



Reliability, validity, and responsiveness of the Japanese version of the EORTC QLQ-ELD14 in evaluating the health-related quality of life of elderly patients with cancer

Yumiko Kinoshita^{1,2} · Rieko Izukura³ · Junji Kishimoto⁴ · Maki Kanaoka¹ · Hayato Fujita^{5,6} · Koji Ando⁷ · Shuntaro Nagai^{6,8} · Sayuri Akiyoshi⁷ · Tetsuzo Tagawa^{7,9} · Makoto Kubo⁶ · Junichi Inokuchi¹⁰ · Kenoki Ohuchida⁶ · Eiji Oki⁷ · Kentaro Tanaka¹¹ · Masatoshi Eto¹⁰ · Tomoharu Yoshizumi⁷ · Masafumi Nakamura⁶ · Akiko Chishaki^{2,12}

Received: 10 July 2022 / Accepted: 11 October 2022 / Published online: 28 October 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Purpose This study evaluated the reliability, validity, and responsiveness of the Japanese version of the European Organization for Research and Treatment of Cancer (EORTC) QLQ-ELD14 and measured the health-related quality of life (HRQOL) of elderly Japanese patients with cancer aged ≥ 60 and ≥ 70 years.

Methods The study recruited elderly Japanese patients with cancer aged ≥ 60 (≥ 70) years ($n = 1803$ [$n = 1236$]). The EORTC QLQ-ELD14 was evaluated for reliability, validity, responsiveness, and correlations of changes in score between the EORTC QLQ-ELD14 and the EORTC QLQ-C30 before and after the commencement of the COVID-19 pandemic.

Results In both age groups, the proportion of missing items was low ($< 3\%$). Cronbach's α was good at ≥ 0.70 , except for two of the seven items. All the intraclass coefficient constants were good at ≥ 0.70 . The concurrent validity was good but correlation with the EORTC QLQ-C30 was not strong, except for the hypothesis items. Regarding the assessment of responsiveness, only one item ("maintaining purpose") of the EORTC QLQ-ELD14 worsened (-6.14 ± 29.20 , standard response of mean > 0.2) after the commencement of the COVID-19 pandemic. The changes in score between the EORTC QLQ-ELD14 and the "global health status/QOL" and "summary score" of the EORTC QLQ-C30 had moderate-to-high negative correlations for all items, except two. Hypotheses to evaluate construct validity were accepted at 90%, while responsiveness was accepted at 80%.

Conclusion The Japanese version of the EORTC QLQ-ELD14 questionnaire appears to have acceptable reliability, validity, and responsiveness to evaluate HRQOL in elderly Japanese people with cancer.

Keywords Validation study · Quality of life · Japanese · Aged · Neoplasms · COVID-19 pandemic

✉ Yumiko Kinoshita
yumiko_kinoshita@med.miyazaki-u.ac.jp

¹ School of Nursing, Faculty of Medicine, University of Miyazaki, 5200 Kiyotakecho Kihara, Miyazaki 889-1692, Japan

² Department of Health Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

³ Department of Basic Medicine, Faculty of Medical Sciences, Kyushu University, Fukuoka, Japan

⁴ Center of Clinical and Translational Research, Kyushu University Hospital, Fukuoka, Japan

⁵ Department of Surgery, Japan Community Health Care Organization Fukuoka Central Yutaka Hospital, Fukuoka, Japan

⁶ Department of Surgery and Oncology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

⁷ Department of Surgery and Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

⁸ Department of Surgery, Kitakyushu Municipal Medical Center, Fukuoka, Japan

⁹ Department of Thoracic Surgery, Clinical Research Institute, National Hospital Organization, Kyushu Medical Center, Fukuoka, Japan

¹⁰ Department of Urology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

¹¹ Department of Respiratory Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

¹² Faculty of Nursing, Fukuoka Nursing College, Fukuoka, Japan

Introduction

Cancer is a major cause of morbidity and mortality in elderly populations worldwide, particularly in developed countries, owing to their proportionately large aging population (Pilleron et al. 2019). Age is the strongest risk factor for the development of cancer. In comparison to younger people with cancer, older people with cancer tend to prefer a better health-related quality of life (HRQOL) over an increased span of life (Tester et al. 2004). Moreover, the concept of HRQOL in elderly patients is not linked to any specific medical condition and is a broad concept (Wilson and Cleary 1995). In an original study on HRQOL, Wilson and Cleary (1995) divided health outcomes into the following five levels: biological and physiological factors, symptoms, functioning, general health perceptions, and overall quality of life. They determined that, at each level, there are an increasing number of traditional clinical variables that cannot be controlled by clinicians or the health care system (Wright et al. 2005).

The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) is the primary questionnaire used to evaluate HRQOL in cancer patients worldwide (Aaronson et al. 1993). However, elderly individuals with cancer (aged ≥ 70 years) have a different HRQOL profile compared with other age groups, and the EORTC QLQ-C30 does not address their specific needs (Fitzsimmons et al. 2009; Johnson et al. 2010). There are significant differences in EORTC QLQ-C30 responses according to age (Hjermstad et al. 1998; Michelson et al. 2000; Schwarz and Hinz 2001). The EORTC QLQ-ELD15 and its revised version (ELD14) were developed to complement the QLQ-C30 and to account for age-specific issues that are of relevance and importance for older cancer patients (Wright et al. 2005); these two questionnaires help assess the HRQOL of elderly people with cancer (Johnson et al. 2010). Studies using the EORTC QLQ-ELD14 were conducted with patients with cancer from 10 countries ($n=518$) and showed good reliability and validity (Wheelwright et al. 2013). Moreover, it was recently validated in Korea (Goo et al. 2017), Poland (Wrazen et al. 2014), Spain (Arraras et al. 2019), and Chile (Lorca et al. 2021).

The term “elderly” does not have a clear definition, and it is important to assess the definition individually. The Korean version of the scale was validated for reliability and validity in elderly patients, ≥ 60 years, with cancer (Goo et al. 2017). From the societal point of view, the age of retirement entitlement in elderly people is approximately 55–70 years (World Health Organization 2021). Japanese people retire at approximately 60 years

in most companies, but in recent years the retirement age has increased. The World Health Organization (WHO) described aging populations as those aged ≥ 65 years (World Health Organization 2021). In politics, the definition applies to ages ranging from 70 or 75 years and above. The EORTC QLQ-ELD15 paper described the definition of elderly as age ≥ 70 years as one of the potential limitations of the module (Johnson et al. 2010). We aimed to assess biological age rather than chronological age. A thorough search of the literature revealed that there were no studies evaluating HRQOL changes in Japanese people with cancer, especially among the elderly, using the EORTC QLQ-ELD14. Therefore, we evaluated the psychometric properties of the Japanese version of the EORTC QLQ-ELD14 with the aim of confirming its applicability to Japanese individuals with cancer ≥ 60 and ≥ 70 years of age and of comparing the results to the Korean EORTC QLQ-ELD14 study.

The COVID-19 pandemic had not begun when we commenced this study. After the pandemic began in December 2019, we were unsure of the feasibility of the investigation, but we soon found it to be a valuable opportunity to evaluate how HRQOL changes in people with cancer.

The pandemic has had a significant health impact on the elderly. Consequently, we also attempted to assess the changes in the EORTC QLQ-ELD14 score due to the influence of the COVID-19 pandemic, in elderly individuals with cancer. To the best of our knowledge, this is the first study to compare the HRQOL of elderly people before and during the COVID-19 pandemic, using the EORTC QLQ-ELD14.

Participants and methods

Participants

Of the 4166 cancer patients (the five most common cancer types in the Japanese population being colorectal, gastric, lung, breast, and prostate) with data in the hospital-based cancer registry (registered from 2013 at Kyushu University Hospital), 1803 (43.3%) with data from June 2019 to March 2021 were included in this study. Initially, we met patients at outpatient clinics whenever possible and asked them to complete the questionnaires at home and return them, but after the emergence of the COVID-19 pandemic, we conducted the survey by mail. Eligibility criteria included (i) a diagnosis of primary cancer, (ii) age ≥ 60 years, (iii) physician’s approval, (iv) mentally healthy as far as the physicians were aware, and (v) ability to complete the questionnaire. Secondary cancers were excluded from the study.

