

Patient-reported outcome measures in presbyopia: a literature review

Garima Sharma ,¹ Sima Chiva-Razavi,² Daniel Viriato,² Christel Naujoks,² Francesco Patalano,² Sarah Bentley,³ Amy Findley,³ Chloe Johnson,³ Rob Arbuckle,³ James Wolffsohn ⁴

To cite: Sharma G, Chiva-Razavi S, Viriato D, *et al.* Patient-reported outcome measures in presbyopia: a literature review. *BMJ Open Ophthalmology* 2020;**5**:e000453. doi:10.1136/bmjophth-2020-000453

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjophth-2020-000453>).

Received 13 February 2020
Revised 18 May 2020
Accepted 12 June 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Novartis Healthcare, Hyderabad, India

²Novartis Pharma AG, Basel, Switzerland

³Adelphi Values Patient-Centered Outcomes, Bollington, UK

⁴Optometry, Aston University, Birmingham, UK

Correspondence to

Ms Sima Chiva-Razavi; sima.chiva-razavi@novartis.com

ABSTRACT

Presbyopia is the age-related loss of near-distance focusing ability. The aim of this study was to identify patient-reported outcome measures (PROMs) used in clinical trials and quality-of-life studies conducted in individuals with presbyopia and to assess their suitability for use in individuals with phakic presbyopia. Literature searches were performed in Medline and Embase up until October 2017. Specific search terms were used to identify presbyopia studies that included a PROM. All clinical trials with PROM-supported endpoints in presbyopia were identified on ClinicalTrials.gov. Further searches were conducted to retrieve articles documenting the development and psychometric evaluation of the PROMs identified. A total of 703 records were identified; 120 were selected for full-text review. Twenty-one clinical trials employed PROMs to support a primary or secondary endpoint. In total, 13 PROMs were identified; a further 23 publications pertaining to the development and validation of these measures were retrieved. Most PROMs were developed prior to release of the Food and Drug Administration (FDA) 2009 patient-reported outcome guidance and did not satisfy regulatory standards. The Near Activity Visual Questionnaire (NAVQ) was identified as the most appropriate for assessing near-vision functioning in presbyopia. While the NAVQ was developed in line with the FDA guidance, the items do not reflect changes in technology that have occurred since the questionnaire was developed in 2008 (eg, the increase in smartphone use), and the measure was not validated in a purely phakic presbyopia sample. Further research is ongoing to refine the NAVQ to support trial endpoints related to changes in near-vision functioning associated with phakic presbyopia.

BACKGROUND

Presbyopia is a common age-related vision disorder characterised by a progressive inability to focus on near objects. Presbyopia is hypothesised to be caused by either a weakening of the ciliary muscles or a loss of lens elasticity preventing focal point change, the latter considered the primary causative mechanism. This condition is found to occur both in individuals with intact natural crystalline lens (phakic presbyopia), as well as those who have undergone an invasive surgical procedure involving the extraction of the natural crystalline lens (pseudophakic presbyopia).^{1–4}

It was estimated there were 1.3 billion people living with presbyopia worldwide in 2017, which is projected to increase to 1.8 billion by 2050.⁵ The lens of the human eye is usually able to change shape in order to focus light onto the retina, enabling individuals to see objects at both near and far distances. The lens decreases in flexibility throughout life until, after the age of 40 years, it cannot change shape easily, leading to difficulty focusing on near-distance objects and performing near-vision activities, such as reading or threading a needle.¹ There are currently no approved therapies that reverse the normal ageing process that causes presbyopia; instead, current management focuses on either optical correction with medical devices (eg, spectacles, contact lenses) or surgical intervention (eg, corneal inlay, corneal refractive procedures and intraocular lens (IOL) replacement).⁶ Until recently, the assessment of treatment outcomes in both clinical trials and clinical practice has primarily relied on clinician-reported biomedical parameters, mostly based on Snellen or logarithm of the minimum angle of resolution (logMAR) acuity testing. However, there is now an increasing emphasis from regulators and other relevant stakeholders on the incorporation of patient voice in clinical trials.⁷ One way to include patient perspective when evaluating treatment efficacy is through inclusion of a patient-reported outcome measure (PROM)-supported endpoint.^{8–10}

A patient-reported outcome (PRO) is a health outcome reported directly by the patient (ie, study participant) about the status of the patient's health without amendment or interpretation of the patient's response by a clinician or anyone else.¹¹ In 2009, the US Food and Drug Administration (FDA) released their influential 'Guidance for Industry on Patient-Reported Outcome Measures', which describes the review and evaluation criteria for PROs used to support claims in approved medical product labelling.¹² The



European Medicines Agency also published a reflection paper in 2005 to provide broad recommendations on health-related quality-of-life (HRQoL) evaluation in the context of clinical trials.¹³ Aside from having the potential to support regulatory evaluation, marketing authorisation and product label claims, the data generated by PROM-supported endpoints can also be of value for reimbursement authorities, clinicians and patients to quantify the added benefit of a treatment.

In order for PROM data to support an FDA product label claim, the PROM must assess relevant and important concept(s) of interest and has been developed and validated to the standards specified in regulator guidelines (including evidence of both content validity and psychometric validity in the population of interest) and included in a well-designed and adequately controlled clinical trial. Hence, the review and critical evaluation of existing PROMs have become an important first step in identifying a PROM that assesses the concept of interest (the ‘thing’ being measured; ie, near-vision functioning) in a defined context of use (ie, as a clinical trial endpoint in a sample of adults with presbyopia with intact natural crystalline lens or phakic presbyopes). The objective of the current literature review was to identify the most commonly cited PROMs designed for use in presbyopia or similar conditions and critically evaluate the evidence of content validity and psychometric properties, and as such their adequacy for use to support endpoints in a Novartis’s pivotal trial testing a new pharmacological therapy for phakic individuals with presbyopia.

