Proxy responses regarding quality of life of patients with terminal lung cancer: preliminary results from a prospective observational study

Tomoyuki Takura,1 Tomoko Koike,2 Yoko Matsuo,3 Asuko Sekimoto,4 Masami Mutou5

ABSTRACT
Objective This prospective study used the EQ-5D utility and Visual Analogue Scale (VAS) scores to analyse the potential usefulness of proxy responses in quality of life assessments of Japanese patients with terminal lung cancer sufficiently healthy to communicate and reply by themselves. We did not investigate the potential usefulness of using proxy responses for patients who could not respond by themselves.

Design A prospective observational study.

Setting Single centre.

Participants The EQ-5D and VAS responses were gathered from 30 in-hospital patients with lung cancer for a total of three observation points. At nearly the same time, two nurses responded by providing proxy responses.

Primary and secondary outcome measures EQ-5D and VAS responses.

Results There were no significant differences between the patients’ and nurses’ responses for EQ-5D utility and VAS scores. For the five dimensions of the EQ-5D, significant differences were found between the patients’ and nurses’ responses for usual activities (patients’ response 1.64±0.07, nurses’ response 1.41±0.05, p=0.03) and anxiety/depression (patients’ response: 1.40±0.05, nurses’ response: 1.19±0.03, p=0.02). There was a significant weak positive correlation between patients’ and nurses’ responses regarding changes in responses from the first to the third observation point (Spearman’s rank correlation coefficient ρ=0.228; p<0.01).

Conclusion The results suggest that proxy responses are useful because there were no significant differences between the patients’ and nurses’ responses for EQ-5D utility and VAS scores at the three observation points. These findings should, however, be verified in future large-scale trials.

BACKGROUND
Lung cancer has become a leading cause of death in Japan, causing more deaths than gastric cancer since 1998. The incidence of lung cancer has been increasing steadily, a trend that is expected to continue. Although evidence-based clinical guidelines show that first-line treatment for patients with early-stage lung cancer is surgery interventions, the treatment outcomes have been insufficiently successful. Therefore, many patients continue to suffer from poor health, despite new medical technologies, procedures, medicines and nonsurgical interventions. Complications such as respiratory failure, malnutrition, anaemia, infectious diseases, cerebral hypertension and multiple organ failure occur in the terminal stage, for which palliative care remains the primary treatment. Several studies have evaluated the efficacy of palliative care using various methods, such as conducting surveying patients and patients’ families1 and using various health technology assessments and cost-effectiveness analyses.2 However, the quality of life (QoL) of patients with terminal cancer undergoing palliative care has not been sufficiently evaluated. The lack of empirical data has led to concerns that social and financial support for palliative care for patients with lung cancer might stagnate, spurring an interest in patient-reported outcome measures (PROMs) in palliative care.
While everyone agrees that healthcare professionals should provide excellent, compassionate care for terminal patients, there is less consensus on the nature and extent of that care. Limited healthcare and insurance resources necessitate prioritising clinical efficacy when selecting and combining therapies and procedures to maximise medical resources to provide the best care possible without placing an undue financial burden on patients. As the number of patients with lung cancer receiving end-of-life (EOL) care is increasing, so is the need for a cost–benefit performance analysis focusing on the ratio of utility obtained and the expense incurred to quantitatively clarify the clinical and economic value of palliative care for patients with lung cancer. In short, we need to measure patients with lung cancer health-related QoL (HRQoL).

HRQoL is indispensable for calculating patients’ quality-adjusted life-year, a generic economic measure of people’s disease burden, quality and quantity of life lived. In 1987, the EuroQol Group, an international network of multidisciplinary researchers, began developing the Euro Qol 5 Dimension (EQ-5D), a family of instruments to describe and value health. The EQ-5D has been used in previous studies to assess QoL in patients with lung cancer. However, many terminal patients need assistance to respond to questions, including those related to QoL, due to deterioration of their physical and cognitive functions. Thus, the need sometimes arises for proxy responses but with patients with conditions other than terminal lung cancer, such as dementia, whose acuity and urgency of the patients’ clinical needs significantly affected the ratios. Furthermore, other studies have used the EQ-5D to investigate proxy responses rather than family members for the following reasons: limited healthcare and insurance resources necessitate prioritising clinical efficacy when selecting and combining therapies and procedures to maximise medical resources to provide the best care possible without placing an undue financial burden on patients, as the number of patients with lung cancer receiving end-of-life (EOL) care is increasing, so is the need for a cost–benefit performance analysis focusing on the ratio of utility obtained and the expense incurred to quantitatively clarify the clinical and economic value of palliative care for patients with lung cancer. In short, we need to measure patients with lung cancer health-related QoL (HRQoL).

