MACHINE LEARNING CLASSIFIERS DETECT SUBTLE FIELD DEFECTS IN EYES OF HIV INDIVIDUALS

BY Igor Kozak MD, Pamela A. Sample PhD, Jiucang Hao MS, William R. Freeman MD, **Robert N. Weinreb MD**, Te-Won Lee PhD, AND **Michael H. Goldbaum MD***

ABSTRACT

Purpose: To test the following hypotheses: (1) eyes from patients with human immunodeficiency virus (HIV) have retinal damage that causes subtle field defects, (2) sensitive machine learning classifiers (MLCs) can use these field defects to distinguish fields in HIV patients and normal subjects, and (3) the subtle field defects form meaningful patterns. We have applied supervised MLCs—support vector machine (SVM) and relevance vector machine (RVM)—to determine if visual fields in patients with HIV differ from normal visual fields in HIV-negative controls.

Methods: HIV-positive patients without visible retinopathy were divided into 2 groups: (1) 38 high-CD4 (H), 48.5 ± 8.5 years, whose CD4 counts were never below 100; and (2) 35 low-CD4 (L), 46.1 ± 8.5 years, whose CD4 counts were below 100 at least 6 months. The normal group (N) had 52 age-matched HIV-negative individuals, 46.3 ± 7.8 years. Standard automated perimetry (SAP) with the 24-2 program was recorded from one eye per individual per group. SVM and RVM were trained and tested with cross-validation to distinguish H from N and L from N. Area under the receiver operating characteristic (AUROC) curve permitted comparison of performance of MLCs. Improvement in performance and identification of subsets of the most important features were sought with feature selection by backward elimination.

Results: SVM and RVM distinguished L from N (L: AUROC = 0.804, N: 0.500, P = .0002 with SVM and L: .800, P = .0002 with RVM) and H from N (H: 0.683, P = .022 with SVM and H: 0.670, P = .038 with RVM). With best-performing subsets derived by backward elimination, SVM and RVM each distinguished L from N (L: 0.843, P < .00005 with SVM and L: 0.870, P < .00005 with RVM) and H from N (H: 0.695, P = .015 with SVM and H: 0.726, P = .007 with RVM). The most important field locations in low-CD4 individuals were mostly superior near the blind spot. The location of important field locations was uncertain in high-CD4 eyes.

Conclusions: This study has confirmed that low-CD4 eyes have visual field defects and retinal damage. Ranking located important field locations superiorly near the blind spot, implying damage to the retina inferiorly near the optic disc. Though most fields appear normal in high-CD4 eyes, SVM and RVM were sufficiently sensitive to distinguish these eyes from normal eyes with SAP. The location of these defects is not yet defined. These results also validate the use of sensitive MLC techniques to uncover test differences not discernible by human experts.

INTRODUCTION

Humans are innately able to recognize patterns such as faces by sight, voices by sound, and a comb by feel. Our brains are constructed to learn to do these pattern recognition tasks. Pattern recognition is the act of taking in a sample of data represented as a pattern and making a reasoned decision about the category of the sample. Machine learning is the use of algorithms and methods by which computers are set up to learn. Machine learning classifiers (MLCs) comprise a number of computer algorithms that learn to map a pattern of features (eg, perimetry) into a choice of classes. State-of-the-art MLCs are approaching the performance of the theoretical optimal classifier for a given set of data. MLCs have been previously applied to ophthalmologic problems, such as the interpretation and classification of visual fields,[1,2] detection of visual field progression,[3,4] assessment of the structure of the optic nerve head,[5,6] measurement of retinal nerve fiber layer thickness,[7,8] and separation of noise from visual field information.[9] We found support vector machine (SVM) and relevance vector machine (RVM) to be particularly effective MLCs for discriminating between normal and glaucomatous visual fields.[1,10]

Machine learning classifiers can be trained to distinguish the group identity of patterns, sometimes with greater sensitivity than a human expert. An important problem concerning patients with human immunodeficiency virus (HIV) is whether the retina is damaged by HIV infections in eyes that have no visible evidence of retinopathy. However, human perimetric experts have difficulty finding evidence for visual field loss in patients with HIV. Visual fields can detect retinal damage; we use the sensitivity of MLCs to detect eyes with retinal damage and to locate the damage.

Highly active antiretroviral therapy (HAART) has increased life expectancy and the quality of life with the persistent suppression of HIV viremia and has dramatically decreased the incidence of cytomegalovirus (CMV) retinitis.[11,12] Even if patients on HAART do not develop infectious retinitis, with prolonged low CD4 cell status they are prone to developing HIV retinopathy.[13] It is an ischemic retinopathy ophthalmoscopically characterized by cotton-wool spots, intraretinal hemorrhages, and capillary nonperfusion.[14]

Previous studies have shown that individuals with HIV retinopathy without infectious retinitis, particularly those with low CD4

counts, show deficits in visual function, even though the central visual acuity, as measured by Snellen and similar visual acuity charts, is preserved, likely because of the redundancy in this visual pathway.[15-20] Several centers, including our own, have demonstrated reduced sensitivity in the field of vision,[15,16] color and contrast sensitivity test of central vision,[17-20] and electrophysiological testing[21] of individuals from the preHAART era. Further studies in the HIV-positive patient population without retinitis have shown that there are particular topographic patterns of this visual field (VF) loss.[22] Involvement of the inner retina from retinovascular disease in HIV patients, including retinal cotton-wool spots, microaneurysms, capillary dropout, and ischemia, is assumed to damage the ganglion cell layer and retinal nerve fiber layer.[13,23,24] We have recently reported thinning in the retinal nerve fiber layer of patients with a history of low CD4 counts, using high-resolution optical coherence tomography.[25] HIV patients with high CD4 counts may also have retinopathy; some patients report changes in the visual field. We hypothesize that optimized pattern recognition methods can learn to distinguish between fields from eyes in HIV patients with previously undetected field loss and fields from normal eyes.

