Australian Orthoptic Journal 1992; 28: 11-17

RETINAL PHOTOGRAPHIC GRADING: THE ORTHOPTIC PICTURE

ROBERT SPARKES, DipAppSc(Cumb), DOBA, MOAA

ASSOC. PROF. PAUL MITCHELL, MD, FRACO, FRACS, FCOphth

PETERIS DARZINS, BMBS, FRACP
Department of Clinical Ophthalmology, Westmead Hospital, Westmead, Sydney, 2145.

Abstract

Signs of the major retinal causes of blindness in Australia are readily identified from stereoscopic fundus photographs. This study involved the masked grading of 922 photographic slide sets which included 400 cases of advanced Age-related Macular Degeneration (AMD) and a similar number of age and sex-matched controls.

Photographic colour slides of the retina showing AMD lesions, including drusen (size, number and confluence), focal pigment, geographic atrophy and neovascular disease, were graded individually by an ophthalmologist, orthoptist and a nurse using the protocol set by the Waterman Chesapeake Bay study. These grades were then used to define four categories of AMD severity.

Kappa statistics were generated from the data to show inter-observer variability between the ophthalmologist and orthoptist and the orthoptist and nurse. The resulting high level of agreement in all areas graded (between 0.45 . . . 0.83) indicated that the orthoptist compares favourably to the ophthalmologist and the nurse in grading expertise.

Orthoptic involvement in this area of ophthalmic research opens new spheres of interest and professional development.

Key words: *AMD, Grading, Fundus photographs, kappa, inter-observer variability.*

INTRODUCTION

Masked grading, that is grading which obscures other subject characteristics, is becoming standard in modern clinical trials throughout the world. Many present studies in ophthalmology, as well as many planned studies, utilise non-ophthalmological staff to grade lesions of the eye. This has the effect of reducing observer bias of the investigating party.

This paper gives some insight into how two non-ophthalmologists, an orthoptist and a nurse, were trained in the grading of retinal lesions. It attempts to address the viability of using non-ophthalmologists in studies of this and similar types by comparing the inter-observer variability between an orthoptist and ophthalmologist. Kappa statistics were also calculated between the orthoptist and nurse comparing the results to a similar study done in Beaver Dam U.S.A.[1]

Age related macular degeneration is the most common blinding disease of the elderly community in Australia. Up to 55% of registrations for the blind pension now have vision loss attributable to AMD.[2] In 80% of cases, where blindness is attributable to AMD, a haemorrhagic lesion appears which then passes through the stages of fibrosis and atrophy. This is termed a disciform lesion or "wet" AMD. The other

Address for correspondence: Robert Sparkes, Department of Clinical Ophthalmology, Westmead Hospital, Westmead, Sydney, 2145.

20% of AMD presentations are entirely atrophic and can be termed as geographic atrophy or "dry" AMD.[3] Early AMD lesions that precede these blinding stages include the presence of drusen and other abnormalities of the retinal pigment epithelium in the macular region.[4] The aetiology and pathogenesis of AMD is poorly understood.[3]

Grading standards in this study were based on the AMD protocol of the Waterman Chesapeake Bay study.[5] This involved grading lesions of the fundus which may be related to AMD. These included drusen characteristics such as type and size, drusen confluence and the space they occupied, focal hyperpigmentation, geographic atrophy and non-geographic atrophy. Neovascular and non-neovascular lesions, as well as any other fundal anomalies, were noted and classified as per the protocol.

MATERIALS AND METHODS

Nine hundred and twenty two photographic slide sets from the Newcastle AMD case control study[6] were randomly selected for grading. These included 400 cases of advanced AMD and a similar number of age and sex-matched controls selected from the Newcastle age-restricted electoral roll.

The involvement of the orthoptist and the nurse in the grading process was as follows.

(i) Establishment of a slide grading centre

A slide grading centre was established within Westmead Hospital where slide grading, computer entry of data and storage of slides and sheets occurred.

