

Visual impairment registration: evaluation of agreement among ophthalmologists

E Guerin¹, G Bouliotis² and A King³

Abstract

Purpose To evaluate consistency among consultant ophthalmologists in registration of visual impairment of patients with glaucoma who had a significant visual field component to their visual loss.

Method Thirty UK NHS consultant ophthalmologists were asked to grade data sets comprising both visual acuity and visual fields as severely sight impaired, partially sight impaired, or neither. To assess intra-consultant agreement, a group of graders agreed to repeat the process.

Results Kappa for inter-consultant agreement ($n = 30$) for meeting the eligibility criteria for visual impairment registration was 0.232 (95% CI 0.142–0.345), the corresponding inter-class correlation score was 0.2 (95% CI 0.172 to 0.344). Kappa for intra-consultant agreement ($n = 16$) ranged from 0.007 to 0.9118.

Conclusions When presented with the clinical data necessary to decide whether patients with severe visual field loss are eligible for vision impairment registration, there is very poor intra- and inter-observer agreement among consultant ophthalmologists with regard to eligibility. The poor agreement indicates that these criteria are open to significant subjective interpretation that may be a source of either under- or over-registration of visual impairment in this group of patients in the UK. This inconsistency will affect the access of visually impaired glaucoma patients to support services and may result in inaccurate recording of the prevalence of registerable visual disability among glaucoma patients with severe visual field loss. More objective criteria with less potential for misclassification should be introduced.

Eye (2014) 28, 808–813; doi:10.1038/eye.2014.66; published online 4 April 2014

Introduction

In the United Kingdom, registration as visually impaired is of practical importance to those with severe visual disability. For those eligible, it triggers access to a range of services that can facilitate independent living or continued participation in the workforce.¹ The process of registration of a patient as visually impaired requires completion of a certificate of vision impairment by an ophthalmologist who recognises that the patient meets the visual criteria for its completion. It has been estimated that a large proportion of patients eligible for vision impairment registration remain unregistered.²

Visually impaired individuals in the UK may be registered as being either severely sight impaired (previously known as blind) or sight impaired (previously known as partially sighted). The criteria for each category are based on the extent of loss of both the patients' visual acuity and visual field. However, the visual field criteria for blind registration are only imprecisely defined with such terms as 'very restricted' and 'gross defect'.³ It has been found that patients are less likely to be registered if they predominantly have visual field loss rather than visual acuity loss.^{2,4,5}

As the evidence suggests that visual impairment is likely to be under-registered. This study was designed to evaluate the consistency within and across consultant ophthalmologists in deciding whether a patient is eligible for visual impairment registration.

Materials and methods

Study design

Thirty consultant ophthalmologists with experience of registration of patients as visually impaired were invited to act as raters.

¹WrexhamMaelor Hospital, Wrexham, UK

²Midland Hub for Trials Methodology Research, University of Birmingham, Birmingham, UK

³Queens Medical Centre, Nottingham, UK

Correspondence: E Guerin, Wrexham Maelor Hospital, Croesnewydd Road, Wrexham LL13 7TD, UK
Tel: +44 (0)1978 291100;
Fax: +44 (0)1978 725379;
E-mail: eoinguerin@gmail.com

Received: 24 October 2013
Accepted in revised form: 25 February 2014
Published online: 4 April 2014

This work was partly presented to the UK and Eire Glaucoma Society.

Consultants based at Teaching and District General Hospitals including those with glaucoma and other subspecialty interests were invited to participate. The evaluating material was a folder consisting of 35 fully anonymised visual acuity and visual field examination data sets, condensed onto separated single sheets which showed the right and left visual fields (24-2 Humphrey visual fields) with related visual acuity for each eye (Snellen visual acuity) (Figure 1). All examinations presented were from patients with glaucoma, all had a visual acuity of 6/12 or better in both eyes and severe visual field loss in at least one eye with varying degrees of visual field loss in the fellow eye. The strict visual acuity criteria were applied to ensure that decision-making regarding vision impairment registration was based primarily on the extent of visual field loss.