METHODS

Patient and public involvement

As this is a review of the available literature on PROMs used in presbyopia, patients were not involved in this research project.

Phase 1 (search strategy and screening)

A comprehensive review of bibliographic databases and other sources (ClinicalTrials.gov, Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID)) was conducted to identify PROMs that have been used in individuals with presbyopia. Literature searches were performed in Medline, Embase and Evidence-Based Medicine Reviews (Cochrane Database of Systematic Reviews and Database of Abstracts of Reviews of Effects) via Ovid SP up until October 2017 (when the search was conducted) using specific search strings

(table 1). Limits were placed on searches to exclude articles in languages other than English or studies not conducted in human subjects.

Conference abstracts, case studies and case reports were not considered for further screening. Studies of populations not exclusively comprising individuals with presbyopia were excluded, as were studies that did not include a PROM. Studies assessing treatment satisfaction in patients with presbyopia using a non-standardised measure, such as a single-item visual analogue scale or a numeric analogue scale, were excluded. PROMs designed to assess only dry eye symptoms as a result of the use of contact lens in individuals with presbyopia were also excluded from further analysis.

The first-level screening was performed based on the title and abstract of the citations, and full-text copies of the studies with individuals with presbyopia were obtained for the next round of review, which involved critical full-text appraisal based on the aforementioned criteria.

Phase 2 (evaluation of PROMs)

In order to assess the PROMs identified in phase 1 based on available information regarding the development and psychometric validation of the measures, additional publications were accessed via a PROM-specific search using Ovid, bibliographic searches and information provided by the PROM developers (either online or on request). Information about the development and psychometric validation was extracted and compared across the measures. Only PROMs where information on the development and validation history was available via an accessible full-text publication were chosen for the review. The properties of the PROMs were assessed against the FDA guidance for PROMs submitted as a clinical endpoint for drug approval to support a labelling claim (online supplementary table S1).¹² Information related to the use of modern psychometric techniques, such as Rasch analysis, was also reviewed where available. Rasch models convert ordinal scores into linear, interval-level data allowing for easier interpretation, presenting both item difficulty and respondent ability on a common scale. Further evaluation of the face validity, including examination of the item wording and missing concepts, was undertaken for PROMs that satisfied most of the psychometric assessment criteria.

Table 1 Search strings used in the targeted literature search

Search strings used in the targeted literature search

Patient-reported outcomes and quality of life	("quality of life" or "quality-of-life" or QoL or disability or "patient reported outcome*" or "patient-reported outcome*" or PRO or "Quality adjusted life year*" or QALY or "Quality-adjusted life year*" or ((activit* adj2 daily adj2 living) or prefer* or satisfact* or questionnaire or scale or score) or (burden adj5 (societ* or parent* or caregiver or humanistic))).mp.
Disease	presbyo*.mp.

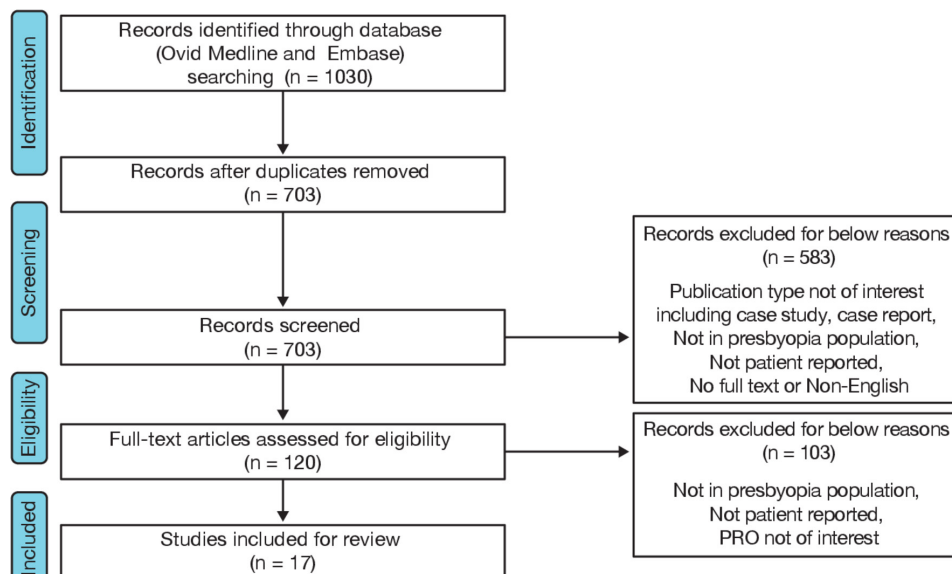


Figure 1 Flow chart illustrating the study selection process. PRO, patient-reported outcome.

RESULTS

A total of 703 unique records were identified from the literature search during the first round of abstract screening, and 120 were selected for full-text review (figure 1). Using ClinicalTrials.gov, a total of 121 clinical studies in presbyopia were identified, of which only 21 mentioned a PROM used to support either a primary or secondary endpoint. Seventeen studies with varied study designs and sample sizes ranging from 26 to 7890 participants were found to meet the inclusion criteria for full-text review (online supplementary table S2).

In total, 13 PROMs designed to assess vision outcomes (including symptoms and/or HRQoL impact) of presbyopia or similar vision conditions were identified. Of the 13 measures, only one was presbyopia-specific; 11 were generic eye disease measures used across a range of patient groups, including those with refractive correction (refractive surgery, spectacles and contact lenses) and cataracts; and one was a numeric rating scale (NRS) assessing overall vision satisfaction and ocular comfort. Based on the relevance of content (excluding PROMs measuring the effect of dry eye from contact lens use and NRS), nine PROMs were deemed suitable for further assessment of psychometric properties (table 2). The nine-item version of the National Eye Institute Visual Function Questionnaire (NEI VFQ-9) is an abbreviated version of NEI VFQ-25 and so was not counted separately within the nine PROMs identified.

