BMJ Open Ophthalmology

Diagnostic accuracy of teleretinal screening for detection of diabetic retinopathy and age-related macular degeneration: a systematic review and meta-analysis

Parsa Mehraban Far [1], Felicia Tai,[2] Adeteju Ogunbameru,[3,4] Petros Pechlivanoglou,[4,5] Beate Sander,[3,4] David T Wong,[2,6] Michael H Brent,[2,7] Tina Felfeli [2,3,4]

Correspondence to
Dr Tina Felfeli; tina.felfeli@mail.utoronto.ca

ABSTRACT

Objective To evaluate the diagnostic accuracy of teleretinal screening compared with face-to-face examination for detection of diabetic retinopathy (DR) and age-related macular degeneration (AMD).

Methods and analysis This study adhered to the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA). A comprehensive search of OVID MEDLINE, EMBASE and Cochrane CENTRAL was performed from January 2010 to July 2021. QUADAS-2 tool was used to assess methodological quality and applicability of the studies. A bivariate random effects model was used to perform the meta-analysis. Referrable DR was defined as any disease severity equal to or worse than moderate non-proliferative DR or diabetic macular oedema (DMO).

Results 28 articles were included. Teleretinal screening achieved a sensitivity of 0.91 (95% CI: 0.82 to 0.96) and specificity of 0.88 (0.74 to 0.95) for any DR (13 studies, n=7207, Grading of Recommendations, Assessment, Development and Evaluation (GRADE) low). Accuracy for referrable DR (10 studies, n=6373, GRADE moderate) was lower with a sensitivity of 0.88 (0.81 to 0.93) and specificity of 0.86 (0.79 to 0.90). After exclusion of ungradable images, the specificity for referrable DR increased to 0.95 (0.90 to 0.98), while the sensitivity remained nearly unchanged at 0.85 (0.76 to 0.91). Teleretinal screening achieved a sensitivity of 0.71 (0.49 to 0.86) and specificity of 0.88 (0.85 to 0.90) for detection of AMD (three studies, n=697, GRADE low).

Conclusion Teleretinal screening is highly accurate for detecting any DR and DR warranting referral. Data for AMD screening is promising but warrants further investigation.

PROSPERO registration number CRD42020191994.

Key messages

What is already known about this subject?
► With continuous advances in telecommunication technology and ophthalmic imaging in the last decade, teleretinal imaging is being relied on to identify patients with sight-threatening diabetic retinopathy (DR) and age-related macular degeneration (AMD).

What are the new findings?
► This meta-analysis of diagnostic test accuracy validates teleretinal screening as a highly accurate modality for diagnosis of any or referrable DR, while evidence for diagnosis of AMD is more limited.

How might these results change the focus of research or clinical practice?
► More research is required to determine the diagnostic accuracy of teleretinal screening for diagnosis of AMD.
► The role of teleretinal screening relying on artificial intelligence should be examined in future research.

diagnosis and monitoring is essential for the prevention of irreversible vision loss.[3] Traditional office-based face-to-face examination is effective for screening patients, but there are associated challenges in regions with limited accessibility to resources and eyecare specialists.[4] Over the past decade with technological improvements, teleretinal screening has been explored as a cost-effective strategy to meet the increasing needs of the population worldwide.[5 6]

To date, a few systematic reviews have assessed the accuracy of teleretinal screening using human graders; however, there has been a limited number of meta-analyses with the use of correct hierarchical models to quantitatively summarise these results.[7 8] Hierarchical methods are recommended as

INTRODUCTION

Diabetic retinopathy (DR) and age-related macular degeneration (AMD) are among the leading causes of vision impairment in both low-income and high-income countries.[1 2] Despite the significant advancements in therapeutics for both disorders, timely

they account for the inherent correlation between sensitivity and specificity and allow for a greater degree of heterogeneity in the results. Moreover, to our knowledge, no meta-analysis to date has specifically assessed the diagnostic accuracy of teleretinal imaging by human graders to identify referral-warranted cases.[9] Furthermore, current estimates available from review papers exclude ungradable images from the analysis, which could falsely over-estimate the accuracy of teleretinal screening programmes.[7 8 10] Herein, we present a systematic review and meta-analysis on the diagnostic accuracy of teleretinal screening for detection of AMD and DR that specifically addresses these deficiencies.

