Automated Computer System Grading of Diabetic Retinopathy from Fundus Photographs*

Otomatik Bilgisayarlı Sistemle Gözdibi Fotoğraflarından Diabetik Retinopatinin Derecelendirilmesi

Yaprak Banu ÜNVER[1], Gülderen AKTAN YAVUZ[2], Stephen H. SINCLAIR[3]

Original Article *Klinik Çalışma*

ABSTRACT

Purpose: To evaluate the efficacy of computer algorithms in the detection and grading of retinopathy derived from screening of diabetic photographs.

Methods: Screening fundus photographs (5 per eye) taken with a mydriatic camera in a general ophthalmology clinic for 52 sequential diabetic patients were graded by a general ophthalmologist and a retina specialist and compared with computer grading of digitized images acquired from color slides.

Results: Among the 104 eyes, 64 (61.5%) had no retinopathy lesions (Grade 10), 6 (5.8%) eyes had only 1 or 2 microaneurysms, 21 (20.2%) eyes had more than 2 microaneurysms but no other lesions (grade 21), and 18 (17.3%) had additional lesions (grade 21+). Among the ophthalmologists, there was exact agreement in the grading of 76 eyes; in 21 eyes grading by each was within 1 grade level, and in 7 the grades differed by 2 grade levels. Computer algorithm grading agreed with the ophthalmologists' grade in 51 (50%) eyes; in 24 (23%) eyes the computer under-graded by one grade level, and by 2 grades in 2 eyes. In 22 (21%) eyes, computer grading was greater by one grade level than the ophthalmologists' grading and in 5 (5%) eyes the algorithms overgraded by 2 grades.

Conclusions: Computer algorithms were able to detect and grade retinopathy from fundus photographs taken of a set of sequentially screened diabetics with consistency similar to the agreement in grading between a general ophthalmologist and a retina specialist.

Key Words: Diabetic Retinopathy, Screening, Digitized Fundus Photographs, Computer Algorithm.

ÖZ

Amaç: Diabetik hastaların fotografları tarayarak elde edilen retinopati bulunmasında ve derecelendirmesinde bilgisayar algoritimalarının etkinliğini araştırmak.

Gereç ve Yöntem: Genel göz kliniğinde ardışık 52 diabetik hastanın midriatik kamera ile alınan gözdibi tarama fotografları (her göz için 5) genel göz doktorları ve retina uzmanları tarafından derecelendirildi ve renkli resimlerden elde edilen dijital imajların bilgisayar derecelendirilmesi ile karşılaştırıldı.

Bulgular: Toplam 104 gözden 64 (%61.5) gözde retinopati lezyonu yoktu (Grade 10), 6 (%5.8) göz sadece 1 yada 2 mikroanevrizmaya, 21 (%20.2) göz 2'den fazla mikroanevrizmaya sahipti fakat başka lezyon yoktu. (grade 21) ve 18 (%17.3) gözde ek lezyonlar vardı (grade 21+). Tüm göz doktorları, 76 gözün derecelendirilmesinde bilgisayar derecelendirmesi ile tamamen aynı fikirde idi, her biri arasında 21 gözün derecelendirmesinde bir seviye, ve 7'sinde dereceler arasında 2 derece seviyesi fark vardı. 51 (%50) gözün bilgisayar algoritma derecelendirmesi ile göz doktoru derecelendirmesi aynı idi; 24 (%23) gözde bilgisayar derecelendirmesi göz doktoru derecelendirmesinden bir seviye aşağı ve 2 gözde de iki seviye aşağı idi. 22 (%21) gözde bilgisayar değerlendirmesi göz doktoru derecelendirmesinden bir seviye yukarı ve 5 (%5) gözde 2 derece fazla idi.

Sonuç: Bilgisayar algoritmaları önceden hazırlanmış bir ardışık diabetik hasta taramasından alınan gözdibi fotograflarından, genel göz doktorları ve retina uzmanları arasında derecelendirmedeki karar uygun benzerlikte, retinopatiyi tespit edebildi ve derecelendirebildi.

Anahtar Kelimeler: Diabetik retinopati, tarama, dijital fundus fotoğrafları, bilgisayar algoritimi.

