used as

Table 1. Variables Used in the pSVM.

63 Variables	84 Variables
Age	Age
Sex	Sex
Better BCVA	Better BCVA
Worse BCVA	Worse BCVA
52 TD values	52 TD values
mTD$_{sup-right}$	mTD$_{sup-right}$
mTD$_{sup-left}$	mTD$_{sup-left}$
mTD$_{sup-center}$	mTD$_{sup-center}$
mTD$_{inf-right}$	mTD$_{inf-right}$
mTD$_{inf-left}$	mTD$_{inf-left}$
mTD$_{inf-center}$	mTD$_{inf-center}$
mTD	mTD
	No care
	Avoid night
	Avoid rain
	Avoid fog
	Avoid highway
	Avoid high-speed driving
	Avoid lane change
	Car distance
	Signal
	Taking care of VF
	Difficulty of driving
	HT
	DM
	anti HT drug
	Smoking habit
	Alcohol intake
	Years with driver's license
	Distance driven per week (km)
	BMI
	Use of sleep aid / sedatives, yes/no
	Past history of MVC
	no/once/more than two times

A total of 62 variables were used in the predpenSVM_basic model, and 84 variables were used in the predSVM and predpenSVM_all models. sup, superior; inf, inferior.

validation data, and all data from the remaining patients are used as training data ($N = 184/185$). This procedure is repeated until each patient in the original sample is used once as validation data (i.e., 185 times). In other words, for each individual, only the data from all other subjects are used in the prediction model. Finally, the relationship between the pSVM-predicted outcome and the self-reported actual incidence of MVCs was analyzed using logistic

Table 2. Characteristics of the Study Population

Variables, mean ± SD (range)	Value
Age, y	63.6 ± 10.2 (40–84)
Sex male:female	137:48
Better BCVA, logMAR	0.0036 ± 0.018 (0.00–0.15)
Worse BCVA, logMAR	0.015 ± 0.038 (0.00–0.15)
mTD, dB	−1.7 ± 3.2 (−20.3–3.2)
BMI, kg/m^2	22.7 ± 2.9 (15.6–31.2)

regression. For comparison, we also generated a pSVM using only a subset of 63 variables (see Table 1), namely the 52 TD values, the six sector average TDs, mTD, VAs, age, and sex. This model is denoted, $pred_{penSVM\_basic}$.

All statistical analyses were performed using the statistical programming language R (ver. 3.1.3, The R Foundation for Statistical Computing, Vienna, Austria). The R packages "penalizedSVM," "kernlab," and "DMwR" were used to generate penalized SVM models, standard SVM models, and the SMOTE method, respectively. Benjamini's method was used to correct P values for the problem of multiple testing.[27] Numerical values between two groups were compared using the Mann-Whitney U test and the relationship between two categorical values was analyzed using the χ^2 test.

Results

We studied 185 patients with POAG: 137 (74%) male and 48 (26%) female patients, average age 63.6 ± 10.2 SD (range, 40–84; Table 2). Average BMI (kg/m^2) was 22.7 ± 2.9 (range, 15.6–31.2). There were 119 rural, and 51 (Iidabashi eye clinic) and 15 (Keio University Hospital) urban patients. Average better and worse BCVA was 0.0036 ± 0.018 (0.00–0.15) and 0.015 ± 0.038 (0.00–0.15, logMAR), respectively. Mean TD value was −1.7 ± 3.2 (−20.3–3.2) dB. Among the 185 patients, 28 experienced a MVC during the study period. The self-reported MVC incidence rate was 5.0%/y.

Table 3 shows a comparison of variables between the MVC+ and MVC− groups. There was no significant difference between any of the numerical values, except for worse VA ($P = 0.010$, Mann-Whitney U test). Further, none of the categorical variables showed a significant relationship with MVC incidence.

Figures 2 to 4 show the results of logistic regression associated with three models: $pred_{penSVM\_basic}$,

$pred_{SVM}$, and $pred_{penSVM\_all}$. There was a significant relationship between predicted MVC in $pred_{penSVM\_all}$ (odds ratio [OR] = 1.21; 95% confidence interval [CI], 1.08–1.31; $P = 0.0015$, logistic regression, Fig. 4), but not in $pred_{penSVM\_basic}$ (OR = 1.15; 95% CI, 0.99–1.29; $P = 0.064$, Fig. 2) and not in $pred_{SVM}$ (OR = 1.26; 95% CI, 0.97–1.49; $P = 0.085$, Fig. 3).