Phase 2 searches were conducted to evaluate and compare the psychometric properties of the nine PROMs meeting the inclusion criteria from phase 1. PROM-specific Ovid searches identified 116 abstracts. Of these 116 abstracts, 20 had been identified in the phase 1 searches, 32 were not relevant and the remaining were selected for full-text review, of which 23 were included in the final review. Information related to the psychometric properties of these nine PROMs is presented in table 3.

Content validity

Although patient input was sought in the development of most PROMs reviewed, individuals with presbyopia were only involved in the item generation process of six PROMs: Near Activity Visual Questionnaire (NAVQ), National Eye Institute Visual Function Questionnaire (NEI VFQ-25), Quality of Vision (QoV), National Eye Institute Refractive Error Quality of Life Instrument-42 (NEI RQL-42), Freedom From Glasses Value Scale (FGVS), and Refractive Status And Vision Profile (RSVP). While the NAVQ was the only presbyopia-specific PROM identified, it should be noted that individuals with pseudophakic presbyopia were included in the NAVQ development along with individuals with phakic presbyopia.^{14 15} The FGVS¹⁶ included individuals diagnosed with cataracts or presbyopia who had undergone IOL implantation surgery. Refractive error focus groups were used to derive items for the NEI RQL-42¹⁷ and RSVP,¹⁸ whereas content for the NEI VFQ-25^{19 20} was informed by conducting 26 condition-specific focus groups to meet its intended objective of evaluating vision-related quality of life across several common eye conditions, including cataracts, age-related macular degeneration, diabetic retinopathy and glaucoma. Participants involved in the item refinement procedure of the QoV²¹ included individuals with and without refractive correction. The remaining three PROMs, Catquest-9SF, Visual Function Index (VF-14) and Cataract Type Spec, were developed with input from individuals with cataract only.^{22–24}

Reliability

The internal consistency of the NAVQ met the acceptability threshold of >0.70 ,¹⁴ as did the *near vision* subscales of the NEI VFQ-25, NEI RQL-42 and Cataract Type Spec (ranging from 0.85 to 0.95).^{24–26} While high internal consistency for the NAVQ (0.945) and Cataract Type Spec (0.94) indicates that the items are highly correlated

**Table 2** Patient-reported outcome measures in presbyopia considered for evaluation of psychometric characteristics

Patient-reported outcome measures	Objective	Number of items and available domain scores	Response options	Recall period
Near Activity Visual Questionnaire	To assess near-visual function and satisfaction to complement visual function evaluations of presbyopic corrections.	11 items relating to domains: <ul style="list-style-type: none"> ▶ Near-vision tasks (10 items). ▶ Overall satisfaction with near vision (1 item). 	<ul style="list-style-type: none"> ▶ 4-point verbal descriptor scale ranging from 'no difficulty' to 'extreme difficulty', along with option of 'N/A or stopped for non-visual reasons'. ▶ 5-point verbal descriptor scale ranging from 'completely satisfied' to 'completely dissatisfied'. 	None specified.
National Eye Institute Visual Function Questionnaire-9	To measure the influence of visual impairment on quality of life.	Abbreviated version of NEI VFQ-25 with 9 items in 7 domains: <ul style="list-style-type: none"> ▶ General vision. ▶ Well-being/mental health. ▶ Near vision. ▶ Distance vision. ▶ Driving. ▶ Role limitation. ▶ Peripheral vision. 	▶ 5-point to 6-point verbal descriptor scales with variable response options.	Present time and in general.
National Eye Institute Visual Function Questionnaire-25 (NEI VFQ-25)	To measure the influence of visual impairment on quality of life.	25 (+1) items in 11 domains: <ul style="list-style-type: none"> ▶ Global vision rating (1 item). ▶ Difficulty with near-vision activities (3 items). ▶ Difficulty with distance vision activities (3 items). ▶ Limitations in social functioning due to vision (2 items). ▶ Role limitations due to vision (2 items). ▶ Dependency to others due to vision (3 items). ▶ Mental health symptoms due to vision (4 items). ▶ Driving difficulties (3 items). ▶ Limitations with peripheral vision (1 item). ▶ Colour vision (1 item). ▶ Ocular pain (2 items). ▶ General health (1 item). 	<ul style="list-style-type: none"> ▶ Yes/no. ▶ 3-point to 6-point verbal descriptor scales. 	Present time and in general.
Quality of Vision	To measure outcome of refractive correction and cataract surgery.	30 items made up of 10 symptoms each assessed in 3 dimensions: <ul style="list-style-type: none"> ▶ Frequency. ▶ Severity. ▶ Bothersome. 	▶ 4-point verbal descriptor scale.	Past week.
Visual Function Index	To measure functional impairment caused by cataract.	18 items in 1 domain: <ul style="list-style-type: none"> ▶ Visual impairment. 	<ul style="list-style-type: none"> ▶ Yes/no. ▶ 4-point verbal descriptor scale. 	Currently.

