

EDITORIAL



Measuring atrophy in age-related macular degeneration: why does reliability matter?

© The Author(s), under exclusive licence to The Royal College of Ophthalmologists 2022

Eye (2022) 36:1527–1528; <https://doi.org/10.1038/s41433-022-02010-9>

INTRODUCTION

Geographic atrophy (GA) is an advanced stage of age-related macular degeneration (AMD) that has no course of intervention to manage progression. GA leads to significant vision impairments and quality of life decline [1, 2]. Currently, there are no viable treatment options to manage GA; however, this is an area of high clinical importance, and a number of Phase 3 trials are assessing various treatment strategies to manage the progression of GA [3–5]. In this context, it is vital that researchers and clinicians are able to identify atrophy associated with AMD at various stages and assess the impact of treatment on progression from one stage to the next. This is important not only for accurate assessment of treatment effect in clinical trials, but also for effective adoption of these treatment modalities in clinical practice [2, 6]. Clinicians need to be able to identify the appropriate patient population who will benefit from intervention and monitor the progression of atrophy over time. A very important recent development in this field has been the work done by the Classification of Atrophy Meetings (CAM) program where international experts have provided new consensus definitions for atrophy using multimodal imaging modalities. Recent advancement in the early diagnosis of GA progression has been published by the CAM program, which focuses on the use of spectral domain optical coherence tomography (SD-OCT) to assess and classify changes in the retinal and choroidal tissue, as well as identify precursor lesions of GA [2].

HOW CAN WE EVALUATE ATROPHY USING SD-OCT?

The CAM group developed a system in which SD-OCT could be used to identify changes in the outer retina and retinal pigment epithelium (RPE), leading to the classification of patients into distinct groups that may be of higher, or lower, degrees of GA progression [2]. The framework proposed by the CAM group employs atrophy classification based on the anatomic layers impacted on OCT. The key categories include incomplete retinal pigment (RPE) and outer retinal atrophy (iRORA) and complete RPE and outer retinal atrophy (cRORA) [2].

Devised as a research tool to facilitate AMD related research, it is vital that the classification system is highly reliable. The CAM group recently published a study assessing inter-rater agreement by assessing twelve readers from six reading centers after formal training and demonstrated that most of the features of iRORA and cRORA could be assessed relatively consistently and robustly [6].

As an important tool to assess and monitor disease progression and response to treatment in AMD, the CAM classification system will gain increasing clinical applicability if it can demonstrate high reliability not only among “expert” reading center evaluators, but

also among clinicians who will need to make treatment decisions. In this issue of EYE, Chandra et al. aim to address this important topic by assessing the reliability of the cRORA classification by clinical experts and provide important insights into some of the potential challenges that must be addressed as we try to bring the learnings from the research realm into clinical practice [2].

WHAT IS RELIABILITY?

Reliability is a psychometric property of measurement tools that is often discussed alongside validity. While validity refers to the accuracy of a measurement tool to appropriately evaluate the construct that it is planned to measure, reliability refers to the ability for assessors to consistently use the tool. In the case of the CAM classification system, validity refers to the accuracy in which the system evaluates changes that are appropriately considered markers for GA progression. Reliability, on the other hand, accounts for how consistently clinicians are able to measure the markers of GA progression and come to the appropriate conclusions [2]. There are a few different types of reliability, which are broadly categorized as internal and external reliability. Internal reliability relates to the consistency of individual items within a measurement tool [7]. In the CAM classification system, this would be consistency between outer retina readings and RPE readings to come to the same classification conclusions. External reliability, on the other hand, focuses on the ability in which the classification system can come to the same result between different evaluators (inter-rater), as well as when the same evaluator repeats their measurement (intra-rater) [2, 7, 8].

A specific form of intra-rater reliability—test-retest reliability—refers to the degree in which a clinician can rate a patient’s SD-OCT images using the CAM classification system, then (after some time has passed) re-rate the same images. High test-retest reliability would occur if the clinician can effectively reproduce the same classifications as they had given on the first instance. When classifications differ between the test and retest classifications, there are issues with the reliability of the classification tool itself [2].

INTER- AND INTRA-RATER RELIABILITY OF THE CAM CLASSIFICATION SYSTEM

As discussed above, the investigation of the inter- and intra-rater reliability of the CAM classification system determined that well trained reading center readers demonstrated reliable classification. However, in a more broad clinician group, they struggled to consistently rate cRORA with this system [2, 6]. Both inter-rater and intra-rater reliability was drastically improved in individuals who intimately understood the CAM criteria; however, the other clinical experts were unable to reliably measure and classify patients with GA as being within the cRORA classification [2, 6]. This proves to be an important goal for the CAM group to address.

Received: 20 February 2022 Revised: 23 February 2022 Accepted: 23 February 2022
Published online: 8 March 2022

Given the growing importance and impact of the CAM classification in the field of AMD, it will be vital that not only reading center experts, but also clinicians are able to consistently and appropriately classify atrophy in AMD to ensure appropriate treatment decisions and monitoring for response to intervention in their patients.

WHAT CAN BE DONE TO FACILITATE EFFECTIVE ADOPTION OF CAM CLASSIFICATION SYSTEM INTO CLINICAL PRACTICE?