Examinations were randomly mixed into a folder and were rated by each participant in the same order and without a time limit in an unsupervised environment without restrictions.

In the first part of the evaluation, an assessment of inter-rater consistency was undertaken. Each rater was

asked to make a judgement as to whether to rate each patient's data set as, severely sight impaired, sight impaired, or neither, based exclusively on the combination of visual field and visual acuity information presented in the data sets. No additional information was provided.

The rating was completed on a *pro forma* for each examination. Each ophthalmologist undertook the assessment independently of each other. No location for completion of the rating was stipulated. The evaluation process was thoroughly explained prior to undertaking the evaluation and followed the normal process that would be undertaken when evaluating a patient for visual impairment registration in clinical practice.

In the second part of the evaluation, an assessment of intra-rater consistency was undertaken. Raters were asked to repeat the evaluation after a minimum of 1 month but within 3 months.

Statistical analysis

This is a typical rating-agreement study involving multiple experts acting as raters. Analysis involved both

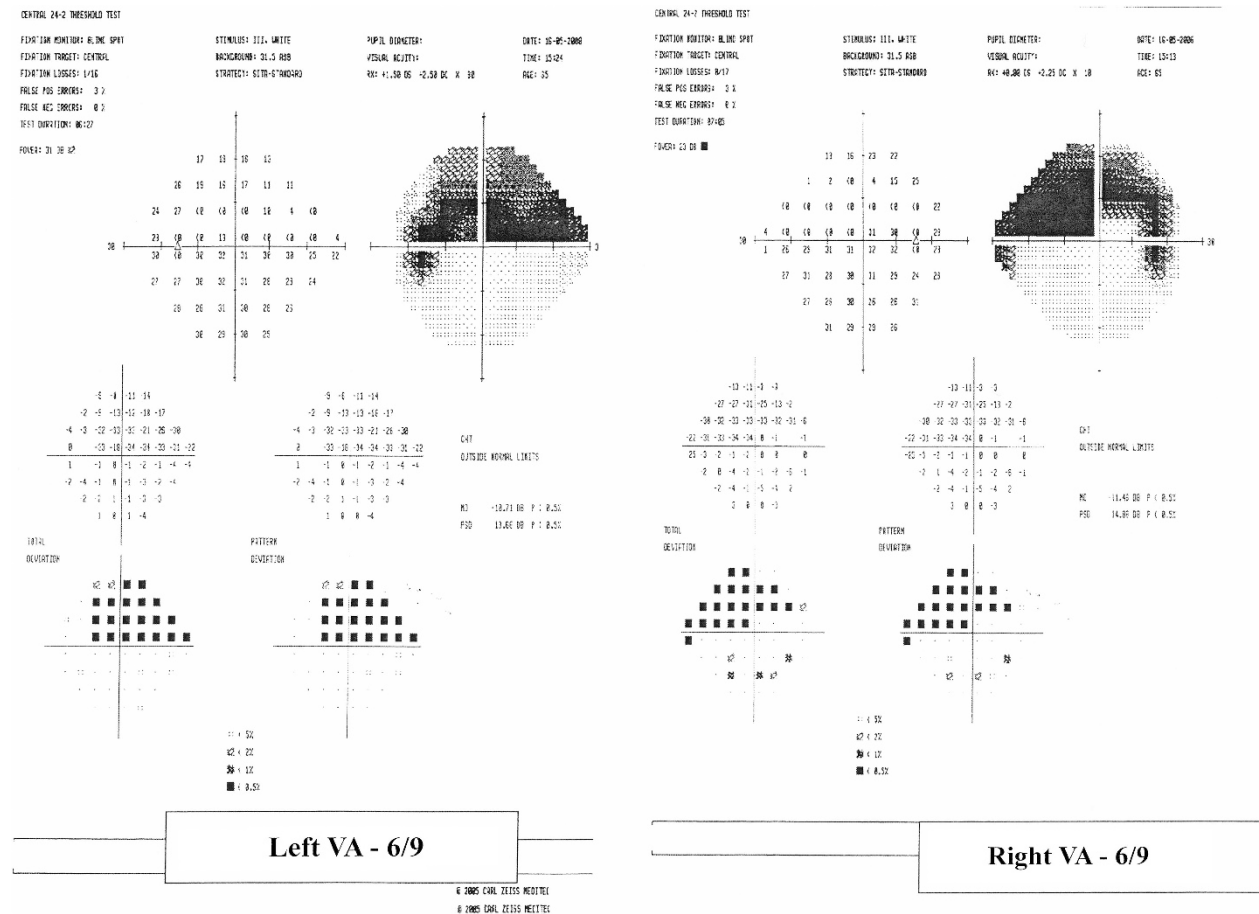


Figure 1 Sample data set used for rating.

agreement omnibus indexes and regression models. The inter-rater agreement was assessed by estimating three of the most popular agreement indexes (the Fleiss generalised kappa,^{6,7} the Krippendorff's alpha,^{8,9} and the interclass consistency (ICC) scores).^{10,11} Kappa was used also for assessing intra-rater consistency. Such scores are univariate statistics suitable for designs involving multiple raters that allow estimation of the size of disagreement.^{12,13}

Although kappa and alpha provide agreement score for only the involved participants, ICC computes the intra-class correlation for random effects repeated-measures ANOVA by considering both the raters and patient data as sampled from a larger population. Finally, Krippendorff's alpha emphasises a more efficient analysis of agreement expected by chance.¹⁴