Questionnaires

EORTC QLQ-C30 (version 3.0)

The 30-item EORTC QLQ-C30 (version 3.0) assesses the functional status and symptoms of cancer patients and contains the following items: a “global health status/QOL” scale, five functional scales/items (physical, role, cognitive, emotional, and social), and nine symptom scales/items (fatigue, nausea, pain, dyspnea, insomnia, anorexia, constipation, diarrhea, and financial difficulties) (Kobayashi et al. 1998; Schwarz and Hinz 2001; Arraras et al. 2019). “Global health status/QOL” questions include the following: “How would you rate your overall health during the past week?” and “How would you rate your overall quality of life during the past week?” A mean difference of ≥ 10 points is regarded as clinically meaningful for all EORTC QLQ-C30 scales (Osoba et al. 1998). Total scores range from 0 to 100. Higher scores on the “global health status/QOL” and functional scales indicate better functioning. Higher scores on the symptom scales/items indicate more severe problems. The EORTC QLQ-C30 “summary score” is calculated from the mean of 13/ 15 item scores of the QLQ-C30 scales; except the items “global health status/QOL”, and “financial difficulties” on the symptom scales/items (Kasper 2020). In recent years, the EORTC QLQ-C30 “summary score” was indicated to have the highest prognostic significance in predicting the survival of patients with cancer compared with individual item and scale scores.

EORTC QLQ-ELD14

The Japanese version of the EORTC QLQ-ELD14 was authorized by the EORTC Quality of Life Group Translation Unit (Fayers and Machin 2000) and was translated in accordance with the EORTC quality of life group translation procedure (Young et al. 1999). The translated version of the EORTC QLQ-ELD14 was then pilot-tested with 10 patients with cancer. Patients were asked if they found any of the questions difficult, confusing, upsetting, and so forth. Based on these suggestions, the questionnaire was modified. The EORTC QLQ-ELD14 was subjected to a standard forward–backward translation process. Two translators translated the EORTC QLQ-ELD14 into Japanese with forward translations. After that, two professional translators translated the Japanese version of the EORTC QLQ-ELD14 back into English. The translation report was then reviewed by the EORTC translation coordinator. The 14-item QLQ-ELD14 is designed to assess the HRQOL of elderly individuals with cancer, complementing the QLQ-C30. The QLQ-ELD14 has the following two functional scales: “maintaining purpose” (2 items) and “family support” (1 item); it also has five multi-item symptom scales: “mobility” (3 items), “worries

about others” (2 items), “future worries” (3 items), “burden of illness” (2 items), and “joint stiffness” (1 item) (Wheeler et al. 2013). Total scores range from 0 to 100. Higher scores on the functional scales indicate better functioning. Higher scores on the symptom and single-item scales indicate more severe problems.

G8 frailty screening tool

The G8 was developed for elderly individuals with cancer (Bellera et al. 2012). It assesses age, appetite changes, weight loss, mobility, neuropsychological problems, body mass index, medication intake, and self-reported health, by single-item scales. Total scores range from 0 (poor score) to 17 (good score). Higher scores indicate a “not at all impaired” status, with a score cut-off for potential frailty of ≤ 14 (Mohile et al. 2018).

Analysis

First, to establish a 95% confidence interval of the estimated proportion within $\pm 2.5\%$, it was necessary to collect data from 1600 elderly people. Descriptive statistics were used to analyze socio-demographic and clinical data.

We followed the EORTC QLQ-C30 Scoring Manual (Osoba et al. 1998; Fayers and Machin 2000; Goo et al. 2017). For missing items, a straightforward method used by many HRQOL instruments for imputing items from multi-item scales was used as following: if the participant had answered at least half of the items from the scale, it was assumed that the missing items would have values equal to the average of those items that had been answered by that particular respondent. Moreover, we used the COSMIN Risk of Bias checklist (Mokkink et al. 2018a, b).

Reliability

The reliability domain consists of the following three measurement properties: (i) internal consistency, (ii) reliability, and (iii) measurement error. First, Cronbach α was used to assess the level of reliability of internal consistency, and $\alpha \geq 0.7$ was considered statistically significant (Cronbach 1951). Furthermore, it indicates the level of internal consistency with respect to the specific sample. Second, for the test–retest analysis, conventionally, an intraclass coefficient constant (ICC) of ≥ 0.70 was considered sufficient (Shrout and Fleiss 1979). Third, measurement error refers to the systematic and random error in an individual patient’s score that is not attributable to true changes in the construct being measured; the statistic used to assess measurement error was the smallest detectable change (SDC) (Fayers and Machin 2000). It is directly related to the standard error of measurement (SEM). Changes within the SDC, and smaller

than the SDC, were attributed to measurement error, and changes outside the SDC, and larger than the SDC, were considered to be real changes in individual patients. SEM is a measure of how much the measured test scores are spread around a “true” score. The SDC is a measure of the variation in a scale due to measurement error. Thus, a change score can only be considered to represent a real change if it is larger than the SDC. It is the same as a minimum detectable change.

Validity

Validity was measured as construct and criterion validity.

Although we first attempted to use factor analysis, we were unable to do so because the structure of the items was different. We performed multi-trait analysis (Pearson’s correlation uses parametric equivalents to evaluate the linear relationship between two continuous variables; 0–100) to assess structural validity. Convergent validity was considered when an item correlated highly with its own hypothesized scale (correlation ≥ 0.40 ; corrected for overlap). Discriminant validity was considered when an item did not correlate highly with the scales it was not a part of. Discriminant validity was supported when the correlation between an item and its hypothesized scale (corrected for overlap) was > 2 standard errors higher than its correlation with other scales.

We developed the following hypotheses to evaluate construct validity:

- (1) The direction and magnitude of differences in the EORTC QLQ-ELD14 between groups: known-group comparisons.
 - (a) Disease stage; patients with stage 0–II would have significantly superior scores compared with those in stages III–IV; “Worries about others” on the EORTC QLQ-ELD14.
 - (b) Comorbidities; patients with no comorbidities would have significantly superior scores than those with > 1 comorbidity; “Mobility”, “Burden of illness”, “Joint stiffness”.
 - (c) G8; Patients who were not frail would have significantly superior scores compared to frail patients; “Mobility”, “Worries about others”, “Future worries”, and “Burden of illness”.
- (2) Correlations with instruments measuring similar constructs; criterion convergent validity
 - (a) Between “Mobility” on EORTC QLQ-ELD14 and “Physical functioning” on the EORTC QLQ-

C30, there would be a high negative correlation of ≥ -0.70 .

- (b) Between “Mobility” on the EORTC QLQ-ELD14 and “Role functioning” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40
- (c) Between “Mobility” on the EORTC QLQ-ELD14 and “Fatigue” on the EORTC QLQ-C30, there would be a moderate-to-high positive correlation of ≥ 0.40 .
- (d) Between “Future worries” on the EORTC QLQ-ELD14 and “Emotional functioning” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40 .
- (e) Between “Future worries” on the EORTC QLQ-ELD14 and “Social functioning” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40 .
- (f) Between “Burden of illness” on the EORTC QLQ-ELD14 and “Social functioning” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40 .
- (g) Between “Burden of illness” on the EORTC QLQ-ELD14 and “Financial difficulties” on the EORTC QLQ-C30, there would be a moderate-to-high positive correlation of ≥ 0.4 .

Hypotheses to evaluate construct validity are shown in Supplementary Table S1. The results were confirmed if 75% of the results were in accordance with the hypotheses and the testing and responsiveness results.

The hypotheses about discriminative or known group validity were made about the expected differences in EORTC QLQ-ELD14 scores between patients according to Union for International Cancer Control (UICC) stage, comorbidity, and G8 scores (a cut-off for potential frailty), by known group comparisons. Known group comparisons were assessed by comparing elderly people with cancer using independent *t* tests.

Concurrent validity was analyzed using the EORTC QLQ-ELD14 scale, and the correlation coefficient between single items, to evaluate the capabilities of each scale and of single items. In addition, criteria-related validity was assessed using a Pearson correlation analysis between each scale of the EORTC QLQ-ELD14 and EORTC QLQ-C30. These scales correlated weakly (only “maintaining purpose” on the EORTC QLQ-ELD14 did not correlate significantly with most of the items on the EORTC QLQ-C30). We hypothesized with respect to the expected magnitude and orientation of the relationship between the EORTC QLQ-ELD14 and EORTC QLQ-C30 subscales in Supplementary Table S1.

Responsiveness

The hypotheses used to evaluate responsiveness are shown in Supplementary Table S2. The results were confirmed if 75% of the results were in accordance with the hypotheses, and the testing and responsiveness of the results.

The hypotheses to evaluate responsiveness were as follows:

- (1) Responsiveness to change on the EORTC QLQ-ELD14 would worsen after the commencement of the COVID-19 pandemic; $SRM > 0.2$. Four of Five hypotheses were in accordance with the hypothesis. Only “Future worries” on EORTC QLQ-ELD14 was not in accordance with the hypotheses.
 - (a) “Maintaining of purpose” on the EORTC QLQ-ELD14.
 - (b) “Future worries” on the EORTC QLQ-ELD14.
- (2) Correlations with changes in instruments measuring similar constructs. All three results were in accordance with the hypotheses.
 - (a) Between “mobility” on the EORTC QLQ-ELD14 and “Global health stats QOL” and “Summary Score” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40 .
 - (b) Between “future worries” on the EORTC QLQ-ELD14 and “Global health stats QOL” and “Summary Score” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40 .
 - (c) Between “burden of illness” on the EORTC QLQ-ELD14 and “Global health stats QOL” and “Summary Score” on the EORTC QLQ-C30, there would be a moderate-to-high negative correlation of ≥ -0.40 .