Discussion

In the current study, self-reported MVCs were recorded in 185 patients with POAG. The pSVM model including all variables ($pred_{penSVM\_all}$), namely visual function data, driving behavior, systemic status, and past history of MVCs, resulted in a model that was associated significantly with the correct prediction of a future MVC. This was not the case when only the subset of 63 mainly vision-related (VF/VA) variables were included in the model ($pred_{penSVM\_basic}$).

In a real world clinic, not all variables included in the $pred_{penSVM\_all}$ model, such as questions regarding driving behavior and past history of MVCs, are readily available to a clinician. However, it appears that these variables are important to predict accurately the risk of a future MVC. Many studies have suggested a significant relationship between the deterioration in patients' VAs (and/or their VF) and driving ability.[28–37] However most of these studies assessed driving ability in an artificial situation.[28,37–41] For example, Crabb et al.[39] investigated eye movements during a PC-based driving simulation and suggested that superior VF damage is disadvantageous for driving ability. However, it is not entirely clear whether this could contribute to a person's actual risk of a MVC. Tatham et al.[37] investigated the relationship between MVCs in drivers with glaucoma and visual function, and found that the reaction times to low contrast divided attention tasks during driving simulation were significantly associated with a history of MVC. Kunimatsu et al.[28] reported that patients with advanced glaucoma were involved in a significantly higher number of collisions in a driving simulator than controls. Szylk et al.[40] reported that the number of accidents, measured in a driving simulator, in a glaucoma group was correlated significantly with peripheral VF extent. Szykl et al.[41] also reported that lower contrast sensitivity (in the eye with better contrast sensitivity) in subjects with glaucoma correlated with driving skills, again evaluated by a driving simulator. All of these previous studies have

Table 3. Comparison of Ocular and Systemic Variables between the MVC+ and MVC− Groups

Variables Mean ± SD (range)	MVC+	MVC−	P Value
Age, y	66.9 ± 11.4 (46–83)	63.2 ± 10.0 (40–84)	0.032
Sex, male:female	20:8	112:40	0.91
Better BCVA	0.00 ± 0.00 (0.00–0.00)	0.00 ± 0.02 (0.00–0.15)	0.20
Worse BCVA	0.01 ± 0.04 (0.00–0.15)	0.03 ± 0.04 (0.00–0.15)	0.010
mTD$_{sup-right}$, dB	−3.5 ± 6.1 (−21.5–3.3)	−2.4 ± 5.5 (−29.1–3.9)	0.98
mTD$_{sup-left}$, dB	−2.3 ± 5.4 (−23.5–3.5)	−2.6 ± 6.2 (−27.8–4.6)	0.31
mTD$_{sup-center}$, dB	−3.4 ± 7.0 (−23.2–2.7)	−2.2 ± 5.3 (−32.3–3.3)	0.56
mTD$_{inf-right}$, dB	−0.79 ± 2.9 (−11.0–3.0)	−0.43 ± 2.7 (−18.0–3.3)	0.88
mTD$_{inf-left}$, dB	−1.5 ± 3.9 (−11.8–3.3)	−0.91 ± 2.9 (−13.7–3.6)	0.74
mTD$_{inf-center}$, dB	−1.0 ± 2.9 (−10.3–2.9)	−1.0 ± 3.6 (−19.0–3.7)	0.58
mTD, dB	−2.1 ± 3.9 (−11.8–2.8)	−1.6 ± 3.7 (−20.0–3.2)	0.58
Care while driving, Yes:No	3:20	20:137	0.77
Avoid night, Yes:No	7:21	50:107	0.47
Avoid rain, Yes:No	5:23	30:127	0.87
Avoid fog, Yes:No	5:23	20:137	0.47
Avoid highway, Yes:No	2:26	23:134	0.28
Avoid high speed, Yes:No	18:10	89:68	0.45
Avoid lane change, Yes:No	6:22	16:141	0.092
Car distance, Yes:No	16:12	76:81	0.39
Signal, Yes:No	0:28	10:147	0.17
Care of VF, Yes:No	3:25	31:126	0.26
Difficulty of driving, Yes:No	3:25	25:132	0.48
Systemic hypertension, Yes:No	7:21	55:102	0.30
Diabetes mellitus, Yes:No	4:24	27:130	0.70
Anti hypertensive drug, Yes:No	6:22	52:105	0.22
Smoking habit, Yes:Previously:No	2:8:18	26:54:77	0.26
Alcohol intake, Daily:Sometimes:No	10:6:12	41:43:73	0.55
How long, y	38.8 ± 13.5 (8.0–60.0)	38.6 ± 10.0 (17.0–60.0)	0.70
How far (km)	98 ± 89 (5–300)	90 ± 125 (1 to 600)	0.076
BMI (kg/m^2)	22.4 ± 2.5 (16.9–25.8)	22.8 ± 3 (15.6–31.2)	0.74
Use of Sleeping aid, Yes:No	1:27	4:153	0.76
Sedative, Yes:No	0:28	27:130	0.67

