OPEN ACCESS

Clinical science

Expert CONsensus on Visual Evaluation in Retinal disease manaGEment: the CONVERGE study

Roger S Anderson ,[1,2] Mark Roark,[3] Rose Gilbert,[2,4] Dayyanah Sumodhee[2,4]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bjo-2024-325310).

[1]Centre for Optometry and Vision Science, Ulster University, Coleraine, UK
[2]NIHR Moorfields Biomedical Research Centre, London, UK
[3]Allisonville Eye Care Center, Inc., Indiana USA, Allisonville, Indiana, USA
[4]UCL Institute of Ophthalmology, London, UK

Correspondence to
Professor Roger S Anderson; rs.anderson@ulster.ac.uk

Mark Roark passed away shortly before acceptance of the manuscript

Received 14 February 2024
Accepted 15 July 2024

To cite: Anderson RS, Roark M, Gilbert R, *et al*. *Br J Ophthalmol* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjo-2024-325310

ABSTRACT

Background/Aims Recent decades have seen significant advances in both structural and functional testing of retinal disease. However, the current clinical value of specific testing modalities, as well as future trends, need to be clearly identified in order to highlight areas for further development in routine care and clinical trials.

Methods We designed a modified two-round Delphi study to obtain the opinion of a multidisciplinary group of 33 international experts involved in the field of retinal disease management/research to determine the level of agreement and consensus regarding the value and performance of specific structural and functional testing methods for retinal disease. On a Likert scale, a median of 1–2 indicated disagreement with the statement, and 5–6 indicated agreement with the statement. An IQR of ≤2 indicated consensus in the responses. Several questions also allowed comments on responses.

Results There was overall agreement that structural testing currently predominates for detection and monitoring. There was moderate agreement that functional testing remains important and will continue to do so in the future because it provides complementary information. Certain respondents considered that properly designed and applied psychophysical tests are as reliable and repeatable as structural observations and that functional changes are the most important in the long run. Respondents considered future care and research to require a combination of structural and functional testing with strong consensus that the relative importance will depend on disease type and stage.

Conclusion The study obtained important insights from a group of international experts regarding current and future needs in the management of retinal disease using a mix of quantitative and qualitative approaches. Responses provide a rich range of opinions that will be of interest to researchers seeking to design tests for future patient care and clinical trials.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ While recent decades have seen great advances in both structural and functional testing for retinal disease, the relative importance of these modalities, both now and in future, remains poorly explored and understood, with little guidance on what future requirements will be.

WHAT THIS STUDY ADDS

⇒ Using a mix of quantitative and qualitative approaches, this Delphi study acquired the opinions of 33 international experts (ophthalmologists, optometrists, psychophysicists) and found high levels of agreement and consensus about the relative importance of structural versus functional testing, both now and in future, as well as opinions about the future importance of home testing, artificial intelligence and patient reported outcome measures.

HOW THIS RESEARCH MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The study provides timely insights regarding current and future needs/priorities in the management of retinal disease. Responses also provide a rich range of opinions that will be of interest to researchers seeking to design tests for future patient care and clinical trials.

INTRODUCTION

The past two decades have seen important developments in the diagnosis and monitoring of retinal disease. Advances in structural imaging technology, most notably optical coherence tomography (OCT) now permit the in-vivo visualisation and study of retinal and subretinal layers, previously invisible to fundus photography. Along with further structural imaging developments such as OCT angiography and autofluorescence imaging, this has aided both disease detection and monitoring, but also improved understanding of the disease process itself.

Alongside these advances in structural imaging, functional measures of visual integrity have also, in parallel, significantly improved. In fact, functional testing still remains an essential component of retinal disease detection and monitoring in clinical care. Crucially, it measures what the patient cares the most about, namely the preservation of their vision. The current and future relative preference for structural or functional testing in the clinic is therefore very important.

Conventional visual acuity (VA) remains the most widely adopted test of visual function in clinical care and the most commonly adopted endpoint in clinical trials,[1 2] this despite its widely acknowledged high test–retest variability (TRV)[3–7] and poor sensitivity to conditions such as early age-related macular degeneration (AMD) and early retinal disease stages.[8 9] The realisation that even logMAR charts display high

TRV and poor sensitivity to early AMD has led to further attempts to improve VA reliability by different scoring and termination rules,[10 11] computerised control of termination and scoring[12] and the use of high-pass filtered letters.[13 14] Tests to measure disease-specific aspects of visual loss such as contrast sensitivity (CS), low contrast letter acuity, cone function, flicker sensitivity, retinal adaptation and shape discrimination[15] have also been designed. Although these tests displayed early promise in detecting deficits of visual function not detected by conventional VA in clinical trials, they have not yet made their way into routine clinical care. It is therefore crucial to identify and understand the barriers to wider clinical adoption.

CS in particular has existed for more than 40 years and is reported to be more sensitive to subtle changes of visual function than VA[16–18] and to better relate to subjective visual impairment and quality of life.[18–20] Despite its availability in various commercial forms, the test continues to be largely absent from routine clinical care.[21] Whether this is due to poor repeatability[22] or some other combination of factors is not entirely clear.

Similarly, microperimetry is commonly used as an endpoint in trials and research[23] as it is effective at measuring functional loss in conditions such as AMD[24–26] and other retinal conditions.[27] Again, exploring the potential barriers to its wider adoption in a clinical setting is essential.

The shortcomings of VA have led to the adoption of structural imaging as a secondary (surrogate) endpoint in clinical trials. Increasingly there have even been attempts to develop structural primary endpoints[8 28] often because, as Schaal *et al*[28] note, 'it is unrealistic to use visual acuity as a clinical trial endpoint in non-exudative AMD because vision loss takes many years to develop'. More recently there have been attempts to develop combined/composite endpoints which combine different types of functional and structural endpoints.[29] However, Terheyden *et al*[29] point out that there remains no agreement with regards to their implementation.