We have applied supervised MLCs—SVM and RVM—to determine if visual fields in HIV patients differ from normal visual fields in HIV-negative controls. Since the immune function presumably was better in the high-CD4 group, we expected HIV retinopathy damage to be less than in the low-CD4 group. The hypotheses tested are (1) eyes from HIV patients have retinal damage that causes subtle field defects, (2) sensitive MLCs can find and use these subtle field defects to distinguish between fields in HIV patients and normal subjects, and (3) the subtle field defects in eyes from HIV patients form meaningful patterns.

METHODS

SUBJECTS

Patients with positive serology for HIV were recruited from the University of California, San Diego, AIDS Ocular Research Unit at the Jacobs Retina Center in La Jolla, California. Patients were enrolled in an institutional review board–approved longitudinal study of HIV disease (HAART protocol), and informed consent for imaging and data collection was obtained from the patients. Non-HIV controls were age-matched healthy participants in the HAART study and nonglaucomatous healthy controls from the ongoing National Eye Institute–sponsored longitudinal Diagnostic Innovations in Glaucoma Study (DIGS) of visual function in glaucoma.

The patients were divided into 3 groups. Group H consisted of HIV-positive patients whose medical records showed that their CD4 cell counts were never below 100 (1.0×10^9/L). Included in group L were HIV-positive patients with CD4 cell counts below 100 at some point of time in their medical history lasting for at least 6 months. All HIV patients were treated with HAART therapy prior to and at the time of the examination. Group N consisted of HIV-negative patients who served as normal control subjects. The HIV individuals had no significant ocular disease or eye surgery. Group N was age-matched to the HIV groups. The L and H groups were combined, and the patients were put into five 10-year bins between 10 and 60 years. The N group was adjusted with random removal of eyes from each of the bins to make the bin proportions equivalent to the HIV group. One eye was randomly selected from each patient.

OPHTHALMOLOGIC EVALUATION

Each patient was given a complete eye examination with indirect ophthalmoscopy and intraocular pressure measurement. Included in the study were patients with central visual acuity equal to or better than 20/40 and good ocular media. Excluded were patients with a history of ocular disease or surgery. Concurrent or healed CMV retinitis in an eye was an exclusion criterion. Excluded also were eyes with spherical refraction higher than ±5 diopters, amblyopia, glaucoma, or suspicion of glaucoma by disc or field.

Visual Field Testing

A Humphrey Visual Field Analyzer (model 620; Carl Zeiss Meditec, Dublin, California), standard automated perimetry (SAP) program 24-2, and routine settings for evaluating visual fields were used. Visual fields were taken within 1 week of the ophthalmologic examination. Patients were given a brief test before the second eye was examined. Only reliable VFs, defined as those with less than 33% false-positives, 33% false-negatives, and 33% fixation losses, were used.

Data Input and Class Output for the Classifier

The absolute sensitivity (in decibels) of the 54 visual field locations, excluding the two located in the blind spot, formed a vector in 52-dimensional input space for each of the 125 SAP fields of normal and HIV eyes.[1,11] SAPs from the left eye were flipped horizontally to make all the fields appear as right eyes for presentation as input to the machines. The output was the prediction of class (N or L, N or H) for SVM or probability of class for RVM.

MACHINE LEARNING CLASSIFIERS

Support vector machine is a machine classification method that seeks the boundary that best separates the data into 2 classes.[26,27] SVM uses a kernel to map the feature space (eg, a pattern of field locations) into a high-dimensional transformed space. We selected a Gaussian kernel because it surpassed the linear kernel with visual field data in previous research on SAP.[10] In high-dimensional space, a linear hyperplane can be an effective boundary. The samples that are closest to the separation boundary are the most difficult to classify. These samples, called support vectors (SVs), determine the location and orientation of the separation boundary. The margin is defined as the distance between the hyperplane and SVs in the transformed space. The SVM learns by locating the hyperplane to maximize the margin between the 2 classes. Because easy-to-classify examples do not help to determine the location and orientation of the separating boundary, the boundary based on the more-difficult-to-classify examples can separate the 2 classes more effectively.

The learned hyperplane in the transformed space maps back to the feature space as a nonlinear separation boundary.[27] Thus SVM is very adaptable and often outperforms other classifiers. By using the kernel, SVM is able to handle data with high dimensions because it overcomes difficulties induced by an input with a large number of features (the curse of dimensionality). Usually only a very small portion of samples become SVs, yielding sparse representation that simplifies and speeds the learning. SVMs have been used for various clinical medicine classification applications including the detection of glaucoma.[10,28]

The SVM was implemented by using Platt's sequential minimal optimization algorithm in commercial software (MatLab, version 7.0; The MathWorks, Natick, Massachusetts). For classification of the SAP data, Gaussian (nonlinear) kernels of various widths were tested, and the chosen Gaussian kernel width was the one that gave the highest area under the receiver operating characteristic (AUROC) curve using 10-fold cross-validation.