Some researchers have questioned whether proxy responses accurately reflect patients’ QoL. One study comparing the responses of palliative care patients and proxies reported that the proxies less frequently (39%) reported patients’ symptom distress than the patients’ themselves (61%), although the healthcare setting, diagnosis, and the acuity and urgency of the patients’ clinical needs significantly affected the ratios. Furthermore, Clapham et al reported that PROMs being increasingly recognised as feasible in most clinical scenarios in inpatient palliative care, including proxy reporting. Overall, however, that study used a different assessment scale and did not focus on patients with terminal lung cancer in palliative care.

In this study, we selected nurses as proxy respondents rather than family members for the following reasons: First, in other disease areas, there were a few studies in which nurses were selected as substitute respondents, and we used them only as a reference for observation and analysis. In Japan, the number of older adults living alone increases due to the declining birth rate and ageing population. There are a certain number of patients who do not have family support even at the treatment site. In addition, family backgrounds vary according to age, occupation, lifestyle, and time spent with the patient. We speculated that this causes variability in observations and effort required for survey training. In considering the disease characteristics and study objectives, regular observation was desirable in this study. Hence a survey
by a nurse with a transparent nursing system was deemed appropriate.

The patient inclusion criteria were patients who could answer the EQ-5D questionnaire, were in the EOL stage of stage IV cancer, were in the hospital for at least a week, and included patients receiving radiation and other cancer therapies. The exclusion criteria were: (1) patients with brain metastasis, (2) patients with underlying psychiatric disorders, (3) patients below 18 years of age and (4) patients who did not provide informed consent. Incidentally, all patients who were approached to participate in the study provided their consent for the same.

We collected patient background information and EQ-5D and VAS scores for all 30 patients who agreed to participate in the study. We obtained the EQ-5D and VAS scores by asking the patients during a routine afternoon temperature check, asking once a day for three consecutive days.

The inclusion criteria for the nurses (n=2) who participated in the study were as follows: (1) team leaders; (2) at least 5 years’ experience with patients with cancer at the EOL; and (3) identified by the department chief as having sufficient experience to measure the EQ-5D. The two participating nurses provided the proxy responses at nearly the same time as the patients. The nurses measured each dimension of the EQ-5D based on their observations; however, we allowed them to ask the patients about general pain and anxiety for their responses as part of their normal nursing management duties. These nurses provided the necessary care for the participating patients during their standard shift work.

The patients’ and nurses’ responses were evaluated using Wilcoxon’s rank-sum test and a correlation analysis using Spearman’s rank correlation. Consistency between patients’ and nurses’ responses was assessed using the κ coefficient (kappa statistic) and weighted kappa statistic using a cross-tabulation table. That is, the agreement rate between patients’ and nurses’ responses to the five dimensions of the EQ-5D was calculated using a weighted score. The scores were applied to the patients’ and nurses’ responses: 1.00 if the patient and nurse’s answers agreed, 0.75 if the scores differed by 1, and 0.00 if the scores varied by 2, and the weighted κ coefficient was calculated accordingly. Generally, if the κ coefficient was ≥0.6, the responses’ consistency was considered extremely high.

We set the statistical significance at p<0.05. Peyre et al found that replacing missing items in QoL questionnaires is often necessary. Applying their methodology to maintain the sample size, we supplemented the missing values using a single imputation method at random based on the average value of the missing observation time points and the weighted ratio of the answers by the patient or nurse to the average value of the previous or subsequent observation time point. Our statistical analysis was IBM SPSS Statistics for Windows, V.26.0 (IBM).