Ret-Vit 2009;17:114-119

Geliş Tarihi : 12/03/2009
Kabul Tarihi : 14/04/2009

Received : March 12, 2009
Accepted : April 14, 2009

* This study was Supported by TUBİTAK.

1- Beyoğlu Göz Eğitim ve Araştırma Hastanesi, 2. Klinik, İstanbul, Uzm. Dr.
2- Düzce Üniversitesi Tıp Fakültesi, Göz A.D., Düzce, Prof. Dr.
3- Drexel Üniversitesi Tıp Fakültesi, Göz A.D., Philadelphia, PA, ABD, Prof. Dr.

1- M.D, Beyoğlu Eye Research and Education Hospital, Kuledibi İstanbul/TURKEY
ÜNVER Y.B., yaprakbanu@yahoo.com
2- M.D Proffessor, Düzce University Faculty of Medicine, Department of Ophthalmology Düzce/TURKEY
AKTAN G., aktangulderen@hotmail.com
3- M.D Proffessor, Drexel University Faculty of Medicine, Department of Ophthalmology Philadelphia PA USA
SINCLAIR S.H.,

Correspondence: M.D. Yaprak Banu ÜNVER
Beyoğlu Eye Research and Education Hospital, Kuledibi İstanbul/TURKEY

INTRODUCTION

Despite the availability of recognized effective therapy, diabetic retinopathy has become the leading cause of new blindness and the leading overall cause of blindness in the United States for persons of age 25 to 65.[1] This appears to be primarily because of the lack of adequate screening to detect retinopathy lesions early in the course of the disease.[2] If not discovered until the patient has sufficient vision problems to initiate a visit to the physician, the disease most often is severely progressed and treatment, although successful at preventing further progression of the disease, is only rarely able to restore vision.[3,4] Regular, fundus screening has been recommended by many health care agencies, but only approximately 40% to 58% of primary care physicians either perform ophthalmoscopy or refer their patients to an eye-care provider.[2] A further problem is that most often such recommendations are heeded by only approximately 40-50% of asymptomatic diabetic patients[5,6] because of multiple factors, including procrastination driven by the fear of discovering a potentially blinding condition, their difficulty in scheduling the time for physician appointments, unreimbursed costs of the examination, and their desire in general to avoid physicians after years with the disease.

Published guidelines have traditionally recommended that retinopathy screening be fulfilled by an annual fundus examination performed with pupillary dilation. However, this method has significant deficiencies: it is expensive, it is extremely variable in the accuracy of detection among providers[7,8] (many of whom do not dilate the pupils[9]), and has been shown to be less accurate than the evaluation of fundus photographs.[10] Photographic screening is not only more convenient, less costly, and demonstrates improved detection rates[10-12], it appears in early studies to be less confrontational to the patient and to improve screening compliance when provided as a walk-in, non-scheduled convenience in the primary-care center

The human evaluation of fundus photographs, however, is not without variability between trained examiners[13,14] and among examiners over time due to fatigue. This variability limits the sensitivity of the data that can be utilized in the evaluation of changes occurring over time, and heretofore has limited the evaluation to relatively gross changes occurring in the severity of retinopathy.

We have attempted to enhance the lesion detection and have superimposed an artificial intelligence layer to allow the set to grade the retinopathy. Previous attempts have been made to develop digital image processing techniques in order to automatically detect retinal lesions[15,16] but these authors have not attempted to evaluate the sensitivity of such algorithms to grade photographs acquired under routine, office, screening conditions. Herein, we report the ability of a set of expert system algorithms to detect the lesions and grade retinopathy within a set of unselected screening photographs taken of diabetics.

MATERIALS AND METHODS

The screening fundus photographs, which had been previously taken in a general ophthalmology clinic for a sequential series of 52 diabetic patients, were reviewed. The photographs consisted of 5 fields taken for each eye (total 104 eyes) after pupillary dilation (with tropicamide 1% and phenylephrine 2.5%) utilizing a Topcon 50X camera at 30° on Ektachrome 35-mm color slides. The 5 fields photographed were as follows:

1) centered on fovea,

2) centered on disc,

3) fovea located near inferior periphery of photograph,

4) fovea located near right or left (depending on whether right or left eye) periphery of photograph (in order to view the temporal perifoveal retina), and

5) fovea located near superior periphery of photograph). Examples of the average photograph quality are demonstrated in figures 1-3A.