Responsiveness was evaluated as standardized effect size (ES) and standardized response mean (SRM) (Beaton et al. 2001), using paired *t*-tests to compare the changes in the scores from before and after the commencement of the COVID-19 pandemic. The ES was determined by calculating the difference between the means before and after treatment and dividing it by the standard deviation of the same measure before treatment. The ES is sensitive to between-subject variability, and SRM is sensitive to within-subject variability. The larger the SRM, the greater the responsiveness. An SRM of 0.20 was considered small, 0.50 as moderate, and 0.80 or higher as large. Next, correlations between changes in the EORTC QLQ-ELD14 and EORTC QLQ C30 scales were calculated with Pearson’s correlations.

All tests were two-tailed. Statistical significance was set at $p < 0.05$. Analyses were performed with IBM SPSS statistics version 26.0 software (IBM Corp, Armonk, NY, USA).

Ethics

This study was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and was approved by the Ethics Committee of Kyushu University (2019-450). All persons gave informed consent prior to inclusion in the study. Prior to the study, participants were asked to sign a written informed consent form to confirm their willingness to participate in this study.

Results

The proportion of missing items on the EORTC QLQ-ELD14 was 0.3–2.3% for those aged ≥ 60 years and 0.3–2.0% for those aged ≥ 70 years. The highest scored item was item number 12 (Have you felt motivated to continue with your normal hobbies and activities?) in “maintaining of purpose”. Missing data of $< 5\%$ for each item was considered acceptable (Fayers and Machin 2013).

Sociodemographic characteristics

The final analysis included 1803 individuals (aged 60 years and older) and 1236 individuals (aged 70 years and older). The characteristics of the participants are described in Table 1. The mean age of the patients was 73.04 ± 6.858 in the ≥ 60 years age group and 76.55 ± 5.093 in the ≥ 70 years age group. There were 62.3% men (aged ≥ 60 years) and 65.8% men (aged ≥ 70 years). The period of diagnosis was 2–5 years in 46.4% of those aged ≥ 60 years and 45.6% in those aged ≥ 70 years.

Reliability

Internal consistency—Cronbach’s α

Table 2 shows that all the measures met the criteria for internal coherence (Cronbach α). Cronbach’s α for the scales was considered sufficient for all items with a value higher than 0.70; the exception was “maintaining purpose” with values of 0.65 in the ≥ 60 years age group and 0.62 in the ≥ 70 years age group, and “worries about others” with values of 0.67 in the ≥ 60 years age group and 0.68 in the ≥ 70 years age group.

Table 1 Baseline information

Baseline questions	Answers	≥ 60 years (N=1803)		≥ 70 years (N=1236)	
Age	Age (mean ± SD)	73.04 ± 6.86		76.55 ± 5.09	
		Frequency	Percent	Frequency	Percent
Sex	Male	1124	62.3	813	65.8
	Female	679	37.7	423	34.2
Marital status	Married	1599	88.7	1113	90.0
	Others	203	11.3	122	9.9
	Missing data	1	0.1	1	0.1
Living with others	No	306	17.0	219	17.7
	Yes	1496	83.0	1016	82.2
	Missing data	1	0.1	1	0.1
Occupational status	Unemployed	1239	68.7	971	78.6
	Employed	563	31.2	264	21.4
	Missing data	1	0.1	1	0.1
Income	< 20,000 JPY	475	26.3	325	26.3
	≥ 20,000 JPY	890	49.4	596	48.2
	Missing data	438	24.3	315	25.5
Education	College or less	1219	67.6	850	68.8
	University or above	476	26.4	297	24.0
	Missing data	108	6.0	89	7.2
Diabetes	No	1456	80.8	973	78.7
	Yes	347	19.2	263	21.3
Hypertension	No	1405	77.9	941	76.1
	Yes	398	22.1	295	23.9
Cardiovascular disease	No	1700	94.3	1153	93.3
	Yes	103	5.7	83	6.7
Kidney disease	No	1701	94.3	1160	93.9
	Yes	102	5.7	76	6.1
Liver disease	No	1736	96.3	1183	95.7
	Yes	67	3.7	53	4.3
Pulmonary disease	No	1660	92.1	1137	92.0
	Yes	143	7.9	99	8.0
Cerebrovascular disease	No	1700	94.3	1153	93.3
	Yes	103	5.7	83	6.7
Number of comorbidities	0	1038	57.6	671	54.3
	1	401	22.2	280	22.7
	2	228	12.6	174	14.1
	3	102	5.7	82	6.6
	4	29	1.6	24	1.9
	5	5	0.3	5	0.4
Number of comorbidities < 1	0	1038	57.6	671	54.3
	≤ 1	765	42.4	565	45.7
Cancer type	Colorectal	379	21.0	266	21.5
	Gastric	262	14.5	183	14.8
	Lung	334	18.5	239	19.3
	Breast	289	16.0	156	12.6
	Prostate	539	29.9	392	31.7
UICC stage	0	44	2.4	23	1.9
	I	768	42.6	537	43.4
	II	454	25.2	305	24.7
	III	370	20.5	256	20.7
	IV	164	9.1	113	9.1

Table 1 (continued)

Baseline questions	Answers	≥ 60 years ($N=1803$)		≥ 70 years ($N=1236$)	
Period since diagnosis (years)	< 1	89	4.9	56	4.5
	1–2	434	24.1	288	23.3
	2–5	836	46.4	564	45.6
	5 and above	444	24.6	328	26.5
Surgery	No	351	19.5	263	21.3
	Yes	1452	80.5	973	78.7
Chemotherapy	No	1348	74.8	967	78.2
	Yes	455	25.2	269	21.8
Radiotherapy	No	1515	84.0	1041	84.2
	Yes	288	16.0	195	15.8

UICC, Union for International Cancer Control

Table 2 Cronbach's α and multitrait scaling analysis on the EORTC QLQ-ELD14

EORTC QLQ-ELD14		≥ 60 years ($N=1760$)				≥ 70 years ($N=1205$)			
Scale	Number of items	Mean \pm SD	Cronbach's α	Item-own scale correlation	Item-other scale correlation	Mean \pm SD	Cronbach's α	Item-own scale correlation	Item-other scale correlation
<i>Functional scale/item</i>									
Maintaining purpose	2	40.94 \pm 24.56	0.65	0.48	–0.02–0.19	39.86 \pm 24.08	0.62	0.45	0.00–0.23
Family support	1	29.20 \pm 31.16	NA	NA	0.19–0.52	29.90 \pm 31.06	NA	NA	0.23–0.53
<i>Symptom scales/items</i>									
Mobility	3	17.50 \pm 21.25	0.84	0.59–0.72	–0.02–0.60	19.81 \pm 22.35	0.85	0.61–0.72	0.00–0.53
Worries about others	2	30.05 \pm 23.76	0.67	0.51	0.18–0.66	30.95 \pm 24.01	0.68	0.51	0.21–0.65
Future worries	3	34.67 \pm 24.27	0.89	0.71–0.78	0.13–0.66	35.45 \pm 24.51	0.89	0.69–0.78	0.18–0.65
Burden of illness	2	31.69 \pm 25.27	0.85	0.74	0.26–0.63	32.96 \pm 25.35	0.86	0.75	0.06–0.63
Joint stiffness	1	21.23 \pm 25.20	NA	NA	0.05–0.60	21.99 \pm 24.95	NA	NA	0.06–0.61

Reliability—ICC

Table 3 shows the ICCs, and ICCs ≥ 0.7 were considered sufficient. ICCs for all the scales were above 0.80 in the ≥ 60 years age group and above 0.70 in the ≥ 70 years age group.

Measurement error-SEM

In the ≥ 60 years age group, the SDC scores ranged between 14.63 (mobility) to 29.84 (family support) on the 100 points scale. In the ≥ 70 years age group, the SDC scores ranged from 19.92 (mobility) to 36.80 (family support) on the 100 point scale. The SEMs/SDCs of all scales on the EORTC QLQ-ELD14 were less than 1.

Validity

Construct validity

Construct validity is shown in Table 2. Item-own scale correlations were 0.48–0.78 for the ≥ 60 years age group and 0.45–0.78 for the ≥ 70 years age group. Additionally, item-other scale correlations were –0.02–0.66 for the ≥ 60 years age group and 0.01–0.65 for the ≥ 70 years age group. Item 7, i.e., “worries about others” correlated more highly with items 8, 9, and 10 in the “future worries” scale.

