Combined automated screening for age-related macular degeneration and diabetic retinopathy in primary care settings

Alauddin Bhuiyan[1,2], Arun Govindaiah[1], Sharmina Alauddin[2], Oscar Otero-Marquez[2], R. Theodore Smith[2]

[1]Research & Development Department, iHealthScreen Inc., Richmond Hill, USA; [2]Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, New York, USA

Contributions: (I) Conception and design: A Bhuiyan, A Govindaiah; (II) Administrative support: A Bhuiyan; (III) Provision of study materials or patients: RT Smith; (IV) Collection and assembly of data: RT Smith, A Bhuiyan; (V) Data analysis and interpretation: A Govindaiah, O Otero-Marquez, S Alauddin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Alauddin Bhuiyan. Associate Professor, Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, New York, USA. Email: alauddin.bhuiyan@gmail.com.

Background: Age-related macular degeneration (AMD) and diabetic retinopathy (DR) are among the leading causes of blindness in the United States and other developed countries. Early detection is the key to prevention and effective treatment. We have built an artificial intelligence-based screening system which utilizes a cloud-based platform for combined large scale screening through primary care settings for early diagnosis of these diseases.

Methods: iHealthScreen Inc., an independent medical software company, has developed automated AMD and DR screening systems utilizing a telemedicine platform based on deep machine learning techniques. For both diseases, we prospectively imaged both eyes of 340 unselected non-dilated subjects over 50 years of age. For DR specifically, 152 diabetic patients at New York Eye and Ear faculty retina practices, ophthalmic and primary care clinics in New York city with color fundus cameras. Following the initial review of the images, 308 images with other confounding conditions like high myopia and vascular occlusion, and poor quality were excluded, leaving 676 eligible images for AMD and DR evaluation. Three ophthalmologists evaluated each of the images, and after adjudication, the patients were determined referrable or non-referrable for AMD DR. Concerning AMD, 172 were labeled referable (intermediate or late), and 504 were non-referable (no or early). Concurrently, regarding DR, 33 were referable (moderate or worse), and 643 were non-referable (none or mild). All images were uploaded to iHealthScreen's telemedicine platform and analyzed by the automated systems for both diseases. The system performances are tested on per eye basis with sensitivity, specificity, accuracy, and kappa scores with respect to the professional graders.

Results: In identifying referable DR, the system achieved a sensitivity of 97.0% and a specificity of 96.3%, and a kappa score of 0.70 on this prospective dataset. For AMD, the sensitivity was 86.6%, the specificity of 92.1%, and a kappa score of 0.76.

Conclusions: The AMD and DR screening tools achieved excellent performance operating together to identify two retinal diseases prospectively in mixed datasets, demonstrating the feasibility of such tools in the early diagnosis of eye diseases. These early screening tools will help create an even more comprehensive system capable of being trained on other retinal pathologies, a goal within reach for public health deployment.

Keywords: Diabetic retinopathy; age-related macular degeneration; primary care

Introduction

Age-related macular degeneration (AMD) and diabetic retinopathy (DR) (shown in *Figure 1*) are some of the most common blinding diseases affecting millions of people worldwide (1-4). AMD is the leading cause of vision loss in those over 50 years of age in the developed world (5-7). The number of people with AMD is expected to increase 1.5-fold over ten years due to our aging population, hypertension, and other causes (8). It is often too late to mitigate the complications by the time a person visits an ophthalmologist, as the treatments cannot regenerate vision (9,10). Further, such treatments are expensive, typically costing up to $65,000 for one eye, depending on the drug used, for a 2-year course of treatment (11). While the total (direct and indirect in the USA) cost of AMD is $225 billion per year (5) and is expected to increase, the indirect cost is even greater due to injury, depression, and social dependency resulting from blindness (12).

Diabetic Retinopathy (4) is one of the leading causes of blindness, regardless of age, in the developed world. In the US, the number of patients suffering from DR is expected to reach 6 million by 2020 and 11.3 million by 2030 (13). Early detection of the disease is key to its effective treatment and subsequent reduction of associated economic burdens. The total annual economic burden of eye diseases in the US is about $139B (13).