Both trichotomous (the three categories of severe sight impairment, sight impairment, and neither) and dichotomous (sight impaired and severely sight impaired condensed into a single category) univariate analysis were undertaken. Statistical software used was R (www.r-project.org), SAS (SAS Institute Inc., Cary, NC, USA) and Stata (StataCorp LP, College Station, TX, USA).

A power determination (R CRAN package 'kappaSize') was performed for Fleiss kappa assuming six raters, three ordinal ratings, expected agreement of 0.35 (expected lower level of 0.20 and upper level of 0.45). To achieve an alpha of 0.05, it was calculated 95 examinations would be required. Given the higher number of participating raters in our study, 35 examinations were considered to be an appropriate sample size.

Results

One data set was excluded from analysis as it was noted that it did not fulfil the original study criteria (visual acuity of 6/12 or better in both eye). As a consequence of this, instead of undertaking analysis on 35 sets of data as initially intended (and on which the study had been powered), only 34 data sets were evaluated.

Descriptive statistics revealed substantial individual variation across the consultants performing the ratings. The proportion of patients rated as not suffering from any sight impairment *vs* those rated as severely sight impaired varied substantially between raters. More specifically, rater 21 classified 80% of the patients to be sight impaired, whereas raters 9 and 10 judged only 29.4% to be sight impaired. Rater 17 judged 67.6% of patients as severely sight impaired when four other raters found no patients merited registration for this (Table 1).

Not surprisingly, agreement indexes confirmed a limited agreement with kappa, alpha, and ICC scores as 0.160, 0.295, and 0.307, respectively. When looking at the subgroup of consultant ophthalmologists specialising in

glaucoma, the kappa and ICC scores were 0.226 and 0.399, respectively, and when looking at those with >10 years consultant experience the kappa and ICC scores were 0.193 and 0.323, respectively (Table 2).

