

REVIEW

Open Access

Patient-reported outcome measures for retinoblastoma: a scoping review



Ana Janic^{1,2,3,4}, Sylvie Bowden^{1,2}, Sarah Levy^{2,3}, Jennifer Stinson^{3,4,5} and Helen Dimaras^{1,2,3,6,7*} 

Abstract

Background: Retinoblastoma is a childhood retinal cancer with lifelong consequences such as vision loss and increased risk of second cancer. Patient-reported outcome measures (PROMs) are instruments that measure outcomes related to health directly reported by patients. The purpose of this study was to determine the scope, characteristics and quality of PROMs used in retinoblastoma and related fields of pediatric ophthalmology and pediatric oncology.

Methods: Databases MEDLINE and Embase were searched for studies in the English language that reported on PROMs used in retinoblastoma, pediatric oncology, or pediatric ophthalmology; grey literature and studies reporting on developmental PROM phases were excluded. PROMs were grouped by the construct measured and domains assessed, and classified as condition-specific or generic. A subsequent search was then conducted in MEDLINE and Embase for studies assessing measurement properties of the identified PROMs. PROMs with associated studies were assessed for their methodologic quality using the COnsensus-based standard for the Selection of health Measurement INstruments (COSMIN) strategy.

Results: Among 110 eligible studies uncovered by the database searches, 143 PROMs were identified: one retinoblastoma-specific, 56 ophthalmology- and 86 oncology-related. The most common construct measured was 'health-related quality of life' and the most common domain assessed was emotional well-being. Of the 143 PROMs, 100 had associated validation studies; the one retinoblastoma-specific PROM was not validated. Quality assessment revealed 34/100 PROMs received a score of sufficient quality in both subcategories of 'overall content validity'; 3/100 received a score of sufficient quality in both subcategories of 'internal structure'; 0/100 received a score of sufficient quality in all three subcategories of 'remaining measurement properties'. The Patient-Reported Outcome Measure Information System (PROMIS) Pediatric Profile-25 was the highest-scoring PROM identified, meeting COSMIN standards for 2/3 measurement property categories (and 5/7 subcategories). Eleven additional PROMs were identified which had sufficient scores in 1/3 measurement property categories (and 5/7 subcategories).

(Continued on next page)

* Correspondence: helen.dimaras@sickkids.ca

¹Faculty of Medicine, University of Toronto, 1 King's College Circle, Toronto M5S 1A8, Canada

²Department of Ophthalmology & Vision Sciences, The Hospital for Sick Children, 555 University Ave, Toronto M5G 2R3, Canada

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

(Continued from previous page)

Conclusion: The study identified several PROMs from the pediatric ophthalmology and pediatric oncology literature that could be relevant to the retinoblastoma population, but many have limits to their validation. Future development of a retinoblastoma-specific PROM, performed in partnership with retinoblastoma patients to support optimal content validity, could first focus on the selection and definition of the optimal construct to measure, followed potentially by adaptation and further validation of the relevant PROMs with strong methodologic quality identified in this study.

Keywords: Patient-reported outcome measures, Retinoblastoma, Pediatric ophthalmology, Pediatric oncology, Measurement properties, Quality assessment

Introduction

Retinoblastoma is a rare, intraocular cancer that arises in the developing retina, primarily due to bi-allelic inactivation of the *RBI* gene [1, 2]. With uniform global incidence, approximately 8000 children are diagnosed with the disease annually [1]. In high-income countries, the retinoblastoma survival rate is 96%, yet the cancer carries lifelong consequences that can affect vision and quality of life [3].

Treatment options for patients diagnosed with retinoblastoma vary depending on tumor size and location [3]. Smaller tumors located away from the center of vision are generally treated with focal laser therapy, while larger tumors could warrant enucleation (surgery to remove the eye), or more aggressive therapies to save the eye, like chemotherapy, or external beam radiation therapy [1, 3]. Treatment decisions also depend on whether or not tumor is present in one (unilateral) or both (bilateral) eyes [1, 3]. Unsurprisingly then, long-term clinical outcomes can differ among patients based on their initial diagnosis and treatment [3–6].

In addition to clinical outcomes, there are other consequences for patients that can impact a retinoblastoma survivor. For example, research indicates that survivors who have undergone enucleation or radiation therapy experience distress regarding their cosmetic appearance [3, 7, 8]. Having to wear a prosthetic eye or live with cranial-facial deformities induced by radiation, can foster self-consciousness and induce social anxiety [1, 3]. Furthermore, approximately half of survivors carry a constitutional *RBI* mutation (heritable retinoblastoma) [1, 2], which confers an increased risk of second malignancies and potential to pass the disease causing allele to future offspring [1, 3, 4]. Survivors and their caregivers have indicated that the risks associated with heritable retinoblastoma are a source of worry and anxiety [4–6, 9].

There is an increasing emphasis on patient-focused evaluation of disease impact [10–12]. This has led to the creation of instruments intended to capture the subjective assessment of health outcomes from the patient's perspective. Known as patient-reported outcome measures (PROMs), these instruments capture a variety of health-

related outcomes and are directly reported on by patients themselves [13–15]. PROMs can be administered via hard copy or electronic self-report questionnaires, as well as through in-person or telephone interviews [13–15].

Over the past 50 years, PROMs have been used in clinical trials [16–18], national audits [19, 20], and registers for condition-specific databases [21]. During the past two decades, widespread interest in using PROMs to track and monitor outcomes in clinical practice has further increased [12, 13]. PROMs can facilitate open communication between patients, families, and health care professionals [12, 18]. They provide unique information on the impact of a medical condition and its treatment from the patients' perspective [13, 18, 22], and can be used for continual monitoring of patient progress, offering the possibility of early detection of secondary problems [18, 22, 23]. Further, PROMs can provide researchers and clinicians with valuable data that can catalyze more focused research, answer currently pending research questions, or inform clinical decision making [22].

It follows then, that a retinoblastoma-specific PROM could be relevant for a population with such significant and varied outcomes over their life course. There are a number of PROMs used in pediatric oncology and pediatric ophthalmology [24–27], yet it is unclear if (i) these are used or relevant for retinoblastoma or (ii) condition-specific measures for retinoblastoma exist. Before developing a retinoblastoma-specific PROM, it is important to first identify PROMs that might be relevant for retinoblastoma patients. It is equally important to assess which outcomes are evaluated by these PROMs and if the PROMs are methodologically sound – that is, which PROMs measure what is meant to be measured accurately (validity) and do they measure what is meant to be measured consistently (reliability).

Methods

Primary objective

To identify and characterize PROMs currently used for retinoblastoma, pediatric ophthalmology and pediatric oncology research.

Search strategy

A search strategy was developed with the assistance of an information scientist. The search was conducted in databases MEDLINE and Embase. These databases were posited to be most relevant to the study aims due to their focus on biomedical life science and therefore ensured comprehensive capture of articles. Search terms included “Patient-Reported Outcome” AND “Patient-Reported Outcome Measure” AND [“exp retinoblastoma” OR “exp Ophthalm” OR “exp Neoplasm”], in studies on populations aged 0–18 published in 2004 or later (the search terms “Patient-Reported Outcome” and “Patient-Reported Outcome Measure” were not indexed in MEDLINE or Embase prior to 2004, precluding the capture of relevant articles before this date). The search took place on February 10th, 2019 and was updated June 6th, 2020.

Eligibility criteria

Records in English that reported on PROMs used in retinoblastoma, pediatric oncology, or pediatric ophthalmology were included in the review. Grey literature, including conference abstracts and non-peer reviewed text, and records that reported on the preliminary phases of PROM development were excluded.

Study selection

The database searches identified study titles and abstracts which were screened for eligibility. All relevant full-text articles were assessed based on inclusion criteria by two independent reviewers (AJ, SL). Disagreements were resolved by a tie-breaker decision from a third study author (HD). Included full-text articles were read by study authors (AJ, SL) to identify unique PROMs.

Descriptive analysis

All unique PROMs underwent descriptive analysis. First, PROMs were classified by the construct they measured, based on the information provided in the included full-text articles (the research team did not independently assign construct labels). PROMs measuring the same constructs were grouped and the domains assessed by these PROMs were identified and tabulated.

PROMs were also classified as generic or condition-specific; generic PROMs are those that can assess a wide variety of conditions, and condition-specific PROMs are as those that examine the specific outcomes of a particular condition or aspect of care [10].

Secondary objective

To evaluate the methodologic quality of the identified PROMs through application of the COSMIN ‘Good Measurement Property’ checklist.

Search strategy

Each PROM identified in the primary objective was subject to a subsequent, individual search. This separate search was necessary, in order to identify articles that assess the measurement properties of the identified PROMs. Databases MEDLINE and Embase were searched using terms “Patient-Reported Outcome”, the specific name of the PROM, and “exp Validation”, all limited to ages 0–18. The searches were conducted between May 20th and June 10th, 2019.

To audit the search strategy, one study author (AJ) randomly selected 10% of PROMs and determined the number of articles assessing measurement properties associated with each PROM. Performing an audit on 10% of included PROMs was guided by recommendations in literature [28, 29]. An additional search was then performed in MEDLINE and Embase, using an alternative search strategy recommended by COSMIN and created with the help of an information scientist. New search terms included “Patient-Reported Outcome” AND “Patient-Reported Outcome Measure” AND [“exp Retinoblastoma” or “exp Ophthalm” or “exp Neoplasm”] AND “specific PROM name” AND [“exp Validation” OR “exp Psychometrics”] for populations aged 0–18. The articles identified by the alternative search strategy were compared with the articles originally identified.

Eligibility criteria and study selection

Studies that reported on the quantitative evaluation of measurement properties for any of the previously identified PROMs were included and passed on to quality assessment.

