Randomized Controlled Trial of Paper-Based at a Hospital versus Continual Electronic Patient-Reported Outcomes at Home for Metastatic Cancer Patients: Does Electronic Measurement at Home Detect Patients’ Health Status in Greater Detail?

Takeru Shiroiwa, Yasuhiro Hagiwara, Naruto Taira, Takuya Kawahara, Keiko Konomura, Tetsuya Iwamoto, Shinichi Noto, Takashi Fukuda, and Kojiro Shimozuma

Purpose. This study aimed to determine whether continual electronic patient-reported outcome (ePRO) measurements at home can capture the fluctuations in health-related quality of life (HRQOL) scores between visits. Methods. We performed a randomized controlled trial to compare the scores obtained by standard practice (paper-based measurements in the hospital) to scores by continuous measurement of ePRO at home. Metastatic cancer patients were randomly assigned to either the paper-based (n = 50) or the ePRO group (n = 52). EQ-5D-5L and EORTC QLQ C-30 scores were obtained on 3 different chemotherapy days in the paper-based group. Meanwhile, scores were obtained on the chemotherapy day and on days 3, 7, 10, and 14 in the ePRO group during 2 cycles. The first hypothesis of our study was that both scores at the same time points would be equivalent despite different measurement frequency, place, or mode of measurement. The second hypothesis was that PRO score-adjusted time would be different between the groups. For equivalence, the endpoint was the mean EQ-5D-5L index value on the chemotherapy day before the outpatient treatment. Only if equivalence was shown, quality-adjusted life-days (QALDs) were considered using all the data. Results. The adjusted mean difference in the EQ-5D-5L index was determined to be 0.013 (95% confidence interval [CI]: -0.049 to 0.022); the 95% CI did not exceed the equivalence margin. Similarly, the mean difference in global health status (2.28 [95% CI: 2.55 to 7.11]) also showed equivalence. However, the QALD by EQ-5D-5L was significantly lower in the ePRO group by 1.36 per 30 d (95% CI: -2.22 to -0.51; P = 0.0021). Conclusions. Continual measurements of the HRQOL at home by ePRO may yield more detailed profiles of the HRQOL.

Keywords
ePRO, patient-reported outcome, EQ-5D-5L, EORTC QLQ C-30, QALY

Introduction

Patient-reported outcomes (PROs), including health-related quality of life (HRQOL), are often collected in clinical trials to measure the patients’ health status and
evaluate the usefulness of interventions. However, the current practice of PRO measurement in clinical trials has some problems. In particular, in the case of chemotherapy, cancer patients normally feel worse for some days after undergoing chemotherapy, and they can only receive the subsequent chemotherapy after they recover from the adverse effects of the anticancer drugs administered in the previous cycle. However, the patients’ HRQOL is normally measured only on the chemotherapy days, when they would be expected to have recovered from adverse effects. This may cause over- or underestimation of the patients’ HRQOL influenced by chemotherapy.

Electronic PRO (ePRO) may resolve such problems associated with HRQOL measurements in clinical trials. Although almost all of the PRO methods were originally developed in a paper-based format, they have been adapted to an electronic (personal computer-, smartphone-, or tablet-based) format. In some cases, this type of data collection is more convenient (e.g., reminders for responses, no missing data, no data entry, and immediate responses) than the paper-based format. With ePRO, continual collection of the patients’ HRQOL data at home is now possible between the visits of the patients to the hospital. Patients can simply respond to the instrument on the scheduled days in their own homes. This strategy may overcome the limitations of the currently used PRO measurement in cancer patients.

The feasibility of this strategy is supported by the environment surrounding portable electrical devices (e.g., smartphones and tablets). Smartphones are rapidly becoming universally used gadgets. According to Japan’s statistics for 2018, about 80% of families and 65% of individuals own a smartphone. Younger people can access the internet more effectively than older people can, despite it being available for more than half of the people who are older than 70 y. In the context of such conditions, it is expected that respondents, including older cancer patients, can easily operate the application and respond to the PRO instrument through the smartphone. Nevertheless, there are no comparisons as yet of the scores obtained by frequent measurement using ePRO at home and by paper-based measurements conducted at the hospital.