### Table 1 Patients’ demographic and clinical characteristics (excluding one patient who failed to respond)

|                          | N=29 |
|--------------------------|------|
| Sex (n)                  |      |
| Male                     | 23 (79.3%) |
| Female                   | 6 (20.7%) |
| Age (year)               | 67.3±8.2 |
| Mean±SD                  |      |
| Male                     | 65.7±7.8 |
| Female                   | 73.7±6.6 |
| Primary disease (n)      |      |
| Lung cancer (stage-IV)   | 29 (100%) |
| Treatment (n)            |      |
| Cancer chemotherapy      | 26 (89.7%) |
| Radiation therapy        | 0 (0%) |
| Type of care (n)         |      |
| Respiratory care         | 10 (34.5%) |
| Pain care                | 9 (31.0%) |
| Outcome (n)              |      |
| Discharge                | 24 (82.8%) |
| Changing hospital        | 2 (6.9%) |
| Death                    | 1 (3.4%) |
| Unknown                  | 2 (6.9%) |

### Results

Patients were enrolled between September 2013 and April 2014. A total of 30 patients with stage IV lung cancer who satisfied the study inclusion criteria were enrolled in the study. The patients’ background characteristics are shown in table 1. Background information was not available for one patient. Information was obtained for 29 patients, of whom 23 were men (79.3%) and 6 were women (20.7%), with a mean age of 67.3±8.2 years. The treatments which the patients received were cancer chemotherapy in 26 patients (89.7%), respiratory care in 10 patients (34.5%), and pain care in 9 patients (31.0%). Outcomes were discharged in 24 patients (82.8%), changing hospitals in 2 patients (6.9%), death in 1 patient (3.4%) and unknown outcome in 2 patients (6.9%). All the participating patients completed the target observation period. None of the participants left the questionnaire completely blank. According to the question item base, the missing values of EQ-5D were 1.3% in the patient population and 1.2% in the nurse population. The VAS scores were 4.4% and 2.2%, respectively.

Patient and public involvement
Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our study.
For the EQ-5D utility and VAS scores, we obtained responses from 30 patients and two nurses at three observation points. Figure 1 shows the mean EQ-5D utility scores based on the levels of the five dimensions. The mean EQ-5D utility scores of the patients’ responses showed a slight decrease from the first to the third time point (EQ-5D utility score, mean±SE: first time point, 0.77±0.03; second time point, 0.73±0.04; and third time point, 0.73±0.04). The nurses’ responses remained almost unchanged from the first to the third time point (EQ-5D utility score, mean±SE: first time point, 0.81±0.02; second time point, 0.81±0.03; and third time point, 0.81±0.02). The mean EQ-5D utility scores were higher for the nurses’ responses than the patients’ responses, and the changes in the nurses’ responses did not perfectly mirror the small changes in the patients’ responses. Nevertheless, we found no significant differences between the groups at each observation time point (first time, p=0.36; second time, p=0.11; third time, p=0.16).

The mean VAS score is shown in Figure 2. The mean VAS scores of the patients’ responses remained almost unchanged from the first to the third time point (VAS scores, mean±SE: first time, 65.1±4.0; the second time, 64.6±3.3; the third time, 65.0±3.7), but the nurses’ responses decreased slightly from the first time point to the third time point (VAS scores, mean±SE: first time, 70.1±2.1; the second time, 71.3±2.1; the third time, 67.9±2.1). Although the VAS scores were higher for the nurses’ responses than the patients’ responses, there were no significant differences in the VAS scores between the groups at each observation point (first time, p=0.40; the second time, p=0.18; third time, p=0.74).

Subsequently, we compared the patients’ and nurses’ responses for each of the five dimensions of the EQ-5D. There were no significant differences in the following three dimensions: mobility (EQ-5D score, mean±SE: patients’ responses, 1.33±0.05; nurses’ responses, 1.30±0.04; p=0.83), self-care (patients’ responses, 1.27±0.05; nurses’ responses, 1.19±0.04; p=0.60), and pain/discomfort (patients’ responses, 1.51±0.06; nurses’ responses, 1.45±0.05; p=0.52). On the other hand, there were significant differences in two dimensions, usual activities (patients’ responses, 1.64±0.07; nurses’ responses, 1.41±0.05; p=0.03) and anxiety/depression (patients’ responses, 1.40±0.05; nurses’ responses, 1.19±0.03; p=0.02) (Figure 3). Significant differences between the patients’ and nurses’ responses were observed in the items relating to the daily activity level and emotional aspect.