Next, we analysed dichotomous ratings by looking at the consistency for grading between the two categories for any form of visual impairment *vs* no visual impairment. The achieved agreement was increased to 0.232 (kappa), 0.233 (Krippendorff), and 0.2 (ICC). When looking at the subgroup of consultant

Table 1 The proportion of data sets that were rated by each rater as either not meeting the criteria for any sight impairment registration, meeting the criteria for sight impairment registration, or meeting the criteria for severe sight impairment registration

Rater	Neither (%)	Sight impaired (%)	Severely sight impaired (%)
1	32.4	52.9	14.7
2	17.6	61.8	20.6
3	35.3	58.8	5.9
4	2.9	41.2	55.9
5	58.8	35.3	5.9
6	38.2	58.8	2.9
7	38.2	50	11.8
8	11.8	47.1	41.2
9	70.6	29.4	0
10	52.9	29.4	17.6
11	20.6	47.1	32.4
12	41.2	55.9	2.9
13	32.4	32.4	35.3
14	41.2	58.8	0
15	35.3	58.8	5.9
16	29.4	70.6	0
17	0	32.4	67.6
18	32.4	58.8	8.8
19	35.3	58.8	5.9
20	32.4	50	17.6
21	11.8	79.4	8.8
22	50	32.4	17.6
23	61.8	38.2	0
24	0	58.8	41.2
25	17.6	41.2	41.2
26	8.8	67.6	23.5
27	5.9	47.1	47.1
28	41.2	55.9	2.9
29	32.4	58.8	8.8
30	38.2	41.2	20.6

Table 2 Trichotomous analysis of inter-rater agreement scores for the three categories of not sight impaired, sight impaired, and severely sight impaired

Sample	Kappa	95% CI	Alpha (ordinal)	95% CI	ICC	95% CI
Overall	0.160	0.101–0.252	0.295	0.167–0.415	0.307	0.212–0.446
Glaucoma	0.226	0.146–0.369	0.386	0.269–0.495	0.399	0.278–0.552
Experienced	0.193	0.118–0.292	0.207	0.065–0.336	0.323	0.133–0.373

ophthalmologists specialising in glaucoma, the kappa and ICC scores were 0.316 and 0.3, respectively, and when looking at those with >10 years consultant experience the Kappa and ICC scores were 0.28 and 0.2, respectively (Table 3).

For assessing the intra-rater analysis, only 16 of the original 30 raters performed the rating for a second time. On average, raters achieved <60% agreement between the first and the second review of the same data (range 29.41–94.12%). Kappa scores ranged from 0.007 to 0.9118 for the 16 raters (Table 4). About 37% of the estimated kappa scores were <0.5.

The subgroup analysis for glaucoma experts and experienced raters found that the probability of a false rating (rating a non-blind as blind) was considerably lower indicating more homogenous ratings. For those with additional expertise in glaucoma, the probability rates ranged from 0 to 19%, whereas the probability for those having higher than 10 years expertise ranged from 0 to 65%.

Discussion

It has been estimated that approximately a third of those eligible for blind registration in the UK are not registered

Table 3 Dichotomous analysis of inter-rater agreement scores for the two categories of not sight impaired and any sight impairment whether severe or not

Sample	Kappa	95% CI	Alpha	95% CI	ICC	95% CI
Overall	0.232	0.142–0.345	0.233	(–0.006 to 0.438)	0.241	0.173–0.344
Glaucoma	0.316	0.194–0.479	0.317	(0.106–0.53)	0.327	0.236–0.451
Experienced	0.28	0.094–0.274	0.170	(–0.062 to 0.39)	0.189	0.119–0.293

Table 4 Intra-rater agreement analysis calculated from the grading results of those raters who completed rating the same data set a second time

Rater	Kappa	% agreement
1	0.441	64.71
2	0.71	85.29
3	0.113	50
4	0.5258	70.59
5	0.835	91.18
7	0.7043	82.35
8	0.007	29.41
9	0.2835	67.65
10	0.658	79.41
13	0.9118	94.12
14	0.4354	73.53
22	0.2082	52.94
23	0.6577	82.35
25	0.2314	50
27	0.6741	82.35
28	0.5318	76.47

and that a third of those registered as blind are inappropriately registered.² A hospital-based cross-sectional survey estimated the rate of blind under registration to be as high as 41%. In the community, this rate has been estimated at 69% by the RNIB.⁴ The fact that patients are less likely to be registered if they predominantly have visual field loss rather than visual acuity loss, may be a consequence of the visual field criteria used for visual impairment registration.^{2,4,5} These criteria are subjective and vague, whereas the criteria for registration on the basis of visual acuity loss are much more objective.

