Interrater reliability and agreement between children with visual impairment and their parents on participation and quality of life

Ellen BM Elsman, [iD] Mitchel LC Koel, Ruth MA vanNispen and Ger HMB vanRens

Department of Ophthalmology, Amsterdam UMC, Vrije Universiteit Amsterdam, the Amsterdam Public Health Research Institute, MB Amsterdam, The Netherlands

ABSTRACT.

Purpose: To investigate interrater reliability and agreement between children with visual impairment (VI) and their parents on participation and quality of life and factors associated with disagreement.

Methods: Children 7–17 years and their parents completed the PAI-CY 7–12 ($n = 180$) and 13–17 ($n = 65$), the KIDSCREEN-27 ($n = 250$) and the CASP ($n = 70$). Mean scores of children and parents were compared, with effect sizes for the differences. Interrater reliability was evaluated using intraclass correlation coefficients (ICCs), whereas agreement was assessed using the Bland–Altman limits of agreement. Linear regression analyses examined child- and proxy-related factors associated with discrepancies.

Results: On average, children rated their participation and quality of life as significantly better than their parents on most (sub)scales, but with wide range of disagreement. Effect sizes were large for the PAI-CY 7–12 (0.86) and 13–17 (0.86) and small for the CASP (0.36) and KIDSCREEN-27 (0.18–0.28). Interrater reliability was poor for the PAI-CY 7–12 (ICC = 0.29) and most KIDSCREEN-27 subscales (ICC =0.18–0.32), moderate for the PAI-CY 13–17 (ICC =0.43) and the KIDSCREEN-27 Physical Wellbeing subscale (ICC = 0.46) and good for the CASP (ICC = 0.63). Comorbidity was significantly associated with greater discrepancies on participation scales.

Conclusion: Children with VI and their parents have different perspectives on the child's participation and quality of life. Disagreement was largest on participation scales and smallest on quality of life subscales, while opposite results were found for interrater reliability. Reports of children and parents seem to be complementary and are both relevant to obtain a complete picture of the burden of VI and relevant to inform healthcare decisions.

Introduction

Patient-reported outcomes, such as participation and health-related quality of life, have become increasingly important in health care (Matza et al. 2004; Black et al. 2016; Greenhalgh et al. 2018; Calvert et al. 2019). The use of patient-reported outcomes can facilitate patient-centred care, shared-decision making and the evaluation of intervention effectiveness (Varni et al. 2005; Black et al. 2016; Greenhalgh et al. 2018; Calvert et al. 2019). Obtaining patient-reported outcomes is particularly important for chronic health conditions, for which cure is often not possible (Kaplan 2001; Ingerski et al. 2010). Childhood visual impairment (VI) is an example of such a condition, and participation and quality of life are often considered important outcomes of rehabilitation for children with VI.

To measure these outcomes in children with VI, both generic paediatric patient-reported outcome measures (PROMs) (Varni et al. 2001; Ravens-Sieberer et al. 2007a; Bedell 2009) and vision-specific paediatric PROMs can be used. Several vision-specific paediatric PROMs have been developed in recent years (Cochrane et al. 2011; Tadić et al. 2013; Tadic et al. 2016; Elsman et al. 2020a; Elsman et al. 2020b), and these instruments might be more sensitive to the specific problems children with VI encounter compared to their generic counterparts.

Administering PROMs to children is subject to various challenges, including differences in children's cognitive and linguistic abilities to respond to questions at different ages and parents' influence on experiences of children (Le Coq et al. 2000; Upton et al. 2008; Gothwal et al. 2018). Hence, there is substantial debate who the most appropriate respondent is to obtain paediatric patient-reported outcomes (Theunissen et al. 1998; Matza et al. 2004), although in general, self-report is considered most accurate (Upton et al. 2008), and can be obtained from children as young as five years (Varni et al. 2007). Nevertheless, several paediatric PROMs contain self-report alongside proxy-report versions (Ravens-Sieberer et al. 2007a; McDougall et al. 2013; Elsman et al. 2020a; Elsman et al. 2020b).

A large body of literature shows that there might be discrepancies between child and parent reports of health outcomes, although results are mixed (Theunissen et al. 1998; Verrips et al. 2000; Eiser & Morse 2001; Waters et al. 2003; Matza et al. 2004; Cremeens et al. 2006; Upton et al. 2008). The level of agreement may be dependent on several aspects, including the observability of the domain being measured, gender, age, health status and education, but again with inconclusive results (Theunissen et al. 1998; Verrips et al. 2000; Eiser & Morse 2001; Waters et al. 2003; Cremeens et al. 2006; Upton et al. 2008).

When reports of children and parents deviate, it raises the question how the results should be interpreted, and whose reports are more accurate (Eiser & Morse 2001; Matza et al. 2004). This is particularly important whether healthcare decisions are made based upon these reports (De Los Reyes et al. 2011; Eiser & Morse 2001; Janicke et al. 2001). Parent reports may be expected to be more reliable, but the subjective nature of outcomes such as participation and quality of life suggests the child reports to be more valid (Eiser & Morse 2001; Matza et al. 2004).

Agreement in health outcomes between children with VI and their parents has been investigated in several studies (Chak & Rahi 2007; Van Dijk et al. 2007; Hamblion et al. 2011; Tadić et al. 2017; Gothwal et al. 2018). Most of these studies focused on specific conditions, such as retinoblastoma

(Van Dijk et al. 2007), congenital glaucoma (Gothwal et al. 2018), congenital cataract (Chak & Rahi 2007) and hereditary retinal disorders (Hamblion et al. 2011), and all of them have only evaluated agreement on health-related quality of life, either with generic (Varni et al. 2001; Ravens-Sieberer et al. 2007a) or with vision-specific PROMs (Tadić et al. 2013; Tadic et al. 2016). Only one study also reported interrater reliability (Tadić et al. 2017), which is often used interchangeably with agreement, but conceptually and technically different (de Vet et al. 2006; Kottner et al. 2011; Gisev et al. 2013). None of the studies have focused on participation, while reported limitations in activities and participation are often used to inform assignment to low vision, pedagogical or behavioural interventions. Moreover, not much is known about factors associated with the level of disagreement between children with VI and their parents (Tadić et al. 2017; Gothwal et al. 2018).