The photographs were graded in a masked fashion by a general ophthalmologist and a retina specialist, and any differences in grading were arbitrated to produce a final grade among three retinopathy levels: no retinopathy (grade 10), microaneurysms alone (grade 21)

Table 1: Comparison of ophthalmologists' final grading with computer grading of 104 screened eyes.

		Computer Grade		
		10	**21**	**21+**
	10	41 (39.4%)	21 (20.2%)	2 (1.9%)
	21	14 (13.5%)	7 (6.7%)	1 (1.0%)
Human Grade	**21+**	5 (4.8%)	10 (9.6%)	3 (2.9%)

Table 2: Comparison of ophthalmologists' final grade with computer grade of 104 eyes.

Difference in grading of all eyes	number
Exact agreement	51 (50%)
Computer one grade less than ophthalmologist arbitrated grade	24 (23%)
Computer one grade higher than ophthalmologist arbitrated grade	22 (21%)
Computer two grades less than ophthalmologist arbitrated grade	5 (5%)
Computer two grades higher than ophthalmologist arbitrated grade	2 (2%)

Table 3: Comparison of ophthalmologists' final grading with computer grading of 52 patients based on the grading of the worst eye.

Difference in grading of worst eye	Computer vs Ophthalmologist Arbitrated Grade
Exact agreement	20 (38%)
Computer one grade less than ophthalmologist arbitrated grade	15 (28%)
Computer one grade higher than ophthalmologist grade	14 (26%)
Computer two grades less than ophthalmologist grade	2 (4%)
Computer two grades higher than ophthalmologist grade	1 (2%)

or microaneurysms with other lesions (grade 21+).[17] After grading the color photographs, the retina specialist also marked the following retinopathy lesions on the digital images after comparing them with the color slides (using NIH Image on a Macintosh): dot hemorrhages, blot hemorrhages, striate hemorrhages, lipid exudates, and nerve-fiber-layer infarcts (examples are shown in figures 1B and 1C).

The photographs were digitized at 1024x1024 resolution and 8-bit density (LS-3500 slide scanner, Nikon,) using a Wratten green filter. The digital images were subjected to the algorithm set to detect the same lesions and then grade the retinopathy based on the sum of the lesions identified in each of the five fields. The computer-derived retinal grading was compared with the final retinopathy grade determined by the two ophthalmologists. In addition, the identification of the individual lesions indicated by the computer was compared with that provided by the retina specialist (examples figure 2, for lipid exudates; and figure 3, for cotton-wool spots).

RESULTS

Among the 104 diabetic eyes that underwent screening photography, 64 (61.5%) had no retinopathy lesions (Grade 10) where as 6 (5.8%) eyes had only 1 or 2 microaneurysms, 14 eyes (13.5%) had more than 2 microaneurysms but no other lesions (totaling 22 eyes with grade 21, (21.2%), and 18 (17.3%) had additional lesions (grade 21+). Differences in the grading of eyes were noted between the retina specialist and general ophthalmologist. Among the 104 eyes, there was exact agreement in the grade given in 76, whereas in 21 eyes grades assigned by the two were within 1 grade level. In 7 eyes, the grades assigned by the retina specialist and general ophthalmologist differed by 2 grade levels. This was similar to the number of eyes in which the ophthalmologists differed by 2 grade levels and for the most part represented similar difficulties in detecting (or over detecting) cotton-wool spots or in differentiating lipid granular or punctate exudates from drusen in which the computer algorithms demonstrated the same difficulty as the ophthalmologists in these screening photographs. Therefore the ophthalmologists agreed exactly on the retinopathy grade in 75% of the eyes. The comparison of the computer grading with the final grade level given by the two ophthalmologists is demonstrated in table-1, and the results are summarized in table-2. In 51 (50%) eyes, the computer grading exactly matched that of the arbitrated grade of the ophthalmologists. In 24 (23%) eyes the computer under-graded an eye by one grade level, and in 22 (21%) eyes the computer grading was greater by one grade. Most all of these eyes represented either missing or detecting 1 or 2 microaneurysms or missing small foci of lipid exudates (or labeling drusen as exudates when the ophthalmologists decided otherwise). In 21 eyes the ophthalmologists also disagreed by 1 grade level in their grading for similar reasons.

