

1       **Systematic review recommends the European Organization for**  
2       **Research and Treatment of Cancer colorectal cancer–specific**  
3       **module for measuring quality of life in colorectal cancer patients**

4  
5       Correspondence Author:

6       Name: Carlos King Ho Wong, PhD, MPhil, BSc

7       Institution: Department of Family Medicine and Primary Care, The University of Hong Kong

8       Address: 3/F, Ap Lei Chau Clinic, 161 Ap Lei Chau Main Street, Ap Lei Chau, Hong Kong

9       Contact: +852-25185688 (tel); +852-28147475 (fax) [carlosho@hku.hk](mailto:carlosho@hku.hk) (email)

10  
11       Order of Author: Carlos K.H. Wong, PhD\*<sup>1</sup>, Jing Chen, PhD<sup>2</sup>, Charlotte L.Y. Yu<sup>1</sup>, Mansy  
12       Sham<sup>1</sup>, Cindy L.K. Lam, MD<sup>1</sup>

13       \* First and correspondence Author

14       <sup>1</sup> Department of Family Medicine and Primary Care, The University of Hong Kong

15       <sup>2</sup> School of Public Health, The University of Hong Kong

25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

**Abstract**

**Objective:** To critically appraise the measurement properties of standardized health-related quality of life (HRQOL) instruments for colorectal cancer (CRC) patients, and to provide recommendations on the choice of HRQOL instruments.

**Study Design and Setting:** Systematic review of English-language literature published between January 1985 and May 2014 identified through a database search of Pubmed, Web of Science, Embase, and OVID Medline. HRQOL instruments were rated on methodological quality and overall levels of evidence using a COSMIN checklist.

**Results:** Internal consistency and hypothesis testing were evaluated most frequently in 63 studies identified. The Functional Assessment of Cancer Therapy-Colorectal (FACT-C) was the most extensively evaluated. The highest number of positive ratings in the overall level of evidence was found in the colorectal cancer-specific quality of life questionnaire module (QLQ-CR38) in European Organization for Research and Treatment of Cancer (EORTC) module, followed by the Memorial Sloan Kettering Cancer Center (MSKCC) Bowel instrument, FACT-C and Quick-FLIC. The EORTC QLQ-CR38 had the most positive ratings on measurement property and was recommended.

**Conclusion:** The EORTC QLQ-CR38 was recommended to assess HRQOL in patients with CRC, regardless of disease stage and primary tumour site.

**Running Title:** Recommendation for HRQOL Instruments in CRC

**Keywords:** systematic review; colorectal cancer; quality of life, measurement property, psychometrics; COSMIN



49 **Abbreviations:** CRC=colorectal cancer; HRQOL=health-related quality of life;  
50 MESH=Medical Subject Heading; COSMIN=Consensus-based Standards for the selection of  
51 health Measurement Instruments; PRISMA=Preferred Reporting Items for Systematic  
52 Reviews and Meta-Analyses; ICC=intraclass correlation coefficient; EORTC=European  
53 Organization for Research and Treatment of Cancer; QLQ=Quality-of-Life Questionnaire;  
54 FACIT=Functional Assessment of Chronic Illness Therapy; FACT-G=Functional  
55 Assessment of Cancer Therapy-General; FACT-C=Functional Assessment of Cancer  
56 Therapy-Colorectal; FCSI-9=Functional Assessment of Cancer Therapy Colorectal Symptom  
57 Index; FACIT-F=Functional Assessment of Chronic Illness Therapy-Fatigue; FLIC=Functional  
58 living index for cancer; GIQLI=Gastrointestinal Quality of Life Index; RSCL-M=Modified  
59 Version of Rotterdam Symptom Checklist; MSAS-SF=Memorial Symptom Assessment  
60 Scale; CMSAS=Condensed MSAS; FIQL=Fecal Incontinence Quality of Life; mCOH-QOL-  
61 Ostomy=Modified City of Hope Ostomy questionnaire; SRQS=Social Relational Quality  
62 Scale; MSKCC=Memorial Sloan Kettering Cancer Centre; CRQ=Cancer Rehabilitation  
63 Questionnaire; HLQ=Herdecke Quality of Life questionnaire; QLACS=Quality of Life in  
64 Adult Cancer Survivors; MACFS= Modified Ambulatory Care Flow Sheet; Mini-MAC  
65 Scale= Mini-Mental Adjustment to Cancer Scale; QLICP-GM=Quality of Life Instruments  
66 for Cancer Patients-General Module; QLICP-GM=Quality of Life Instruments for Cancer  
67 Patients-General Module; QLICP-CR=Quality of Life Instruments for Cancer Patients-  
68 Colorectal Cancer; EQ-5D=EuroQol-5 dimension

69

70

71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89

**What is new?**

**Key finding:**

- The EORTC QLQ-CR38 had the greatest number of positive ratings and the most positive ratings on measurement property according to quality assessment criteria
- The EORTC QLQ-CR38 was recommended to measure HRQOL in patients with CRC, regardless of disease stage and location of primary tumour site.

**What this adds to what was known:**

- There is a shortage of a perfect methodological quality for measurement property of HRQOL instrument used in CRC patients, despite large amount of instruments available.

**What is the implication, what should change now:**

- There is a need for an improvement in the reporting quality of measurement properties in newly developed or translated instruments. Efforts on the universal consensus on whether the measurement properties met the acceptable quality criteria, especially construct validity and responsiveness, should be commenced.

90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114

**Manuscript Text**

**Introduction**

Colorectal cancer (CRC) is one of the major cancer deaths worldwide [1], being the third leading cause of cancer death in the US [2, 3] Previous studies suggested that there is an growing trend of colorectal cancer incidence rates in some economically developing countries that involves a rapid transition in dietary style and exercise patterns [2, 3]. Emerging medical treatment of CRC has contributed to the extension on prolonged survivals. In addition to disease survivals, health-related quality of life (HRQOL) is becoming standard outcome measurement of the impact of illness and treatment for CRC in clinical care and research. Among cancer survivors living with disease, certain aspects of HRQOL (i.e. physical, psychological, social, etc.) were challenged by the impairment in the ability to perform daily activities and the presentation of symptoms arising from disease and treatment[4]. The appropriate usage and adoption of instruments to evaluate HRQOL were considered important for the better assessment of rehabilitation needs and treatment benefits from the standpoint of patients, clinicians and health policy makers[5].

Classical literature of HRQOL has classified instruments [6, 7] into two major types: generic and condition-specific (or disease-specific) instruments. Generic instruments are designed for assessing HRQOL in a wide range of clinical settings and utilized in various areas of health conditions in a population. They allowed for cross-cultural and international comparisons of HRQOL from one population to another. Condition-specific instruments measure more aspects unique to the HRQOL in patients with specific condition, and have generally been reported to be more responsive than generic instruments [8]. Measurement

115 properties are important to support the evidence-based selection of the best instrument for a  
116 particular purpose or population, and quality assurance criteria[9] have been recommended  
117 by Scientific Advisory Committee of the Medical Outcomes Trust[10]. Comprehensive  
118 quality assessment of measurement properties for available HRQOL instruments is presented  
119 based on a wide range of cancer populations[11-13] but there is little evidence concerning  
120 HRQOL instruments in colorectal cancer (CRC) [14]. In spite of multiple instruments had  
121 developed for the measurement of HRQOL in patients with CRC, clinicians and researchers  
122 urged for the most recommended and appropriate HRQOL instruments under a  
123 comprehensive process of psychometric quality assessment. A recent literature review[15]  
124 across a wide range of HRQOL instruments administered in CRC patients, sourcing from  
125 Ovid searching engine and several key journals, recommended SF-12 for generic measure,  
126 EQ-5D for preference-based measure and European Organization for Research and Treatment  
127 of Cancer Core Quality-of-Life Questionnaire (QLQ-C30) and Functional Assessment of  
128 Cancer Therapy-Colorectal (FACT-C) for condition-specific measures. The aforementioned  
129 review was, however, limited by the inclusion of studies with English-speaking populations  
130 only. Papers concerning the target population of other language speakers were still relevant.  
131 Yet, no systematic review synthesized evidence on the critical appraisal of the measurement  
132 properties of generic and condition-specific HRQOL instruments that have been validated for  
133 use in patients with CRC. The aim of this paper was to conduct a systematic literature review  
134 on the measurement properties of standardized HRQOL instruments for CRC, providing  
135 recommendations on the HRQOL instrument through collective evidence from previous  
136 studies.

137

## 138 **Methods**

139

140 Literature Search Methods

141

142 *Search Engines and Strategies*

143

144 A series of systematic literature search was conducted in databases of PubMed, Web  
145 of Science using Web of Knowledge platform, Embase and MEDLINE using OVID  
146 searching platform, to identify studies that investigated the HRQOL of colorectal neoplasm  
147 patients. The Medical Subject Heading (MESH) ‘quality of life’ was combined with ‘colon  
148 neoplasm’, ‘colon cancer’, ‘rectal cancer’, ‘rectal neoplasm’ and ‘colorectal cancer’. Studies  
149 were limited to English language, and the years between January 1985 and May 2014. The  
150 earliest year was chosen as 1985 because the conceptual framework of health-related quality  
151 of life emerged around mid of 1980s[16]. Systematic searches were conducted in May 2014  
152 with electronic search strategies shown in Appendix 1. No additional hand search was done.  
153 After the initial check for duplicated articles, the abstracts of remaining articles were  
154 screened to rule out the introductions, editorials, letters, commentaries, study protocols, case  
155 reports, pure literature reviews and meta-analyses, conference `proceedings, past and current  
156 clinical guidelines and recommendations that were not recognized as original articles.  
157 Articles were also excluded if no abstract available.