There has been an increase in studies examining the impact of retinal disease on patient quality of life from both a functional and psychological perspective,[30] and the adoption of questionnaire-based patient-related outcome measures (PROMs).[15 18 26]

Several studies have indicated that patient reports of visual difficulty in low light/low contrast environments are predictive of disease progression across a range of AMD stages.[31 32] However, Finger *et al*[33] state that current PROMs need to be improved to detect or stage disease, and that these are not yet accepted by regulators as clinical endpoints, in part due to insufficient adherence to development guidelines. The existing and potential role of PROMs in clinical care is therefore a question of interest.

We attempted to provide answers to some of these important points by designing a Delphi study during which we obtained the opinion of internationally recognised experts involved in the field of retinal disease management and research. The aim of this study was to determine the level of agreement and consensus to a series of questions to identify current deficiencies in the assessment of retinal disease and how these need to change in the future, especially for the evaluation of novel therapies.

METHODS

An initial literature search was undertaken of research studies involving different retinal or optic nerve diseases that employed various structural and/or functional outcome measures. Conditions included AMD, diabetic retinopathy, optic neuritis (ON), inherited retinal dystrophy, cystoid macular oedema, central or branch retinal vein occlusion as well as several glaucoma studies. The purpose was to initially review the abstracts and/or manuscripts to identify the most common primary and secondary outcome measures employed for various conditions with a view to formulating the questions for the first round of the study. 453 studies/papers were identified and reviewed in total by the authors.

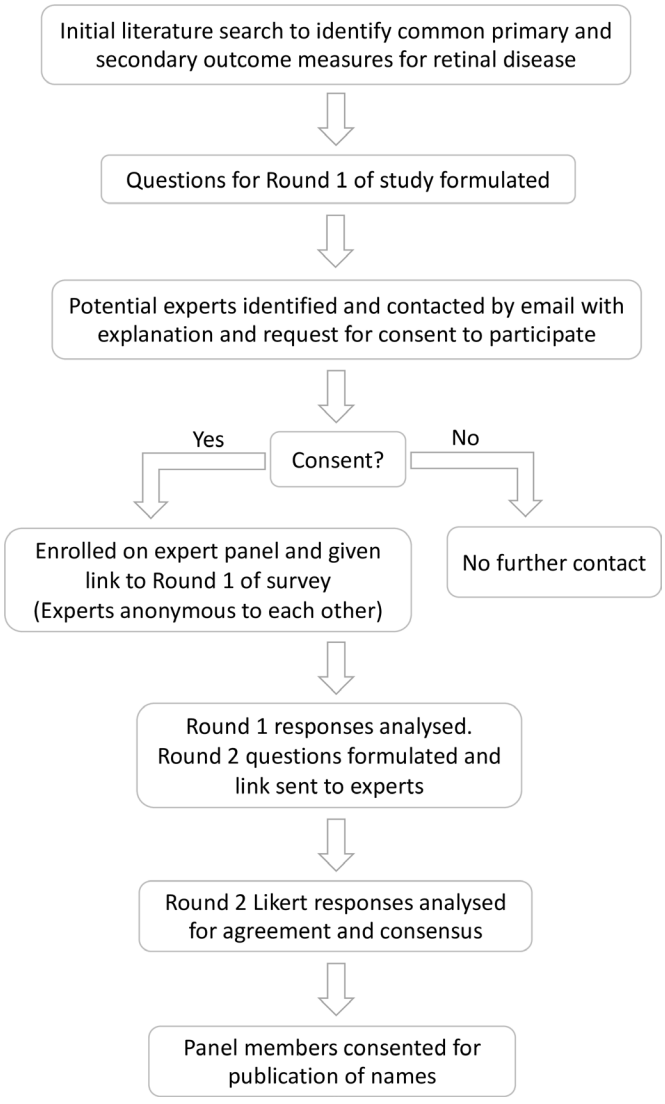

Figure 1 Flowchart demonstrating the Delphi process employed by the study.

Participants

A modified online Delphi study was designed that closely followed the recommendations of previously published 'how to' studies.[34 35] A flowchart of the process can be seen in figure 1. 60 individuals were initially identified by the authors as potential expert participants in the study. Criteria for inclusion were that they were either:

- Senior clinicians (ophthalmologists, optometrists, nurses) involved in the care of retinal disease, including active involvement in research and/or clinical trials of such conditions, evidenced by a significant publication record in the