Our literature review found a number of screening models (14-16) for detecting AMD automatically but without validation from external data or derived from very few test images. Among recent advances in deep learning (DL) is a method proposed by Liu *et al.* (17), which utilized multiple instances of learning to produce a model built with under 5,000 fundus images and an area under the curve (AUC) of 0.79. Several studies focused on automated screening of DR and have achieved varying performance

(18,19). Gulshan *et al.* (20) applied deep learning for DR detection and concluded that further research was needed to bring it to a clinical setting. Abràmoff *et al.* (21) proposed a similar algorithm with 87% sensitivity and 90% specificity. Ting *et al.* (22) proposed and validated a deep learning system built with data from multiethnic populations and compared it with professional human graders. Ting's results showed a sensitivity of 90.5%, and a specificity of 91.6% for detecting referable DR. Gargeya *et al.* (23) proposed a similar deep learning-based model with an AUC of 0.94, a sensitivity of 93%, and a specificity of 87% on a public dataset. However, these models, all built on retrospective datasets, still need validation in real-world primary care clinical settings and a prospective study, which is the paper's motivation.

Telemedicine platforms using cloud-based applications have helped increase the rate of screening for eye diseases, with one study reporting an increase of diabetes-related retinal exams from 37% to 87% (24). Studies have concluded that cloud-based DR screening can identify up to 25% more cases in the diabetic population (25). It has also been shown that telemedicine screening of diseases reduces costs significantly (26). Therefore, we have combined the AMD and DR screening tools on a secure HIPAA compliant telemedicine platform (27,28) to screen patients for the two eye diseases without additional imaging or visits.

This study demonstrates the validity and suitability of the screening system for AMD and DR on prospective data in clinical settings. We present this study in accordance with the STARD reporting checklist (available at http://dx.doi.org/10.21037/aes-20-114).

Methods

A patient's retinal color fundus images are taken in a clinical

Figure 1 Normal vision (A), DR patient vision (B), and vision with AMD (C).
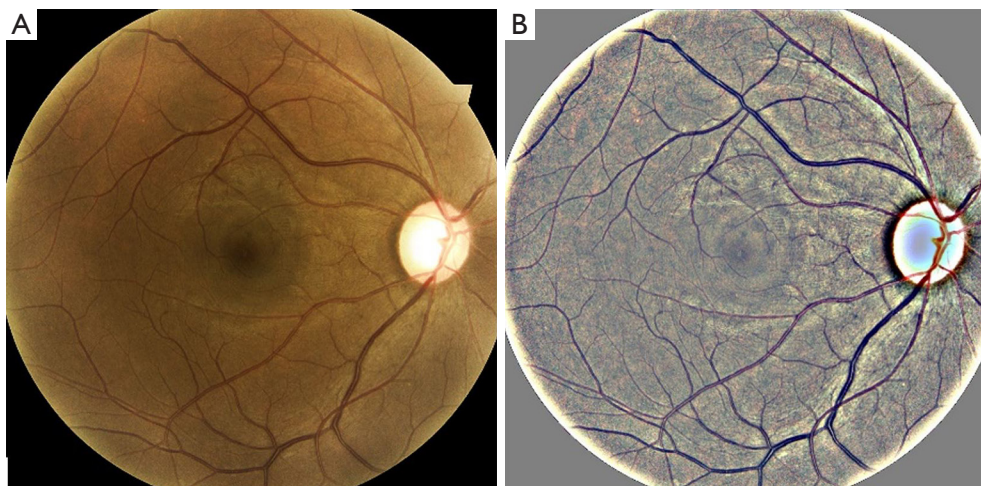
Figure 2 Original RGB fundus image (A), preprocessed image (B), with brightness gradient correction.

setting then uploaded to the iHealthScreen developed cloud-based telemedicine platform by the healthcare worker. If the image is deemed ungradable by the system, it prompts the user to upload a new image. Once the image is accepted, the AI-powered automated AMD and DR screening algorithms perform evaluations and send back reports about the referability with respect to the two diseases individually. Based on this report, the patient would be referred to an ophthalmologist if needed. The method is explained in detail in the following paragraphs. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional review board of Mount Sinai (# IRB-18-00778) and informed consent was taken from all the patients.

Image acquisition and preprocessing

The screening system, built using recent advances in machine learning and artificial intelligence, accepts retinal color fundus images, which can be taken from a wide variety of cameras and imaging conditions. The images used in testing the systems were captured without pharmacological dilation with a Topcon TRC NW6 non-mydriatic fundus camera (Topcon Corp., Tokyo, Japan) with a 45-degree field of view, the DRS camera from CenterVue Inc. and the Eidon camera (Eidon, CenterVue Inc., Fremont, CA) with a 45-degree field of view. Images were preprocessed before training the models to enhance the robustness of the systems. A local color averaging technique was used to eliminate lighting gradients in the fundus image. An

example of such a technique is shown in *Figure 2*. The screening systems referenced later in this paper have performed well under different imaging conditions.