158

159 *Inclusion and Exclusion Criteria*

160

161 After the review of the full-text of screened articles, the eligibility criteria of studies  
162 were 1) to involve original articles 2) to use standardized HRQOL instruments with items  
163 rating on point Likert scales or on linear analogue scales, 3) to carry out in human subjects  
164 and 4) to evaluate the measurement properties of HRQOL instruments in a mixture of CRC



165 patients (i.e. CRC with other types of cancer). Articles without available full-text were  
166 excluded. Two reviewers (CW and JC) independently screened the eligibility criteria of study  
167 titles, abstracts, selected full-texts, and reference lists of the studies retrieved by the literature  
168 search. To standardize the appraisal criteria amongst reviewers, the methodological quality of  
169 the included studies was assessed using the Consensus-based Standards for the selection of  
170 health Measurement Instruments (COSMIN) checklist [14, 17, 18] which was previously  
171 adopted for the evaluation of oncological HRQOL instruments[11]. Assessment of the  
172 methodological quality per property was performed by two reviewers independently.  
173 Disagreements regarding the procedures of database search, study selection and eligibility  
174 were resolved by discussion.

175

#### 176 *Quality Criteria of Measurement Properties*

177

178 According to the international consensus COSMIN taxonomy and definitions[17], the  
179 review evaluated nine measurement properties including: 1) internal consistency, 2)  
180 reliability, 3) measurement error, 4) content validity, 5) structural validity, 6) hypothesis  
181 testing, 7) cross-cultural validity, 8) criterion validity, and 9) responsiveness. The first three  
182 measurement properties are in the subset of reliability category, whereas the subsequent six  
183 measure properties are contained in the validity category. Internal consistency was supported  
184 if the Cronbach's alpha was equal to or greater than 0.70 and the factor analysis was  
185 conducted with adequate sample size for the support of unidimensionality of the scales.  
186 Reliability was supported if the test-retest reliability coefficient represented by intraclass  
187 correlation coefficient (ICC) and weighted kappa was equal to or greater than 0.70 between  
188 two administrations over short period of time among subjects with stable health condition.

189 Measurement error was considered adequate if the smallest detectable change was smaller  
190 than the minimal important change, or if the minimal important change reached the limits of  
191 agreement. Content validity was supported if the target population considers all items in the  
192 questionnaire to be relevant and the questionnaire to be complete. Structural validity was  
193 considered to be adequate if the factors explained at least 50% of the total variance.  
194 Hypothesis testing was assessed by testing a priori hypotheses specific to the expected  
195 correlations between scores representing similar concepts or expected differences in scores  
196 between known groups. Cross-cultural validity was supported if the original factor structure  
197 was confirmed or there was no important differential item functioning between language  
198 versions. Criterion validity was considered to be present if the gold standard for HRQOL  
199 measure existed as the full-length version and tested for the correlations with the shortened  
200 version of the instrument. Responsiveness was examined using different statistics to detect  
201 important changes over time.

202

203 *Data Synthesis on Methodological Quality Evaluation*

204

205 For each study, each measurement property was rated as ‘adequate’ (+, positive sign)  
206 or ‘not adequate’ (-, negative sign) if the quality criterion was met or was not met for each  
207 measurement property respectively. If the information given to the measurement property  
208 was unclear or ambiguous, it was rated as ‘doubtful’ (?). Given no information was found on  
209 that measurement property, zero (0) rating was assigned to that quality assessment. The  
210 measurement properties of HRQOL instruments were evaluated based on the explicit quality  
211 criteria proposed by Terwee et al.[9, 19]. A summary of the quality criteria for measurement  
212 properties of HRQOL instruments is presented in the Appendix 2.

213

214           Furthermore, the quality criteria of each measurement property were accompanied  
215 with the level of evidence scoring on a 4-point Likert scale in an ascending order of ‘poor’,  
216 ‘fair’, ‘good’ or ‘excellent’. To consolidate the grading of measurement properties of multiple  
217 instruments, the overall rating for a measurement property was synthesized by taking the  
218 quality ratings of each measurement property, consistency of results between studies, and its  
219 evidence level for measurement properties. One of the five possible rating options  
220 representing ‘unknown’ (?), ‘conflicting’ (+/-), ‘limited’ (+ or -), ‘moderate’ (++) or (--) or  
221 ‘strong’ (+++ or ---) were assigned if the measurement property of instrument was graded at  
222 least one. Rating summary of the overall levels of evidence for the quality of each  
223 measurement property is displayed in Appendix 3.

224

## 225 **Results**

226           Figure 1 shows the process of literature identification, screening for eligibility and  
227 selection of studies during the literature search presented in a Preferred Reporting Items for  
228 Systematic Reviews and Meta-Analyses (PRISMA) flow diagram [20]. The literature search  
229 was completed in June 2014 and identified a total of 7553 potentially relevant studies  
230 (PubMed: 1349; Web of Science: 2318; MEDLINE: 1735; Embase: 2151) that met the  
231 searching criteria in four bibliographic databases. After the removal of duplicated (n=3332)  
232 and non-original articles (n=1439) by abstract screening, the abstract content of 2782 studies  
233 were reviewed for eligibility. The full-text articles of all (n=65) eligible studies were  
234 reviewed. The earliest study that assessed measurement properties of HRQOL instruments  
235 relevant to CRC patients was published in 1993. We found 63 studies which investigated the  
236 measurement properties of HRQOL instruments in CRC patients or in a variety of cancer  
237 patients including CRC patients.

238

239           This review identified 3 generic and 34 condition-specific instruments: six  
240 instruments in EORTC module, seven instruments in FACIT module and 21 instruments in  
241 other modules. All identified instruments were evaluated by whether the psychometric  
242 properties met the quality criteria. The names of instruments in each module are shown in  
243 Table 1. The Functional Assessment of Cancer Therapy-Colorectal (FACT-C) was the most  
244 evaluated HRQOL instrument. The second most evaluated instrument was QLQ-CR38[21-27]  
245 which was evaluated by seven studies. General characteristics of the HRQOL instruments  
246 and the respective measurement properties evaluated are summarized in Appendix 4. Some of  
247 HRQOL instruments evaluated was designed for CRC patients exclusively (14/37, 37.8%)  
248 whereas a majority of them was designed for a wide range of cancer patients or patients with  
249 colostomy or ileostomy (23/37, 62.1%). The total number of items varied from five to 47  
250 whilst the number of subscales or domains varied from one (unidimensional) to 12. The  
251 response options were predominantly a 4-point or 5-point Likert scale.

252

### 253 *Characteristics of included instruments*

254           Characteristics of the 63 studies included in this review were reported in Appendix 5.  
255 Forty-one studies were originated from countries located in North America and Europe [5, 7,  
256 23-25, 27-54] and the rest of them were conducted in Asian countries. Most of the  
257 instruments were evaluated in the language versions of English, French and German. Four  
258 studies reported the measurement properties of instruments were translated to more than one  
259 language versions and evaluated in CRC patients recruited from the respective countries [33,  
260 46, 50, 54].

261

262           Thirty-five studies focused on CRC patients solely, ranging from patients with liver  
263 metastatic CRC to patients undergoing surgery [21-27, 33-35, 39, 42, 45, 46, 50, 52, 54-62],  
264 16 studies assessed patients with CRC and other diseases, including breast, ovarian or  
265 prostate cancer [28-30, 43, 44, 51, 59, 63-67]. The remaining nine studies were mostly  
266 targeted at patients suffering from rectal cancer, ostomy or unspecified or mixed types of  
267 cancer [7, 31, 36-38, 40, 41, 47, 49, 53, 68-71].

268

269           The stage of diseases was also different in the selected studies. Twenty-five of the  
270 studies included patients from stage one to four according to American Joint Committee on  
271 Cancer classification system[23, 24, 27, 28, 30, 32-34, 42, 44, 45, 51, 52, 54, 55, 57, 58, 65-  
272 67] and patients in these studies were largely over 60 years old.

273

#### 274 *Methodological Quality of Each Study*

275

276           Each eligible study was assessed on nine measurement properties based on the  
277 aforementioned criteria on a 4-point Likert Scale, ranging from 'excellent' to 'poor'. Details  
278 of the measurement properties of each study are summarized in Table 2. The three most  
279 frequently reported properties were hypothesis testing (58 studies reported), structural  
280 validity (56 studies reported) and internal consistency (49 studies reported), whilst  
281 measurement error (1 study reported), criterion validity (8 studies reported) and cross-cultural  
282 validity (11 studies reported) were the three least addressed ones. The best performed area  
283 was content validity with 10 out of 12 articles reported achieved excellent quality. However,

284 more than half of the articles reported internal consistency (55.1%), structural validity  
285 (64.3%), cross-cultural validity (72.7%) and criterion validity (62.5%) were found to have  
286 poor methodological quality.

