Title:

An age- and stage-appropriate patient-reported outcome measure of vision-related quality of life (VQoL) of children and young people with visual impairment

Short title:

Vision-related quality of life of children/young people

Authors:

Valerija Tadić (MSc, PhD)[1,2]
Alexandra O Robertson (MSc, PhD)[2]
Mario Cortina-Borja (MSc, PhD)[2]
Jugnoo S Rahi (PhD, FRCOphth) [2,3,4,5] *for the Child Vision PROMs group**
 *Members of the Child Vision PROMs group are listed in the Acknowledgements. Professor Rahi is the study chair.
** Both Dr. Tadić and Dr. Robertson contributed equally as first authors.

Affiliations:

1 School of Human Sciences, University of Greenwich
2 Population, Policy and Practice Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, UK
3 Great Ormond Street Hospital NHS Foundation Trust, UK
4 National Institute for Health Research (NIHR) Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK
5 Ulverscroft Vision Research Group, London, UK

Conflict of interest:

No conflicting relationship exists for any author.

Corresponding author (and address for reprints):

Jugnoo S Rahi, UCL Great Ormond Street (GOS) Institute of Child Health, Life Course Epidemiology and Biostatistics Section, Population, Policy and Practice Programme, 30 Guilford Street, London WC1N 1EH, UK; Telephone: 44 (0)20 7905 2250; Email: j.rahi@ucl.ac.uk

Online supplemental materials:

The following should appear online-only: Table 4 and Table 5.

1   **ABSTRACT**

2   **Purpose:** Developmentally sensitive measures of vision-related quality of life (VQoL) are

3   needed to capture age-specific concerns about the impact of living with visual impairment

4   (VI) in children and young people. Our objective was to use our validated vision-related

5   quality of life instrument for children and young people aged 10-15 years (the VQoL_CYP) as

6   the foundation for development of age-specific extensions.

7   **Design:** Questionnaire development

8   **Participants:** A representative sample of children and young people aged 6-19 years with

9   visual impairment, visual acuity of the logarithm of the minimum angle of resolution

10 (LogMAR) worse than 0.50 in the better eye. They were recruited from Paediatric

11 Ophthalmology clinics at Great Ormond Street Hospital and Moorfields Eye Hospital and in

12 the final phase of the study from 20 additional UK hospitals.

13 **Methods:** Standard instrument development processes were followed across four phases.

14 29 semi-structured interviews with children and young people permitted draft age-appropriate

15 extensions. 28 cognitive interviews informed items and response options. Age-appropriate

16 extensions were pre-piloted with 49 participants to ensure feasibility, and administered via a

17 postal survey to a national sample of 160 for psychometric evaluation using Rasch analysis.

18 Construct validity was evaluated through correlations with the Pediatric Quality of Life

19 Inventory (PedsQL).

20 **Main Outcome measures:** Psychometric indices of validity and reliability of the instrument

21 versions.

22 **Results:** Interviews confirmed the existing VQoL_CYP content and format were relevant

23 across a wider age-range. Age-appropriate extensions were drafted for children (8-12 years)

24 and young people (13-17 years). Psychometric item reduction produced 20-item child and

25 22-item young person versions, each with acceptable fit values, no notable differential item

26   functioning, good measurement precision, ordered response categories and acceptable

27   targeting, and no notable differential item functioning on items common to both. Construct

28   validity was demonstrated through correlations with health-related quality of life ($r = 0.698$).

29   **Conclusions:** Using an efficient child/young person-centred approach we have developed

30   two robust, age-appropriate versions of an instrument capturing VQoL that can be used

31   cross-sectionally or sequentially across the age-range of 8-17 years in research and clinical

32   practice. This approach may be applicable in other rare childhood ophthalmic disorders.

The use of patient-reported outcomes measures (PROMs) is now well established in both clinical practice and in research evaluating new treatments.[1] Patient-reported outcome measures enabling self-report of health-related quality of life (HRQoL), which cannot be captured through objective clinical assessments, are particularly important. Generic HRQoL measures[2,3] designed with developmental differences in mind, have followed the standard approach of concurrent development of age-appropriate instrument versions across different age groups, by drawing on the whole population. This approach is challenging in populations with rare ophthalmic disorders such as those causing visual impairment or blindness (VI for brevity throughout).