Therefore, this study investigates interrater reliability and agreement between children with VI and their parents and the factors associated with disagreement. Because both generic and vision-specific instruments were included, differences in child–parent agreement and interrater reliability between these instruments were investigated.

Methods

The study protocol was approved by the Medical Ethical Committee of Amsterdam UMC, the Netherlands. The study adhered to the tenets as laid down in the Declaration of Helsinki. Parents of children, as well as children aged 13 years and older, provided written informed consent. Data were collected as part of a study to validate the Participation and Activity Inventory – Children and Youth (PAI-CY) for children aged 7–12 and 13–17 years (Elsman et al. 2020a; Elsman et al. 2020b).

Participants

Children aged 7–17 years who were registered at two Dutch nationwide low vision services, and their parents, were invited to participate. The following inclusion criteria applied: (1) sufficient

knowledge and understanding of the Dutch language; (2) children with VI from any cause, without restrictions regarding visual performance; (3) children without profound cognitive impairment, as registered in their patient file; and (4) children with mild cognitive impairment as reported by parents could participate.

Procedure

Participating children completed questionnaires via face-to-face interviews at their homes, while their parents completed questionnaires via a web-based survey (or a paper-and-pencil version on request). Children aged 7–12 years and their parents completed the PAI-CY 7–12 (Elsman et al. 2020a) and the KIDSCREEN-27 (Ravens-Sieberer et al. 2007a). Children aged 13–17 years and their parents completed the PAI-CY 13–17 (Elsman et al. 2020b), the KIDSCREEN-27 (Ravens-Sieberer et al. 2007a) and the Child and Adolescent Scale of Participation (CASP) (Bedell 2009). Parents additionally completed a questionnaire regarding sociodemographic and clinical characteristics of their child and its family. Ophthalmic diagnoses, decimal visual acuity and visual field were retrieved from the patient files at the low vision services. Missing data were complemented by parents' self-reported data. Decimal visual acuity of the better eye was transformed into logMAR and classified according to criteria of the World Health Organization (WHO) (WHO 2010): no VI (logMAR ≤ 0.3), mild VI (logMAR 0.31-0.52), moderate VI (logMAR 0.53-1), severe VI (logMAR 1.01-1.30) and blind (logMAR ≥ 1.31).

Instruments

The PAI-CY is a series of questionnaires to measure limitations in participation and activities of children and youth with VI. The PAI-CY was developed in the Netherlands and as such is available in Dutch. The PAI-CY 7–12 and 13–17 consist of a self-report version alongside a proxy-report version, which both contain, respectively, 47 and 55 items (Elsman et al. 2020a; Elsman et al. 2020b; Elsman et al. 2020c). Items are scored on a 4-point Likert scale with response options 'not difficult', 'slightly difficult', 'very difficult' and 'impossible'. The

response option 'not applicable' is treated as a missing value. Sum scores were calculated based on the mean of scores (at least 75% of the items had to be completed) which were then multiplied by the total number of items. Sum scores were subsequently rescored to 0–100. Higher scores represent fewer limitations in participation and activities, but there is no cut-off that signifies suboptimal scores.

The CASP measures the degree of participation of a child in home, school and community activities and asks respondents to compare their participation to the participation of children of the same age. Youth-report and proxy-report versions of the Dutch CASP were used for children aged 13–17 years (Bedell 2009, 2011; McDougall et al. 2013). The CASP contains 20 items, which are scored on a 4-point Likert scale with response options 'age expected', 'somewhat limited', 'very limited' and 'unable'. The response option 'not applicable' is treated as a missing value. Previous analyses indicated the CASP comprises a unidimensional scale (Elsman et al.), and as such, sum scores for the total scale were calculated ranging from 0 to 100, using the same method as to calculate scores for the PAI-CY. Higher scores represent greater age-expected participation, but there is no cut-off that signifies suboptimal scores.

The KIDSCREEN-27 measures health-related quality of life in children and adolescents in five relevant domains: Physical Wellbeing (five items), Psychological Wellbeing (seven items), Autonomy & Parent Relation (seven items), Social Support & Peers (four items) and School Environment (four items) (KIDSCREEN 2006; Ravens-Sieberer et al. 2007b). The Dutch self-report and proxy-report versions were used for children aged 7–17 years. All items are scored on a 5-point Likert scale with response options referring to frequency or degree of feeling. Negatively formulated items were recoded, and a scoring algorithm was used to calculate T-scores on each subscale, with a mean of 50 and a standard deviation of 10 in the international survey population. In practice, T-scores range from 20 to 80 (i.e. 99.7% of the population will score in this range). Scores that deviate more than half a standard deviation from the mean score of 50 are classified as

noticeably different. Higher T-scores represent better quality of life.

Statistical analyses

Sociodemographic and clinical characteristics of participants were analysed using descriptive statistics. Mean scores of children and parents on all (sub) scales were compared using paired-samples t-tests. Cohen's effect sizes of the differences were calculated and interpreted: ≤0.49 small, 0.50–0.79 moderate and ≥0.80 large (Cohen 1988). Interrater reliability on (sub) scales was evaluated using intraclass correlation coefficients (ICCs), in a two-way random effects model, single measure, absolute agreement. An ICC ≤0.40 was considered poor, 0.41–0.60 moderate, 0.61–0.80 good and 0.81–1.00 excellent (Varni et al. 2007). Agreement between child and parent scores on (sub)scales was assessed using the Bland–Altman's limits of agreement. Therefore, the difference between child and parent scores (displayed on the y-axis) is plotted against the mean of the scores (displayed on the x-axis). The limits of agreement were calculated with the following formula: mean difference ± 1.96 x standard deviation of the difference.