AI-based telemedicine platform

iHealthScreen Inc. developed an artificial intelligence (AI) based telemedicine platform (28) that integrates the server-side programs for image analysis and deep-learning (29) modules intended for screening system, with local remote computer or mobile devices for collecting patient data and images. The images are first checked for gradability automatically by AI developed in-house that achieved over 99% accuracy on 3,000 fundus images. Once the check is passed, the remote devices in primary care will upload images and data to the server for automatic analysis, as shown in *Figure 3*. The telemedicine platform is compatible with both web and mobile platforms. It sends a report to the remote devices with an individual's screening results of the two eye diseases and further recommendations to visit an ophthalmologist. The entire process from data entry to image analysis report is determined to take only a few minutes, depending on the user's experience in handling the equipment, saving time for both the doctor and the patient. The client-side app will call the clinical decision support system to access the data, perform automated screening, and decide if a referral to an ophthalmologist is necessary.

Deep learning (29) is a popular technique that has been recently used for eye disease screening. Deep learning is a class of machine learning techniques that allows systems
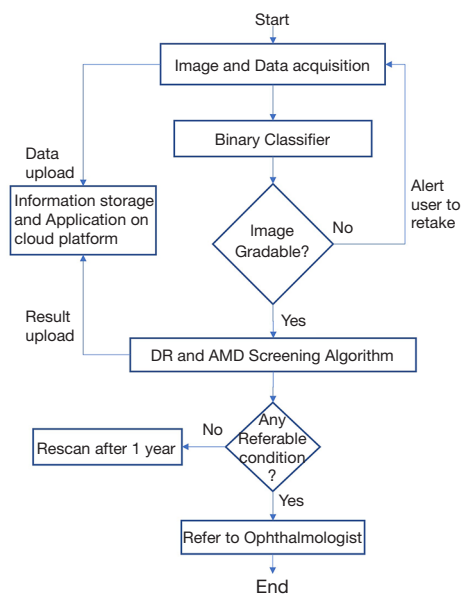
Figure 3 The screening system workflow showing image acquisition, gradability testing, image analysis, referral to an ophthalmologist, and the upload of results to the telemedicine application.

to learn features directly from the data without having to specify any rules or conditions about predictive parameters if there is sufficiently large labeled data input. Deep learning has also been applied in medical applications to detect various diseases such as macular degeneration (30), melanoma (31), and others. Our two models use recent advances in deep learning and artificial intelligence to produce highly accurate classifiers.

AMD screening system

The AMD screening system (32-34) was developed, tested, and validated by iHealthScreen Inc. for identifying referable AMD patients. An ensemble of deep learning screening methods was trained and validated on 116,875 color fundus photos from 4,139 participants in the Age-Related Eye Diseases study (35) to classify them as normal (healthy), early, intermediate, or advanced AMD based on the presence and extent of retinal abnormalities. This study evaluated the system's performance as a binary classifier—referable (intermediate/late) and non-referable (normal/early). For the identification of referable AMD over non-referable, the system achieved 99.2% accuracy, with a sensitivity of 98.9% and a specificity of 99.5%.

DR screening system

The DR screening model (36,37) was developed by iHealthScreen Inc. using deep learning techniques and tested using 88702 images from the Kaggle dataset (38) and externally validated using 1,748 high-resolution fundus images from the Messidor-2 (39) dataset. The images were uploaded in the cloud-based software for testing the automated DR screening platform. The system accepts a fundus image and automatically grades it on a five-point scale—normal, mild, moderate, severe non-proliferative with diabetic macular edema and proliferative DR. An image is considered as referable DR if the grade is moderate or worse. Otherwise, it is considered as non-referable. The automated referable and non-referable DR system evaluation is compared against the expert ophthalmologists' evaluation. The screening system, used on the Kaggle dataset, achieved a sensitivity of 99.2%, a specificity of 97.6%, and an AUC of 0.99 when identifying referable DR. The system was also externally validated in Messidor-2, where it achieved a sensitivity of 97.6%, a specificity of 99.5%, and an AUC of 0.99.