Table 1 Members of the CONVERGE expert panel

CONVERGE panel of experts	
Aslam, Tariq, MD, PhD Manchester Royal Eye Hospital & University of Manchester, UK	**Mahroo, Omar, PhD, FRCOphth** Moorfields Eye Hospital & UCL, London, UK
Balaskas, Konstantinos, MD Moorfields Eye Hospital, London, UK	**Mukherjee, Rajarshi, MD, FRCOphth** St James's University Hospital, Leeds, UK
Bex, Peter, PhD Northeastern University, Boston, MA, USA	**Mulholland, Pádraig, PhD, MCOptom** Moorfields Eye Hospital & UCL, London, UK
Crabb, David, PhD City, University of London, UK	**Owsley, Cynthia, PhD, MSPH** University of Alabama at Birmingham, AL, USA
Crosby-Nwaobi, Roxanne, RN, PhD Moorfields Eye Hospital & UCL, London, UK	**Patel, Praveen, MD, FRCOphth** Moorfields Eye Hospital, London, UK
De Silva, Samantha, DPhil, FRCOphth Oxford Eye Hospital, Oxford, UK	**Pearce, Ian, MD, FRCOphth** Royal Liverpool University Hospital, Liverpool, UK
Dierker, Damon, OD, FAAO Eye Surgeons of Indiana, Indianapolis, IN, USA	**Peto, Tunde, MD, PhD** Queens University Belfast, UK
Harper, Robert, DPhil, FCOptom Manchester Royal Eye Hospital & University of Manchester, UK	**Rafieetary, Mohammad R, OD, FAAO, FORS** Pres, Optometric Retinal Society/Charles Retina Institute, TN, USA
Hogg, Ruth, PhD, MCOptom Queen's University Belfast, UK	**Robson, John, ScD, FRS** University of Cambridge, UK
Hood, Donald, PhD, FAAAS Columbia University, New York, NY, USA	**Rodman, Julie, OD, MS, FAAO** Nova Southeastern University College of Optometry, Fort Lauderdale, FL, USA
Karpecki, Paul, OD, FAAO Univ. of Pikeville College of Optometry/ Kentucky Eye Institute, KY, USA	**Shah, Nilpa, PhD, MCOptom** Moorfields Eye Hospital, London, UK.
Keane, Pearse, MD, FRCOphth Moorfields Eye Hospital & UCL, London, UK	**Silvestri, Guiliana, CBE, MD, FRCOphth** Royal Victoria Hospital, Belfast, UK
Laidlaw, Alistair, MD, FRCOphth Guy's and St. Thomas's Hospital, London, UK	**Sutton, Brad, OD, FAAO** Indiana University School of Optometry, IN, USA
Latham, Keziah, PhD, FCOptom Anglia Ruskin University, Cambridge, UK	**Tolentino, Michael, MD** Palm Beach Eye Center, Lake Worth, FL, USA
Lotery, Andrew, MD, FRCOphth University of Southampton, Southampton, UK	**Vingrys, Algis, PhD, FAAO** University of Melbourne, Australia
Loughman, James, PhD, FAOI Technical University Dublin, Ireland	

field, or recognised as key opinion leaders in retinal disease management. The majority of candidates were in this group.

- Senior scientific researchers (psychologists, psychophysicists, neurophysiologists) involved in the research and/or design of tests of visual structure or function in retinal disease. Again, these individuals were required to be experienced experts in their field and display significant involvement in clinical retinal research and its publication.

Potential expert panel members were first approached by email to explain the purpose of the study and obtain their consent to participate. They were then given a link to a survey where they could access and respond to the questions for the first round. 33 of the identified experts consented to participate in the study and become part of the CONVERGE expert panel. The names and affiliations of those who consented to be identified are listed in table 1, but for the duration of the study they were anonymous to each other. Of these, 14 were ophthalmologists, 13 were optometrists, 1 was a specialist ophthalmic research nurse, 3 were psychologists specialising in psychophysics and 2 others were from a neurophysiology and quantitative methods background. 22 were senior clinicians specialising in medical retina care and/or research who held the rank of consultant/head of service. Others were clinically qualified academics working in clinical visual research (psychophysics/visual function) related to medical retina. The non-clinicians consisted of senior academics specialising in clinical psychophysics or neurophysiology research. 21 of the total group also held the rank of full professor at their institution.

Summary graphs relating to the expert panel profile can be seen in figure 2. Throughout the duration of the study the responses were collected and curated by staff from the study sponsor and made available to the authors in such a way that individual responses or comments were anonymous. In this way, the authors could not be influenced in their subsequent design or interpretation of questions by responses from individual panel members whom they regarded as authoritative opinion leaders, or with whom they were more personally acquainted.

Round 1 methods

The first round of the study was comprised of more general, open-ended questions to determine the initial panel member opinions and experience with regard to the sort of questions posed in the introduction above, and to identify issues requiring further specific probing.

Questions for this round related to such topics as the importance of structural versus functional testing in the future, any experience of VA being poorly correlated with symptoms/retinal appearance, barriers to the use of functional tests other than VA, reasons why CS is not tested more commonly and how this could be improved. Questions also explored functional tests that panel members considered most valuable in both clinical and research environments, and the importance of PROMs. The panel members were also afforded the opportunity to expand on their responses using comments in text boxes.

Round 2 methods

Based on the responses to round 1, a series of more focused questions was drawn up for round 2 with the goal of determining the level of agreement and consensus to various statements. We followed the design recommendation of Trevelyan and Robinson[35] where panel members were forwarded the (unattributed) responses and comments from all panel member in round 1, and were then asked to indicate their level of agreement to the round 2 questions on a 6-point Likert scale where, for most questions, a score of 1 meant 'strongly disagree' and a score of 6 meant 'strongly agree' with the statement. The responses were quantitatively analysed and a group median of 1–2 was taken to indicate disagreement with the statement, and 5–6 was taken to indicate agreement with the statement. An IQR of ≤2 was taken to indicate consensus in the responses, that is, the degree to which the experts agreed with each other, with 1 taken to mean strong consensus and 2 taken to mean moderate consensus. Several questions also afforded the opportunity to comment on foregoing responses.

