Eye (2020) 34:600–602
https://doi.org/10.1038/s41433-019-0624-7

The ROYAL COLLEGE of OPHTHALMOLOGISTS

CORRESPONDENCE

Validation of RetmarkerAMD as a semiautomatic grading software for AMD

João Pedro Marques[1,2,3] · João Pires[2] · Jorge Simão[1] · Marco Marques[3] · João Q. Gil[1,2,3] · Inês Laíns[2,4] · Dalila Alves[3] · Sandrina Nunes[3] · Maria Luz Cachulo[1,2,3] · John B. Miller[4] · Demetrios G. Vavvas[4] · Joan W. Miller[4] · Deeba Husain[4] · Rufino Silva[1,2,3]

Received: 12 May 2019 / Revised: 24 September 2019 / Accepted: 27 September 2019 / Published online: 23 October 2019

Introduction

RetmarkerAMD® (Retmarker SA, Coimbra, Portugal) is a semiautomatic grading software developed specifically for age-related macular degeneration (AMD). Pilot studies demonstrated higher sensitivity and specificity than manual grading and attested its capacity to decrease grading time and identify more AMD features, thus reducing human error [1]. The aim of this study was to validate RetmarkerAMD® as an AMD grading tool, by comparing it with an already validated and widely used platform—Topcon IMAGEnet2000®.

Methods

Multicentre, cross-sectional study. A set of 202 colour fundus photographs (CFPs) randomly selected from a pool of eyes with and without AMD were used. All images had previously been graded by a senior retina specialist (gold standard) using Topcon IMAGEnet2000®. Two certified graders with different experience independently classified all CFPs using both platforms (Fig. 1; Supplemental Fig. 1) after brightness, contrast and colour balance standardization [2]. Conversion of AMD staging from the Rotterdam classification to the AREDS [3] classification (Supplemental Tables 1 and 2) was achieved to allow a comparison between platforms.

Intra- and inter-grader agreement was evaluated by the percentage of agreement and the weighted Kappa coefficient considering linear weights [4].

Results

The inter-grader analysis for all features analysed with RetmarkerAMD® is shown in Table 1. For AMD staging alone, an almost perfect agreement (93.0%; Kappa = 0.95, $p < 0.001$) was observed. The same was true for AMD staging using Topcon IMAGEnet2000® (90.1%; Kappa = 0.87, $p < 0.001$). Both graders showed a high agreement with the gold standard (90.1%; Kappa = 0.88, $p < 0.001$ and 87.1%; Kappa = 0.86, $p < 0.001$ for graders 1 and 2, respectively). Regarding the inter-modality analysis (Supplemental Table 3), a 76.8% agreement (Kappa = 0.73, $p < 0.001$) and a 70.8% agreement (Kappa = 0.67, $p < 0.001$) was observed for graders 1 and 2, respectively.

Discussion

This study aimed to validate RetmarkerAMD® as a semiautomatic grading software for the fundoscopic changes associated with AMD. We used CFP because of its reproducibility and loyalty to clinical fundoscopy, making it easy to extrapolate results to clinical practice.

Regarding AMD staging, a considerably high agreement was found in both platforms. Subtle differences between graders may have been influenced by a higher level of experience of grader 1. In fact, grader experience is a factor used across studies to justify discrepancies on agreement analyses [5].

Supplementary information The online version of this article (https://doi.org/10.1038/s41433-019-0624-7) contains supplementary material, which is available to authorized users.

