Clinical Validation Study of an automated DR Screening System against 7-field ETDRS Stereoscopic Reference Standard

Kaushal Solanki, Malavika Bhaskaranand, Chaithanya Ramachandra, Sandeep Bhat

PURPOSE AND BACKGROUND: Comprehensive clinical validation of an automated diabetic retinopathy (DR) screening system, EyeArt v2.0 for detecting referable diabetic eye disease (DED) (moderate non-proliferative DR (NPDR) or higher on the ICDR scale and/or surrogate markers for clinically significant macular edema (CSME)). To the best of our knowledge, this is the first time an automated DED screening system has been comprehensively validated against ETDRS 7-field stereoscopic reference standard

STUDY DESIGN AND METHODS: Retrospective observational study evaluating automated screening system on anonymized fundus images. 7-field ETDRS stereoscopic fundus images of 755 eyes from DRCRnet were graded on the ETDRS scale. EyeArt was evaluated on 2 sets of 10,000 Monte-Carlo experiments, each with 10,000 eyes, with first set using 3-retinal monoscopic fields (macula centered, optic nerve centered, and temporal) and the second set using 1-retinal monoscopic field (macula centered). Both sets simulated a screening population: 68% no DR cases, 10% mild NPDR cases, 14% moderate NPDR cases, 4% severe NPDR cases, and 4% PDR cases.

RESULTS: EyeArt results are summarized in table below.

Protocol	AUROC [95% CI]	Sensitivity % [95% CI]	Specificity % [95% CI]	Treatable DR Sensitivity % [95% CI]
3-field	**0.974** [0.971-0.977]	**94.4** [93.4-95.3]	**87.7** [87.0-88.3]	**98.8** [97.9-99.5]
1-field	**0.967** [0.963-0.971]	**91.5** [90.4-92.7]	**85.8** [85.0-86.5]	**95.4** [95.4-97.9]

PRÉCIS: Automated screening using EyeArt is safe and effective as demonstrated by high sensitivity/specificity when analyzing simulated screening populations with monoscopic 3-fields and 1-field per eye (typical in screening programs) against the gold standard grading of 7-field stereoscopic ETDRS fundus images. Sensitivity for treatable DED is high though EyeArt detects CSME only based on surrogate markers (standard practice in DR screening programs that use non-stereoscopic images).

CONCLUSION: Automated screening with EyeArt achieves high sensitivity and specificity on image sets with 3-fields, 1-field per eye against the gold standard 7-field ETDRS grading.

Comprehensive Clinical Validation Study of a Fully-automated Diabetic Retinopathy Screening System using Color Fundus Images Against 7-field ETDRS Stereoscopic Reference Standard

Kaushal Solanki, Malavika Bhaskaranand, Chaithanya Ramachandra, Sandeep Bhat [Eyenuk, Inc, Woodland Hills, CA, USA]

PURPOSE: Comprehensive clinical validation of an automated diabetic retinopathy (DR) screening system, EyeArt-v2.0 for detecting referable diabetic eye disease (DED) defined as presence of (i) moderate non-proliferative DR (NPDR) or higher on the International Clinical DR (ICDR) scale and/or (ii) surrogate markers for clinically significant macular edema (CSME) defined as hard exudates within one disc diameter of the macular center. Evaluation dataset comprises 755 eyes imaged using the 7-field Early Treatment DR Study (ETDRS) stereoscopic fundus photography protocol and graded on ETDRS scale. EyeArt is evaluated on two separate imaging protocols involving 2 and 3 retinal fields per eye respectively.

SETTING/VENUE: Patients were recruited at multiple clinical sites that were part of the Diabetic Retinopathy Clinical Research Network (DRCRnet) and underwent mydriatic 7-field ETDRS stereoscopic fundus photography. The fundus photographs were acquired using different commercially available table-top fundus cameras

METHODS: The DR severity level and CSME status were determined as per the ETDRS scale by certified reading-center graders for each eye using the 7-field ETDRS stereo photographs/images. The reference standard for evaluating EyeArt performance was obtained by determining the presence of referable DED using CSME presence and DR severity on the ICDR scale (mapped from the ETDRS DR severity level). The anonymized fundus images underwent automated analysis by EyeArt. EyeArt independently analyzed the following two imaging protocols with monoscopic (non-stereoscopic) images: (i) 3-field protocol with 3 images comprising ETDRS fields 1 (centered at the optic disc), 2 (centered at the macula), and 3 (temporal to the macula) and (ii) 2-field protocol with 2 images comprising ETDRS fields 1 and 2. A "refer" recommendation was provided for patients with apparent signs of referable DED (moderate NPDR or higher on the ICDR scale or with surrogate markers for CSME). A "no refer" recommendation was provided for patients with no apparent signs of DR or signs of mild DR without surrogate markers for CSME. EyeArt's referral recommendation performances was evaluated by using sensitivity, specificity, and area under the receiver operating characteristic curve (AUROC) along with corresponding 95% confidence intervals (CI).

RESULTS[*]: As per the ETDRS reference standard, 95.5% of the eyes have referable DED and 62.3% of the eyes have treatable DED (with CSME and/or severe NPDR or higher on ICDR scale). EyeArt analysis using the 3-field protocol for each eye achieved screening sensitivity of 94.3% [95% CI: 92.6% – 96.0%] at specificity of 84.4% [95% CI: 69.7% – 94.6%]. This corresponds to 687 "refer" recommendations and 41 false negatives of which 5 had treatable DED. The screening AUROC was 0.970 [95% CI: 0.953 – 0.985]. EyeArt's sensitivity for detecting treatable DED was 98.9% i.e. of the 470 eyes with treatable DED as per reference standard, 465 eyes were correctly provided "refer" recommendations. EyeArt analysis using the 2-field protocol for each eye achieved screening sensitivity of 90.5% [95% CI: 88.3% – 92.6%] at specificity of 87.5% [95% CI: 75.0% – 97.1%]. This corresponds to 658 "refer" recommendations and 69 false negatives of which 10 had treatable DED. The screening AUROC was 0.965 [95% CI: 0.943 – 0.982]. EyeArt's sensitivity for detecting treatable DED was 97.9% i.e. of the 470 eyes with treatable DED as per reference standard, 460 eyes were correctly provided "refer" recommendations.

CONCLUSIONS: Automated screening using EyeArt achieves high sensitivity and specificity when analyzing image sets with 3 images or 2 images per eye (typically used in DR screening programs) against the gold standard grading of 7-field stereoscopic ETDRS fundus photographs. To the best of our knowledge, this is the first time an automated screening system has been comprehensively validated against an ETDRS 7-field stereoscopic reference standard. EyeArt's sensitivity for referring treatable DED is high even though EyeArt detects CSME based on surrogate markers (hard exudates within 1 disc diameter of macular center) as is standard practice in screening programs that use non-stereoscopic fundus images. However, the reference standard for CSME grading is based on retinal thickening near the macula as discernable from stereoscopic fundus images. It is to be noted that the evaluation dataset used in this study has a significantly higher prevalence of referable and treatable DED (95.5% and 62.3% respectively) than in typical screening populations where the typical prevalence is about 20% and 5% respectively.

---

[*] Results updated since the original EURETINA submission in Mar 2016.