Methodologic quality assessment

The COSMIN ‘Good Measurement Property’ checklist was used to determine the methodological quality of each PROM [30]. The ‘Good Measurement Property’ checklist assesses PROM quality across 10 measurement properties, distributed among three categories: (i) ‘overall content validity’ with subcategories ‘PROM development’ and ‘content validity’; (ii) ‘internal structure’ with subcategories ‘structural validity’ and ‘internal consistency’; and (iii) ‘remaining measurement properties’ with subcategories ‘reliability’, ‘measurement error’, ‘construct validity’, ‘criterion error’, ‘cross-cultural validity’, and ‘responsiveness’. COSMIN guidelines allow authors to determine which of the 6 ‘remaining measurement properties’ to include in their review; ‘reliability’, ‘measurement error’, and ‘construct validity’ were chosen for this study. The definition and judgment criteria for each measurement property followed COSMIN guidelines, as previously published [31].

All identified studies evaluating a given PROM’s measurement properties were read in full and measurement

properties were identified. Sufficiency of each measurement property was judged against established criteria [31] and given a score of '+', '-', or '?'; if an individual measurement property fulfilled the established criteria it received '+' indicating sufficient quality, if the property was not reported on it received '?', and if the property did not fulfill the established criteria, it received '-' to indicate insufficient quality. If identified studies reported discordant measurement property values for an instrument, the review team used the lower-score, i.e., '?' is higher than '-', consistent with COSMIN recommendations [31]. Two reviewers (AJ, SB) independently applied the good measurement property checklist across all identified studies evaluating measurement properties, for all PROMs. Discrepancies were resolved by a third reviewer (HD).

Finally, all PROMs, stratified by construct, were ranked based on the quality of individual measurement properties; a '+' value was determined to be the highest value and '-' the lowest value. Within a construct group, PROMs that received the most '+' values across 7 individual properties were ranked highest.

Results

Primary objective

Study selection

The MEDLINE and Embase search identified 533 unique manuscripts and after initial screening 145 were

excluded (142 were grey literature and three were non-English studies). Full-text review of the remaining 388 articles yielded 110 manuscripts that met inclusion criteria (Fig. 1). These articles identified 143 unique PROMs: one used for retinoblastoma, 56 used in pediatric ophthalmology broadly, and 86 used in pediatric oncology broadly (Fig. 2). The inter-rater agreement between reviewers on screening and full-text review was 83% (444/533 manuscript agreement), adhering to standards in literature [32, 33].

Descriptive analysis

Constructs and domains Descriptive analysis revealed the 6 constructs measured by 143 PROMs included: 'Health-Related Quality of Life' (HRQoL) (35%, 50/143), 'Psychological Impact' (29%, 42/143), 'Symptom Assessment' (22%, 37/143), 'Functional Vision' (FV) (4%, 6/143), 'Vision-Related Quality of Life' (VRQoL) (4%, 5/143), and 'Functional Ability' (2%, 3/143) (Fig. 3). One retinoblastoma-specific PROM was identified, measuring 'HRQoL'. The majority of PROMs used in pediatric oncology measured 'HRQoL' (43%, 37/86), followed by 'Symptom Assessment' (30%, 26/86), 'Psychological Impact' (23%, 20/86), and 'Functional Ability' (4%, 3/86). Of PROMs used in pediatric ophthalmology, the majority measured 'Psychological Impact', (39%, 22/56) followed by 'HRQoL' (21%, 12/56), 'Symptom

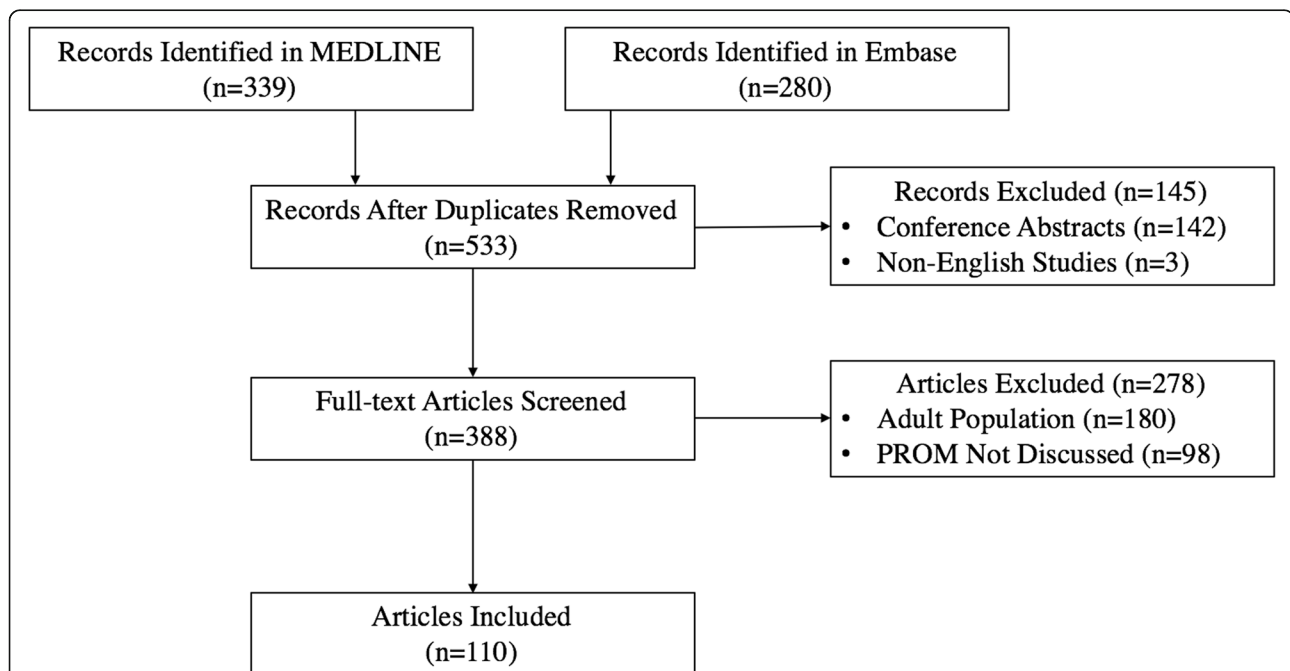
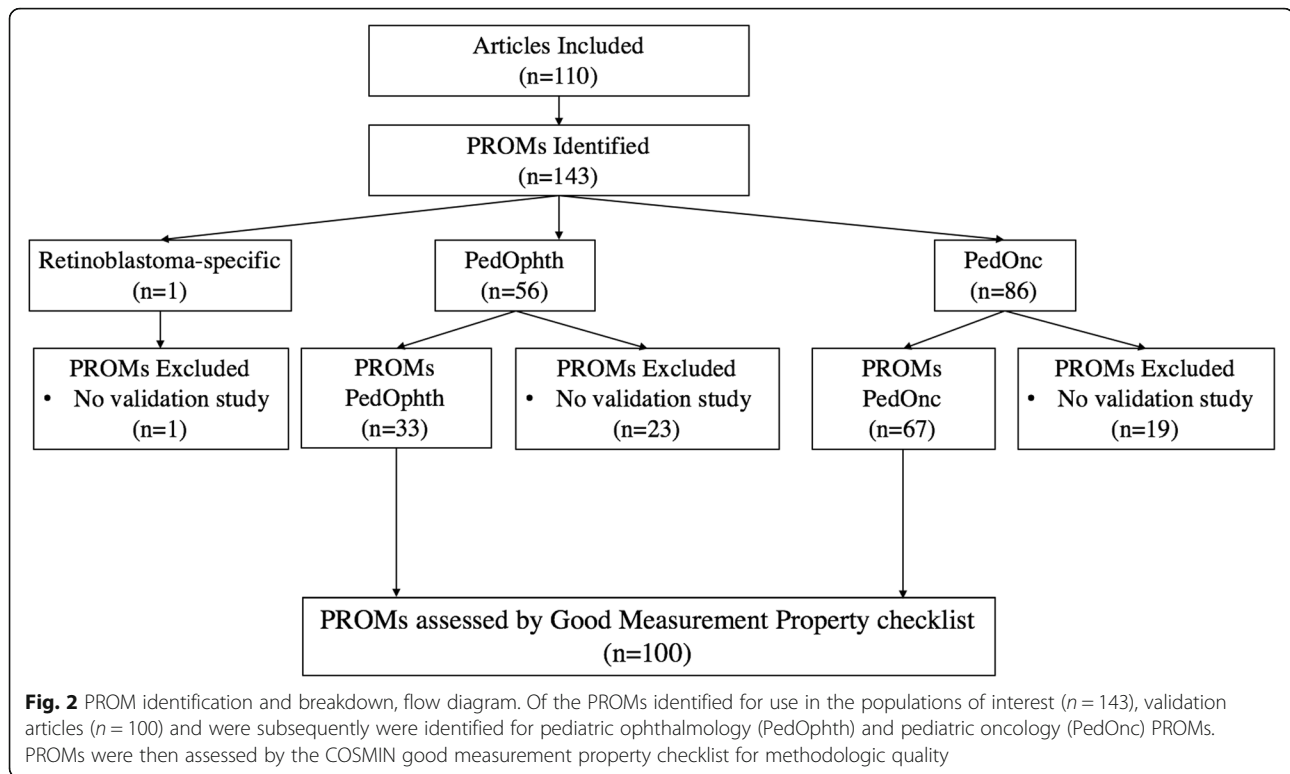


Fig. 1 Process of study selection, consort diagram. Unique articles identified by comprehensive database searches ($n = 533$) were screened by title and abstract against inclusion and exclusion criteria. Full text articles ($n = 388$) were then reviewed against inclusion and exclusion criteria. Full-text review resulted in 110 included articles



Assessment' (20%, 11/56). 'FV' (11%, 6/56), and 'VRQoL' (9%, 5/56).

Identification and tabulation of domains indicated emotional wellbeing to be assessed most often (80%, 110/143), followed by general health (63%, 90/143), symptom assessment (50%, 71/143), functioning (35%, 51/143), relationships (34%, 48/143), daily impact (23%, 33/143), family impact (17%, 25/143), future (16%, 23/143), independence (15%, 22/143), treatment (13%, 19/143), appearance (8%, 11/143) and FV (7%, 10/143) (Fig. 4).