The validity of the registration criteria has not been previously assessed and furthermore there is no absolute standard against which to assess the correctness of a consultant's opinion to register a patient as visually impaired or not. The only way to assess the validity of the process is by measurement of the agreement between consultant ophthalmologists who are tasked with the role of determining eligibility for blind registration in the UK. The loose definition of the visual field loss required for visual impairment registration allows for subjective interpretation by consultants which generates inconsistency in registration of those with primarily visual field loss causing their visual disability. It is difficult to generate strict criteria for registration if there is no gold standard to measure it against, however it should be possible to produce criteria which when applied result in a more consistent agreement among those making registration decisions. Published literature looking at the problem of under-registration has either used an objective measurement of visual field impairment devised by the researchers themselves or the author's subjective assessment of the visual field.^{2,4}

The dramatic discrepancies within the overall results and the intra-rater variability demonstrate that patient access to services will vary largely depending on the ophthalmologist they attend, and the vagaries of his or her decision process for putting patients onto the blind register at the time of the consultation. Despite the clear indication of greater inter-rater agreement among glaucoma specialists as well as consultants with greater than 10-year experience, we believe that the level of agreement was still extremely poor. It also must be borne in mind that when the raters graded the visual fields they were not in the normal clinical situation which is likely to be more time pressured and have more confounding factors that could bias the reliability of a decision to register a patient. It is therefore likely that in the clinical situation the registration process will be even less consistent and prevent patients with visual disability from accessing supporting services to allow them to function independently.

Our results indicate that when presented with sufficient information to decide if a patient with severe visual field loss meets the eligibility criteria for vision or severe vision impairment registration, there is a marked lack of agreement among consultant ophthalmologists between each other and between themselves. This suggests a problematic definition of what different stages of impairment mean to different raters, due to the obvious lack of standardised objective criteria and the absence of gold standards.

Latent class analysis allowed us to group the consultants into two (latent) categories, those who persistently favour blindness outcomes and those who are reluctant to go that way, indicating a subjective bias widening or lowering the thresholds (stages) for the definition of blindness. Furthermore, analysis also indicated that additional training in glaucoma is able to decrease such disagreement by encouraging more homogenous definitions among the raters. Finally, experience does contribute to some extent to this homogeneity but not as strongly as expertise suggesting that glaucoma expertise allows more consistent visual field evaluation, which underpinned the evaluation of registration status in this cohort of data sets.

Alternative methods or criteria for registering patients with predominantly visual field loss should be investigated. Some of the inconsistency may arise from presentation of the visual field data independently from each eye. This requires the assessor to combine the information mentally into an assessment of binocular function to reflect the patient's true visual field function. The ability to do this may vary considerably between raters as it is not a commonly undertaken clinical exercise. A solution would be to utilise readily available software to combine the visual fields electronically into a single representative 'binocular visual field',¹⁵ which can then be assessed for blind registration purposes.

Another option would be to impose more guidance on the visual criteria necessary to register patients as visually impaired. Making criteria more rigid and less open to the interpretation of the examining physician may reduce the variability resulting from subjective interpretation of vague definitions. The United States has an objective federal legal definition of blindness of a limitation in the field of vision such that the widest diameter of the visual field subtends an angle no greater than 20°. However, there is no data to determine if the more objective definition of blindness in the United States increases reporting or improves consistency in registration of visually impaired patients.^{16,17}

Although we have attempted to replicate the information available to consultants when making decisions about sight impairment registration, we have

not replicated the environment of an outpatient clinic in which these decisions are made or the dynamic of the patient interaction. Both of these may potentially influence the decision-making the clinic due to the time pressurised environment for decision-making and other clinic-related distractions and the patient interaction which may bias a clinician one way or another depending upon patient specific factors such as other visible disabilities patient and family requests and a desire to 'help' patients who may be lacking other types of social support.