(ii) Training

Grader training took place over a two month period. This involved attending ophthalmology registrar training lectures and many hours of tutoring in retinal disease. Training was performed by Associate Professor Paul Mitchell at Westmead Hospital who acted as an adjudicator throughout the grading process. Experience was gained with the grader training programme of the "Beaver Dam Eye Study" that took place in Wisconsin U.S.A.[1]

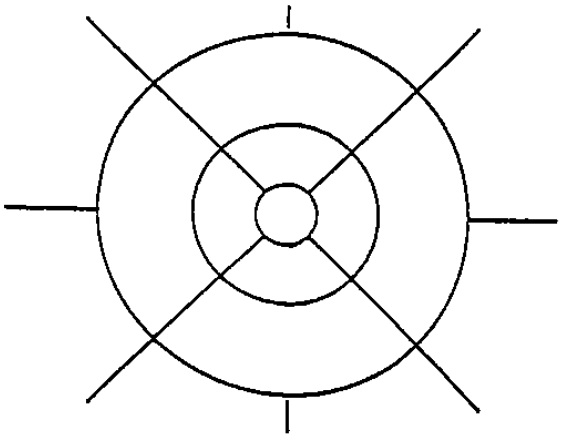


Figure 1: The grid which centred onto the slide by aligning the lines at 12 and 3 o'clock

(iii) Assembling the required materials

Before any grading could be carried out the following materials were needed:

(a) Viewing Box

This had Kelvin rating of 6200° which is bluer than daylight. The Wisconsin Age-Related Maculopathy grading system states that "light of a lesser rating emits a yellowish hue that makes it difficult to define more subtle drusen".[1] Placement of grids and the grading process was carried out on this box.

(b) Slide mounting sheets

These allowed the placement of stereoscopically paired slides ready for viewing.

(c) Viewer

Accurate placement of the grid and all grading was performed using a Donaldson viewer, which provides 5×magnification which combined with the camera's 3×totalled some 15× magnification.

(d) Gridding templates

Consisted of three concentric circles, which were supplied by Professor R. Klein, Wisconsin U.S.A. (Fig. 1). The gridding circles had diameters of 1000 μm, 3000 μm and 6000 μm. The two inner most circles represented the central zone and the space between the 3000 μm and 6000 μm circles represented the pericentral zone. The grid also contained two diagonal lines which divided each zone into segments.

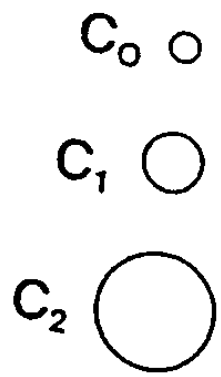


Figure 2: The measuring tool which was used to determine drusen size.
Circle C_0 represents 63 μm
Circle C_1 represents 125 μm
Circle C_2 represents 250 μm

(e) Measuring tool template
These consisted of three concentric circles, varying in size (as specified by the protocol), superimposed on transparent sheeting (Fig 2). These are introduced between the slide and the Donaldson viewer by the grader in order to measure the various characteristics found.

(f) Grading forms
The recording form is based on the Waterman Chesapeake Bay study grading sheet template. This form has identifying information, eye graded, photographic quality and sections that list specific grading characteristics.

(iv) Pre-Grading
Prior to grading, a grid was placed on one of the stereoscopically paired slides (usually the left slide). The grid was then centred on the fovea using lines at 12 o'clock and 6 o'clock as positioning tools. Finally the grid was secured to the cardboard mount of the slide using sticky tape and placed back into the mounting sheet ready for grading.

(v) Grading
The following characteristics were assessed when grading:
(1) Photo and Stereo Quality. These were graded as adequate, fair, poor or missing. Poor photo and/or stereo quality were often caused by opacity present in the media of the eye (eg cataract, corneal disease or asteroid hyalosis). If the photo was graded as poor this automatically precluded the slide from further grading poor Stereo quality, however, did not preclude the eye from grading but was noted if no stereo view could be attained (eg. missing slide).
(2) Neovascular disease. All neovascular diseases were identified and categorised into the following types.

Disciform lesions
These consist of subretinal elevation beneath or adjacent to the fovea, surrounded by an area of serous retinal detachment. The lesion is usually associated with a subretinal haemorrhage. It is common for exudates to appear around the margins of the serous detachment and in late stage disciforms massive exudation and fibrosis is prominent.[7]

The disciform lesions were classified into categories of:
(a) Haemorrhagic
(b) Fibrous
(c) Atrophic
(d) Treated
(e) unclassified.