287

288 *Overall quality of measurement properties*

289

290 Table 3 summarizes the overall levels of evidence per measurement property and  
291 HRQOL instrument. None of the HRQOL instruments has been evaluated and rated on all the  
292 nine measurement properties recommended by the quality assessment criteria. Most of the  
293 instruments provided information on internal consistency, structural validity, and hypothesis  
294 testing. Measurement error, cross-cultural validity, and criterion validity were only assessed  
295 in a minority of instruments as limited information was found in these studies. Two  
296 instruments, QLQ-CR38 [21-27] and FACT-C [23, 27, 33, 34, 55, 59-61] were the most  
297 evaluated instruments with each of them had eight reviewed measurement properties. Among  
298 all the assessed items, FACT-G (Version 2) [7], FACT-G (Version 3) [29] and mCOH-QOL-  
299 Ostomy [37-39] were the instruments scored limited to strong positive evidence on all their  
300 measured properties without any negative or uncertain rating. Seven evaluated instruments  
301 were found to have at least one negative rating in any one of their reviewed measurement  
302 properties (Version 3 of QLQ-C30 [27, 29, 30, 65, 66], Coping with CRC [58], Quality-  
303 Quantity Questionnaire [43], Mini-MAC Scale [49], SF-12 (Version 2) [60], EQ-5D [72] and  
304 SF-6D [73]).

305 Internal Consistency

306 Thirty-one instruments were assessed on this property. QLQ-C30 (Version 3)[27, 29,  
307 30, 65, 66], FACT-G (Version 3) [5, 29], Quick-FLIC[63] and Modified FIQL [70] were  
308 rated as strong positive and none of the instruments were rated negative.

#### 309 Reliability

310 Seventeen instruments were assessed on this property where seven were rated as  
311 limited positive (Version 3 of QLQ-C30 [5, 27, 30, 65, 66], Version 2 of FACT-G [7],  
312 Version 4 of FACT-G [23, 27, 69], FACT-C [23, 27, 33, 34, 55, 59-61], QLICP-GM [67],  
313 QLICP-CR [62] and MSKCC Bowel Function [40, 41]). Quick-FLIC [63] was the only  
314 instrument that was rated as strong positive by the criteria.

#### 315 Measurement Error

316 Only one instrument (FACT-C) [23, 27, 33, 34, 55, 59-61] was assessed on this  
317 property and rated as limited positive.

#### 318 Content Validity

319 Twelve instruments were assessed on this property and ten of them (QLQ-CR38 [21-  
320 27], QLQ-LMC21 [46, 50], QLQ-SWB36[53], Version 2 of FACT-G [32], FACT-C [23, 27,  
321 33, 34, 55, 59-61], QLICP-GM [67], mCOH-QOL-Ostomy [37-39], MSKCC Bowel  
322 Function[40, 41], Coping with CRC [58], QLACS [44]) were rated as strong positive. Two  
323 instruments (CRQ [42] and QLICP-CR [62]) were rated as unknown.

#### 324 Structural Validity

325 Thirty-two instruments were assessed on this property but only nine were rated as  
326 limited to moderate positive (QLQ-CR38 [21-27], Version 2[32] and Version 3[29] of  
327 FACT-G, mCOH-QOL-Ostomy [37-39], SRQS [57], MSKCC Bowel Function [40, 41],

328 Coping with CRC [58], HLQ [51], QLACS [44]). The rest of them were rated as unknown,  
329 and one item (Quality-Quantity Questionnaire [43]) was rated as limited negative.

### 330 Hypothesis Testing

331 Thirty-two instruments were rated on this property with eleven rated as strong  
332 positive (Version 1 of QLQ-C30 [28, 64], Version 3 of QLQ-C30 [5, 27, 30, 65, 66], QLQ-  
333 CR38 [21-27], QLQ-LMC21 [46, 50], Version 4 of FACT-G [23, 27, 69], FACT-C [23, 27,  
334 33, 34, 55, 59-61], FCSI-9 [35], FACIT-F [68], QLICP-CR [62], MSKCC Bowel  
335 Function[40, 41] and EQ-5D [72]).

### 336 Cross-cultural validity

337 All ratings given to the nine instruments which were assessed on this property  
338 (Version 3 of QLQ-C30 [27, 29, 30, 65, 66], QLQ-SWB36 [53], QLQ-CR38 [21-27], QLQ-  
339 CR29 [31, 54], FACT-C [23, 27, 33, 34, 55, 59-61], MSAS-SF [56], CMSAS [56], SRQS  
340 [57], MSKCC Bowel Function [40, 41]) were unknown as inadequate information was  
341 provided by these studies.

### 342 Criterion Validity

343 Eight instruments were assessed on this property with only two instruments, QLQ-  
344 LMC21 [46, 50] and FACT-G7 [71] were rated as moderate positive. The rest of them were  
345 rated as unknown due to unclear methodological quality.

### 346 Responsiveness

347 Fourteen instruments were assessed on this property and only one instrument was  
348 rated strong positive (FCSI-9 [35]), whilst three instruments were rated as moderate positive  
349 (QLQ-CR38 [21-27], FACT-G (Version 4) [23, 27, 69] and FACT-C [23, 27, 33, 34, 55, 59-



350 61]) and other three moderate negative (Version 3 of QLQ-C30 [5, 27, 30, 65, 66], SF-12  
351 (Version 2) [60] and SF-6D [73]).

352

### 353 **Discussions**

354

355 This systematic review evaluated the measurement properties of 37 standardized HRQOL  
356 instruments used in CRC patients among 63 eligible studies identified in the full-text  
357 assessment stage. None of the instruments were adequately evaluated for all nine  
358 measurement properties recommended by COSMIN[17]. Compared to other instruments,  
359 EORTC QLQ-CR38 and FACT-C were more comprehensively evaluated as each of them  
360 obtained eight ratings in all of the nine measurement properties. Moreover, the number of  
361 positive sign as an indication of satisfactory performance of measurement property was  
362 obtained to compare the overall performance of the 37 instruments. QLQ-CR38 in EORTC  
363 module received the highest number ('+': 14) of positive signs among all the instruments,  
364 followed by MSKCC Bowel Function ('+': 11), FACT-C ('+': 10) and Quick-FLIC ('+': 10).  
365 MSKCC bowel function was limited for the designed for the measurement of HRQOL in  
366 rectal cancer. While the QLQ-CR38 and FACT-C were the most assessed instruments in our  
367 review, QLQ-CR38 seized more positive ratings and had a better rating in internal  
368 consistency and structural validity than the FACT-C. Therefore, current review suggested  
369 that QLQ-CR38 was the most comprehensive and positive rated instrument in this review,  
370 with positive ratings in internal consistency, reliability, content validity, structural validity,  
371 hypothesis testing and responsiveness. Despite the QLQ-CR38 was not assessed on  
372 measurement error and rated unknown on both cross-cultural validity and criterion validity,  
373 these three properties were either not evaluated or rated as unknown for other instruments.  
374 Aside from the highest number of positive sign, the measurement properties of QLQ-CR38

375 were evaluated by seven studies, greater than two studies for MSKCC bowel function  
376 instrument and one study for Quick-FLIC. Though FACT-G (Version 2) was rated as positive  
377 on the same properties rated for the QLQ-CR38, the ratings were generally poorer than those  
378 for the latter. The QLQ-CR38 was therefore recommended regardless of disease stage and  
379 location of primary tumour site because of its positive performance on most of the assessed  
380 properties.

381

382 In line with another structured review evaluated for CRC patients in the UK[15], the  
383 measurement properties of generic instruments among CRC patients was identified in current  
384 systematic review. The methodological evaluation of that structured review[15]  
385 recommended the use of two generic instruments, EQ-5D and SF-12, and two condition-  
386 specific instruments, QLQ-C30 and FACT-C, based on eight categories of appraisal criteria  
387 that were used in earlier studies[13]. Unlike current systematic review, not all studies selected  
388 in previous review were aimed and dedicated to demonstrate the measurement properties of  
389 HRQOL instruments. It was worthwhile noting that the study selection criteria and the  
390 appraisal criteria in current review differed from with that in previous review. For instance, in  
391 case of methodological quality of responsiveness property of HRQOL instruments, COSMIN  
392 checklist focused on the detailed description of gold standard or comparator instrument that  
393 indicated the responsiveness property whereas the appraisal criteria applied in previous  
394 review focused on the statistically significant changes in HRQOL scores over time.