Table 3 Test/re-test analysis on the EORTC QLQ-ELD14

Scale	≥ 60 years (N=101)					≥ 70 years (N=85)				
	Test	Re-test	ICC	SEM	SDC	Test	Re-test	ICC	SEM	SDC
	Mean ± SD	Mean ± SD			Mean ± SD	Mean ± SD				
<i>Functional scale/item</i>										
Maintaining purpose	42.73 ± 24.77	43.56 ± 25.49	0.87	9.13	25.31	41.76 ± 24.48	41.37 ± 24.61	0.75	12.32	34.16
Family support	23.10 ± 26.56	25.41 ± 26.72	0.85	10.18	28.22	21.18 ± 25.64	23.14 ± 26.24	0.74	13.28	36.80
<i>Symptom scales/items</i>										
Mobility	10.78 ± 13.47	10.78 ± 14.94	0.86	5.28	14.63	11.11 ± 13.06	10.46 ± 13.97	0.72	7.19	19.92
Worries about others	24.59 ± 19.94	27.55 ± 20.61	0.82	8.51	23.58	24.31 ± 20.01	26.08 ± 19.83	0.71	10.76	29.84
Future worries	29.48 ± 22.52	29.92 ± 20.05	0.85	8.30	23.00	28.24 ± 22.65	29.41 ± 20.55	0.74	10.95	30.36
Burden of illness	26.73 ± 22.61	26.73 ± 22.74	0.85	8.72	24.18	25.49 ± 21.91	24.90 ± 22.07	0.70	12.12	33.61
Joint stiffness	14.85 ± 20.26	15.18 ± 20.29	0.82	8.65	23.98	14.90 ± 20.90	14.51 ± 20.22	0.74	10.58	29.33

All correlation coefficients were calculated using a pair-wise deletion method.

NA, not available; ICC, intraclass correlation coefficient; SEM, standard error of measurement; SDC, smallest detectable change

Criterion validity

Hypotheses to evaluate construct validity (known group and concurrent validity) are shown in Supplementary Table S1. Regarding sufficient hypothesis testing for construct validity, 90% (9/10 items) of the results were in accordance with the hypotheses.

Two of three results were in accordance with the hypotheses. Only (a) Disease stage; patients with stage 0–II would have significantly superior scores than those in stages III–IV; "Worries about others" on EORTC QLQ-ELD14 was not in accordance with the hypotheses.

Known-group validity is presented in Tables 4 and 5. All scales having comorbidities and the G8, except the cancer stage of the EORTC QLQ-ELD14, showed known-group validity in both age groups. Analyses for the clinical cancer stage showed no significant differences among the EORTC QLQ-ELD14 scales. Patients with comorbidities had significantly worse EORTC QLQ-ELD14 scales except for the item "family support" in the ≥ 60 years age group. The ≥ 70 years age group had significantly inferior EORTC QLQ-ELD14 scales except for items such as "family support", "worries about others", and "future worries". According to the G8, frail patients had significantly worse EORTC QLQ-ELD14 scale scores except for the item "family support" in both groups. The most significantly different scale was "burden of illness", 19.42 ± 20.40 (not frail) vs. 37.89 ± 26.66 (frail) in the ≥ 70 years age group.

All seven results were in accordance with the hypotheses.

Table 6 shows the concurrent validity between the QLQ-C30 and the QLQ-ELD14. Convergent validity was established between "mobility" and "physical functioning" ($r = -0.797$, ≥ 60 years age group; $r = -0.799$, ≥ 70 years age group), between "mobility" and "role functioning"

($r = -0.797$, 60 years age group; $r = -0.799$, ≥ 70 years age group), and "mobility" and "fatigue" ($r = -0.688$, 60 years age group; $r = -0.714$, ≥ 70 years age group). Convergent validity was established between "future worries" and "emotional functioning" ($r = -0.514$, ≥ 60 years age group; $r = -0.519$, ≥ 70 years age group), and between "future worries" and "social functioning" ($r = -0.425$, ≥ 60 years age group; $r = -0.419$, ≥ 70 years age group).

Convergent validity was established between "burden of illness" and "social functioning" ($r = -0.498$, ≥ 60 years age group; $r = -0.487$, ≥ 70 years age group) and between "burden of illness" and "financial difficulties" ($r = -0.471$, ≥ 60 years age group; $r = -0.444$, ≥ 70 years age group).

Responsiveness (responsiveness to change analysis, RCA)

Hypotheses to evaluate responsiveness are presented in Supplementary Table S2. The results were confirmed if 75% of the results were in accordance with the hypotheses, and the testing and responsiveness results. Our hypotheses were accepted by approximately 80% (4/5 items). Responsiveness has been calculated and presented in Table 7. Our hypothesis was that "maintaining purpose" and "future worries" of the EORTC QLQ ELD-14 would worsen after the onset of the COVID-19 pandemic compared to before the pandemic. The indicators of responsiveness were calculated by ES and SRM. ES and SRM are unrelated to sample size. In our results, "maintaining purpose", "future worries" (in both groups), and "burden of illness" (≥ 60 years age group) were significantly worse ($p < 0.05$), but ES and SRM were very low (SRM < 0.2) except for the item "maintaining purpose" (SRM = 0.21 for both groups).

Table 4 Known-group comparison using stage and comorbidities

EORTC QLQ-ELD14	Comorbidities																			
	UICC Stage				≥ 60 years (N = 1800)				≥ 70 years (N = 123)				≥ 60 years (N = 1803)				≥ 70 years (N = 1236)			
	0, I, II	III, IV	Mean ± SD	p-value	0, I, II	III, IV	Mean ± SD	p-value	0	≤ 1	Mean ± SD	p-value	0	≤ 1	Mean ± SD	p-value	0	≤ 1	Mean ± SD	p-value
	n = 1266	n = 534			n = 865	n = 369			n = 1038	n = 765			n = 671	n = 565						
	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
<i>Functional scales/item</i>																				
Maintaining purpose	40.7 ± 24.5	41.8 ± 24.9	0.518	39.8 ± 24.4	40.3 ± 23.3	0.207	42.5 ± 24.4	38.9 ± 24.8	0.003	41.6 ± 23.7	38.0 ± 24.5	0.013	29.1 ± 31.1	31.0 ± 31.1	0.397	29.1 ± 31.1	31.0 ± 31.1	0.284		
Family support	29.6 ± 31.2	28.7 ± 31.2	0.525	29.1 ± 30.8	32.1 ± 31.8	0.742	28.7 ± 31.3	30.1 ± 31.1												
<i>Symptom scales/items</i>																				
Mobility	16.9 ± 21.0	18.7 ± 21.3	0.198	19.2 ± 22.1	19.5 ± 22.2	0.225	15.2 ± 19.6	20.5 ± 23.0	< 0.001	17.5 ± 20.9	22.5 ± 23.8	< 0.001	20.8 ± 24.0	23.5 ± 26.0	0.002	20.8 ± 24.0	23.5 ± 26.0	0.029		
Worries about others	30.0 ± 23.7	30.1 ± 24.3	0.656	29.9 ± 24.3	33.5 ± 23.6	0.123	29.3 ± 23.7	31.1 ± 24.0	0.049	30.5 ± 24.1	31.6 ± 24.2	0.215	30.0 ± 24.8	36.0 ± 25.2	0.048	35.1 ± 24.2	36.0 ± 25.2	0.374		
Future worries	34.2 ± 24.3	33.0 ± 24.0	0.164	35.0 ± 24.6	36.8 ± 24.6	0.752	33.7 ± 24.0	36.0 ± 25.0	0.048	35.1 ± 24.2	36.0 ± 25.2	0.374	30.0 ± 24.8	36.5 ± 25.8	< 0.001	30.0 ± 24.8	36.5 ± 25.8	< 0.001		
Burden of illness	31.4 ± 25.1	32.3 ± 26.1	0.290	31.4 ± 24.9	36.5 ± 26.3	0.173	29.0 ± 24.5	35.4 ± 26.0	< 0.001	30.0 ± 24.8	36.5 ± 25.8	< 0.001	20.8 ± 24.0	23.5 ± 26.0	0.002	20.8 ± 24.0	23.5 ± 26.0	0.029		
Joint stiffness	20.7 ± 24.7	22.5 ± 26.3	0.084	22.2 ± 25.1	21.8 ± 24.7	0.977	19.8 ± 24.0	23.3 ± 26.8												

SD, standard deviation; UICC, Union for International Cancer Control

The largest change in the responsiveness score was for the item “maintaining purpose” of the EORTC QLQ ELD-14; it worsened from 40.97 ± 24.26 to 35.00 ± 25.58 (change score: 6.14 ± 20.20) between before and after the onset of the COVID-19 pandemic ($ES = 0.26$, $SRM = 0.21$, $p < 0.001$). The responsiveness values of “global health status/QOL” on the EORTC QLQ-C30’s were $SRM = 0.09$ (≥ 60 years age group) and $SRM = 0.13$ (≥ 70 years age group), and these were smaller than the values of the item “maintaining purpose” of the EORTC QLQ ELD-14.