METHODS
The primary objective of this review was to assess the accuracy (sensitivity and specificity) of teleretinal screening for detection of DR and AMD compared with face-to-face clinical examination as a real-world reference standard. DR and AMD were chosen as the target conditions given that they account for the majority of cases in retina practices and have overlapping pathogenesis.[11 12] The secondary objective was to formally assess the influence of exclusion of ungradable images in diagnostic accuracy calculations. This study adhered to the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA).[13]

Search strategy
A comprehensive search of the literature was conducted on CENTRAL, Ovid MEDLINE and OVID Embase databases from 1 January 2010 to 25 July 2021. Detailed search strategy is included in online supplemental table 1. The time restriction for the search was placed to capture the most recent technological advances in teleretinal imaging modalities. No further restrictions were placed based on location, type and language of the publications. Retrieved studies were imported into Covidence (Melbourne, Australia), where duplicates were removed and article screening was performed.

Disease definition
Referrable DR was defined as any severity equivalent or worse than moderate non-proliferative DR or diabetic macular oedema (DMO).[14] DMO was defined as any retinal thickening or the presence of hard exudate in the macula.[14] Referrable AMD was defined as disease with features suggestive of intermediate and advanced state such as extensive intermediate drusen (<125 μm), any large drusen (>125 μm), neovascularisation or geographic atrophy.[15]

Reference standard
The reference standard for determining diagnostic accuracy of teleretinal screening was chosen to be face-to-face clinical examination as opposed to the seven-field Early Treatment of Diabetic Retinopathy Study (ETDRS protocol. The selection of this reference standard was based on three important justifications. First, face-to-face clinical examination is an established real-world modality for diagnosis and monitoring of patients with DR.[16–18] Second, in-person examination is in keeping with the clinical pathway of most teleretinal screening programmes where selected patients with high-risk features or ungradable patients are ultimately referred for office-based examination.[9] Lastly, face-to-face examination is a reputable consistent reference standard for all of the diagnoses included in the review including AMD, whereas the evidence supporting the ETDRS protocol stems from DR literature.[17 19 20]

Study selection criteria
Two reviewers (PMF and FT) independently assessed eligibility first by the title and abstracts and then the full text of the retrieved studies. Conflicts not resolved were discussed with a third reviewer (TF) to reach consensus. Eligible studies included all comparative studies assessing the diagnostic accuracy using any form of retinal imaging modality for DR, DMO and AMD, where the reference standard was face-to-face examination using dilated funduscopy using direct/indirect ophthalmoscopy or slit lamp biomicroscopy. Retinal imaging modalities were defined broadly to include any form of fundus imaging device including handheld and table-top instruments regardless of the quality, field of view and wavelength of light used to capture the image.

Studies were excluded if they met any of the following exclusion criteria: (1) no reference standard of face-to-face ophthalmic examination, (2) no full-text available or insufficient information to allow for independent calculation of sensitivity and specificity and (3) grading of images not performed by human graders with specialty in ocular health (optometry/ophthalmology/special training attendee).

Data extraction and risk of bias assessment
Two independent reviewers (PMF, FT) were responsible for the extraction of relevant data from all included studies and assessing their methodological quality. Extracted data included: authors; year of publication; country or countries where the study was done; imaging devices used; imaging protocol and credential of image graders. For determination of diagnostic test accuracy, a two-by-two table was generated and the values corresponding to true positive, false positive, true negative and false negative were extracted for each calculation. In cases where information was insufficient, the data were requested via email from the corresponding authors of the publications. The QUADAS-2 tool was used independently by two reviewers (PMF, FT) to assess the methodological validity and applicability of each included study.[21]

Data synthesis
The unit of analysis was number of eyes; however, number of patients was also accepted for analysis if

number of eyes screened was not reported. In cases where data were reported as patients only, each patient was counted as one unit of analysis. A sensitivity analysis was planned to assess the influence of using patients or eyes as the unit of analysis in the final diagnostic accuracy calculations. The random effects bivariate binomial model in R using the lme4 package was used to perform a meta-analysis and generate sensitivity, specificity, likelihood ratios (LR) and diagnostic OR associated with each test.[22 23] The random effects bivariate binomial model was selected as it is a hierarchical method and is suitable for deriving summary sensitivity/specificity data at a specific threshold.[24] Further details with regard to statistical analysis can be found in online supplemental appendix 1.

For the primary analysis, ungradable images were classified as having the target condition being assessed to simulate real-life patient care in a screening programme where undetermined results are further analysed and assumed positive. A sensitivity analysis was planned in advance, where ungradable images were excluded from analysis and diagnostic accuracy was calculated for only gradable images. A p value lower than 0.05 was used as the threshold for statistical significance. Summary estimates from statistical analyses were presented with their respective 95% CIs along with p values where applicable.