395

396 The EORTC and FACIT groups of instruments shared the same evolvement of a few  
397 versions through long-lasting transitions and variations in the item wordings, item responses,  
398 scale structure, and scoring algorithm. In EORTC modules based on their group website, the

399 QLQ-CR38 and version 1 of QLQ-C30 was superseded by the shortened QLQ-CR29 and  
400 version 3 of QLQ-C30, respectively. The QLQ-CR38 and version 3 of QLQ-C30 were rated  
401 better HRQOL instrument than the QLQ-CR29 and superseded version of QLQ-C30,  
402 possibly due to the substantive amount of measurement property evaluation for QLQ-CR38  
403 and version 3 of QLQ-C30 over the past two decades. Nevertheless, the currently available  
404 version of QLQ-CR29 was not rated more adequate than QLQ-CR38, in part explained by the  
405 fact that the first publication year of latter one was in 1999, at least one decade earlier than  
406 the end of search date in current review. Possible explanation was likely to be extrapolated to  
407 the FACIT modules, in the situation when the version 2 and 3 of FACT-G were superseded  
408 by its version 4 which was firstly introduced in year 2003[74]. The instruments deserved  
409 significant advantages on their overall levels of evidence when the instruments have been  
410 developed at least ten years up to the search date in current review. Among the 39 eligible  
411 studies, the HRQOL instruments developed by EORTC and FACIT modules were directly  
412 compared and contrasted in three studies in France [23, 27, 29]. Aside from the one study [27]  
413 limited by small sample size ( $n < 30$ ), the rest of them reported the acceptability and  
414 preference of instruments in two groups in addition to the measurement properties of  
415 instruments. CRC patients elicited patient preference for core HRQOL instruments of two  
416 modules, and preferred the QLQ-C30 over the FACT-G and FLIC instruments in part  
417 reflecting their response burden and completion time (7.9 vs 10.2 vs 8.4 minutes,  
418 respectively)[29]. Likewise, in light of comparing the patient preference for the QLQ-CR38  
419 and FACT-C instruments, no preference difference between QLQ-C30 in supplement with  
420 QLQ-CR38 and FACT-C was observed[23].

421

422 It is noteworthy to point out that included studies of measurement property evaluation  
423 had two major shortcomings of quality criteria reporting. Firstly, the CRQ and Quality-

424 Quantity Questionnaire [42, 43] did not present a priori hypothesis for testing the construct  
425 validity although construct validity has been introduced in the included studies. Secondly,  
426 data on responsiveness were lacking. Out of 22 studies that had investigated responsiveness,  
427 seven studies were rated as ‘poor’ in methodological quality of responsiveness property [21,  
428 45, 46, 54, 62, 66, 67] in part due to the controversy over whether the mean change exceeded  
429 the minimum clinically important difference or discriminated between groups defined by  
430 anchor using area under the curve. Besides, the methodological quality were rated as ‘poor’  
431 in the evaluation of properties of internal consistency, structural validity and cross-cultural  
432 validity with a lack of details on the factor analysis or item response theory test employed.  
433 Further studies investigating measurement properties of instruments are necessary to  
434 adequately report clear priori hypothesis for construct validity and statistics for  
435 responsiveness. Standard reporting of factor analysis is also recommended to provide  
436 comprehensive information for researchers to review the reliability and validity of the  
437 instruments.

438

#### 439 *Limitations*

440 A plausible limitation of this review aroused from the process of literature search and  
441 selection of studies. Abstracts from conference were also precluded in this review which  
442 implies that not all the CRC studies that had used HRQOL were included. Besides, several  
443 instruments (e.g. FCSI-9 [35]) were only used in one or a few studies, which may not be  
444 supportive enough to provide reliable evidence to the instruments being evaluated. Further  
445 studies are recommended to ensure the evidence comparing different HRQOL instruments is  
446 the most updated and comprehensive. The second limitation is inadequate coverage of  
447 measurement properties suggested by COSMIN checklist. Other uncovered measurement

448 criteria, including floor/ceiling effects, predictive validity, and interpretability were useful in  
449 assessing the HRQOL of CRC patients. Alternative checklists for general cancer patients[12,  
450 13, 15] were helpful to provide evidence of quality criteria in HRQOL instruments. However,  
451 the adoption of COSMIN checklist, the standardized and structured guideline, would  
452 facilitate a head-to-head comparison in the measurement properties evidence of HRQOL  
453 instruments between CRC patients and other disease populations. Last but not least, some of  
454 the studies[28-30, 32, 36-38, 44, 51, 64-67, 69] involved in this systematic review were based  
455 on patients with mixed types of cancer, but not limited to CRC only. Studies that involved  
456 CRC patients also did not report sufficient measurement properties. In Nicholas et al.'s study  
457 [42], the level of evidence for nine measurement properties was unknown due to flawed  
458 methodological quality and the lack in reporting. This may confound the results of this  
459 review that targeted on evaluating HRQOL of CRC patients.

460

## 461 **Conclusions**

462

463 This systematic review draws out an attention to the shortage of a perfect  
464 methodological quality for measurement property of HRQOL instrument used in CRC  
465 patients, despite large amount of instruments available. There is a need for an improvement in  
466 the reporting quality of measurement properties in newly developed or translated instruments.  
467 Efforts on the universal consensus on whether the measurement properties met the acceptable  
468 quality criteria, especially construct validity and responsiveness, should be commenced.  
469 Concerning our review of HRQOL instruments, the EORTC QLQ-CR38 had the greatest  
470 number of positive ratings according to quality assessment criteria. Concerning with the  
471 number of measurement properties with positive ratings, the EORTC QLQ-CR38 was

472 recommended to measure HRQOL in patients with CRC, regardless of disease stage and  
473 primary tumour site.

474

475 **Acknowledgements: None**

476

477 **Funding:** This work was supported by Health and Health Services Research Fund (project  
478 number 08090851) from the Food and Health Bureau of the Hong Kong SAR.

479

480 **Disclosure: The authors have declared no conflicts of interest.**

481

482 **Reference**

- 483 [1] Ferlay J, Shin H-R, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of  
484 cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127:2893-917.
- 485 [2] Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. *Cancer*  
486 *Epidemiology Biomarkers & Prevention*. 2009;18:1688-94.
- 487 [3] Hagggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk  
488 factors. *Clinics in colon and rectal surgery*. 2009;22:191.
- 489 [4] Ramsey SD, Andersen MR, Etzioni R, Moinpour C, Peacock S, Potosky A, et al. Quality of life in  
490 survivors of colorectal carcinoma. *Cancer*. 2000;88:1294-303.
- 491 [5] Conroy T, Bleiberg H, Glimelius B. Quality of life in patients with advanced colorectal cancer:  
492 what has been learnt? *European Journal of Cancer*. 2003;39:287-94.
- 493 [6] Patrick DL, Deyo RA. Generic and Disease-Specific Measures in Assessing Health Status and  
494 Quality of Life. *Med Care*. 1989;27:S217-S32.
- 495 [7] Cella DF, Tulsky DS. Quality of Life in Cancer: Definition, Purpose, and Method of Measurement.  
496 *Cancer Investigation*. 1993;11:327-36.
- 497 [8] Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic  
498 and specific quality-of-life instruments. *J Clin Epidemiol*. 2003;56:52-60.
- 499 [9] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria  
500 were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*.  
501 2007;60:34-42.
- 502 [10] Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and  
503 quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*. 2002;11:193-205.
- 504 [11] Lockett T, King MT, Butow PN, Oguchi M, Rankin N, Price MA, et al. Choosing between the  
505 EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical  
506 research: issues, evidence and recommendations. *Annals of Oncology*. 2011;22:2179-90.
- 507 [12] Pearce NJM, Sanson-Fisher R, Campbell HS. Measuring quality of life in cancer survivors: a  
508 methodological review of existing scales. *Psycho-Oncology*. 2008;17:629-40.
- 509 [13] Fitzsimmons D, Gilbert J, Howse F, Young T, Arrarras J-I, Brédart A, et al. A systematic review  
510 of the use and validation of health-related quality of life instruments in older cancer patients.  
511 *European Journal of Cancer*. 2009;45:19-32.
- 512 [14] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN  
513 checklist for assessing the methodological quality of studies on measurement properties of health  
514 status measurement instruments: an international Delphi study. *Quality of Life Research* 2010;19:539-  
515 49.
- 516 [15] Patient-reported Outcome Measurement Group, Department of Public Health, University of  
517 Oxford. A Structured Review of Patients-Reported Outcome Measures for Colorectal Cancer. 2010.
- 518 [16] Aaronson NK, Meyerowitz BE, Bard M, Bloom JR, Fawzy FI, Feldstein M, et al. Quality of life  
519 research in oncology. Past achievements and future priorities. *Cancer*. 1991;67:839-43.
- 520 [17] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN  
521 study reached international consensus on taxonomy, terminology, and definitions of measurement  
522 properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737-45.
- 523 [18] Mokkink L, Terwee C, Knol D, Stratford P, Alonso J, Patrick D, et al. The COSMIN checklist  
524 for evaluating the methodological quality of studies on measurement properties: A clarification of its  
525 content. *BMC Medical Research Methodology*. 2010;10:22.
- 526 [19] Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the  
527 methodological quality in systematic reviews of studies on measurement properties: a scoring system  
528 for the COSMIN checklist. *Quality of Life Research* 2012;21:651-7.
- 529 [20] Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for  
530 Systematic Reviews and Meta-Analyses: The PRISMA Statement. *J Clin Epidemiol*. 2009;62:1006-  
531 12.
- 532 [21] Kong D, Yang Z, Wang Y, Meng Q, Tang X, Cun Y, et al. Development and validation of a  
533 simplified chinese version of EORTC QLQ-CR38 to measure the quality of life of patients with  
534 colorectal cancer. *Oncology*. 2012;83:201-9.

