Plus Disease in Retinopathy of Prematurity: Development of Composite Images by Quantification of Expert Opinion

Michael F. Chiang,[1,2] *Rony Gelman,*[1] *Steven L. Williams,*[1] *Joo-Yeon Lee,*[1,3] *Daniel S. Casper,*[1] *M. Elena Martinez-Perez,*[4] *and John T. Flynn*[1]

Purpose. To demonstrate a methodology for generating composite wide-angle images of plus disease in retinopathy of prematurity (ROP), using quantitative analysis of expert opinions.

Methods. Thirty-four wide-angle retinal images were independently interpreted by 22 ROP experts as "plus" or "not plus." All images were processed by the computer-based Retinal Image multiScale Analysis (RISA) system to calculate two parameters: arterial integrated curvature (AIC) and venous diameter (VD). Using a reference standard defined by expert consensus, sensitivity and specificity curves were calculated by varying the diagnostic cutoffs for AIC and VD. From these curves, individual vessels from multiple images were identified with particular diagnostic cutoffs, and were combined into composite wide-angle images using graphics-editing software.

Results. The values associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff) were AIC 0.061 and VD 4.272, the values associated with 50% underdiagnosis of true plus disease (i.e., a 50% sensitivity cutoff) were AIC 0.049 and VD 4.088, and the values associated with 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff) were AIC 0.042 and VD 3.795. Composite wide-angle images were generated by identifying and combining individual vessels with these characteristics.

Conclusions. Computer-based image analysis permits quantification of retinal vascular features, and a spectrum of abnormalities is seen in ROP. Selection of appropriate vessels from multiple images can produce composite plus disease images corresponding to expert opinions. This method may be useful for educational purposes, and for development of future disease definitions based on objective, quantitative principles. (*Invest Ophthalmol Vis Sci.* 2008;49:4064–4070) DOI: 10.1167/iovs.07-1524

From the Departments of [1]Ophthalmology and [2]Biomedical Informatics, Columbia University College of Physicians and Surgeons, New York, New York; the [3]Department of Ophthalmology, Hallym University College of Medicine, Seoul, South Korea; and the [4]Department of Computer Science, National Autonomous University of Mexico (IIMAS-UNAM), Mexico City, Mexico.

Supported by a Career Development Award from Research to Prevent Blindness, New York, NY (MFC), and by Grant EY13972 from the National Institutes of Health, Bethesda, MD (MFC).

Submitted for publication November 27, 2007; revised January 22, 2008; accepted July 21, 2008.

Disclosure: **M.F. Chiang**, Clarity Medical Systems (C); **R. Gelman**, None; **S.L. Williams**, None; **J.-Y. Lee**, None; **D.S. Casper**, None; **M.E. Martinez-Perez**, None; **J.T. Flynn**, None

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Michael F. Chiang, Columbia University College of Physicians and Surgeons, 635 West 165th Street, Box 92, New York, NY 10032; chiang@dbmi.columbia.edu.

Retinopathy of prematurity (ROP) is a vasoproliferative disease affecting low-birth-weight infants and is a leading cause of childhood blindness throughout the world.[1–3] The International Classification for ROP (ICROP) provides a universal system for describing the ophthalmoscopic appearance of retinal vessels.[4,5] Plus disease, which is characterized by arterial tortuosity and venous dilation, is a major component of this system. The minimum amount of vascular abnormality required for the presence of plus disease is defined by a standard photograph, which is selected by an expert committee.[6] More recently, major clinical trials have explicitly required ≥2 quadrants of this amount of vascular change for the diagnosis of plus disease.[7,8]

Accurate clinical assessment of plus disease is critical. The multicenter Cryotherapy for ROP (CRYO-ROP) and Early Treatment for ROP (ETROP) trials established that plus disease is a necessary feature of threshold disease and a sufficient feature for diagnosis of type-1 ROP, both of which have been shown to benefit from cryotherapy or laser photocoagulation of the peripheral avascular retina.[6,8,9] Because the definition of plus disease is based on a standard image with descriptive qualifiers, diagnosis may be heavily subjective. The appearance of vessels in the published standard photograph is not uniform, and there are no guidelines regarding which particular vessels represent the minimum requisite dilatation and tortuosity.[6] It is not clear which vascular features are actually used by examiners in making diagnoses. The published standard photograph has a much narrower field of view than is afforded by indirect ophthalmoscopy or wide-angle retinal imaging devices, and these differences in magnification and perspective may result in differing interpretations among examiners. We have shown that the accuracy and agreement of plus disease diagnosis from wide-angle posterior pole images are imperfect, even among recognized experts.[10–12] These factors raise significant concerns, because errors in classification may result in overtreatment or undertreatment of disease. This problem has important implications for ROP classification, treatment, and research.

Objective, quantitative measurement of retinal vascular changes has the potential to improve the accuracy and reproducibility of plus disease diagnosis.[13,14] Several studies have demonstrated the feasibility of this approach through application of automated image-analysis systems for detection of plus disease.[11,12,15–19] In particular, the computer-based Retinal Image multiScale Analysis (RISA) system defines quantitative parameters reflecting the curvature and diameter of vessels. We recently showed that this system can diagnose plus disease with accuracy comparable to that of human experts.[11,12,15] This research has created a large database relating the values of measurable vascular parameters to the overall diagnostic impressions of numerous recognized ROP experts. In principle, these data could be used for systematic examination of vascular features that are considered most important by experts for identifying the presence of plus disease in wide-angle retinal images.