Visually impairing disorders collectively affect about 2 per 1000 children and young people in industrialised countries.[4,5] Most children and young people with VI are affected from infancy. All will face significant lifelong challenges through the impact on development, education, social and emotional wellbeing alongside high economic costs for affected individuals, their families and society.[6] In the industrialised world and increasingly in developing countries, most affected individuals have disorders that are currently neither preventable nor curable. There is therefore a strong focus on maintaining residual vision and functional abilities in order to maximise vision-related quality of life (VQoL). However reliable and valid measures of VQoL in children and young people remain scarce, partly due to the challenges of research on populations with rare disorders.[7] Hitherto, most PROMs for children and young people with ophthalmic conditions, including those designed to assess VQoL, comprised either a single instrument used across a very wide age-range[8,9] or age-specific versions without age-appropriate items or response formats.[10] Thus, they do not take account of the development of children's understanding of illness, health and quality of life (QoL) and how this changes as they mature, and cannot capture developmental differences or age-specific needs in terms of content, response options and ability to complete independently.[11,12]

We recently reported the first stage psychometric validation of a 35-item instrument measuring self-reported VQoL in children and young people with VI aged 10-15 years - the

60 VQoL_CYP.[13,14] To ensure content validity, we undertook semi-structured and cognitive

61 debriefing interviews. In the absence of both an existing conceptual framework and an

62 established methodology for developing measures for this numerically small population, we

63 deliberately targeted the 10-15 years age-group in this foundation research, as most capable

64 of identifying the impact of living with VI through individual interviews and self-completing the

65 instrument with ease. We now report our planned extension and adaption of that foundation

66 instrument[13,14] to a broader age-range, including our novel approach of calibrating the new

67 age-appropriate versions so that they can be used and compared in different age-groups at

68 any given point but also be used to follow subjects over time as they grow older i.e.

69 sequentially. Our decision to set the minimum age threshold at 8 years reflects the age from

70 which self-report becomes reliable, and our maximum age threshold reflects the age of

71 transition into adult services.

72 **METHODS**

73 The study was approved by the National Health Service (NHS) Research Ethics Committee

74 for Essex and East of England, United Kingdom (UK) and followed tenets of the Declaration

75 of Helsinki. Participants gave informed individual assent (if <16 years) or consent and

76 parents gave informed consent to their child's participation (if <16 years).

77 Sample

78 Children and young people were eligible if they were *i)* visually impaired, severely

79 visually impaired or blind (visual acuity in the better eye of LogMAR 0.50 or worse or Snellen

80 worse than 6/18 or additional visual defects causing visual impairment) due to any visual

81 disorder, but without any other significant impairment (i.e., learning, sensory or motor); and *ii)*

82 aged 6-19 years (with age boundaries for the instrument determined later). They were drawn

83 from 2 patient populations between September 2014 and May 2017 comprising those

84 attending the Department of Ophthalmology at Great Ormond Street Hospital and the

85 Pediatric Glaucoma Service and Genetic Eye Disease Service at Moorfields Eye Hospital,

86 London UK supplemented (final phase only) by patients attending 20 other hospitals across

87 Britain (see Acknowledgments). By sampling across multiple sources nationally in the final

88 phases, in which the largest samples are needed, we ensured our sample was as

89 representative as possible of the UK population of children and young people with VI with

90 respect to ethnic and socio-economic status.

91 <u>Procedures</u>

92 Instrument adaptation followed standard instrument development phases, with our

93 'foundation' research with 10-15 year olds[13,14] as the framework.

94 *Phase 1: Item development and adaptation*

95 To investigate whether the issues covered by the existing VQoL_CYP items (from the 10-15

96 year olds' instrument[13,14]) were relevant to children/young people outside the age-range of

97 10-15 years and identify any new age-specific issues not already included, we conducted

98 individual in-depth semi-structured interviews with children younger than 10 and young

99 people older than 15 years. Building on the foundation of the existing VQoL_CYP instrument,

100 which was based on 32 interviews with 10-15 year olds, we reached data saturation after 29

101 interviews (12 with children aged 6-9 years, 17 with young people aged 16-19 years).

102 Interviews were transcribed and coded using NVivo10.[15] We used the thematic framework

103 developed through qualitative thematic analysis in the foundation study that produced the

104 existing VQoL_CYP instrument for 10-15 year olds, to identify areas of overlap and

105 discrepancy between the new interview data and the existing instrument. Where omissions

106 were identified, new, age-appropriate items were developed.

