Evaluating measurement properties of Patient Reported Outcome Measures (PROMs) in glaucoma using COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN): a systematic review

Anastasiya Vinokurtseva[1], MD, Matthew P. Quinn[2,3], MD, PhD, Mandy Wai[4], BSc, Victoria Leung[5], BMSc(c), Monali Malvankar-Mehta[1,6], PhD, Cindy M. L. Hutnik[1,7], MD, PhD

1 Department of Ophthalmology, Schulich School of Medicine, Western University, London ON, 2 Department of Ophthalmology, Queen's University, Kingston ON, 3 Department of Ophthalmology, University of Ottawa, Ottawa, ON, 4 School of Optometry and Vision Science, University of Waterloo, Waterloo ON, 5 Department of Physiology and Pharmacology, Western University, London ON, 6 Department of Epidemiology and Biostatistics, Schulich School of Medicine, University of Western Ontario, London ON, 7 Ivey Eye Institute, St Joseph's Health Care, London ON

BACKGROUND

- **Glaucoma** is the leading cause of irreversible blindness worldwide.[1]
- Visual impairment has an established effect on **quality of life** (QOL), and progressive vision loss undermines patients' functioning and QOL, affecting their mobility, employment, independence, mental health and social functioning.[2]
- While there are multiple objective measures of patients' clinical status and progression, including but not limited to optical coherence tomography, visual fields, visual acuity and intraocular pressure,[3] none of these traditional measures capture patient's perception of their **illness experience**.[4,5]
- Understanding and incorporating patient preferences into decision-making is now recognized as critical for optimal resource allocation, especially in technologically advancing areas such as microinvasive surgeries.
- **Patient-Reported Outcome Measures (PROMs)** are instruments designed to evaluate the health outcomes that are most important to the patients.[6]
- Despite their recognized importance, especially in the era of **patient-centered care**, their routine use in clinical setting remains low.[7]
- The purpose of the current study was to identify and evaluate PROMs for clinical ophthalmology practice.

METHODS

- Study registered with PROSPERO (registration number CRD42020176064).
- Systematic literature search in EMBASE, MEDLINE, PsycINFO, Scopus, BIOSIS, Web of Science databases from the date of inception for articles that report measurement properties of PROMs.
- References screened independently in duplicate using following inclusion criteria:
  1. discusses PROM/PREM or any related self-report QoL instrument;
  2. at least 50% of study patient population are diagnosed with glaucoma and are at least 18 years of age,
  3. published in English language,
  4. describes instrument development, validation, or psychometric properties (Table 1).
- COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) were used for methodological quality and measurement property assessment of the included PROMs.
- Evidence was synthesized using modified GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach.[8]

Table 1. Measurement properties included in COSMIN methodology.

DOMAIN	MEASUREMENT PROPERTY	DEFINITION & IMPORTANT ASPECTS OF MEASUREMENT PROPERTIES
RELIABILITY	Reliability	Consistency of responses in similar circumstances for a consistent sample (includes test-retest, inter-rater and intra-rater reliability)
	Internal consistency	Interrelatedness among items of the scale assessing the same construct
	Measurement error	The degree to which changes in responses are attributed to reasons other than a true difference in construct
VALIDITY	Content validity	Relevance, comprehensiveness, and comprehensibility of items in a scale for the construct being measured (includes face validity)
	Construct validity	The degree of unidimensionality of each scale (structural validity), ability to detect differences between subjects known to be different in assessed construct (hypothesis testing) and performance of translated or culturally adapted versions of an instrument (cross-cultural validity)
	Criterion validity	Consistency with a gold-standard instrument
RESPONSIVENESS	Responsiveness	Sensitivity to change
INTERPRETABILITY*		Ability to assign qualitative meaning to the quantitative scores
FEASIBILITY*		Ease of instrument use in the intended context given practical constraints (such as time or finances)

REFERENCES

1 Tham, Y. C. et al. Ophthalmology 121, 2081–2090 (2014) **2** National Academies of Sciences, E. and M. et al. (2016) **3** Miki, A. Journal of Current Glaucoma Practice 6, 62 (2012) **4** McKenna, S. P. BMC Medicine 9, 86 (2011) **5** Fenwick, E. K. et al. Progress in Retinal and Eye Research 76, 100801 (2020) **6** Somner, J. E. A. et al. Investigative Ophthalmology and Visual Science 53, 5940–5947 (2012) **7** Gazzard, G. et al. Journal of Glaucoma 30, 732–743 (2021) **8** Prinsen, C. A. C. et al. Quality of Life Research 27, 1147–1157 (2018)

RESULTS

Demographic characteristics
Age: ranged between 11 and 93 years, with reported mean (SD) of 63.96 (12.55).
Diagnosis: Glaucoma, including primary open angle glaucoma (POAG), pseudoexfoliative glaucoma (PXF), pigmentary glaucoma (PG), normal tension glaucoma (NTG), primary angle closure glaucoma (PACG) and secondary glaucoma (SG).
Number of participants: varied between 8 and 1349 per study, with smaller sample sizes being common for questionnaire development and pilot studies, and psychometric properties being assessed in larger studies.