suggested that damage to visual function is associated with driving performance, but none investigated the actual impact in real life. We have shown previously that the relationship between VAs and VF, against actual MVC is weak.[5]

There is no doubt that motor vehicle driving is a complicated task; drivers must be aware of hazards and focus on their future direction.[5] Furthermore, age,[6,42] DM,[9] BMI,[8] use of sleep aids,[11] and physiologic disabilities all have an influence on driving performance. Driving accidents can cause fatalities and, hence, drivers usually are very concentrated, and seek to detect hazards by saccadic eye and head movements.[43] Careful consideration is required

when investigating patients' driving ability from only clinical information.

For many glaucoma patients around the world, motor vehicle driving is central to their daily life, and they may depend on driving for commuting to work and grocery shopping. This is especially the case for patients who live in rural areas, where public transport may not be a viable option. Banning these patients, in particular, from driving could seriously affect their quality of life. Motor vehicle driving also is imperative for those whose professions depend on it, such as taxi drivers. The current study suggests that assessing the deterioration of glaucoma patients' visual function is not sufficient to predict accurately
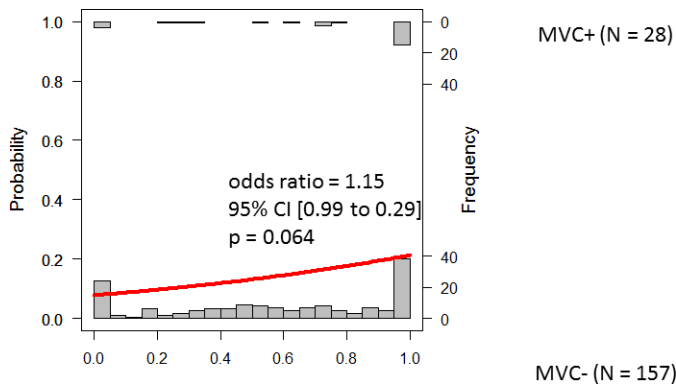
Figure 2. Histograms of future MVCs with fitted logistic regression line. The histograms in the *top* and *bottom* represent the possibility of a future MVC in the MVC+ and MVC− groups, respectively. This Figure shows results obtained with the pred$_{penSVM\_basic}$ model; this model included up to 62 variables (as shown in Table 2): 52 TD values, VAs, age, sex, and the six sectorial average TDs.

the risk of future MVC. Indeed, many other risk factors should be considered (see 84 variables list in Table 1) and the pSVM method could offer clinicians a useful tool to better understand patients' risk of MVC. In addition, none of the nonvisual function parameters was significantly different between MVC+ and MVC− groups. However, these multiple parameters were useful to improve the prediction accuracy when used simultaneously with the pSVM. This may be because these nonvisual function parameters are not independent of each other and there may be a pattern (combination) related to MVC+ or MVC−. In the pSVM, multiple parameters were interpreted simultaneously on the projected Kernel dimension.
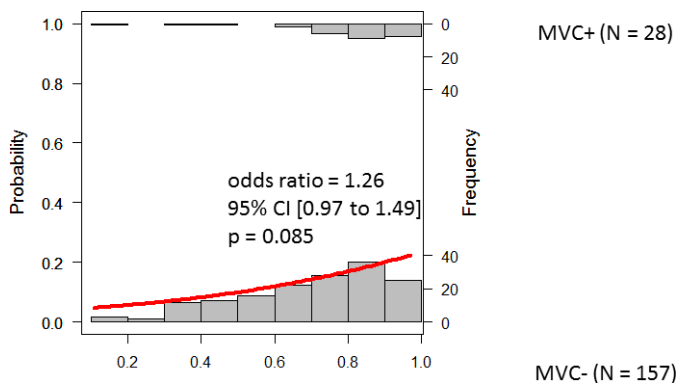