Investigative Ophthalmology & Visual Science, September 2008, Vol. 49, No. 9

The present study demonstrates a technique for generating composite images of plus disease, based on original experimental data that we have gathered regarding expert clinical impressions.[11,12] Arteries and veins with specified levels of curvature and diameter are digitally combined into wide-angle composite images. These composite images may provide guidance in calibrating plus disease diagnosis for clinicians. In addition, this methodology could eventually be applied to other diseases in which findings are represented by quantitative parameters.

METHODS

This study was approved by the Institutional Review Board at Columbia University Medical Center, included waiver of consent for use of deidentified retinal images, and followed tenets of the Declaration of Helsinki. Informed consent was obtained from expert participants using a click-through web form.

Interpretation of Original Images by Experts

Methods for interpretation of the original images have been described.[10–12] Briefly, a set of 34 images was compiled from premature infants during routine ROP care, using a wide-angle camera (RetCam-II; Clarity Medical Systems, Pleasanton, CA). Images were selected that, in the opinion of the authors, reflected some vascular change from baseline. Each photograph displayed the posterior retina, with any visible peripheral disease cropped out.

A group of 22 ROP experts was invited to participate. Eligible experts were defined as practicing pediatric ophthalmologists or retina specialists who met ≥1 of three criteria: having been a study center principal investigator for the CRYO-ROP or ETROP studies, having been a certified investigator for either study, or having coauthored ≥5 peer-reviewed ROP manuscripts. Each participant provided one of four mutually exclusive diagnoses ("plus," "pre-plus," "neither," or "cannot determine") for each original image, using Web-based software developed by the authors.[10] For data analysis, diagnoses of "pre-plus" or "neither" were considered to be "not plus" and diagnoses of "cannot determine" were excluded. Participants were not provided with references or standards for plus disease, although it was assumed that they would be intimately familiar with these definitions. For each image, a reference standard diagnosis was defined based on the response ("plus" or "not plus") that was given by the majority of experts. In case of a tie, the more severe diagnosis was selected as the reference standard.

Computer-Based Analysis of Original Images

Retinal vessels in each of the 34 original images were identified, classified by author consensus as arteries or veins, and analyzed by the RISA computer-based system using previously described methods.[11,12,15,16] *Arterial integrated curvature* (AIC), which is defined as the sum of angles along the vessel divided by vessel length (radians/pixel), was calculated for every artery (Fig. 1). *Venous diameter* (VD), which is defined as the total area of the vessel cross-section divided by its length (pixels), was calculated for every vein. In vessels with multiple branches, the RISA system required that only the major branch be identified and selected for analysis.

The mean value of each system parameter was calculated for each image. Sensitivity and specificity for each parameter at different cutoffs were calculated by using the previously described reference standard. Sensitivity and specificity curves were plotted as a function of cutoffs for each system parameter. All data analysis was performed with statistical and computational software (Excel 2003; Microsoft, Redmond, WA; SPSS ver. 14, SPSS Inc., Chicago, IL; R programming language ver. 2.4.0, Free Software Foundation, Boston, MA).

Generation of Composite Images

AIC and VD parameters were used for generation of images, because they have been shown to model the clinical diagnosis of plus disease accurately,[11,12,15] and because the narrative definition of plus disease refers to "venous dilation" and "arterial tortuosity."[4,5] As shown in Figure 2, sensitivity and specificity curves were used to identify three diagnostic cutoffs for AIC and VD: (1) values associated with 25% underdiagnosis of true plus disease according to the reference standard (i.e., 75% sensitivity cutoff), (2) values associated with 50% underdiagnosis of true plus disease (i.e., 50% sensitivity cutoff), and (3) values associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff). Linear interpolation was used to calculate exact cutoff values for each level of sensitivity.

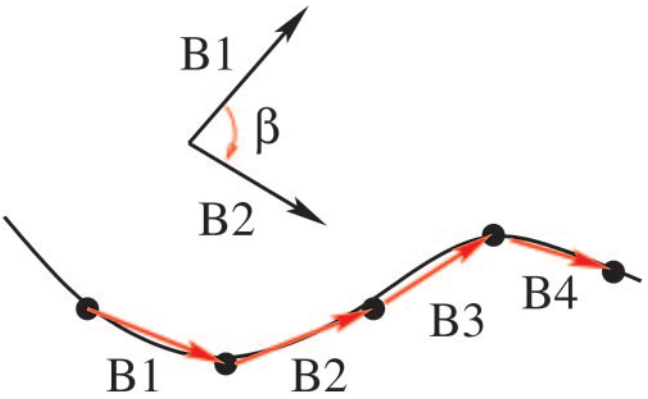

FIGURE 1. Computer-based RISA system. The IC of vessels is computed by forming vectors along the vessel, calculating the cumulative sum of angles (β) between vectors and normalizing by vessel length.

From among the 34 original study images, individual vessels were identified that matched these cutoff values. One artery and one vein were selected from each quadrant (superotemporal, inferotemporal, superonasal, inferonasal). These vessels were combined using graphics editing software (Photoshop CS2; Adobe Systems, San Jose, CA) to produce composite images at 75%, 50%, and 25% sensitivity cutoff values for AIC and VD. If no suitable vessel was available with AIC or VD within 5% of the cutoff value, then a vessel from the same side of the retina was substituted (e.g., a superotemporal vessel was substituted for the missing inferotemporal vessel). Vessels were combined on a uniform background consisting of the optic disc and choroidal system. Arteries and veins were designated as bright red and dark red, respectively. Vascular appearance was smoothened using graphics-editing software. Although only major branches of each vessel were used for RISA, other branches were included in the final composite images to produce a more realistic appearance.