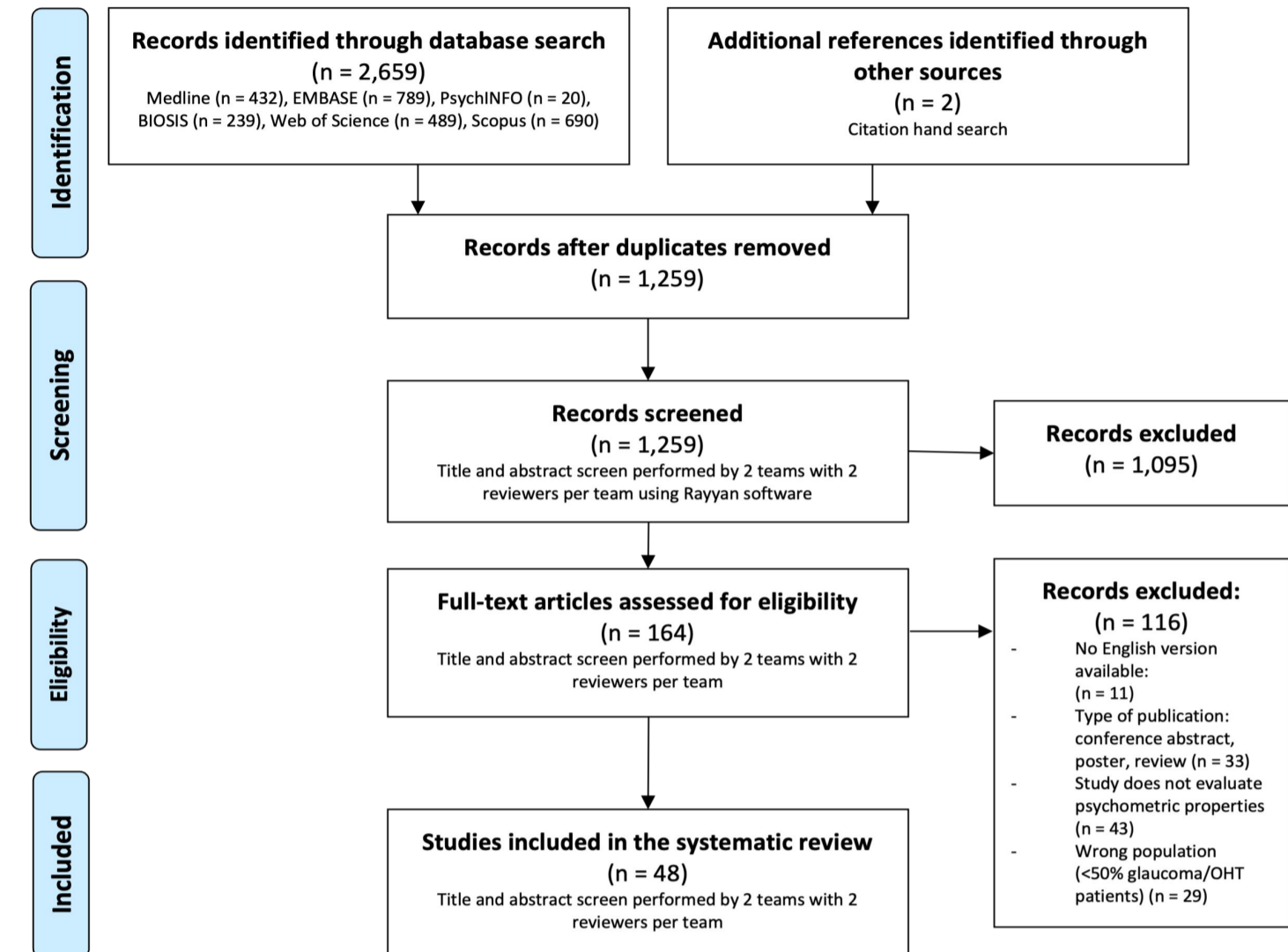Figure 1. PRISMA flow diagram of selection of studies for inclusion in the systematic review.