Quality of evidence

Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach was used to assess the quality of evidence.[25] Summary of findings tables along with GRADE evidence profiles were generated using GRADEpro software.[26]

RESULTS

A total of 28 articles met the inclusion criteria for qualitative and quantitative synthesis. PRISMA flowchart of study selection process is depicted in figure 1. The most prevalent geographical location where the studies were conducted was the USA accounting for 36% (10/28), followed by Europe 25% (7/28). For studies reporting on patients with DR, the mean number of years from initial diagnosis of diabetes was greater than or equal to 5 years for all studies at the time of publication where this was reported. The pooled cohort included both type I and II diabetes; however, the relative proportion of each type was not consistently presented in the included studies. All studies, except for 11% (3/28) used table-top fundus cameras for screening. The reference standard of choice was dilated fundus examination with slit-lamp biomicroscopy or binocular indirect ophthalmoscopy for 82% (23/28) of studies. Descriptive details of the eligible studies are provided in online supplemental table 2.
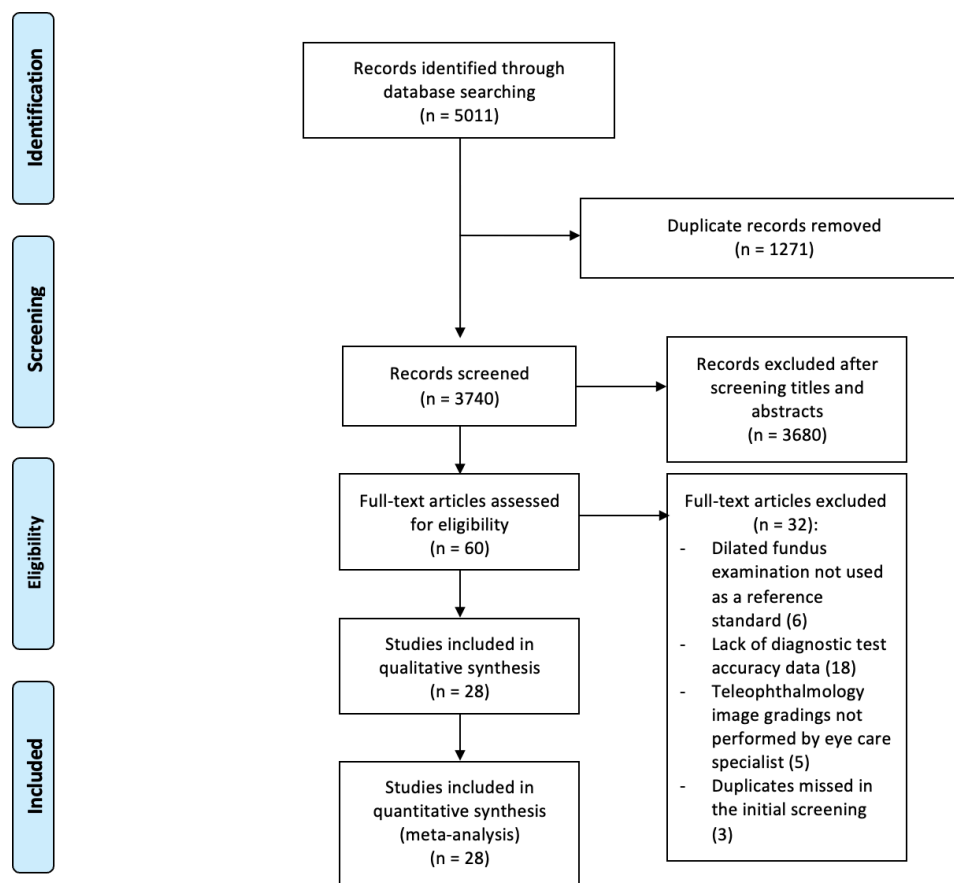