Table 8 shows the correlations with the changes between the EORTC QLQ-C30 and EORTC QLQ-ELD14 after the onset of the COVID-19 pandemic, between the items. “Mobility” on the EORTC QLQ-ELD14 scale and “global health status/QOL” and “summary score” on the EORTC QLQ-C30 scale had a moderate-to-high negative correlation of ≥ -0.40 (“global health status/QOL” -0.730 and “summary score” -0.571 for the ≥ 60 years age group; “global health status/QOL” -0.599 and “summary score” -0.737 for the ≥ 70 years age group). Between the items “future worries” on the EORTC QLQ-ELD14 and “global health status/QOL” and “summary score” on the EORTC QLQ-C30 there was a moderate-to-high negative correlation of ≥ -0.40 (“global health status/QOL” -0.550 and “summary score” -0.489 for the ≥ 60 years age group; “global health status/QOL” -0.502 and “summary score” -0.544 for the ≥ 70 years age group).

There was a moderate-to-high negative correlation (≥ -0.40) between the “burden of illness” on the EORTC QLQ-ELD14 and the “global health status/QOL” and “summary score” on the EORTC QLQ-C30 (“global health status/QOL” -0.604 and “summary score” -0.526 for the ≥ 60 years age group; “global health status/QOL” -0.535 and “summary score” -0.605 for the ≥ 70 years age group). Outside of our hypotheses, there was a moderate-to-high negative correlation (≥ -0.40) between “worries about others” and “joint stiffness”, but not for “maintaining purpose” and “family support” of the EORTC QLQ-ELD14 and “global health status/QOL” and “summary score” of the EORTC QLQ-C30.

Discussion

The purpose of this study was to evaluate the reliability, validity, and responsiveness of the EORTC QLQ-ELD14 for cross-cultural adaptation to elderly Japanese patients with cancer. The Japanese version of the EORTC QLQ-ELD14 questionnaire is considered to be an acceptable, reliable, and effective tool for assessing the HRQOL of elderly Japanese patients with cancer. A study of the original version, EORTC QLQ-ELD15, described the definition of a study population aged ≥ 70 years as “elderly”

Table 5 Known-group comparison using the G8 score

EORTC QLQ-ELD14	G8 score					
	≥ 60 years (N= 1215)			≥ 70 years (N= 827)		
	Normal (N=540)	Abnormal (N=675)	P-value	Normal (N=346)	Abnormal (N=481)	P-value
	Mean ± SD	Mean ± SD		Mean ± SD	Mean ± SD	
<i>Functional scale/item</i>						
Maintaining purpose	36.79 ± 27.01	35.82 ± 24.71	0.001	35.79 ± 26.96	35.15 ± 24.28	0.009
Family support	20.00 ± 27.69	31.17 ± 31.18	0.228	20.47 ± 27.55	33.26 ± 31.05	0.974
<i>Symptom scales/items</i>						
Mobility	9.72 ± 12.40	22.81 ± 23.78	<0.001	11.16 ± 13.42	25.88 ± 24.73	<0.001
Worries about others	24.10 ± 19.78	32.36 ± 24.12	<0.001	25.05 ± 19.71	34.49 ± 24.95	<0.001
Future worries	27.37 ± 20.03	39.05 ± 25.71	<0.001	27.99 ± 20.93	41.38 ± 26.48	<0.001
Burden of illness	18.89 ± 19.68	35.22 ± 26.25	<0.001	19.42 ± 20.40	37.89 ± 26.66	<0.001
Joint stiffness	15.30 ± 18.73	24.86 ± 26.46	<0.001	15.12 ± 18.28	25.47 ± 25.89	<0.001

SD, standard deviation

as a potential limitation (Johnson et al. 2010). Moreover, the definition of “elderly” varies and it also differs on an individual level. Therefore, in this study, we grouped the study population to observe differences between those over 60 years of age and those over 70 years of age.

In our study, the five most common cancer types, among the Japanese population, were colorectal, gastric, lung, breast, and prostate. From a study of 10 countries with populations aged ≥ 70 years, Wheelwright et al. (2013) found additional types of cancer, including ovarian, hematological, and others. When compared to the Korean study (Goo et al. 2017), except for breast and prostate cancer, the clinical background of the patients was similar to that of our study. Moreover, in our study, EORTC QLQ-ELD14 scores were similar to that of Korean studies, except for the item “family support”. For this item, where participants were asked, “Have you felt able to talk to your family about your illness?”, the result was inferior in elderly Japanese people. The results were 29.20 ± 31.16 in Japanese patients aged ≥ 60 years, 29.90 ± 31.06 in Japanese patients aged ≥ 70 years, 45.9 ± 33.7 in Korean patients aged ≥ 60 years, and 66.2 to 77.8 (based on the Eastern Cooperative Oncology Group [ECOG] score) in patients aged ≥ 70 years in the study of 10 countries, most of which were in Europe. Based on this trend, social background and cultural difference may be one of the points to investigate in future studies with elderly Japanese people with cancer.

Depending on the EORTC QLQ-C30 reference values, for all cancer patients at all stages and all ages, the global health status/QOL’s mean ± SD was 61.3 ± 24.2; and for all cancer patients aged ≥ 70 years it was 60.6 ± 25.1, so our data may be clinically superior to that of the EORTC QLQ-C30 reference values.

Reliability

Internal consistency—Cronbach α

Internal consistency (Cronbach α) was at a similar level to that of other language versions. The internal consistency of the Japanese version had a Cronbach α ≥ 0.7. However, the Cronbach α value for the item “maintaining purpose” in Japanese patients aged ≥ 60 years and ≥ 70 years was 0.65 and 0.62, respectively. It was similar to the results of the 10 countries study (Cronbach α = 0.68 for individuals aged ≥ 70 years), but lower than that of the Korean population aged ≥ 60 years, where Cronbach α was good (0.79). Next, for the item “worries about others” in the Japanese population aged ≥ 60 years and in those aged ≥ 70 years, Cronbach α was 0.67 and 0.68, respectively. It was similar to that of the Korean study (Goo et al. 2017), where Cronbach α was 0.65 (age ≥ 60 years), but not to the original study of the 10 countries, where Cronbach α was good (0.72) (age ≥ 70 years). “Maintaining purpose” and “worries about others” consisted of only two items, and we evaluated them to demonstrate acceptable internal consistency (Cronbach α).

Reliability-ICC

With regards to ICCs, patients with prostate cancer (clinical stages I–III; 5-year survival rate of 100%; patients were stable during the study period) had an ICC of ≥ 0.70 for the test–retest analysis, which was considered good. ICCs for the scales were 0.80 for the ≥ 60 years old age group, which was better than that of the ≥ 70 years age group at 0.70. In the study of 10 countries (Wheelwright et al. 2013), after 1 week, the ICCs were low for “burden of illness” and