Evaluation of Composite Images

To evaluate whether these computer-generated composite images truly represented a spectrum of vascular abnormalities that would be judged as plus disease by experts, seven ROP experts (3 pediatric ophthalmologists and four retinal specialists) independently examined the three composite images. These evaluating ophthalmologists were different from the 22 experts who reviewed the original retinal photographs, and thus bias was avoided. Evaluators were asked (1) whether each image in their opinion represented plus disease, pre-plus disease, or neither; and (2) to rank the three images based on increasing severity of vascular abnormality. All the evaluating ophthalmologists performed ROP examinations at major medical centers, and six of seven had served as certified ETROP investigators.

RESULTS

Interpretation of Original Images by Experts and a Computer-Based System

From the 34 original study images, 118 arteries and 138 veins were analyzed. On average, 3.47 arteries and 4.03 veins were

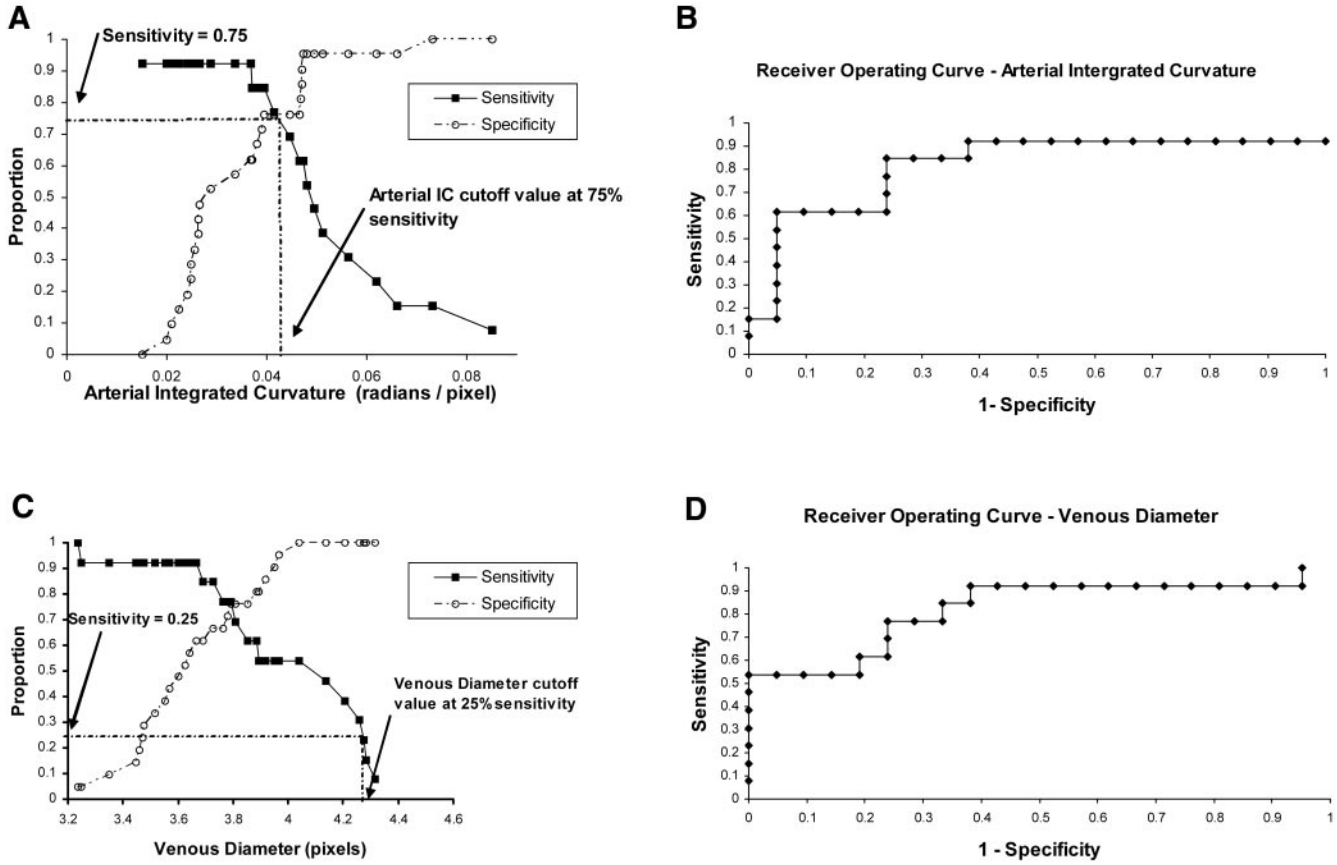

FIGURE 2. Sensitivity, specificity, and receiver operating characteristic (ROC) curves of individual computer system parameters for plus disease diagnosis, compared to reference standard diagnosis of the majority vote among 22 ROP experts. Curves are displayed as a function of parameter cutoff criteria for detection of plus disease: (**A**) sensitivity and specificity and (**B**) ROC for AIC, and (**C**) sensitivity and specificity and (**D**) ROC for VD. (**A**, **C**, *dotted lines*) AIC cutoff associated with 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff) = 0.0423, and the VD cutoff value associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff) = 4.2717.

analyzed per image. Table 1 summarizes the values of individual system parameters from computer-based analysis of the 34 retinal study images. The median (range) AIC was 0.037 (0.010–0.107) radians/pixel, and the median (range) VD was 3.737 (2.014–4.967) pixels. Figure 2 demonstrates examples of how exact sensitivity cutoffs for AIC and VD were determined. For AIC, the cutoff value associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff) was 0.061 radians/pixel, and the cutoff value associated with 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff) was 0.042 radians/pixel. For VD, the cutoff associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff) was 4.272 pixels, and the cutoff associated with 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff) was 3.795 pixels.