Linear regression models were computed to assess whether discrepancies between child and parent scores were associated with sociodemographic and clinical characteristics of the child and its family. The absolute difference between children and parents was used as a dependent variable, which was square root transformed because of non-normality. The following variables were included as independent variables: score of the child, gender, presence of comorbidity, level of VI and educational level of the parent. In regression models of the KIDSCREEN-27 subscales, age of the child was included additionally. Variables at the $p < 0.05$ level were considered significantly associated with discrepancies. Relevant regression assumptions were examined, and regression coefficients and confidence intervals were back-transformed to aid interpretation.

Results

In total, 291 participants provided written informed consent. Due to incomplete questionnaires and missing

items, scores on (sub)scales were calculated for the following numbers of child–parent dyads: PAI-CY 7–12 $n = 180$, PAI-CY 13–17 $n = 65$, CASP $n = 70$ and KIDSCREEN-27 $n = 250$. Table 1 presents their sociodemographic and clinical characteristics.

Table 2 presents scores of children and parents on each of the (sub)scales and their mean difference and inter-rater reliability. On average, children rated their participation and quality of life as significantly better than their parents, except on the KIDSCREEN-27 Social Support & Peers subscale, where they scored significantly worse. Effect sizes for the differences were large for the PAI-CY 7–12 and 13–17 and small for (subscales of) the CASP and KIDSCREEN-27. Interrater reliability was poor for the PAI-CY 7–12 and most KIDSCREEN-27 subscales, but moderate for the PAI-CY 13–17 and the KIDSCREEN-27 Physical Wellbeing subscale. Good interrater reliability was found for the CASP.

The Bland–Altman plots show that the level of agreement was wide and bidirectional for all subscales, indicating that parents both over- and underestimate their child's participation and quality of life (Fig. 1, Table 2). Visual interpretation, to examine whether the difference in scores between children and their parents tend to get smaller or larger as the mean score increases, of the Bland–Altman plots suggests the following:

1 When the average participation score on the CASP and PAI-CY 7–12 and 13–17 is higher (x-axis), parents seem to rate their children as having relatively higher participation, as indicated by the cluster below the mean difference line on the right side of the graphs (Fig. 1A-C, solid-line circled clusters);

2 When the average participation score on the PAI-CY 13–17 is lower (x-axis), parents seem to rate their children as having relatively lower participation, as indicated by the cluster above the mean difference line on the left side of the graph (Fig. 1B, dashed-line circled cluster);

3 When the average quality of life score on the KIDSCREEN-27 Physical Wellbeing and Psychological Wellbeing is lower (x-axis), parents seem to rate their children as having relatively lower quality of life, as indicated by the

Table 1. Sociodemographic and clinical characteristics participants

Participant characteristic	PAI-CY 7–12 (n = 180)	PAI-CY 13–17 (n = 65)	CASP (n = 70)	KIDSCREEN-27 (n = 250)
Age in years, mean ± SD (range)	9.5 ± 1.6 (7–12)	14.7 ± 1.5 (13–17)	14.7 ± 1.5 (13–17)	11.0 ± 2.8 (7–17)
Male gender, n (%)	103 (57)	38 (59)	44 (63)	148 (59)
Category of VI*, n (%)				
Blind: LogMAR ≥1.31	12 (7)	8 (12)	10 (14)	22 (9)
Severe VI: logMAR 1.01–1.30	1 (1)	3 (5)	3 (4)	4 (2)
Moderate VI: LogMAR 0.53–1	49 (27)	20 (31)	19 (27)	69 (28)
Mild VI: logMAR 0.31–0.52	35 (19)	12 (19)	12 (17)	47 (19)
No VI: logMAR ≤0.30	81 (45)	20 (31)	23 (33)	103 (41)
Unknown	2 (1)	2 (3)	3 (4)	5 (2)
Comorbidity, n (%)	86 (48)	24 (37)	29 (41)	118 (47)
Parent who completed questionnaire, n (%)				
Mother	138 (77)	42 (65)	47 (67)	185 (74)
Father	19 (11)	12 (19)	12 (17)	31 (12)
Together	19 (11)	8 (12)	8 (11)	27 (11)
Caregiver	4 (2)	3 (5)	3 (4)	7 (3)
Dutch nationality parent, n (%)	168 (93)	62 (95)	67 (96)	235 (94)
Financial situation parent, n (%)				
Usually enough money	80 (44)	33 (51)	37 (53)	118 (47)
Just enough money	46 (26)	15 (23)	15 (21)	60 (24)
Not enough money	10 (6)	7 (11)	6 (9)	19 (8)
I'd rather not say	44 (24)	10 (15)	12 (17)	53 (21)
Educational level parent, n (%)				
Low	19 (11)	12 (19)	13 (19)	32 (13)
Middle	72 (40)	32 (49)	30 (43)	107 (43)
High	86 (48)	20 (31)	26 (37)	108 (43)
Unknown	3 (2)	1 (2)	1 (1)	3 (1)

* VI categories were based on acuity loss in the better-seeing eye following the World Health Organization (WHO 2010).

Table 2. Scores of children and parents, their mean difference and interrater reliability coefficients

(sub)scale	Child score, mean ± SD	Parent score, mean ± SD	Mean paired difference (95% CI)	P-value paired difference	Effect size	ICC (95% CI)	Lower and upper limit of agreement
PAI-CY 7–12	87.5 ± 8.4	78.7 ± 11.8	8.8 (7.1 to 10.4)	<0.001	0.86	0.29 (0.03 to 0.50)	−13.3; 30.8
PAI-CY 13–17	89.5 ± 7.3	80.3 ± 13.2	9.2 (6.8 to 11.6)	<0.001	0.86	0.43 (−0.01 to 0.69)	−9.8; 28.2
CASP	89.4 ± 10.5	85.0 ± 15.5	4.4 (1.9 to 6.9)	<0.001	0.33	0.65 (0.46 to 0.78)	−16.2; 25.0
KIDSCREEN-27 Physical Wellbeing	51.2 ± 9.7	49.3 ± 11.7	1.9 (0.5 to 3.3)	0.007	0.18	0.46 (0.35 to 0.55)	−19.9; 23.8
KIDSCREEN-27 Psychological Wellbeing	53.0 ± 9.4	50.0 ± 11.6	3.0 (1.5 to 4.5)	<0.001	0.28	0.32 (0.21 to 0.43)	−20.8; 26.8
KIDSCREEN-27 Autonomy & Parent Relation	54.9 ± 10.5	52.2 ± 9.7	2.8 (1.2 to 4.3)	<0.001	0.27	0.20 (0.08 to 0.31)	−22.2; 27.7
KIDSCREEN-27 Social Support & Peers	50.1 ± 11.2	52.0 ± 10.3	−1.9 (−3.6 to −0.2)	0.030	−0.18	0.18 (0.06 to 0.30)	−28.9; 25.1
KIDSCREEN-27 School Environment	56.8 ± 10.3	54.7 ± 10.7	2.1 (0.4 to 3.8)	0.014	0.20	0.18 (0.06 to 0.30)	−24.2; 28.4