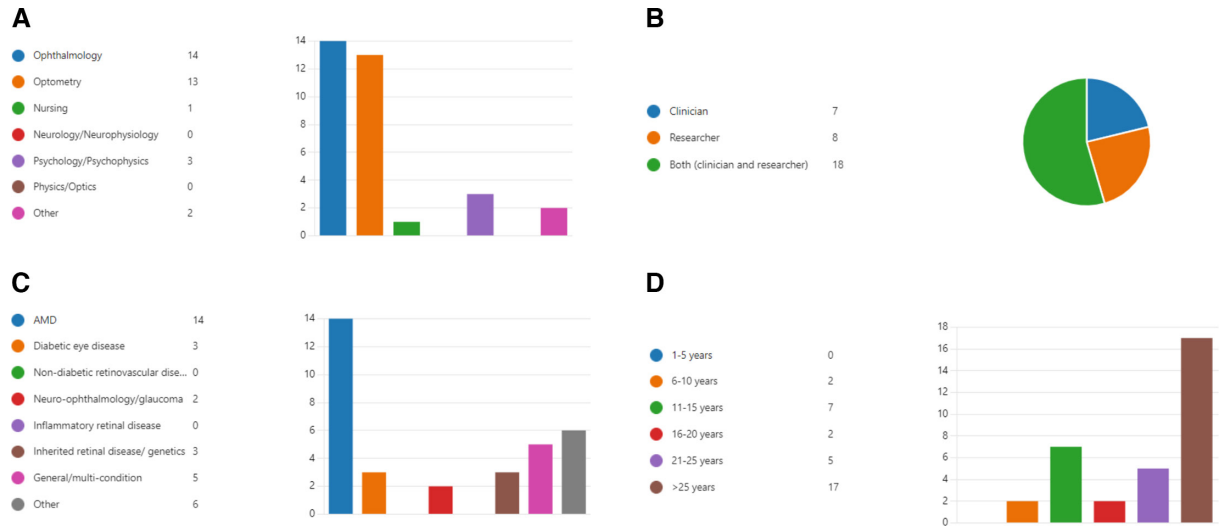

Figure 2 Plots displaying the background and specialist interests of the CONVERGE expert panel, including: (A) professional background, (B) clinical and/or research involvement, (C) specialist area of interest, (D) years in specialty area. AMD, age-related macular degeneration.

RESULTS

Round 1 results

All 33 panel members completed round 1. We did not seek to quantitatively analyse the responses at this stage but the questions, summary of responses and unattributed additional comments can be seen in online supplemental material.

Notable outcomes included:

- Majority agreement (26/33, 79%) that both structural and functional testing will be required in the future. The themes identified in the comments are illustrated with 11 quotes in table 2 (online supplemental material). One of the dominant themes was that panel members considered the two types of testing to each have their relative strengths and weaknesses and could potentially complement each other (quote

Table 2 Selected round 1 and round 2 quotes

Quote number	Round	Quotes
1	1	'Structural testing, especially with the enhancement supported by deep learning will continue to predominate for detection and monitoring of retinal disease.' (P1, Consultant in Ophthalmology)
2	1	'Detection will probably be more based on structural testing such as OCT. For monitoring, structural change will probably remain the medical mainstay, but the importance of functional change in terms of the impact on the person (as assessed by clinical measures, and also PROMs (patient reported outcome measures)) should be incorporated'. (P2, Professor in Optometry)
3	1	'Often visual acuity scores appear unaffected, even in advanced macular degeneration, until the fovea is affected. Another problem is that test retest variability with visual acuity measurements, reported to be higher in those with ocular disease, makes it more challenging to correlate with visual symptoms and/or disease severity'. (P3, Professor in Optometry)
4	1	'Functional measures require attention to measurement techniques and quantitative detail with which most clinicians are unfamiliar. A major problem is variability of function within the normal population, especially if the testing situation and the testing methods are not as well controlled as they need to be. Longitudinal measurements on individuals are likely to be significantly more reliable and informative than single determinations and comparisons with (dubious) norms'. (P4, Professor Neurophysiology) 'Poor understanding of normative values, lack of reimbursement, and weak recommendations from professional organizations regarding clinical utility of CS testing'. (P5, Professor in Optometry)
5	1	'Visual function testing assesses impairment. PROMs assess activity limitation, or the impact the impairment has on the person. Both are necessary to understand the patient's needs and impact of any interventions'. (P6, Professor in Optometry)
6	1	'Modality most appropriate for detection and monitoring depends upon stage of disease. Early and very early disease is much more likely to be picked up by functional testing (but not visual acuity)'. (P4, Professor in Neurophysiology)
7	2	'Some of my previous research points to high variability of functional tests in patients with macular disease (particularly in AMD). This variability is heightened in busy clinical settings. I therefore place more reliance on structural tests or at least i look hard for correlates of functional change in imaging test results'. (P6, Medical Retina Consultant)
8	2	'Functional tests are often the domain of specialised clinics. This should not be so. They should be simple to administer, easy to perform and cheap. Because they will need to be repeated often they should have the capacity for self administration in order to reduce clinical chair time'. (P7, Professor in Optometry)
9	2	'The Pelli Robson CS chart is probably the most widely used test to measure CS in a clinical setting, but it can be difficult to illuminate it evenly and consistently. It also measures CS in one spatial frequency so could potentially miss deficits with certain diseases and test-retest variability may be an issue due to the coarse measurement scale. More comprehensive CS measurements take a long time and may require special equipment'. (P8, Optometrist)
10	2	'Elderly patients struggle with new technologies' I've been involved in testing this previously so I don't think it will be useful'. (P9, Professor in Optometry)
11	2	'The monitoring of chronic eye disease will demand greater involvement as the general population ages and places greater demands one hospitals. To prevent this situation from leading to sight loss due to inappropriate review schedules home monitoring must be adopted to identify patients with true change in need of hospital visit against time who could have their visit delayed because they are stable at present'. (P7, Professor in Optometry)

1). Generally, most agreed that structural testing would predominate over functional to detect early stage disease (quotes 1, 2 and 3). Nevertheless, their respective usage varied depending on the type and stage of the disease with some disagreement as to which one performs better in the early stages (quote 2 vs quote 6).

- Unanimous agreement that VA often poorly correlates with symptoms and/or disease severity.
  Panel members further commented that VA does not necessarily mirror patients' reported day-to-day visual function. They noted that a poor correlation is more likely to be found in certain diseases where the regions of the retina affected differ from that determining VA, usually the fovea (quote 3).
- A majority of panel members agreed with the potential of CS to improve understanding of visual function in retinal disease due to its high sensitivity to early vision loss and gradual change in visual function. Nevertheless, various barriers were given for why it is not as widely used. They highlighted the complexity of employing the test including staff training, repeatability, lack of standardisation and time constraints (quote 4).
- A wide range of functional tests, such as VA, colour vision, CS, visual field and microperimetry, are used in clinical care and research/clinical trials. However, the relative ranking of the tests in terms of importance is different in clinical care compared with research/clinical trials. For example, colour vision is considered more important in a clinical care setting, whereas in research/clinicals it is CS (round 1 online supplemental material, questions 13 and 15). Various other tests were suggested as important such as electrophysiology, reading function and low luminance VA.
- Most respondents considered PROMs to be important in future assessment but the level of importance ranged from 'essential' to 'fairly unimportant'. They further explained in the comments that PROMs, a subjective measure of the visual impairment impact on patients, should be used in conjunction with visual function testing (quote 5).