Figure 1 Preferred Reporting Items for a Systematic Review and Meta-analysis flow diagram for transparent reporting of study selection process and meta-analysis.
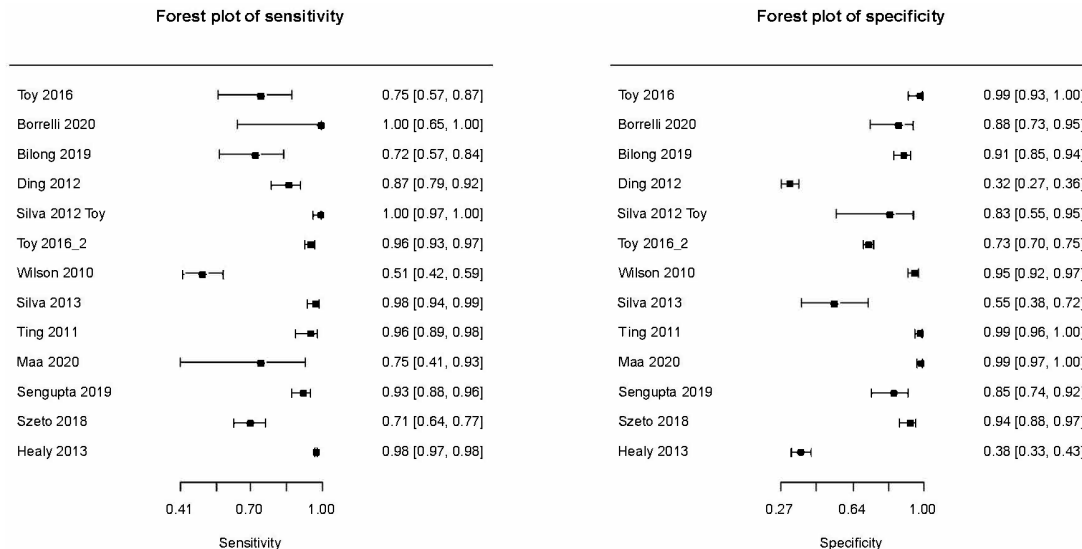
Figure 2 Forest plots depicting sensitivity and specificity of teleretinal screening for detection of any level of diabetic retinopathy.

Additional information such as diagnostic criteria for referrable disease in each study as well as the number and process of grading are presented in online supplemental tables 3 and 4.

Diagnostic accuracy for detectable DR

Thirteen studies including 7207 eyes contributed to the meta-analysis (figure 2). Teleretinal screening was found to have a specificity of 0.88 (95% CI: 0.74 to 0.95) and sensitivity of 0.91 (95% CI: 0.82 to 0.96) for detection of DR. Diagnostic OR, LR+ and LR− associated with the accuracy of fundus imaging were calculated to be 77.59 (95% CI: 29.88 to 201.50), 7.78 (95% CI: 3.43 to 17.65) and 0.1 (95% CI: 0.05 to 0.202), respectively. There was a large degree of heterogeneity in the results from individual studies corresponding to a large predictive region in the summary receiver operating characteristic (sROC) curve (online supplemental figure 1).

Investigation of heterogeneity through meta-regression was only possible for mydriasis due to paucity of data for other covariates. Both sensitivity and specificity were higher in dilated eyes in comparison to undilated eyes with a sensitivity of 0.91 (95% CI: 0.82 to 0.95) versus 0.89 (95% CI: 0.66 to 0.97) and specificity of 0.89 (95% CI: 0.63 to 0.97) versus 0.85 (95% CI: 0.71 to 0.93), respectively. Addition of pupil status as a covariate for both sensitivity and specificity to the model did not achieve statistical significance ($\chi^2(5)$=10.31, p=0.07).

Diagnostic accuracy for referrable DR

Ten studies including 6373 eyes contributed to the meta-analysis (online supplemental figure 2). Diagnostic accuracy was lower than that of any DR, with a sensitivity of 0.88 (95% CI: 0.81 to 0.93) and specificity of 0.86 (95% CI: 0.79 to 0.90) for detection of referrable DR. Diagnostic OR, LR+ and LR− were calculated to be 45.27 (95% CI: 28.79 to 71.19), 6.16 (95% CI: 4.37 to 8.67) and 0.14 (95% CI: 0.09 to 0.22), respectively. No formal

meta-regression could be performed due to paucity of data. There was a mild-moderate degree of heterogeneity as depicted in the sROC curve in online supplemental figure 3.

Diagnostic accuracy for AMD

Due to paucity of data, the bivariate binomial model could not be fitted for analysis. A univariate random effects model was fitted by removing the correlation term between sensitivity and specificity. The raw diagnostic accuracy values for each individual study are provided in online supplemental table 5. Three studies including 697 eyes contributed to the univariate random effects meta-analysis for detection of any AMD with an overall sensitivity of 0.71 (95% CI: 0.49 to 0.86) and specificity of 0.88 (95% CI: 0.85 to 0.90).

Diagnostic accuracy for referrable AMD

No meta-analysis for referrable AMD was performed due to significant interstudy variability in the definition of referrable AMD.

Use of ancillary imaging

Only one study assessed the influence of addition of optical coherence tomography (OCT) to a teleretinal screening programme.[17] The authors concluded that OCT did not improve the diagnostic accuracy of their teleophthalmology programme for the detection of glaucoma, DR and AMD.

