AES ANNALS OF EYE SCIENCE
AN OPEN ACCESS, PEER-REVIEWED JOURNAL COVERING BASIC AND CLINICAL RESEARCH IN OPHTHALMOLOGY

Peer Review File

Article information: https://dx.doi.org/10.21037/aes-21-53

Reviewer A

This prospective study enrolled 265 diabetic patients seen at the endocrinology clinic and obtained Optomap images of 241 patients. They had 2 ophthalmologists review the images and triage the patients for annual comprehensive exam (unclear if this was with optometry), general ophthalmology, or retinal specialists. < 50% of the study patients actually had follow-up eye exam at the study center. The fundus images were also used for Deep Learning.
The manuscript is not very well referenced or written. The authors claim to be the first to use Optos for DR screening and for DL but neither is true. The authors should go into more detail in the introduction and discussion about prior work in this area and how their study is new and different from prior published work.

Comment 1:
Abstract:
Results:
Should specify that < 50% of enrollees were imaged. Also should add % of images with pathology, DR or others.
Reply 1: We thank the reviewer for their thoughtful consideration of our work and helpful comments for improvements. Of the 265 patients who were enrolled, 241 patients were imaged. Of the 241 patients who were imaged, 19% were referred to comprehensive ophthalmology and 27% to a retina specialist for referable retinal pathology.
Changes in the text: We have clarified the following in the abstract: **“**265 patients were enrolled, of which 433 eyes of 241 patients were imaged” and “64 patients (27%) were referred to a retina specialist, and 46 (19%) were referred to comprehensive ophthalmologist for a referable retinal pathology on remote diagnosis.”

Comment 2:
Introduction
Remote diagnosis is part of teleophthalmology. I don’t agree with the authors’ definition of teleophthalmology vs remote diagnosis.
Reply 2: Thank you for this clarifying point. While remote diagnosis is indeed a subset of teleophthalmology, we have tried to distinguish it from other approaches to teleophthalmology in that the imaging does not need to be done in a “brick and mortar” setting, and does not require an expert photographer, and preferably (not in

our case) be done with an inexpensive and easy-to-use unit. To better explain our view of remote diagnosis, we cited our paper published in JAMA Ophthalmology in 2019, that explains the concept in further detail.

Changes in the text: We changed the phrasing of how we distinguish this method from other methods of teleophthalmology in our introduction, "While there are many approaches to teleophthalmology, one type hereby referred to as remote diagnosis, has been proposed as a particularly effective form of screening (17). In contrast to other forms of teleophthalmology screening, remote diagnosis uses imaging devices permanently located at the point of service (e.g., primary care or endocrinology clinics) and operated by non-expert imagers (e.g., office clinical medical assistants, CMAs)." We also made similar changes throughout the manuscript."

Comment 3: Line 75; reference?

The authors do not site reference to several prior work done using Ultrawide field fundus imaging to diagnose Diabetic retinopathy compared to ETDRS and using deep learning.

Reply 3: We have added the relevant citation to lines 82 and 94, including previous studies using Optos images to train DL algorithms to predict DR.

Changes in the text: We have added the relevant citation to lines 82 and 94.

Comment 4:

Methods

Unclear how quickly the Optomap images were reviewed and how quickly participants were referred for eye exam. Was a patient with any level of retinopathy sent to a retinal specialist? Was there any plan to have high risk retinopathy patients be seen sooner?

Reply 4: All images acquired during the week were remotely reviewed by the end of that work week. For our purposes, patients with any level of diabetic retinopathy were sent to a retinal specialist. Those patients of all range of severity were instructed to follow up with a retinal specialist as soon as they can.

Changes in the text: We have added to our methods that the images were graded in a timely manner, by the end of the week they were acquired: "All images were graded by the end of the work week during which they were acquired."

Comment 5:

It appears any image with a good view of the macula was considered acceptable quality. If a significant portion of the peripheral retina is not imaged, can you accurately access level of retinopathy?

Reply 5: We were trying to use all acquired images for the automated image interpretation tool, and in that way, mimic as much as possible the real-world circumstances. One of the primary purposes of the screening for DR is to provide the services to remote rural and underserved populations with automated image capture

and interpretation. In those circumstances, the percentage of sub-optimally taken images or images with limited presence of image biomarkers is expected to be high. That is why we wanted to include all images that could provide us with any information. We fully agree with your comment that we cannot with high accuracy stage diabetic retinopathy without imaging the peripheral retina, and we are aware that that has been the biggest issue with using 30° and 45° field of view. This was our way to try to mitigate this problem.
Changes in the text: No changes made.