Round 2 results

All 33 experts from round 1 also participated in round 2. The complete round 2 questionnaire with responses displayed graphically can be found in online supplemental material. The median and IQR results of the Likert responses are displayed in table 3, with selected quotes in table 2 (online supplemental material). Missing question numbers in the table relate to questions asking for follow-up comments.

Structural versus functional testing for detection and monitoring

Questions 2–16 dealt with the relative importance of structural versus functional testing both now and in the future. For question 2, there was moderate agreement and consensus among panel members that structural testing dominated *current* detection and monitoring of retinal disease but not necessarily future (median 2.5–3). Panel members seemed to have greater confidence in structural than functional testing as it offers more reliable and objective testing (Q4 and quote 7). There was strong consensus for current dominance of structural testing (Q2.a,b, median 2, IQR 1), but more moderate consensus for the future (Q2.c,d, median 2.5/3, IQR 2).

The notion that the dominance of structural testing would increase in the future with the inclusion of artificial intelligence (AI) reached agreement and consensus, but affordability, software interface and compactness of equipment will remain limitations (Q5).

There was agreement and consensus that functional testing remains essential for some conditions that structural testing cannot detect (Q7) but, conversely, panel members did not agree that structural testing has reached its zenith, thus requiring new functional tests to detect changes that are invisible to structural testing (Q8). Panel members highlighted the need to improve functional testing to make it easier to administer and more cost effective (Q10).

There was agreement and consensus that functional testing provides information that is complementary to structural testing (Q9) and that future detection (Q10) and monitoring (Q11) will require both for a fuller picture of retinal health. There was also agreement and consensus that the relative value of each will depend on both the disease type (Q12) and particularly stage (Q13).

The idea that the value of either test type was dependent on its ability to assess the efficacy of future treatment also reached agreement (Q14). There was also agreement and strong consensus that the relative value of each could change as new tests of each type are developed (Q15).

Interestingly, the notion that more emphasis should be placed on developing new functional rather than structural tests showed moderate agreement and consensus (Q16).

Visual acuity

Question 18 explored the reasons why VA did not always correlate with patient symptoms or retinal appearance. Panel members suggested three reasons for this, namely TRV, relative vulnerability to optical blur/retinal changes and the localised nature of retinal disease. Only the last two reasons reached agreement with moderate consensus.

Contrast sensitivity

Questions 20–24 investigated the role of CS testing. Question 20 interrogated the *potential* versus current usage of CS which is less wide than VA in *clinical practice*. There was agreement and consensus that it was time-consuming and required special equipment (Q20.a), that it was non-specific and thus susceptible to other conditions such as cataract (Q20.c), that it is poorly understood by clinicians (Q20.d), and that current tests lack uniformity in design and testing methods (Q20.f). This latter suggestion (Q20.f) was the only answer that had an IQR of 0. The idea that either it has poor repeatability (Q20.b) or that target sizes do not correspond to patient difficulty did not reach agreement (Q20.e and quote 9).

In question 22 panel members elaborated on how CS might change if these problems were resolved, in particular the possibility that CS could detect early change not seen by imaging (Q22.a), monitor advanced retinal disease (Q22.b), predict real-world impact on patient tasks (Q22.c) or better correlate with structural change (Q22.d). Agreement was reached only for Q22.c (median 5, IQR 2).

Question 24 examined why CS is more commonly used in clinical *research or trials*. There was agreement and consensus that this was because it was easier to administer in research because there is more time to devote to careful CS testing (Q24.a) but not necessarily that researchers have a better understanding of CS testing than the average clinician (Q24.b), or because clinical trial subjects are carefully selected (Q24.c).

Other functional tests

Question 26 asked about the importance of other functional tests in the management of retinal disease, both now and in future.

Table 3 Summary responses to questions from round 2, indicating the median (agreement) and IQR (consensus) of the responses (n=33)