clusters above the mean difference line on the left side of the graphs (Fig. 1D and E, solid-line circled clusters);

4 When the average quality of life score on the KIDSCREEN-27 Autonomy & Parent Relation is lower (x-axis), parents seem to rate their children as having relatively higher quality of life, as indicated by the clusters below the mean difference line on the left side of the graph (Fig. 1F, solid-line circled cluster);

5 No clear trends for smaller or larger differences in scores between children and their parents with increasing mean scores could be identified from Fig. 1G and H.

Linear regression analyses (Table 3) showed that comorbidity of the child was significantly associated with greater child–parent discrepancies across all instruments measuring participation (i.e. PAI-CY 7–12, PAI-CY 13–17 and CASP) and smaller child–

parent discrepancies on the KIDSCREEN-27 School Environment subscale. Moreover, moderate VI was significantly associated with greater child–parent discrepancies on the CASP and KIDSCREEN-27 Social Support & Peers, compared to no VI. For the CASP, severe VI was also significantly associated with greater discrepancies, compared to no VI. Higher (i.e. better) child scores were significantly associated with greater
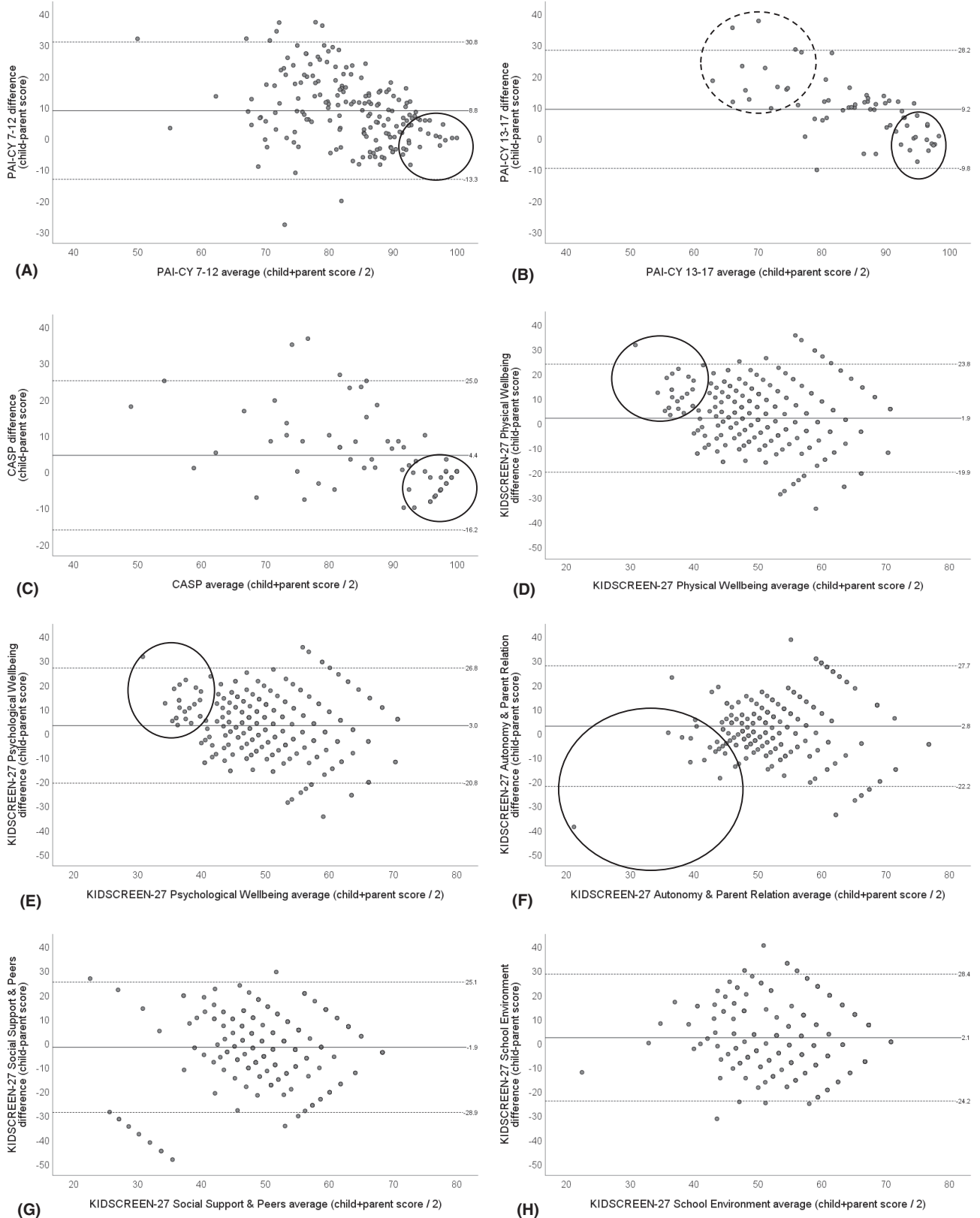
Figure 1. Bland–Altman plots for child–parent agreement on the PAI-CY 7–12 (A), PAI-CY 13–17 (B), CASP (C), and subscales of the KIDSCREEN-27 (D–H).