Relevance vector machine employs a kernel function similar to SVM. Unlike SVM, which focuses on the separation boundary, RVM employs a data-generative model that gives a fully probabilistic output that is desired in many clinical applications. The probability of disease is often more intuitive than a declaration of class. RVM also makes the assumption of probability distribution for some parameters. The learning and classification are achieved through Bayesian inferences. There are fewer support vectors than with SVM, which can improve learning efficiency. In this study a Gaussian kernel was used. With the Gaussian kernel function, we have observed similar performances of SVM and RVM.[8] As with the SVM, we optimize the kernel width with grid search, trying several values and choosing the one that performed best.[29,30] For classification of the SAP data, a Gaussian kernel width was chosen that gave the highest AUROC with 10-fold cross-validation. The RVM was implemented using a commercially available algorithm (SparseBayes verison 1.0; Microsoft Research, Cambridge, United Kingdom, for MatLab, The MathWorks).

Training and Testing Machine Learning Classifiers

Cross-validation is a *k*-partition rotating resampling technique that seeks an unbiased evaluation of MLC performance by using test data separate from the training data. In this study, 10-fold cross-validation was used, which randomly split each class into 10 equal subsets. The classifier was trained on a set that combined 9 of the 10 partitions, and the 10th partition of previously unseen samples served as the testing set. This procedure was performed 10 times with each subset having a chance to serve as the test set.

Performance Measure of Trained Machine Learning Classifiers

Receiver operating characteristic curves displayed the discrimination of each classifier as the separation threshold was moved from one end of the data to the other (Figure 1). The AUROC was used as a single measure of classifier performance to compare different classifiers; the larger the AUROC, the better the classifier.[31] We generated an ROC curve to represent a chance decision to permit comparison of the MLCs against chance; the predictor with performance equal to chance should have AUROC = 0.5, whereas the ideal classifier would give an AUROC = 1.0. The method of DeLong[31] tested the null hypothesis (*P* value) for comparing the AUROCs of classifiers.

We trained and tested 2 types of classifiers, RVM and SVM, to distinguish fields from normal subjects (group N) and high-CD4 patients (group H) and between normal subjects and low-CD4 patients (group L). We trained these MLCs with the full-feature set, a 10-feature subset, and a performance-peaking subset of optimal features derived by backward elimination.

The principal comparisons selected in advance were normal vs high CD4 and normal vs low CD4, each with RVM. Hence the Bonferroni adjustment was a denominator of 2. Other comparisons displayed in Tables 1 and 2 were for interest.

Optimization by Feature Selection

A high ratio of the number of samples (*n*) to the number of features (*f*) reduces the uncertainty of the location of the separation boundary between the classes. The *n*/*f* ratio can be achieved by increasing the number of samples or by reducing the number of features. We reduced the number of features with feature selection. To create small subsets of the best features, we used backward elimination with RVM. Previous research found backward elimination to work better than forward selection on visual field data.[6] Backward elimination started with the full-feature set. Each feature was given the opportunity to be the one removed to determine the effect on the RVM performance. During feature selection for each subset, 10-fold *internal* cross-validation was used to identify the subset with the highest AUROC for each size. The feature that either maximally increased or minimally decreased the RVM performance was removed, the AUROC was noted, and the process was repeated sequentially down to one feature (Figure 2, upper, bold dashed curve). A near optimal feature set could be determined, for example, by the peak performance along that curve. At each step, after selecting the feature to be removed, we employed 5-fold *external* cross-validation to give a more conservative estimate of performance of the feature subset. As each partition acted as the test set, the AUROC was noted and the results were averaged to obtain the conservative subset performance (Figure 2, lower, bold continuous curve).

RESULTS

The study included 125 eyes of 125 subjects (111 men, 14 women). Group N consisted of 52 HIV-negative individuals with the mean age ± standard deviation (SD) of 46.3 ± 7.8 years. Group H consisted of 38 high-CD4 HIV-positive patients with mean age of 48.5 ± 8.5 years. Included in group L were 35 low-CD4 HIV-positive patients with mean age of 46.1 ± 8.5 years.

FIELDS FROM HIV PATIENTS WITH LOW CD4 COUNTS

Table 1 shows the AUROCs with SDs for the various combinations of CD4 level, MLC, and feature set size. Table 2 demonstrates the *P* values for chosen comparisons.

Analysis With SVM

The purpose of this task was to evaluate whether MLCs could learn to discriminate fields from HIV patients with low CD4 counts, most of which have no visible field defects, from true normals. The AUROC was 0.804 ± 0.052 using SVM trained with the full-feature set (Figure 1). It significantly differed from chance decision with $P = .0002$. With an optimized 10-feature feature set, which was close to the peaking subset of 12, (upper, bold dashed curve in Figure 2), the AUROC increased to 0.843 ± 0.043, and this AUROC was greater, but not significantly so, than the result with the full-feature set ($P = .243$) and significantly better than chance ($P < .00005$).