Question	Median	IQR
Structural vs functional testing		
2. Please indicate the extent to which you consider one to predominate over the other in (1—strongly favour structure, 2—strongly favour function)		
a. The *current* early detection of retinal disease	2	1
b. The *current* monitoring of retinal disease	2	1
c. The *future* early detection of retinal disease	2.5	2
d. The *future* monitoring of retinal disease	3	2
4. Structural testing is more dependable than functional testing due to its ability to gather repeatable, objective data, and will continue to be so in the future	5	2
5. Structural testing will increasingly dominate in both early detection and in monitoring progression of known retinal disease because it will continue to improve by way of		
a. Resolving power	5	1
b. Availability/affordability	4	1
c. The implementation of artificial intelligence/deep learning	5	1.25
d. Software interface	4	1.5
e. Equipment size/compactness	4	2
7. Functional testing is essential for the detection of some retinal diseases that current structural testing cannot detect	5	2
8. Structural testing has reached its zenith and current, new or better functional tests will be required to detect changes that imaging devices will never see	3	2
9. Functional testing provides information that is different and complementary to that provided by structural testing in many retinal diseases	5	1
10. The future *detection of early retinal disease* will continue to require both structural and functional testing to provide a complete picture of retinal health	5	2
11. The future *monitoring of retinal disease* will continue to require both structural and functional testing to provide a complete picture of retinal health	5	2
12. The value of structural or functional testing will depend on the *disease* in question	5	2
13. The value of structural or functional testing will depend on the *stage* of disease, and both are often important in staging	5	1
14. The value of structural or functional testing will depend on the ability to determine the *efficacy of any current/future treatment* for the disease	5	1
15. The value of structural or functional testing will change as new testing methods are developed, with one group perhaps overtaking the other	5	1
16. Overall, more emphasis should be placed on research and development of new and better functional rather than structural tests in retinal disease management	4	2
Visual acuity		
18. Round 1 of the study indicated strong agreement that high-contrast VA does not always correlate well with visual symptoms or retinal appearance. Is this		
a. Because its test–retest variability is high?	4	2
b. Because high-contrast letter targets are more vulnerable to optical changes (blur/cataract) but less vulnerable to retinal changes associated with disease?	5	2
c. Because of the localised nature of damage from retinal disease?	5	2
Question	**Median**	**IQR**
Contrast sensitivity		
20. Round 1 of the study indicated quite strong agreement that contrast sensitivity (CS) had the *potential* to improve understanding of visual function in patients with retinal disease, even though it is not as widely used in clinical practice as VA. Thinking in terms of *clinical practice*, is this because		
a. CS is time-consuming and complicated to set up and measure properly?	5	1
b. CS has poor repeatability?	4	1
c. CS is non-specific and so susceptible to other factors such as cataract?	5	1
d. CS is poorly understood by clinicians making management decisions (eg, in terms of what is 'normal')?	5	1
e. Current tests do not reliably measure CS at relevant target sizes where patients experience most difficulty?	4	1
f. Current tests lack uniformity in testing methods such as background illumination, target type (sine waves vs. letters), and spatial frequency (target size), or nomenclature to display results?	5	0
22. If the above problems were resolved, to what extent do you believe CS would become important to		
a. Detect the earliest changes in retinal disease, perhaps not yet seen by imaging?	4	2
b. Monitor advanced retinal disease	4	3
c. Predict patient difficulties with real-world tasks?	5	2
d. Correlate more consistently with structural changes?	4	2
24. Round 1 of the study indicated that contrast sensitivity (CS) is perhaps the most widely used 'non-VA' functional test in *research or clinical trials*. Is this because		
a. There is more time to devote to careful CS testing in research/trials?	5	2

Continued

Table 3 Continued

Question	Median	IQR
b. Researchers have a better understanding of the value of CS than the average clinician?	4	1
c. Patient subjects are more carefully selected for research trials?	3	2
Other functional tests **26. Please indicate to what degree you consider the following tests to be important (either now or in future) in the management of retinal disease**		
a. Reading speed/function.	5	1
b. Retinal densitometry.	3	2
c. Critical flicker frequency (CFF)	3	2
d. Electrophysiology (eg, VEP/FF ERG/PERG/Multifocal ERG)	4	2
e. Low luminance/contrast VA	5	1
f. Glare disability testing (effects on CS)	4	2
Home testing **28. Several respondents in round 1 indicated support for increased home testing in the future Please indicate the degree to which you think the following will be important in the future**		
a. Increased use of *functional* testing (eg, using tablet, VR headset) for home monitoring	5	1
b. Increased use of *structural* testing (eg, smart-phone camera, hand-held OCT) for home monitoring	5	2
PROMs **30. Patient reported outcome measures (PROMs) have increased in usage in recent years as a way of evaluating 'real world' problems experienced by patients Please indicate the degree to which you agree with the following statements**		
a. PROMs are essential in detecting functional problems that other structure/function tests often miss	4	1
b. PROMs are essential in tailoring treatment regimens to individual patient needs	4	3
c. The value of PROMs will continue to increase as management becomes more personalised	4	2
d. PROMs provide little extra information to help the clinician treat the patient more effectively as long as appropriate functional and structural testing has been done	3	2
Unless indicated, the Likert scale ranges from 1 (strongly disagree) to 6 (strongly agree).		

Only reading speed and low luminance/contrast VA reached agreement and consensus. There was no agreement for retinal densitometry, CFF, electrophysiology or glare testing.

Home testing

For question 28, there was agreement and consensus that there would be a future increase in both functional (Q28.a) and structural testing (Q28.b) in the home in the future, using technologies such as tablet/VR and smart-phone/hand-held OCT, respectively. Although, this may lead to digital exclusion of elderly people (quote 10), home monitoring offers the opportunity to ease the burden on the hospital and prevent better sight loss (quote 11).

Patient-related outcome measures

Question 30 indicated no agreement with regard to the use of PROMs to either (a) detect problems that other tests miss, (b) tailor treatment regimens to individual patient needs or (c) the notion that they will become more valuable as care becomes more personalised. However, responses did not agree that PROMs provide little additional clinical information over current structure/function testing.

DISCUSSION

The study was successful in gaining the opinion of an expert group of clinicians and scientists, from a range of different professional groupings. The panel responses were analysed for both agreement, defined as whether the experts agreed with the question as it was posed, and consensus, indicating the degree to which they agreed with each other. Both are important but the latter is, we believe, more indicative of present and future trends.

Structure versus function

The overall position of the panel seems to be that structural testing currently holds sway in terms of detection and monitoring, owing to its less variable results, but functional testing remains important. Clinicians (ophthalmologists and optometrists) strongly agreed that structural testing is more dependable and repeatable, whereas psychophysicists were not so sure.

However, several respondents had a strong opinion that imaging will in future be able to predict function.

On the other hand, others consider that robustly designed and applied psychophysical tests are as reliable and repeatable as structural observations and that functional changes are the most important in the long run. One respondent noted that structural testing cannot yet detect colour vision changes. Some respondents acknowledged that there was a lack of research and development into practical functional tests, particularly tests that can detect cellular dysfunction rather than cell death.

Overall, panel members considered future care and research to require a combination of structural and functional testing as these provide complementary information, with strong consensus that the relative importance will depend on disease type and stage.

There was consensus that AI would continue to grow in importance and several comments related to this, with some panel members responding that AI could potentially help provide better prediction of function from structure and also understanding of individual disease progression.

The optometrists and psychophysicists were in strong agreement that future detection of retinal disease will require a combination of structural and functional testing, whereas

the ophthalmologists did not quite reach agreement. It may be that the clinicians, and ophthalmologists in particular, are more familiar with, and better trained in, structural rather than functional testing.