Composite Plus Disease Images

Figure 3 displays composite retinal images constructed with this methodology. Each image is composed of two smaller black-and-white images showing the major branches of arteries or veins used for RISA analysis, and one larger color image, which is a composite of all vessels, along with their minor branches. The image generated from vessels associated with 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff) is shown in Figure 3A. Among vessels in this image,

TABLE 1. Individual System Parameters from Computer-Based Analysis of 34 Study Images

System Parameter	Median (Range)	Sensitivity Cutoff Values			Specificity Cutoff Values		
		75%	50%	25%	75%	50%	25%
Arterial Integrated Curvature*	0.037 (0.010–0.107)	0.042	0.049	0.061	0.039	0.028	0.025
Venous Diameter†	3.737 (2.014–4.967)	3.795	4.088	4.272	3.788	3.615	3.473

Sensitivity and specificity curves were constructed by varying the cutoff value of each individual parameter, compared with the reference standard of the majority consensus among 22 recognized ROP experts. Median (range) values for each parameter are shown, along with the cutoff values corresponding to sensitivity and specificity of 25%, 50%, and 75%. Units of arterial integrated curvature are in radians/pixel, and units of venous diameter are in pixels.

* Integrated curvature is in units of radians/pixel.

† Diameter is in units of pixels.

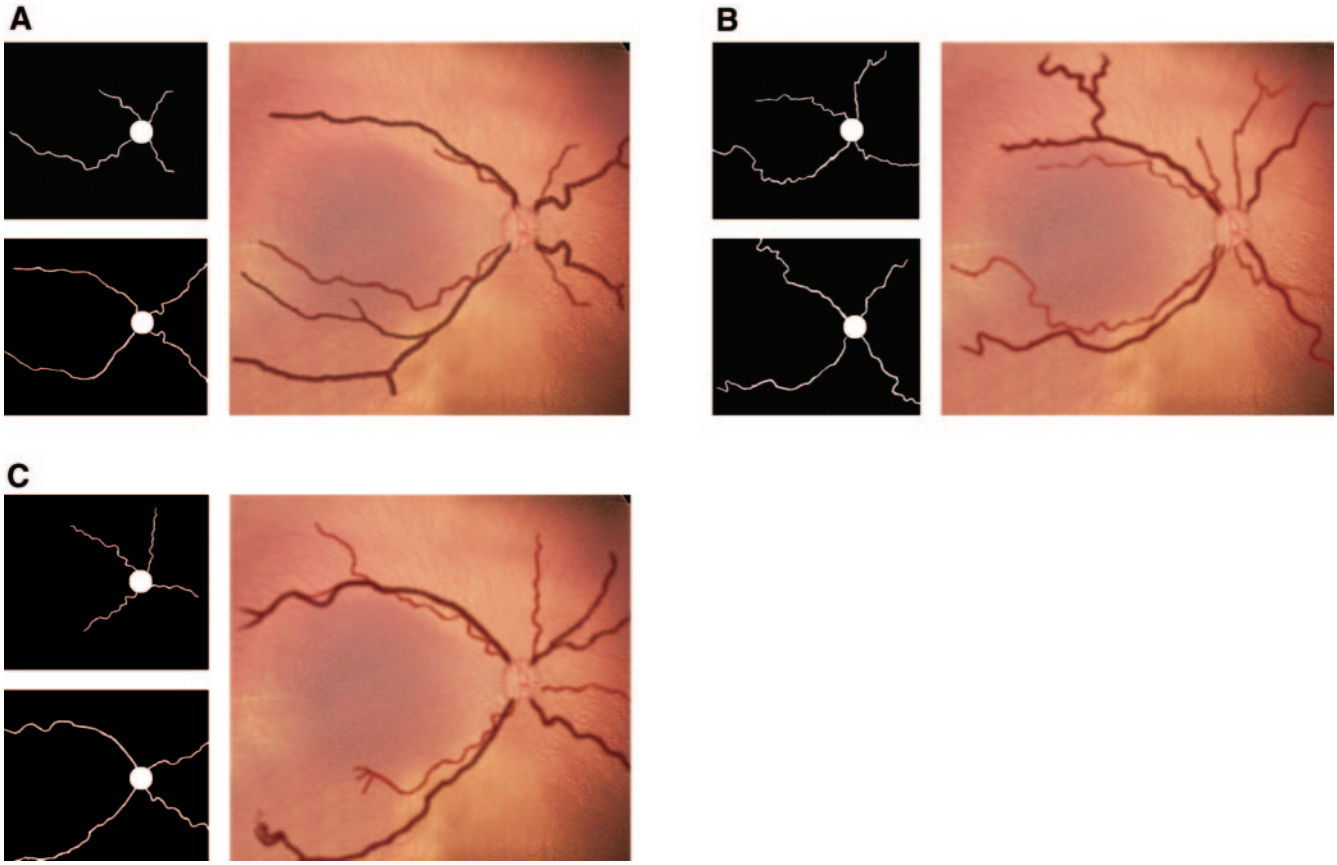

FIGURE 3. Composite images generated demonstrating AIC and VD values reflecting (**A**) 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff), (**B**) 50% underdiagnosis of true plus disease (i.e., 50% sensitivity cutoff), and (**C**) 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff). Black-and-white images display major branches of the arteries (*top*) and veins (*bottom*) used for analysis by the computer-based system. Color images display complete vessels superimposed on a choroidal background, with arteries shown in *light red* and veins in *dark red*. Vascular appearance was smoothed with graphics-editing software.