TABLE 1. COMPARISON OF VISUAL FIELD FINDINGS IN HIGH-CD4 AND LOW-CD4 GROUPS DETECTED BY SUPPORT VECTOR MACHINE (SVM) AND RELEVANCE VECTOR MACHINE (RVM)*

VISUAL FIELD	HIGH CD4		LOW CD4	
	SVM	RVM	SVM	RVM
Top 10 field locations	0.708 ± 0.056	0.704 ± 0.056	0.843 ± 0.043	**0.870 ± 0.038**
Top 12 field locations			0.837 ± 0.045	0.824 ± 0.048
Top 22 field locations	0.695 ± 0.056	**0.726 ± 0.053**		
All 52 field locations	0.683 ± 0.057	0.670 ± 0.060	0.804 ± 0.052	0.800 ± 0.054
Chance	0.500 ± 0.062	0.500 ± 0.062	0.500 ± 0.064	0.500 ± 0.064

*Values are expressed as area under ROC curve $\pm$ SD. Best performance is shown in bold type.

TABLE 2. COMPARISON OF *P* VALUES IN HIGH-CD4 AND LOW-CD4 GROUPS USING SUPPORT VECTOR MACHINE (SVM) AND RELEVANCE VECTOR MACHINE (RVM)*

COMPARISON	HIGH CD4		LOW CD4	
	SVM	RVM	SVM	RVM
10 vs chance	**.012**	**.011**	**<.00005**	**<.00005**
12 vs chance			**<.00005**	**.0001**
22 vs chance	**.015**	**.007**		
All 52 vs chance	**.022**	**.038**	**.0002**	**.0002**
10 vs all 52	.213	.455	.243	.057
12 vs all 52			.321	.502
22 vs all 52	.472	.277		

*P values < .05 are in bold type.

Analysis With RVM

The performance with RVM was similar (Figures 1 and 2). AUROC with the full-feature set was 0.800 ± 0.054. RVM trained on the full-feature set performed better than chance with $P = .0001$. RVMs trained with the optimized 10-feature subset yielded 0.870 ± 0.038. The optimized feature set was not quite significantly better than the full-feature set ($P = .057$) but was significantly better against chance ($P < .00005$).

Feature Selection by Backward Elimination

The upper, bold dashed curve was the average of curves generated by the standard method of backward elimination, without the extra step of external validation. Peaking of this curve determined the optimal size feature set of 12. The lower, bold continuous curve averaged the curves generated with the extra step of external validation to verify performance as each feature was removed. There was minimal improvement in performance with feature sets larger than 4.

FIELDS FROM HIV PATIENTS WITH HIGH CD4 COUNTS

Analysis With SVM

The AUROC was 0.683 ± 0.057 with SVM trained on the full dataset (Figure 1). It was significantly better than chance ($P = .022$). Backward elimination produced subsets that performed optimally with the best 22 features; also tested were the best 10 features. AUROC with 22 features was 0.695 ± 0.056 and with 10 features was 0.708 ± 0.056, which was not significantly better better ($P = .213$) than the AUROC derived from the full-feature set.
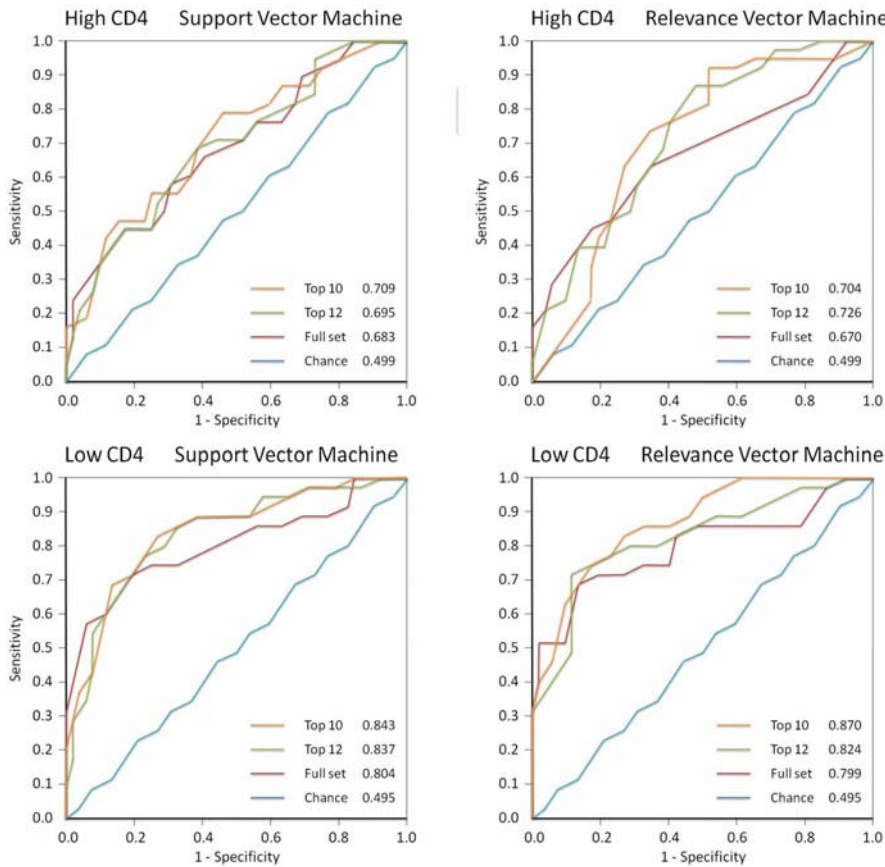
FIGURE 1

Receiver operating characteristic curves for support vector machine, relevance vector machine, high CD4, and low CD4. Within each graph are curves generated for machine learning classifiers trained to distinguish HIV eyes from normal eyes using all 52 field locations, the subset with peak performance, and the 10-location feature set; the chance curve is the attempt to learn classes with equivalent data.