- 535 [22] Law CC, Tak Lam WW, Fu YT, Wong KH, Sprangers MAG, Fielding R. Validation of the  
536 Chinese version of the EORTC colorectal cancer-specific quality-of-life questionnaire module (QLQ-  
537 CR38). *Journal of Pain & Symptom Management*. 2008;35:203-13.
- 538 [23] Rotonda C, Conroy T, Mercier M, Bonnetain F, Uwer L, Miny J, et al. Validation of the French  
539 version of the colorectal-specific quality-of-life questionnaires EORTC QLQ-CR38 and FACT-C.  
540 *Quality of Life Research*. 2008;17:437-45.
- 541 [24] Sprangers MA, te Velde A, Aaronson NK. The construction and testing of the EORTC colorectal  
542 cancer-specific quality of life questionnaire module (QLQ-CR38). *European Organization for  
543 Research and Treatment of Cancer Study Group on Quality of Life*. *European Journal of Cancer*.  
544 1999;35:238-47.
- 545 [25] Thaysen HV, Jess P, Laurberg S, Groenvold M. Validation of the Danish version of the disease  
546 specific instrument EORTC QLQ-CR38 to assess health-related quality of life in patients with  
547 colorectal cancer. *Health & Quality of Life Outcomes*. 2012;10:150.
- 548 [26] Tsunoda A, Yasuda N, Nakao K, Yokoyama N, Marumori T, Hashimoto H, et al. Validation of  
549 the Japanese version of EORTC QLQ-CR38. *Quality of Life Research*. 2008;17:317-22.
- 550 [27] Uwer L, Rotonda C, Guillemin F, Miny J, Kaminsky M-C, Mercier M, et al. Responsiveness of  
551 EORTC QLQ-C30, QLQ-CR38 and FACT-C quality of life questionnaires in patients with colorectal  
552 cancer. *Health & Quality of Life Outcomes*. 2011;9:70.
- 553 [28] Apolone G, Filiberti A, Cifani S, Ruggiata R, Mosconi P. Evaluation of the EORTC QLQ-C30  
554 questionnaire: a comparison with SF-36 Health Survey in a cohort of Italian long-survival cancer  
555 patients. *Annals of Oncology*. 1998;9:549-57.
- 556 [29] Conroy T, Mercier M, Bonnetterre J, Luporsi E, Lefebvre JL, Lapeyre M, et al. French version of  
557 FACT-G: validation and comparison with other cancer-specific instruments. *European Journal of  
558 Cancer*. 2004;40:2243-52.
- 559 [30] Koukouli S, Stamou A, Alegakis A, Georgoulas V, Samonis G. Psychometric properties of the  
560 QLQ-C30 (version 3.0) in a sample of ambulatory Cretan cancer patients. *European Journal of Cancer  
561 Care*. 2009;18:447-56.
- 562 [31] Arraras JI, Suarez J, Arias de la Vega F, Vera R, Asin G, Arrazubi V, et al. The EORTC Quality  
563 of Life questionnaire for patients with colorectal cancer: EORTC QLQ-CR29 validation study for  
564 Spanish patients. *Clinical & Translational Oncology*. 2011;13:50-6.
- 565 [32] Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment  
566 of Cancer Therapy scale: development and validation of the general measure. *Journal of Clinical  
567 Oncology*. 1993;11:570-9.
- 568 [33] Ward WL, Hahn EA, Mo F, Hernandez L, Tulsky DS, Cella D. Reliability and validity of the  
569 Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrument. *Quality of  
570 Life Research*. 1999;8:181-95.
- 571 [34] Yost KJ, Cella D, Chawla A, Holmgren E, Eton DT, Ayanian JZ, et al. Minimally important  
572 differences were estimated for the Functional Assessment of Cancer Therapy-Colorectal (FACT-C)  
573 instrument using a combination of distribution- and anchor-based approaches. *J Clin Epidemiol*.  
574 2005;58:1241-51.
- 575 [35] Colwell HH, Mathias SD, Turner MP, Lu J, Wright N, Peeters M, et al. Psychometric evaluation  
576 of the FACT Colorectal Cancer Symptom Index (FCSI-9): reliability, validity, responsiveness, and  
577 clinical meaningfulness. *Oncologist*. 2010;15:308-16.
- 578 [36] Stein KD, Denniston M, Baker F, Dent M, Hann DM, Bushhouse S, et al. Validation of a  
579 modified Rotterdam Symptom Checklist for use with cancer patients in the united states. *Journal of  
580 Pain and Symptom Management*. 2003;26:975-89.
- 581 [37] Grant M, Ferrell B, Dean G, Uman G, Chu D, Krouse R. Revision and psychometric testing of  
582 the City of Hope Quality of Life Ostomy Questionnaire. *Quality of Life Research*. 2004;13:1445-57.
- 583 [38] Krouse RS, Mohler MJ, Wendel CS, Grant M, Baldwin CM, Rawl SM, et al. The VA ostomy  
584 health-related quality of life study: objectives, methods, and patient sample. *Current Medical  
585 Research and Opinion*. 2006;22:781-91.
- 586 [39] Mohler MJ, Coons SJ, Hornbrook MC, Herrinton LJ, Wendel CS, Grant M, et al. The health-  
587 related quality of life in long-term colorectal cancer survivors study: objectives, methods and patient  
588 sample. *Current Medical Research & Opinion*. 2008;24:2059-70.



- 589 [40] Temple LK, Bacik J, Savatta SG, Gottesman L, Paty PB, Weiser MR, et al. The development of a  
590 validated instrument to evaluate bowel function after sphincter-preserving surgery for rectal cancer.  
591 *Diseases of the Colon & Rectum*. 2005;48:1353-65.
- 592 [41] Zotti P, Del Bianco P, Serpentine S, Trevisanut P, Barba MC, Valentini V, et al. Validity and  
593 reliability of the MSKCC Bowel Function instrument in a sample of Italian rectal cancer patients.  
594 *European Journal of Surgical Oncology*. 2011;37:589-96.
- 595 [42] Nicholas PK, Leuner JDM, Hatfield JM, Corless IB, Marr KH, Mott MK, et al. Using the Cancer  
596 Rehabilitation Questionnaire in patients with colorectal cancer. *Rehabilitation Nursing*. 2006;31:106-  
597 13.
- 598 [43] Stiggelbout AM, De Haes JCJM, Kiebert GM, Kievit J, Leer JWH. Tradeoffs between quality  
599 and quantity of life: Development of the QQ questionnaire for cancer patient attitudes. *Medical*  
600 *Decision Making*. 1996;16:184-92.
- 601 [44] Avis NE, Smith KW, McGraw S, Smith RG, Petronis VM, Carver CS. Assessing Quality of Life  
602 in Adult Cancer Survivors (QLACS). *Quality of Life Research*. 2005;14:1007-23.
- 603 [45] Whistance RN, Gilbert R, Fayers P, Longman RJ, Pullyblank A, Thomas M, et al. Assessment of  
604 body image in patients undergoing surgery for colorectal cancer. *International Journal of Colorectal*  
605 *Disease*. 2010;25:369-74.
- 606 [46] Blazeby JM, Fayers P, Conroy T, Sezer O, Ramage J, Rees M, et al. Validation of the European  
607 Organization for Research and Treatment of Cancer QLQ-LMC21 questionnaire for assessment of  
608 patient-reported outcomes during treatment of colorectal liver metastases. *British Journal of Surgery*.  
609 2009;96:291-8.
- 610 [47] Canova C, Giorato E, Roveron G, Turrini P, Zanotti R. Validation of a stoma-specific quality of  
611 life questionnaire in a sample of patients with colostomy or ileostomy. *Colorectal Disease*.  
612 2013;15:e692-e8.
- 613 [48] Horsman S, Olson K, Au H, Ghosh S. Symptom assessment in ambulatory oncology: initial  
614 validation of the nurse-developed Modified Ambulatory Care Flow Sheet (MACFS). *Quality of Life*  
615 *Research*. 2012;21:899-908.
- 616 [49] Hulbert-Williams NJ, Hulbert-Williams L, Morrison V, Neal RD, Wilkinson C. The mini-Mental  
617 Adjustment to Cancer Scale: Re-analysis of its psychometric properties in a sample of 160 mixed  
618 cancer patients. *Psycho-Oncology*. 2012;21:792-7.
- 619 [50] Kavadas V, Blazeby JM, Conroy T, Sezer O, Holzner B, Koller M, et al. Development of an  
620 EORTC disease-specific quality of life questionnaire for use in patients with liver metastases from  
621 colorectal cancer. *European Journal of Cancer*. 2003;39:1259-63.
- 622 [51] Kroz M, Bussing A, Girke M, Heckmann C, Ostermann T. Adaptation of the Herdecke Quality  
623 of Life questionnaire towards quality of life of cancer patients. *European Journal of Cancer Care*.  
624 2008;17:593-600.
- 625 [52] Schwenk W, Neudecker J, Haase O, Raue W, Strohm T, Muller JM. Comparison of EORTC  
626 quality of life core questionnaire (EORTC-QLQ-C30) and gastrointestinal quality of life index  
627 (GIQLI) in patients undergoing elective colorectal cancer resection. *International Journal of*  
628 *Colorectal Disease*. 2004;19:554-60.
- 629 [53] Vivat B, Young T, Efficace F, Siguradottir V, Arraras JI, Asgeirsdottir GH, et al. Cross-cultural  
630 development of the EORTC QLQ-SWB36: A stand-alone measure of spiritual wellbeing for palliative  
631 care patients with cancer. *Palliative Medicine*. 2013;27:457-69.
- 632 [54] Whistance RN, Conroy T, Chie W, Costantini A, Sezer O, Koller M, et al. Clinical and  
633 psychometric validation of the EORTC QLQ-CR29 questionnaire module to assess health-related  
634 quality of life in patients with colorectal cancer. *European Journal of Cancer*. 2009;45:3017-26.
- 635 [55] Yoo HJ, Kim JC, Eremenco S, Han OS. Quality of life in colorectal cancer patients with  
636 colectomy and the validation of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C),  
637 Version 4. *Journal of Pain & Symptom Management*. 2005;30:24-32.
- 638 [56] Lam WWT, Law CC, Fu YT, Wong KH, Chang VT, Fielding R. New Insights in Symptom  
639 Assessment: The Chinese Versions of the Memorial Symptom Assessment Scale Short Form (MSAS-  
640 SF) and the Condensed MSAS (CMSAS). *Journal of Pain and Symptom Management*. 2008;36:584-  
641 95.
- 642 [57] Hou WK, Lam WWT, Chi CL, Yiu TF, Fielding R. Measuring social relational quality in  
643 colorectal cancer: The Social Relational Quality Scale (SRQS). *Psycho-Oncology*. 2009;18:1097-105.