Considering such current practices of HRQOL measurements and progress in ePRO technologies, the objective of our study was to determine whether ePRO measurements at home can capture the fluctuations in the HRQOL of cancer patients between their visits to the hospital for chemotherapy. For this purpose, we carried out a parallel-group randomized controlled trial (RCT) to compare the scores obtained by the standard practice (paper-based PRO measurements conducted at the hospital) and those obtained by continual measurements by ePRO (using tablets) at home in metastatic cancer patients.

The first hypothesis of our study was that both scores measured at the same time points would be equivalent, despite the differences in the frequency (number of measurement points), place (home or hospital), and mode (electronic or paper-based) of the measurements. If this equivalence was supported, then the effects of the aforementioned could be neglected.

Second, to consider the influence of continual measurements at home, our hypothesis was that the PRO score-adjusted time (area under the curve [AUC] of longitudinal scores) would be different between the 2 groups. If fluctuations in the patients’ HRQOL scores between their visits to the hospital can be detected by continual ePRO measurements, different patterns of longitudinal PRO scores would be obtained, and this would reflect the PRO score-adjusted times. If both the first and second hypotheses were supported, then we could conclude that continual measurement by ePRO at home is superior to conventional paper-based measurement for evaluation of the patients’ health status in greater detail, regardless of the differences in the frequency, place, and mode of measurement.

Methods

Study Design and Participants

In this multicenter, open-label RCT, we enrolled patients aged 20 y or older (definition of adults in Japan) with inoperable or metastatic cancer in Japan (UMIN000034863). The patients were receiving outpatient chemotherapy treatments, including immunotherapy, every 2, 3, or 4 wk. The chemotherapy and immunotherapy regimens were selected...
based on the physicians’ choices. Patients who could not use tablets by themselves and inpatients were excluded.

The trial was conducted with the approval of the institutional review board or ethics committee at every center, and the study protocol complied with the Declaration of Helsinki and local (Japanese) laws. Written informed consent was obtained from all patients before any trial procedure.

**Procedures**

Patients who were assigned to the paper-based PRO group responded to both the EQ-5D-5L and quality of life questionnaires (QLQ) C-30 using hard-copy questionnaires and a pen at the hospital before they received the outpatient chemotherapy.

EQ-5D-5L is one of the preference-based measures and comprises 5 items: “mobility,” “self-care,” “usual activities,” “pain/discomfort,” and “anxiety/depression.” Each item was assessed at 5 levels of description. Responses were converted to an EQ-5D-5L index value using a predetermined algorithm based on the societal preferences of the general population. The scoring system was designed as follows: a score of 1 indicates full health, and a score of 0 corresponds to death. Health statuses with negative scores also exist.

EORTC QLQ C-30 is a cancer-specific measure that consists of 30 items. It incorporates 5 functional scales (physical, role, cognitive, emotional, and social scales), 3 symptom scales (fatigue, pain, and nausea and vomiting), and the global health status. It also includes 6 single items (dyspnea, loss of appetite, insomnia, constipation, and diarrhea, and financial difficulties). Scores on all the scales (global health status, functional scales, and symptom scales) are converted to scores ranging from 0 to 100. Higher scores on the functional scales and global health status indicate a better health status.

Participants in the paper-based group were asked to respond at 3 different points in time during 2 consecutive chemotherapy cycles (baseline, the day of the next cycle of chemotherapy, and the day of the next-to-next cycle chemotherapy). The paper-based instruments were distributed by a clinical research coordinator (CRC) or a health care practitioner (such as a nurse or pharmacist), and the responses were collected in the hospital on the day of chemotherapy. This was repeated for another cycle (in total: 9 time points of measurements). Reminders were sent by e-mail, and alerts were displayed on the screen of the tablet on the day of the measurement. When the patients had problems operating the tablet to provide their responses, they called a support call center.

**Statistical Analysis**

We considered 2 hypotheses: 1) the HRQOL scores in both groups at the same time points will be equivalent and (2) the PRO score–adjusted lifetime will be different between the 2 groups. The primary endpoint was the EQ-5D-5L index value. Only if the first hypothesis was proven, the PRO score–adjusted life-days (i.e., quality-adjusted life-days [QALDs] in the case of EQ-5D-5L) were considered. The secondary endpoint was the EORTC QLQ C-30 score (global health status). The PRO score–adjusted life-days was used to consider the fluctuations in the EQ-5D-5L index and EORTC QLQ C-30 scores of the patients between the 2 visits.