Methodological quality

Online supplemental figure 4 depicts a graphical representation of summary quality and applicability to our research question of the included studies. Selection bias was noted based on exclusion of some data for patients with corneal disorders, media opacity and ungradable images from the analysis in some studies. Online
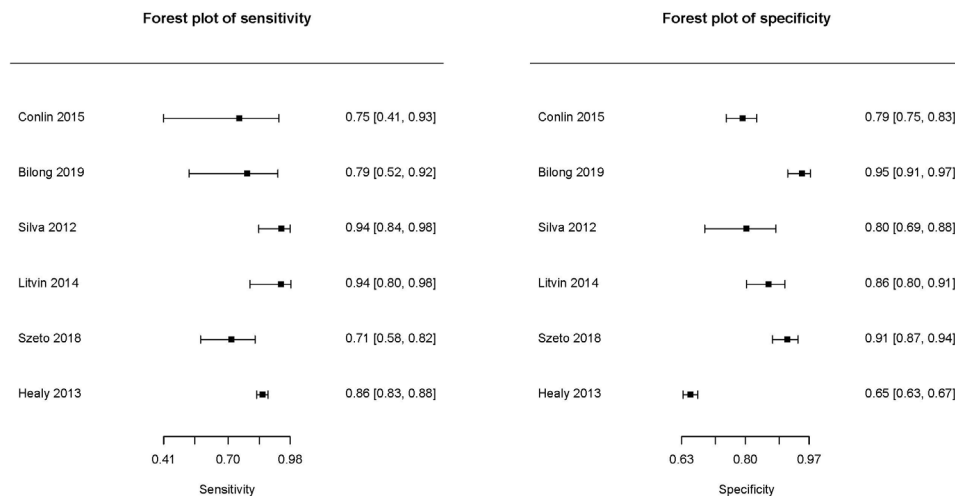
Forest plot of sensitivity

Conlin 2015		0.75 [0.41, 0.93]
Bilong 2019		0.79 [0.52, 0.92]
Silva 2012		0.94 [0.84, 0.98]
Litvin 2014		0.94 [0.80, 0.98]
Szeto 2018		0.71 [0.58, 0.82]
Healy 2013		0.86 [0.83, 0.88]

0.41   0.70   0.98

Sensitivity

Forest plot of specificity

Conlin 2015		0.79 [0.75, 0.83]
Bilong 2019		0.95 [0.91, 0.97]
Silva 2012		0.80 [0.69, 0.88]
Litvin 2014		0.86 [0.80, 0.91]
Szeto 2018		0.91 [0.87, 0.94]
Healy 2013		0.65 [0.63, 0.67]

0.63   0.80   0.97

Specificity

Figure 3 Forest plots depicting sensitivity and specificity of teleretinal screening for detection of diabetic macular oedema.

supplemental figure 5 depicts the detailed assessment of quality in each individual study.

Strength of evidence

Overall, quality of body of evidence was low to moderate for the reported outcomes. Summary of findings tables and GRADE profiles are demonstrated in online supplemental table 6.

Sensitivity analyses

Diagnostic accuracy for DMO

A sensitivity analysis was performed to quantify the accuracy of teleretinal screening for identifying DMO. Six studies including 4255 eyes contributed to the meta-analysis (figure 3). teleretinal screening maintained a favourable diagnostic accuracy, but lower than that of any DR, with a sensitivity of 0.84 (95% CI: 0.76 to 0.90) and specificity of 0.85 (95% CI: 0.75 to 0.91) for detection of DMO. Diagnostic OR, LR+ and LR– were calculated to be 30.57 (95% CI: 15.20 to 61.48), 5.61 (95% CI: 3.39 to 9.28) and 0.18 (95% CI: 0.12 to 0.29), respectively. No formal meta-regression could be performed due to the limited number of studies contributing to this analysis. There was a mild-moderate degree of heterogeneity with a relatively small predictive region in sROC curve (online supplemental figure 6).

Exclusion of ungradable images

Twelve studies including 8452 eyes contributed to the meta-analysis for detection of any level of DR (online supplemental figure 7). After the exclusion of ungradable images, diagnostic accuracy for detection of any levels DR remained high with a sensitivity of 0.88 (95% CI: 0.75 to 0.94) and specificity of 0.90 (95% CI: 0.81 to 0.96). A total of 13 studies, including 6481 eyes, were analysed for meta-analysis of referrable DR (online supplemental figure 8). After exclusion of ungradable cases, the specificity increased to 0.95 (95% CI: 0.90 to 0.98), while the sensitivity remained nearly unchanged at 0.85 (95% CI: 0.76 to 0.91).