- 644 [58] Rinaldis M, Pakenham KI, Lynch BM, Aitken JF. Development, confirmation, and validation of  
645 a measure of coping with colorectal cancer: a longitudinal investigation. *Psycho-Oncology*.  
646 2009;18:624-33.
- 647 [59] Wong CK, Lam CL, Law WL, Poon JT, Chan P, Kwong DL, et al. Validity and reliability study  
648 on traditional Chinese FACT-C in Chinese patients with colorectal neoplasm. *Journal of evaluation in*  
649 *clinical practice*. 2012;18:1186-95.
- 650 [60] Wong CK, Lam CL, Law W-L, Poon JT, Kwong DL, Tsang J, et al. Condition-specific measure  
651 was more responsive than generic measure in colorectal cancer: all but social domains. *J Clin*  
652 *Epidemiol*. 2013;66:557-65.
- 653 [61] Wong CKH, Lam CLK, Mulhern B, Law W-L, Poon JTC, Kwong DLW, et al. Measurement  
654 invariance of the Functional Assessment of Cancer Therapy-Colorectal quality-of-life instrument  
655 among modes of administration. *Quality of Life Research*. 2013;22:1415-26.
- 656 [62] Xu C, Yang Z, Tan J, Meng Q, Cun Y, Tang X, et al. Development and validation of the system  
657 of quality of life instruments for cancer patients: colorectal cancer (QLICP-CR). *Cancer Investigation*.  
658 2012;30:732-40.
- 659 [63] Cheung YB, Goh C, Wong LC, Ng GY, Lim WT, Leong SS, et al. Quick-FLIC: Validation of a  
660 short questionnaire for assessing quality of life of cancer patients. *British Journal of Cancer*.  
661 2004;90:1747-52.
- 662 [64] King MT, Dobson AJ, Harnett PR. A comparison of two quality-of-life questionnaires for cancer  
663 clinical trials: the functional living index--cancer (FLIC) and the quality of life questionnaire core  
664 module (QLQ-C30). *J Clin Epidemiol*. 1996;49:21-9.
- 665 [65] Luo N, Fones CSL, Lim SE, Xie F, Thumboo J, Li SC. The European Organization for Research  
666 and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30): Validation of English  
667 version in Singapore. *Quality of Life Research*. 2005;14:1181-6.
- 668 [66] Wan C, Meng Q, Yang Z, Tu X, Feng C, Tang X, et al. Validation of the simplified Chinese  
669 version of EORTC QLQ-C30 from the measurements of five types of inpatients with cancer. *Annals*  
670 *of Oncology*. 2008;19:2053-60.
- 671 [67] Wan C, Yang Z, Meng Q, Feng C, Wang H, Tang X, et al. Development and validation of the  
672 general module of the system of quality of life instruments for cancer patients. *Int J Cancer*.  
673 2008;122:190-6.
- 674 [68] Butt Z, Lai J-s, Rao D, Heinemann AW, Bill A, Cella D. Measurement of fatigue in cancer,  
675 stroke, and HIV using the Functional Assessment of Chronic Illness Therapy - Fatigue (FACIT-F)  
676 scale. *Journal of Psychosomatic Research*. 2013;74:64-8.
- 677 [69] Cheung YB, Goh C, Wee J, Khoo KS, Thumboo J. Measurement properties of the chinese  
678 language version of the functional assessment of cancer therapy-general in a singaporean population.  
679 *Annals of the Academy of Medicine Singapore*. 2009;38:225-9.
- 680 [70] Hashimoto H, Shiokawa H, Funahashi K, Saito N, Sawada T, Shirouzu K, et al. Development  
681 and validation of a modified fecal incontinence quality of life scale for Japanese patients after  
682 intersphincteric resection for very low rectal cancer. *Journal of Gastroenterology*. 2010;45:928-35.
- 683 [71] Yanez B, Pearman T, Lis CG, Beaumont JL, Cella D. The FACT-G7: A rapid version of the  
684 functional assessment of cancer therapy-general (FACT-G) for monitoring symptoms and concerns in  
685 oncology practice and research. *Annals of Oncology*. 2013;24:1073-8.
- 686 [72] Kim SH, Hwang JS, Kim TW, Hong YS, Jo M-W. Validity and reliability of the EQ-5D for  
687 cancer patients in Korea. *Supportive Care in Cancer*. 2012;20:3155-60.
- 688 [73] Wong CKH, Mulhern B, Wan Y-F, Lam CLK. Responsiveness was similar between direct and  
689 mapped SF-6D in colorectal cancer patients who declined. *J Clin Epidemiol*. 2014;67:219-27.
- 690 [74] Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT)  
691 Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes*.  
692 2003;1:1-7.

693

694 Table 1. Names of each evaluated HRQOL instrument

| Abbreviation                   | Full name   |
|--------------------------------|---|
| <i>EORTC</i>                   |   |
| QLQ-C30                        | Core Quality-of-Life Questionnaire  |
| QLQ-CR38                       | Colorectal Cancer-specific Quality-of-Life Questionnaire Module                       |
| QLQ-CR29                       | Colorectal Cancer-specific Quality-of-Life Questionnaire Module                       |
| QLQ-LMC21                      | Liver Metastases from Colorectal Cancer-specific Quality-of-Life Questionnaire Module |
| QLQ-SWB36                      | Spiritual Wellbeing-specific Quality-of-Life Questionnaire Module                     |
| <i>FACIT</i>                   |   |
| FACT-G                         | Functional Assessment of Cancer Therapy-General                                       |
| FACT-C                         | Functional Assessment of Cancer Therapy-Colorectal                                    |
| FCSI-9                         | Functional Assessment of Cancer Therapy Colorectal Symptom Index                      |
| FACT-G7                        | Rapid Version of the Functional Assessment of Cancer Therapy-General                  |
| FACIT-F                        | Functional Assessment of Chronic Illness Therapy-Fatigue                              |
| FLIC                           | Functional living index for cancer  |
| Quick-FLIC                     | Quick-Functional living index for cancer  |
| RSCL-M                         | Modified Version of Rotterdam Symptom Checklist                                       |
| QLICP-GM                       | Quality of Life Instruments for Cancer Patients-General Module                        |
| QLICP-CR                       | Quality of Life Instruments for Cancer Patients-Colorectal Cancer                     |
| GIQLI                          | Gastrointestinal Quality of Life Index  |
| MSAS-SF                        | Memorial Symptom Assessment Scale-Short Form  |
| CMSAS                          | Condensed Memorial Symptom Assessment Scale   |
| Modified FIQL                  | Modified Fecal Incontinence Quality of Life   |
| mCOH-QOL-Ostomy                | Modified City of Hope Ostomy questionnaire  |
| SRQS                           | Social Relational Quality Scale   |
| MSKCC Bowel Function           | Memorial Sloan Kettering Cancer Centre Bowel Function                                 |
| CRQ                            | Cancer Rehabilitation Questionnaire   |
| Coping with CRC                | Coping with Colorectal Cancer   |
| Quality-Quantity Questionnaire | Quality-Quantity Questionnaire  |
| HLQ                            | Herdecke Quality of Life Questionnaire  |
| QLACS                          | Quality of Life in Adult Cancer Survivors   |
| Body-image Scale               | Body-image Scale  |
| Stoma Care QoL scale           | Stoma Care Quality of life Scale  |
| MACFS                          | Modified Ambulatory Care Flow Sheet   |
| Mini-MAC Scale                 | Mini-Mental Adjustment to Cancer Scale  |
| <i>Generic</i>                 |   |
| EQ-5D                          | EuroQol-5 dimension   |
| SF-12 (Version 2)              | Version 2 of the 12-Item Short Form Health Survey                                     |
| SF-6D                          | 6-Item Short Form Health Survey   |