Clinical validation

For AMD and DR, 340 non-dilated subjects aged over 50 years had retinal color imaging of both eyes performed randomly at New York Eye and Ear faculty retina practices with an FDA approved color fundus camera (Eidon, CenterVue Inc., Fremont, CA, USA) between October 2019 and April 2020, and 152 diabetic patients at New York Eye and Ear faculty retina practices, ophthalmic and primary care clinics, yielding 984 images. After exclusion of 308 images with other confounding conditions like myopia and vascular occlusion, and those of poor quality, a total of 676 images were selected and evaluated for AMD and DR. It should be noted that, while in practice, patient referral is based on the worse eye, the models are trained and evaluated on per eye image basis. All images were uploaded to the telemedicine platform and analyzed by the appropriate screening systems.

Three expert graders also classified patients' eyes as referable AMD (intermediate or late AMD) or non-referable (healthy macula or early AMD). Separately, they were also classified as referable DR and non-referable DR. After adjudication of disagreement in the grading to consensus, 172 were referable, and 504 eyes were non-referable AMD. Similarly, 33 were referable, and 643 were

Table 1 Inter-grader reliability calculated using kappa scores against the majority grade (two of three grader agreement)

Graders/eye diseases	Kappa score against majority grade
Graders 1	0.71
Graders 2	0.95
Graders 3	0.88

Table 2 Performance evaluation of the diabetic retinopathy (DR) and age-related macular degeneration (AMD) screening systems on the prospective dataset with accuracy, sensitivity, specificity, and kappa scores measured against adjudicated human grading

Metric	AMD screening	DR screening
Sensitivity	86.6% (80.6% to 91.3%)	97.0% (84.2% to 99.9%)
Specificity	92.1% (89.4% to 94.3%)	96.3% (94.5% to 97.6%)
Accuracy	90.7% (88.2% to 92.8%)	96.3% (94.6% to 97.6%)
Kappa score	0.76 (0.71 to 0.82)	0.70 (0.59 to 0.81)

Table 3 Confusion matrix of the AMD screening system on prospective clinical data—adjudicated human grading *vs.* system grading

AMD grading	Predicted referable	Predicted non-referable
Human graded referable	149	23
Human graded non-referable	40	464

AMD, age-related macular degeneration.

Table 4 Confusion matrix of the DR screening system on prospective clinical data—adjudicated: human grading *vs.* system grading

DR grading	Predicted referable	Predicted non-referable
Human graded referable	32	1
Human graded non-referable	24	619

DR, diabetic retinopathy.

non-referable DR.

Statistical analysis and measures

With a referable case as "positive" and non-referable case as "negative", clinically relevant measures such as sensitivity, specificity, accuracy, and kappa scores were calculated, with

95% confidence intervals. For the two diseases, 2×2 tables were generated to characterize the algorithm's sensitivity and specificity with respect to the reference standard. The reference standard was defined as the majority decision of the experts' grading. The 95% confidence intervals for the sensitivity and specificity of the algorithm were calculated to be exact Clopper-Pearson intervals (40), which corresponded with individual coverage probabilities of sqrt (0.95) ~0.975.

Intergrader reliability was measured among the human graders using kappa scores. The graders' disagreements were adjudicated by taking the majority grade among the graders (two of three grader agreement). The human grader disagreement is an important measure to compare system performance with that of humans. The grades thus obtained are used to measure the performance of our DR and AMD screening system.

Results

To assess disagreement among the three graders, we chose the AMD dataset and kappa score measures. The majority grading, two of three graders agree, was compared to each of the graders, and kappa scores were calculated. Against majority grading, grader 1 had a kappa score of 0.71, while grader 2 scored 0.95, and grader 3 scored 0.88 (see *Table 1*).

Against the majority agreement grading, the AMD screening system (see *Tables 2* and *3*) achieved a sensitivity of 86.6% (80.6% to 91.3%), a specificity of 92.1% (89.4% to 94.3%), an accuracy of 90.7% (88.2% to 92.8%), and a kappa score of 0.76 (0.71 to 0.82) on the prospective dataset. The DR screening system achieved (see *Tables 2* and *4*) a sensitivity of 97.0% (84.2% to 99.9%), a specificity of 96.3% (94.5% to 97.6%), an accuracy of 96.3% (94.6% to 97.6%), and a kappa score of 0.70 (0.59 to 0.81).

Discussion

In this study, we have used two prospective datasets and demonstrated the suitability of an AI-based automated combined screening platform for AMD and DR in clinical settings. The screening platform can fill an unmet need for screening individuals with AMD and DR through a regular primary care doctor's visit. A study from the National Eye Institute showed that half of the patients do not obtain eye examinations recommended by their general physicians. Optometrists and ophthalmologists who screen for DR and AMD are often limited geographically and have limited time. Specialist visits are also time-consuming. In these

instances, automated screening tools within the primary care setting will help mitigate these issues and provide better care for patients with these eye diseases.