Comment 6:
For deep learning dataset, what other pathology was included besides diabetic retinopathy?
Explain in more detail what was exactly done for "flipping, cropping and rotation" to assist in deep learning.
Reply 6: Thank you for this clarifying question. For the deep learning dataset, all images were classified as 1) having referable retinal pathology, or 2) not needing referral for retinal pathology. Thus, we included all the other retinal pathologies in our sample as part of (1) having referrable retinal pathology. Flipping, cropping, and rotation are data augmentation tools commonly used in image-based convolutional neural networks that increase the variation in data. Thus, they allow for more generalizable learning of image features. For example, whether the Optos image was acquired with a slight degree of rotation should not affect how the image is classified. By introducing rotation in a deliberate manner, we thus allow the DL model to learn features that we care about more generally and not fixate on certain characteristics (e.g., degree of rotation) that are not important. The augmentation operations involved: (i) Cropping 0 to 16 pixels from each side of the images randomly; (ii) Flipping the images horizontally or vertically; (iii) Rotating the angles of the images between negative 25 degrees to positive 25 degrees.
Changes in the text: We have clarified these points in our Methods: "and the remaining 164 were labeled as having some retinal pathology, whether DR or other pathologies" and "The augmentation operations involved: 1) Cropping 0 to 16 pixels from each side of the images randomly, 2) Flipping the images horizontally or vertically, and 3) Rotating the angles of the images between negative 25 degrees to positive 25 degrees. The purposes of augmentation were to increase the number of training data to reduce overfitting and improve the generalizability of our model."

Comment 7:
Results:
Lines 277-279: were these incidental findings noted on exam or Optomap or both?
Reply 7: These incidental findings were those noted on Optomap remote diagnosis. The incidental findings found on Optomap relative to clinical exam are shown in Table 4.

Changes in the text: No changes made.

Comment 8:
Lines 256-9: Did the severity of diabetic retinopathy noted on Optomap correlate with clinical exam? Table 4 lists missed diagnosis. For example, was PDR missed as NPDR?
Reply 8: While most of the cases of DR correlated in severity between Optomap and exam, there were some cases as the reviewer notes in which a diagnosis was missed or graded as less severe than on exam as reported in Table 4. For example, 1 case of PDR was indeed missed as NPDR.
Changes in the text: No changes made.

Comment 9:
Only < 50% of imaged patients were seen for eye exam—what was the timeline for getting an exam after remote imaging?
Reply 9: Patients were instructed to follow-up with the appropriate provider, whether a retinal specialist or comprehensive ophthalmologist as soon as they could. While all patients agreed as such at study enrollment, not all followed up for standard clinical exam. We did not and could not enforce a strict timeline for getting an exam following remote imaging. Regardless, the reviewer makes a very important point, and we will be conducting a follow-up study, involving all of our screening programs on patients' compliance (lost to follow), clinical (visual outcomes), and cost-effectiveness in the near future.
Changes in the text: No changes made.

Comment 10:
Line 270-2: Is the sensitivity and specificity for DR or any retinal pathology?
Reply 10: For our DL model, the sensitivity and specificity reported were for any referable retinal pathology.
Changes in the text: We further clarified this point in our results: "The accuracy of our DL model for classifying referable retinal pathology was…"

Comment 11:
Discussion:
Lines 287-290: I don't think this paper is the first to use DL and Optos for DR detection. Please site prior published studies in this topic and specify how this study defers from prior studies and why the difference is important.
Reply 11: While there have been other work using deep learning on Optos images for DR detection and staging, there have not been any studies that applied this approach to images acquired in a prospective manner to detect *referable retinal pathologies*, in addition to DR. We believe that this distinction is important given the context of this approach's implementation. Given that detecting any referable retinal pathology in a

busy endocrinology clinic would be important at an institution such as ours, we intentionally used the DL model to detect those pathologies that our retinal specialists would refer to be seen in clinic given their practice patterns.
Changes in the text: We have adjusted the language on this issue at several points throughout the manuscript, including the following: "However, it is important to emphasize that, while previous studies have developed DL models to detect or stage DR and other retinal pathologies, this is the first computational model for automated analysis for identification of any referable pathology using Optomap images acquired in a prospective manner at an endocrinology clinic."

Comment 12:
Lines 293-296: I don't think this is always the case for telemedicine.
Reply 12: We have softened the relevant language since the statement in question applies to varying extents for the many different types of telemedicine.
Changes in the text: We have modified the lines in question: **"**In some forms of teleophthalmology screening, color fundus photographs (CFP) are taken by expert photographers in mobile vans or dedicated brick-and-mortar imaging centers, then transmitted to grading experts. These approaches suffer from high costs, requiring dedicated vehicles or office space, as well as photographers and graders."

Comment 13:
Line 298-9: not true. One can get good quality posterior pole images without dilation and minimal training of photographer but limited view of periphery without dilation.
Reply 13: The reviewer is correct that a limited view of the fundus can be achieved without dilation. We have changed the language in our manuscript to reflect this statement. The novelty of our study was to validate a semi-novel unit at the point of service (e.g., endocrinology clinics), where patients are screened by clinical medical assistants (without any experience in ophthalmology or retinal image capture) on their way out without disturbing the clinic flow. We had a great patient capture and pretty good sensitivity of the automated image interpretation tool.
Changes in the text: We have changed the language in our manuscript to reflect your statement.

Comment 14:
Lines 329-31; Not first study of this type. Please site prior work in this area.
Reply 14: We have deleted the line in question and added more citations of prior work in this area in the introduction and discussion.
Changes in the text: Deleted: **"**To the best of our knowledge, this is the first DL-based approach to identify referable retinal pathology using Optomap" from the lines in question. We have also added citations of prior work in the Introduction as appropriate.