mean AIC was 0.042 radians/pixel (sensitivity 74.9%) and mean VD was 3.792 radians/pixel (sensitivity 72.3%). The image formed by vessels associated with 50% underdiagnosis of true plus disease (i.e., 50% sensitivity cutoff) is shown in Figure 3B. Among the vessels in this image, mean AIC was 0.048 (sensitivity 51.5%) and mean VD was 4.077 (sensitivity 50.9%). The image formed by vessels associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff) is shown in Figure 3C. Among the vessels in this image, mean AIC was 0.060 (sensitivity 25.7%) and mean VD was 4.275 (sensitivity 23.5%).

At the 75% sensitivity cutoff, there was no superonasal artery with the appropriate integrated curvature and no inferonasal vein with the appropriate diameter in our vessel database. Therefore, the vessels in Figure 3A were obtained by reflection of the inferonasal artery and superonasal vein, respectively.

Evaluation of Composite Images

Responses provided by the seven expert evaluators are displayed in Table 2. The image formed by vessels associated with 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff, Fig. 3A) was diagnosed as pre-plus by six of the seven (85.7%) evaluators, and was judged least severe among the

TABLE 2. Evaluation of Composite Images

	Vessels at 75% Sensitivity Cutoff (Fig. 3A)	Vessels at 50% Sensitivity Cutoff (Fig. 3B)	Vessels at 25% Sensitivity Cutoff (Fig. 3C)
Expert responses			
Plus	1/7 (14.3)	5/7 (71.4)	7/7 (100)
Pre-plus	6/7 (85.7)	2/7 (28.6)	0/7 (0)
Neither	0/7 (0)	0/7 (0)	0/7 (0)
Ranking			
Least severe	7/7 (100)	0/7 (0)	0/7 (0)
Intermediate	0/7 (0)	2/7 (28.6)	5/7 (71.4)
Most severe	0/7 (0)	5/7 (71.4)	2/7 (28.6)

Seven experts (three pediatric ophthalmologists and four retinal specialists) were shown composite images and were asked whether images in their opinion represented plus disease, pre-plus disease, or neither, and to rank the three images based on increasing severity of vascular abnormality. Results are shown as number (percentage) of experts.

composite images by seven (100%). The image formed by vessels associated with 50% underdiagnosis of true plus disease (i.e., 50% sensitivity cutoff, Fig. 3B) was diagnosed as plus by five (71.4%) evaluators and pre-plus by two (28.6%). The image formed by vessels associated with 75% underdiagnosis of true plus disease (i.e., 25% sensitivity cutoff, Fig. 3C) was diagnosed as "plus" by seven (100%) evaluators, and was judged most severe among the composite images by two (28.6%).

Discussion

To our knowledge, this is the first study to use quantitative vascular parameters to generate composite wide-angle retinal images at different diagnostic thresholds of recognized ROP experts. The key findings were that (1) computer-based image analysis permits quantification of retinal vascular features, and a spectrum of abnormalities in arterial curvature and venous diameter is seen in ROP; and (2) selection of appropriate vessels from multiple photographs can produce composite plus disease images corresponding to expert opinions.

Accurate and reliable detection of plus disease is critical to the management of ROP. A standard published narrow-angle photograph displays the minimum amount of vascular abnormality required for presence of plus disease.[4–6] However, our previous studies have demonstrated that there is significant variability in plus disease diagnosis among recognized ROP experts reviewing wide-angle retinal photographs.[10–12,15] The variability among responses provided by the evaluators in this study (Table 2) supports this notion. This highlights the subjective nature of diagnostic judgments by experts and suggests that there are not always clear distinctions between "plus" and "not plus." Instead, there appears to be a spectrum of findings that consists of retinas that clearly do not reflect plus disease, those that clearly do reflect plus disease, and those that do not clearly fall into either category. Although the new "pre-plus" categorization in the international classification of ROP may be intended to represent this latter category,[5] we have shown in a previous study that disagreement among experts persists even with a three-level (plus, pre-plus, neither) system.[10]

The spectrum of changes in arterial curvature and venous diameter, as a function of diagnoses given by experts, may be seen by comparing the images in Figure 3. This could provide insight into how the definition of plus disease is interpreted by experts and may improve the way in which ophthalmologists are taught to interpret the range of vascular changes. As shown in Table 1 and Figure 3A, the image at the 75% sensitivity cutoff illustrates the level associated with 25% underdiagnosis of plus disease and therefore has less AIC and VD than do the images at the 50% and 25% sensitivity cutoffs (Figs. 3B, 3C). Of note, the AIC value at the 25% sensitivity cutoff (0.061) is 45.2% greater than the corresponding value at the 75% sensitivity cutoff (0.042). In comparison, the venous diameter value at the 25% sensitivity cutoff (4.272) is only 12.6% greater than the corresponding value at the 75% sensitivity cutoff (3.795). This is consistent with findings of several recent studies involving automated ROP image analysis systems, which have suggested that the curvature of retinal vessels may be a more useful measure of plus disease than is vascular dilation.[17,18,20] It is conceivable that vascular diameter values may be confounded by effects such as photographic blurring, variances in axial length or corneal power, and differences in image acquisition technique that result in variable image magnification. However, if arterial curvature is truly shown in future studies to be more important than venous diameter, then an optimal approach for image generation may require unequal weighting of parameters. We have explored the use of linear combinations of multiple parameters to improve accuracy of computer-based ROP image analysis systems,[11–12,15] and this may provide a mechanism for integrating and weighting multiple features.