We have developed and validated the telemedicine-ready, AI-based, and fully automated screening tools that can be used for screening these diseases with just one image per eye in one visit. Evaluation of the systems in primary care clinics validated the high accuracy of these screening tools, comparable to that of human graders. We note that we used different retinal color fundus cameras (Topcon TRC NW6 non-mydriatic cameras) for developing the model than those used to obtain the prospective data (DRS and Eidon cameras from CenterVue Inc.). Thus, our results are already robust to various types of cameras in the development and test phases and, with larger image sets from each camera, can surely produce even better results.

These AI-based screening tools should be tested prospectively in more diverse clinical settings with cost and time analyses for establishing reliability, consistency, and quantifying individual cost and time benefits. While two eye diseases can be screened simultaneously, the system is limited to only DR and AMD. Other eye diseases can potentially be screened using the same fundus image, such as glaucoma and hypertensive retinopathy. A non-referable grade in either AMD or DR speaks only to its referability for the said diseases. That means other non-AMD and non-DR related pathologies, which the ophthalmologists would have picked up while reviewing the fundus images, can be missed by the more targeted automated grading systems. Making the automated systems more versatile in detecting other pathologies is warranted.

As evident from the disagreements in specialist human graders and their comparisons with the screening systems' kappa scores, it can be concluded that the automated tools perform as well as the human graders, if not better. The performance of the tools is confirmed in the context of real-world image acquisition and analysis. The physical system and the telemedicine software were tested for usability, convenience, and security. The screening systems were deployed on HIPAA-compliant telemedicine platforms and built for minimum interaction with the interface. By using such a secure, fast, reliable, and low-cost system, millions of eyes can potentially be saved from preventable vision-loss, with significant healthcare savings.

Footnote

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional review board of Mount Sinai (# IRB-18-00778) and informed consent was taken from all the patients.

formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. AMD-Defined. Age-Related Macular Degeneration (AMD). Projections for AMD (2010-2030-2050), National Eye Institute (NEI), last accessed on August 3, 2020.
2. Age-Related-Eye-Disease-Study-Research-Group. The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no 1. Control Clin Trials 1999;20(6):573-600.
3. Bartlett H, Eperjesi F. Use of Fundus Imaging in Quantification of Age-related Macular Change. Surv. Ophthalmol 2007;52:655-71.
4. American Academy of Ophthalmology. Diabetic Retinopathy PPP – 2014. Available online: http://one.aao.org/preferred-practice-pattern/diabetic-retinopathy-ppp--2014, last accessed on October 1, 2020.
5. Schmier JK, Jones ML, Halpern MT. The burden of age-related macular degeneration. Pharmacoeconomics 2006;24:319-34.
6. Hassell JB, Lamoureux EL, Keeffe JE. Impact of age related macular degeneration on quality of life. Br J Ophthalmol 2006;90:593-6.
7. Australia CfER. Centrally Focused: The Impact of Age Related Macular Degeneration. Prepared by Access Economics for the Centre for Eye Research Australia, Melbourne 2006.
8. Wong TY, Rogers SL. Age Related Macular Degeneration: time for a randomized controlled trial. Am J Ophthalmol 2007;144:117-9.
9. AMD.Org - Macular Degeneration Partnership. Available online: http://www.amd.org/what-is-macular-degeneration/dry-amd/, last accessed on October 1, 2020.
10. Kanagasingam Y, Bhuiyan A, Abràmoff M, et al. Progress on retinal image analysis for age related macular degeneration. Prog Retin Eye Res 2014;38:20-42.
11. Macular Degeneration Treatment Cost. How Much Does Macular Degeneration Treatment Cost? Available online: https://healthcosthelpercom/macular-degenerationhtml, last accessed on August 10, 2020.
12. Age-Related Eye Disease Study Research Group, SanGiovanni JP, Chew EY, et al. The Relationship of Dietary Carotenoid and Vitamin A, E, and C Intake With Age-Related Macular Degeneration in a Case-Control Study. Arch Ophthalmol 1999;20:573-600.
13. Eye Disease Statistics, National Eye Insititute 2014.