Table 6 Correlations between the EORTC QLQ-C30 and EORTC QLQ-ELD14

	≥ 60 years (N = 1760)						≥ 70 years (N = 1205)							
	Functional scale/item			Symptom scales/items			Functional scale/item			Symptom scales/items				
	Maintain- ing purpose	Family sup- port	Mobility	Worries about oth- ers	Future wor- ries	Burden of illness	Joint stiff- ness	Mobility	Worries about oth- ers	Future wor- ries	Burden of illness	Joint stiffness		
Global health status/QOL	0.142	-0.231	-0.515	-0.345	-0.428	-0.488	-0.374	0.125	-0.236	-0.526	-0.344	-0.424	-0.498	-0.374
Summary Score	0.047	-0.306	-0.733	-0.437	-0.528	-0.571	-0.543	0.032	-0.332	-0.740	-0.439	-0.520	-0.580	-0.539
<i>Functional scale/item</i>	0.047	-0.298	-0.797	-0.373	-0.407	-0.495	-0.504	0.021	-0.331	-0.799	-0.372	-0.396	-0.502	-0.508
Physical functioning	0.058	-0.290	-0.688	-0.365	-0.410	-0.506	-0.444	0.055	-0.330	-0.714	-0.373	-0.392	-0.509	-0.453
Emotional functioning	0.026	-0.204	-0.459	-0.409	-0.514	-0.431	-0.379	0.017	-0.226	-0.459	-0.410	-0.519	-0.451	-0.354
Cognitive functioning	0.020	-0.192	-0.485	-0.322	-0.403	-0.344	-0.369	-0.003	-0.234	-0.479	-0.335	-0.408	-0.351	-0.335
Social functioning	0.023	-0.260	-0.523	-0.396	-0.425	-0.498	-0.337	0.019	-0.294	-0.519	-0.397	-0.410	-0.487	-0.316
<i>Symptom scales/items</i>														
Fatigue	-0.021	0.300	0.676	0.396	0.483	0.514	0.485	-0.016	0.316	0.684	0.388	0.481	0.524	0.469
Nausea or vomiting	-0.015	0.136	0.240	0.146	0.179	0.192	0.153	-0.007	0.150	0.252	0.151	0.192	0.202	0.166
Pain	0.022	0.228	0.550	0.300	0.354	0.401	0.540	0.050	0.245	0.558	0.306	0.353	0.418	0.537
Dyspnea	-0.046	0.156	0.463	0.242	0.285	0.333	0.281	-0.020	0.161	0.444	0.227	0.285	0.339	0.265
Insomnia	-0.051	0.145	0.355	0.239	0.317	0.308	0.308	-0.042	0.142	0.344	0.228	0.302	0.297	0.311
Loss of appetite	-0.076	0.170	0.389	0.240	0.298	0.345	0.286	-0.075	0.158	0.398	0.242	0.302	0.364	0.299
Constipation	-0.017	0.119	0.303	0.217	0.266	0.245	0.225	-0.008	0.141	0.307	0.244	0.262	0.237	0.254
Diarrhea	0.001	0.146	0.209	0.162	0.192	0.209	0.179	0.004	0.177	0.239	0.196	0.213	0.251	0.204
Financial difficulties	-0.010	0.182	0.405	0.352	0.389	0.471	0.280	-0.003	0.192	0.396	0.333	0.364	0.444	0.267

For bold values, correlation is significant at the 0.05 level (2-tailed)

QOL: quality of life

Table 7 Responsiveness of the EORTC QLQ-C30 (Global health status/QOL) and EORTC QLQ-ELD14

	T1 Mean ± SD	T3 Mean ± SD	Change Mean ± SD	Change 95% CI	p-value	ES	SRM
<i>≥ 60 years (N = 1287)</i>							
<i>EORTC QLQ-C30</i>							
Global health status/QOL	73.75 ± 20.21	71.97 ± 20.64	1.85 ± 20.18	0.68, 2.88	0.002	0.09	0.09
Summary score	87.37 ± 11.23	86.21 ± 12.19	1.16 ± 8.55	0.68, 1.63	<0.001	0.10	0.14
<i>EORTC QLQ-ELD14</i>							
<i>Functional scale/item</i>							
Maintaining purpose	41.89 ± 24.92	36.00 ± 25.81	6.12 ± 29.02	4.30, 7.49	<0.001	0.25	0.21
Family support	27.94 ± 30.91	26.19 ± 29.94	1.88 ± 32.16	-0.04, 3.55	0.056	0.06	0.06
<i>Symptom scales/items</i>							
Mobility	15.55 ± 19.81	17.26 ± 20.66	-1.55 ± 15.34	-2.55, -0.89	<0.001	-0.08	-0.10
Worries about others	28.75 ± 23.21	28.54 ± 22.56	0.47 ± 22.79	-1.03, 1.45	0.739	0.02	0.02
Future worries	33.16 ± 23.56	33.82 ± 23.91	-0.45 ± 20.76	-1.78, 0.47	0.253	-0.02	-0.02
Burden of illness	29.49 ± 24.26	28.06 ± 25.01	1.59 ± 22.92	0.17, 2.68	0.026	0.07	0.07
Joint stiffness	20.40 ± 24.67	20.84 ± 24.02	-0.13 ± 23.75	-1.74, 0.85	0.498	-0.01	-0.01
<i>≥ 70 years (N = 888)</i>							
<i>EORTC QLQ-C30</i>							
Global health status/QOL	73.22 ± 20.66	71.13 ± 21.13	2.09 ± 20.55	0.73, 3.45	0.003	0.09	0.13
Summary score	86.69 ± 11.63	85.34 ± 12.74	1.34 ± 8.85	0.74, 1.94	<0.001	0.12	0.15
<i>EORTC QLQ-ELD14</i>							
<i>Functional scale/item</i>							
Maintaining purpose	40.97 ± 24.26	35.00 ± 25.58	6.14 ± 29.20	4.04, 7.91	<0.001	0.26	0.21
Family support	28.30 ± 30.45	27.75 ± 30.00	0.61 ± 31.83	-1.59, 2.69	0.614	0.02	0.02
<i>Symptom scales/items</i>							
Mobility	17.57 ± 20.80	20.00 ± 21.97	-2.27 ± 16.14	-3.49, -1.38	<0.001	-0.11	-0.14
Worries about others	29.66 ± 23.47	30.21 ± 23.25	-0.28 ± 22.99	-2.06, 0.95	0.470	-0.01	-0.01
Future worries	34.03 ± 23.85	35.61 ± 25.01	-1.45 ± 21.11	-2.96, -0.21	0.024	-0.06	-0.07
Burden of illness	30.98 ± 24.50	30.23 ± 26.03	0.85 ± 23.59	-0.81, 2.30	0.347	0.04	0.04
Joint stiffness	20.77 ± 24.13	21.34 ± 23.87	-0.34 ± 23.77	-2.14, 0.99	0.472	-0.02	-0.02

SD, standard deviation; ES, effect size (mean change score/SD baseline); SEM, standardized response mean (mean change score/SD change score); QOL, quality of life

Table 8 Relationship between the EORTC QLQ-C30 (Global health status/QOL and summary score) and EORTC QLQ-ELD14 under COVID-19

	<i>≥ 60 years (N = 1760)</i>		<i>≥ 70 years (N = 1205)</i>	
	Global health status/QOL	Summary Score	Global health status/QOL	Summary Score
<i>EORTC QLQ-ELD14</i>				
<i>Functional scale/item</i>				
Maintaining purpose	-0.066*	-0.022	-0.060	-0.101**
Family support	-0.355**	-0.266**	-0.312**	-0.391**
<i>Symptom scales/items</i>				
Mobility	-0.730**	-0.571**	-0.599**	-0.737**
Worries about others	-0.441**	-0.333**	-0.344**	-0.426**
Future worries	-0.550**	-0.489**	-0.502**	-0.544**
Burden of illness	-0.604**	-0.526**	-0.535**	-0.605**
Joint stiffness	-0.526**	-0.415**	-0.432**	-0.544**

QOL, quality of life

**Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

“family support” in patients aged ≥ 70 years. We agree that EORTC QLQ-ELD14 may be related to both cancer evaluation and non-clinical events.

Standard error of measurement (SEM)

The SEM was not analyzed in the following studies: ten countries of Europe (Wheelwright et al. 2013), Korea (Goo et al. 2017), Spain (Wrazen et al. 2014), Spain (Arraras et al. 2019), and Chile (Lorca et al. 2021). This study clarified that all the scale/items of the SEM and SDC in the ≥ 70 years age group were larger than those for the ≥ 60 years age group, especially “family support” and “maintaining purpose” on the functional scale/items. Moreover, the SEMs/SDCs of all the scales on the EORTC QLQ-ELD14 were less than 1, indicating good quality.

Validity

Construct validity

As we were unable to perform a factor analysis, we carried out a multi-trait scaling analysis, and it proved to be a good construct, except for some scales. For the item “worries about others,” other scale correlations were higher than the item’s scale correlation. Item 7 in the “worries about others” (i.e., “have you worried about the future of people who are important to you?”) correlated more highly with item 8 in the “future worries” subscale ($r=0.63$, ≥ 60 years age group, $r=0.62$, ≥ 70 years age group). This may be because the word “future” is shared in items 7, 8, and 9. It was similar to the results of the Korean study (Goo et al. 2017).

Criterion validity

We concluded that we had consistent findings for sufficient hypotheses to test for construct validity, as 90% of the results were in accordance with the hypotheses. We found that the hypotheses were based on literature and on theoretical considerations.

Known-group comparisons were demonstrated by the ability to distinguish subgroups of patients with different clinical profiles. This study demonstrated the clinical validity of the EORTC QLQ-ELD14 by its ability to distinguish between different comorbidities, and the G8 score, in particular.

The ≥ 60 years age group had superior scores for all scales, except “family support”, compared to the ≥ 70 years age group. In our study hypotheses, similar to that of the Korean study (Goo et al. 2017), clinical stages III and IV patients were assumed to have more worries about others. However, this hypothesis was not supported in this study.

Patients with comorbidities in the ≥ 60 years age group had significantly inferior scores except for “family support”. Patients with comorbidities in the ≥ 70 years old age group had significantly inferior scores except for “family support”, “worries about others”, and “future worries”. Moreover, our results indicated that frailty, identified by G8 scores, was strongly associated with poor EORTC QLQ-ELD14 scale results. Both patient groups with frailty (≥ 60 and ≥ 70 years age groups) had significantly inferior scores except for “family support”. The QLQ-ELD14 identified patient groups defined by comorbidities and G8 scores.