Target population Of the 143 PROMs, 61 (43%) were condition-specific measures and 82 (57%) were generic measures (AdditionalDataTable). Condition-specific measures were designed for retinoblastoma (1/61, 1%), other eye conditions (27/61, 44%), other cancers (18/61, 29%), and miscellaneous conditions (16/61, 26%). Of the PROMs assessing eye conditions, 13/27 (48%) assessed strabismus, 9/27 (33%) assessed amblyopia, and 5/27 (19%) assessed others (i.e. rhinoconjunctivitis, cataracts, nasolacrimal pathway obstruction, exotropia, and vernal keratoconjunctivitis). Of PROMs assessing cancers, 16/18 (89%) were used for non-specific invasive malignancies such as leukemia or lymphoma, while 2/18 (11%) were used for patients undergoing bone marrow transplant (AdditionalDataTable).

Secondary objective

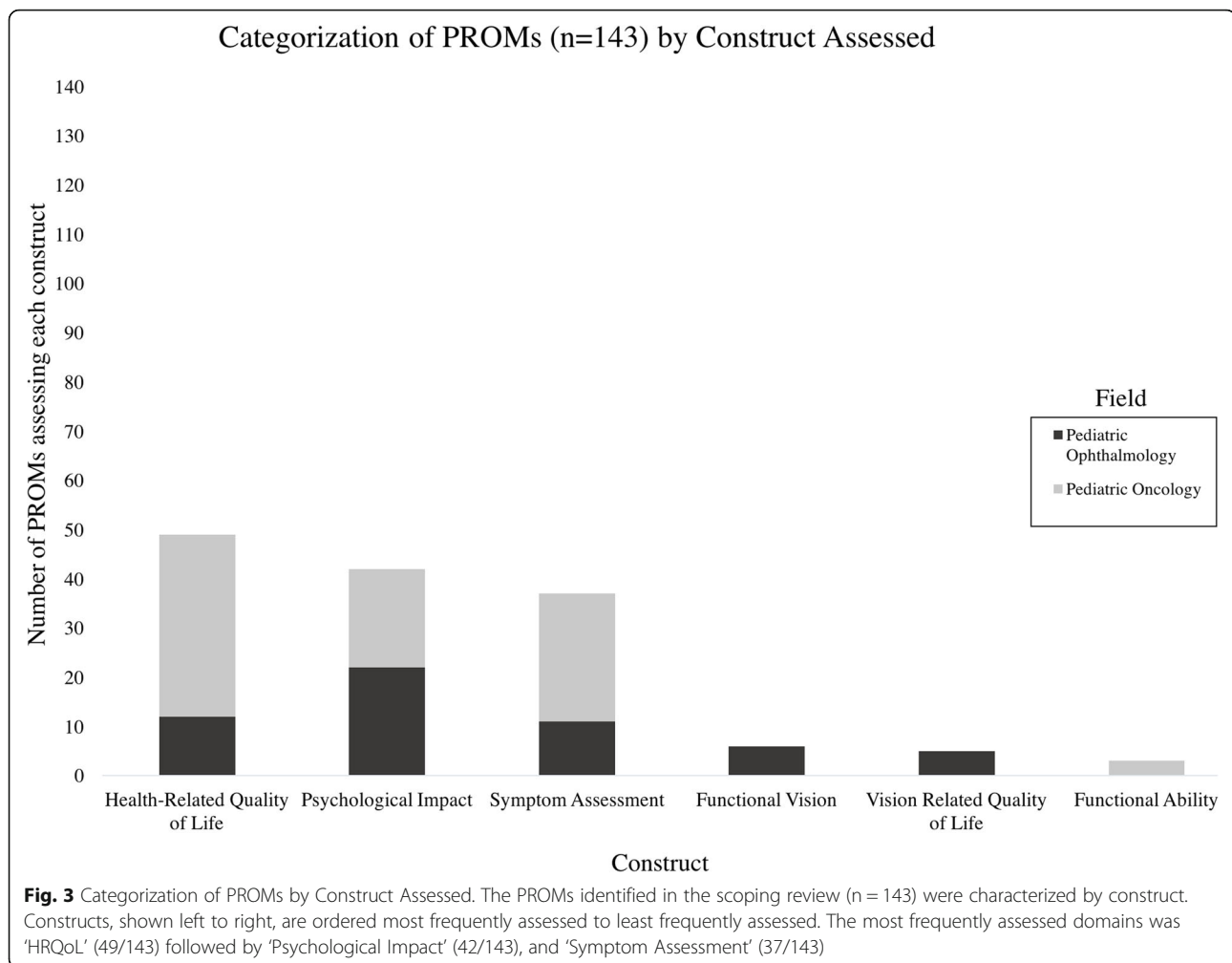
Study selection

Of the 143 identified PROMs, the secondary MEDLINE and Embase searches identified articles assessing measurement properties for 100 PROMs – 33 used in pediatric ophthalmology and 67 used in pediatric oncology (Fig. 2). Notably, the retinoblastoma-specific PROM did not have any associated studies that assessed any measurement properties. A total of 191 studies (mean 1.91, median 2) were identified from database searches for the 100 PROMs.

The audit performed on 10% of PROMs revealed 95% agreement between the two search strategies, identifying 2 additional studies assessing the measurement properties of 2 PROMs; 33 studies were identified by through the initial method and 35 identified through the subsequent search. The two newly identified studies underwent COSMIN quality analysis and the results did not impact the original quality score.

Methodologic quality assessment results

The quality score of all PROMs that had associated measurement property assessment studies in both pediatric ophthalmology and oncology, stratified by construct, is summarized in Table 1. Analysis using the COSMIN 'Good Measurement Property' checklist revealed 34/100 PROMs received a score of sufficient quality in both subcategories of 'overall content validity':

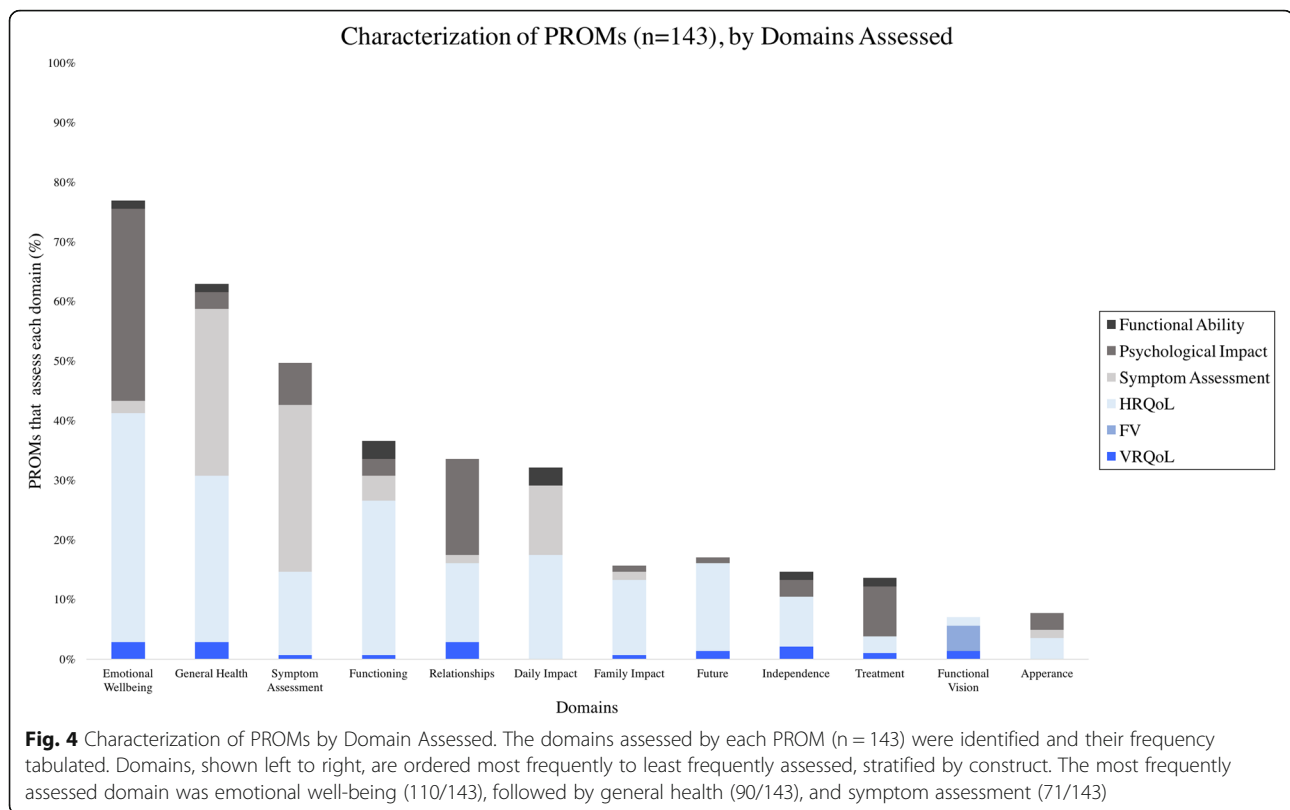


15 used in pediatric ophthalmology and 19 used in pediatric oncology. Additionally, 3/100 received a score of sufficient quality in both subcategories of 'internal structure': one used in pediatric ophthalmology and two used in pediatric oncology. No identified PROM received a score of sufficient quality in all three subcategories of 'remaining measurement properties', however, 19/100 PROMs received a score of sufficient quality in two of its three subcategories, reliability and construct validity.