Figure 3. Histograms of future MVCs with fitted logistic regression line. The histograms in the *top* and *bottom* part represent the possibility of a future MVC in the MVC+ and MVC− groups, respectively. This Figure shows results obtained with the pred$_{SVM}$ model; this model included up to 84 variables (as shown in Table 2).
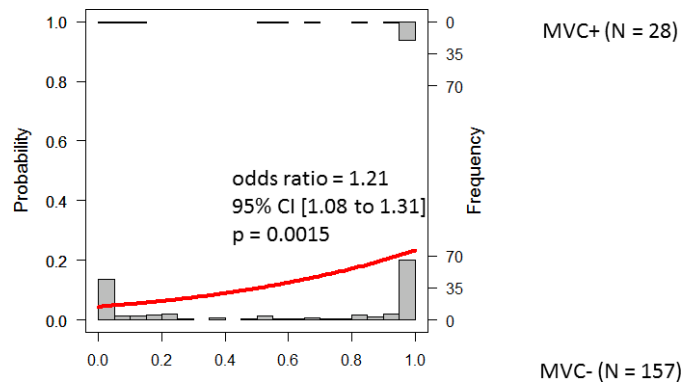

Figure 4. Histograms of future MVC with fitted logistic regression line. The histograms in the *top* and *bottom* part represent the possibility of a future MVC in the MVC+ and MVC− groups, respectively. This Figure shows the results obtained with the pred$_{penSVM\_all}$ model; this model included up to 84 variables (as shown in Table 2).

In other words, the current results suggested the Kernel plane could determine a useful pattern consisting of various visual and nonvisual parameters for predicting future MVC. Indeed, adding none of the single nonvisual function parameters to penSVM_basic did result in the improvement of prediction accuracy (data not shown in the Results).

It remains controversial whether better-eye or worse-eye VA is more influential on MVCs. In the current study, a significant difference was not observed between better VA in the MVC+ and MVC− groups, but worse VA was significantly healthier in the MVC+ compared to the MVC− group. The controversy and contradictory results on this issue may be attributed to differences in the studied populations. For instance, in a population that mainly consists of early stage glaucoma patients it is likely there will be a wider range of measurements in worse than in better VA. The opposite tendency would be seen in a population with very advanced glaucoma. Such differences will have an influence on the results coming from an analysis of the relationship between better and worse VAs on MVCs.

Classification models sometimes can be over-fitted to the sample dataset, particularly when the number of independent variables is large. Variable selection models with penalization have been proposed to overcome this problem. This may be a reason why pred$_{penSVM\_all}$ was shown to provide a significant prediction, but pred$_{SVM}$ was not, because 84 variables were considered. If clinicians had access to an accurate MVC risk model they could offer an appropriate warning to patients about their individual

risk, and it also may act as an appropriate intervention point to offer driving re-education lessons and advice, such as using a driving simulation tool. However, further investigations are necessary looking at the full effects of glaucoma on MVC risk.

A limitation of our study is that we relied on self-reporting of MVCs as the main outcome.[44] Another limitation is that we were unable to follow all the participants over the 3-year period; subjects who were lost to follow-up could introduce a bias in our results if the reasons for leaving the study were associated with the study outcome. In addition, the studied patients mainly consisted of those with relatively early stage glaucoma, as suggested by the average mTD values of −2.1 and −1.6 dB in the MVC+ and MVC− groups, respectively. A further study is needed to investigate the relevance of the current results in more advanced glaucomatous cases. Other limitations include the predominance of male study subjects and a limited number of patients with systemic complications. It would be interesting to see whether the current results change in patients with systemic complications. Nonetheless, the purpose of the current study was to develop an algorithm to predict future MVC in "real world" patients. We recruited patients without bias to systemic conditions. Also, we focused on a single disease rather than collecting mixed data. Other forms of glaucoma, such as angle closure glaucoma, may have different findings.