Furthermore, to see the correlation between the responses of patients and nurses, the change in responses from the first to third observations for all EQ-5D five dimensions was confirmed. There was a significant weak positive correlation (Spearman’s rank correlation coefficient: p=0.228; p<0.01) (Figure 4). In addition, the agreement rate between patients’ and nurses’ responses to the five dimensions of the EQ-5D was 72.2% (Table 2). The weighted κ coefficient was 0.922 (92.2%).

**DISCUSSION**

This study examined whether proxy responses to the EQ-5D and patient health level (VAS) utility scores based on preference were valid for 30 patients with terminal lung cancer receiving palliative care at one facility in Japan. We found no significant differences between the
patients’ and nurses’ responses for EQ-5D utility or VAS scores at the three observation points. When we checked the answers to the EQ-5D utility scores in detail, we found that the changes in the nurses’ responses for the three time points did not perfectly mirror the small changes in the patients’ responses. We theorised that the main reason for this was that the patients were more attuned to and more likely to recognise and acknowledge their physical symptoms (e.g., pain, fatigue) and psychological symptoms (e.g., anxiety, depression) than the nurses. However, the changes were not statistically significant. The nurses’ responses more closely paralleled the patients’ responses in the VAS scores. These findings suggest that proxy responses could be considered valid surrogates for patient data.

The EQ-5D and VAS are based on different interval measures: the EQ-5D utility score is calculated using a conversion table, while the VAS is answered using a health status. Since the EQ-5D-3L consists of three levels, insufficient sensitivity and the ceiling effect for high scores answers are considered problems.21–24 EQ-5D is theoretically calculated in the range of 0–1, but if it becomes ‘harder than dying’, a negative value is assigned accordingly. Additionally, the maximum of good health was limited to scale 1. Based on this study’s results, the utility scores did not tend to concentrate at 1, which is likely due to the disease severity of the target patients. Moreover, there was no significant difference between the patients’ and nurses’ responses for both the utility values and VAS scores obtained. Thus, proxy responses seem useful for these two indicators.

In previous reports, such as comparing QOL evaluations between self and proxy responses for elementary and junior high school students who experienced childhood cancer and their parents, it was reported that there was a significant difference only in ‘self-esteem’.25 Further, there was a tendency for a little discrepancy in the items that can be objectively observed. There was some deviation for items related to patients’ emotional aspects. A similar trend was also shown in this study. There was a significant difference concerning ‘anxiety/depression’. However, unlike the previous report, there was no significant difference in ‘pain/discomfort’. This is, presumably, because the nurses who provided the proxy responses had expert knowledge about cancer pain, so the deviation was small.

Regarding the correlation between patients’ and nurses’ responses, the changes in the responses at the first and third observation time points were confirmed for all five dimensions of the EQ-5D, and the correlation coefficient was low (p=0.228). However, a significant weak positive correlation was shown. This result suggests that the agreement of the nurses’ responses to the patients’ responses was not weak, as shown in figures 1 and 3. In this study, the correlation coefficient was low, which might be influenced by the fact that the change in the EQ-5D level from the first observation to the third was 0 in many patients.

A cross-tabulation table is a statistical method that can be used only for discrete data, and its use is limited. However, it was suitable for evaluating the consistency of the EQ-5D responses in this study, indicating that the agreement of the responses in this study was extremely high.