This methodology for image generation may be applied toward developing a definition for plus disease based on quantitative parameters. To illustrate the feasibility of this approach, Figure 4 compares the appearance of two images generated in this study to published photographs selected by the expert committee. Figures 4A and 4B display the composite image reflecting 50% underdiagnosis of true plus disease (i.e., 50% sensitivity cutoff), which has been cropped and magnified to match the perspective of the standard photographic definition of plus disease.[6] Figures 4C and 4D display the composite image reflecting 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff), compared with a published example of pre-plus disease.[5] This allows direct comparison of features from these images, and illustrates potential limitations of a photographic definition with a smaller field-of-view than indirect ophthalmoscopy or wide-angle imaging. The comparison also illustrates that pre-plus disease and plus disease may represent a spectrum of disease severity that can be quantified using parameters such as integrated curvature and diameter to represent vessel characteristics.

The development of a quantitative definition for plus disease could improve diagnostic accuracy and reliability, but would require identification of the actual vascular features that most closely correlate with the presence of plus disease as judged by experts. One practical difficulty is that there is no consensus regarding which exact aspects of the published standard photograph should actually be considered during diagnosis.[4–6] This study used AIC and VD for image generation, because these parameters have been shown to correlate closely with detection of plus disease by experts and because the narrative description of plus disease is characterized by "arterial tortuosity" and "venous dilation."[4–6,11–12,15] Responses from evaluators appear to support the notion that these composite images represent various diagnostic thresholds of disease. For example, all evaluators ranked Figure 3A as having the least severe vascular abnormality, and all diagnosed Figure 3C as "plus disease," as might be predicted. However, more evaluators ranked Figure 3B as having the most severe vascular changes among the three composite images (Table 2). This raises the possibility that other characteristics beyond arterial tortuosity and venous dilation may be considered by experts while assessing plus disease. Techniques from cognitive science such as think-aloud methodologies[21] may provide insight about retinal features that are truly perceived as important for diagnosis and whether other attributes such as vascular branching or congestion are considered during real-world disease management. Because the current photographic standard has been shown in major studies to have prognostic significance,[6,8] development of a quantitative definition would very likely require prospective validation in a clinical trial or retrospective validation using images from premature infants in whom the natural history of untreated ROP is known.

Quantitative interpretation of anatomic characteristics is often useful for image-based diagnosis, and methodologies similar to those described in this article might be applied to other diseases. Parameters of optical coherence tomography (OCT) imaging have been used to detect the presence of glaucoma with good sensitivity and specificity.[22] Structural features of the optic nerve head, such as disc size and retinal vascular arrangement, have correlated with expert opinion by using Rasch analysis to determine whether certain anatomic characteristics show greater heritance than others.[23] Mathematical techniques such as fractal models have been used to simulate neovascularization in corneal disease and diabetic retinopathy, and to generate computer-simulated images.[24,25] Analysis of diabetic retinopathy images, using artificial neural networks