107 Additionally, to ensure that the subsequent first draft version of the instrument version for

108 younger children was developmentally appropriate, participants <10 years were asked to

109 complete the existing VQoL_CYP (10-15 years)[13,14] with parental assistance and provide

110 feedback to inform development of the subsequent age-appropriate version. This was not

111 considered necessary for participants older than 15 years, who were developmentally well

112 placed to comprehend the existing VQoL_CYP (10-15 years) items.

113 *Phase 2: Pre-testing*

114 The upper and lower age boundaries of each new age-appropriate VQoL instrument version

115 were developed empirically throughout Phase 2, whilst considering data also from the early

116 interview phases of the VQoL_CYP (10-15 years) development.[13] Due to the extensive

117 foundation work in development of the original instrument for 10-15 year olds and the

118 resemblance of the new age-appropriate drafts to the published instrument, recruitment in

119 this phase was focused primarily on participants younger than 10 and older than 15 years.

120 Individual cognitive interviews with 12 children aged 7-10 years and 16 young people aged

121 13-18 years ensured comprehensibility of the new age-appropriate draft instrument versions.

122 This was supplemented by parental feedback on the same items presented to children and

123 young people and study group consensus. Items were refined accounting for importance,

124 comprehensibility, difficulty and response format. Alongside re-reading of the original

125 individual interviews with 10-15 year olds,[13] feedback from children and young people, their

126 parents, and study group consensus was used to determine the age thresholds for the new

127 instrument versions as 8-12 years (VQoL_Child) and 13-17 years (VQoL_Young Person).

128 *Phase 3: Pre-piloting*

129 Pre-piloting of the modified new instrument versions comprised a postal survey of 26 children

130 aged 8-12 years and 23 young people 13-17 years, to ensure feasibility with respect to

131 missing data and administration burden and to inform initial decisions about subsequent item

132 reduction.

133 Participants received a pack comprising invitation letters, child and parent information sheets

134 and consent/assent forms, the age-appropriate instrument versions in large print (including a

135 link to an electronic version) and a postage-paid envelope for return of the completed

136   materials. Participants were invited to provide written qualitative feedback. Questionnaire

137   data were verified by checking the study database, with no errors detected.

138   *Phase 4: Piloting*

139   Formal piloting comprised a large-scale postal survey of a national sample (UK) of 87

140   children aged 8-12 years and 73 young people aged 13-17 years to confirm psychometric

141   properties of the two new instrument versions. The VQoL_Child and the VQoL_Young

142   Person were administered alongside the Child (8-12 years) and Teenager (13-18 years)

143   versions of the Pediatric Quality of Life Inventory (PedsQL[3]) to assess construct validity. The

144   PedsQL, a validated generic HRQoL instrument, produces Total, Physical Health and

145   Psychosocial Health Scores, with higher scores indicating better HRQoL.[3,16]

146   Participants received study packs, as per previous phases. Questionnaire data were verified

147   through double-checking the study database and any data-entry errors were corrected.

148   *Psychometric evaluation*

149   In keeping with published criteria,[17] data from participants with >25% of item responses

150   missing were excluded, as were items for which >50% of participant responses were

151   missing.

152   Rasch analysis[18-22] was used for item reduction and psychometric assessment using

153   Andrich's Rasch Rating Scale model.[23] Established criteria were used to assess the

154   appropriateness of the two instruments,[17,24] as detailed in Table 2 and Figures 1 and 2. Prior

155   to conducting Rasch analysis negatively worded items were reversed, and 1-4 responses

156   were coded into scores of 0 to 3.

157   *Calibration of VQoL_Child and VQoL_Young Person.*

158   The model resulting from equating both instruments, as outlined by Lincacre[25] ensured that

159   the age-appropriate instrument versions were capable of measuring the same construct in

160   children and young people. This model, based on the overlapping items on both age-

9

161 dependent instruments provides continuity of measurement for ages 8 to 17 years, ensuring

162 the instruments can be used in cross-sectional studies. It also allows comparisons of

163 summary scores measured during follow-up of individuals as they grow older (i.e. *sequential*

164 use). These scores are obtained as the sum of all individual item raw scores, and can be

165 transformed into Rasch person measures using Table 5 (available at www.aaojournal.org).