discrepancies on the KIDSCREEN-27 Psychological Wellbeing and Autonomy & Parent Relation subscales and smaller discrepancies on the KIDSCREEN-27 Social Support & Peers subscale. Last, as compared to high parental education, low education was significantly associated with greater discrepancies on the KIDSCREEN-27

Table 3. Linear regression results examining factors associated with the absolute score discrepancy across (sub)scales

	PAI-CY 7–12	PAI-CY 13–17	CASP	KIDSCREEN-27 Physical Wellbeing	KIDSCREEN-27 Psychological Wellbeing	KIDSCREEN-27 Autonomy & Parent Relation	KIDSCREEN-27 Social Support & Peers	KIDSCREEN-27 School Environment
N	175	62	66	242	242	242	242	242
R^2	0.1	0.25	0.37	0.02	0.07	0.22	0.2	0.07
Characteristic	B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)
Score child	0.16 (−0.04; 0.22)	−0.19 (−0.29; 0.10)	−0.13 (−0.22; 0.13)	0.11 (−0.07; 0.17)	**0.14 (0.07; 0.19)**	**0.24 (0.20; 0.27)**	**−0.22 (−0.25; −0.18)**	0.12 (−0.04; 0.17)
Age 13–17 child*				0.31 (−0.53; 0.69)	−0.57 (−0.82; 0.13)	0.14 (−0.57; 0.60)	−0.46 (−0.75; 0.37)	−0.27 (−0.66; 0.54)
Female sex child	−0.60 (−0.88; 0.23)	−0.28 (−0.85; 0.75)	−0.63 (−1.03; 0.51)	−0.30 (−0.66; 0.50)	−0.26 (−0.61; 0.49)	−0.49 (−0.74; 0.25)	−0.38 (−0.68; 0.41)	−0.33 (−0.66; 0.46)
Comorbidity child	**0.86 (0.55; 1.09)**	**0.99 (0.55; 1.28)**	**1.02 (0.59; 1.32)**	0.26 (−0.53; 0.64)	0.29 (−0.48; 0.63)	0.06 (−0.55; 0.56)	−0.44 (−0.71; 0.36)	**−0.62 (−0.84; −0.23)**
VI child[†]								
Mild	0.37 (−0.65; 0.84)	0.7 (−0.63; 1.18)	0.73 (−0.63; 1.21)	−0.27 (−0.73; 0.63)	−0.35 (−0.74; 0.55)	−0.5 (−0.82; 0.42)	0.38 (−0.53; 0.76)	−0.47 (−0.81; 0.47)
Moderate	0.12 (−0.71; 0.73)	−0.2 (−0.91; 0.86)	**0.99 (0.39; 1.34)**	0.46 (−0.45; 0.79)	−0.44 (−0.76; 0.43)	0.37 (−0.49; 0.71)	**0.69 (0.30; 0.92)**	−0.61 (−0.87; 0.15)
Severe	0.68 (−0.62; 1.14)	0.52 (−0.83; 1.11)	1.36 (0.94; 1.67)	0.33 (−0.69; 0.83)	−0.52 (−0.89; 0.51)	−0.5 (−0.88; 0.53)	0.63 (−0.39; 0.97)	−0.44 (−0.86; 0.61)
Education parent[‡]								
Low	−0.23 (−0.86; 0.80)	−0.51 (−1.11; 0.84)	−0.45 (−1.09; 0.88)	0.58 (−0.45; 0.93)	−0.51 (−0.87; 0.48)	0.2 (−0.67; 0.73)	−0.23 (−0.74; 0.67)	**0.82 (0.40; 1.09)**
Middle	0.48 (−0.46; 0.82)	−0.16 (−0.86; 0.83)	0.62 (−0.58; 1.05)	0.2 (−0.57; 0.63)	0.5 (−0.28; 0.76)	**0.59 (0.13; 0.82)**	0.46 (−0.34; 0.74)	0.42 (−0.42; 0.72)

* Reference group is 7–12, only included in analyses KIDSCREEN-27. [†]Reference group is no VI (WHO 2010). [‡]Reference group is high education. Bold is significant at p < 0.05.

School Environment subscale, whereas middle education was significantly associated with greater discrepancies on the KIDSCREEN-27 Autonomy & Parent Relation subscale. The explained variance was low on the KIDSCREEN-27 Physical Wellbeing ($R^2 = 0.02$), Psychological Wellbeing ($R^2 = 0.07$) and School Environment ($R^2 = 0.07$) subscales, but higher on the CASP ($R^2 = 0.37$).

Discussion

In this study, we investigated both agreement between children with VI and their parents and interrater reliability for generic and vision-specific instruments of participation and quality of life. Although often used interchangeably, there is a conceptual and technical distinction. Reliability relates to the extent of variability and error inherent in a measurement, whereas agreement is defined as the degree to which scores or ratings are identical (de Vet et al. 2006; Kottner et al. 2011; Gisev et al. 2013). Both agreement and interrater reliability are important in the evaluation of the quality of the measurements (Kottner et al. 2011).

On average, parents perceived their child's participation and quality of life as worse than children themselves on all but one (sub)scale. This finding is consistent with general results from literature, showing that in children with chronic health conditions, parents rate the well-being of their child to be worse than children themselves, which is opposite to findings in healthy children (Upton et al. 2008). The underestimation of parents is also in line with most other studies in ophthalmic populations (Van Dijk et al. 2007; Hamblion et al. 2011; Tadić et al. 2017), although in children with congenital glaucoma, parents reported better quality of life of their child than children themselves (Gothwal et al. 2018), and in children with congenital cataract, quality of life reports of children and parents were comparable (Chak & Rahi 2007). The Bland–Altman plots provided an informative visual representation giving information about the discrepancies across the range of scores. From these plots, it was evident that the range of agreement was wide, and parents both overestimated and underestimated their child's report. Moreover, effect sizes for the differences were mostly small, indicating limited clinical relevance. Large effect sizes were found for the PAI-CY 7–12 and 13–17, which might suggest that the difference is not only statistically significant, but also clinically relevant. Besides these distribution-based indicators for differences, it is important to have anchor-based indicators for differences, such as minimally important differences in child–parent discrepancies (Eiser & Varni 2013), which are currently lacking. If these indicators are available, they can indicate whether discrepancies are sufficiently large to warrant further investigation prior to healthcare decisions. Nonetheless, it is notable that agreement between children and parents was smaller for scales measuring participation (as indicated by the larger effect sizes), as in general, parents are more in agreement on objective

physical domains, but less on emotional and social domains (Eiser & Varni 2013). The PAI-CY and CASP are instruments to assess limitations in activities and participation and as such contain more objective, physical items, than, for example, the various subscales of the KIDSCREEN-27.