Continued

Table 2 Continued

Patient-reported outcome measures	Objective	Number of items and available domain scores	Response options	Recall period
Cataract TyPE Spec	To measure outcomes of cataract extraction.	12 items assessing visual functioning in 5 domains: <ul style="list-style-type: none"> ▶ Distance vision. ▶ Near vision. ▶ Daytime driving. ▶ Night-time driving. ▶ Glare. 	▶ 5-point verbal descriptor scale ranging from 'totally disabled' to 'not at all'.	Present time.
Catquest-9SF	To measure the benefit of cataract surgery.	9 items in 2 domains: <ul style="list-style-type: none"> ▶ Disability (7 items). ▶ Global evaluation (2 items). 	▶ 4-point verbal descriptor scale ranging from 'very great difficulty' to 'no difficulty' for 6 items, and from 'very dissatisfied' to 'very satisfied' for 1 item.	Past 4 weeks.
National Eye Institute Refractive Error Quality of Life Instrument-42	To assess the impact of refractive error and its correction on day-to-day life.	42 items in 13 domains: <ul style="list-style-type: none"> ▶ Clarity of vision (4 items). ▶ Expectations (2 items). ▶ Near vision (4 items). ▶ Far vision (5 items). ▶ Diurnal fluctuations (2 items). ▶ Activity limitations (4 items). ▶ Glare (2 items). ▶ Symptoms (7 items). ▶ Dependence on correction (4 items). ▶ Worry (2 items). ▶ Suboptimal correction (2 items). ▶ Appearance (3 items). ▶ Satisfaction with correction (1 item). 	<ul style="list-style-type: none"> ▶ Yes/no. ▶ 3-point to 6-point verbal descriptor scale. 	Depending on the items (At this time, The last 4 weeks, In general, Now, ...).
Freedom From Glasses Value Scale	To assess the added value of complete independence from spectacles after multifocal intraocular lens surgery.	20 items with 17 items grouped under 2 domains: <ul style="list-style-type: none"> ▶ Global evaluation. ▶ Advantages. 	▶ 5-point verbal descriptor scale ranging from 'much better' to 'much worse', 'very positive' to 'very negative', 'no, not at all' to 'yes, absolutely', 'totally agree' to 'totally disagree', or 'definitely better without glasses' to 'definitely better with glasses'.	Present time.
Refractive Status and Vision Profile	For the measurement of patient-reported outcomes of refractive surgery and other treatments for refractive error.	42 items in 8 domains: <ul style="list-style-type: none"> ▶ Concern (6 items). ▶ Expectations (2 items). ▶ Physical/social functioning (11 items). ▶ Driving (3 items). ▶ Symptoms (5 items). ▶ Optical problems (5 items). ▶ Glare (3 items). ▶ Problems with corrective lenses (7 items). 	▶ 5-point, 7-point and 11-point verbal descriptor scales with variable response options.	During the past month.



Table 3 Evaluation of psychometric characteristics of the patient-reported outcome measures in presbyopia

PROMs	Validity: content		Reliability		Validity: construct				Known groups	Responsiveness in presbyopia]	Comments
	Patient input	Cognitive debriefing	Item-level performance using Rasch analysis	Internal consistency*	Test-retest†	Hypothesis specified	Convergent/divergent‡	Convergent/divergent‡			
NAVQ (2008)	✓	NR	✓	✓	✓	✓	✓	✓	✓	<ul style="list-style-type: none"> ▲ Only measure found to be specific to presbyopia. ▲ Has rigorous and well-documented development and has robust psychometric properties. 	
NEI VFQ-9 (2010)	✓	NR	✓	✓	NR	✓	✓	✓	NR	<ul style="list-style-type: none"> ▲ Demonstrates stronger psychometric properties than the original 25-item version. 	
NEI VFQ-25 (2001)	✓	NR	X	X	X	✓	✓	✓	X	<ul style="list-style-type: none"> ▲ Failed to demonstrate validity as it did not fit the Rasch model; several subscales found to be psychometrically unsound. ▲ Lengthy and not easy to complete. 	

Continued

Table 3 Continued

PROMs	Validity: content		Item-level performance using Rasch analysis	Reliability		Validity: construct			Comments
	Patient input	Cognitive debriefing		Internal consistency*	Test-retest†	Hypothesis specified	Convergent/divergent‡	Known groups§	
QoV (2010)	√	√	√	√	NR	√	√	√	<ul style="list-style-type: none"> Although this measure has strong psychometric properties, the validation study suffered from methodological limitations, such as the average value of the measurements for both eyes was used for correlation analysis for construct validity assessment.
VF-14 (1994)	NR	NR	√	NR	NR	NR	NR	√	<ul style="list-style-type: none"> Originally designed for measuring functional impairments caused by only cataracts. Not validated in individuals with presbyopia. Has a two-stage question structure, which is likely to make it more confusing to complete.

Continued



Table 3 Continued

PROMs	Validity: content		Item-level performance using Rasch analysis	Reliability		Validity: construct			Comments
	Patient input	Cognitive debriefing		Internal consistency*	Test-retest†	Hypothesis specified	Convergent/divergent‡	Known groups§	
Cataract Type Spec (2003)	X	NR	X	NR	NR	NR	NR	NR	<ul style="list-style-type: none"> ▲ Developed to measure cataract surgery outcome. ▲ Not validated in a presbyopia sample. ▲ Significant floor effect in that the items did not target participants with low levels of visual disability (ie, at bottom of the item-person map).
Catquest-9SF (2008)	X	NR	√	NR	NR	NR	NR	NR	<ul style="list-style-type: none"> ▲ Not developed using input from individuals with presbyopia. ▲ Small level of differential item functioning was seen for a few items.
NEI RQL-24 (2003)	√	NR	X	√	X	NR	NR	NR	<ul style="list-style-type: none"> ▲ Was shown to be deficient for all psychometric properties tested using Rasch analysis, with issues related to lack of precision in scores, item misfit and poor targeting.

Continued

Table 3 Continued

PROMs	Validity: content		Item-level performance using Rasch analysis	Reliability		Validity: construct			Comments	
	Patient input	Cognitive debriefing		Internal consistency*	Test-retest†	Hypothesis specified	Convergent/divergent‡	Known groups§		Responsiveness in presbyopia¶
FGVS (2010)	√	√	X	√	NR	√	√	√	NR	▶ A ceiling effect was observed for the dimension as well as for all subdimension scores, that is, majority of individuals selecting highest response option for many of the items.
RSVP (2000)	√	NR	X	√	NR	√	NR	√	NR	▶ Subscales suffer from poor targeting. ▶ 1 of 3 items in each scale demonstrates differential item functioning. This is a problem because it indicates that people of different characteristics answer questions differently.