The Patient-Reported Outcome Measure Information System (PROMIS) Pediatric Profile-25 [34–36], a measure commonly used in pediatric oncology, was the only PROM to receive a score of sufficient quality in both 'overall content validity' and 'internal structure'. The PROMIS Pediatric Profile-25 was also one of only 5 PROMs used in pediatric oncology to receive a score of sufficient quality in 5 of 7 measurement property subcategories (Table 1). The other PROMs meeting this score included; two 'HRQoL' measures: Pediatric Rhinoconjunctivitis Quality of Life Questionnaire (PRQLQ) [37,

38] and the Early-onset Scoliosis Questionnaire (EOSQ) [39–41], one 'Functional Ability' measure: Toronto Extremity Salvage Score (TESS) [42], and one 'Psychological Impact' measure: the Derriford Appearance Scale (DAS) [43, 44] (Table 1). Additionally, all of these PROMs had good 'overall content validity'; however, unlike PROMIS Pediatric Profile-25, they did not score well on internal structure.

The 6 highest-scoring PROMs used in pediatric ophthalmology each received a score of sufficient quality in four of 7 measurement property subcategories (Table 1). These included four 'VRQoL' measures: National Eye Institute-Visual Function Questionnaire (NEI-VFQ) [45, 46], Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) [47, 48], Vision-Related Quality of Life for Children and Young People (VQoL_CYP) [49, 50], Impact of Vision Impairment for Children (IVI_C) [51, 52] and two measures of 'Functional Vision': Functional Vision Quality of Life for Children and Young People (FVQ_CYP) [53] and Cardiff Visual Ability Questionnaire for Children (CVAQC) [54]. Of these, 4/6 (i.e. VQoL_CYP,



IVI_C, FVQ_CYP and CVAQC) received a score of sufficient quality in 'overall content validity'.

The inter-rater percent agreement and Kappa coefficient between quality scores from independent reviewers across all measurement properties were 92% and 0.79, respectively; both values adhere to standards in literature [32, 33, 55]. The percent agreement and Kappa coefficients for individual categories was as follows: PROM development (91%, 0.82), content validity (90%, 0.79), structural validity (92%, 0.65), internal consistency (95%, 0.87), reliability (78%, 0.65), measurement error (96%, 0.78), and construct validity (85%, 0.98). There were no discordant measurement property values reported between studies analyzing the same PROM.

Discussion

General discussion

This review identified 143 PROMs that have been used to assess outcomes in retinoblastoma, pediatric ophthalmology, and pediatric oncology populations. One retinoblastoma-specific PROM was identified, however, it was not built with help of patients and its methodologic quality has yet to be assessed [56]. Several PROMs with variable validity and reliability were identified that are currently used in pediatric ophthalmology and oncology. No PROMs received a perfect quality score indicating that all could benefit from further psychometric

improvements, such as improving their content validity, increasing reliability, or defining measurement error associated with use.

Patient participation in the development of PROMs is of paramount importance and is becoming more commonplace [30, 57–59]. Conversations with patients can help highlight important domains of the construct to be assessed, which is essential for strong content validity, i.e. the degree to which the content of an instrument reflects the underlying construct [13, 60]. The results of this research indicated that of all PROMs that received a sufficient score in 'overall content validity, over two-thirds measured 'HRQoL'. In contrast, few PROMs measuring 'Symptom Assessment' and 'Psychological Impact' received a sufficient score in 'overall content validity'; the vast majority of these PROMs did not meet COSMIN criteria in either subcategory of 'overall content validity', implying insufficient patient involvement in their development. Possible explanations for low patient involvement during the development of the 'Symptom Assessment' construct have been suggested previously. For example, patient input might be deemed unnecessary when the specific condition in question is well documented clinically, with the presumption that no new information would be gleaned from patient participation [61]. More broadly in PROM development, absence of patient involvement could be due to time and

Table 1 Quality analysis of all identified PROMs ($n = 100$). The COSMIN ‘good property checklist’ was applied on 100 PROMs, across 7 measurement properties stratified by construct. PROMs were ranked highest score of sufficient quality to lowest, within a construct. ‘+’ indicates a score of sufficient quality, ‘?’ indicates that the specific measurement property was not reported on, and ‘-’ indicates insufficient quality

PROM	COSMIN Good Measurement Properties Checklist						
	Overall Content Validity		Internal Structure		Remaining Measurement Properties		
	PROM Development	Content Validity	Structural Validity	Internal Consistency	Reliability	Measurement Error	Hypotheses testing for construct validity
VRQoL							
Vision Related Quality of Life (VQoL_CYP)	+	+	?	+	-	?	+
Impact of Vision Impairment for Children (IV_C)	+	+	?	+	-	?	+
National Eye Institute-Visual Function Questionnaire (NEI-VFQ)	?	+	?	+	+	?	+
Effects of Youngsters’ Eyesight on Quality of Life (EYE-Q)	+	-	?	+	+	?	+
Childrens Visual Function Questionnaire (CVFQ)	-	-	?	-	-	?	?
FV							
Functional Vision Quality of Life (FVQ_CYP)	+	+	?	+	?	?	+
Cardiff Visual Ability Questionnaire for Children (CVAQC)	+	+	-	+	+	?	?
LV Prasad-Functional Vision Questionnaire (LVP-FVQ II)	+	+	?	?	+	?	?
Visual Analogue Scale	?	?	?	?	+	?	?
Visual Function Index-14	-	-	?	+	?	?	?
HRQoL							
Patient Reported Outcome Measure Information System (PROMIS) Pediatric Profile-25	+	+	+	+	?	?	+
Pediatric Rhinoconjunctivitis Quality of Life Questionnaire (PRQLQ)	+	+	?	+	+	?	+
Early-Onset Scoliosis Questionnaire (EOSQ)	+	+	?	+	+	?	+
Adult Strabismus-20 (AS-20)	+	+	?	+	+	?	?
Amblyopia and Strabismus Questionnaire (A&SQ)	+	+	?	+	+	?	?
Bt-Dux	+	+	?	+	?	?	+
Minneapolis-Manchester Quality of Life (MMQL) Questionnaire	+	+	?	+	-	?	+
KIDSCREEN 52	+	+	?	+	-	?	+
Child Health Ratings Inventories (CHRIs - HSCT)	+	+	?	+	-	?	+
Lee Chronic Graft Versus Host Disease (Lee-cGVHD)	+	+	?	+	-	?	+
Child Health and Illness Profile (CHIP)	+	+	?	+	-	?	+
Quality of Life in Children with Vernal Keratoconjunctivitis (QUICK)	+	+	?	+	-	?	+
Intermittent Exotropia Questionnaire (ITXQ)	+	+	-	+	+	?	?
EuroQol 5-Dimension Youth (EQ-5D-Y)	+	+	?	?	+	?	+
Royal Marsden Hospital Pediatric Oncology Quality of Life Questionnaire (RMH-PQLQ)	+	+	?	+	?	?	?
Nasolacrimal Duct Obstruction (NLDO) Questionnaire	+	+	?	+	?	?	?
The Cancer Needs Questionnaire-Young People (CNQYP)	+	+	?	+	-	?	?

Table 1 Quality analysis of all identified PROMs ($n = 100$). The COSMIN ‘good property checklist’ was applied on 100 PROMs, across 7 measurement properties stratified by construct. PROMs were ranked highest score of sufficient quality to lowest, within a construct. ‘+’ indicates a score of sufficient quality, ‘?’ indicates that the specific measurement property was not reported on, and ‘-’ indicates insufficient quality (*Continued*)

PROM	COSMIN Good Measurement Properties Checklist						
	Overall Content Validity		Internal Structure		Remaining Measurement Properties		
	PROM Development	Content Validity	Structural Validity	Internal Consistency	Reliability	Measurement Error	Hypotheses testing for construct validity
Quality of Life for Children with Cancer (QOLCC)	+	+	-	+	?	?	?
Children’s Amblyopia Treatment Quality of Life Questionnaire (CAT-QoL)	+	+	-	+	?	?	-
Health Utilities Index (HUI)	+	+	?	?	?	?	+
Coddington Life Events Questionnaire (CLES)	+	+	?	?	-	?	+
Perceived Illness Experience (PIES)	+	+	?	-	?	?	+
TNO-AZL Preschool Children Quality of Life (TAPQOL)	+	+	?	-	-	?	+
Pediatric Quality of Life Inventory—Cancer Module (PedsQL Cancer)	+	?	?	+	+	?	?
MPQOLQ Miami Pediatric Quality of Life Questionnaire	+	?	?	+	+	?	?
M.D. Anderson Symptom Inventory	-	-	?	+	+	?	+
PainDETECT	+	-	?	+	?	?	?
World Health Organisation Quality of Life assessment (WHOQoL) BREF	-	+	-	+	?	?	+
The Hope Scale	-	+	?	+	?	?	?
The Adolescent Quality of Life Questionnaire	-	+	?	?	?	?	+
KINDL	?	?	?	+	+	?	?
Medical Outcomes Study Short-Form 36 (SF-36)	-	-	?	+	?	?	+
Infant Toddler Quality of Life Inventory (ITQOL)	-	-	?	+	?	?	+
Functional Assessment of Cancer Therapy - General (FACT - G)	?	?	?	+	?	?	+
16-Dimensional Health-related Measure (16D)	-	+	?	?	?	?	?
Child Health Questionnaire (CHQ)	?	?	?	+	?	?	?
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30)	?	?	-	?	?	?	+
Pediatrics Outcomes Data Collection Instrument (PODCI)	?	?	-	?	?	?	?
Behavioral, Affective, and Somatic Experiences Scale (BASES)	-	-	?	+	?	?	?
Scoliosis Research Society [SRS]-30	?	?	?	-	?	?	?
Pediatric Quality of Life (PEDQOL)	-	-	?	-	?	?	?
Functional Ability							
Toronto Extremity Salvage Score (TESS)	+	+	?	+	+	?	+
DISABKIDS	+	+	?	+	?	?	+
The Activities Scale for Kids Performance Version (ASKp)	+	-	?	?	+	?	+
Symptom Assessment							
Amblyopia Treatment Index (ATI)	+	+	?	+	+	?	?
Chronic Otitis Media Questionnaire-12 (COMQ-	-	+	+	+	-	?	+