✉ João Pedro Marques
marquesjoaopedro@gmail.com

1 Centro Hospitalar e Universitário de Coimbra (CHUC), Coimbra, Portugal

2 Faculty of Medicine University of Coimbra (FMUC), Coimbra, Portugal

3 Association for Innovation and Biomedical Investigation in Light and Image (AIBILI), Coimbra, Portugal

4 Department of Ophthalmology, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA

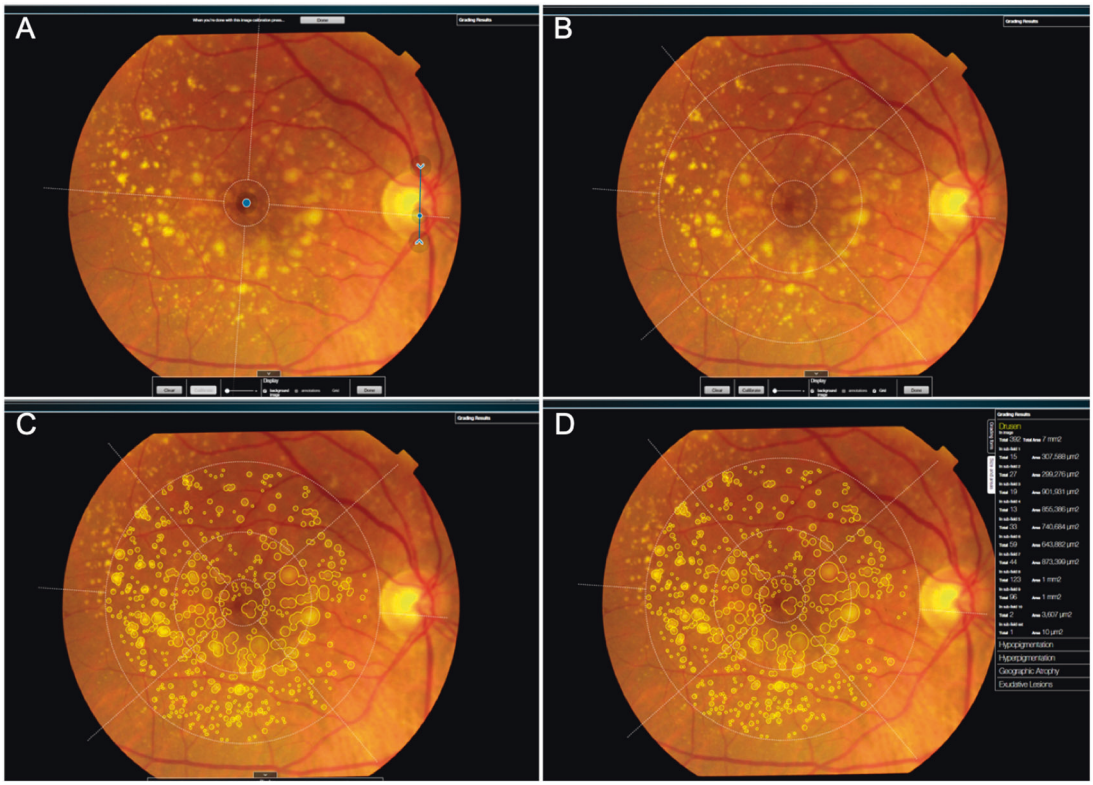

Fig. 1 Real-time demonstration of the grading process using RetmarkerAMD®. **a** First, it is necessary to manually identify the fovea and establish the optic disc diameter (blue dot and arrows, respectively) to achieve the calibration. The software then generates an automated grid **b** according to International Grading System for AMD. **c** Fundus abnormalities including drusen, pigmentary changes, geographic atrophy and neovascular AMD can be quantified using free-forms or pre-defined circles (63, 125, 175, 250 and 500 µm). **d** Grading results (including automatic AMD staging according to the Rotterdam classification—Supplementary Table 1) are depicted on screen and can be exported to a Microsoft Excel® file for posterior statistical analysis. RetmarkerAMD® also allows to categorically quantify the number (0, 1–9, 10–19, ≥20) and area (<1%, <10%, <25%, <50%, ≥50%) of drusen, and additionally to determine the real number and area of drusen (μm^2), total or by semi-field

Table 1 Inter-grader (grader 1 vs. grader 2) agreement analysis using the RetmarkerAMD® software

Variable	% Agreement	Kappa coefficient	Strength of agreement	*P*-value
Number of drusen	85.2%	0.86	Almost perfect	<0.001
Number of drusen <63 µm	82.7%	0.83	Almost perfect	<0.001
Number of drusen 63–125 µm	76.7%	0.78	Substantial	<0.001
Number of drusen >125 µm	82.7%	0.80	Substantial	<0.001
Predominant drusen type within the outer circle	83.2%	0.25	Fair	<0.001
Total area occupied by drusen	90.0%	0.75	Substantial	<0.001
Area covered by drusen in subfield 1	85.2%	0.84	Almost perfect	<0.001
Area covered by drusen in inner circle	83.0%	0.69	Substantial	<0.001
Area covered by drusen in outer circle	86.5%	0.66	Substantial	<0.001
Confluence of drusen	48.8%	0.32	Fair	<0.001
Hyperpigmentation	97.0%	0.93	Almost perfect	<0.001
Hypopigmentation	99.5%	0.98	Almost perfect	<0.001
Geographic Atrophy	99.5%	0.96	Almost perfect	<0.001
Neovascular AMD	100.0%	1.00	Perfect	<0.001
Stage AMD	93.0%	0.95	Almost perfect	<0.001

Kappa coefficient and its correspondent strength of agreement according to Landis & Koch: <0.00 = poor; 0.00–0.20 = slight; 0.21–0.40 = fair; 0.41–0.60 = moderate; 0.61–0.80 = substantial; 0.81–0.99 = almost perfect; 1.00 = perfect

AMD age-related macular degeneration

The comparison of intra-grader analyses using the two software systems revealed a substantial agreement. The clear overlap between early and intermediate stages where subtle and somewhat arbitrary differences are seen may have influenced the results. Classification bias may also have been introduced due to the conversion of the Rotterdam classification into the AREDS classification.

Despite several advantages, RetmarkerAMD® presents some limitations. First, it demands lesion identification by a human grader, a rate limiting and fatigable process. Second, this platform is not prepared to grade optical coherence tomography, a rapidly growing technology with promising outcomes in automatic and semiautomatic AMD grading.

By conducting a carefully planned study, we were able to demonstrate that RetmarkerAMD® is a reliable and consistent semiautomated grading tool for AMD. The validation of RetmarkerAMD® may prompt its use both in clinical studies and in clinical trials.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Marques JP, Costa M, Melo P, Oliveira CM, Pires I, Cachulo ML, et al. Ocular risk factors for exudative AMD: a novel semiautomated grading system. ISRN Ophthalmol. 2013; 2013:464218.
2. Tsikata E, Lains I, Gil J, Marques M, Brown K, Mesquita T, et al. Automated brightness and contrast adjustment of color fundus photographs for the grading of age-related macular degeneration. Transl Vis Sci Technol. 2017;6:3.
3. Age-Related Eye Disease Study Research G. 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.
4. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
5. Scholl HP, Peto T, Dandekar S, Bunce C, Xing W, Jenkins S, et al. Inter- and intra-observer variability in grading lesions of age-related maculopathy and macular degeneration. Graefes Arch Clin Exp Ophthalmol. 2003;241:39–47.