√: meets the criteria for good psychometric performance; X: fails to satisfy the criteria for good psychometric performance.

*Range for acceptable Cronbach's alpha: above 0.70 but not higher than 0.95.⁴⁴

†Threshold for acceptable test-retest reliability: interclass correlation coefficient of 0.75 or greater.⁴⁵

‡At least Pearson's correlation coefficient (r) value was categorised as moderate (0.25–0.50) or strong (>0.50).⁴⁶

§Discriminant validity demonstrated by statistically significant (p<0.05) difference in at least one comparison of patient subgroups with differing clinical features.

¶Responsiveness demonstrated by statistically significant (p<0.05) changes in scores following an intervention in at least one longitudinal validation study or one randomised controlled trial in individuals with presbyopia.

FGVS, Freedom From Glasses Value Scale; NAVQ, Near Activity Visual Questionnaire; NEI RQL-42, National Eye Institute Refractive Error Quality Of Life Instrument-42; NEI VFQ-9, National Eye Institute Visual Function Questionnaire-9; NEI VFQ-25, National Eye Institute Visual Function Questionnaire-25; NR, not reported; PROMs, patient-reported outcome measures; QoV, Quality of Vision; RSVP, Refractive Status And Vision Profile; VF-14, Visual Function Index.



and therefore measure the common concept of near vision, such high alpha coefficients suggest there may be some level of redundancy and that item reduction may be possible. Acceptable test–retest reliability was found for the NAVQ and near vision subscales of the NEI VFQ-25 and NEI RQL-42 (intraclass correlation coefficient ranging from 0.72 to 0.91), although the same was not reported for TyPE Spec. Among other measures that do not include a separate near-vision dimension, the internal consistency was found to be acceptable (>0.70) for FGVS, QoV and RSVP, whereas the test–retest reliability was strong for QoV (>0.70), was modest in the case of RSVP (intraclass correlation of 0.66) and was not reported for FGVS and NEI VFQ-9.^{18 21 27 28} As the values of reliability for Catquest-9SF, VF-14 and Cataract TyPE Spec were not reported specifically for individuals with presbyopia, these were not considered in the analysis.

Construct validity

The assessment of convergent validity showed that the NAVQ demonstrated moderate correlations with near visual acuity (Pearson's correlation coefficient (r)=0.32) and critical print size (r =0.27), which provided evidence of convergent validity.¹⁴ Strong correlations (r =0.65–0.70) were observed between Early Treatment Diabetic Retinopathy Study visual acuity and NEI VFQ-25 subscales of near vision, distance vision and general vision.²⁵ In the case of the NEI RQL-42, there was a significant association between most of the NEI RQL-42 subscales and the refractive error in both better-seeing and worse-seeing eyes.²⁶ For the QoV, moderate correlations were observed between logMAR visual acuity and the three QoV subscales (frequency scale r =0.72, severity scale r =0.64 and bothersome scale r =0.35).^{21 29} No evidence of the FGVS correlating with objective clinical parameters or subjective measures of visual function was reported.³⁰ RSVP scores were moderately associated with patient satisfaction (r =−0.41) and rating of vision (r =−0.42).³¹

Discriminant validity for the NAVQ was also supported by the high area under the receiver operator characteristic curve (0.91), a Rasch separation index of 2.92, and an overall good fit to the Rasch model with negligible ceiling and floor effects.¹⁴ The NAVQ scores also discriminated between individuals with different kinds of correction, such as IOL, contact lenses and varifocal spectacles, supporting known groups validity. In the case of the NEI VFQ-25, the near-vision and distance-vision scores significantly discriminated between participants in the reference group (better vision) and those with poor vision.²⁵ Although the psychometric performance of the NEI VFQ-25 based on analysis using classic test theory has been shown to be adequate, analyses using Rasch validation methodology have demonstrated performance limitation in the near vision subscale and general design of the NEI VFQ-25 structure, including a substantial ceiling effect.^{32–34} Scores of the NEI RQL-42 differed significantly for participants in subgroups based on refractive error correction (no correction;

postsurgery—no correction; glasses; multifocal glasses; contact lenses).²⁶ In a Rasch validation study, the NEI RQL-42 near vision subscale demonstrated poor discriminative ability as suggested by a person separation index of 0.71, which was below the acceptable threshold of 2.0.³⁵ For the QoV, discriminant ability of the scores was demonstrated by acceptable value of person separation indexes: 2.08, 2.10 and 2.01 for frequency, severity and bothersome scales, respectively.^{21 29} FGVS score suggested discrimination between individuals wearing glasses after surgery and those who gained spectacle independence.³⁰ RSVP scores were shown to distinguish among individuals based on the degree of refractive error. However, the Rasch analysis revealed several redundant and misfitting items and poor item to person targeting that could diminish its discriminative ability.³¹ The construct validity of the remaining measures was not evaluated in a presbyopia-specific patient sample nor in a diverse group that included a subset of individuals with presbyopia.^{22–24}

Responsiveness

In a randomised, controlled, cross-over trial assessing performance of a commercially available contact lens, Biofinity users demonstrated significantly greater improvements in NAVQ scores compared with OASYS contact lens users (p =0.047).³⁶ Gundersen and Potvin³⁷ reported that QoV scores of frequency, severity and degree to which the symptoms were bothersome were more sensitive than when using the NEI VFQ-25 in detecting between multifocal and monofocal IOLs. Furthermore, a cross-over study comparing two presbyopic soft contact lens modalities showed that, out of 13, only the 3 subscales of clarity, vision and appearance showed significant improvements in scores compared with the habitual correction method on the NEI RQL-42 scale.³⁸ Gierek-Ciaciura *et al*³⁹ reported an improvement in VF-14 scores for 83% of the individuals after implantation of multifocal IOLs, although there was no significant difference in postoperative scores between individuals with different IOLs. Following the bilateral Laser Anterior Ciliary Excision (LaserACE) procedure, a significant change in scores was registered on Catquest-9SF from a mean patient satisfaction score of −1.00 to 0.33 after surgery (p =0.000016).⁴⁰ Responder definition or minimal important difference thresholds were not reported for any of the reviewed questionnaires in individuals with presbyopia.