Since patients are referred from screening for further care based upon the level of grading of the worst

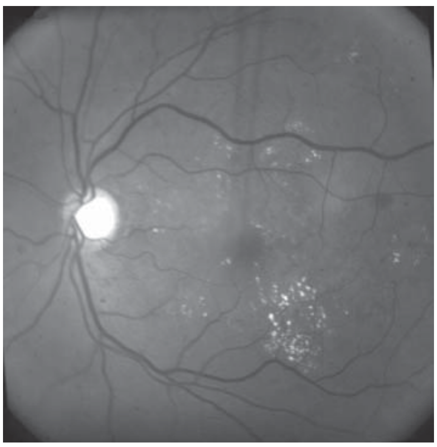

Figure 1A: Red-free, retinal photograph digitized from color slide demonstrating non-proliferative diabetic retinopathy.

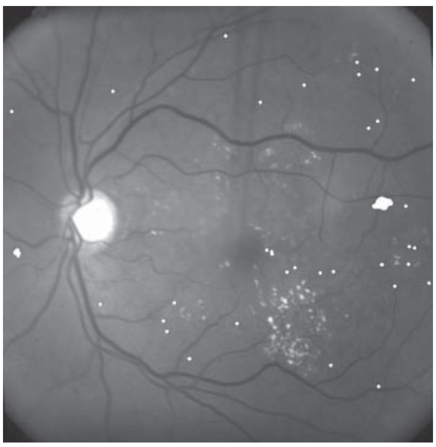

Figure 1B : The digital image has been marked by a retinal specialist with a painting program identifying the dot and blot hemorrhages.

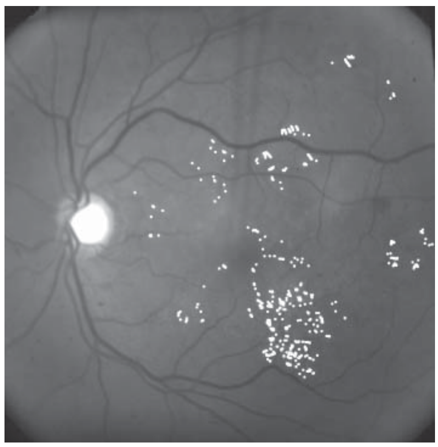

Figure 1C: lipid exudates.

Figure 2A: Red-free, retinal photograph digitized from color slide demonstrating non-proliferative diabetic retinopathy.

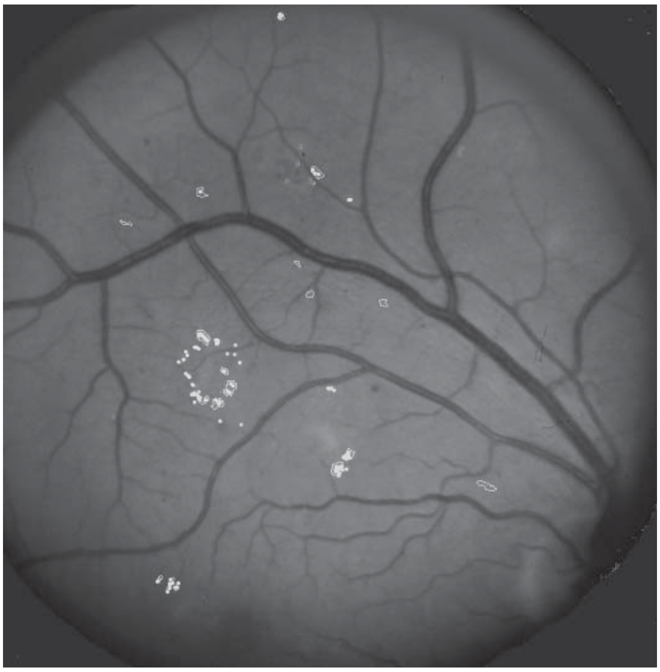

Figure 2B: The digital image has been marked by a retinal specialist with a painting program to identify the lipid exudates. Those exudates identified by the computer algorithms are circled.