Participants were randomly assigned at the ratio of 1:1, using the minimization method (computerized random numbers) with 2 factors (i.e., institutions and Eastern Cooperative Oncology Group [ECOG] performance status). Central randomization was applied, and the details of the series were unknown to the investigators and the coordinator. The patients and investigators were not blinded to the allocation. The required sample size was then calculated to test the hypothesis that both scores would be equivalent when the equivalence margin was 0.1 (based on the minimal clinical difference for the EQ-5D-5L) and the standard deviation was 0.15 (based on the SELECE BC trial). The equivalence margin method was used to assess the measurement equivalence. To obtain a statistical power of 80% power and significance level of 2.5%, the minimum sample size required per group was 48, and the sample size for this study was determined to be 50 per group.

The full analysis set population was used for statistical analysis of the PRO scores, which means we excluded patients with no response data from the first survey. The responses to the EQ-5D-5L were converted to utility scales using the Japanese value set. The score for each scale of EORTC QLQ C-30 was standardized to a scale ranging from 0 to 100. After the summary statistics of the EQ-5D-5L index and each scale of EORTC QLQ
C-30 were calculated, the results were analyzed using a mixed-effects model for repeated measures to confirm equivalence. The scores (at 2 points in time: on the day of the second and third chemotherapy visits, which was commonly measured in both the groups) were adjusted to the baseline score (measured on the day of the first cycle), gender, age, location of cancer, ECOG performance status, and stage. The boundary of equivalence was set using the minimally important difference of each instrument (i.e., both 0.1 and the more strict 0.05 for the EQ-5D-5L and 10 for each scale of the QLQ C-30). Missing data were imputed using the linear relation of scores between the 2 time points.

The PRO score–adjusted life-days for each instrument were obtained by computing the AUC. The AUC was calculated using a trapezoidal method from the data at all the time points (9 points in the ePRO group and 3 points in the paper-based group) and converted to the AUC per 30 days. This was done to adjust for the length of the chemotherapy cycles, which was different for each patient. The AUC was analyzed using a model including the same covariance with analysis of the scores for equivalence.

All statistical analyses were performed using SAS version 9.3.

Results

Between August 2019 and October 2019, 102 patients with inoperable or metastatic cancer at 6 sites were randomly assigned to either the paper-based PRO or ePRO measurement group. In total, 52 patients were allocated to the ePRO group and 50 to the paper-based PRO group. Among the participants, 59 patients (57.8%) had lung cancer, and the remaining 41 patients (40.2%) had breast cancer. The demographic and baseline characteristics were well balanced between the paper-based PRO and ePRO measurement groups (Table 1). The reasons for discontinuation are shown in a CONSORT diagram (Figure 1). Treatment-related adverse events during the 2 cycles of treatment occurred in 50 patients (51.0%), of whom 27 patients (26.5%) experienced grade 2 or more severe adverse events. The most common adverse events were neutropenia (6 patients), elevation of the serum glutamic oxaloacetic transaminase (4 patients), nausea (3 patients), and diarrhea (3 patients). The responses are summarized in Table 2. The global response rate was approximately 95%, which shows the feasibility and good compliance of the frequent ePRO measurements at home.

| Table 1 Characteristics of Patients |
|-------------------------------------|
| Paper-Based Group (n = 50) | ePRO Group (n = 52) |
| n | % | n | % |
|---|---|---|---|
| Gender | | | |
| Male | 24 | 48.0 | 24 | 46.2 |
| Female | 26 | 52.0 | 28 | 53.8 |
| Median age, y (Q1–Q3) | 59.5 (53.0–72.0) | 60.0 (54.0–69.0) |
| Location | | | |
| Lung | 28 | 56.0 | 31 | 59.6 |
| Breast | 20 | 40.0 | 21 | 40.4 |
| Other | 2 | 4.0 | 0 | 0.0 |
| ECOG performance status | | | |
| 0 | 27 | 54.0 | 34 | 65.4 |
| 1 | 23 | 46.0 | 17 | 32.7 |
| 2 | 0 | 0.0 | 1 | 1.9 |
| Cycle | | | |
| Biweekly | 4 | 8.0 | 5 | 9.6 |
| Triweekly | 32 | 64.0 | 33 | 63.5 |
| 4 times a wk | 14 | 28.0 | 14 | 26.9 |
| Stage | | | |
| I | 1 | 1.0 | 1 | 1.9 |
| II | 2 | 4.0 | 3 | 5.8 |
| III | 10 | 20.0 | 9 | 17.3 |
| IV | 37 | 74.0 | 39 | 75.0 |
| Surgery | | | |
| Yes | 19 | 38.0 | 15 | 28.8 |
| No | 31 | 62.0 | 37 | 71.2 |
| Radiotherapy | | | |
| Yes | 23 | 46.0 | 18 | 34.6 |
| No | 27 | 54.0 | 34 | 65.4 |

ECOG, Eastern Cooperative Oncology Group; ePRO, electronic patient-reported outcome.