Patients and eyes as unit of analysis

Separate meta-analysis in studies which used patients versus eyes as unit of analysis for detection of any level of DR was performed. Meta-analysis using eyes as unit of analysis for detection of any level of DR was performed using data from eight studies including 4299 eyes. Teleretinal screening had a sensitivity of 0.95 (95% CI: 0.87 to 0.98) and specificity of 0.88 (95% CI: 0.70 to 0.96). Using patients as the unit of analysis based on five studies and 2908 patients, teleretinal screening achieved a sensitivity of 0.80 (95% CI: 0.59 to 0.91) and specificity of 0.88 (95% CI: 0.60 to 0.98).

DISCUSSION

Here, we presented our findings from a systematic review and meta-analysis on estimates of sensitivity and specificity of teleretinal screening for detection of DR and AMD when compared with face-to-face clinical examination as the real-world reference standard.[17 18] We found that teleretinal screening achieved a high accuracy for detection of any DR with a sensitivity of 0.91 (95% CI: 0.82 to 0.96) and specificity of 0.88 (95% CI: 0.74 to 0.95) and referrable DR with a sensitivity of 0.88 (95% CI: 0.81 to 0.93) and specificity of 0.86 (95% CI: 0.79 to 0.90).

Our results are in keeping with a recent review, which showed that teleretinal screening can achieve very high accuracy with a sensitivity of 0.84 (95% CI: 0.79 to 0.88) and specificity of 0.95 (95% CI: 0.94 to 0.96) for detection of any level of DR.[7 27] Based on our findings, data on AMD are limited but the diagnostic accuracy was calculated to be lower with a sensitivity of 0.71 (95% CI: 0.49 to 0.86) and specificity of 0.88 (95% CI: 0.85 to 0.90). The diagnostic accuracy teleretinal screening has been previously characterised at specific levels of DR severity; however, data on the overall accuracy for detection of referrable cases have not been consistently reported. Another review estimated a sensitivity of 0.76 (0.69 to 0.82) and specificity of 0.95 (0.93 to 0.96) for detection of DMO.[8] In our study, we also noted a lower level

of accuracy for detection of DMO and referrable DR in comparison to any DR.

Given that the previous reviews to date on this topic have typically excluded ungradable images from their analysis, our sensitivity analysis with the exclusion of ungradable images showcases a cautionary message for future investigators. In fact, the specificity of fundus imaging for identification of referral-warranted DR improved by nearly 10% after ungradable images were removed from analysis. This observation is expected and can be explained by spectrum effect, whereby systematic removal of a patient subgroup, such as difficult to diagnose cases with media or corneal opacity, leads to an easier diagnosis and detection of referrable and non-referrable cases.[28]

Based on our findings, evidence in support of implementation of teleretinal screening for detection of AMD was limited in comparison to DR. Only three studies provided diagnostic accuracy data for detection of any AMD.[17 19 29] Although these results are encouraging, with an overall sensitivity of 0.71, more research is required to establish a role for fundus imaging for diagnosis and treatment of patients with AMD. One strategy to generate more diagnostic accuracy data for AMD detection is to implement AMD detection into the already existing teleretinal screening infrastructure for DR. If teleretinal screening proves to be a highly accurate tool within the structure of the pre-existing teleophthalmology programmes, further emphasis may be placed on detection of AMD.

The role of OCT for diagnosis and monitoring of retinal disorders is well established.[30] However, whether its incorporation into teleophthalmology screening programmes is beneficial remains controversial.[17] Only one study in this systematic review provided a direct comparison in the diagnostic accuracy of fundus photography combined with OCT and fundus photography alone.[17] In the paper by Maa *et al*, despite the detailed cross-sectional analysis of the macula, optic nerve head and retinal nerve fibre layer that is provided in OCT scans, the authors did not detect any improvement in sensitivity or specificity of fundus photography for the detection of glaucoma and retinal disorders with the addition of OCT.[17] In contrast to these results, other groups have clearly demonstrated a role for OCT in addition to fundus photographs, especially for the detection of diabetic macula oedema which requires a stereoscopic view.[31–35] In the current meta-analysis, only a fraction of the studies used OCT alone or as an adjunct modality in addition to fundus photographs and we were unable to perform a formal meta-regression. Due to widespread use of OCT as well as recent advances in OCT technology such as swept-source OCT and OCT angiography, it is inevitable that more data will become available in the near future.[36]

Meta-regression based on pupil status showed a sizeable improvement in diagnostic accuracy when eyes were dilated prior to capturing of the image which approached statistical significance. Although we are unable to identify the exact reason behind this observation, it can be hypothesised that a larger pupil diameter allows for increased capture of light by the camera and therefore generates a higher quality image.[7] This finding should be verified in different ethnic groups as well as in individuals with difference in iris colour which could elicit different levels of response to pharmacological dilation.[37]