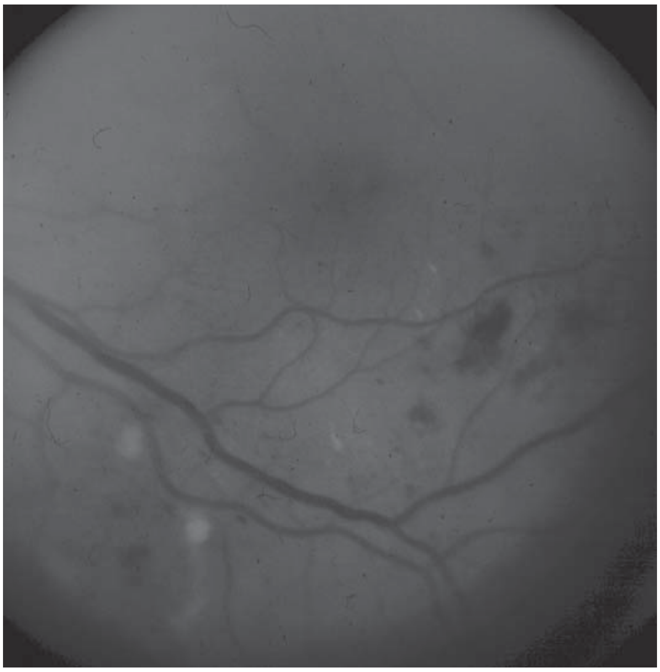

Figure 3A: Red-free, retinal photograph digitized from color slide demonstrating non-proliferative diabetic retinopathy.

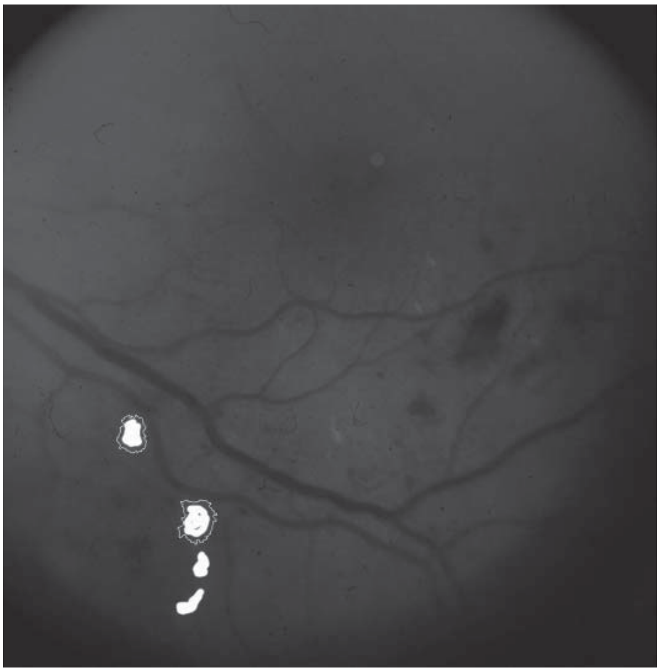

Figure 3B: The digital image has been marked by a retinal specialist with a painting program to identify cotton-wool spots. The infarcts identified by the computer algorithms are circled.

eye, comparison was examined between the computer algorithms and the arbitrated grading of the ophthalmologists for the grade level given the worst eye. The results are presented in table-3. There was exact agreement in 20 (38%) eyes, computer grading was one grade less than the ophthalmologist in 15 (28%) eyes, computer grading was one grade higher in 14 (26%) eyes, computer grading was two grades less than the ophthalmologists grade in 2 (4%) eyes, and computer grading was two grades higher in 1 (2%) eye.

DISCUSSION

In this study, we have evaluated an automated image-processing and expert-systems algorithm set in the accuracy of detecting lesions and grading retinopathy from digitized images acquired from color slides taken under screening circumstances with a mydriatic camera. Against the red background of the choroid on film images, the human eye as well as the computer algorithms appears to have similar difficulty in differentiating whether minimally contrasting lesions represent true hemorrhagic lesions especially when they are few in number. We utilized a Wratten green filter in an attempt to improve the detection of the hemorrhagic lesions since we subjectively have observed that such filters improve their observation. However, using the green filter, while enhancing the detection of hemorrhagic lesions, was observed to reduce the texture patterns of others, making the differentiation of lipid exudates and cotton wool infarcts more difficult. This suggests that texture mapping to detect the different lesions is perhaps best done at different wavelengths. Improved performance of the algorithm set, either by taking multiple monochromatic images at different wavelengths, or by the analysis of digital color images, is being evaluated.