In conclusion, it was useful to predict future MVCs from visual function metrics, driving behavior, and MVC history, using the penalized support vector machine. A robust classification model with a wide number of different variables is needed to predict MVCs, because the relationship between visual function and other factors that influence the risk of MVCs is not straightforward.

Acknowledgments

References

1. Desapriya E, Harjee R, Brubacher J, et al. Vision screening of older drivers for preventing road traffic injuries and fatalities. *Cochrane Database Syst Rev*. 2014; Cd006252.
2. Quigley HA. Glaucoma. *Lancet*. 2011;377:1367–1377.
3. Coleman AL, Miglior S. Risk factors for glaucoma onset and progression. *Surv Ophthalmol*. 2008;53(suppl 1):S3–S10.
4. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014; 121:2081–2090.
5. Yuki K, Asaoka R, Tsubota K. The relationship between central visual field damage and motor vehicle collisions in primary open-angle glaucoma patients. *PloS One*. 2014;9:e115572.
6. Brown LH. Senior drivers: risks, interventions, and safety. *Nurse Pract*. 2006;31:38–40, 43-34, 49.
7. Braver ER, Trempel RE. Are older drivers actually at higher risk of involvement in collisions resulting in deaths or non-fatal injuries among their passengers and other road users? *Inj Prev*. 2004;10:27–32.
8. Carter PM, Flannagan CA, Reed MP, Cunningham RM, Rupp JD. Comparing the effects of age, BMI and sex on severe injury (AIS 3+) in motor-vehicle crashes. *Accid Anal Prev*. 2014;72: 146–160.
9. Inkster B, Frier BM. Diabetes and driving. *Diabetes Obes Metab*. 2013;15:775–783.
10. Kripke DF. Mortality Risk of Hypnotics: Strengths and limits of evidence. *Drug Saf*. 2016;39:93–107.
11. Hansen RN, Boudreau DM, Ebel BE, Grossman DC, Sullivan SD. Sedative hypnotic medication use and the risk of motor vehicle crash. *Am J Public Health* 2015;105:e64–e69.
12. Hirasawa H, Murata H, Mayama C, Araie M, Asaoka R. Evaluation of various machine learning methods to predict vision-related quality of life from visual field data and visual acuity in patients with glaucoma. *Br J Ophthalmol*. 2014; 98:1230–1235.
13. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. *J R Stat Soc Ser B*. 2005;301–320.
14. Bengtsson B, Heijl A. False-negative responses in glaucoma perimetry: indicators of patient perfor-

mance or test reliability? *Invest Ophthalmol Vis Sci.* 2000;41:2201–2204.

15. Anderson DR, Patella VM. *Automated Static Perimetry*. Mosby, St. Louis, MO; 1999.

16. Owsley C, Stalvey B, Wells J, Sloane ME. Older drivers and cataract: driving habits and crash risk. *J Gerontol A Biol Sci Med Sci.* 1999;54:203–211.

17. McGwin G, Jr., Mays A, Joiner W, Decarlo DK, McNeal S, Owsley C. Is glaucoma associated with motor vehicle collision involvement and driving avoidance? *Invest Ophthalmol Vis Sci.* 2004;45:3934–3939.

18. Crabb DP, Viswanathan AC. Integrated visual fields: a new approach to measuring the binocular field of view and visual disability. *Graefes Arch Clin Exp Ophthalmol.* 2005;243:210–216.

19. Cristianini N. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, UK: Cambridge University Press; 2000.

20. Bermingham ML, Pong-Wong R, Spiliopoulou A, et al. Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Sci Rep.* 2015;5:10312.

21. Asaoka R. Measuring visual field progression in the central 10 degrees using additional information from central 24 degrees visual fields and 'lasso regression'. *PloS One.* 2013;8:e72199.

22. Fujino Y, Murata H, Mayama C, Asaoka R. Applying "lasso" regression to predict future visual field progression in glaucoma patients. *Invest Ophthalmol Vis Sci.* 2015;56:2334–2339.