Table 1 Quality analysis of all identified PROMs ($n = 100$). The COSMIN ‘good property checklist’ was applied on 100 PROMs, across 7 measurement properties stratified by construct. PROMs were ranked highest score of sufficient quality to lowest, within a construct. ‘+’ indicates a score of sufficient quality, ‘?’ indicates that the specific measurement property was not reported on, and ‘-’ indicates insufficient quality (*Continued*)

PROM	COSMIN Good Measurement Properties Checklist						
	Overall Content Validity		Internal Structure		Remaining Measurement Properties		
	PROM Development	Content Validity	Structural Validity	Internal Consistency	Reliability	Measurement Error	Hypotheses testing for construct validity
12)							
Oral Health Impact Profile (OHIP-14)	-	+	?	+	-	?	+
Hip disability Osteoarthritis Outcome Score (HOOS)	+	-	?	+	+	?	+
The Cancer and Treatment Distress (CTXD)	-	+	?	+	?	?	?
Children’s Vision for Living Scale (CVLS)	+	+	-	+	?	?	?
Oral Mucositis Daily Questionnaire (OMDQ)	-	+	?	-	?	?	+
Brief Fatigue Inventory	?	?	?	+	+	?	+
Impact of Events Scale-Revised (IES-R)	?	?	-	+	+	?	+
Memorial Symptom Assessment Scale	-	+	?	+	?	?	+
Fatigue Symptom Inventory	-	+	?	+	?	?	+
Adolescent Pediatric Pain Tool (APPT)	+	-	?	+	?	?	+
Knee-injury osteoarthritis outcome score (KOOS)	+	-	?	?	+	?	+
Hearing Measurement Scale	-	+	?	?	?	?	?
Short Form McGill Pain Questionnaire	-	+	-	+	?	?	?
Expectations of Strabismus Surgery Questionnaire	-	+	?	?	?	?	?
Kessler-6 Two Factor	-	-	?	+	?	?	+
Colored Analogue Scale	-	-	?	?	+	?	+
Edmonton Symptom Assessment Scale (ESAS)	?	?	?	+	?	?	?
Wong-Baker FACES Pain Rating Scale	?	?	?	?	?	?	+
Symptom Checklist 90 R	-	-	?	?	?	?	+
Brief Pain Inventory Short Form	-	-	?	?	?	?	?
Emotional Impact of Amblyopia Questionnaire (EIAQ)	-	-	?	?	?	?	?
Psychological Impact							
Derriford Appearance Scale (DAS)	+	+	?	+	+	?	+
Multidimensional Scale of Perceived Social Support (MSPSS)	?	+	?	+	+	?	+
Hospital Anxiety and Depression Scale (HADS)	-	+	+	+	?	?	+
Screen for Child Anxiety Related Emotional Disorders (SCARED)	-	+	?	+	+	?	+
Zung Self-Rating Depression Scale	+	?	?	+	?	?	+
Beck Anxiety Inventory	-	+	-	+	?	?	+
Hopkins Symptom Check List (HSCL)	-	-	?	+	+	?	+
Fear of Negative Evaluation (FNE)	-	-	?	+	+	?	+
Liebowitz Social Anxiety Scale	-	-	-	+	+	?	+
Psychological Impact Questionnaire (PIQ)	+	+	?	?	?	?	?
Childrens Manifest Anxiety Scale (CMAS)	-	+	?	+	?	?	?
Hamilton Anxiety Rating Scale	?	?	-	+	?	?	+

Table 1 Quality analysis of all identified PROMs ($n = 100$). The COSMIN ‘good property checklist’ was applied on 100 PROMs, across 7 measurement properties stratified by construct. PROMs were ranked highest score of sufficient quality to lowest, within a construct. ‘+’ indicates a score of sufficient quality, ‘?’ indicates that the specific measurement property was not reported on, and ‘-’ indicates insufficient quality (*Continued*)

PROM	COSMIN Good Measurement Properties Checklist						
	Overall Content Validity		Internal Structure		Remaining Measurement Properties		
	PROM Development	Content Validity	Structural Validity	Internal Consistency	Reliability	Measurement Error	Hypotheses testing for construct validity
Distressed Personality Questionnaire (DS-14)	?	?	?	+	-	?	+
Self-Perception Profile for Children (SPPC)	?	?	-	+	-	?	+
Behavioural Assessment System for Children (BAS-C)	-	-	+	+	?	?	-
Social Anxiety Scale for Children-Revised	-	-	?	+	?	?	+
Mental Health Inventory (MHI)	-	-	?	+	?	?	+
Center for Epidemiologic Studies Depression Scale – Revised (CESD-R)	-	-	-	+	?	?	+
Perceived Psychosocial Questionnaire (PPQ)	+	-	?	?	?	?	?
Children’s Depression Inventory (CDI)	?	?	?	+	-	?	-
State-Trait Anxiety Inventory	-	-	?	+	-	?	?
Pediatric Camp Outcome Measure (PCOM)	-	-	?	+	?	?	?
The Psychosocial Assessment Tool (PAT)	-	-	?	-	?	?	+
Faces Anxiety Scale	-	-	?	?	?	?	?

budget constraints that researchers face [61] when developing a PROM; the financial costs of patient involvement, frequently recognized in literature as a challenge of qualitative research broadly [62], can contribute to a researcher team’s decision to not involve patients. However, despite these challenges, researchers have started to shift their thinking and recognize the positive implications of patient participation and the value of patient experience [11, 24, 27, 30]. Additionally, several funding bodies and regulation bodies, such as the Canadian Institute for Health Research and the US Food and Drug Administration have made patient involvement in PROM development mandatory [63, 64].

The review identified no PROMs that received a sufficient quality score across all three ‘remaining measurement properties’ (i.e. reliability, construct validity, and measurement error). One explanation for this could be due to the strict statistical criteria outlined by the ‘Good Measurement Property’ checklist. The checklist uses well-established means for evaluating measurement properties [65], yet it provides a very narrow list of options for evaluation, often only accepting specific statistical tests. For instance, in practice, an acceptable evaluation of a PROM’s reliability can include determining the interclass coefficient, weighted kappa, or Pearson’s r Coefficient, yet the COSMIN ‘Good Measurement Property’ checklist only acknowledges the former two as acceptable values [30]. Therefore, PROMs demonstrating strong reliability solely through Pearson’s

r Coefficient would not meet COSMIN criteria. Additionally, the definition of reliability involves examining internal consistency, reliability, and measurement error; however, we found that while internal consistency and reliability were commonly reported, measurement error was often omitted. Beyond COSMIN, the recommendations regarding reporting of measuring property provides mixed direction; some experts recommend assessment and reporting of all measurement properties of broad reliability and validity [66], while others argue that related measurement properties should not be assessed or reported as separate dimension [67]. Thus, it remains unclear if the oft-omitted measurement error in the studies we assessed was actually assessed but not reported (to avoid redundancy), or if it was not assessed at all.

As the name implies, generic PROMs are intended for use on any disease population to allow for comparability across patient populations [10]. However, since generic measures are not designed to capture areas of concern to specific patient populations [12, 22, 68], they are likely to include irrelevant questions for certain patient groups or omit issues that are specific to a particular condition [68–70]. Use of generic measures to assess specific conditions may result in i) capturing data that are not meaningful to patients [22, 68], ii) failure to accurately measure what the PROM purports to measure [57, 58, 68], and iii) producing inaccurate comparisons between groups [68–70]. Where the content validity of generic PROMs may be questionable, condition-specific PROMs

are preferred [71]. Literature regarding development of condition-specific PROMs recommends patient co-development [12, 22, 57, 68], i.e. developing questions by means of qualitative methods with relevant patients and thoroughly testing the measurement properties of novel PROMs with new populations of patients [71, 72]. This results in PROMs that address aspects of an outcome that are important for a particular patient population, facilitating good validity and responsiveness of the measure. Yet, despite the described importance of patient co-development, our study identified several condition-specific self-report outcome measures which had poor PROM development and content validity. In particular, patients were not included in the development of many of these condition-specific PROMs (Table 1), therefore they may be lacking a focus on topics or questions meaningful to the patient experience.

Patient co-development of PROMs is likewise encouraged by theoretical frameworks used to guide construct development, as has been robustly demonstrated in development of several PROMs assessing 'HRQoL' [72, 73]. When there is an absence of clarity surrounding the initial definition of a construct, this can lead to: development of measures that provide inaccurate inferences about outcomes; difficulty deciding on what domains to include in a measure; and challenges in comparing, generalizing, and replicating research findings. Examples of such confusion include inconsistent definitions of 'HRQoL' used by several PROMs [74, 75], and conflation of related, but distinct, constructs, such as 'VRQoL' and 'FV' [24, 58, 76]. Prior research on pediatric oncology PROMs highlighted measures, of which several are assessed in this study, that were developed without precisely conceptualizing and defining the construct 'HRQoL' [75, 77]. One such measure is the Pediatric Quality of Life Inventory – Cancer Module (PedsQL - Cancer) [77]; this is consistent with our study, in which there was insufficient published data to allow evaluation of content and construct validity (Table 1). Perhaps then, informing the construct definition through insights gleaned from qualitative conceptual discussions with patients, parents, researchers, and health-care professionals, may improve several psychometric measurement properties [73, 78].

Looking towards the development of a retinoblastoma-specific PROM, care must be taken to ensure that the constructs and domains measured are relevant and valuable to the patient population. Some research groups have argued that 'HRQoL' is the most important construct to assess for retinoblastoma [7, 56], however the patient voice is largely missing from the literature. RetinoQuest [56], the retinoblastoma-specific PROM identified by our study, is an 'HRQoL' measure which has yet to be assessed for its methodologic quality. Furthermore,

RetinoQuest lacks a conceptual definition of 'HRQoL' and retinoblastoma patients were not consulted during its development. A way forward for this measure could be to work alongside retinoblastoma patients to first determine if this is the appropriate construct to assess, then refine the definition and theoretical underpinning of 'HRQoL'. Another approach could be to adapt an existing well-validated 'HRQoL' measure to the retinoblastoma context; for example, the PROMIS Pediatric Profile-25 proved strong methodological quality as assessed in our study, and it also demonstrates a well-defined theoretical underpinning of 'HRQoL' [75, 77].