Critical evaluation of the NAVQ

As the NAVQ was the only disease-specific measure identified at the time of literature review, further critical assessment of its face validity was undertaken (figure 2). It was concluded that the language used in the instructions could be interpreted differently by respondents who use contact lenses instead of reading spectacles. Furthermore, the absence of a recall period mentioned in the instructions could lead to participants using different timeframes for recollection.

THE NEAR ACTIVITY VISUAL QUESTIONNAIRE (NAVQ)

Name: _____ DOB: ___/___/___ Gender: Male/Female Date: _____

Please answer ALL questions for the situation IF/WHEN YOU DO THE DESCRIBED ACTIVITY WITHOUT EXTRA READING SPECTACLES.

Circle the relevant option.

If you do not do the described activity or you have stopped for reasons that are not related to your vision then please circle the 'N/A' option.

How much difficulty do you have:	N/A or stopped for non-visual reasons	No Difficulty	A little difficulty	Moderate difficulty	Extreme Difficulty
1. Reading small print, such as: newspaper articles, items on a menu, telephone directories?	x	0	1	2	3
2. Reading labels/ instructions/ ingredients/ prices such as on: medicine bottles, food packaging?	x	0	1	2	3
3. Reading your post/ mail, such as: electric bill, greeting cards, bank statements, letters from friends & family?	x	0	1	2	3
4. Writing and reading your own writing, such as: greeting cards, notes, letters, filling in forms, checks, signing your name?	x	0	1	2	3
5. Seeing the display & keyboard on a computer or calculator?	x	0	1	2	3
6. Seeing the display & keyboard on a mobile or fixed telephone?	x	0	1	2	3
7. Seeing objects close to you and engaging in your hobbies, such as: playing card games, gardening, seeing photographs?	x	0	1	2	3
8. Seeing objects close to you in poor or dim light?	x	0	1	2	3
9. Maintaining focus for prolonged near work?	x	0	1	2	3
10. Conducting near work?	x	0	1	2	3
OVERALL	Completely Satisfied	Very Satisfied	Moderately Satisfied	A little satisfied	Completely Unsatisfied
How satisfied are you with your near vision?	0	1	2	3	4

Figure 2 Near Activity Visual Questionnaire.

Some of the issues identified with the content included use of examples of differing difficulty level in the same item (items 1, 5, 6 and 7); use of examples not relevant to modern times (such as telephone directories, postal electricity bills and letters); and examples measuring concepts other than near-vision ability (such as writing, playing card games and gardening) that assess manual dexterity and may lead to spurious scores. In addition, more relevant examples assessing ease of typing on a smartphone or tablet, reading on an electronic device (such as a smartphone, tablet or computer screen) and impact of lighting conditions on performing routine activities were missing from the items.

DISCUSSION

This literature review identified PROMs developed for use in individuals with presbyopia and compared their properties in the context of regulatory requirements for supporting product label claims.¹² Of the nine unique measures that were evaluated, only NAVQ was found to be developed to measure difficulties in near-vision function specifically in individuals with presbyopia. The other PROMs were found to have limitations related to lack of focus on presbyopia and insufficient evidence to support strong psychometric properties using modern psychometrics. Modern psychometric methods based on the item response theory, such as the Rasch measurement theory, provide a robust approach to examine validity and to help overcome two key limitations associated with traditional validation methods based on classic test theory: (1) scores are sample-dependent and scale-dependent; and (2) SE of measurement around individual subjects' scores is assumed to be a constant value regardless of the person's location on the range of a scale.^{41 42}

While the NAVQ has undergone rigorous psychometric analysis and was developed with input from individuals with presbyopia, its validation study was performed in a sample of individuals with pseudophakic and phakic presbyopia, and therefore the results may not be generalisable specifically to phakic presbyopia. Face validity analyses revealed the requirement for rewording the measure instructions, including a short recall period of 1 week; reassessment of the relevance of examples; and replacement of irrelevant examples with others more applicable in present times.

As per FDA's PRO guidance, examples of modifications to the original measure that alter responses of participants to the same set of questions and therefore require qualitative evidence to establish content validity include (1) making changes in the order of items, item wording, response options or recall period or deleting portions of a questionnaire; (2) using the questionnaire in a different setting, population or condition from the one for which it was originally developed; (3) changing instructions or the placement of instructions within the PROM; (4) switching mode of administration from paper to electronic format; and (5) changing the timing of, or procedures for, PROM administration within the clinic visit.¹² Thus, based on the proposed modifications to the NAVQ, a qualitative study confirming the content validity and conceptual framework of the revised measure along with evidence generated in the context of a well-designed trial to support its psychometric properties, both in the targeted population of interest, would be necessary for the NAVQ to be qualified as a fit-for-purpose measure for supporting product labelling claims.