Table 2. Instrument characteristics.

PROM	Construct(s)	Mode of Administration	(Sub)Scale(s), Number of Items	Response Options / Range of Scores & Scoring (If stated)	Available Languages
GENERAL HEALTH QOL					
CES-D	Depression	Interview-based	8 items	1 to 5 / 8 to 40	English
EQ-5D	General health status, utility (preference-based measure QOL)	Interview-based, self-report	5 items	1 to 3 / scored as a 5-digit health status code (ie. 11111)	English
HPI	General health perceptions	Interview-based	4 items	NS	English
PHQ-9	Depression	Interview-based	9 items	0 to 3 / higher score represents greater depression	English, Telugu, Hindi
SF-36	Generic health-related QOL	Self-administered, interview-administered	8 subscales, 36 items	Subscales scored from 0 to 100	English
SF-6D	Utility values	Interview-based	6 of 8 dimensions of the 36-item SF-36	0-100 scale	English
SIP	General impairment	Interview-based	136 items	0 to 100	English
VISION-SPECIFIC QOL					
IVI	Participation in daily living assessment	Self-administered, interview-administered	28 items	Rated as "not at all" (0), "hardly at all" (1), "a little" (2), "a fair amount" (3), "a lot" (4), "can't do because of eyesight" (5), don't do because of other reasons (8)	English
NEI VFQ-25	Vision-targeted QOL	Self-administered, oral interview in person or over the telephone, computer-assisted telephone interviewing system	12 subscales, 25 items	0-100 for subscale scores / lower score means lower QOL	English, Swedish, Greek, French, Yoruba, Igbo, Hausa
Sumi VDQ	Visual disability	Interview-administered	30 items	0 to 2 / 0 to 60	Japanese
TTO	Utility values	Interview-based	One 2-part question	Yes, No	English
VAQ	Disease-specific functional status	Interview-based	33 items	1 to 5 / 33 to 165	English
VDQ	Visual disability assessment	Self-administered with or without assistance, interview-administered	58 items	5-point scale (no difficulty at all to severe difficulty)	English
VF-14	Vision-targeted QOL	Self-report	14 items	0 to 5 / averaged overall score then transformed to a 0 to 100-point scale	English
VF-14(modified)	Vision-targeted QOL	Self-report	14 items + 6 for peripheral vision, contrast sensitivity, and adaptation to darkness + 5 more related to treatment satisfaction	0 to 5 quantification scale (0-none, 1-some, 2-plenty, 3-a lot, 4-incapable of doing it, 5-cannot do it for reasons not related to eyesight) and AVS 0-100	English, Spanish
VFQ-UI	Vision-related function and preference-based measure of utility	Self-administered, with or without assistance	6 items	1 to 5 / 6 to 30; 0.16 (worst health state) to 1.0 (full health)	English
GLAUCOMA-SPECIFIC QOL					
AGQ	Glaucoma-specific instrument for PRO assessment in RCTs	Interview administered; mailed questionnaire	68 items	Ordinal values (1-4, 1-5), dichotomous responses (yes/no)	English
AL domain of the glaucoma module of the Eye-tem Bank	Glaucoma-specific activity limitation	Interview-based	88 items	5 category scale	English
CGVFT	Glaucoma-specific QOL	Self-report (computer-based)	59 items	Yes, no; 1-6 scale	English
COMTOL	Tolerability of topical medications	Interview-based	12 (but only questions 4-12 are scored)	0 to 5 or 0 to 6	English
EDSQ	Patients' satisfaction/compliance with glaucoma treatment	Self-report	6 dimensions, 43 items	Variable response scales: 1 continuous, 8 dichotomous, 2 categorical, 32 ordinal / scoring converted to 0-100 scale; higher score reflects more of the attribute referred to in the dimension	French, English
GAL-10	Glaucoma-specific QOL	Mix of self-report and interview-based	10 items	1 to 5	English, Telugu, Hindi
GAL-9	Glaucoma-specific QOL	Self-report	9 items	1 to 5 / 9 to 45	English
GHPI	Glaucoma-specific health perceptions	Interview-based	6 items	NS	English
Glau-QoL	Glaucoma-specific QOL	Self-administered, mailed, interview-administered	7 subscales, 36 items	Variable: 0 to 3, 0 to 4, or 0 to 5, depending on domain / each domain transformed into a 0 to 100 scale	French, English, Mandarin
Glau-U	Glaucoma specific QOL	Interview-based	6 items	ordinal scale, 3 levels: no difficulty, some difficulty, severe difficulty	English, Mandarin
Glaucoma medication self-efficacy scale	Glaucoma-specific self-efficacy	Self-administered, mailed	35 items (21 related to overcoming barriers to glaucoma medication use; 14 related to correct use of eye drops)	4-level ordinal scale: not at all confident, somewhat confident, very confident, does not apply	English
Glaucoma outcome expectations scale	Glaucoma-specific self-efficacy	Self-administered, mailed	4 items	9-point Likert scale ranging from 'not at all' to 'somewhat' to 'extremely'	English
Glausat	Patient satisfaction with glaucoma treatment	Self-report	7 dimensions, 22 items	5-point Likert scale / higher score reflects more satisfaction with therapy	Spanish, English
GQL-15	Glaucoma-specific QOL	Self-report or interview-based	4 subscales, 15 items	1 to 5 / higher score indicates lower QOL	English, Telugu, Hindi, Yoruba, Igbo, Hausa, Mandarin, German, Persian, Serbian
GSI	Glaucoma symptoms, QOL	Self-report, Online survey	32 items	Degree of difficulty: (1) none or I do not do this for nonvisual reasons (2) a little or some difficulty (3) yes or I no longer do this for visual reasons	English
GSS	Quantify complaints/functional impairment	Self-administered, interview-administered	2 subscales, 10 items	0 to 4, then converted on a 0 to 100 scale / lower score means lower HRQoL	English, Italian, Serbian
GTCAT	Adherence to glaucoma therapy	Self-report	27 items	5-interval Likert scale (from 'disagree a lot' to 'agree a lot')	English, Brazilian Portuguese
GUI	Utility-based glaucoma health outcome measure	Self-report	6 dimensions	4 levels	English
HUG-5	Glaucoma-specific health status / QOL, preference-based measure	Self-administered, interview-administered	5 domains, 5 items	5-level ordinal scale: none, slight, moderate, very much, severe (1 to 5) / composite score of 5 to 25	English
Item bank based on 23 PROMs	Glaucoma-specific QOL	Self-administered, with or without assistance	187 items	1 to 5	Japanese
Japanese questionnaire	Glaucoma-specific QOL	Self-report, interview-based (if couldn't read)	31 items	Yes (4), sometimes (2), no (0)	Japanese, English
MIGS questionnaire	Patient preference on MIGS	Self-report	52 items	NS	English
POEM	Glaucoma-specific QOL	Interview-based	6 items	NS	English
SHPC-18	Symptoms useful for patient care	Self-report (questionnaire administered before initiation of treatment for glaucoma)	2; Local Eye Symptom (7 items) & Visual Function Symptom (11 items)	5-point scale (a lot = 5, not at all = 1)	English
SHPC/SIG	Disease-specific impairment	Interview-based	43 items	1 to 5 / 43 to 215	English
TSS-IOP	Treatment satisfaction for ocular hypotensive medications	Self-report (questionnaire administered before initiation of treatment for glaucoma)	5 factors, 15 items	NS	English
Viswanathan 10	Glaucoma-specific QOL	Interview-based	10 items	NS / lower score means lower QOL	NS

