Diagnosis from fundus photographs

E G Dallas, R A Clement, D S I Taylor

Br J Ophthalmol 2007;**91**:608–612. doi: 10.1136/bjo.2006.105726

Aim: To investigate the way in which ophthalmologists observe fundi and make diagnoses from their observations.

Methods: A set of 12 test photographs was presented to 9 ophthalmologists. The subjects were asked to identify the features in the photographs that are important for forming a diagnosis and were also asked to form differential diagnoses. The scanpaths of the subjects were recorded during their inspection of the photographs. Subsequently, they were asked to trace over the important features of four of the photographs.

Results: The correctness of the diagnoses was described by weighted numerical scores. Differential diagnoses made after 30 s of inspection were significantly better than those made after 5 s. Irrespective of correctness, the reported diagnoses were dominated by the most obvious features of the photograph. Incorrect diagnoses were made either because the subjects failed to identify the significant features of the photograph or because they failed to comprehend the significance of the identified features.

Conclusion: Accurate funduscopy involves both perception of diagnostic features and cognitive interpretation of these features. Verbal reports, eye movement recordings and tracings reveal the features and interpretations used to make a diagnosis. These techniques will be used in a subsequent study to evaluate the relative contributions of formal training and experience to the development of diagnostic skills.

Since the invention of the ophthalmoscope in 1851, funduscopy has played a key role in the diagnosis of eye conditions. Fundus photographs are often included in ophthalmological papers and presentations to disseminate knowledge that can be used to formulate a diagnosis in clinical cases. They are also effective images for probing visual perception because they have both salient features such as the optic nerve head and macula, and considerable variation in the contours, colour and contrast in the image. Because the making of a diagnosis from a fundus photograph is a well-defined task, it also provides an attractive paradigm for investigating how visual perception is linked to cognition.

The most straightforward technique for probing visual cognition is that of asking the subjects to report what they see. Verbal reports can be corroborated by asking the subjects to trace over the important diagnostic features of a photograph. The drawback of these techniques is that subjects are not always aware of the features that they have used in recognising the objects present in the picture. This drawback can be overcome by complementing the verbal reports with eye movement recording, which identifies the relatively small visual areas selected for detailed processing.

Our aim in this study is to identify the aspects of visual perception and cognition relevant to making diagnoses from fundus photographs, in order that these can be emphasised during training. We begin by verifying that the information used in making a diagnosis accrues over prolonged viewing,

and then identify the processes used by ophthalmologists to make the diagnosis.

METHODS

Experimental procedure

Nine ophthalmologists (four women and five men, aged 26–49 years, median 33 years), varying in ophthalmological experience from senior house officer to consultant, were recruited from the Ophthalmology Department, Great Ormond Street Hospital, London, UK. The subjects carried out three tasks.

First, the subjects looked at 12 fundus photographs, presented for 5 s with a 5 s gap between each presentation, and had to voice a diagnosis that was recorded on a dictaphone and transcribed subsequently. Subjects were allowed to voice “not known” as a diagnosis.

Second, the subjects were shown the same fundus photographs, but this time each picture was displayed for 30 s with no gap in between presentations, and their eye movements were recorded. After completion of this task, the subjects were shown the pictures again and asked to describe the features used to form a diagnosis, and were given the opportunity to revise the diagnosis given in the first task. The eye movements of the subjects were recorded throughout tasks 1 and 2.

Finally, the subjects were given a sheet of acetate covering four of the photographs (5, 7, 9 and 11) and were asked to trace over the features that they identified as being relevant to the diagnosis.

Written instructions were given to the subjects at the outset, which described the three tasks, the responses required from the subjects and the approximate times that the tasks would take. Subjects were given the opportunity to ask questions about the procedures before the experimental session began. Ethical approval for the study was obtained from the Institute of Child Health/Great Ormond Street Hospital Research Ethics Committee before commencement.

Eye movement recording and analysis

The subjects viewed the photographs with both eyes open. Head movements were restrained with a chin rest. The movements of the right eye were recorded with a video recording system (Model 504, Applied Science Laboratories, Bedford, Massachusetts, USA), which has a resolution of approximately 0.5° of visual angle. The photographs were displayed on a plasma screen and subtended an angle of 44.8° horizontally and 34.7° vertically. The Eyenal software supplied with the recording system was used to identify the fixation points in the recording of horizontal and vertical eye positions.