Note:

695 EORTC=European Organization for Research and Treatment of Cancer; FACIT=Functional

696 Assessment of Chronic Illness Therapy

Table 2. Methodological Quality of Each Study per Measurement Property and HRQOL Instrument

| Instrument/Reference | Measurement Properties |             |                   |                  |                     |                    |                         |                    |                |
|----------------------|------------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|----------------|
|                      | Internal Consistency   | Reliability | Measurement error | Content validity | Structural validity | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
| <i>EORTC</i>         |                        |             |                   |                  |                     |                    |                         |                    |                |
| QLQ-C30 (Version 1)  |                        |             |                   |                  |                     |                    |                         |                    |                |
| [64]                 |                        |             |                   |                  | Poor                | Good               |                         |                    |                |
| [28]                 | Poor                   |             |                   |                  | Poor                | Excellent          |                         |                    |                |
| QLQ-C30 (Version 3)  |                        |             |                   |                  |                     |                    |                         |                    |                |
| [29]                 | Poor                   |             |                   |                  | Poor                | Good               |                         |                    |                |
| [65]                 | Poor                   |             |                   |                  | Poor                | Good               |                         |                    |                |
| [66]                 | Excellent              | Fair        |                   |                  | Excellent           | Excellent          | Poor                    | Poor               | Poor           |
| [30]                 | Good                   |             |                   |                  | Good                | Poor               |                         |                    | Good           |
| [27]                 |                        | Fair        |                   |                  | Poor                | Poor               |                         |                    | Fair           |
| QLQ-CR38             |                        |             |                   |                  |                     |                    |                         |                    |                |
| [24]                 | Poor                   | Poor        |                   | Excellent        | Poor                | Good               |                         |                    | Good           |
| [23]                 | Poor                   | Fair        |                   |                  | Poor                | Excellent          |                         |                    |                |
| [22]                 | Poor                   |             |                   |                  | Poor                | Fair               |                         |                    |                |
| [26]                 | Poor                   | Good        |                   |                  | Poor                | Fair               | Poor                    |                    | Fair           |
| [27]                 |                        | Fair        |                   |                  | Poor                | Poor               |                         |                    | Fair           |
| [25]                 | Poor                   |             |                   |                  | Poor                | Excellent          |                         |                    |                |
| [21]                 | Good                   | Fair        |                   |                  | Good                | Excellent          |                         | Poor               | Poor           |
| QLQ-CR29             |                        |             |                   |                  |                     |                    |                         |                    |                |
| [54]                 | Poor                   | Good        |                   |                  | Poor                | Good               | Poor                    |                    | Poor           |
| [31]                 | Poor                   |             |                   |                  | Poor                | Good               | Poor                    |                    |                |
| QLQ-SWB36            |                        |             |                   |                  |                     |                    |                         |                    |                |
| [53]                 |                        |             |                   | Excellent        |                     |                    | Poor                    |                    |                |
| QLQ-LMC21            |                        |             |                   |                  |                     |                    |                         |                    |                |
| [50]                 |                        |             |                   | Excellent        | Poor                |                    |                         |                    |                |

| Instrument/Reference            | Measurement Properties |             |                   |                  |                     |                    |                         |                    |                |
|---------------------------------|------------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|----------------|
|                                 | Internal Consistency   | Reliability | Measurement error | Content validity | Structural validity | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
| [46]                            | Poor                   |             |                   |                  | Poor                | Excellent          |                         | Good               | Poor           |
| <i>FACIT</i>                    |                        |             |                   |                  |                     |                    |                         |                    |                |
| FACT-G (Version 2)              |                        |             |                   |                  |                     |                    |                         |                    |                |
| [32]                            | Fair                   | Fair        |                   | Excellent        | Good                | Fair               |                         |                    | Fair           |
| FACT-G (Version 3)              |                        |             |                   |                  |                     |                    |                         |                    |                |
| [29]                            | Excellent              |             |                   |                  | Good                | Good               |                         |                    |                |
| FACT-G (Version 4)              |                        |             |                   |                  |                     |                    |                         |                    |                |
| [23]                            | Poor                   | Fair        |                   |                  | Poor                | Excellent          |                         |                    |                |
| [69]                            | Poor                   | Fair        |                   |                  | Poor                | Good               |                         |                    | Good           |
| [27]                            |                        | Fair        |                   |                  | Poor                | Poor               |                         |                    | Fair           |
| <i>FACT-C</i>                   |                        |             |                   |                  |                     |                    |                         |                    |                |
| [33]                            | Poor                   |             |                   | Excellent        | Poor                | Good               | Poor                    |                    | Good           |
| [34]                            | Poor                   |             | Fair              |                  |                     | Good               |                         |                    |                |
| [55]                            | Poor                   |             |                   |                  | Poor                | Excellent          | Fair                    |                    |                |
| [23]                            | Poor                   | Fair        |                   |                  | Poor                | Excellent          |                         |                    |                |
| [27]                            |                        | Fair        |                   |                  | Poor                | Poor               |                         |                    | Fair           |
| [59]                            | Poor                   | Fair        |                   |                  | Poor                | Excellent          |                         |                    |                |
| [60]                            |                        |             |                   |                  |                     |                    |                         |                    | Good           |
| [61]                            |                        |             |                   |                  | Excellent           | Excellent          |                         |                    |                |
| <i>FCSI-9</i>                   |                        |             |                   |                  |                     |                    |                         |                    |                |
| [35]                            | Poor                   | Good        |                   |                  | Poor                | Excellent          |                         |                    | Excellent      |
| <i>FACT-G7</i>                  |                        |             |                   |                  |                     |                    |                         |                    |                |
| [71]                            | Poor                   |             |                   |                  | Poor                | Fair               |                         | Good               |                |
| <i>FACIT-F</i>                  |                        |             |                   |                  |                     |                    |                         |                    |                |
| [68]                            | Poor                   |             |                   |                  | Poor                | Excellent          |                         |                    |                |
| <i>Other Condition-specific</i> |                        |             |                   |                  |                     |                    |                         |                    |                |
| <i>FLIC</i>                     |                        |             |                   |                  |                     |                    |                         |                    |                |

| Instrument/Reference | Measurement Properties |             |                   |                  |                     |                    |                         |                    |                |
|----------------------|------------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|----------------|
|                      | Internal Consistency   | Reliability | Measurement error | Content validity | Structural validity | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
| [64]                 | Poor                   |             |                   |                  | Poor                | Good               |                         |                    |                |
| [29]                 | Poor                   |             |                   |                  | Poor                | Good               |                         |                    |                |
| Quick-FLIC           |                        |             |                   |                  |                     |                    |                         |                    |                |
| [63]                 | Excellent              | Excellent   |                   |                  | Poor                | Good               |                         |                    | Good           |
| RSCL-M               |                        |             |                   |                  |                     |                    |                         |                    |                |
| [36]                 | Fair                   |             |                   |                  | Fair                | Fair               |                         |                    |                |
| QLICP-GM             |                        |             |                   |                  |                     |                    |                         |                    |                |
| [67]                 | Fair                   | Fair        |                   | Excellent        | Fair                | Fair               |                         | Poor               | Poor           |
| QLICP-CR             |                        |             |                   |                  |                     |                    |                         |                    |                |
| [62]                 | Poor                   | Fair        |                   | Poor             | Poor                | Excellent          |                         | Poor               | Poor           |
| GIQLI                |                        |             |                   |                  |                     |                    |                         |                    |                |
| [52]                 |                        |             |                   |                  | Poor                | Fair               |                         |                    |                |
| MSAS-SF              |                        |             |                   |                  |                     |                    |                         |                    |                |
| [56]                 | Poor                   |             |                   |                  | Poor                | Fair               | Poor                    |                    |                |
| CMSAS                |                        |             |                   |                  |                     |                    |                         |                    |                |
| [56]                 | Poor                   |             |                   |                  | Poor                | Fair               | Poor                    |                    |                |
| Modified FIQL        |                        |             |                   |                  |                     |                    |                         |                    |                |
| [70]                 | Excellent              |             |                   |                  | Excellent           | Fair               |                         |                    |                |
| mCOH-QOL-Ostomy      |                        |             |                   |                  |                     |                    |                         |                    |                |
| [37]                 | Fair                   |             |                   | Excellent        | Fair                | Fair               |                         |                    |                |
| [38]                 | Fair                   |             |                   |                  | Poor                | Fair               |                         |                    |                |
| [39]                 | Fair                   |             |                   |                  | Poor                | Fair               |                         |                    |                |
| SRQS                 |                        |             |                   |                  |                     |                    |                         |                    |                |
| [57]                 | Good                   |             |                   |                  | Good                | Good               | Fair                    |                    |                |
| Stoma Care QoL Scale |                        |             |                   |                  |                     |                    |                         |                    |                |
| [47]                 | Good                   |             |                   |                  | Excellent           | Fair               |                         |                    |                |
| MACFS                |                        |             |                   |                  |                     |                    |                         |                    |                |