Analysis With RVM

The performance with RVM was similar (Figure 1). The AUROC from RVM trained with the full-feature set was 0.670 ± 0.060. The AUROC with the 22-feature subset was 0.726 ± 0.053 and with the 10-feature subset was 0.0704 ± 0.056. The 22- and 10-feature subsets were not significantly better than the full-feature set ($P = .277$ and $P = .455$, respectively). These optimized 22- and 10-feature subsets were significantly better than chance, with $P = .007$ and $.011$, respectively.
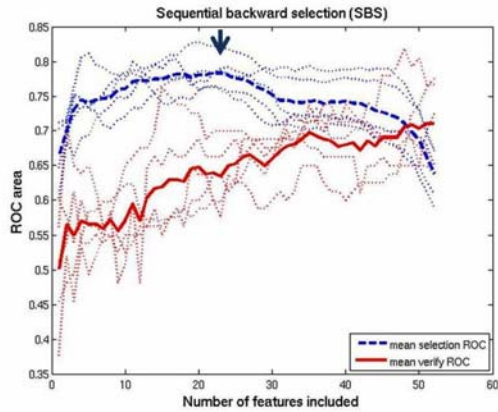
Feature Selection by Backward Elimination

The upper dashed curve peaked at 22 features. The lower continuous, more conservative curve revealed poor performance with small feature sets, with performance gradually improving to maximum with the full-feature set.

THE LOCATION OF RANKED FEATURES

For both the high-CD4 and the low-CD4 eyes, the top 10 locations were plotted for analysis (Figure 3). Three of the top 4 field locations for distinguishing low CD4 eyes from normals were superior nasal and superior temporal to the blind spot. Seven of the top 10 field locations were mostly distributed in the hemifield temporal to fixation. In the high-CD4 group, the distribution of the top 10 locations differed from that with the low-CD4 group.

High CD4 Peaking
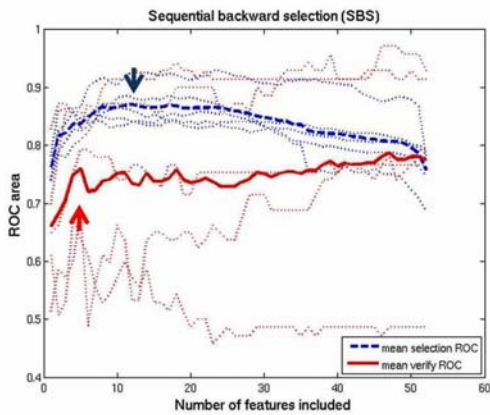

Low CD4 Peaking

FIGURE 2

Performance curves measuring area under the receiver operating characteristic (AUROC) for each size subset of near optimal combinations of features generated by backward elimination from all 52 features down to 1 feature. The upper, dashed blue curve averages curves derived from standard backward elimination, which graphs the AUROC of the selected set for each set size. The peak (blue arrow) is the subset size with the best performance. The lower, continuous red curve averages the results after the extra step of external cross-validation to give a more conservative estimate of performance at each step of backward elimination. In the low-CD4 eyes, performance did not improve with extra-validated feature sets larger than the set with 4 dependent features. In high-CD4 eyes, small subsets performed poorly and performance increased up to the full 52-location feature set.

Distribution of top 10 field locations

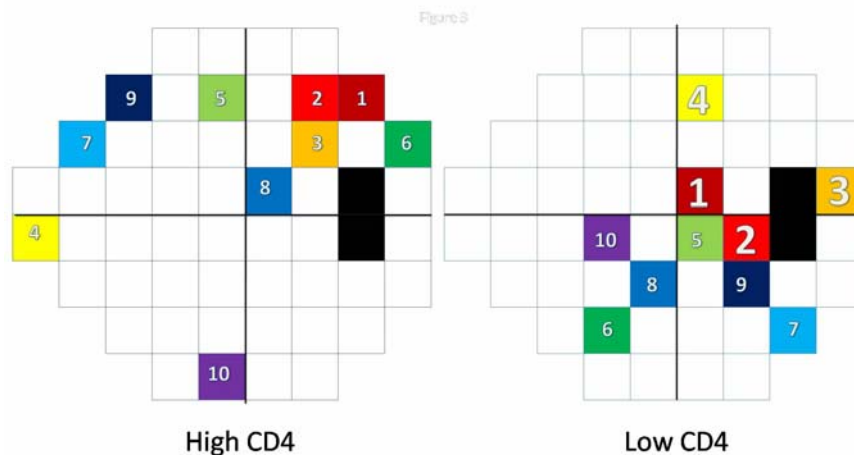

High CD4   Low CD4

FIGURE 3

Plot of the 10 best features in combination, from backward elimination. In the high-CD4 eyes, small sets did not perform well, indicating uncertain distribution of the top features (small numbers). In the low-CD4 eyes, feature sets larger than 4 did not improve performance, indicating the top 4 field locations (large numbers) were strong predictors of HIV.