Equivalence between the Paper-Based PRO Measurement in the Hospital and ePRO Measurement at Home

The simple mean scores of the EQ-5D-5L and global health status of the EORTC QLQ C-30 are shown in Figure 2 and Figure 3, respectively. In the ePRO group, these scores were found to be low on day 7 or day 10/14 and then recovered before the start of the next chemotherapy. The least square (LS) mean of EQ-5D-5L during 2 cycles was estimated to be 0.81 (95% confidence interval [CI]: 0.73 to 0.88) in the ePRO group and 0.82 (95% CI: 0.75 to 0.89) in the paper-based group. The difference between the 2 groups was −0.013 (95% CI: −0.049 to 0.022). The 95% CI did not exceed the equivalence margin; that is, the equivalence of the EQ-5D-5L between the paper-based PRO measurements in the hospital and ePRO measurements at home was confirmed.
Similarly, the LS mean of the global health status was estimated to be 66.8 (95% CI: 56.6 to 77.0) in the ePRO group and 64.5 (95% CI: 54.8 to 74.3) in the paper-based group. The difference (2.28 [95% CI: 2.2 to 7.1]) also showed equivalence between the ePRO at home and the paper based at the hospital group, as the equivalence margin was 10. Similar results were found for the other functional scales (physical, role, cognitive, emotional, and social scales). The difference between the 2 groups in physical function was 1.50 (95% CI: −2.07 to 5.08), in role function was 0.14 (95% CI: −6.54 to 6.81), in emotional function was 0.068 (95% CI: −3.27 to 3.41), in cognitive function was −0.0079 (95% CI: −4.78 to 4.76), and in social function was 4.53 (95% CI: −1.41 to 10.5). Functional scales, except for social function, showed similar results to those of the global health status.

**Comparison of the AUC between the Paper-Based PRO and ePRO Groups**

The LS mean of the AUC for the EQ-5D-5L index (i.e., QALD per 30 d) was 23.3 (95% CI: 21.5 to 25.2) in the ePRO group and 24.7 (95% CI: 23.0 to 26.4) in the paper based in the hospital group. The difference between the 2 groups was 1.36 (95% CI: 2.2 to 0.51), which was statistically significant (**P** = 0.0021). The AUC in the ePRO group was lower than that in the paper-based group, when the scores measured at home were reflected.

The AUC for the global health status was estimated to be 1853 (95% CI: 1615 to 2090) in the ePRO group and 1964 (95% CI: 1738 to 2190) in the paper-based group. The estimated difference was −111 (95% CI: −223 to 1.04), but the difference was not statistically significant (**P** = 0.0521), although the AUC in the ePRO group tended to be lower than that in the paper-based PRO group. No significant differences in the AUC were observed for any of the other functional scales of the EORTC QLQ C-30, except physical function.

**Table 2 Response Rate**

|               | Paper-Based Group | ePRO Group |
|---------------|-------------------|------------|
|               | n     | %    | n     | %    |
| **First cycle** |       |      |       |      |
| Baseline      | 48    | 100.0| 51    | 100.0|
| Day 3         | 50    | 98.0 | 50    | 98.0 |
| Day 7         | 50    | 98.0 | 50    | 98.0 |
| Day 10 or 14  | 51    | 100.0| 51    | 100.0|
| On the day of next chemotherapy | 46 | 100.0 | 51 | 100.0 |
| **Second cycle** |       |      |       |      |
| Day 3         | 48    | 94.1 | 49    | 96.1 |
| Day 7         | 49    | 96.1 | 49    | 96.1 |
| Day 10 or 14  | 50    | 98.0 | 50    | 98.0 |
| On the day of next chemotherapy | 45 | 97.8 | 50 | 98.0 |

ePRO, electronic patient-reported outcome.