Using second capture data from a film intermediary, rather than direct CCD digital system, results in an image with poorer contrast, detail and texture. Direct digital capture retinal camera systems, in our experience, produce a higher-grade image leading to improved analysis. We conducted the analysis on digital images acquired from color film slides for this paper since we felt the quantitative software analysis would be of value to screening centers that utilized film, although digital systems, despite upfront capital investment, with sufficient utilization appear to offer significant reduction in costs per screening than when done with film.

The individual lesions identified by the image analysis software were subjected to an artificial intelligence expert systems envelope that graded the retinopathy into 3 levels. The decision of the grading envelope was optimized first to minimize the number of more severe lesions missed (lipid exudates, cotton-wool spots and hemorrhages in that order) that would lead to under-referral and secondly to reduce the number of retinopathy overgrades that would result in inappropriate over-referral after screening. The grading was forced into 3 levels to separate those patients that need only repeated screening at lengthy intervals (no retinopathy or occasional microaneurysms) from those that need more careful evaluation. We envision this as the primary purpose of retinopathy screening, removing the system from the task of micromanagement of the retinopathy, once discovered. However, after an individual is photographed several times, the quantitated history of particular lesions over time (e.g. changes in the number of microaneurysms or blot hemorrhages) may provide a truer indication of the risk for progression than changes in the composite grades. Our intent is to collect follow-up data on patients screened with the same system to evaluate lesion specific history data. In addition, the use of image processing software allows the evaluation of other characteristics of the fundus that cannot be evaluated by human observation but which are known markers of disease and may be beneficial in establishing risk for the development of retinopathy or its progression especially in the early stages. These may include, but are not limited to the evaluation of vascular diameter and variation, tortuosity, and branching angles.

In conclusion, we have demonstrated the utility of a computer image processing algorithm set that independently detects and quantitates the retinal lesions of early diabetic retinopathy and, using the accumulation of the lesions, grades the retinopathy. The system appears to have sufficient sensitivity and specificity to provide for a methodology that can be placed in the primary care setting to allow economical retinopathy screening. We believe that this will significantly improve the current screening rates that are limited by poor patient compliance.[18-20]

Undilated digital-video images using the Joslin Vision Network (JVN)[21,22] system were comparable photographs for the determination of diabetic retinopathy level. [23] The results validate the agreement between nonmydriatic JVN images and dilated ETDRS photographs and suggest that this digital technique may be an effective telemedicine tool for remotely determining the level of diabetic retinopathy, suggesting timing of next retinal evaluation and identifying the need for prompt referral to ophthalmology specialists.[23] Furthermore, by providing real-time, onsite analysis of image quality, the photographer can be advised of those needing dilation or further ophthalmologic examination. Moreover, by providing onsite, rapid examination of the photographs for lesions, the patient can be advised immediately of whether he/she has retinopathy requiring further evaluation, which we believe will improve compliance with the recommended follow-up.[24,25]

We would also envision the development of an electronic feedback response from the ophthalmologist to the primary care screening center that could be used in the database to provide "learning" by the screening algorithm set and improve the proficiency of the screening.

REFERENCES/KAYNAKLAR

1. American Diabetes Association PS.: Implications of the diabetes control and complications trial. Clin Diabetes. 1993;11:91-96.
2. Klein R.: Barriers to prevention of vision loss caused by diabetic retinopathy. Arch Ophthalmol. 1997;115:1073-1074.
3. Early Treatment Diabetic Retinopathy Study Research Group.: Photocoagulation for diabetic macular edema: Early Treatment Diabetic Retinopathy Study Report no. 4. Int Ophthalmol Clin. 1987;27:265-272.
4. Diabetic Retinopathy Study Research Group.: Photocoagulation treatment of proliferative diabetic retinopathy. Ophthalmology. 1981;88:583.
5. Donovan J.: Patient decision making: The missing ingredient in compliance research. International journal of technology assessment in health care. 1995;11:443-455.