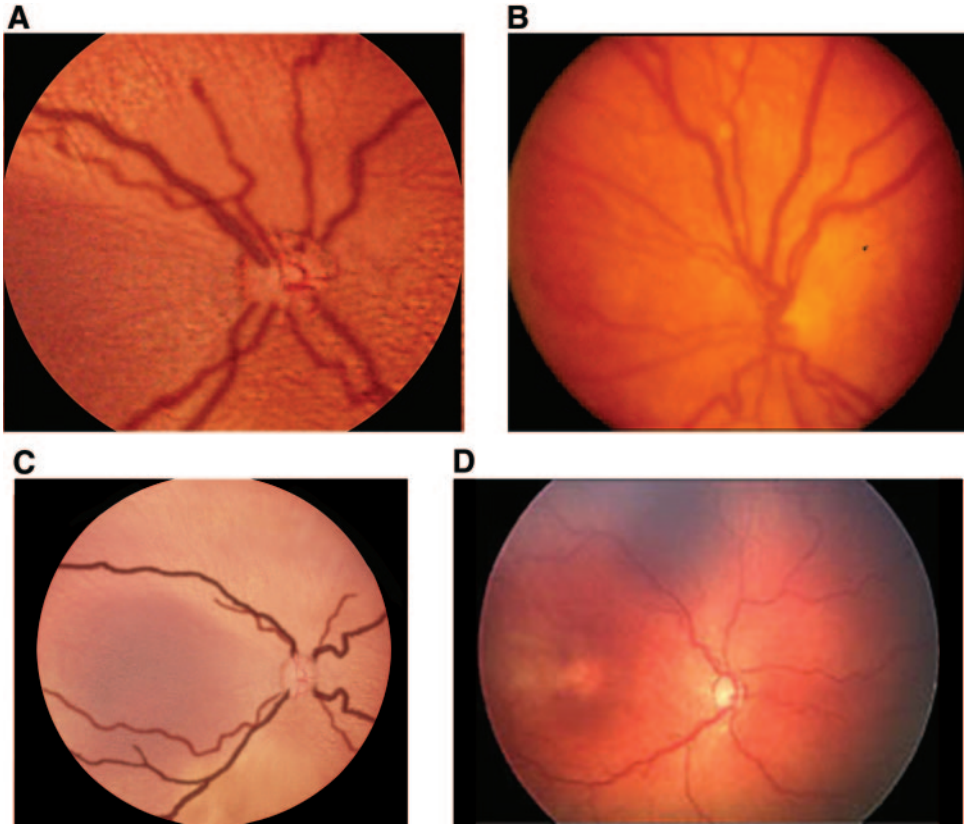

FIGURE 4. The *top* pair of images shows comparison of (**A**) composite generated image reflecting 50% underdiagnosis of true plus disease (i.e., 50% sensitivity cutoff) and (**B**) published standard photographic definition of plus disease selected by expert committee. Reprinted with permission from Multicenter Trial of Cryotherapy for Retinopathy of Prematurity. Preliminary results. Cryotherapy for Retinopathy of Prematurity Cooperative Group. *Arch Ophthalmol.* 1988;106: 471–479. Copyright © 1988, American Medical Association. All rights reserved. The *bottom* pair of images shows comparison of (**C**) composite generated image reflecting 25% underdiagnosis of true plus disease (i.e., 75% sensitivity cutoff), and (**D**) published example photograph of pre-plus disease selected by the expert committee. Reprinted with permission from An International Committee for the Classification of Retinopathy of Prematurity. The International Classification of Retinopathy of Prematurity Revisited. *Arch Ophthalmol.* 2005;123:991–999. Copyright © 2005, American Medical Association. All Rights reserved. The size of the composite generated images was adjusted to allow comparison at a similar level of magnification.

and quantification of image features, has demonstrated high sensitivity for detecting disease.[26–28] In general medicine, computer-assisted detection of quantitative features of breast masses can increase the sensitivity of mammography for some cancerous breast lesions.[29,30] Quantitative algorithms using mammographic features such as spiculation, border shape, and density may been used to classify breast masses.[30,31] Computer-based analysis of three-dimensional thoracic computed tomography images has been used to extract features such as pulmonary nodule structure, and these features have been used to predict the likelihood of malignancy with very high sensitivity.[32,33] Combining results from quantitative image analysis with diagnostic responses from multiple experts may make it feasible to develop similar composite images in diseases such as these for diagnosis, classification, and educational purposes.

Several limitations should be noted: (1) This study relied on a single set of 34 images that were examined by both the experts and the computer-based system. Larger studies may be necessary to validate the findings. (2) Our study did not account for any potential differences in magnification within the set of 34 images, which could have affected the measurement of system parameters such as vascular diameter. Of course, some variability in magnification may also be seen with standard indirect ophthalmoscopy. (3) System parameter thresholds were derived from analysis of vessels from all four quadrants in each image. This method may not necessarily be equivalent to clinical plus disease diagnosis, which is defined as the requisite amount of vascular abnormality in ≥2 quadrants.[5,7] Further work to derive thresholds based on the two quadrants with greatest vessel tortuosity and dilatation may be warranted.[4] Only sensitivity cutoffs were used to generate composite images, because the specificity curves were steeper than the sensitivity curves. As shown in Figure 2, there was a larger absolute slope from 25% to 75% specificity (AIC = 34.48, VD = 1.59) than from 25% to 75% sensitivity (AIC = 26.74, VD = 1.05). For this reason, composite images based on expert specificity over this range would look more similar to one another than composite images based on expert sensitivity.[5] Although reference standard diagnoses in this study were based on majority consensus of recognized ROP experts, they do not necessarily reflect the true presence of plus disease. Use of alternative reference standards based on indirect ophthalmoscopy or other methodologies for obtaining expert consensus may be informative.[6] It could be argued that the use of quantitative cutoff points as a diagnostic tool for plus disease could result in cases lost to therapy (false negatives). However, it has been shown that experienced ophthalmologists often disagree about the presence of plus disease,[10–12,15,34] presumably because they have different thresholds for diagnosis. An aggressive cutoff point could be selected so that no false-negative cases would occur in the opinion of any examiner, but this would result in many false-positive referrals because of the inherent tradeoff between sensitivity and specificity as the cutoff point is shifted. Therefore, we feel that it is most useful to visualize vascular changes over a range of cutoff values.

In conclusion, this study describes a methodology for quantifying characteristics of retinal images and generating composite plus disease images over a range of disease severities, based on the opinions of recognized ROP experts. The method may have application as an educational tool for ophthalmologists, may provide a mechanism for developing future quantitative definitions of plus disease, and may eventually be generalized to other image-based diseases.

Acknowledgments

The authors thank each of the 29 expert participants for their contribution to the study.