Specific functional tests

Visual acuity

Despite its recognised limitations, no panel member commented that VA would become less important as a functional test in the future. Responses indicated its main limitations to be its lack of ability to discriminate between neural and optical losses of vision, and its poor correlation with symptoms/retinal appearance. The clinicians were close to agreement that the poor correlation between VA, visual symptoms and retinal appearance is because of VA's poor TRV, whereas the psychophysicists were much closer to disagreement. Comments commonly referred to the lack of localisation of retinal disease compared with the foveally dominant nature of VA as a test.

Nonetheless, a poor correlation between two tests may merely indicate that they are each measuring a different aspect of vision and are providing complementary information about ocular health. However, a VA test that employs a stimulus that better taps into the parafoveal damage associated with early AMD would appear to be very welcome.

Contrast sensitivity

The results and comments from both rounds of the study indicate that CS has the potential to provide significant further understanding of retinal disease and its impact, but is currently let down by its perceived complicated technical nature, lack of testing uniformity and susceptibility to optical/cortical changes. Several comments also alluded to the poor resolution of early stage disease.

Comments also indicated that it has the ability to provide additional clinical information and better predict real world task performance, but remains more associated with clinical trials where there is more time for testing and more technical assistance with set-up. Many clinicians perhaps continue to lack understanding of the test.

Other functional tests

Reading speed and low luminance VA seemed more significant than others with one comment highlighting the importance of understanding what people see beyond distance VA. The variation in answers was perhaps not surprising given the variation in disease specialty of the panel members, and the likelihood that different tests are more appropriate for different retinal diseases, for example, electrophysiology.

Home testing

There was agreement and consensus that home testing will likely become more prevalent in the future as both structural and functional tests become more compact and suitable for patient operation.

Some respondents who previously considered structural tests to have greater value considered functional home testing to be important in future. Others, however, considered home structural testing to be the future solution to reducing clinic visits and better detecting change.

Patient-related outcome measures

There was consensus that PROMs will be of limited importance in future in terms of detecting and managing patient problems. This result may be due to the fact that most panel members are clinicians or psychophysicists who by training rely more on their examinations than what the patient reports. There were no noticeable differences in the responses between the various professional groups.

Strengths and weaknesses of the study

This study managed to obtain important insight from a group of international experts regarding current and future needs in the management of retinal disease using a mix of quantitative and qualitative approaches. The strength of a Delphi study is that it can rapidly identify the core problems/solutions without resorting to lengthy experimental research or clinical trials. The weakness is that it relies on subjective opinions, and even experts are not always correct. Several of the experts engaged in interesting discussions before consenting to take part. One concern expressed was that smaller discussion panels or working groups, typically convened to define guidelines are often dominated by a few prominent individuals with strong opinions. To address this risk, our study, (a) recruited a large panel of experts, (b) the panel members were unlikely to be influenced by others as they remained unknown to each other for the duration of the study. Another strength of our study was that we included experts representing all relevant professional groupings and specialist interests, allowing for a diverse range of opinions.

CONCLUSION

Our findings suggest a moderate overall agreement that structural rather than functional testing dominates current practice, but not necessarily future practice. Overall, respondents tend to favour the tests they are most familiar with. Clinicians (ophthalmologists, optometrists) tend to favour structural testing in the future, whereas those with a psychophysical background tend to favour functional testing. Trained psychophysicists may see future solutions to current functional testing limitations and the potential for more targeted tests, particularly for different conditions and disease stages, and thus substantial untapped potential in functional testing. New improved tests may have subsequently come on stream, and will continue to do so in the future, and the need for expert psychophysical understanding and input will only grow if appropriate decisions are to be made. Most of our respondents also agreed that AI and home testing will become increasingly relevant in retinal disease management.

Acknowledgements The authors would like to specifically thank Sophie Fairweather and Grace Evans for their support in facilitating meetings and assisting with administrative details.

Contributors RSA, RG and MR were involved in the initial design of the study and formulation of the questionnaires and recruitment of the subjects. All authors were involved in the analysis of the results and the writing of the paper. RSA is the guarantor.

Competing interests The study was funded and facilitated by Alliance Pharmaceuticals Ltd. The authors received financial payment from the funder, but the study objectives, design, analysis and subsequent write-up was entirely undertaken by the authors. Alliance Pharmaceuticals Ltd played no part in the study design, the formulation of the questions, or the analysis and respondents were not paid for their responses. The online questionnaires were administered by the funder and the data curated by Alliance staff to render the authors blind to the identity of respondents until final analysis.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and the methods for this study were reviewed and approved by the institutional legal and ethics committee of Alliance Pharmaceuticals Ltd. All participating experts granted personal approval for their name to be included in the manuscript. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD

Roger S Anderson http://orcid.org/0000-0003-3981-707X

REFERENCES

1 Coassin M. Dry AMD: the regulatory view. *EMA Ophthal Workshop*; 2011.
2 Lesmes LA, Jackson ML, Bex P. Visual function endpoints to enable dry AMD clinical trials. *Drug Discov Today Ther Strateg* 2013;10:e43–50.
3 Rosser DA, Cousens SN, Murdoch IE, *et al*. How sensitive to clinical change are ETDRS logmar visual acuity measurements? *Invest Ophthalmol Vis Sci* 2003;44:3278.
4 Yu HJ, Kaiser PK, Zamora D, *et al*. Visual acuity variability: comparing discrepancies between snellen and ETDRS measurements among subjects entering prospective trials. *Ophthalmol Retina* 2021;5:224–33.
5 Patel PJ, Chen FK, Rubin GS, *et al*. Intersession repeatability of visual acuity scores in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2008;49:4347.
6 Lovie-Kitchin JE, Brown B. Repeatability and intercorrelations of standard vision tests as a function of age. *Optom Vis Sci* 2000;77:412–20.
7 falkenstein IA, cochran DE, azen SP, *et al*. Comparison of visual acuity in macular degeneration patients measured with snellen and early treatment diabetic retinopathy study charts. *Ophthalmology* 2008;115:319–23.
8 Csaky K, Ferris F, Chew EY, *et al*. Report from the NEI/FDA endpoints workshop on age-related macular degeneration and inherited retinal diseases. *Invest Ophthalmol Vis Sci* 2017;58:3456.
9 Owsley C, Huisingh C, Clark ME, *et al*. Comparison of visual function in older eyes in the earliest stages of age-related macular degeneration to those in normal macular health. *Curr Eye Res* 2016;41:266–72.
10 Carkeet A. Modeling logmar visual acuity scores: effects of termination rules and alternative forced-choice options. *Optom Vis Sci* 2001;78:529–38.
11 Shah N, Dakin SC, Whitaker HL, *et al*. Effect of scoring and termination rules on test–retest variability of a novel high-pass letter acuity chart. *Invest Ophthalmol Vis Sci* 2014;55:1386.
12 Laidlaw DAH, Tailor V, Shah N, *et al*. Validation of a computerised logmar visual acuity measurement system (complog): comparison with ETDRS and the electronic ETDRS testing algorithm in adults and amblyopic children. *Br J Ophthalmol* 2008;92:241–4.
13 Shah N, Dakin SC, Redmond T, *et al*. Vanishing optotype acuity: repeatability and effect of the number of alternatives. *Ophthalmic Physiol Opt* 2011;31:17–22.
14 Shah N, Dakin SC, Dobinson S, *et al*. Visual acuity loss in patients with age-related macular degeneration measured using a novel high-pass letter chart. *Br J Ophthalmol* 2016;100:1346–52.
15 Hogg RE, Chakravarthy U. Visual function and dysfunction in early and late age-related maculopathy. *Prog Retin Eye Res* 2006;25:249–76.
16 Preti RC, Ramirez LMV, Pimentel SLG, *et al*. Single intravitreal bevacizumab injection effects on contrast sensitivity in macular edema from branch retinal vein occlusion. *Arq Bras Oftalmol* 2012;75:29–32.
17 Preti RC, Ramirez LMV, Pimentel SLG, *et al*. Effect of a single intravitreal bevacizumab injection on contrast sensitivity and macular thickness in eyes with macular edema from central retinal vein occlusion: a prospective, nonrandomized, three-month follow-up study. *Ophthalmic Res* 2014;51:140–5.
18 Pondorfer SG, Terheyden JH, Heinemann M, *et al*. Association of vision-related quality of life with visual function in age-related macular degeneration. *Sci Rep* 2019;9:15326.
19 Ivers RQ, Mitchell P, Cumming RG. Visual function tests, eye disease and symptoms of visual disability: a population-based assessment. *Clin Exp Ophthalmol* 2000;28:41–7.
20 Rubin GS, Bandeen-Roche K, Huang GH, *et al*. The association of multiple visual impairments with self-reported visual disability: SEE project. *Invest Ophthalmol Vis Sci* 2001;42:64–72.
21 Latham K. Who uses contrast sensitivity in optometric practice? *Ophthalmic Physiol Opt* 1998;18 Suppl 1:S2–13.
22 Patel PJ, Chen FK, Rubin GS, *et al*. Intersession repeatability of contrast sensitivity scores in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2009;50:2621.
23 Yang Y, Dunbar H. Clinical perspectives and trends: microperimetry as a trial endpoint in retinal disease. *Ophthalmologica* 2021;244:418–50.
24 Wu Z, Ayton LN, Luu CD, *et al*. Longitudinal changes in microperimetry and low luminance visual acuity in age-related macular degeneration. *JAMA Ophthalmol* 2015;133:442–8.
25 Steinberg JS, Fitzke FW, Fimmers R, *et al*. Scotopic and photopic microperimetry in patients with reticular drusen and age-related macular degeneration. *JAMA Ophthalmol* 2015;133:690–7.
26 Mehat MS, Sundaram V, Ripamonti C, *et al*. Transplantation of human embryonic stem cell-derived retinal pigment epithelial cells in macular degeneration. *Ophthalmology* 2018;125:1765–75.
27 Horie S, Corradetti G, Esmaeilkhanian H, *et al*. Microperimetry in retinal diseases. *Asia Pac J Ophthalmol (Phila)* 2023;12:211–27.
28 Schaal KB, Rosenfeld PJ, Gregori G, *et al*. Anatomic clinical trial endpoints for nonexudative age-related macular degeneration. *Ophthalmology* 2016;123:1060–79.
29 Terheyden JH, Schmitz-Valckenberg S, Crabb DP, *et al*. Use of composite end points in early and intermediate age-related macular degeneration clinical trials: state-of-the-art and future directions. *Ophthalmologica* 2021;244:387–95.
30 Taylor DJ, Jones L, Binns AM, *et al*. "You've got dry macular degeneration, end of story": a qualitative study into the experience of living with non-neovascular age-related macular degeneration. *Eye (Lond)* 2020;34:461–73.
31 Ying G-S, Maguire MG, Liu C, *et al*. Night vision symptoms and progression of age-related macular degeneration in the complications of age-related macular degeneration prevention trial. *Ophthalmology* 2008;115:1876–82.
32 Ying G-S, Maguire MG. Development of a risk score for geographic atrophy in complications of the age-related macular degeneration prevention trial. *Ophthalmology* 2011;118:332–8.
33 Finger RP, Schmitz-Valckenberg S, Schmid M, *et al*. MACUSTAR: development and clinical validation of functional, structural, and patient-reported endpoints in intermediate age-related macular degeneration. *Ophthalmologica* 2019;241:61–72.
34 Hsu C-C, Sandford BA. The delphi technique: making sense of consensus. *Pract Assess Res Eval* 2007;12.
35 Trevelyan EG, Robinson PN. Delphi methodology in health research: how to do it? *Eur J Integr Med* 2015;7:423–8.