166 This transformation assumes that all items have equal importance, and that response

167 categories are scaled accordingly to yield an equal value with uniform increments between

168 consecutive categories. To examine whether the equated Rasch person measures from the

169 two age groups (8-12 and 13-17 years) were comparable in this way, a final differential item

170 functioning (DIF) analysis was conducted using the 'core' set of items common to both.[26]

171 Unidimensionality was assessed using infit and outfit statistics, and the criteria described in

172 Table 2.[17] DIF statistics, shown in Table 2 represent the effect size, in logits of the difference

173 between the two classifications of persons.[27]

174 *Construct validity*

175 VQoL summary scores were calculated and converted into Rasch person measures ranging

176 from 0 (severely reduced VQoL) to 100 (excellent VQoL) using the score-to-measure tables

177 for each age-appropriate version (Table 5, available at www.aaojournal.org), ensuring the

178 derived measures can be compared between age-appropriate versions despite differences in

179 the number and wording of items.

180 Construct validity (i.e. instrument's ability to truly measure an intended outcome) was

181 assessed through correlations between Rasch person measures on the VQoL_Child and

182 VQoL_Young Person and scores on the Child and Teen PedsQL (Total and Psychosocial

183 subscale summaries). Participants with any missing responses were excluded from the

184 analyses. Additionally correlation between Rasch person measures on the VQoL_Child and

185 VQoL_Young Person and visual acuity was examined, without anticipation of a correlation, in

186 keeping with the 'disability paradox'.[28]

187 Correlations with PedsQL were examined using the Rasch person measures for each new

188 VQoL version individually, before combining scores from both age-appropriate versions.

189 Spearman's Rank correlations were reported.

190 Rasch analysis was conducted using Winsteps, 4.0.1.[29] All other analyses were completed

191 using SPSS.

192 **RESULTS**

193 Table 1 shows the participant characteristics across the study phases, illustrating an

194 unbiased representation of the overall UK population of children and young people with VI

195 with respect to clinical and socio-demographic characteristics and ophthalmic diagnoses

196 (given the exclusion of participants with any other significant impairment).[5,13,14]

197 *Phase 1: Item development and adaptation*

198 Analysis of the new interview data revealed significant overlap between the issues raised by

199 children younger than 10 and young people older than 15 years, and the issues covered by

200 the existing VQoL_CYP instrument for 10-15 year olds.[13,14] Where age-related variation

201 emerged it was in descriptions/and attributions of issues to QoL, rather than differences in

202 the type of issues experienced, necessitating some adaptations. For the older age group, 11

203 items removed during the foundation research were reinstated based on views expressed in

204 the interviews regarding relevance. A new item on tiredness and impact on sleep, as flagged

205 by participants, was added.

206 The format involving the illustrative child/3rd person vignette was changed as a result of

207 significant skew in VQoL_CYP items presented on the 'ideal status' scale in the foundation

208 study.[14] All items were re-worded as first person statements (e.g. 'I feel left out because of

209 my eyesight') and response categories amended accordingly whereby the responding

210 child/young person reported how true each statement was about him/her. Four response

211 categories were developed and refined, considering children and young people's natural

212 vocabulary used during interviews (1-Not at all true, 2-A little bit true, 3-Mostly true, 4-

213 Completely true). Instrument instructions ensured the respondent children and young people

214 made their responses in relation to their eyesight.

215 The resulting draft 31-item VQoL_Child and 37-item VQoL_Young Person versions for

216 children aged <10 years and young people aged >15 years, were pre-tested.

217 *Phase 2: Pre-testing*

218 A small number of items considered ambiguous by participants were re-phrased or removed.

219 The minimum age threshold was agreed as 8 years and age boundaries re-adjusted as 8-12

220 years and 13-17 years, thus aligning to other child PROMs.[3] The resulting 29-item

221 VQoL_Child and 39-item VQoL_Young Person extensions were pre-piloted.

222 *Phase 3: Pre-piloting*

223 The participation rates were 44.1 % and 31.1% for children and young people respectively.

224 Median completion time was 15 minutes (IQR=13) for children and 10 minutes (IQR=23.75)

225 for young people, with 86% and 95% of children and young people respectively rating

226 instrument completion as easy/very easy, and 95% and 100% respectively rating the

227 instructions as easy/very easy.

228 Data from one child were excluded due to 76% missing data. There were no missing

229 responses in the child dataset and a small (≤10.26%) number of missing values per item in

230 the young people's dataset.

231 The number of items with over 50% of responses or 0% responses in an 'end' category were

232 8 and 4 respectively in the child and 5 and 13 in the young person dataset. Items with

233 problematic distribution were flagged for potential removal during formal piloting of the 30-

234 item VQoL_Child and 39-item VQoL_Young Person.