Available online: https://nei.nih.gov/sites/default/files/nei-pdfs/NEI_Eye_Disease_Statistics_Factsheet_2014_V10.pdf, last accessed on October 1, 2020.
14. Zheng Y, Hijazi MHA, Coenen F. Automated "disease/no disease" grading of age-related macular degeneration by an image mining approach. Invest Ophthalmol Vis Sci 2012;53:8310-8.
15. Jayasakthi J, Rajaselvi VM. editors. Detection of Macular Degeneration in Retinal Images Based on Texture Segmentation. Proceedings of the International Conference on Soft Computing Systems. Springer, 2016.
16. Bhuiyan A, Kawasaki R, Sasaki M, et al. Drusen Detection and Quantification for Early identification of age related macular degeneration using color fundus imaging. J Clin Exp Ophthalmol 2013;4:305.
17. Liu H, Wong DW, Fu H, et al. DeepAMD: detect early age-related macular degeneration by applying deep learning in a multiple instance learning framework. Asian Conference on Computer Vision. Springer, 2018.
18. Xiao D, Bhuiyan A, Frost S, et al. Major automatic diabetic retinopathy screening systems and related core algorithms: a review. Machine Vision and Applications 2019;30:423-46.
19. Saha SK, Xiao D, Bhuiyan A, et al. Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: a review. Biomedical Signal Processing and Control 2019;47:288-302.
20. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402-10.
21. Abràmoff MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci 2016;57:5200-6.
22. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA 2017;318:2211-23.
23. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology 2017;124:962-9.
24. Cloud-based imaging, telemedicine platform increases rate of diabetic retinal exams: Healio Primary care optometry news; 2018. Available online: https://www.healio.com/news/optometry/20181206/cloudbased-imaging-telemedicine-platform-increases-rate-of-diabetic-retinal-exams, last accessed on October 1, 2020.

25. Naik S, Wykoff CC, Ou WC, et al. Identification of factors to increase efficacy of telemedicine screening for diabetic retinopathy in endocrinology practices using the Intelligent Retinal Imaging System (IRIS) platform. Diabetes Res Clin Pract 2018;140:265-70.

26. Deschenes S. 5 ways telemedicine is reducing the cost of healthcare July 12, 2012. Available online: https://www.healthcarefinancenews.com/news/5-ways-telemedicine-reducing-cost-healthcare, last accessed on October 1, 2020.

27. Assistance HC. Summary of the hipaa privacy rule. Office for Civil Rights 2003.

28. iHealthScreen. iPredict application 2019. Available online: https://ipredict.health/, last accessed on October 1, 2020.

29. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436.

30. Govindaiah A, Hussain A, Smith R, et al. Deep Convolutional Neural Network-Based Screening and Assessment of Age-Related Macular Degeneration from Fundus Images. In the Proceedings of IEEE International Symposium on Biomedical Imaging 2017;10.1109/ISBI.2018.8363863:1525-8.

31. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115.

32. Govindaiah A, Smith T, Bhuiyan A. A New and Improved Method for Automated Screening of Age-Related Macular Degeneration Using Ensemble Deep Neural Networks. In the proceedings of IEEE EMBC 2018;2018:702-5.

33. Alauddin S, Bhuiyan A, Govindaiah A, et al. A Prospective Trial of an Artificial Intelligence based Telemedicine Platform to Stratify Severity of Age-related Macular Degeneration (AMD). Invest Ophthalmol Vis Sci 2020;61:1843.

34. Bhuiyan A, Wong TY, Ting DSW, et al. Artificial Intelligence to Stratify Severity of Age-Related Macular Degeneration (AMD) and Predict Risk of Progression to Late AMD. Transl Vis Sci Technol 2020;9:25.

35. The Age-Related Eye Disease Study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the Age-Related Eye Disease Study Report Number 6. Am J Ophthalmol 2001;132:668-81.

36. Bhuiyan A, Govindaiah A, Hossain M, et al. Development and Validation of an Automated AI-based Diabetic Retinopathy Screening Tool for Primary Care Setting. Invest Ophthalmol Vis Sci 2020;61:3821.

37. Bhuiyan A, Govindaiah A, Deobhakta A, et al. Development and validation of an automated diabetic retinopathy screening tool for primary care setting. Diabetes Care 2020;43:e147-8.

38. Kaggle Diabetic Retinopathy Detection competition. Available online: https://www.kaggle.com/c/diabetic-retinopathy-detection, last accessed on October 1, 2020.

39. Messidore-2. Kaggle - Adjudicated DR Severity, DME, and Gradability for the MESSIDOR-2 fundus dataset. Available online: https://wwwkagglecom/google-brain/messidor2-dr-grades (last accessed on October 1, 2020).

40. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 1934;26:404-13.