| Instrument/Reference           | Measurement Properties |             |                   |                  |                     |                    |                         |                    |                |
|--------------------------------|------------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|----------------|
|                                | Internal Consistency   | Reliability | Measurement error | Content validity | Structural validity | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
| [48]                           |                        |             |                   |                  |                     | Fair               |                         |                    |                |
| Mini-MAC Scale                 |                        |             |                   |                  |                     |                    |                         |                    |                |
| [49]                           | Fair                   | Fair        |                   |                  | Good                | Fair               |                         |                    |                |
| MSKCC Bowel Function           |                        |             |                   |                  |                     |                    |                         |                    |                |
| [40]                           | Good                   | Poor        |                   | Excellent        | Good                | Excellent          |                         |                    |                |
| [41]                           | Good                   | Fair        |                   |                  | Good                | Excellent          | Fair                    |                    |                |
| CRQ                            |                        |             |                   |                  |                     |                    |                         |                    |                |
| [42]                           | Poor                   | Poor        |                   | Poor             | Poor                | Fair               |                         |                    |                |
| Coping with CRC                |                        |             |                   |                  |                     |                    |                         |                    |                |
| [58]                           | Fair                   | Fair        |                   | Excellent        | Fair                | Fair               |                         | Poor               |                |
| Quality-Quantity Questionnaire |                        |             |                   |                  |                     |                    |                         |                    |                |
| [43]                           | Fair                   |             |                   |                  | Fair                | Fair               |                         |                    |                |
| HLQ                            |                        |             |                   |                  |                     |                    |                         |                    |                |
| [51]                           | Fair                   | Poor        |                   |                  | Fair                | Fair               |                         |                    |                |
| QLACS                          |                        |             |                   |                  |                     |                    |                         |                    |                |
| [44]                           | Fair                   |             |                   | Excellent        | Fair                | Fair               |                         | Fair               |                |
| Body-image Scale               |                        |             |                   |                  |                     |                    |                         |                    |                |
| [45]                           | Fair                   | Poor        |                   |                  | Fair                | Fair               |                         |                    | Poor           |
| Generic                        |                        |             |                   |                  |                     |                    |                         |                    |                |
| SF-12 (Version 2)              |                        |             |                   |                  |                     |                    |                         |                    |                |
| [60]                           |                        |             |                   |                  |                     |                    |                         |                    | Good           |
| EQ-5D                          |                        |             |                   |                  |                     |                    |                         |                    |                |
| [72]                           |                        | Fair        |                   |                  |                     | Excellent          |                         |                    |                |
| SF-6D                          |                        |             |                   |                  |                     |                    |                         |                    |                |
| [73]                           |                        |             |                   |                  |                     |                    |                         |                    | Good           |

Note: Rating of methodological quality according to 4-point Likert scale: Excellent, Good, Fair or Poor.

Table 3. Overall Levels of Evidence per Measurement Property and HRQOL Instrument

| Instrument                      | Measurement Properties |             |                   |                  |                     |                    |                         |                    | Reference |                             |
|---------------------------------|------------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|-----------|-----------------------------|
|                                 | Internal Consistency   | Reliability | Measurement error | Content validity | Structural validity | Hypothesis testing | Cross-cultural validity | Criterion validity |           | Responsiveness              |
| <i>EORTC†</i>                   |                        |             |                   |                  |                     |                    |                         |                    |           |                             |
| QLQ-C30 (Version 1)             | ?                      |             |                   |                  | ?                   | +++                |                         |                    |           | [28, 64]                    |
| QLQ-C30 (Version 3)             | +++                    | +           |                   |                  | ?                   | +++                | ?                       | ?                  | --        | [27, 30, 65, 66]            |
| QLQ-CR38                        | ++                     | ++          |                   | +++              | ++                  | +++                | ?                       | ?                  | ++        | [22-27]                     |
| QLQ-CR29                        | ?                      | ++          |                   |                  | ?                   | ++                 | ?                       |                    | ?         | [31, 54]                    |
| QLQ-SWB36                       |                        |             |                   | +++              |                     |                    | ?                       |                    |           | [53]                        |
| QLQ-LMC21                       | ?                      |             |                   | +++              | ?                   | +++                |                         | ++                 | ?         | [46, 50]                    |
| <i>FACIT</i>                    |                        |             |                   |                  |                     |                    |                         |                    |           |                             |
| FACT-G (Version 2)              | +                      | +           |                   | +++              | ++                  | +                  |                         |                    | +         | [32]                        |
| FACT-G (Version 3)              | +++                    |             |                   |                  | ++                  | ++                 |                         |                    |           | [29]                        |
| FACT-G (Version 4)              | ?                      | +           |                   |                  | ?                   | +++                |                         |                    | ++        | [23, 27, 69]                |
| FACT-C                          | ?                      | +           | +                 | +++              | ?                   | +++                | ?                       |                    | ++        | [23, 27, 33, 34, 55, 59-61] |
| FCSI-9                          | ?                      | ++          |                   |                  | ?                   | +++                |                         |                    | +++       | [35]                        |
| FACT-G7                         | ?                      |             |                   |                  | ?                   | +                  |                         | ++                 |           | [71]                        |
| FACIT-F                         | ?                      |             |                   |                  | ?                   | +++                |                         |                    |           | [68]                        |
| <i>Other Condition-specific</i> |                        |             |                   |                  |                     |                    |                         |                    |           |                             |
| FLIC                            | ?                      |             |                   |                  | ?                   | ++                 |                         |                    |           | [64]                        |
| Quick-FLIC                      | +++                    | +++         |                   |                  | ?                   | ++                 |                         |                    | ++        | [63]                        |
| RSCL-M                          | +                      |             |                   |                  | ?                   | +                  |                         |                    |           | [36]                        |
| QLICP-GM                        | +                      | +           |                   | +++              | ?                   | +                  |                         | ?                  | ?         | [67]                        |
| QLICP-CR                        | ?                      | +           |                   | ?                | ?                   | +++                |                         | ?                  | ?         | [62]                        |



| Instrument                     | Measurement Properties |             |                   |                  |                     |                    |                         |                    | Reference |                |
|--------------------------------|------------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|-----------|----------------|
|                                | Internal Consistency   | Reliability | Measurement error | Content validity | Structural validity | Hypothesis testing | Cross-cultural validity | Criterion validity |           | Responsiveness |
| GIQLI                          |                        |             |                   |                  | ?                   | +                  |                         |                    |           | [52]           |
| MSAS-SF                        | ?                      |             |                   |                  | ?                   | +                  | ?                       |                    |           | [56]           |
| CMSAS                          | ?                      |             |                   |                  | ?                   | +                  | ?                       |                    |           | [56]           |
| Modified FIQL                  | +++                    |             |                   |                  | ?                   | +                  |                         |                    |           | [70]           |
| mCOH-QOL-Ostomy                | +                      |             |                   | +++              | +                   | +                  |                         |                    |           | [37-39]        |
| SRQS                           | ++                     |             |                   |                  | ++                  | ++                 | ?                       |                    |           | [57]           |
| Stoma Care QoL scale           | ++                     |             |                   |                  | ?                   | +                  |                         |                    |           | [47]           |
| MACFS                          |                        |             |                   |                  |                     | +                  |                         |                    |           | [48]           |
| Mini-MAC Scale                 | +                      | -           |                   |                  | --                  | +                  |                         |                    |           | [49]           |
| MSKCC Bowel Function           | ++                     | +           |                   | +++              | ++                  | ++                 | ?                       |                    |           | [40, 41]       |
| CRQ                            | ?                      | ?           |                   | ?                | ?                   | ?                  |                         |                    |           | [42]           |
| Coping with CRC                | +                      | -           |                   | +++              | +                   | +                  |                         | ?                  |           | [58]           |
| Quality-Quantity Questionnaire | +                      |             |                   |                  | -                   | ?                  |                         |                    |           | [43]           |
| HLQ                            | +                      | ?           |                   |                  | +                   | +                  |                         |                    |           | [51]           |
| QLACS                          | +                      |             |                   | +++              | +                   | +                  |                         | ?                  |           | [44]           |
| Body-image Scale               | +                      | ?           |                   |                  | ?                   | +                  |                         |                    | ?         | [45]           |
| <i>Generic</i>                 |                        |             |                   |                  |                     |                    |                         |                    |           |                |
| SF-12 (Version 2)              |                        |             |                   |                  |                     |                    |                         |                    | --        | [60]           |
| EQ-5D                          |                        | -           |                   |                  |                     | +++                |                         |                    |           | [72]           |
| SF-6D                          |                        |             |                   |                  |                     |                    |                         |                    | --        | [73]           |

Note:

Overall levels of evidence: +++/--- , strong evidence positive/negative result; ++/-- , moderate evidence positive/negative result; +/- , limited evidence positive/negative result; ?, unknown due to poor methodological quality.

Figure 1. PRISMA Flow Diagram of the literature search and selection process