DISCUSSION

This study establishes that sensitive MLCs, such as SVM and RVM, are powerful classification methods that approach the performance of the theoretical optimal classifier for classifying visual fields, especially when the feature sets are optimized.[1,3,9,10] Optimized MLCs appear to be a valid approach to detecting otherwise hidden abnormalities in medical tests that generate complex multidimensional output.

This study confirmed the finding in other publications that eyes from HIV patients with low CD4 counts have retinopathy damage that significantly affects the visual field.[15,16,22] The top 10 field locations were mostly temporal to fixation, with 3 of the top 4 locations superiorly near the blind spot (Figure 3). These locations correlated with the inferior retinal nerve fiber thinning found by high-resolution optical coherence tomography (OCT) in low-CD4 eyes.[23] The lack of performance improvement beyond 4 features in the lower, continuous, more conservative backward elimination curve (Figure 2) indicated that the top 4 features truly were the most important. These features showed damage superior to the blind spot, which implied damage most commonly to the inferior retinal nerve fibers.

In addition, this study demonstrated that patients with high CD4 counts also had retinal damage that affected the visual fields. Whereas the fields from high-CD4 eyes mostly appeared normal to human experts, the trained MLCs could distinguished between eyes from high-CD4 patients and normal eyes, suggesting that there were field defects indicative of retinal damage that were not apparent to human observers. Although it appears that the top-ranked field locations are superior near the blind spot (Figure 3), the poor performance with small-feature subsets and continued improvement up to the full-feature set on the lower, more conservative backward elimination curve indicated that the ranking of the features was not reliable and thus that the locations of the important features were uncertain. In eyes from HIV individuals with high CD4 counts, we have not yet found a reliable pattern of vision loss.

These findings correlate with other methods of retinal evaluation. There is evidence of oxidative and inflammatory endothelial cell damage in AIDS vasculopathy.[32] The HIV retinopathy is a frequent clinical finding in these patients during the course of their immunosuppression. It is a microvasculopathy that can present with multiple peripapillary hemorrhages or cotton-wool spots, the remnants of which can be detected using imaging technology long after their disappearance.[14,24,25] Although the retinopathy is generally asymptomatic, even patients with high CD4 counts sometimes observe visual changes. HIV retinopathy preferentially affects the posterior pole and major vascular arcades.[33] We assume that these retinal microinfarctions may be responsible for subsequent field defects, which constitute biologic plausibility for this observation.

The result of such changes seems to be thinning of the retinal nerve fiber layer (RNFL) in HIV-positive patients. Our original descriptions of RNFL thinning using OCT[25] and later scanning laser polarimetry[34] with variable corneal compensation (GDx-VCC)[35] were confirmed by Besada and associates.[36] Using GDx-VCC and Heidelberg retinal tomography II, they found reduced RNFL thickness in HIV-positive patients receiving HAART, even in patients with good immune status. Their study also agrees with the topographic location of RNFL damage, although it is not clear whether the damage is indeed most prominent in the superior and inferior retina or whether it is diffuse but most easily detected there. HAART, therefore, may not fully prevent these patients from ongoing subtle vascular damage, which may persist in spite of successful immune recovery.

In summary, this study (1) confirmed that eyes from low-CD4 HIV patients have retinal damage and discovered that eyes from high-CD4 HIV patients also have retinal damage; (2) demonstrated that MLCs are sensitive enough to uncover the field defects produced by the retinal damage; and (3) revealed that a likely location of retinal damage with low CD4 counts is inferior and near to the disc, producing field defects superior and near to the blind spot.

ACKNOWLEDGMENTS

REFERENCES

1. Goldbaum MH, Sample PA, White H, et al. Interpretation of automated perimetry for glaucoma by neural network. *Invest Ophthalmol Vis Sci* 1994;35:3362-3373.
2. Mutlukan E, Keating K. Visual field interpretation with a personal computer based neural network. *Eye* 1994;8:321-323.
3. Brigatti L, Nouri-Mahdavi K, Weitzman M, Caprioli J. Automatic detection of glaucomatous visual field progression with neural networks. *Arch Ophthalmol* 1997;115:725-728.
4. Sample PA, Boden C, Zhang Z, et al. Unsupervised machine learning with independent component analysis to identify areas of progression in glaucomatous visual fields. *Invest Ophthalmol Vis Sci* 2005;46:3684-3692.
5. Brigatti L, Hoffman BA, Caprioli J. Neural networks to identify glaucoma with structural and functional measurements. *Am J Ophthalmol* 1996;121:511-521.
6. Bowd C, Chan K, Zangwill LM, et al. Comparing neural network and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Invest Ophthalmol Vis Sci* 2002;49:3444-3454.