Limitations and future directions

It is important to note that there is a large degree of heterogeneity in the diagnostic criteria for referrable DR and AMD. Additionally, there is also a large degree of heterogeneity in the sample including patients with type I and II diabetes of differing durations. Similar to all review papers, publication bias may be present whereby studies that achieve high diagnostic accuracy are preferentially published in comparison to studies where the accuracy is lower which could lead to an overestimation of sensitivity and specificity. Our study results are only applicable to teleretinal programmes using human graders. Recent diagnostic test accuracy meta-analyses have provided very promising accuracy estimates for machine-learning-based teleretinal screening programmes for DR.[38 39] Future studies should assess the diagnostic accuracy of automated systems using artificial intelligence and deep-learning algorithms in teleophthalmology screening programmes for ocular diseases.[40] Lastly, the focus of this review was on teleretinal screening for the most common retinal pathologies. As more data become available, future investigations should assess the utility of teleglaucoma screening programmes.[41]

Author affiliations
[1]Department of Ophthalmology, Queen's University, Kingston, Ontario, Canada
[2]Department of Ophthalmology and Vision Sciences, University of Toronto, Toronto, Ontario, Canada
[3]Toronto Health Economics and Technology Assessment (THETA) Collaborative, University Health Network, Toronto, Ontario, Canada
[4]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[5]Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
[6]Department of Ophthalmology, St. Michael's Hospital, Toronto Unity Health, Toronto, Toronto, Ontario, Canada
[7]Retina Service, Donald K Johnson Eye Institute, University Health Network, Toronto, Ontario, Canada

ORCID iDs
Parsa Mehraban Far http://orcid.org/0000-0003-4543-4155
Tina Felfeli http://orcid.org/0000-0002-0927-3086

REFERENCES

1 Centers for Disease Control and Prevention. Common eye disorders and diseases. secondary common eye disorders and diseases, 2020. Available: https://www.cdc.gov/visionhealth/basics/ced/index.html

2 World Health Organization. Blindness and vision impairment. Secondary blindness and vision impairment.

3 Flaxman SR, Bourne RRA, Resnikoff S, *et al*. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. *Lancet Glob Health* 2017;5:e1221–34.

4 Wong TY, Sabanayagam C. The war on diabetic retinopathy: where are we now? *Asia Pac J Ophthalmol* 2019;8:448–56.

5 Rathi S, Tsui E, Mehta N, *et al*. The current state of Teleophthalmology in the United States. *Ophthalmology* 2017;124:1729–34.

6 Yoon S, Kim HYS, Kim J, *et al*. A current status of teleophthalmology in low- and middle-income countries: literature review. *J Glob Health Sci* 2019;1.

7 Piyasena MMPN, Murthy GVS, Yip JLY, *et al*. Systematic review and meta-analysis of diagnostic accuracy of detection of any level of diabetic retinopathy using digital retinal imaging. *Syst Rev* 2018;7:182.

8 Shi L, Wu H, Dong J, *et al*. Telemedicine for detecting diabetic retinopathy: a systematic review and meta-analysis. *Br J Ophthalmol* 2015;99:823–31.

9 Joseph S, Kim R, Ravindran RD, *et al*. Effectiveness of Teleretinal imaging-based Hospital referral compared with universal referral in identifying diabetic retinopathy: a cluster randomized clinical trial. *JAMA Ophthalmol* 2019;137:786–92.

10 Silva PS, Horton MB, Clary D, *et al*. Identification of diabetic retinopathy and ungradable image rate with ultrawide field imaging in a national teleophthalmology program. *Ophthalmology* 2016;123:1360–7.

11 Rosenblatt TR, Vail D, Saroj N, *et al*. Increasing incidence and prevalence of common retinal diseases in retina practices across the United States. *Ophthalmic Surg Lasers Imaging Retina* 2021;52:29–36.

12 He M-S, Chang F-L, Lin H-Z, *et al*. The association between diabetes and age-related macular degeneration among the elderly in Taiwan. *Diabetes Care* 2018;41:2202–11.

13 Salameh J-P, Bossuyt PM, McGrath TA, *et al*. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. *BMJ* 2020;370:m2632.

14 Flaxel CJ, Adelman RA, Bailey ST, *et al*. Diabetic retinopathy preferred practice Pattern®. *Ophthalmology* 2020;127:P66–145.

15 Ferris FL, Davis MD, Clemons TE, *et al*. A simplified severity scale for age-related macular degeneration: AREDS report No. 18. *Arch Ophthalmol* 2005;123:1570–4.