Anyone who knows the patient well and has the patient’s best interest at heart could be a good candidate for being a patient’s proxy or surrogate evaluator. In most cases, the best candidates will be partners or other close family members. QoL surveys conducted at the end of adulthood have found that family members living with

| Patients’ response | Level | 1 | 2 | 3 | Total |
|--------------------|------|---|---|---|------|
|                    | 1    | 469 (53.4%) | 68 (7.7%) | 0 (0.0%) | 537 |
|                    | 2    | 150 (17.1%) | 158 (18.0%) | 4 (0.5%) | 312 |
|                    | 3    | 10 (1.1%) | 12 (1.4%) | 8 (0.9%) | 30 |
| **Total**          |      | 629 | 238 | 12 | 879 |

EQ-5D, Euro QoL 5 Dimension.
the patient or close friends or family who visited daily provided responses most consistent with the patient’s response. However, previous research has pointed out that close family members and friends of EOL patients will be experiencing increased mental distress that could significantly affect their objectivity. Some close family members’ or friends’ responses might be motivated more by the desire to postpone their own bereavement (by prolonging the patient’s life no matter what) than to reflect the patient’s genuine feelings. This is a known problem well documented in the literature. Thus, ethical constraints and biases should also be considered in surrogate’s evaluations. This study did not consider family members as proxy respondents. However, future studies should compare close family or friends as proxy respondents to caregiver proxies and actual patient responses to clarify the validity of proxy responses.

The results of this study’s cross-tabulations can be interpreted as follows by further organising the clinically relevant differences. Appropriate follow-up of changes in the patient’s condition is essential for the judgement of clinical interventions. In this study, the proxy responses by the experienced nurses were generally similar to the patients’ responses. Regarding the proxy responses for controlling pain and discomfort, a significant part of those on providing palliative care for patients in the terminal stage were highly consistent. As for cancer, there were a few reports that assessed the minimal clinically important difference (MCID) of EQ-5D. The differences in measure response values tended to be smaller than in previous studies on MCIDs. From the above, the findings, which are closely related to nursing assessment, were effective from a clinical point of view based on previous studies on MCIDs in patients with cancer using the EQ-5D. Although we did not specifically apply the MCID concept to our examination of proxy responses to the EQ-5D, we did note all the statistically significant differences between the patient and proxy responses. Future studies could add the MCID framework to augment the clinical findings. When measuring the utility of palliative care, the most important perspective is the EOL patients’ perceptions of the smallest change in a treatment outcome that they would consider important. We found that the evaluations of anxiety and depression were inconsistent. Future studies should also consider offering participants (both nurses and patients) some standard evaluation measure or tool for anxiety and depression to ensure meaningful comparisons.

This study had several limitations. First, the study was limited to 30 patients from a single centre. Second, we did not investigate the potential usefulness of using proxy responses for QoL assessments for patients who could not communicate or replay by themselves. Third, since we did not carry out an equivalence test for statistical differences, we could not verify the consistency of the proxy responses thoroughly. Fourth, because there were only two observers, we could not confirm the impact of the nursing staff’s diversity (lack of diversity). Finally, there was the possibility that processing any missing values could cause bias. We tried to minimise this by supplementing the missing values based on the average value of the previous observation time points and the weighted ratio of the nurses’ or patients’ answers to the average value of the previous or subsequent observation time point. This is a substantial limitation of the current study. This was a pilot study, and we hope to address these issues in a future large-scale study.

CONCLUSIONS

The results suggest that proxy responses can be useful because there were no significant differences between the patients’ and nurses’ responses in EQ-5D utility and VAS scores at the three observation points. This result is expected to be verified in future large-scale trials.

Contributors Conceptualisation: TT and TK; methodology: TT; software: YM; data curation: AS and MM; writing: TT and YM; visualisation: YM; project administration: TK; guarantor: TT. All authors read and agreed to the published version of the manuscript.

Funding This study was funded by Grants-in-Aid for Scientific Research (JSPS KAKENHI, Grant Number JP 25460620).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The ethics committee of National Cancer Center Hospital East approved this study (No.20121101).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The datasets analysed during the current study will be made available by the corresponding author on reasonable request.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) licence, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ids Tomoyuki Takura http://orcid.org/0000-0002-5987-9254 Yoko Matsuo http://orcid.org/0000-0001-7055-9996