7. Zangwill LM, Chan K, Bowd C, et al: Heidelberg retina tomograph measurements of the optic disc and parapapillary retina for detecting glaucoma analyzed by machine learning classifiers. *Invest Ophthalmol Vis Sci* 2004;45:3144-3151.
8. Bowd C, Medeiros FA, Zhang et al. Relevance vector machine and support vector machine classifier analysis of scanning laser polarimetry retinal nerve fiber layer measurements. *Invest Ophthalmol Vis Sci* 2005;46:1322-1329.
9. Henson DB, Spenceley SE, Bull DR. Artificial neural network analysis of noisy visual field data in glaucoma. *Artif Intell Med* 1997;10:99-113.
10. Goldbaum MH, Sample PA, Chan K, et al. Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. *Invest Ophthalmol Vis Sci* 2002;43:162-169.
11. Deayton JR, Wilson P, Sabin CA, et al. Changes in the natural history of cytomegalovirus retinitis following the introduction of highly active antiretroviral therapy. *AIDS* 2000;14:1163-1170.
12. Jabs DA, Van Natta ML, Kempen JH, et al. Characteristics of patients with cytomegalovirus retinitis in the era of highly active antiretroviral therapy. *Am J Ophthalmol* 2002;133:48-61.
13. Goldberg DE, Smithen LM, Angelilli A, Freeman WR. HIV-associated retinopathy in the HAART era. *Retina* 2005;25:633-649.
14. Kupperman BD, Petty JG, Richman DD, et al. The correlation between CD4+ counts and the prevalence of cytomegalovirus retinitis and human immunodeficiency virus-related noninfectious retinal vasculopathy in patients with acquired immunodeficiency syndrome. *Am J Ophthalmol* 1993;115:575-582.
15. Plummer DJ, Sample PA, Arevalo JF, et al. Visual field loss in HIV-positive patients without infectious retinopathy. *Am J Ophthalmol* 1996;122:542-549.
16. Geier SA, Nohmeier C, Lachenmayr BJ, Klauss V, Goebel FD. Deficits in perimetric performance in patients with symptomatic human immunodeficiency virus infection or acquired immunodeficiency syndrome. *Am J Ophthalmol* 1995;119:335-344.
17. Quiceno JI, Capparelli E, Sadun AA, et al. Visual dysfunction without retinitis in patients with acquired immunodeficiency syndrome. *Am J Ophthalmol* 1992;113:8-13.
18. Geier SA, Kronawitter U, Bogner JR, et al. Impairment of color contrast sensitivity and neuroretinal dysfunction in patients with symptomatic HIV infection or AIDS. *Br J Ophthalmol* 1993;77:716-720.
19. Geier SA, Hammel G, Bogner JR, Kronawitter U, Berninger T, Goebel FD. HIV-related ocular microangiopathic syndrome and color contrast sensitivity. *Invest Ophthalmol Vis Sci* 1994;35:3011-3021.
20. Shah KH, Holland GN, Yu F, et al. Contrast sensitivity and color vision in HIV-infected individuals without infectious retinopathy. *Am J Ophthalmol* 2006;142:284-292.
21. Iragui VJ, Kalmijn J, Plummer DJ, et al. Pattern electroretinograms and visual evoked potentials in HIV infection: evidence of asymptomatic retinal and postretinal impairment in the absence of infectious retinopathy. *Neurology* 1996;47:1452-1456.
22. Sample PA, Plummer DJ, Mueller AJ, et al. Pattern of early visual field loss in HIV-infected patients. *Arch Ophthalmol* 1999;117:755-760.
23. Kozak I, Bartsch DU, Cheng L, Freeman WR. In vivo histology of cotton-wool spots using high-resolution optical coherence tomography. *Am J Ophthalmol* 2006;141:748-750.
24. Kozak I, Bartsch DU, Cheng L, Freeman WR. Hyperreflective sign in resolved cotton wool spots using high resolution optical coherence tomography and OCT ophthalmoscope. *Ophthalmology* 2007;114:537-543.
25. Kozak I, Bartsch DU, Cheng L, Kosobucki BR, Freeman WR. Objective analysis of retinal damage in HIV-positive patients in the HAART era using OCT. *Am J Ophthalmol* 2005;139:295-301.
26. Vapnik V. *Statistical Learning Theory*. New York: Wiley; 1998.
27. Vapnik V. *The Nature of Statistical Learning Theory*. 2nd ed. New York: Springer; 2000.
28. Bowd C, Chan K, Zangwill LM, et al. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Invest Ophthalmol Vis Sci* 2002;43:3444-3454.
29. Bishop CM, Tipping ME. Variational relevance vector machines. In: Boutilier C, Goldszmidt , eds. *Uncertainty in Artificial Intelligence.* Cambridge, UK: Microsoft Research; 2000:45-63.
30. Tipping ME. Sparse Bayesian learning and the relevance vector machine. *J Mach Learn Res.* 2001;1:211-244.
31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845.
32. Baliga RS, Chaves AA, Jing L, Ayers LW, Bauer JA. AIDS-related vasculopathy: evidence for oxidative and inflammatory pathways in murine and human AIDS. *Am J Physiol Heart Circ Physiol* 2005;289:1373-1380.
33. Newsome DA, Green WR, Miller ED, et al. Microvascular aspects of acquired immunodeficiency syndrome retinopathy. *Am J Ophthalmol* 1984;98:590-601.
34. Kozak I, Bartsch DU, Cheng L, et al. Scanning laser polarimetry demonstrates retinal nerve fiber layer damage in human immunodeficiency virus positive patients without infectious retinitis. *Retina* 2007. Forthcoming.
35. Weinreb RN, Bowd C, Greenfield DS, Zangwill LM. Measurement of the magnitude and axis of corneal polarization with scanning laser polarimetry. *Arch Ophthalmol* 2002;120:901-906.
36. Besada E, Shechtman D, Black G, Hardigan PC. Laser scanning confocal ophthalmoscopy and polarimetry of human immunodeficiency virus patients without retinopathy, under antiretroviral therapy. *Optom Vis Sci* 2007;84:189-196.