Although this literature review was confined to assessing PROMs used in presbyopia studies, it complements

findings of a review by Kandel *et al*,⁴³ which also identified the NAVQ as one of the superior quality questionnaires for measuring refractive surgery outcomes related to activity limitations. With regard to the limitations, it is possible that some relevant literature may have been missed due to restrictions of the literature search to only a few selected databases, although a bibliographic search was also performed to ensure identification of all relevant PROMs in presbyopia. Furthermore, the properties of the culturally adapted versions of the PROMs were not assessed, and it is possible that some of the identified measures might have performed better in a non-English-speaking population. However, such data would be unlikely to change the conclusions here, as NAVQ is the only presbyopia-specific measure that has the potential to support product label claims, although with further modification. Given the lack of availability of a conceptual model based on patient and clinician input, the final selection of PROMs identified for detailed review was not informed by a concept-mapping exercise. The results of a recently conducted qualitative study in phakic presbyopes and healthcare practitioners to develop a conceptual model for the NAVQ modification will be reported in a future publication. Finally, it is also acknowledged that this literature review did not meet the criteria to be considered as a 'systematic review'. However, the authors believe that the approach was sufficiently structured and rigorous (eg, through the use of search terms and inclusion criteria) and that 'systematic' literature review methods are not typically judged necessary or appropriate to identify and evaluate adequately PROMs.

CONCLUSION

No single PROM fully adhered to the quality standards detailed in the US FDA Guidance for Industry on Patient-Reported Outcome Measures (2009), which represents best practice methods for PROMs. The NAVQ has the most potential to support trial endpoints related to changes in near-vision functioning associated with presbyopia, but with modification. Further research is ongoing to confirm the content validity and psychometric validity of a revised NAVQ in this specific population.

Acknowledgements Editorial support was provided by Santanu Bhadra, PhD (Novartis), in accordance with Good Publication Practice (GPP3) guidelines (<http://www.ismpp.org/gpp3>).

Contributors GS performed the initial literature search and analysis of data in April 2017. CN did the reanalysis and reporting. SB, AF, CJ and RA did the update of the literature search, along with analysis and reporting, and contributed to critical analysis of NAVQ. DV had the idea of this study. SC-R took over DV's role as the guarantor. FP helped with planning the publication. JW actively collaborated with Novartis to provide all available evidence for NAVQ and permission to use copy of the NAVQ in the manuscript.

Funding This study was funded by Novartis Pharma, Basel, Switzerland.

Competing interests GS is an employee of Novartis Pharma. SC-R, DV, CN and FP are employees and shareholders of Novartis Pharma. SB, AF and RA are employees of Adelphi Values, which received funding from Novartis to conduct some aspects of the study. CJ was an employee of Adelphi Values during the conduct of this

study. While JW was one of the developers of the NAVQ, he did not influence the conclusions and outcome of this review.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Garima Sharma <http://orcid.org/0000-0002-9075-0248>

James Wolffsohn <http://orcid.org/0000-0003-4673-8927>

REFERENCES

- 1 Glasser A, Campbell MC. Biometric, optical and physical changes in the isolated human crystalline lens with age in relation to presbyopia. *Vision Res* 1999;39:1991–2015.
- 2 Heys KR, Cram SL, Truscott RJW. Massive increase in the stiffness of the human lens nucleus with age: the basis for presbyopia? *Mol Vis* 2004;10:956–63.
- 3 Garner WH, Garner MH. Protein disulfide levels and lens elasticity modulation: applications for presbyopia. *Invest Ophthalmol Vis Sci* 2016;57:2851–63.
- 4 Mathebula SD, Makunyane PS. Amplitude of accommodation is reduced in pre-presbyopic diabetic patients. *Journal of Endocrinology, Metabolism and Diabetes of South Africa* 2017;22:12–16.
- 5 Holden BA, Fricke TR, Ho SM, *et al*. Global vision impairment due to uncorrected presbyopia. *Arch Ophthalmol* 2008;126:1731–9.
- 6 Patel I, West SK. Presbyopia: prevalence, impact, and interventions. *Community Eye Health* 2007;20:40–1.
- 7 FDA. Patient-Focused drug development draft guidance 1: collecting comprehensive and representative input, 2018. Available: <https://www.fda.gov/media/113653/download> [Accessed 04 May 2020].
- 8 McDonnell PJ, Lee P, Spritzer K, *et al*. Associations of presbyopia with vision-targeted health-related quality of life. *Evidence-Based Eye Care* 2004;5:100–1.
- 9 Lu Q, Congdon N, He X, *et al*. Quality of life and near vision impairment due to functional presbyopia among rural Chinese adults. *Invest Ophthalmol Vis Sci* 2011;52:4118–23.
- 10 Williams S, Brian G, du Toit R. Measuring vision-specific quality of life among adults in Fiji. *Ophthalmic Epidemiol* 2012;19:388–95.
- 11 Weldring T, Smith SMS. Patient-Reported outcomes (pros) and patient-reported outcome measures (PROMs). *Health Serv Insights* 2013;6:61–8.
- 12 FDA. Patient-Reported outcome measures: use in medical product development to support labeling claims, 2009. Available: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims> [Accessed 06 Feb 2020].
- 13 EMA. Reflection paper on the regulatory guidance for the use of health related quality of life (HRQL) measures in the evaluation of medicinal products, 2005. Available: https://www.ema.europa.eu/en/documents/presentation/presentation-reflection-paper-use-patient-reported-outcome-pro-measures-oncology-studies-daniel_en.pdf [Accessed 06 Feb 2020].
- 14 Buckhurst PJ, Wolffsohn JS, Gupta N, *et al*. Development of a questionnaire to assess the relative subjective benefits of presbyopia correction. *J Cataract Refract Surg* 2012;38:74–9.
- 15 Gupta N, Wolffsohn JS, Naroo SA, *et al*. Development of a near activity vision questionnaire to assess accommodating intraocular lenses. *Cont Lens Anterior Eye* 2007;30:134–43.
- 16 Lévy P, Elies D, Dithmer O, *et al*. Development of a new subjective questionnaire: the freedom from glasses value scale (FGVS). *J Refract Surg* 2010;26:438–46.
- 17 Berry S, Mangione CM, Lindblad AS, *et al*. Development of the National eye Institute refractive error correction quality of life questionnaire: focus groups. *Ophthalmology* 2003;110:2285–91.
- 18 Vitale S, Schein OD, Meinert CL, *et al*. The refractive status and vision profile: a questionnaire to measure vision-related quality of life in persons with refractive error. *Ophthalmology* 2000;107:1529–39.
- 19 Mangione CM, Lee PP, Gutierrez PR, *et al*. Development of the 25-item National eye Institute visual function questionnaire. *Arch Ophthalmol* 2001;119:1050–8.