Table 3. Measurement properties' assessment for most frequently used PROMs among generic, vision-specific and glaucoma-specific instruments.
Legend: measurement property grading (+), sufficient, (?), indeterminate, (-), insufficient; quality of evidence high, moderate, low, very low.

	PROM	Structural validity	Internal consistency	Cross-cultural validity/ measurement invariance	Reliability	Measurement error	Criterion validity	Hypotheses testing	Responsiveness
GENERAL HEALTH QOL	EQ-5D	NA	NA	NA	NA	NA	NA	NA	- (Moderate)
	SF-36	NA	NA	NA	NA	NA	NA	NA	- (Low)
VISION-SPECIFIC QOL	NEI VFQ-25	- (Low)	+ (Low)	- (Low)	+ (Low)	NA	NA	+ (High)	NA
	VDQ	? (High)	+ (Moderate)	+ (High)	? (Moderate)	? (High)	NA	+ (Moderate)	NA
GLAUCOMA-SPECIFIC QOL	GQL-15	? (Moderate)	+ (High)	+ (Moderate)	+ (High)	? (High)	NA	+ (High)	NA
	GSS	- (Low)	+ (Moderate)	+ (Moderate)	+ (Moderate)	? (High)	NA	+ (Moderate)	NA

CONCLUDING REMARKS

GQL, GSS and NEI-VFQ are the 3 most commonly used questionnaires in research setting, having considerable validation in a glaucoma patient population. Limited reports on interpretability, responsiveness, and feasibility in all 43 identified instruments make identification of a single optimal questionnaire for clinical use challenging and in need of further studies.