A measure of the similarity between the fixation positions of different subjects was provided by the least-squares index introduced to the analysis of scanpaths by Mannan *et al.*¹ In this index, the differences between the fixation points are normalised with respect to the differences, with random distributions of fixation points. If there are no differences between the fixation points of different subjects then the similarity index is 100%, and if the differences are random then the index is 0%.

Scoring of diagnoses

The verbal reports of diagnoses were converted to numerical scores to enable quantitative comparison of the results. The initial score was given by the number of correct differential diagnoses. The score for correctness of diagnosis was weighted to ensure that subjects who gave many incomplete differential diagnoses did not have scores higher than subjects who gave succinct correct diagnoses. A complete differential diagnosis, such as papilloedema, was given a score of 1, and an incomplete differential diagnosis, such as swollen disk, was given a score of 0.5. The scores of differential diagnoses that could occur conjointly (eg, papilloedema and retinal hamartoma) were added, whereas the maximum score was taken of diagnoses that are alternatives (eg, chronic severe papilloedema or neuroretinitis or severe hypertension). The possible values for the differential diagnosis score ranged from 0 to 2.5.

RESULTS

Effect of viewing time

A two-way analysis of variance with repeated measures was used to compare the diagnosis score after 5 and 30 s of inspection. There was a significant difference between the diagnostic scores with the two viewing durations ($F_{11,88} = 10.5$, $p < 0.05$). There was no significant interaction effect between the inspection time and the photograph ($F_{11,88} = 2.31$, NS), which implies that the difference due to viewing time was independent of the photograph being inspected. The average diagnostic score was 18.8% of the maximum after 5 s and 38.6% after 30 s of inspection.

Verbal reports

Table 1 summarises the verbal reports of diagnostic features and differential diagnoses. The choice of differential diagnosis

Table 1 Summary of reported diagnostic features and differential diagnoses

Image	Reported features	Most commonly reported diagnoses	Possible differential diagnoses
1	Swollen disc, exudates Retinal hamartoma (tuberos sclerosi), engorged veins	Papilloedema Hamartoma	Papilloedema Retinal hamartoma (tuberos sclerosi) Optic neuritis
2	Abnormal macula, swollen disc Haemorrhage, exudates, swollen veins	Hypertension Papilloedema, neuroretinitis	Chronic severe papilloedema, neuroretinitis Severe hypertension
3	Choroidal atrophy Abnormal macula, small vessels	Choroideraemia	Choroideraemia Gyrate atrophy Myopia
4	Striae in nerve fibre layer, retinal pigment epithelium mottled with exudates, pale and swollen optic discs Swelling, hazy vessels		Hypertension Posterior uveitis Choroideraemia Retinal vasculitis
5	Haemorrhage, pigmented and flat macula, large amount of optic nerve atrophy Reflection indicating nerve fibre loss	Optic atrophy Diabetic retinopathy	Optic atrophy Diabetic retinopathy Retinal haemorrhages Hypertension Diabetes insipidus, diabetes mellitus, optic atrophy and deafness
6	Choroidal vessels, pale fundus	Albinism	Optic nerve aplasia
7	Pre-retinal fibrosis Pale disc, straight vessels to macula, mild coloboma	Pre-retinal fibrosis	Pre-retinal fibrosis Epiretinal membrane Neurofibromatosis type 2 Trauma
8	Small optic disc Normal	Normal	Optic nerve hypoplasia "Bowtie" hypoplasia Homonymous hemioptic hypoplasia
9	Mottled background Crystalline bodies Normal pigment? angioid streak, pale disc?	Retinal dystrophy	Pseudo Xanthoma Elasticum Angioid streaks
10	Fishtail lesions in retinal pigment epithelium Macular pigment disturbance Pale temporal segment of optic disc	Stargardt's disease/fundus flavimaculatus	Stargardt's disease Fundus flavimaculatus
11	Tiny disk, pale surround Pigmented macula, engorged vessels, normal	Optic nerve hypoplasia	Optic nerve hypoplasia
12	Abnormal macula Streaking, abnormal optic disk, normal		Macular retinoschisis

The reported features and diagnoses have been recorded in decreasing order of frequency (top row 9–7, middle row 6–4 and bottom row 3–1). Incorrect reported diagnoses are in *italics*.