The external reliability concerns with the CAM classification methodology when adopted by clinicians can be potentially addressed in a number of ways. First and foremost, advanced education and training on how to utilize the classification system has shown to aid in reliability—as clinicians who were accustomed to the CAM criteria had much better reliability than non-familiar ophthalmologists within the CAM reliability study [2, 6]. Although difficult to implement, widespread education and training would naturally improve the ability for clinicians to utilize this tool effectively. This can prove to be a difficult task, as widespread implementation of training and education would likely require resource and buy-in from major ophthalmology organizations with broad reach in the field. Despite these difficulties, training and education are always an important aspect of implementing novel clinical methodologies.

Another consideration for improving reliability would require iterative improvements to the CAM classification system itself, in which the specific areas that create difficulties in consistent rating are addressed. Within Chandra et al.'s evaluation of the CAM classification system, they evaluated the differences that white-on-black vs black-on-white SD-OCT images had on the reliability [2]. This did identify some areas in which reliability may be improved, as agreement was typically improved using white-on-black images; with the exception of classification components that measured RPE attenuation/loss. It is these types of iterative investigations and subsequent improvements to the CAM classification system that may aid in improving the overall reliability of the tool.

CONCLUSION

The CAM classification system is a major step forward toward an objective and staged classification of atrophy in AMD. Similar to the DRSS classification for Diabetic Retinopathy and the AREDS classification for AMD, the CAM classification system will play a major role in our understanding of disease progression and allow for careful assessment of treatment effect for new therapeutics in the management of atrophy in AMD. As a research tool, the CAM classification has demonstrated high reliability when utilized in the context of trained readers at reading centers. However, the applicability of this classification system in real-world clinical practice to guide clinical decision making appears to be limited due to reliability concerns based on current evidence in the literature. The work by Chandra et al. is an important step forward to highlight some of the challenges around widespread adoption of this classification system into clinical practice and to provide insights into how to improve reliability of this important tool in the hands of clinicians. Educational and training efforts may prove to be the vital next step in widespread adoption and improvements to the reliability of the CAM classification system in the real-

world; however, improvements to the system itself may help future iterations to be better deployed within broad ophthalmologic practice.

Mark R. Phillips¹ and Varun Chaudhary^{1,2}

¹Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada. ²Department of Surgery, McMaster University, Hamilton, ON, Canada. ✉email: vchaudh@mcmaster.ca

REFERENCES

1. Fleckenstein M, Mitchell P, Freund KB, Sadda S, Holz FG, Brittain C, et al. The progression of geographic atrophy secondary to age-related macular degeneration. *Ophthalmology*. 2018;125:369–90. <https://doi.org/10.1016/j.ophtha.2017.08.038>.
2. Chandra S, Rasheed R, Sen P, Menon D, Sivaprasad S. Inter-rater reliability for diagnosis of geographic atrophy using spectral domain OCT in age-related macular degeneration. *Eye*. 2022;36:392–7. <https://doi.org/10.1038/s41433-021-01490-5>.
3. Heier JS, Pieramici D, Chakravarthy U, Patel SS, Gupta S, Lotery A, et al. Visual function decline resulting from geographic atrophy: results from the chroma and spectri phase 3 trials. *Ophthalmol Retin*. 2020;4:673–88. <https://doi.org/10.1016/j.oret.2020.01.019>.
4. Liao DS, Grossi FV, El Mehdi D, Gerber MR, Brown DM, Heier JS, et al. Complement C3 inhibitor pegcetacoplan for geographic atrophy secondary to age-related macular degeneration: a randomized phase 2 trial. *Ophthalmology*. 2020;127:186–95. <https://doi.org/10.1016/j.ophtha.2019.07.011>.
5. Steinle NC, Pearce I, Monés J, Metlapally R, Saroj N, Hamdani M, et al. Impact of baseline characteristics on geographic atrophy progression in the FILLY trial evaluating the complement C3 inhibitor pegcetacoplan. *Am J Ophthalmol*. 2021;227:116–24. <https://doi.org/10.1016/j.ajo.2021.02.031>.
6. Wu Z, Pfau M, Blodi BA, Holz FG, Jaffe GJ, Liakopoulos S, et al. OCT signs of early atrophy in age-related macular degeneration: interreader agreement: classification of atrophy meetings report 6. *Ophthalmol Retin*. 2022;6:4–14. <https://doi.org/10.1016/j.oret.2021.03.008>.
7. Fayers PM, Hays R, Hays RD. Assessing quality of life in clinical trials: methods and practice. New York (USA): Oxford University Press; 2005.
8. Lachin JM. The role of measurement reliability in clinical trials. *Clin Trials*. 2004;1:553–66. <https://doi.org/10.1191/1740774504cn0570a>.

AUTHOR CONTRIBUTIONS

MRP and VC were responsible for conception of idea, writing, critical review and feedback on manuscript.

COMPETING INTERESTS

MRP: nothing to disclose. VC: Advisory board member: Alcon, Roche, Bayer, Novartis; Grants: Bayer, Novartis—unrelated to this study.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Varun Chaudhary.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.