Our analyses of factors associated with discrepancies between child and parent reports showed that comorbidity is a factor consistently associated with larger discrepancies on participation scales. Comorbidity was assessed through an open question (i.e. 'does your child have any other conditions'). This resulted in a large variety of comorbidities reported, ranging from allergies to attention-deficit/ hyperactivity disorder to various syndromes, hindering quantification in subgroups. We also found that on the CASP, moderate and severe VI was associated with larger discrepancies, compared to no VI. Tadić and colleagues also found that several ophthalmic factors had an influence on child–parent discrepancy, although they used classifications of ICCs as an index of agreement. ICCs were higher for those with more severe VI, late-onset VI and progressive VI (Tadić et al. 2017). The other factors significantly associated with discrepancies in our study had a more random pattern, and their influence on discrepancy was bidirectional (i.e. score of the child was significantly associated with greater discrepancies on the KIDSCREEN-27 Psychological Wellbeing and Autonomy & Parent Relation subscales, but with smaller discrepancies on the KIDSCREEN-27 Social Support & Peers subscale). Also, some of the groups included as independent variable were relatively small, such as the severe VI group or the low education of the parent group. Studies show that the factors associated with child–parent discrepancies vary considerably by health condition and (sub)scale, making it difficult to interpret these results in the context of existing literature (Upton et al. 2008; Eiser & Varni 2013). Child age and gender were not associated with discrepancies on any of the (sub)scales, which is contrasting the findings of Gothwal and colleagues, who found greater discrepancies for younger children and girls (Gothwal et al. 2018).

Reliability between children and parents was poor for the PAI-CY 7–

12 and most KIDSCREEN-27 subscales, whereas moderate reliability was found for the PAI-CY 13–17 and KIDSCREEN-27 Physical Wellbeing subscale, and good reliability for the CASP. This might suggest that reliability is better for those scales consisting of more visible concepts (e.g. physical well-being, participation in activities), especially when children get older, as the PAI-CY 13–17 and CASP were only administered to children from 13 years onwards and their parents.

It is important to collect patient-reported outcomes to evaluate the burden of VI and to inform healthcare decisions. The PAI-CY has been developed as an intake instrument, to identify difficulties with respect to activities and participation in children and youth with visual impairment. It can be used to map the needs and difficulties of children, youth and their parents prior to a rehabilitation programme at low vision services. It might also be possible to administer the PAI-CY after rehabilitation, to determine the effectiveness, although responsiveness of the PAI-CY is yet to be investigated. Additionally, the PAI-CY can be used in broader clinical practice or as outcome instrument in research (Elsman et al. 2020c). The KIDSCREEN-27 can also be used in all types of epidemiological, paediatric and clinical studies, in health services research and health reporting, and in integrated outcome measurement (KIDSCREEN 2006). However, a recent study suggests that it might not be sensitive enough to the specific problems children with VI encounter (Elsman et al. 2021a). Therefore, it might not be recommended to use this instrument to inform healthcare decisions in children with VI. The CASP has been designed as part of the Child and Family Follow-up Survey, to monitor outcomes and needs of children with acquired brain injuries. The CASP has subsequently been used as a separate instrument and in children with other diagnoses and as such can be used for individualized intervention planning, programme evaluation and population-based research (Bedell 2011). It seems suitable for children with VI (Elsman et al. 2021a), although no self-report version for younger children exists alongside the proxy-report version, a feature of the PAI-CY that is highly valued (Elsman et al. 2017).

If PROMs are used to inform healthcare decisions, involvement of parents is vital, which makes it important to know how reports of parents deviate from that of children. The discrepancies between reports of children and their parents might be caused by underreporting or minimising problems of children with VI, who might be denying their condition, or by the inability to report their participation and quality of life. On the other hand, children might report their participation and quality of life accurately to their parents, who then exaggerate or over-report children's problems. Alternatively, parents might report participation and quality of life of children with VI in comparison to children without VI, or while considering future life demands. Conversely, children might report their participation and quality of life while focusing on their current situation, and without comparisons to others. Because neither self-report nor proxy-report is without risk, and deviations can be expected, obtaining information from both children and their parents may provide the most complete picture of the burden of VI. The deviations can additionally give rise to fruitful discussions between children and parents about the origin of discrepancies.

The results of our study should be considered in light of several limitations. First, we had limited information from the proxy responders, which prohibited the exploration of other factors associated with child–parent discrepancies. For example, the parent's own quality of life has shown to have an impact on discrepancy (Janicke et al. 2007; Kobayashi & Kamibeppu 2011; Oltean & Ferro 2019), and not much is known about father-child agreement compared to mother-child agreement (Petsios et al. 2011). The number of fathers completing the proxy-report versions was relatively small, and therefore this was not included in the linear regression analyses. The same was true for the number of respondents not having the Dutch nationality.

Second, the classification into level of VI was based on visual acuity only, as often limited information was available in the patient files regarding visual field, or it was described in qualitative terms such as 'concentrically restricted', 'severe peripheral field loss', or 'right-sided hemianopia'. When visual acuity was

not affected, these children were classified in the no VI group. Similarly, around 10% of the children had cerebral visual impairment, a condition in which visual acuity is often not affected (Dutton & Jacobson 2001), and were therefore also categorized in the no VI group. Thus, it is likely that at least part of the children classified as no VI had in fact more severe VI than resulting from their visual acuity, which might have influenced the results on factors associated with discrepancies. One should keep in mind that the children in our study were all registered at Dutch low vision services. According to the Dutch guidelines, referral to these centres is recommended when visual acuity is <0.3, visual field is <30 degrees, in case of disorders in lower or higher visual functions (e.g. respectively night blindness/photophobia or cerebral visual impairment), in case of a progressive disorder, or in case of a rehabilitation need for which no opportunities in regular ophthalmological care exist (Van Rens et al. 2011). Thus, these children are likely to comply to one of these conditions, but this gets lost in the classification based on visual acuity.