6. Brechner R, Cowie C, Howie L, et al.: Ophthalmic examination among adults with diagnosed diabetes mellitus. JAMA. 1993; 270:1714-1718.
7. Sussman E, Tsiaras W,Soper K.: Diagnosis of diabetic eye disease. JAMA. 1982;247:3231-3234.
8. Kleinstein R, Roseman J, Herman W, et al.: Detection of diabetic retinopathy by optometrists. J Am Optom Assoc. 1987;58:879-882.
9. Kraft S, Marrero D, Lazaridis E, et al.: Primary care physicians' practice patterns and diabetic retinopathy. Current levels of care. Arch FAM Med. 1997;6:29-37.
10. Moss S, Klein R, Kressler S, Richie K.: Comparison between ophthalmoscopy and fundus photography in determining severity of diabetic retinopathy. Ophthalmology. 1985;92:62-67.
11. Valez R, Haffner S, Stern M.: Ophthalmologist vs retinal photographs in screening for diabetic retinopathy. Clin Res. 1987;35:363A.
12. O'Hare J, Hopper A, Madhaven C, et al.: Adding retinal photography to screening for diabetic retinopathy: a prospective study in primary care. BMJ. 1996;312:679-682.
13. Klein R, Moss S, Klien B.: New management concepts for timely diagnosis of diabetic retinopathy treatable by photocoagulation. Diabetes Care. 1987;10:633-638.
14. Early Treatment Diabetic Retinopathy Study Research Group.: Grading diabetic retinopathy from stereoscopic color fundus photographs-an extension of the modified Airlie house classification: ETDRS report number 10. Ophthalmology. 1991;98:786-806.
15. Goldbaum M, Katz N, Nelson M, et al.: The discrimination of similarly colored objects in computer images of the ocular fundus. Invest Ophthalmol Vis Sci. 1990;31:617-623.
16. Cree M, Olson J, McHardy K, et al.: A fully automated comparative microaneurysm digital detection system. Eye. 1997;11:622-628.
17. Early Treatment Diabetic Retinopathy Study Research Group.: Grading diabetic retinopathy from stereoscopic color fundus photographs-an extension of the Modified Airlie House Classification: ETDRS report number 10. Ophthalmology. 1991;98:786-806.
18. Sinclair S, Olarn R, DeVaro J, et al.: Fundus examination of diabetic inpatients. Invest Ophthalmol Vis Sci. 1989;30:79.
19. Jacques C, Jones R, Houts P, et al.: Reported practice behaviors for medical care of patients with diabetes mellitus by primary-care physicians in Pennsylvania. Diabetes Care. 1991;14:712-717.
20. Peters A, Davidson M, Ziel F.: Cost-effective screening for diabetic retinopathy using a non-mydriatic retinal camera in a prepaid health-care setting. Diabetes Care. 1993;16:1193-1195.
21. Ahmed J, Ward TP, Bursell SE, et al.: The sensitivity and specificity of nonmydriatic digital stereoscopic retinal imaging in detecting diabetic retinopathy. Diabetes Care. 2006;29:2205-2209.
22. Chow SP, Aiello LM, Cavallerano JD, et al.: Comparison of nonmydriatic digital retinal imaging versus dilated ophthalmic examination for nondiabetic eye disease in persons with diabetes. Ophthalmology. 2006;113:833-840.
23. Bursell SE, Cavallerano JD, Cavallerano AA, et al.: Joslin Vision Network Research Team. Stereo nonmydriatic digital-video color retinal imaging compared with Early Treatment Diabetic Retinopathy Study seven standard field 35-mm stereo color photos for determining level of diabetic retinopathy. Ophthalmology. 2001;108:572-585.
24. Sprafka J, Fritsche T, Baker R, et al.: Prevalence of undiagnosed eye disease in high-risk diabetic individuals. Arch Intern Med . 1990;150:857-861.
25. Burns-Cox C, Dean Hart J.: Screening of diabetics for retinopathy by ophthalmic opticians. Brit Med J. 1985;290:1052-1054.