Retinal pigment epithelium detachment (RPE)
Fluid collects beneath the RPE and Bruch's membrane to form a localised detachment. It appears as a well demarcated elevation in the macular area and can be distinguished from serous elevation due to this more defined appearance.[7] Any detachment of the RPE, with or without serous elevation, precluded the eye from further grading.

These diseases precluded grading of other characteristics.

(3) Drusen
Drusen are areas of amorphic, non cellular material that lie between Bruch's membrane and the retinal pigment epithelium layer. They are commonly found around the macular region and vary in appearance from fine granular deposits to large confluent areas of deposition.[8] Drusen may be defined as hard, soft, calcific or reticular.[9]

In this study drusen number, size, percentage of grid occupied and confluence were graded regardless of type. Grading of lesions were divided into two zones, central and pericentral.

The options — none, questionable, can't decide and can't grade were available in this section as well as all other sections covered in the grading.

(a) Number

In each zone the amount of drusen present (if seen to be as prominent as determined by the protocol) were classified into the following categories:

(i) 1-4 present
(ii) 5-19 present
(iii) ≥20 present.

(b) Size

Drusen size was measured using a measuring tool consisting of circles of varying sizes. These circles were labelled as follows:

C_0 = circles that measured any drusen ≤63 μm
C_1 = circles that measured any drusen 64-124 μm
any drusen greater than C_1 was categorised as ≥125 μm

The measuring tool was placed over the largest drusen in each zone. The size findings were then categorised as above.

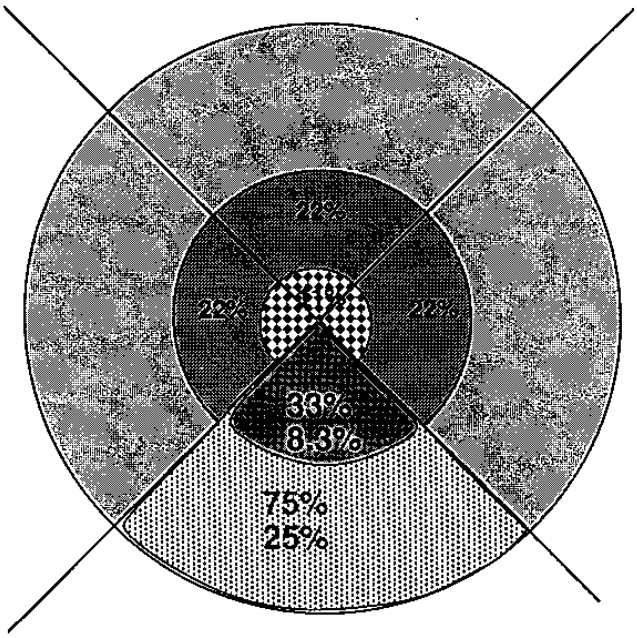
(c) Percentage of area covered

If drusen size was greater than 63 μm (considered large drusen) a measurement was performed by estimating the percentage of each zone covered by large drusen. These were recorded as follows:

(a) not applicable i.e. drusen size ≤64 μm
(b) ≤10%
(c) 11-33%
(d) 34-67%
(e) >67%.

A graph representing the gridding circles was created and allowed quick reference to percentage of area covered. (see Fig 3) This graph was quickly devised by using the formula πr^2 (area of a circle), then dividing the area into appropriate segments as displayed on the grid.

All areas of geographic atrophy and non geographic atrophy (RPE disease) were excluded from the area ie. were thought of as non diseased retina. Areas of reticular drusen were included in the area covered.



Central circle is 3.7% of pericentral zone and
11% of central zone = one central circle
33% of central zone = one 22% pie + central circle
67% of central zone = three 22% pie sectors.
11% of pericentral zone = 3 whole central circles
33% of pericentral zone = whole of central zone
67% of pericentral zone = 2 whole central zones

Figure 3: The gridded area of the slide was broken up into areas (shown here as percentages) by using the formula πr^2.