235 *Phase 4: Piloting*

236 The participation rates were 31.4% and 26.4% for children and young people respectively.

237 Missing data per item (completely at random) were less than 3% for both instrument

238 versions. Two children (but no young people) were excluded from subsequent analysis

239 based on having >25% missing data per person. All remaining missing data per person was

240 found to be missing completely at random (MCAR),[30] and retained for Rasch analyses.[31]

241 *Psychometric evaluation*

242 Six items were removed from the VQoL_Child and 5 from the VQoL_Young Person due to

243 significant skewness, and ceiling effects and a further 4 and 12 respectively during Rasch

244 based on goodness-of-fit, response ordering and DIF statistics (Table 4, available at

245 www.aaojournal.org). The resulting 20-item child and 22-item young person instrument

246 versions showed these statistics to be within acceptable limits. One item fell just outside the

247 acceptable criteria for only goodness-of-fit criterion but was retained in the VQoL_Young

248 Person to preserve content validity and comparability with VQol_Child where it was retained

249 (Table 2). For each version, the item probability plots showed good ordering, and acceptable

250 differentiation between the 4 response categories (Figure 1) and targeting of items to

251 respondents (the difference between person and item means = 0.81 logits (child version) and

252 0.76 (young person version)) although items were clustered around the mid-low end of the

253 item difficulty scale (Figure 2). Each version showed good precision as indicated by indices

254 for person separation (3.64 and 2.74 for child and young person versions respectively).[17,32]

255 The final 20 item VQoL_Child and 22 item VQoL_Young Person scales included 12 common

256 'core' items and 8 and 10 age specific items respectfully.

257 <u>*Calibration of the VQoL_Child and VQol_Young Person instrument versions*</u>

258 Differential item functioning analysis of overlapping core items showed no contrasts greater

259 than 1 logit (Table 2), demonstrating they were not biased to either age group (after adjusting

260 for the overall scores of respondents). Thus, all remaining overlapping items are productive

261 for measurement of VQoL in both instrument versions despite the presence of additional,

262 age-specific items.

263 *Score-to-measure transformation*

264 To enable easy and precise scoring, we developed conversion tables for transforming the

265 summary scores to Rasch person measures, as shown in Table 5 (available at

266 www.aaojournal.org). These can be used to compare Rasch person measures when using

267 either or both versions cross-sectionally or sequentially.

268 <u>*Construct validity*</u>

269 We excluded 6 children and 5 young people with missing data before analysing construct

270 validity. Rasch person measures on the VQoL_Child and VQoL_Young Person correlated

271 positively with Child and Teen PedsQL scores, substantiating the instrument's construct

272 (convergent) validity (Table 3). As anticipated, acuity did not correlate significantly with

273 VQoL.

274 **DISCUSSION**

275 We report an effective, efficient and child/young person-centred approach to developing an

276 age-appropriate PROM for children and young people with VI. Using a novel approach for

277 calibrating instruments and exploiting our prior research and original instrument for those

278 aged 10-15 years,[13,14] we have generated two psychometrically robust versions of this

279 measure that are suitable for a wider age-range, spanning 8-17 years, whilst retaining

280 developmentally appropriate content through a modular structure of common core items

281 alongside age-group specific items. Using this approach, we have improved feasibility for

282 both patients and clinicians. Our final 20- and 22-item VQoL_Child and VQoL_Young Person

283 instrument versions, respectively, are shorter than our original version for 10-15 year olds

284 and reported to be easy to complete without sacrificing comprehensiveness. We have

285 calibrated the two age-specific versions using overlapping core items, so that the correct

286   instrument version can be used based on the age of children in the study at that time point

287   and also so that VQoL can be measured without loss of continuity of measurement as the

288   subjects get older by using the alternative instrument version. Thus, these versions can be

289   used both cross-sectionally (e.g. in trials with a wide age-range of subjects) and sequentially

290   (e.g. in cohort studies or clinical follow up of individual patients) in future studies and

291   research. Our log transformation tables, which convert summary scores into Rasch person

292   measures, provide clinicians the means for using and interpreting scores with precision and

293   ease. We also provide the model-based standard error of each measure, which should be

294   used in future clinical research implementing the instruments.