This study demonstrates extremely poor agreement among consultant ophthalmologists in determining whether a patient with glaucoma and severe visual field loss is eligible for vision impairment registration. This has implications for consistency of access of visually impaired glaucoma patients to supporting services and may influence the accuracy of reported vision impairment rates for patients suffering from glaucoma.¹⁸ The finding that poor agreement exists between observers repeating the grading indicates that the system requires extensive revision.

Summary

What was known before

- Visual impairment is under-recorded in the UK. The factors contributing to this unrecording have not been fully defined.

What this study adds

- The criteria used to determine eligibility for blind registration appear to be overly subjective. This results in lack of consistency among consultant ophthalmologists as to which patients merit blind registration.
-

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- 1 For people who are blind or partially sighted. 2003 Apr 24 [cited 2012 May 23]; Available from: http://webarchive.nationalarchives.gov.uk/+ /www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/browsable/DH_5856365.
- 2 Barry RJ, Murray PI. Unregistered visual impairment: is registration a failing system? *Br J Ophthalmol* 2005; **89**(8): 995–998

- 3 Department for Work and Pensions RH. DWP - A-Z of medical conditions—Vision—Registration of blindness / partially sighted [Internet]. [cited 2012 May 24]. Available from: <http://www.dwp.gov.uk/publications/specialist-guides/medical-conditions/a-z-of-medical-conditions/vision/reg-blind-partial-sight-vision.shtml>.
- 4 King AJ, Reddy A, Thompson JR, Rosenthal AR. The rates of blindness and of partial sight registration in glaucoma patients. *Eye (Lond)* 2000; **14**(Pt 4): 613–619.
- 5 Bunce C, Evans J, Fraser S, Wormald R. BD8 certification of visually impaired people. *Br J Ophthalmol* 1998; **82**(1): 72–76.
- 6 Fleiss JL, Endicott J, Spitzer RL, Cohen J. Quantification of agreement in multiple psychiatric diagnosis. *Arch Gen Psychiatry* 1972; **26**: 168–171.
- 7 Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*. 3rd edn. Wiley-Interscience: Hoboken, NJ, 2003.
- 8 Krippendorff K. Association, agreement, and equity. *Quality Quantity* 1987; **21**: 109–123.
- 9 Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Meas* 2007; **1**: 77–89.
- 10 Fleiss JL, Cohen J. Equivalence of weighted kappa and intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; **33**: 613–619.
- 11 Fleiss JL, Shrouf PE. Approximate interval estimation for a certain intraclass correlation-coefficient. *Psychometrika* 1978; **43**: 259–262.
- 12 Shrouf PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**(2): 420–428.
- 13 Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; **85**(3): 257–268.
- 14 Krippendorff K. Agreement and information in the reliability of coding. *Commun Methods Meas* 2011; **5**: 93–112.
- 15 Asaoka R, Crabb DP, Yamashita T, Russell RA, Wang YX, Garway-Heath DF. Patients have two eyes!: binocular versus better eye visual field indices. *Invest Ophthalmol Vis Sci* 2011; **52**(9): 7007–7011.
- 16 Goldstein H. Blindness register as a research tool. *Public Health Rep* 1964; **79**: 289–295.
- 17 The Eye Diseases Prevalence Research Group. Causes and prevalence of visual impairment among adults in the United States. *Arch Ophthalmol* 2004; **122**(4): 477–485.
- 18 Bunce C, Xing W, Wormald R. Causes of blind and partial sight certifications in England and Wales: April 2007–March 2008. *Eye (Lond)* 2010; **24**(11): 1692–1699.