- 20 Mangione CM, Berry S, Spritzer K, *et al.* Identifying the content area for the 51-item National eye Institute visual function questionnaire: results from focus groups with visually impaired persons. *Arch Ophthalmol* 1998;116:227–33.
- 21 McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the quality of vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci* 2010;51:5537–45.
- 22 Lundström M, Roos P, Jensen S, *et al.* Catquest questionnaire for use in cataract surgery care: description, validity, and reliability. *J Cataract Refract Surg* 1997;23:1226–36.
- 23 Steinberg EP, Tielsch JM, Schein OD, *et al.* The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol* 1994;112:630–8.
- 24 Javitt JC, Jacobson G, Schiffman RM. Validity and reliability of the cataract type spec: an instrument for measuring outcomes of cataract extraction. *Am J Ophthalmol* 2003;136:285–90.
- 25 Mangione CM, Lee PP, Gutierrez PR, *et al.* Development of the 25-item National eye Institute visual function questionnaire. *Evidence-Based Eye Care* 2002;3:58–9.
- 26 Hays RD, Mangione CM, Ellwein L, *et al.* Psychometric properties of the National eye Institute-Refractive error quality of life instrument. *Ophthalmology* 2003;110:2292–301.
- 27 Berdeaux G, Meunier J, Arnould B, *et al.* Measuring benefits and patients' satisfaction when glasses are not needed after cataract and presbyopia surgery: scoring and psychometric validation of the freedom from glasses value scale (FGVS©). *BMC Ophthalmol* 2010;10:15.
- 28 Kodjebacheva G, Coleman AL, Ensrud KE, *et al.* Reliability and validity of abbreviated surveys derived from the National eye Institute visual function questionnaire: the study of osteoporotic fractures. *Am J Ophthalmol* 2010;149:330–40.
- 29 McAlinden C, Skiadaresi E, Gatinel D, *et al.* The quality of vision questionnaire: subscale interchangeability. *Optom Vis Sci* 2013;90:760–4.
- 30 Berdeaux G, Meunier J, Arnould B, *et al.* Measuring benefits and patients' satisfaction when glasses are not needed after cataract and presbyopia surgery: scoring and psychometric validation of the freedom from glasses value scale (FGVS). *BMC Ophthalmol* 2010;10:15.
- 31 Garamendi E, Pesudovs K, Stevens MJ, *et al.* The refractive status and vision profile: evaluation of psychometric properties and comparison of Rasch and summated Likert-scaling. *Vision Res* 2006;46:1375–83.
- 32 Kowalski JW, Rentz AM, Walt JG, *et al.* Rasch analysis in the development of a simplified version of the National eye Institute Visual-Function Questionnaire-25 for utility estimation. *Qual Life Res* 2012;21:323–34.
- 33 Marella M, Pesudovs K, Keeffe JE, *et al.* The psychometric validity of the Nei VFQ-25 for use in a low-vision population. *Invest Ophthalmol Vis Sci* 2010;51:2878–84.
- 34 Pesudovs K, Gothwal VK, Wright T, *et al.* Remediating serious flaws in the National eye Institute visual function questionnaire. *J Cataract Refract Surg* 2010;36:718–32.
- 35 McAlinden C, Skiadaresi E, Moore J, *et al.* Subscale assessment of the NEI-RQL-42 questionnaire with Rasch analysis. *Invest Ophthalmol Vis Sci* 2011;52:5685–94.
- 36 Sivardeen A, Laughton D, Wolffsohn JS. Investigating the utility of clinical assessments to predict success with presbyopic contact lens correction. *Cont Lens Anterior Eye* 2016;39:322–30.
- 37 Gundersen KG, Potvin R. Comparative visual performance with monofocal and multifocal intraocular lenses. *Clin Ophthalmol* 2013;7:1979–85.
- 38 Richdale K, Mitchell GL, Zadnik K. Comparison of multifocal and monovision soft contact lens corrections in patients with low-astigmatic presbyopia. *Optom Vis Sci* 2006;83:266–73.
- 39 Gierek-Ciaciura S, Cwalina L, Bednarski L, *et al.* A comparative clinical study of the visual results between three types of multifocal lenses. *Graefes Arch Clin Exp Ophthalmol* 2010;248:133–40.
- 40 Hipsley A, Ma DH-K, Sun C-C, *et al.* Visual outcomes 24 months after LaserACE. *Eye Vis* 2017;4:15.
- 41 Andrich D. Rating scales and Rasch measurement. *Expert Rev Pharmacoecon Outcomes Res* 2011;11:571–85.
- 42 Petrillo J, Cano SJ, McLeod LD, *et al.* Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health* 2015;18:25–34.
- 43 Kandel H, Khadka J, Lundström M, *et al.* Questionnaires for measuring refractive surgery outcomes. *J Refract Surg* 2017;33:416–24.
- 44 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
- 45 David L, Streiner GRN, Cairney J. *Health measurement scales: a practical guide to their development and use.* Oxford University Press, 1995.
- 46 Cohen EJ. *Statistical power analysis for the behavioral sciences.* 2nd edn. Lawrence Erlbaum Associates, 1988.