295   Our two new instrument versions (like the original VQoL_CYP[13,14]), show good construct

296   validity, correlating strongly with HRQoL on a generic measure (particularly its psychosocial

297   component). As anticipated,[14] the VQoL scores for both children and young people were not

298   associated with visual acuity. These findings align with the 'disability paradox'.[28,33,34] This

299   phenomenon, whereby individuals with severe disabilities or illnesses report good QoL,

300   exemplifies the importance of considering QoL to be a subjective construct.[35] Thus the child

301   or young person with VI will construct his/her perception of their QoL from the subjective day-

302   to-day experience of living with a visual disability and ultimately, their scores on a self-

303   reported QoL measure will reflect this. This has important implications for how the

304   VQoL_Child and VQoL_Young Person, and indeed any child QoL PROMs, should be used.

305   For instance, in the context of trials of new interventions or therapies intended to improve

306   vision, the implications of the 'disability paradox' must be recognised to avoid conclusions

307   about impact of interventions being misconstrued.

308   Although the new VQoL instrument versions are age-group specific (for example, concerns

309   about independent living in the future feature only in the VQoL_Young Person) the significant

310   overlap in common content across the two versions, as well as with our original

311   VQoL_CYP,[13,14] demonstrates the core life trajectory of children with VI whereby concerns

312   (e.g. social inclusion and acceptance) and barriers (e.g. in education) emerge and establish

313   across childhood and adolescence. This is likely to be true also for other child populations.

314   Moreover, issues related to VI align with other disabilities as well as other chronic complex

315   childhood conditions, as evidenced by the content of similar HRQoL measures[2,3,35] and by

316   the significant correlations with the PedsQL in our study, thereby affirming the strong content

317   and construct validity of the VQoL_Child and VQoL_Young Person.

318   Although we achieved a good sized sample relative to the rarity of childhood VI, a more

319   granular examination of the underlying domain structure in the instrument was not possible

320   due to limited power. We followed the conventional approach of using infit and outfit statistics

321   to remove items until all the stringent criteria have been met.[17] Unidimensionality, for each

322   instrument version was sufficiently evidenced by the ranges of infit and outfit statistics which

323   support the derivation of a summary score, and the scale items span the spectrum of aspects

324   of QoL suggested by broader literature, [2,35] demonstrating good face validity.

325   Recognising the lack of instruments suitable for the youngest children with VI and cognisant

326   that some children can self-report reliably from as young as 5 years,[12,36,37] we conducted

327   some semi-structured and cognitive interviews with children younger than 8 years but found

328   both recruitment and information capture challenging despite using different child-appropriate

329   methods. This highlights an important direction for future research. In the meantime, the age-

330   range served by our instrument coincides with that recommended and reported in the

331   literature,[12,16] and enables complementary use of generic HRQoL instruments.

332   We found both the VQoL_Child and VQoL_Young Person to be somewhat better targeted to

333   participants reporting lower VQoL. This is comparable to the targeting pattern we reported for

334   our original instrument for 10-15 year olds[14] as well as that reported in the development of

335   the impact of vision impairment for children (IVI_C),[8] which is a similar instrument developed

336   in Australia to assess VQoL of children and young people with VI. Given that the items seem

337   more suited to children with lower VQoL, these instruments may be particularly useful in

338   assessing VQoL changes in visually impaired children and young people who are at risk of

339   lower QoL, for instance, due to receiving less professional support (e.g. in education) and in

340   the context of relevant interventions aimed at increasing such support.

341   Differential item functioning analyses can be unstable and produce spurious results when

342   applied to small samples. In particular, they often reflect an increased chance of false

343   positive findings (i.e. removal of too many items).[38] In the case of questionnaire

344   development, this means that a shorter scale will be produced. This is not the case for the

345   reduced VQoL_Child and VQoL_Young Person instrument versions which have a good

346   coverage of all elements of VQoL.

347   Ethical and practical considerations involved in re-testing participants precluded examination

348   of test-retest reliability and responsiveness of the measure over time. We will address this in

349   our planned research on optimal approaches to routine implementation of vision PROMs in

350   clinical practice, to assess how our VQoL instrument can best be deployed alongside our

351   other vision PROM assessing functional vision[39] to enable a holistic assessment of impact

352   and thus truly 'personalised' care.

353   It is challenging but possible to generate psychometrically robust and developmentally

354   appropriate instruments usable by the whole age-range of children and young people with VI.

355   Our novel approach for vision specific PROMs enables a measurement model in which

356   instruments can be used cross-sectionally and sequentially in both clinical practice and

357   research. We suggest the approach we have described is transferable to other childhood

358   ophthalmic conditions and is a parsimonious approach useful in research on rare conditions.