Figure 1 Colour photographs shown to the subjects and the pooled fixations of the subjects. The fixations are shown on grey scale versions of the photographs. The colour of the points corresponds to the diagnostic score of the subject (blue = zero, red = maximum score).

seemed to be dominated by the most obvious feature of the photograph. In image 1, the most commonly reported features of swollen disc and exudates were usually linked to a differential diagnosis of papilloedema. Similarly, in image 7, the commonly reported feature of pre-retinal fibrosis was usually linked to a differential diagnosis of pre-retinal fibrosis. This dominance also held with making of incorrect diagnoses. The commonly reported feature of a pale fundus for photograph 6 led to a majority of incorrect diagnoses of albinism.

Eye tracking

Qualitative inspection of the fixation positions of all subjects, which are plotted in fig 1, show that the most commonly reported features were also frequently fixated. In cases where

there were no commonly reported features, the subjects adopted a default pattern of fixation of the optic disc, macula and major blood vessels (photographs 8, 11 and 12).

To investigate the relationship between diagnostic abilities and choice of fixation positions, for each photograph, the subjects were divided into two groups: those that had above average diagnostic scores and those that did not. This was not possible for photograph 8, as no subject gave a correct differential diagnosis for this case. Only one subject gave an incorrect differential diagnosis for photograph 2 and only one subject gave a correct differential diagnosis for photographs 9 and 12. We, therefore, concentrated on photographs 3, 5, 10 and 11 for which the split of subjects below and above average was approximately the same (4–5).

Figure 2 Superimposed tracings made by the subjects who had above average diagnostic scores are shown in yellow and the tracings of the other subjects are shown in blue.

The eye movement recordings were used to distinguish between photographs where both groups of subjects did and did not use similar patterns of fixation. The similarity index of the eye movement fixation points of the two groups of subjects for photographs 3, 5, 10 and 11 were 3.4%, 13.8%, 0.6% and 16.6%, respectively. For the two photographs with the lowest similarity indexes, the patterns of fixation were different, and it can be seen that the two groups look at different regions of the photographs. In the case of photograph 3, subjects who were unsuccessful in the diagnosis made many fixations on the dark region on the right of the photograph, which is not relevant to the diagnosis. In the case of photograph 10, subjects who made the correct diagnosis made many fixations to the flecks on the retina, especially on the right-hand side of the photograph. With photographs 5 and 11, the patterns of fixation were more similar, with subjects looking at the same regions of the retina.

Tracings

Figure 2 shows the tracings that the subjects made of the important diagnostic features. Differences in the features used to make the diagnosis are clearly discernable in the superimposed tracings; the coloboma and preretinal fibrosis in image 7 and the small disc and abnormal macular pigmentation in photograph 11. Although the eye movement recordings show that all subjects looked at the area of abnormal macular pigmentation, none of the subjects who failed to diagnose the disease considered it to be an important diagnostic feature.

DISCUSSION

We wanted to find out how ophthalmologists analyse fundi and how successful they are in deriving diagnoses from funduscopy. This study was designed to reveal the procedures by which ophthalmologists reach differential diagnoses from inspection of fundus photographs, and any drawbacks in the procedures that could be overcome by training. Three techniques were used: verbal reports of diagnostic features and differential diagnoses, eye movements elicited by the photographs and tracings of the important features of the photographs. More than one technique was used because none of them can be guaranteed to give full information about the procedures. Verbal reports and tracings may not include features that the subject was unaware of using, and eye movement recordings only specify points of fixation whose significance has to be subsequently interpreted.

Verbal reports revealed that the ophthalmologists in our study relied on the most obvious features of the photograph to make their diagnoses. This meant that they could not always be specific in their differential diagnoses. A greater level of expertise was evident in subjects who could make use of specific features relevant to a particular diagnosis. For instance, in the case of photograph 9, pseudoxanthoma elasticum is

almost invariably associated with angioid streaks, so these could be used to identify this diagnosis. Improving an ability involves learning how rather than learning that.² The distinction being made is between the learning of facts, such as the equivalence of the names Stargardt's disease and fundus flavimaculatus, and the gradual acquisition of a skill, such as recognising hypoplasia of the optic disc. A consequence of this distinction is that ophthalmologists seeking to improve their diagnostic ability should place less emphasis on learning facts about as many conditions as possible, and more on acquiring "know-how" about how to test out alternative diagnoses.