References

1. Munoz B, West SK. Blindness and visual impairment in the Americas and the Caribbean. *Br J Ophthalmol*. 2002;86:498–504.
2. Steinkuller PG, Du L, Gilbert C, et al. Childhood blindness. *J AAPOS*. 1999;3:26–32.
3. Gilbert C, Foster A. Childhood blindness in the context of VISION 2020: the right to sight. *Bull World Health Organ*. 2001;79:227–232.
4. Committee for the Classification of Retinopathy of Prematurity. An international classification of retinopathy of prematurity. *Arch Ophthalmol.* 1984;102:1130–1134.
5. International Committee for the Classification of Retinopathy of Prematurity. The international classification of retinopathy of prematurity revisited. *Arch Ophthalmol.* 2005;123:991–999.
6. Cryotherapy for Retinopathy of Prematurity Cooperative Group. Multicenter trial of cryotherapy for retinopathy of prematurity: preliminary results. *Arch Ophthalmol.* 1988;106:471–479.
7. STOP-ROP multicenter study group. Supplemental therapeutic oxygen for prethreshold retinopathy of prematurity (STOP-ROP): a randomized, controlled trial, I: primary outcomes. *Pediatrics.* 2000;105:295–310.
8. Early Treatment for Retinopathy of Prematurity Cooperative Group. Revised indications for the treatment of retinopathy of prematurity: results of the early treatment for retinopathy of prematurity randomized trial. *Arch Ophthalmol.* 2003;121:1684–1694.
9. Palmer EA, Hardy RJ, Dobson V, et al. 15-year outcomes following threshold retinopathy of prematurity: final results from the multicenter trial of cryotherapy for retinopathy of prematurity. *Arch Ophthalmol*. 2005;123:311–318.
10. Chiang MF, Jiang L, Gelman R, Du YE, Flynn JT. Interexpert agreement of plus disease diagnosis in retinopathy of prematurity. *Arch Ophthalmol*. 2007;125:875–880.
11. Chiang MF, Gelman R, Jiang L, Martinez-Perez ME, Du YE, Flynn JT. Plus disease in retinopathy of prematurity: an analysis of diagnostic performance. *Trans Am Ophthalmol Soc*. 2007;105:73–85.
12. Gelman R, Jiang L, Martinez-Perez ME, Du YE, Flynn JT, Chiang MF. Plus disease in retinopathy of prematurity: pilot study of computer-based and expert diagnosis. *J AAPOS*. 2007;11:532–540.
13. Martinez-Perez ME. Hughes AD, Thom SA, et al. Segmentation of blood vessels from red-free and fluorescein retinal images. *Med Image Anal*. 2007;11:47–61.
14. Martinez-Perez ME, Hughes AD, Stanton AV, et al. Retinal vascular tree morphology: a semi-automatic quantification. *IEEE Trans Biomed Eng*. 2002;49:912–917.
15. Koreen S, Gelman R, Martinez-Perez ME, et al. Evaluation of a computer-based system for plus disease diagnosis in retinopathy of prematurity. *Ophthalmology*. 2007;114:e59–67.
16. Gelman R, Martinez-Perez ME, Vanderveen DK, et al. Diagnosis of plus disease in retinopathy of prematurity using retinal image multiscale analysis. *Invest Ophthalmol Vis Sci*. 2005;46:4734–4738.
17. Swanson C, Cocker KD, Parker KH, et al. Semiautomated computer analysis of vessel growth in preterm infants without and with ROP. *Br J Ophthalmol*. 2003;87:1474–1477.
18. Wallace DK, Jomier J, Aylward WR, Landers MB. Computer-automated quantification of plus disease in retinopathy of prematurity. *J AAPOS*. 2003;7:126–130.
19. Johnson KS, Mills MD, Karp KA, Grunwald JE. Semiautomated analysis of retinal vessel diameter in retinopathy of prematurity patients with and without plus disease. *Am J Ophthalmol*. 2007;143(4):723–725.
20. Wallace DK, Zhao Z, Freedman SF. A pilot study using "ROPtool" to quantify plus disease in retinopathy of prematurity. *J AAPOS*. 2007;11:381–387.
21. Patel VL, Kaufman DR. Medical informatics and the science of cognition. *J Am Med Inform Assoc*. 1998;5:493–502.
22. Burgansky-Eliash Z, Wollstein G, Chu T, et, al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci.* 2005;46:4147–4152.
23. Hewitt AW, Poulsen JP, Alward WL, et al. Heritable features of the optic disc: A novel method for determining genetic significance. *Invest Ophthalmol Vis Sci*. 2007;48:2469–2475.
24. Landini G, Misson G. Simulation of corneal neovascularization by inverted diffusion limited aggregation. *Invest Ophthalmol Vis Sci*. 1993;34:1872–1875.
25. Xu X, Li B, Florez J, Li H. Simulation of diabetic retinopathy neovascularization in color digital fundus images. *Proceedings of the 2nd International Symposium on Visual Computing*. New York: Springer LNCS; 2006;421–433.
26. Usher D, Dumskyj M, Himaga M, Williamson TH, Nussey S, Boyce JF. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. *Diabet Med*. 2004;21(1):84–90.
27. Sinthanoyothin C, Boyce JF, Williamson TH, et al. Automated detection of diabetic retinopathy on digital fundus images. *Diabet Med*. 2002;19:105–112.
28. Niemeijer M, van Ginneken B, Russell SR, Suttorp-Schulten MS, Abramoff MD. Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. *Invest Ophthalmol Vis Sci*. 2007;48:2260–2267.
29. Birdwell R, Ikeda D, O'Shaughnessy KF, Sickles EA. Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology*. 2001;219:192–202.
30. Warren BL, Wood S, D'Orsi C, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology*. 2000;215(2):554–562.
31. Sahiner B, Chan H, Petrick N, et al. Improvement of mammographic mass characterization using spiculation measures and morphological features. *Med Phys*. 2001;28:1455–1465.
32. Li H, Wang Y, Liu K, et al. Computerized radiographic mass detection-part II: decision support by featured database visualization and modular neural networks. *IEEE Trans Med Imaging*. 2001;20:302–313.
33. Kawata Y, Niki N, Ohmatsu H, Moriyama N. Example-based assisting approach for pulmonary nodule classification in three-dimensional thoracic computed tomography images. *Acad Radiol*. 2003;10(12):1402–1415.
34. Reynolds JD, Dobson V, Quinn GE, et al. Evidence-based screening criteria for retinopathy of prematurity: natural history data from the CRYO-ROP and LIGHT-ROP studies. *Arch Ophthalmol*. 2002;120:1470–1476.