(d) Confluence of pairs (>63 μm)

Confluence is defined as a merging or touching of two of more drusen.[1] Only drusen of size 64 μm or more are graded in this category. Any druse that had become confluent were counted and recorded as being <or>10 pairs.

(4) Focal Hyperpigmentation

These are areas where excessive pigment has been generated. This phenomenon is common in AMD and is associated with RPE disturbances.[8] Any areas where pigment was exuded from the RPE layer were noted and recorded as being either less than or greater than or equal to the standard.

Care was taken to distinguish fine pigment deposits from dirt on the slide or mounting sheet.

(5) Geographic Atrophy

This is an area of well demarcated drop out of the RPE, or complete depigmentation of the RPE, exposing the large choroidal vessels.[9] The area often appears as would a country on a map. Grading consisted as being less than or greater than/equal to the standard. If the geographic atrophy (GA) was greater than 50% of the circles then the slide was precluded from grading other anomalies.

(6) Non-Geographic Atrophy (RPE Degeneration)

Faded drusen often contribute to marked depigmentation of the retina leaving areas that are ill defined but still not advanced enough to see the choroidal layer through them.[9] These are termed as areas of non geographic atrophy (non GA). Again these areas were graded as being less than or greater than/equal to the standard.

(7) Other Macular/Retinal Disease

Any lesions of the retina that did not fit into the categories defined above were noted. Disease such as arteriolar opacity, vein or artery occlusions, micro-aneurysms, haemorrhage or premacular fibrosis were some of the more often encountered lesions.

Data Collection/Calculation and Storage

Information from the slides were recorded onto the grading form and then entered into a computer using a multi-relational data base. Upon entry into the data base, automatic AMD grades were calculated. These grades were developed by Bressler and Taylor and have been used in the Chesapeake Bay Waterman study.[3]

Sheets and slides were then filed away in numerical order to allow easy reference if any problems arose.

TABLE 1
Kappa Scores of Inter-observer Variability between the Orthoptist and Ophthalmologist (n = 65)

Variable	Kappa
Neovascular disease	0.727
Drusen number	0.618
Drusen size	0.453
Focal Hyperpigmentation	0.624
Geographic Atrophy	1.000
AMD Grade	0.578

Key to Cohen Kappa Statistic
>0.75 = strong agreement
0.40-0.74 = fair to very good agreement
<0.40 = poor agreement

TABLE 2
Kappa Scores of Inter-observer Variability between the Orthoptist and the Nurse (n = 922)

Variable	Kappa
Neovascular Disease	0.888
Drusen number	0.695
Drusen size	0.685
Focal Hyperpigmentation	0.833
Geographic Atrophy	0.892
AMD Grade	0.671

Statistical Calculation

All statistical calculations (Cohen Kappa or as commonly referred to as Kappa statistic) were performed using the Systat 5.1 programme which was accessed through a Macintosh SE/30 microcomputer.

RESULTS

Computer analysis of data entered from the grading process compared the agreement levels of some of the various characteristics graded. The statistical analysis of this data using the systat programme provided the following statistics that can be found in the Tables labelled 1, and 2.

Orthoptist and Ophthalmologist inter-observer variability

Kappa statistic was calculated on a sample group ($n = 65$) for a number of variables and the results shown in Table 1.

Grader one and two inter observer-variability

Kappa statistic was calculated on a large sample group ($n = 922$) and the results shown in Table 2.

Inter-observer variability in the Beaver Dam Eye Study

Kappa statistic was used in the Beaver Dam Eye study[1] to determine the inter-observer variability of two of the studies graders.[6] Slides were graded from a large sample group ($n = 857$) which is

TABLE 3
Shows Kappa Scores of Inter-observer Variability between two graders in the Beaver Dam Eye Study (n = 857)

Variable	Kappa
Drusen size	0.51
Focal Hyperpigmentation	0.70
Geographic Atrophy	0.83

Key to Cohen Kappa Statistic
>0.75 = strong agreement
0.40-0.74 = fair to very good agreement
<0.40 = poor agreement

similar in number to the Newcastle study. The results are shown in Table 3 and are comparable to our own results.