Visual fixations are usually described as being concentrated on the informative features of a scene³⁻⁵: those that are unpredictable. Thus, although most of the subjects reported the distinctive horizontal blood vessels running between the optic disc and the macula in photograph 7, they made few fixations to this perceptually unambiguous diagnostic feature. Although the fixation positions are not identical to the positions of all the diagnostic features, the eye movement recording technique enables the comparison of how different subjects attend to a subset of the diagnostic features. A comparison of the similarity of the points of fixation chosen by subjects who made a correct diagnosis and those who did not revealed examples where similar and dissimilar positions were fixated by both groups. A possible interpretation of these findings is that in the case where dissimilar eye movements were made, the subjects who failed to make the correct diagnosis did not identify all regions of the photograph pertinent to the diagnosis. In the case where similar fixations were made, the subjects who subsequently failed to make the correct diagnosis did not interpret the features correctly. Eye movement recording could be used to distinguish between mistaken diagnoses made because the relevant diagnostic features were not identified and because the features were not interpreted correctly, but only when the results of a test can be divided into two groups of subjects of approximately equal size. As there is no a priori way of knowing how difficult subjects will find a particular photograph, we could only use the eye movement recordings in this way for 4 out of 12 of our photographs.

The tracings included all the reported diagnostic features, but also features that were not relevant to the diagnosis, such as the major blood vessels of photograph 5. This may be because such features are perceptually salient and would be included in a line drawing of the photograph made for the purpose of subsequent recognition. In the context of teaching of diagnosis from fundus photographs, an important feature of outline drawings is that subjects can learn to recognise new objects faster from cartoons than from photographs.⁶ The steady development of high-quality digital imagery has led to the quality of a textbook being judged by the quality of the digital imagery it contains,

and the gradual adoption of the image quality criterion has led to a steady reduction in the guidance in interpreting such imagery, which can be given through sketches. Although this trend is appropriate for expert ophthalmologists, it is not appropriate for the trainee, who will benefit from being presented with a sketch of the significant features alongside the photograph. For example, we found that judgement of the boundary of the optic disc is difficult, and current training rests on the trainee acquiring this skill after viewing hundreds of images. If this learning process can be accelerated through guidance by examples, then the burden of training future ophthalmologists would be reduced. This proposal could be tested directly with a web-based distance-learning study.

In common with all medical imaging, interpretation of fundus photographs involves both detection of diagnostic features and decisions about the significance of the features. The relative importance of detection and decision has been investigated in the case of diagnosis of early lung cancer from chest radiographs, and it was found that most of the incorrect diagnoses arose from mistaken decisions not to interpret a feature as a lesion.⁷ To overcome these failures, current research is directed towards optimising the display of the radiographs for cognitive rather than perceptual analysis.⁸ As we have found that both detection and decision errors arise in the interpretation of fundus photographs, this research goal is also appropriate for improving interpretation of ophthalmic imagery.

ACKNOWLEDGEMENTS

We thank our volunteers for their generous cooperation.

Authors' affiliations

E G Dallas, R A Clement, D S I Taylor, Visual Sciences Unit, Institute of Child Health, University College London, London, UK

Competing interests: None.

Correspondence to: Dr R A Clement, Visual Sciences Unit, Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK; r.clement@ich.ucl.ac.uk

Accepted 17 November 2006

Published Online First 13 December 2006

REFERENCES

- 1 **Mannan SK**, Ruddock KH, Wooding DS. The relationship between the locations of spatial features and those fixations made during visual examination of briefly presented images. *Spatial Vision* 1996;**30**:165–88.
- 2 **Ryle G**. *The concept of mind*. Harmondsworth: Penguin, 1949.
- 3 **Buswell T**. *How people look at pictures*. Chicago: University of Chicago Press, 1935.
- 4 **Yarbus A**. *Eye movements and Vision*. New York: Plenum Press, 1967.
- 5 **Mackworth NH**, Morandi AJ. The gaze selects informative details within pictures. *Percept Psychophys* 1967;**2**:547–52.
- 6 **Gooch B**, Reinhard E, Gooch A. Human facial illustrations: creation and psychophysical evaluation. *ACM Trans Graphics* 2004;**23**:27–44.
- 7 **Manning DJ**, Ethell SC, Donovan T. Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *Br J Radiol* 2004;**77**:231–5.
- 8 **Manning DJ**, Gale A, Krupinski EA. Perception research in medical imaging. *Br J Radiol* 2005;**78**:683–5.

Save your favourite articles and useful searches

Use the "My folders" feature to save and organise articles you want to return to quickly—saving space on your hard drive. You can also save searches, which will save you time. You will only need to register once for this service, which can be used for this journal or all BMJ Journals, including the BMJ.