23. Li, Wang Zhu J, Zou H. The doubly regularized support vector machine. *Stat Sin.* 2006;16:589–615.

24. Becker N, Werft W, Toedt G, Lichter P, Benner A. PenalizedSVM: a R-package for feature selection SVM classification. *Bioinformatics.* 2009;25:1711–1712.

25. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. Smote: Synthetic minority over-sampling technique. *J Artif Intell Res.* 2000;16:321–357.

26. Chang YW, Hsieh CJ, Chang KW, Ringgaard M, Lin CJ. Training and testing low-degree polynomial data mappings via linear SVM. *JMLR.* 2010;11:1471–1490.

27. Benjamini Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B.* 1995;57:289–300.

28. Kunimatsu-Sanuki S, Iwase A, Araie M, et al. An assessment of driving fitness in patients with visual impairment to understand the elevated risk of motor vehicle accidents. *BMJ Open.* 2015;5:e006379.

29. Bowers A, Peli E, Elgin J, McGwin G Jr, Owsley C. On-road driving with moderate visual field loss. *Optom Vis Si.* 2005;82:657–667.

30. Charman WN. Vision and driving–a literature review and commentary. *Ophthalmic Physiol Opt.* 1997;17:371–391.

31. Crabb DP, Fitzke FW, Hitchings RA, Viswanathan AC. A practical approach to measuring the visual field component of fitness to drive. *Br J Ophthalmol.* 2004;88:1191–1196.

32. Cross JM, McGwin G Jr, Rubin GS, et al. Visual and medical risk factors for motor vehicle collision involvement among older drivers. *Br J Ophthalmol.* 2009;93:400–404.

33. Gracitelli CP, Tatham AJ, Boer ER, et al. Predicting risk of motor vehicle collisions in patients with glaucoma: a longitudinal study. *PloS One.* 2015;10:e0138288.

34. Huisingh C, McGwin G Jr, Wood J, Owsley C. The driving visual field and a history of motor vehicle collision involvement in older drivers: a population-based examination. *Invest Ophthalmol Vis Sci.* 2015;56:132–138.

35. Johnson CA, Keltner JL. Incidence of visual field loss in 20,000 eyes and its relationship to driving performance. *Arch Ophthalmol.* 1983;101:371–375.

36. Kwon M, Huisingh C, Rhodes LA, McGwin G Jr, Wood JM, Owsley C. Association between glaucoma and at-fault motor vehicle collision involvement among older drivers: a population-based study. *Ophthalmology.* 2016;123:109–116.

37. Tatham AJ, Boer ER, Gracitelli CP, Rosen PN, Medeiros FA. Relationship between motor vehicle collisions and results of perimetry, useful field of view, and driving simulation in drivers with glaucoma. *Trans Vis Sci Technol.* 2015;4:5.

38. Crabb DP, Smith ND, Rauscher FG, et al. Exploring eye movements in patients with glaucoma when viewing a driving scene. *PloS One.* 2010;5:e9710.

39. Glen FC, Smith ND, Crabb DP. Impact of superior and inferior visual field loss on hazard detection in a computer-based driving test. *Br J Ophthalmol.* 2015;99:613–617.

40. Szlyk JP, Mahler CL, Seiple W, Edward DP, Wilensky JT. Driving performance of glaucoma patients correlates with peripheral visual field loss. *J Glaucoma.* 2005;14:145–150.

41. Szlyk JP, Taglia DP, Paliga J, Edward DP, Wilensky JT. Driving performance in patients with mild to moderate glaucomatous clinical vision changes. *J Rehabil Res Dev.* 2002;39:467–482.

42. Hamel J, De Beukelaer S, Kraft A, Ohl S, Audebert HJ, Brandt SA. Age-related changes in visual exploratory behavior in a natural scene setting. *Front Psychol*. 2013;4:339.

43. Kasneci E, Sippel K, Aehling K, et al. Driving with binocular visual field loss? A study on a supervised on-road parcours with simultaneous eye and head tracking. *PloS One*. 2014;9:e87470.

44. Marottoli RA, Cooney LM Jr, Tinetti ME. Self-report versus state records for identifying crashes among older drivers. *J Gerontol A Biol Sci Med Sci*. 1997;52:184–187.