Relationship between the EORTC QLQ-C30 and EORTC QLQ-ELD14

The divergent validity of the EORTC QLQ-ELD14 was assessed by evaluating the correlations between its subscales and that of the EORTC QLQ-C30. Criterion validity was demonstrated by the lack of a strong correlation between the EORTC QLQ-ELD14 and the C30, except for “mobility” (ELD14) and “physical functioning” (C30); this was similar to the results of other studies (Michelson et al. 2000; Wrazen et al. 2014; Goo et al. 2017; Arraras et al. 2019). The hypotheses were based on the literature and theoretical considerations. We assumed a moderate to strong correlation between “mobility” and “physical functioning”; between future anxiety, “future worries” and “emotional functioning”; and between “burden of illness” and “social functioning”. This was similar to other studies (Michelson et al. 2000; Wrazen et al. 2014; Goo et al. 2017; Arraras et al. 2019). Our current Japanese study had a stronger correlation ($r=-0.797$ and -0.799 for the ≥ 60 and ≥ 70 years age groups, respectively) between “mobility” (ELD14) and “physical functioning” (C30); this was similar to the Korean study; $r=-0.72$ for the ≥ 60 years age group. “Family support”, “global health status/QOL”, and the functioning scales/items of the EORTC QLQ-C30 did not have a positive correlation, and these trends were similar to those of a previous Korean study (Goo et al. 2017). This may be attributed to differences in how the questions were perceived.

The results of criterion validity with the EORTC QLQ-C30 indicate that HRQOL issues assessed by the QLQ-ELD14 are different from those assessed by the commonly used QLQ-C30. Our results indicated that the Japanese version of the QLQ-ELD14 is a useful tool for assessing the HRQOL of elderly Japanese people with cancer.

Responsiveness: responsiveness to change analysis (RCA)

We demonstrated consistent findings on sufficient hypothesis testing for responsiveness, as 80% of the results were in accordance with the hypotheses. To determine the sensitivity

to change, the paired-*t* test, ES, and SRM were performed. We analyzed the RCA ($n = 1302$; ≥ 60 years age group, $n = 888$; ≥ 70 years age group) one year later.

Our hypothesis was that the items of “maintaining purpose” and “future worries” of the EORTC QLQ ELD-14 would have worsened after the commencement of the COVID-19 pandemic, compared with the results prior to the pandemic. In our Japanese study, patients had significantly decreased scores (SRM, 0.21 for the ≥ 60 years age group and 0.26 for the ≥ 70 years age group) for the “maintaining purpose” item after the commencement of the COVID-19 pandemic, compared with before. Contrastingly, although items such as “burden of illness” and “joint stiffness” showed a significant difference in the scores, we found no clinical difference because the SRM was very low. However, the *t* change mean score was very low (SRM < 0.2), except for the item “maintaining purpose”. As mentioned earlier, the concept of HRQOL in elderly patients is not linked to any specific medical condition and is a broad concept (Wilson and Cleary 1995). In our study, “maintaining purpose” on the EORTC QLQ-ELD14 and the “global health status/QOL” and “summary score” on the EORTC QLQ C30 had very low correlations to score changes. It seems to be due to different conceptualizations of “maintaining purpose”. In the original study in the 10 countries, the item “maintaining purpose” is complex and covers two concepts: positive outlook and motivation for activities (Wheelwright et al. 2013). Another prospective study using the EORTC QLQ-ELD14 in radiotherapy patients ≥ 80 years, six months after radiotherapy, a significant and clinically relevant deterioration of HRQOL was seen in the EORTC QLQ-ELD14 for “future worries”, “burden of illness”, and “family support”. For the EORTC QLQ-ELD14 questionnaire, older head and neck cancer patients, up to 1 year after curative treatment, reported greater difficulty in “maintaining purpose” 12 months after treatment. Including these points, further studies are needed to evaluate the responsiveness of this scale for elderly Japanese patients with cancer under various circumstances.

In previous studies regarding responsiveness, the study in 10 countries (Wheelwright et al. 2013) found no change in the score of the progress stage group three months later, in the ≥ 70 years age group. The studies in Korea (Goo et al. 2017), Spain (Wrazen et al. 2014), Spain (Arraras et al. 2019), and Chile (Lorca et al. 2021) did not analyze responsiveness data. When assessing responsiveness, one of the most difficult tasks is formulating challenging hypotheses (Mokkink et al., 2018a, b). When we planned the hypotheses of this study, there was no available data on evaluating HRQOL change in people with cancer, especially the elderly. Subsequently, we found two recent studies (Jeppesen et al. 2021; Koinig et al. 2021) that compared pre-COVID-19 pandemic and intra-pandemic

scores and revealed that the “global health status/QOL” of the EORTC QLQ-C30 did not change significantly with the pandemic, and is considered clinically important. The survey (Jeppesen et al. 2021) was published in Denmark in May 2020, with 4571 respondents (mean 66 years, 40% male, 30% breast cancer, 22% intractable cancer). The average “global health status/QOL” score was 71.3 points, which is the same as that of our present study. There was no clinically significant difference in the HRQOL score compared to that before the COVID-19 pandemic.

Another study (Koinig et al. 2021) found that the COVID-19 pandemic only had a slight impact on the “global health status/QOL” (mean global health status/QOL score, $\Delta = -1.95$); this was similar to our study findings ($\Delta = -1.85$ for the ≥ 60 years age group and -2.09 for the ≥ 70 years age group), where all the values were below the threshold for clinically meaningful differences. Therefore, we suggest that our hypotheses might not have been adequate to evaluate the responsiveness of the EORTC QLQ-ELD14.

Our study had certain limitations. Our study participants came from a single hospital with better HRQOL. The G8 frailty score was self-assessed by the elderly participants. We could not use factor analysis to evaluate construct validity or the maximal information coefficient (MIC) to analyze measurement error or to evaluate MIC. Moreover, we were unable to analyze cross-cultural validity by COSMIN methods. Responsiveness was difficult to assess, and so was creating a hypothesis for the effect of the COVID-19 pandemic without reference data. The original version stated that non-clinical factors could affect older patients more than younger ones. Future studies will need to explore this hypothesis in Japan. We propose that HRQOL in older individuals is a wide concept. Japan has the world’s most rapidly aging population. Therefore, the contribution of large datasets to create evidence to aid elderly patients with cancer is necessary.

The strength of our study lies in the fact that this scale seems to be usable even for elderly individuals with cancer aged over 60 and 70 years in the Japanese population. The EORTC QLQ-ELD14 can evaluate generic issues affecting older adults with cancer not covered by the EORTC QLQ-C30 or site-specific modules. Moreover, the low number of questions per scale in the EORTC QLQ-ELD14 helps reduce the burden on older adults. Nonetheless, we collected extensive data despite the adverse circumstances, and for the first time, we could demonstrate the effect of the COVID-19 pandemic on HRQOL using the EORTC QLQ-ELD14, with the cooperation of elderly patients with cancer. In the future, using the EORTC QLQ-ELD14 to evaluate comparisons in many research areas and clinical settings may provide clinically useful conclusions on various relevant topics.

Conclusion

Our findings show that the Japanese version of the EORTC QLQ-ELD14 is an acceptable tool that can be used for cancer research and clinical management to assess the HRQOL of elderly Japanese patients with cancer. Moreover, this scale can also be used for individuals aged over 60 and 70 years. Further, the EORTC QLQ-ELD14 showed worsening of the “maintaining purpose” parameter after the onset of the COVID-19 pandemic compared with before. Elderly people with cancer tend to prefer a better HRQOL over an increased span of life.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00432-022-04414-2>.

Acknowledgements We are grateful to all the patients that participated in the study. We would like to also thank the EORTC Translation Team. This work was funded by the Japan Society for the Promotion of Science KAKENHI (18K10312 and 19K22764), the Foundation for Promotion of Cancer Research, and the SGH Foundation.

Author contributions All authors contributed to the study conception and design. Material preparation was performed by YK, RI, HF, KA, SN, SA, TT, MK, JI, and KO, and KT. Data collection was performed by YK, RI, and MK. Analyses were performed by YK, JK, and the Editage Team. The first draft of the manuscript was written by YK and all the authors commented on previous versions of the manuscript. All the authors read and approved the final manuscript.

Funding This work was supported by the Japan Society for the Promotion of Science KAKENHI (18K10312 and 19K22764), the Foundation for Promotion of Cancer Research, and the SGH Foundation.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest All the authors declare that there are no conflicts of interest relevant to this article.

Consent to participate Prior to the study, participants were asked to sign a written informed consent form to confirm their willingness to participate in this study.

Consent to publish Not applicable.