359   Small sample sizes, inherent in research on rare paediatric populations such as children and

360   young people with VI can preclude *concurrent* de novo development of age-group specific

361   measures. We have overcome the challenges posed by limited sample sizes by starting with

362   a foundation instrument that is anchored to the middle of the overall age-range (10-15

363   years),[13,14] and using this as the basis for extending the age-range in both directions.

364   Figure 1: Category probability curves showing the probability of selecting response

365   categories across the scale of item difficulty for age-appropriate extensions of the

366   VQoL_CYP[40]

367   <u>Figure 1a: Category probability curves for the 20-item VQoL_Child</u>

368   <u>Figure 1b: Category probability curves for the 22-item VQoL_Young Person</u>

369

370   Figure 2: Item-person maps illustrating acceptable targeting of VQoL items (located on the

371   right side of the dashed line) to responders (located on the left side of the dashed line and

372   represented by X).[32] Participants with higher VQoL and items with higher difficulty to endorse

373   as true are at the top half of the map.

374   <u>Figure 2a: Item-Person map for the VQoL_Child</u>

375   <u>Figure 2b: Item-Person map for the VQoL_Young Person</u>

376   M = mean; S = 1 standard deviation from the mean; T = 2 standard deviations from the

377   mean.

REFERENCES

1. Black N. Patient reported outcome measures could help transform healthcare. *BMJ.* 2013;346:f167

2. Ravens-Sieberer U, Gosch A, Rajmil L, et al. KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Rev Pharmacoecon Outcomes Res.* 2005;5(3):353-364.

3. Varni JW, Seid M, Rode CA. The PedsQL™:measurement model for the pediatric quality of life inventory. *Med Care.* 1999;37(2):126-139.

4. Solebo AL, Rahi J. Epidemiology, aetiology and management of visual impairment in children. *Arch Dif Child.* 2014;99(4):375-379.

5. Rahi JS, Cable N. Severe visual impairment and blindness in children in the UK. *Lancet.* 2003;362(9393):1359-1365.

6. Wittenborn JS, Zhang X, Feagan CW, et al. The Economic Burden of Vision Loss and Eye Disorders among the United States Population Younger than 40 Years. *Ophthalmology.* 2013;120(9):1728-1735.

7. Tadić V, Hogan A, Sobti N, Knowles RL, Rahi JS. Patient-reported outcome measures (PROMs) in paediatric ophthalmology: a systematic review. *Br J Ophthalmol.* 2013;97:1369-1381.

8. Cochrane GM, Marella M, Keeffe JE, Lamoureux EL. 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.* 2011;52(3):1632-1640.

9. Khadka J, Ryan B, Margrain TH, Court H, Woodhouse JM. Development of the 25-item Cardiff Visual Ability Questionnaire for Children (CVAQC). *Br J Ophthalmol.* 2010;94(6):730-735.

10. Hatt SR, Leske DA, Castañeda YS, et al. Development of pediatric eye questionnaires for children with eye disease. *Am J Ophthalmol.* 2019;200:201-217.

437   11.   Bevans KB, Riley AW, Moon J, Forrest CB. Conceptual and methodological

438   advances in child-reported outcomes measurement. *Expert Rev Pharmacoecon*

439   *Outcomes Res.* 2010;10(4):385-396.

440   12.   Matza LS, Patrick DL, Riley AW, et al. Pediatric Patient-Reported Outcome

441   Instruments for Research to Support Medical Product Labeling: report of the ISPOR

442   PRO Good Research Practices for the Assessment of Children and Adolescents Task

443   Force. *Value Health.* 2013;16(4):461-479.

444   13.   Rahi JS, Tadić V, Keeley S, Lewando-Hundt G. Capturing Children and Young

445   People's Perspectives to Identify the Content for a Novel Vision-Related Quality of

446   Life Instrument. *Ophthalmology.* 2011;118(5):819-824.

447   14.   Tadić V, Cooper A, Cumberland P, Lewando-Hundt G, Rahi J. Measuring the quality

448   of life of visually impaired children: first stage psychometric evaluation of the novel

449   VQoL_CYP instrument. *PLoS One.* 2016;11(2):e0146225.

450   15.   Castleberry A. NVivo 10 [software program]. Version 10. QSR International. *Am J*

451   *Pharm Educ.* 2014;7(1):25.