DISCUSSION

From the results it has been shown that there is statistically good agreement between the orthoptist and the ophthalmologist. The kappa statistics indicated a good level of agreement in grading neovascular disease, drusen number, focal hyperpigmentation and the final AMD grade (which is of course influenced by the aforementioned). A fair level of agreement was attained in judging drusen size. This slightly worse result (although still statistically acceptable) may be explained by the problems with measuring significant drusen. Some faint large drusen may not have been included by one grader because they felt its appearance was below the protocol standard, yet the drusen may have been included by the other grader. The perfect agreement found in grading geographic atrophy should be overlooked because a very small number of this lesion type were included in the random sample. This makes the result statistically invalid.

The statistical calculations between graders showed a strong level of agreement in judging neovascular disease, focal hyperpigmentation and geographic atrophy. In fact near perfect agreement is considered >0.81. A very good level of agreement was attained for drusen number and size and AMD grade. This indicates that a similar understanding of the protocol and grading procedures existed between the two graders.

A high level of agreement, between graders one and two, can be attributed to the many sessions of "problem solving" that took place. These sessions offered times where the two graders and the adjudicator met and discussed areas of difficulty. One of the most often encountered problems was in the determination of whether or not a characteristic was of protocol standard. This can be attributed to the subjective nature of determining the standard set in the protocol. It was also difficult to assess the standard if the photograph was slightly blurred, due either to a lens opacity or other quality diminishing characteristics.

Inter-observer variability between graders one and two and the orthoptist and ophthalmologist compare favourably to the kappa statistics analysed in the Beaver Dam Eye Study. This study, however, differed from the Newcastle Study in that it used a more complex grading format. The variables shown in Table 3 are the ones most similar to the grading system employed in the Newcastle Study.

It can be interpreted from these results that with adequate training and knowledge a statistically good correspondence, between an orthoptist and ophthalmologist, as well as between other graders, can be attained in the identification of retinal lesions.

Grading retinal photographs can be tedious and frustrating at times but on the other hand is often rewarding as it provides a good knowledge of medical retina. Entering a grader training programme is not suited to everybody but it is certainly more interesting if you have an "eye" background and you are able to utilise the skills learnt.

As there are a growing number of studies that will be utilising a masked grading system then this type of training and knowledge may provide an increasing number of opportunities for orthoptic involvement in ophthalmological research.

CONCLUSIONS

This study has shown that the orthoptist compares favourably with the ophthalmologist and the nurse in grading lesions of the retina. This result was attained following a period of specialised training and the use of a well defined

protocol. Ongoing problem solving and adjudication sessions provided a continuing high standard of agreement through-out the grading process and a comparable agreement rate was shown with that of the Beaver Dam Eye Study.

References

1. Klein R, Davis M, Magli Y, Segal P, Klein B, Hubbard L. The Wisconsin Age-Related Maculopathy Grading System: Ophthalmology 1991; 98: 1128-1134.
2. Cooper RL. Blind registrations in Western Australia: a five year study. Aust NZ J Ophthalmology 1990; 18: 421-426.
3. Ferris FL III. Senile Macular Degeneration: review of epidemiological features. Am J Epidemiology 1983; 118: 132-151.
4. Gass JDM. Stereoscopic Atlas of Macular Diseases: Diagnosis and treatment, 2nd ed. St Louis: CV Mosby, 1977.
5. Bresler NM, Bresler SB, West SK. The grading and prevalence of macular degeneration in Chesapeake Bay Waterman. Arch Ophthal 1989; 107: 847-852.
6. Mitchell RP, Darzins P. The Newcastle Age-related Maculopathy Study . . . unpublished Department of Clinical Ophthalmology, Westmead Hospital, Westmead 2145.
7. Spalton DJ, Hitchings RA, Hunter PA eds. Atlas of Clinical Ophthalmology. London: Gower. 13.44-13.46.
8. Peyman PA, Sanders DR, Goldberg MF eds. Principles and practice of Ophthalmology Volume II. Philadelphia: WB Saunders. pp 927-935.
9. Klein R, Davis M, Magli Y, Segal P, Klein B, Hubbard L. The Wisconsin Age-Related Maculopathy Grading System: Protocol. Department of Ophthalmology, University of Wisconsin Medical school, Madison, U.S.A.