16 Scanlon PH. Update on screening for sight-threatening diabetic retinopathy. *Ophthalmic Res* 2019;62:218–24.

17 Maa AY, McCord S, Lu X, *et al*. The impact of OCT on diagnostic accuracy of the technology-based eye care services protocol: Part II of the technology-based eye care services compare trial. *Ophthalmology* 2020;127:544–9.

18 Scanlon PH, Malhotra R, Greenwood RH, *et al*. Comparison of two reference standards in validating two field mydriatic digital photography as a method of screening for diabetic retinopathy. *Br J Ophthalmol* 2003;87:1258–63.

19 Conlin PR, Asefzadeh B, Pasquale LR, *et al*. Accuracy of a technology-assisted eye exam in evaluation of referable diabetic retinopathy and concomitant ocular diseases. *Br J Ophthalmol* 2015;99:1622–7.

20 Maruyama-Inoue M, Kitajima Y, Mohamed S, *et al*. Sensitivity and specificity of high-resolution wide field fundus imaging for detecting neovascular age-related macular degeneration. *PLoS One* 2020;15:e0238072.

21 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.

22 Macaskill P, Gatsonis C, Deeks J. *Chapter 10: analysing and presenting results. Cochrane Handbook for systematic reviews of diagnostic test accuracy version*, 2010: 1.

23 Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006;59:1331–2.

24 Takwoingi Y, Riley RD, Deeks JJ. Meta-Analysis of diagnostic accuracy studies in mental health. *Evid Based Ment Health* 2015;18:103–9.

25 Schünemann H, Brożek J, Guyatt G, *et al*. Grade Handbook for grading quality of evidence and strength of recommendations. updated October 2013. The grade Working group, 2013. available from guidelinedevelopment. org/handbook 2013.

26 GRADEpro G. *Computer program. McMaster University (developed by evidence prime). GRADEpro GDT*. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015.

27 Shi L, Wu H, Dong J, *et al*. Telemedicine for detecting diabetic retinopathy: a systematic review and meta-analysis. *Br J Ophthalmol* 2015;99:823–31.

28 Willis BH. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract* 2008;25:390–6.

29 Maa AY, Evans C, DeLaune WR, *et al*. A novel tele-eye protocol for ocular disease detection and access to eye care services. *Telemed J E Health* 2014;20:318–23.

30 Chopra R, Wagner SK, Keane PA. Optical coherence tomography in the 2020s-outside the eye clinic. *Eye* 2021;35:236–43.

31 Sanborn GE, Wroblewski JJ. Evaluation of a combination digital retinal camera with spectral-domain optical coherence tomography (SD-OCT) that might be used for the screening of diabetic retinopathy with telemedicine: a pilot study. *J Diabetes Complications* 2018;32:1046–50.

32 Virgili G, Menchini F, Casazza G, *et al*. Optical coherence tomography (OCT) for detection of macular oedema in patients with diabetic retinopathy. *Cochrane Database Syst Rev* 2015;1:Cd008081.

33 Kansal V, Armstrong JJ, Pintwala R, *et al*. Optical coherence tomography for glaucoma diagnosis: an evidence based meta-analysis. *PLoS One* 2018;13:e0190621.

34 Felfeli T, Alon R, Merritt R, *et al*. Toronto tele-retinal screening program for detection of diabetic retinopathy and macular edema. *Can J Ophthalmol* 2019;54:203–11.

35 Ashrafzadeh S, Gundlach BS, Tsui I. Implementation of Teleretinal screening using optical coherence tomography in the Veterans health administration. *Telemed J E Health* 2021;27:898–904.

36 Kostanyan T, Wollstein G, Schuman JS. New developments in optical coherence tomography. *Curr Opin Ophthalmol* 2015;26:110–5.

37 Anderson HA, Bertrand KC, Manny RE, *et al*. Comparison of two drug combinations for dilating dark irides. *Optom Vis Sci* 2010;87:120–4.

38 Wang S, Zhang Y, Lei S, *et al*. Performance of deep neural network-based artificial intelligence method in diabetic retinopathy screening: a systematic review and meta-analysis of diagnostic test accuracy. *Eur J Endocrinol* 2020;183:41–9.

39 Nielsen KB, Lautrup ML, Andersen JKH, *et al*. Deep Learning-Based algorithms in screening of diabetic retinopathy: a systematic review of diagnostic performance. *Ophthalmol Retina* 2019;3:294–304.

40 Abramoff MD, Cunningham B, Patel B. Foundational considerations for artificial intelligence utilizing ophthalmic images. *Ophthalmology*.

41 Thomas S-M, Jeyaraman MM, Jeyaraman M, *et al*. The effectiveness of teleglaucoma versus in-patient examination for glaucoma screening: a systematic review and meta-analysis. *PLoS One* 2014;9:e113779.