REFERENCES
1 Shimizu M, Aoyama M, Morita T, et al. A second time nationwide survey of quality of end-of-life cancer care in general hospitals, inpatient palliative care units, and clinics in Japan: the J-HOPE 2 study. Palliat Care Res 2016;11:254–64.
2 Takura T. An evaluation of clinical economics and cases of cost-effectiveness. Intern Med 2018;57:1191–200.
3 Yorke J, Lloyd-Williams M, Smith J, et al. Management of the respiratory distress symptom cluster in lung cancer: a randomised controlled feasibility trial. Support Care Cancer 2015;23:3373–84.
4 Khue PM, Thom VT, Minh DG, et al. Depression and anxiety as key factors associated with quality of life among lung cancer patients in HAI Phong, Vietnam. Front Psychiatry 2019;10:352.
5 Matsuda T. Methodology of QoL measurement and instrument development. J Natl Inst Public Health 2004;53:181–5.
6 Sopina E, Serensen J, Beyer N, et al. Cost-Effectiveness of a randomised trial of physical activity in Alzheimer’s disease: a secondary analysis exploring patient and proxy-reported health-related quality of life measures in Denmark. BMJ Open 2017;7:e015217.
7 Sneeuw KC, Aaronson NK, Sprangers MA, et al. Evaluating the quality of life of cancer patients: comparison with other patients, significant others, physicians and nurses. Br J Cancer 1999;81:87–94.
8 Clapham S, Davesson BA, Allingham SF, et al. Patient-Reported outcome measurement of symptom distress is feasible in most clinical scenarios in palliative care: an observational study involving routinely collected data. Int J Qual Health Care 2021;33:mzab075.
9 Wennman-Larsen A, Tishelman C, Wengström Y, et al. Factors influencing agreement in symptom ratings by lung cancer patients and their significant others. J Pain Symptom Manage 2007;33:146–55.
10 Schiffczyk C, Romero B, Jonas C, et al. Generic quality of life assessment in dementia patients: a prospective cohort study. BMC Neuro 2010;10:48.
11 Schweikert B, Hahmann H, Leidl R. Validation of the EuroQol questionnaire in cardiac rehabilitation. Heart 2006;92:62–7.
12 Hansen H, Beyer N, Frolich A, et al. Inter-Day test-retest reproducibility of the cat, CCQ, HADS and EQ-5D-3L in patients with severe and very severe COPD. Patient Relat Outcome Meas 2021;12:117–28.
13 Limbers CA, Ripperger-Suhler J, Heffer RW, et al. Patient-Reported pediatric quality of life Inventory™ 4.0 generic core scales in pediatric patients with attention-deficit/hyperactivity disorder and comorbid psychiatric disorders: feasibility, reliability, and validity. Value Health 2011;14:521–30.
14 Klaassen RJ, Barr RD, Hughes J, et al. Nurses provide valuable proxy assessment of the health-related quality of life of children with Hodgkin disease. Cancer 2010;116:1602–7.
15 Care Ferrario S, Cardillo V, Vicario F, et al. Advanced cancer at home: caregiving and bereavement. Palliat Med 2004;18:108–20.
16 Campbell ML, Kiernan JM, Strandmark J, et al. Trajectory of dyspnea and respiratory distress among patients in the last month of life. J Palliat Med 2019;21:194–9.
17 Dawber R, Armour K, Ferry P, et al. Measurement of quality of life using EQ-5D in patients on prolonged mechanical ventilation: comparison of patients, family caregivers, and nurses. Qual Life Res 2010;19:721–7.
18 Peyre H, Coste J, Leplège A. Identifying type and determinants of missing items in quality of life questionnaires: application to the SF-36 French version of the 2003 Decennial health survey. Health Qual Life Outcomes 2010;8:16.
19 Luo N, Chew L-H, Fong K-Y, et al. A comparison of the EuroQol-5D and the health Utilities index mark 3 in patients with rheumatic disease. J Rheumatol 2003;30:2268–74.
20 Pang B, Wyatt G, Given C, et al. Burden and depression among caregivers of patients with cancer at the end of life. Oncol Nurs Forum 2004;31:1105–17.
21 Ringdal GI, Ringdal K, Jordhøy MS, et al. Burden and depression among caregivers of patients with cancer – an analysis of self-assessment and proxy-assessment. Palliat Med 2013;8:7–16.
22 Ringdal GI, Ringdal K, Jordhøy MS, et al. Burden and depression among caregivers of patients with cancer – an analysis of self-assessment and proxy-assessment. Palliat Med 2013;8:7–16.