Several research studies have indicated that 'Psychological Impact' maybe another important construct to consider for a retinoblastoma-specific PROM [3, 5, 9]. The DAS, a generic measure assessing the psychological impact of appearance, scored highest in our study in the 'Psychological Impact' construct category [43, 44], receiving a score of sufficient quality for 'overall content validity'. Since retinoblastoma patients and their families have voiced concerns regarding cosmetic appearance after treatment [3, 5], perhaps with appropriate adaptation and methodologic assessment, the DAS can be explored for use in the retinoblastoma population. Alternatively the Psychological Impact Questionnaire (PIQ) [79], a condition-specific amblyopia measure, could similarly be explored for relevance. The PIQ was the only other 'Psychologic Impact' PROM to receive a score of sufficient quality for 'overall content validity', but did not report on any of the other remaining measurement properties. A future step could be to utilize qualitative methods to determine if 'Psychological Impact' is truly an important construct to evaluate for retinoblastoma, and if it is, what facets of 'Psychological Impact' are valued by retinoblastoma patients.

Strengths and limitations

This review has several important strengths. First, it was conducted with methodologic rigor, adhering to COSMIN PROM evaluation guidelines, and used two independent reviewers for study selection (83% inter-rater agreement) and quality analysis (92%, 0.79 inter-rater percent agreement and Kappa Coefficient), reliability that is comparable to recommendations in literature [32, 33]. Second, our review revealed that numerous PROMs for children suffering from various ophthalmic-conditions or cancer whose measurement properties have not been sufficiently assessed for use in these populations, several with significant gaps in their methodologic quality. In keeping with guidelines in literature for the development of adult PROMs [80], a self-report outcome measure should not be put into practice without careful consideration of its psychometric strengths and limitations. Yet, all of the PROMs identified by this

review are in use today. Moving forward, researchers should consider assessing the methodologic quality of all these self-report outcome measures to follow standardized, international COSMIN guidelines before integrating the PROM into a clinical setting. This review has practical significance. Several generic and condition-specific PROMs with sufficient methodologic quality have been identified that could potentially be relevant for adaptation and application to retinoblastoma patients, with subsequent consultation with patients, experts, and pertinent literature. In this vein, to inform the development of a retinoblastoma-specific PROM, next steps could include qualitative discussions with retinoblastoma patients and their caregivers to uncover the treatment outcomes they find valuable and relevant to their care.

Our findings should be interpreted in light of several limitations. First, for practical reasons, we chose to only include English publications. However, as the initial search strategy identified three non-English manuscripts out of 533, it is reasonable to assume that there are unlikely to be many non-English studies relevant to this topic. Second, the authors acknowledge that more PROMs may have been identified if databases beyond just MEDLINE and Embase had been searched. However, our methods were consistent with search strategy recommendations of the COSMIN guidelines [30], and still yielded a relatively large number of relevant PROMs. Third, the initial keyword search, though developed with an information scientist, could have included more comprehensive search terms. Fourth, identification of validation articles for individual PROMs was performed via a separate search in MEDLINE and Embase, and was not included in the initial search strategy. The decision for the second search was necessary to reflect the changes made to the study protocol, the addition of a methodologic quality assessment. Therefore, this might account for articles discussing the analysis of some measurement properties that were not identified. However, in attempt to determine the accuracy of our search method, we audited 10% of the identified instruments and conducted a search following COSMIN systematic review guidelines [30]. A 10% audit is commonly applied in practice [32, 33]. The audit revealed only two new studies assessing measurement properties associated with two PROMs, suggesting that a larger audit would be unlikely to identify more studies. In addition, after quality analysis was applied to these studies, the associated PROM score did not change, suggesting that a larger audit would be unlikely to change the scores even if additional studies were retrieved. Finally, this review was not registered prior to its initiation. Although registration is not commonly required for scoping reviews, it does increase transparency and reproducibility and adherence to

preregistered protocol helps decrease reporting bias. However, the study protocol underwent review, by several study authors (AJ, JS, and HD), and no deviations to the methodology were made throughout the study.

Conclusion

In summary, this review identified many PROMs with variable reliability and validity that are currently used in retinoblastoma and related fields. Twelve PROMs were identified with good methodological quality, and based on the constructs and domains they assess, their potential application to retinoblastoma has been considered. Only one retinoblastoma-specific PROM, that was not developed with the help of patients and has yet to be validated, was identified. Therefore, moving forward it is critical to develop, validate, and implement outcome measures for retinoblastoma patients. Future research could involve patients and experts to jointly examine patient's perspectives and agree on the specific constructs to be measured by a future retinoblastoma PROM. Additionally, adapting the relevant high-scoring PROMs to include retinoblastoma-specific domains could be explored. This could be done through pursuing qualitative methods of patient participation to help identify which outcome domains are important to patients. Once constructs, domains, and the number of measures need have been identified, validating these measures on a retinoblastoma patient-population, in adherence to COSMIN guidelines, will be important in order to ensure thorough reliability and validity.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s41687-020-00232-7>.

Additional file 1: Additional Data Table. Characterization of Condition-specific PROMs ($n = 61$). Of the condition-specific PROMs identified, majority (27/61) were designed for ophthalmic conditions, followed by cancer-specific measures (18/61), miscellaneous conditions (16/61), and retinoblastoma (1/61).

Abbreviations

COSMIN: COnsensus-based standard for the Selection of health Measurement INstruments; FV: Functional Vision; HRQoL: Health-Related Quality of Life; PROM: Patient-Reported Outcome Measure; VRQoL: Vision-Related Quality of Life

Acknowledgements

We are grateful to Thomasin Webber and Lily Ren, information scientist at The Hospital for Sick Children, who helped develop the search strategy and Dr. Ajoy Vincent, staff ophthalmologist at The Hospital for Sick Children, for his important contributions to the manuscript development.

Authors' contributions

Conception of work: HD. Study design development: AJ, JS, HD. Data acquisition: AJ, HD. Analysis of data: AJ, SB, SL, HD. Interpretation of data and manuscript drafting: AJ, HD. Critical revision of manuscript: AJ, SB, SL, JS, HD. All authors read and approved the final manuscript.

Funding

Ana Janic is supported by The Hospital for Sick Children Research Training Competition (Restracomp) graduate scholarship and the University of Toronto Vision Science Research Program (VSRP) scholarship. This study was supported by The Hospital for Sick Children Ophthalmology Research Fund.

Availability of data and materials

The database records identified by the literature search and their characterization, as indicated in Figs. 1, 2, 3 and 4, are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Faculty of Medicine, University of Toronto, 1 King's College Circle, Toronto M5S 1A8, Canada. ²Department of Ophthalmology & Vision Sciences, The Hospital for Sick Children, 555 University Ave, Toronto M5G 2R3, Canada. ³Child Health Evaluative Sciences Program, SickKids Research Institute, 686 Bay Street, Toronto M5G 0A4, Canada. ⁴Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, 555 University Ave, Toronto M5G 2R3, Canada. ⁵Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, 155 College Street, Toronto M5T 1P8, Canada. ⁶The Centre for Global Child Health, SickKids Research Institute, 686 Bay Street, Toronto M5G 0A4, Canada. ⁷Division of Clinical Public Health, University of Toronto, 150 College Street, Toronto M5T 3M7, Canada.