Reviewer B

The manuscript titled Evaluation of a deep learning supported remote diagnosis model for identification of diabetic retinopathy using wide-field Optomap successfully trained a deep learning model to detect DR (yes/no) using Optomap images. The manuscript reads well, and I recommend acceptance with minor revision.

Comment 15:
I am unable to assess the novelty of this work. The techniques to capture images are outside my area of expertise. After a quick search I found a rather similar idea (https://doi.org/10.1155/2021/6651175). It would be great if the authors can explain the differences between the two.

Maybe other forms of DR detection could be mentioned. E.g.
• https://www.nature.com/articles/s41598-021-89225-0
• https://ieeexplore.ieee.org/document/9103111
Reply 15:
We thank the reviewer for their careful review of our manuscript and helpful comments for improvement. As the reviewer mentions, the paper by Nagasawa et al., 2021 also examines a similar topic, however, there are a few key differences. Unlike many other studies including Nagasawa 2021, the images included in our cohort were collected in a prospectively manner of consecutive patients seen for diabetes at an endocrinology clinic. In addition, while Nagasawa 2021 trained their deep learning algorithm to predict staging of diabetic retinopathy, we sought to predict which eyes would be referable for further management by a retinal specialist. As such, while our approach would also have detected referable retinal pathologies other than diabetic retinopathy as well.

Changes to the manuscript: We have made changes throughout the text, accordingly.

Comment 16:
Figure 1 says “241 images with 433 images uploaded”. I believe to this to be a mistake and the first “images” should be “patients” instead.
Reply 16: Thank you for this correction. We have updated the figure accordingly.
Changes to the manuscript: Changed the relevant caption on that figure to “241 patients with 433 images uploaded”.

Comment 17:
Which are the parameters used for the data augmentation?

Reply 17: For data augmentation, we introduced randomly cropping 0 to 16 pixels from each side of the images, flipping the images horizontally or vertically, and rotating the angles of the images between negative 25 degrees to positive 25 degrees.

Changes to the manuscript:

We have added the following clarification to our Methods: "The augmentation operations involved: 1) Cropping 0 to 16 pixels from each side of the images randomly, 2) Flipping the images horizontally or vertically, and 3) Rotating images between negative 25 degrees to positive 25 degrees. The purposes of augmentation were to increase the number of training data to reduce overfitting and improve the generalizability of our model."

Comment 18:

Which ResNet schema the authors used? Was the model pre-trained? Did the authors use transfer learning?

Reply 19: We built the model totally from scratch since the pre-trained models for transfer learning are pre-trained on public datasets, which are substantially different from our data. We followed the classical ResNet architecture to build a small model for our limited number of images for training. We did not use transfer learning since the standard well-established models are pre-trained on public datasets like ImageNet, which contains substantially different images from our UWF Optomap images. Besides, since our large databases of ultrawidefield images are not as readily available (compared to simple fundus photos, for example), there were not enough images to pre-train the model. Thus, we followed the classical architecture of ResNet to build a small model and additionally added soft-thresholding techniques to reduce the noise.

Changes to the manuscript: We have added the following to our manuscript: "A DL model was built from scratch using Keras to identify the presence or absence of retinal pathologies on Optomap images, using a convolutional neural network (CNN) model built on top of ResNet's architecture" and "We did not use transfer learning since the standard well-established models are pre-trained on public datasets like ImageNet, which contains substantially different images from our UWF Optomap images. Besides, since our large databases of ultrawide field images are not as readily available (compared to simple fundus photos, for example), there were not enough images to pre-train the model. Thus, we followed the classical architecture of ResNet to build a small model and additionally added soft-thresholding techniques to reduce the noise."

Comment 19:

It would be more clarifying if in Figure 2C the authors could highlight the lesions.

Reply 19: As below.

Changes to the manuscript: Unfortunately, we were unable to successfully make this change. When we blow up the figure, besides labeling the visible laser scars, the

image gets too busy in the natural size, and almost all information is lost upon labeling the posterior pole features.

Comment 20:
Is the code and dataset will be available and to which capacity?
Reply 20: Yes, the github repository is cited in our references and can be found at: https://github.com/MingzheHu-Duke/Retinopathy-Detection-from-UWF-Retinal-Images.git
Changes to the manuscript: No changes made.

Comment 21:
The authors failed to disclose whether the individuals in the study had other eye lesions or underwent eye surgery prior to the study.
Reply 21: An important goal of our study was to prospectively enroll any diabetic patients seen in one of the endocrinology clinics. As such we did not have any criteria for excluding patients from being imaged remotely or specifically noting any prior ocular history or surgery. While prior ocular history and surgery are certainly important considerations in the decision to refer a patient for follow-up care, these patients would likely already be established with an eye provider and certainly would've had detailed eye exams for the previous surgery. In addition, our study using remote imaging data ONLY to decide referral, without any consideration of prior medical, ocular, or surgical history. As such, in evaluating our remote diagnosis approach, we wanted to include all the patients that normally attend these clinics (and would be imaged by such a remote diagnosis program).
Changes to the manuscript: No changes made.