Last, children completed questionnaires face-to-face, while their parents completed questionnaires online. These different modes of administration might have contributed to the discrepancies between children and their parents. Face-to-face administration is more susceptible to yes-saying bias and socially desirable responses (Davis et al. 2007). However, administration without involvement of a researcher could have led to parents helping their children, thereby influencing children's responses. Furthermore, the VI would have made it difficult for children to independently complete the questionnaires, especially for the younger children.

Our study is the first to report agreement between children with VI and their parents and interrater reliability on quality of life and participation instruments. Our study showed that children with VI and their parents have significantly different perspectives on the child's participation and quality of life, and parents both overestimated and underestimated their child's report. With the linear regression models that were applied to examine factors associated with disagreement between children with VI and their parents, our

study showed that discrepancies were to an extent associated with sociodemographic and clinical characteristics of the child and its family. Disagreement was largest on instruments measuring participation (i.e. the PAI-CY and, to a lesser extent, the CASP), and smallest on quality of life subscales (i.e. the KIDSCREEN-27). Opposite results were found for interrater reliability. Reports of children and parents should be considered complementary, and both relevant to obtain a complete picture of the burden of VI and inform healthcare decisions. Our results might help healthcare providers to better understand and value deviations in patient-reported outcomes between children with VI and their parents. Future studies should attempt to determine minimally important differences in child–parent discrepancies, and which magnitudes should give rise to more extensive investigation. Additionally, the factors associated with discrepancies should be subject to further research.

References

Bedell G (2009): Further validation of the Child and Adolescent Scale of Participation (CASP). Dev Neurorehabil 12: 342–351.

Bedell G (2011): The Child and Adolescent Scale of Participation (CASP): administration and scoring guidelines. Medford, MA: Tufts University.

Black N, Burke L, Forrest CB, Ravens Sieberer UH, Ahmed S, Valderas JM, Bartlett SJ & Alonso J (2016): Patient-reported outcomes: pathways to better health, better services, and better societies. Qual Life Res 25: 1103–1112.

Calvert MJ, O'Connor DJ & Basch EM (2019): Harnessing the patient voice in real-world evidence: the essential role of patient-reported outcomes. Nat Rev Drug Discovery 18: 731–732.

Chak M & Rahi JS (2007): The health-related quality of life of children with congenital cataract: findings of the British Congenital Cataract Study. Br J Ophthalmol 91: 922–926.

Cochrane GM, Marella M, Keeffe JE & Lamoureux EL (2011): The Impact of Vision Impairment for Children (IVI_C): validation of a vision-specific pediatric quality-of-life questionnaire using Rasch analysis. Invest Ophthalmol Vis Sci 52: 1632–1640.

Cohen J (1988): Statistical power analysis for the behavioural sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cremeens J, Eiser C & Blades M (2006): Factors influencing agreement between child

self-report and parent proxy-reports on the Pediatric Quality of Life Inventory™ 4.0 (PedsQL™) Generic Core Scales. Health Qual Life Outcomes 4: 1–8.

Davis E, Nicolas C, Waters E, Cook K, Gibbs L, Gosch A & Ravens-Sieberer U (2007): Parent-proxy and child self-reported health-related quality of life: using qualitative methods to explain the discordance. Qual Life Res 16: 863–871.

De Los RA, Youngstrom EA, Swan AJ, Youngstrom JK, Feeny NC & Findling RL (2011): Informant discrepancies in clinical reports of youths and interviewers' impressions of the reliability of informants. J Child Adolesc Psychopharmacol 21: 417–424.

de Vet HCW, Terwee CB, Knol DL & Bouter LM (2006): When to use agreement versus reliability measures. J Clin Epidemiol 59: 1033–1039.

Dutton GN & Jacobson LK (2001): Cerebral visual impairment in children. Sem Neonatol 6: 477–485.

Eiser C & Morse R (2001): Can parents rate their child's health-related quality of life? Results of a systematic review. Qual Life Res 10: 347–357.

Eiser C & Varni JW (2013): Health-related quality of life and symptom reporting: similarities and differences between children and their parents. Eur J Pediatr 172: 1299–1304.

Elsman EB, Koel M, van Nispen RM & van Rens GH (2021a): Quality of life and participation of children with visual impairment: comparison with population reference scores. Invest Ophthalmol Vis Sci 62: 14.

Elsman EBM, Peeters CFW, van Nispen RMA & van Rens GHMB (2020a): Network analysis of the Participation and Activity Inventory for Children and Youth (PAI-CY) 7–12 years with visual impairment. Transl Vis Sci Technol 9: 19.

Elsman EBM, van Nispen RMA & van Rens GHMB (2017): Feasibility of the Participation and Activity Inventory for Children and Youth (PAI-CY) and Young Adults (PAI-YA) with a visual impairment: a pilot study. Health Qual Life Outcomes 15.

Elsman EBM, van Nispen RMA & van Rens GHMB (2020b): First stage psychometric testing of a new instrument for adolescents with visual impairment: the Participation and Activity Inventory for Children and Youth (PAI-CY) 13–17 years. J Patient-Rep Outcomes 4: 1–10.

Elsman EBM & Van Nispen RMA & Van Rens GHMB (2020c): The Participation and Activity Inventory for Children and Youth (PAI-CY) and Young Adults (PAI-YA) - Manual and Questionnaires.

Gisev N, Bell JS & Chen TF (2013): Interrater agreement and interrater reliability: key concepts, approaches, and applications. Res Soc Admin Pharm 9: 330–338.

Gothwal VK, Bharani S & Mandal AK (2018): Parent-child agreement on health-related quality of life in congenital glaucoma. Transl Vis Sci Technol 7: 15.

Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J & Black N (2018): How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient-Rep Outcomes 2: 1–28.

Hamblion EL, Moore AT & Rahi JS (2011): The health-related quality of life of children with hereditary retinal disorders and the psychosocial impact on their families. Invest Ophthalmol Vis Sci 52: 7981–7986.

Ingerski LM, Modi AC, Hood KK et al. (2010): Health-related quality of life across pediatric chronic conditions. J Pediatrics 156: 639–644.

Janicke DM, Finney JW & Riley AW (2001): Children's health care use. Med Care 39: 990–1001.

Janicke DM, Marciel KK, Ingerski LM, Novoa W, Lowry KW, Sallinen BJ & Silverstein JH (2007): Impact of psychosocial factors on quality of life in overweight youth. Obesity 15: 1799–1807.

Kaplan RM (2001): Quality of life in children: a health care policy perspective. In: Koot HM & Wallander JL (eds.) . Quality of life in child and adolescent illness: concepts, methods and findings. East Sussex: Brunner-Routledge 89–120.

KIDSCREEN. (2006): The KIDSCREEN. Questionnaires - Quality of life questionnaires for children and adolescents. Handbook. Lengerich: Pabst Science Publishers.

Kobayashi K & Kamibeppu K (2011): Quality of life reporting by parent–child dyads in Japan, as grouped by depressive status. Nurs Health Sci 13: 170–177.

Kottner J, Audige L, Brorson S et al. (2011): Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. Int J Nurs Stud 48: 661–671.

Le Coq EM, Boeke AJP, Bezemer PD, Colland VT & van Eijk JTM (2000): Which source should we use to measure quality of life in children with asthma: the children themselves or their parents? Qual Life Res 9: 625–636.

Matza LS, Swensen AR, Flood EM, Secnik K & Leidy NK (2004): Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. Value in health 7: 79–92.

McDougall J, Bedell G & Wright V (2013): The youth report version of the Child and Adolescent Scale of Participation (CASP): assessment of psychometric properties and comparison with parent report. Child Care Health Dev 39: 512–522.

Oltean II & Ferro MA (2019): Agreement of child and parent-proxy reported health-related quality of life in children with mental disorder. Qual Life Res 28: 703–712.

Petsios K, Priftis KN, Tsoumakas C, Hatziagorou E, Tsanakas JN, Galanis P, Antonogeorgos G & Matziou V (2011): Level of parent–asthmatic child agreement on health-related quality of life. J Asthma 48: 286–297.

Ravens-Sieberer U, Auquier P, Erhart M, Gosch A, Rajmil L, Bruil J, Czemy L (2007a): The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. Qual Life Res 16: 1347–1356.

Ravens-Sieberer U, Auquier P, Erhart M, Gosch A, Rajmil L, Bruil J, Kilroe J (2007b): The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. Qual Life Res 16: 1347–1356.

Tadić V, Cooper A, Cumberland P, Lewando-Hundt G & Rahi JS (2013): Development of the functional vision questionnaire for children and young people with visual impairment: the FVQ_CYP. Ophthalmology 120: 2725–2732.

Tadic V, Cooper A, Cumberland P, Lewando-Hundt G & Rahi JS (2016): Measuring the quality of life of visually impaired children: first stage psychometric evaluation of the novel VQoL_CYP instrument. PLoS One 11: e0146225.

Tadić V, Cumberland PM, Lewando-Hundt G & Rahi JS (2017): Do visually impaired children and their parents agree on the child's vision-related quality of life and functional vision? Br J Ophthalmol 101: 244–250.

Theunissen NCM, Vogels TGC, Koopman HM, Verrips GHW, Zwinderman KAH, Verloove-Vanhorick SP & Wit JM (1998): The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res 7: 387–397.

Upton P, Lawford J & Eiser C (2008): Parent–child agreement across child health-related quality of life instruments: a review of the literature. Qual Life Res 17: 895–913.

van Dijk J, Huisman J, Moll AC, Schouten-van Meeteren AYN, Bezemer PD, Ringens PJ, Cohen-Kettenis PT & Imhof SM (2007): Health-related quality of life of child and adolescent retinoblastoma survivors in the Netherlands. Health Qual Life Outcomes 5: 1–8.

Van Rens GHMB, Vreeken HL & Van Nispen RMA (2011): Guideline visual impairment, rehabilitation and referral [Richtlijn visusstoornissen, revalidatie en verwijzing]. Nijmegen, the Netherlands: Dutch Society of Ophthalmology [Nederlands Oogheelkundig Gezelschap].

Varni JW, Burwinkle TM & Lane MM (2005): Health-related quality of life measurement in pediatric clinical practice: an appraisal and precept for future research and application. Health Qual Life Outcomes 3: 1–9.

Varni JW, Limbers CA & Burwinkle TM (2007): How young can children reliably and validly self-report their health-related quality of life?: An analysis of 8,591 children across age subgroups with the PedsQL™ 4.0 Generic Core Scales. Health Qual Life Outcomes 5: 1–13.

Varni JW, Seid M & Kurtin PS (2001): PedsQL™ 4.0: Reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient populations. Med Care 39: 800–812.

Verrips GHW, Vogels AGC, Ouden ALD, Paneth N & Verloove-Vanhorick SP (2000): Measuring health-related quality of life in adolescents: agreement between raters and between methods of administration. Child Care Health Dev 26: 457–469.

Waters E, Stewart-Brown S & Fitzpatrick R (2003): Agreement between adolescent self-report and parent reports of health and well-being: results of an epidemiological study. Child Care Health Dev 29: 501–509.

WHO (2010): International Statistical Classification of Diseases and Related Health Problems 10th Revision. Version 2010. Chapter VII, H54: Visual impairment including blindness. Geneva: World Health Organization.

Correspondence:
Ellen BM Elsman, PhD
Amsterdam UMC
location VUmc
KTC4-027
PO Box 705
1007
MB Amsterdam
The Netherlands
Tel: +31 20 4444795
Fax: +31 20 4444745
Email: e.elsman@amsterdamumc.nl