Received: 24 April 2020 Accepted: 3 August 2020

Published online: 08 August 2020

References

- Dimaras, H., Corson, T. W., Cobrinik, D., et al. (2015). Retinoblastoma. *Nature Reviews Disease Primers*, 1(1), 15021. <https://doi.org/10.1038/nrdp.2015.21>.
- Thériault, B. L., Dimaras, H., Gallie, B. L., & Corson, T. W. (2014). The genomic landscape of retinoblastoma: A review: The genomic landscape of retinoblastoma. *Clinical & Experimental Ophthalmology*, 42(1), 33–52. <https://doi.org/10.1111/ceo.12132>.
- (2009). National retinoblastoma strategy Canadian guidelines for care. *Canadian Journal of Ophthalmology*, 44, S9–S47. <https://doi.org/10.3129/109-194>.
- Hill, J. A., Gedleh, A., Lee, S., Hougham, K. A., & Dimaras, H. (2018). Knowledge, experiences and attitudes concerning genetics among retinoblastoma survivors and parents. *European Journal of Human Genetics*, 26(4), 505–517. <https://doi.org/10.1038/s41431-017-0027-9>.
- Ford, J. S., Chou, J. F., Sklar, C. A., et al. (2015). Psychosocial outcomes in adult survivors of retinoblastoma. *Journal of Clinical Oncology*, 33(31), 3608–3614. <https://doi.org/10.1200/JCO.2014.60.5733>.
- Batra, A., Kumari, M., Paul, R., Patekar, M., Dhawan, D., & Bakhshi, S. (2016). Quality of life assessment in retinoblastoma: A cross-sectional study of 122 survivors from India: Quality of life in retinoblastoma survivors. *Pediatric Blood & Cancer*, 63(2), 313–317. <https://doi.org/10.1002/pbc.25781>.
- Friedman, D. N., Chou, J. F., Francis, J. H., et al. (2018). Vision-targeted health-related quality of life in adult survivors of retinoblastoma. *JAMA Ophthalmol*, 136(6), 637. <https://doi.org/10.1001/jamaophthalmol.2018.1082>.
- Dimaras, H., Kimani, K., Dimba, E. A., et al. (2012). Retinoblastoma. *Lancet*, 379(9824), 1436–1446. [https://doi.org/10.1016/S0140-6736\(11\)61137-9](https://doi.org/10.1016/S0140-6736(11)61137-9).
- van Dijk, J., Oostrom, K. J., Huisman, J., et al. (2010). Restrictions in daily life after retinoblastoma from the perspective of the survivors. *Pediatric Blood & Cancer*, 54(1), 110–115. <https://doi.org/10.1002/pbc.22230>.
- Patient-reported outcome measures (PROMs). Canadian Institute for Health Information. Accessed November 11, 2019. <https://www.cihi.ca/en/patient-reported-outcome-measures-proms>
- Wilson, R. (2018). Patient led PROMs must take Centre stage in cancer research. *Research Involvement and Engagement*, 4(1), 7, s40900-018-0092-0094. <https://doi.org/10.1186/s40900-018-0092-4>.
- Deshpande, P., Sudeepthi, B., Rajan, S., & Abdul Nazir, C. (2011). Patient-reported outcomes: A new era in clinical research. *Perspectives in Clinical Research*, 2(4), 137. <https://doi.org/10.4103/2229-3485.86879>.
- Dawson, J., Doll, H., Fitzpatrick, R., Jenkinson, C., & Carr, A. J. (2010). The routine use of patient reported outcome measures in healthcare settings. *BMJ*, 340(jan18 1), c186. <https://doi.org/10.1136/bmj.c186>.
- Raine, R., Fitzpatrick, R., Barratt, H., et al. (2016). Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. *Health Services and Delivery Research*, 4(16), 1–136. <https://doi.org/10.3310/hsdr04160>.
- Morris C, Gibbons E, Fitzpatrick R. Child and parent reported outcome measures: A scoping report focusing on feasibility for routine use in the NHS. University of Oxford Accessed November 17, 2019. https://phi.uhce.ox.ac.uk/pdf/PROMs_WithChildren_Oxford_2009.pdf.
- Silverman, L. R., Demakos, E. P., Peterson, B. L., et al. (2002). Randomized controlled trial of Azacitidine in patients with the Myelodysplastic syndrome: A study of the Cancer and leukemia group B. *Journal of Clinical Oncology*, 20(10), 2429–2440. <https://doi.org/10.1200/JCO.2002.04.117>.
- Grant, A. M., Wileman, S. M., Ramsay, C. R., et al. (2008). Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. *BMJ*, 337(dec15 2), a2664. <https://doi.org/10.1136/bmj.a2664>.
- Detmar, S. B., Muller, M. J., Schornagel, J. H., Wever, L. D. V., & Aaronson, N. K. (2002). Health-related quality-of-life assessments and patient-physician communication: A randomized controlled trial. *JAMA*, 288(23), 3027–3034. <https://doi.org/10.1001/jama.288.23.3027>.
- Williams, O., Fitzpatrick, R., Hajat, S., et al. (2002). Mortality, morbidity, and 1-year outcomes of primary elective total hip arthroplasty. *The Journal of Arthroplasty*, 17(2), 165–171. <https://doi.org/10.1054/arth.2002.29389>.
- Malchau, H., Garellick, G., Eisler, T., Kärrholm, J., & Herberts, P. (2005). PRESIDENTIAL GUEST ADDRESS: The Swedish Hip Registry: Increasing the Sensitivity by Patient Outcome Data. *Clinical Orthopaedics*, 441(NA), 19–29. <https://doi.org/10.1097/01.bl.0000193517.19556.e4>.
- Zanoli, G., Nilsson, L. T., & Strömqvist, B. (2006). Reliability of the prospective data collection protocol of the Swedish spine register: Test-retest analysis of 119 patients. *Acta Orthopaedica*, 77(4), 662–669. <https://doi.org/10.1080/17453670610012764>.
- Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ*, (jan28 1), 346, f167. <https://doi.org/10.1136/bmj.f167>.
- Engelen, V., van Zwieten, M., Koopman, H., et al. (2012). The influence of patient reported outcomes on the discussion of psychosocial issues in children with cancer. *Pediatric Blood & Cancer*, 59(1), 161–166. <https://doi.org/10.1002/pbc.24089>.
- Tadic, V., Hogan, A., Sobti, N., Knowles, R. L., & Rahi, J. S. (2013). Patient-reported outcome measures (PROMs) in paediatric ophthalmology: A systematic review. *The British Journal of Ophthalmology*, 97(11), 1369–1381. <https://doi.org/10.1136/bjophthalmol-2013-303350>.
- Papadakos, J. K., Charow, R. C., Papadakos, C. J., Moody, L. J., & Giuliani, M. E. (2019). Evaluating cancer patient-reported outcome measures: Readability and implications for clinical use. *Cancer*, 125(8), 1350–1356. <https://doi.org/10.1002/cncr.31928>.
- Zhang, B., Lloyd, W., Jahanzeb, M., & Hassett, M. J. (2018). Use of patient-reported outcome measures in quality oncology practice initiative-registered practices: Results of a National Survey. *Journal of Oncology Practice/ American Society of Clinical Oncology*, 14(10), e602–e611. <https://doi.org/10.1200/JOP.18.00088>.
- Denniston, A. K., Kyte, D., Calvert, M., & Burr, J. M. (2014). An introduction to patient-reported outcome measures in ophthalmic research. *Eye*, 28(6), 637–645. <https://doi.org/10.1038/eye.2014.41>.
- Bramer, W. M., Rethlefsen, M. L., Mast, F., & Kleijnen, J. (2018). Evaluation of a new method for librarian-mediated literature searches for systematic reviews. *Research Synthesis Methods*, 9(4), 510–520. <https://doi.org/10.1002/rsrm.1279>.
- Bramer, W. M., Giustini, D., Kleijnen, J., & Franco, O. H. (2018). Searching Embase and MEDLINE by using only major descriptors or title and abstract fields: A prospective exploratory study. *Systematic Reviews*, 7(1), 200. <https://doi.org/10.1186/s13643-018-0864-9>.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, 19(4), 539–549. <https://doi.org/10.1007/s11136-010-9606-8>.

31. Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., et al. (2018). COSMIN risk of Bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1171–1179. <https://doi.org/10.1007/s11136-017-1765-4>.
32. Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., & Klassen, T. P. (2006). Single data extraction generated more errors than double data extraction in systematic reviews. *Journal of Clinical Epidemiology*, 59(7), 697–703. <https://doi.org/10.1016/j.jclinepi.2005.11.010>.
33. Interrater Reliability. In: Encyclopedia of Research Design. SAGE Publications, Inc.; 2010. doi:<https://doi.org/10.4135/9781412961288.n194>
34. Tang, E., Ekundayo, O., Peipert, J. D., et al. (2019). Validation of the patient-reported outcomes measurement information system (PROMIS)-57 and -29 item short forms among kidney transplant recipients. *Quality of Life Research*, 28(3), 815–827. <https://doi.org/10.1007/s11136-018-2058-2>.
35. Cella, D., Riley, W., Stone, A., et al. (2010). The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*, 63(11), 1179–1194. <https://doi.org/10.1016/j.jclinepi.2010.04.011>.
36. DeWalt, D. A., Gross, H. E., Gipson, D. S., et al. (2015). PROMIS() pediatric self-report scales distinguish subgroups of children within and across six common pediatric chronic health conditions. *Quality of Life Research International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 24(9), 2195–2208. <https://doi.org/10.1007/s11136-015-0953-3>.
37. Juniper, E. F., Thompson, A. K., Ferrie, P. J., & Roberts, J. N. (1999). Validation of the standardized version of the Rhinoconjunctivitis quality of life questionnaire. *The Journal of Allergy and Clinical Immunology*, 104(2 Pt 1), 364–369. [https://doi.org/10.1016/s0091-6749\(99\)70380-5](https://doi.org/10.1016/s0091-6749(99)70380-5).
38. Juniper, E. F., Howland, W. C., Roberts, N. B., Thompson, A. K., & King, D. R. (1998). Measuring quality of life in children with rhinoconjunctivitis. *The Journal of Allergy and Clinical Immunology*, 101(2 Pt 1), 163–170. [https://doi.org/10.1016/s0091-6749\(98\)70380-x](https://doi.org/10.1016/s0091-6749(98)70380-x).
39. Matsumoto, H., Williams, B., Park, H. Y., et al. (2018). The final 24-item early onset scoliosis questionnaires (EOSQ-24): Validity, Reliability and Responsiveness. *Journal of Pediatric Orthopedics*, 38(3), 144–151. <https://doi.org/10.1097/BPO.0000000000000799>.
40. Corona, J., Matsumoto, H., Roye, D. P., & Vitale, M. G. (2011). Measuring quality of life in children with early onset scoliosis: Development and initial validation of the early onset scoliosis questionnaire. *Journal of Pediatric Orthopedics*, 31(2), 180–185. <https://doi.org/10.1097/BPO.0b013e3182093f9f>.
41. Mladenov, K., Braunschweig, L., Behrend, J., Lorenz, H. M., von Deimling, U., & Hell, A. K. (2019). Validation of the German version of the 24-item early-onset scoliosis questionnaire. *Journal of Neurosurgery. Pediatrics*, 23(6), 688–693. <https://doi.org/10.3171/2019.1.PEDS18704>.
42. Davis, A. M., Wright, J. G., Williams, J. I., Bombardier, C., Griffin, A., & Bell, R. S. (1996). Development of a measure of physical function for patients with bone and soft tissue sarcoma. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 5(5), 508–516. <https://doi.org/10.1007/bf00540024>.
43. Carr, T., Harris, D., & James, C. (2000). The Derriford appearance scale (DAS-59): A new scale to measure individual responses to living with problems of appearance. *British Journal of Health Psychology*, 5(2), 201–215. <https://doi.org/10.1348/135910700168865>.
44. Harris, D. L., & Carr, A. T. (2001). The Derriford appearance scale (DAS59): A new psychometric scale for the evaluation of patients with disfigurements and aesthetic problems of appearance. *British Journal of Plastic Surgery*, 54(3), 216–222. <https://doi.org/10.1054/bjps.2001.3559>.
45. Leske, D. A., Hatt, S. R., & Holmes, J. M. (2010). Test-retest reliability of health-related quality-of-life questionnaires in adults with strabismus. *American Journal of Ophthalmology*, 149(4), 672–676. <https://doi.org/10.1016/j.ajo.2009.11.004>.
46. Mangione, C. M., Lee, P. P., Gutierrez, P. R., et al. (2001). Development of the 25-item National Eye Institute Visual Function Questionnaire. *Archives of ophthalmology*, 119(7), 1050–1058. <https://doi.org/10.1001/archophth.119.7.1050>.
47. Angeles-Han, S. T., & Rabinovich, C. E. (2016). Uveitis in children. *Current Opinion in Rheumatology*, 28(5), 544–549. <https://doi.org/10.1097/BOR.0000000000000316>.
48. Angeles-Han, S. T., Yeh, S., McCracken, C., et al. (2015). Using the effects of youngsters' eyesight on quality of life questionnaire to measure visual outcomes in children with uveitis: EYE-Q measures vision in children with uveitis. *Arthritis Care and Research*, 67(11), 1513–1520. <https://doi.org/10.1002/acr.22627>.
49. Robertson, A., Tadic, V., Hundt, G., & Rahi, J. S. (2016). Patient-reported outcome measures (PROMs) of vision-related quality of life (VQoL) and functional vision (FV) for children and young people (CYP) aged 8-18 years. *Investigative Ophthalmology & Visual Science*, 57(12), 5605.
50. Tadic, V., Robertson, A., Lewando-Hundt, G., & Rahi, J. S. (2016). Measuring vision-related quality of life (VQoL) of children and young people with visual impairment. *Quality of Life Research*, 25(1 Supplement 1), 90. <https://doi.org/10.1007/s11136-016-1390-7>.
51. Cochrane, G. M., Marella, M., Keeffe, J. E., & Lamoureux, E. L. (2011). The impact of vision impairment for children (IV_C): Validation of a vision-specific pediatric quality-of-life questionnaire using Rasch analysis. *Investigative Ophthalmology & Visual Science*, 52(3), 1632. <https://doi.org/10.1167/iovs.10-6079>.
52. Cochrane, G., Lamoureux, E., & Keeffe, J. (2008). Defining the content for a new quality of life questionnaire for students with low vision (the impact of vision impairment on children: IV_C). *Ophthalmic Epidemiology*, 15(2), 114–120. <https://doi.org/10.1080/09286580701772029>.
53. Tadic, V., Cooper, A., Cumberland, P., Lewando-Hundt, G., & Rahi, J. S. (2013). Development of the functional vision questionnaire for children and young people with visual impairment. *Ophthalmology*, 120(12), 2725–2732. <https://doi.org/10.1016/j.ophtha.2013.07.055>.
54. Khadka, J., Ryan, B., Margrain, T. H., Court, H., & Woodhouse, J. M. (2010). Development of the 25-item Cardiff visual ability questionnaire for children (CVAQC). *The British Journal of Ophthalmology*, 94(6), 730–735. <https://doi.org/10.1136/bjo.2009.171181>.
55. Mokkink, L. B., Terwee, C. B., Gibbons, E., et al. (2010). Inter-rater agreement and reliability of the COSMIN (COnsensus-based standards for the selection of health status measurement instruments) checklist. *BMC Medical Research Methodology*, 10(1), 82. <https://doi.org/10.1186/1471-2288-10-82>.
56. McNeill, N. A., Kors, W. A., Bosscha, M. I., et al. (2017). Feasibility of RetinoQuest: E-health application to facilitate and improve additional care for retinoblastoma survivors. *J Cancer Survivorship Research Practice*, 11(6), 683–690. <https://doi.org/10.1007/s11764-017-0642-z>.
57. Rothrock, N. E., Kaiser, K. A., & Cella, D. (2011). Developing a valid patient-reported outcome measure. *Clinical Pharmacology and Therapeutics*, 90(5), 737–742. <https://doi.org/10.1038/clpt.2011.195>.
58. for the Child Vision PROMs group, Tadić, V., & Rahi, J. S. (2017). One size doesn't fit all: Time to revisit patient-reported outcome measures (PROMs) in paediatric ophthalmology? *Eye*, 31(4), 511–518. <https://doi.org/10.1038/eye.2016.316>.
59. Anker, S. D., Agewall, S., Borggrefe, M., et al. (2014). The importance of patient-reported outcomes: A call for their comprehensive integration in cardiovascular clinical trials. *European Heart Journal*, 35(30), 2001–2009. <https://doi.org/10.1093/eurheartj/ehu205>.
60. Brod, M., Tesler, L. E., & Christensen, T. L. (2009). Qualitative research and content validity: Developing best practices based on science and experience. *Quality of Life Research International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 18(9), 1263–1278. <https://doi.org/10.1007/s11136-009-9540-9>.
61. Wiering, B., de Boer, D., & Delnoij, D. (2017). Patient involvement in the development of patient-reported outcome measures: The developers' perspective. *BMC Health Services Research*, 17(1), 635. <https://doi.org/10.1186/s12913-017-2582-8>.
62. Stanisiewska, S., Adebajo, A., Barber, R., et al. (2011). Developing the evidence base of patient and public involvement in health and social care research: The case for measuring impact: Patient and public involvement in health and social care research. *International Journal of Consumer Studies*, 35(6), 628–632. <https://doi.org/10.1111/j.1470-6431.2011.01020.x>.
63. Canadian Institutes of Health Research. Canada's Strategy for Patient-Oriented Research. Accessed September 16, 2019. <http://www.cihr-irsc.ca/e/45851.html>
64. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Administration. US FDA. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims>. Accessed 5 Feb 2020.
65. Mokkink, L. B., Terwee, C. B., Patrick, D. L., et al. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes.

- Journal of Clinical Epidemiology*, 63(7), 737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>.
66. Bialocerkowski, A. E., & Bragge, P. (2008). Measurement error and reliability testing: Application to rehabilitation. *International Journal of Therapy and Rehabilitation*, 15(10), 422–427. <https://doi.org/10.12968/ijtr.2008.15.10.31210>.
 67. Hays, R. D., & Hadorn, D. (1992). Responsiveness to change: An aspect of validity, not a separate dimension. *Quality of Life Research*, 1(1), 73–75. <https://doi.org/10.1007/BF00435438>.
 68. McKenna, S. P. (2011). Measuring patient-reported outcomes: Moving beyond misplaced common sense to hard science. *BMC Medicine*, 9(1), 86. <https://doi.org/10.1186/1741-7015-9-86>.
 69. Nijsten, T., Meads, D., & McKenna, S. (2006). Dimensionality of the dermatology life quality index (DLQI): A commentary. *Acta Dermato-Venereologica*, 86(3), 284–285. <https://doi.org/10.2340/00015555-0075>.
 70. Bjorner, J. B., & Pejtersen, J. H. (2010). Evaluating construct validity of the second version of the Copenhagen Psychosocial Questionnaire through analysis of differential item functioning and differential item effect. *Scandinavian Journal of Public Health*, 38(3_suppl), 90–105. <https://doi.org/10.1177/1403494809352533>.
 71. Braithwaite, T., Calvert, M., Gray, A., Pesudovs, K., & Denniston, A. (2019). The use of patient-reported outcome research in modern ophthalmology: Impact on clinical trials and routine clinical practice. *Patient Relat Outcome Meas*, 10, 9–24. <https://doi.org/10.2147/PROM.S162802>.
 72. Khadka, J., Fenwick, E., Lamoureux, E., & Pesudovs, K. (2016). Methods to Develop the Eye-tem Bank to Measure Ophthalmic Quality of Life. *Optometry and Vision Science*, 93(12), 1485–1494. <https://doi.org/10.1097/OPX.0000000000000992>.
 73. Haywood, K., Lyddiatt, A., Brace-McDonnell, S. J., Staniszewska, S., & Salek, S. (2017). Establishing the values for patient engagement (PE) in health-related quality of life (HRQoL) research: An international, multiple-stakeholder perspective. *Quality of Life Research*, 26(6), 1393–1404. <https://doi.org/10.1007/s11136-016-1465-5>.
 74. Anthony, S. J., Selkirk, E., Sung, L., et al. (2014). Considering quality of life for children with cancer: A systematic review of patient-reported outcome measures and the development of a conceptual model. *Quality of life research*, 23(3), 771–789. <https://doi.org/10.1007/s11136-013-0482-x>.
 75. Davis, E., Waters, E., Mackinnon, A., et al. (2006). Paediatric quality of life instruments: A review of the impact of the conceptual framework on outcomes. *Developmental Medicine and Child Neurology*, 48(04), 311. <https://doi.org/10.1017/S0012162206000673>.
 76. Colenbrander, A. (2010). Assessment of functional vision and its rehabilitation. *Acta Ophthalmologica*, 88(2), 163–173. <https://doi.org/10.1111/j.1755-3768.2009.01670.x>.
 77. Fayed, N., De Camargo, O. K., Kerr, E., et al. (2012). Generic patient-reported outcomes in child health research: A review of conceptual content using World Health Organization definitions: Review. *Developmental Medicine and Child Neurology*, 54(12), 1085–1095. <https://doi.org/10.1111/j.1469-8749.2012.04393.x>.
 78. Ashing-Giwa, K. T. (2005). The contextual model of HRQoL: A paradigm for expanding the HRQoL framework. *Quality of Life Research*, 14(2), 297–307. <https://doi.org/10.1007/s11136-004-0729-7>.
 79. Sabri, K., Knapp, C. M., Thompson, J. R., & Gottlob, I. (2006). The VF-14 and psychological impact of amblyopia and strabismus. *Investigative Ophthalmology & Visual Science*, 47(10), 4386. <https://doi.org/10.1167/iov.05-1365>.
 80. Pesudovs, K., Burr, J. M., Harley, C., & Elliott, D. B. (2007). The development, Assessment, and Selection of Questionnaires. *Optometry and Vision Science*, 84(8), 663–674. <https://doi.org/10.1097/OPX.0b013e318141fe75>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