452   16.   Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: reliability and validity of the Pediatric

453   Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient

454   populations. *Med Care.* 2001;39(8):800-812.

455   17.   Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and

456   selection of questionnaires. *Optom Vis Sci.* 2007;84(8):663-674.

457   18.   Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example

458   using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol.*

459   2007;46(1):1-18.

460   19.   Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is

461   it and why use it? When should it be applied, and what should one look for in a Rasch

462   paper? *Arthritis Care Res.* 2007;57(8):1358-1362.

463   20.   Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: an
464   introduction and illustrative example. *Int J Nurs Stud.* 2009;46(3):380-393.

465   21.   Prieto L, Alonso J, Lamarca R. Classical test theory versus Rasch analysis for quality
466   of life questionnaire reduction. *Health Qual Life Outcomes.* 2003;1(1):1-13.

467   22.   Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago,
468   MESA Press; 1993.

469   23.   Wright BD, Masters GN. *Rating Scale Analysis.* Chicago: MESA Press; 1982.

470   24.   Wright BD, GA Douglas. *Best Test Design and Self-Tailored Testing.* Statistical
471   Laboratory, Department of Education: The University of Chicago; 1975.

472   25.   Linacre J. Equating and linking tests. Help for Winsteps Rasch Measurement and
473   Rasch Analysis Software. https://www.winsteps.com/winman/equating.htm. Accessed
474   July 29, 2019.

475   26.   Wolfe E, Chiu C. Measuring pretest-posttest change with a Rasch Rating Scale
476   Model. *J Outcome Meas.* 1999;3(2):134-161.

477   27.   Dorans NJ, Kulick E. Differential item functioning on the Mini-Mental State
478   Examination: an application of the Mantel-Haenszel and standardization procedures.
479   *Med Care.* 2006:S107-S114.

480   28.   Albrecht GL, Devlieger PJ. The disability paradox: high quality of life against all odds.
481   *Soc Sci Med.* 1999;48(8):977-988.

482   29.   Linacre JM. Winsteps (Version 3.81.0). [Computer software]. Winsteps com: Chicago;
483   2014.

484   30.   Little RJ, Rubin DB. *Statistical Analysis with Missing Data.* John Wiley & Sons: New
485   York; 2019.

486   31.   Waterbury GT. Missing data and the rasch model: the effects of missing data
487   mechanisms on item parameter estimation. *J Appl Meas.* 2019;20(2):154-166.

488   32.   Khadka J, Pesudovs K, McAlinden C, Vogel M, Kernt M, Hirneiss C. Reengineering

489   the glaucoma quality of life-15 questionnaire with rasch analysis. *Invest Ophthalmol*

490   *Vis Sci.* 2011;52(9):6971-6977.

491   33.   Kutner JS, Nowels DE. Confirmation of the"disability paradox" among hospice

492   patients: preservation of quality of life despite physical ailments and psychosocial

493   concerns. *Palliat Support Care.* 2003;1(3):231-237.

494   34.   Watson N. Well, I know this is going to sound very strange to you, but I don't see

495   myself as a disabled person: identity and disability. *Disabil Soc.* 2002;17(5):509-527.

496   35.   WHOQOL Group. The World Health Organization quality of life assessment

497   (WHOQOL): position paper from the World Health Organization. *Soc Sci Med.*

498   1995;41(10):1403-1409.

499   36.   Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. *Health*

500   *Technol Assess (Rockv).* 2001;5(4):1-157.

501   37.   Varni J, Limbers C, Burwinkle T. How young can children reliably and validly self-

502   report their health-related quality of life?: an analysis of 8,591 children across age

503   subgroups with the PedsQLTM 4.0 Generic Core Scales. *Health Qual Life Outcomes.*

504   2007;5(1):1.

505   38.   Scott NW, Fayers PM, Aaronson NK, et al. A simulation study provided sample size

506   guidance for differential item functioning (DIF) studies using short scales. *J Clin*

507   *Epidemiol.* 2009;62(3):288-295.

508   39.   Tadić V, Cooper A, Cumberland P, Lewando-Hundt G, Rahi JS. Development of the

509   Functional Vision Questionnaire for Children and Young People with Visual

510   Impairment: the FVQ_CYP. *Ophthalmology.* 2013;120(12):2725-2732.

511   40.   Linacre J. Optimizing rating scale category effectiveness. *J Appl Meas.* 2002;3(1):85-

512   106.

513

514