Ethics approval This study was approved by the Ethics Committee of Kyushu University (2019-450).

References

Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC (1993) The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use

- in international clinical trials in oncology. *J Natl Cancer Inst* 85:365–376. <https://doi.org/10.1093/jnci/85.5.365>
- Arraras JI, Asin G, Illarramendi JJ, Manterola A, Salgado E, Dominguez MA (2019) The EORTC QLQ-ELD14 questionnaire for elderly cancer patients. Validation study for elderly Spanish breast cancer patients. *Rev Esp Geriatr Gerontol* 54:321–328. <https://doi.org/10.1016/j.regg.2019.05.001>
- Beaton DE, Bombardier C, Katz JN, Wright JG (2001) A taxonomy for responsiveness. *J Clin Epidemiol* 54:1204–1217. [https://doi.org/10.1016/S0895-4356\(01\)00407-3](https://doi.org/10.1016/S0895-4356(01)00407-3)
- Bellera CA, Rainfray M, Mathoulin-Pélessier S, Mertens C, Delva F, Fonck M, Soubeyran PL (2012) Screening older cancer patients: first evaluation of the G-8 geriatric screening tool. *Ann Oncol* 23:2166–2172. <https://doi.org/10.1093/annonc/mdr587>
- Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16:297–334
- Fayers PM, Machin D (2000) Quality of life: assessment analysis and interpretation. Wiley, Chichester
- Fayers PM, Machin D (2013) Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. Wiley, Chichester
- Fitzsimmons D, Gilbert J, Howse F, Young T, Arraras JI, Brédart A, Hawker S, George S, Aapro M, Johnson CD (2009) A systematic review of the use and validation of health-related quality of life instruments in older cancer patients. *Eur J Cancer* 45:19–32. <https://doi.org/10.1016/j.ejca.2008.07.036>
- Goo AJ, Shin DW, Yang HK, Park JH, Kim SY, Shin JY, Kim YA, Kim C, Hong NS, Min YJ, Park KJ (2017) Cross-cultural application of the Korean version of the EORTC QLQ-ELD14 questionnaire for elderly patients with cancer. *J Geriatr Oncol* 8:271–276. <https://doi.org/10.1016/j.jgo.2017.03.001>
- Hjermstad MJ, Fayers PM, Bjordal K, Kaasa S (1998) Using reference data on quality of life—the importance of adjusting for age and gender, exemplified by the EORTC QLQ-C30 (+3). *Eur J Cancer* 34:1381–1389. [https://doi.org/10.1016/s0959-8049\(98\)00136-1](https://doi.org/10.1016/s0959-8049(98)00136-1)
- Jeppesen SS, Bentsen KK, Jørgensen TL, Holm HS, Holst-Christensen L, Tarpgaard LS, Dahlrot RH, Eckhoff L (2021) Quality of life in patients with cancer during the COVID-19 pandemic—a Danish cross-sectional study (COPICADS). *Acta Oncol* 60:4–12. <https://doi.org/10.1080/0284186X.2020.1830169>
- Johnson C, Fitzsimmons D, Gilbert J, Arraras JI, Hammerlid E, Brédart A, Ozmen M, Dilektasli E, Coolbrandt A, Kenis C, Young T, Chow E, Venkitaraman R, Howse F, George S, O’Connor S, Yadegarfar G, EORTC Quality of Life Group (2010) Development of the European Organisation for Research and Treatment of Cancer quality of life questionnaire module for older people with cancer: the EORTC QLQ-ELD15. *Eur J Cancer* 46:2242–2252. <https://doi.org/10.1016/j.ejca.2010.04.014>
- Kasper B (2020) The EORTC QLQ-C30 Summary Score as a prognostic factor for survival of patients with cancer: a commentary. *Oncologist* 25:e610–e611. <https://doi.org/10.1634/theoncologist.2019-0749>
- Kobayashi K, Takeda F, Teramukai S, Gotoh I, Sakai H, Yoneda S, Noguchi Y, Ogasawara H, Yoshida K (1998) A cross-validation of the European Organization for Research and Treatment of Cancer QLQ-C30 (EORTC QLQ-C30) for Japanese with lung cancer. *Eur J Cancer* 34:810–815. [https://doi.org/10.1016/S0959-8049\(97\)00395-X](https://doi.org/10.1016/S0959-8049(97)00395-X)
- Koinig KA, Arnold C, Lehmann J, Giesinger J, Köck S, Willenbacher W, Weger R, Holzner B, Ganswindt U, Wolf D, Stauder R (2021) The cancer patient’s perspective of COVID-19-induced distress—a cross-sectional study and a longitudinal comparison of HRQOL assessed before and during the pandemic. *Cancer Med* 10:3928–3937. <https://doi.org/10.1002/cam4.3950>
- Lorca LA, Sacomori C, Fasce Pineda G, Vidal Labra R, Cavieres Faundes N, Plasser Troncoso J, Leao Ribeiro I (2021) Validation of the EORTC QLQ-ELD 14 questionnaire to assess the health-related

- quality of life of older cancer survivors in Chile. *J Geriatr Oncol* 12:844–847. <https://doi.org/10.1016/j.jgo.2020.12.014>
- Michelson H, Bolund C, Nilsson B, Brandberg Y (2000) Health-related quality of life measured by the EORTC QLQ-C30—reference values from a large sample of Swedish population. *Acta Oncol* 39:477–484. <https://doi.org/10.1080/028418600750013384>
- Mohile SG, Dale W, Somerfield MR, Schonberg MA, Boyd CM, Burhenn PS, Canin B, Cohen HJ, Holmes HM, Hopkins JO, Janelsins MC, Khorana AA, Klepin HD, Lichtman SM, Mustian KM, Tew WP, Hurria A (2018) Practical assessment and management of vulnerabilities in older patients receiving chemotherapy: ASCO guideline for geriatric oncology. *J Clin Oncol* 36:2326–2347. <https://doi.org/10.1200/JCO.2018.78.8687>
- Mokink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB (2018a) COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 27:1171–1179. <https://doi.org/10.1007/s11136-017-1765-4>
- Mokink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, de Vet HCW, Terwee CB (2018b) COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) user manual, version 1.0 (2018b). https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018b.pdf. Accessed 14 May 2022
- Osoba D, Rodrigues G, Myles J, Zee B, Pater J (1998) Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 16:139–144. <https://doi.org/10.1200/JCO.1998.16.1.139>
- Pilleron S, Sarfati D, Janssen-Heijnen M, Vignat J, Ferlay J, Bray F, Soerjomataram I (2019) Global cancer incidence in older adults, 2012 and 2035: a population-based study. *Int J Cancer* 144:49–58. <https://doi.org/10.1002/ijc.31664>
- Schwarz R, Hinz A (2001) Reference data for the quality of life questionnaire EORTC QLQ-C30 in the general German population. *Eur J Cancer* 37:1345–1351. [https://doi.org/10.1016/s0959-8049\(00\)00447-0](https://doi.org/10.1016/s0959-8049(00)00447-0)
- Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Tester S, Hubbard G, Downs M, McDonald C, Murphy J (2004) Growing older: quality of life in old age. In: Hennessy CH, Walker C, Walker A (eds) *Frailty and institutional life*. Open University Press, Maidenhead, pp 209–225
- Wheelwright S, Darlington AS, Fitzsimmons D, Fayers P, Arraras JL, Bonnetain F, Brain E, Bredart A, Chie WC, Giesinger J, Hammerlid E, O'Connor SJ, Oerlemans S, Pallis A, Reed M, Singhal N, Vassiliou V, Young T, Johnson C (2013) International validation of the EORTC QLQ-ELD14 questionnaire for assessment of health-related quality of life elderly patients with cancer. *Br J Cancer* 109(4):852–858. <https://doi.org/10.1038/bjc.2013.407>
- Wilson IB, Cleary PD (1995) Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 273:59–65. <https://doi.org/10.1001/jama.1995.03520250075037>
- World Health Organization (2021). *Aging Health*. <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> Accessed 12 May 2022
- Wrzen W, Golec EB, Tomaszewska IM, Walocha E, Dudkiewicz Z, Tomaszewski KA (2014) Preliminary psychometric validation of the Polish version of the EORTC elderly module (QLQ-ELD14). *Folia Med Cracov* 54(2):35–45
- Wright P, Smith A, Booth L, Winterbottom A, Kiely M, Velikova G, Selby P (2005) Psychosocial difficulties, deprivation and cancer: three questionnaire studies involving 609 cancer patients. *Br J Cancer* 93(6):622–626. <https://doi.org/10.1038/sj.bjc.6602777>
- Young T, De Haes H, Curran D et al (1999) EORTC guidelines for assessing quality of life in clinical trials. European Organization for Research and Treatment of Cancer Quality of Life Group Publication, Brussels, pp 2–930064

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.