PEER DISCUSSION

DR DOUGLAS R. ANDERSON: I am pleased to offer a commentary on the presentation by Dr Goldbaum on behalf of his collaborative group.

Most of us are not initiated into the mathematical details of the various types of machine classifiers (in the past sometimes called "neural networks") presently in use, so they provide the analogy about how we assimilate various characteristics of a face to recognize each to help us understand their general nature. I might add my own analogy. Many of us find it difficult to distinguish individuals of a different racial background from our own, mainly because the features of similarity between them overwhelm our ability to distinguish the features that differ among them. However, we are not overwhelmed by the similarities within the group(s) among which we were raised, and find it easier to distinguish facial patterns of individuals. By analogy, machine classifiers are able to sort through the similarities and recognize identifying differences within a large amount of information. There are other instances in the literature to show the remarkable ability of machine classifiers to see through a barrage of information and discern the relevant information to recognize a diagnostic pattern. This paper exemplifies the capabilities of machine classifiers and the strategy of removing the least helpful data for sake of efficiency, but in some cases also to improve the ability to discern.

The focus today at this meeting, however, is not the methodology, but the disease process studied. From the data it can be inferred that there are "subclinical" manifestations of retinal damage related to HIV. Based on other studies that show loss in the retinal nerve fiber layer in individuals without active retinopathy, and the fact that noninfectious HIV retinopathy, when active, includes cotton-wool spots (transient manifestations of microinfarctions of a group of axons), it is postulated that scattered locations in the visual field could have diminished sensitivity of a degree and pattern that they can be recognized by machine classifiers. These defects might, for example, be the residual of previous unobserved episodes of active retinitis or be the result of ongoing mild retinitis that is not recognized by ophthalmoscopy. This retinopathy seems to be related to vascular compromise of inner retinal layers resulting from the compromised immune state, and not to cytomegalovirus (CMV) retinitis. In this particular study, individuals were excluded if they had a history of infectious retinitis in order to permit focus on this noninfectious retinopathy.

The concept here is that the visual field may have features that permit the detection and recognition of retinal damage and that these field characteristics are so subtle that they not only escape our attention, but may be difficult to recognize and distinguish even if we tried. The authors show that such findings are particularly dominant in a group with seriously low CD4 counts, occur in certain locations of the field most frequently, and, of special note, seem to occur also in individuals with better CD4 counts. The area under the receiver operating characteristic (ROC) curve is smaller in those with high CD4, which implies either that the retinopathy is milder (and identifying features therefore more subtle) or that only some of the individuals in the cohort with high CD4 have any retinopathy at all. The latter explanation would apply if all the individuals in the study cohort had a high CD4 all right, but not all had retinopathy. I wonder whether any characteristics of the data analysis can tell us which explanation (mild disease or low prevalence) is true of this cohort—a question for the authors, possibly unanswerable. In any event, it would seem that the area under the ROC curve and the flatness of the ROC curve in Figure 1 would suggest that among those with high CD4, it may not be possible to diagnose visual field abnormalities in an individual with a desired degree of certainty. However, in the presence of low CD4, with 80% specificity, it may be possible to be nearly 80% sensitive in detecting visual field abnormality thought to signify HIV retinopathy.

In summary, it is thus apparent that some (or possibly all) cases have some degree of residual from past retinal damage that leaves its mark and remains discernible by testing of the visual field, or may have subclinical low-grade chronic retinopathy, more obviously in those with low CD4. Machine classifiers seem able to recognize differences in the visual field of HIV patients from the normal control subjects. When sensitivity and specificity of particular features are derived from test cohorts, it is usual to verify the sensitivity and specificity in a second group of patients and controls subjects; in the present study the technique was to use 90% of the subjects to teach the machine classifier and the remaining 10% to test it, rotating 10 subgroups in a series of cycles and confirming the results are the same in any combination.

I would like to close by posing a second question to the authors: To what degree does visual field testing and analysis by a machine classifier help make or refine any therapeutic or management decisions, or are the decisions about present activity of the retinopathy or the need for changes in management determined by ophthalmoscopy and blood tests?

ACKNOWLEDGMENTS

DR. MICHAEL H.GOLDBAUM: I would like to thank Dr. Anderson for his very helpful comments. One of the issues raised is whether there are many eyes with a little bit of damage or a few eyes with moderate damage. No statistically significant study yet indicates which of these may be true. I suspect that it may be the latter, because human perimetry experts who reviewed the data tended to identify the normal eyes correctly with 95% or 98% specificity. Based on the curves we generated at 95% specificity with high CD4 counts, we determined 20% sensitivity with the machine learning classifiers, and the perimetry experts determined a12% sensitivity. With the low CD4 counts we calculated about 50% sensitivity with the machine learning classifiers, and the human experts had 35% sensitivity. This suggests that the machine is finding something comparable to what the human experts are identifying. There are few patients with discernable disease and many of the eyes without discernable disease may actually be normal, but we do not know for sure. Secondly, we would like to have a larger samples set to evaluate. I believe that in the future we will have more accurate information based on a larger set of examples and more comprehensive types of defects that can occur. This approach was

used not for diagnosis or therapeutic implications, although we wish to explore that in the future. The present study was rather an exercise in data mining to help us understand the disease process and to determine how it affects the retina.