Similarly, the LS mean of the global health status was estimated to be 66.8 (95% CI: 56.6 to 77.0) in the ePRO group and 64.5 (95% CI: 54.8 to 74.3) in the paper-based group. The difference (2.28 [95% CI: −2.55 to 7.11]) also showed equivalence between the ePRO at home and the paper based at the hospital group, as the equivalence margin was 10. Similar results were found for the other functional scales (physical, role, cognitive, emotional, and social scales). The difference between the 2 groups in physical function was 1.50 (95% CI: −2.07 to 5.08), in role function was 0.14 (95% CI: −6.54 to 6.81), in emotional function was 0.068 (95% CI: −3.27 to 3.41), in cognitive function was −0.0079 (95% CI: −4.78 to 4.76), and in social function was 4.53 (95% CI: −1.41 to 10.5). Functional scales, except for social function, showed similar results to those of the global health status.

**Comparison of the AUC between the Paper-Based PRO and ePRO Groups**

The LS mean of the AUC for the EQ-5D-5L index (i.e., QALD per 30 d) was 23.3 (95% CI: 21.5 to 25.2) in the ePRO group and 24.7 (95% CI: 23.0 to 26.4) in the paper based in the hospital group. The difference between the 2 groups was 1.36 (95% CI: 2.2 to 0.51), which was statistically significant (**P** = 0.0021). The AUC in the ePRO group was lower than that in the paper-based group, when the scores measured at home were reflected.

The AUC for the global health status was estimated to be 1853 (95% CI: 1615 to 2090) in the ePRO group and 1964 (95% CI: 1738 to 2190) in the paper-based group. The estimated difference was −111 (95% CI: −223 to 1.04), but the difference was not statistically significant (**P** = 0.0521), although the AUC in the ePRO group tended to be lower than that in the paper-based PRO group. No significant differences in the AUC were observed for any of the other functional scales of the EORTC QLQ C-30, except physical function.
Discussion

We carried out this RCT to compare the scores measured on the day of chemotherapy day between a group who underwent paper-based measurement of the PRO in the hospital and a group who underwent ePRO measurement at home. From the ePRO measurements at home, more continual data between 2 visits were collected to capture the fluctuations of the HRQOL in the patients at home, between the hospital visits for chemotherapy. First, the equivalence of EQ-5D-5L and EORTC QLQ C-30 between the paper-based PRO measurements in the hospital and ePRO measurements at home was demonstrated. Despite the differences in the frequency, place, and mode of measurement, equivalence of the scores at the same time points was shown. Second, the QALD determined by EQ-5D-5L determined by the ePRO measurements at home was smaller than that determined by paper-based measurement of the PRO in the hospital. The difference in the QALD was 1.36 per 30 d, which implies that the utility score was lower by an average of 0.042. The value is comparable with the minimally important difference of EQ-5D-5L. We collected and evaluated both the EQ-5D-5L and EORTC QLQ C-30; however, the recall period is different between the 2 instruments (same day in EQ-5D-5L and 1 wk in EORTC QLQ C-30). In the case of EORTC QLQ C-30, the recall period overlapped between different measurement points, which may have influenced the results of the EORTC QLQ C-30. Considering these factors, we think that the results of EQ-5D-5L are clearer and easier to interpret. In addition, we could not detect any significant differences in the AUCs of the EORTC QLQ C-30 scores between the ePRO and paper-based group; this could be related to the suboptimal recall period in this study.

Deterioration of the health status during chemotherapy may be reflected by the QALD in the ePRO group. This was supported by the “W”-shaped plot of the longitudinal scores in the ePRO group (Figures 2 and 3). Patients normally feel worse for some days after chemotherapy, and they receive the subsequent chemotherapy only after they recover from the effects of the previously administered anticancer drugs. Taking this into consideration, it would be reasonable to expect the worst HRQOL scores to be obtained during the period between the hospital visits for chemotherapy, with recovery of the deteriorated scores on the day of chemotherapy. The “W”-shaped plot of the longitudinal scores may capture the patients’ HRQOL after chemotherapy in a more detailed manner. This might also influence the results of cost-effectiveness analyses using QALY estimates.

Our study showed similar and good response rates in both groups (more than 95%), although approximately 25% of the patients were aged 70 y or older. For some older people, the ePRO survey was acceptable. We had concerns about obtaining cooperation from the patients, as the requests for responses to the ePRO questionnaires at home were frequent, and some patients may have unintentionally forgotten to respond. However, most of the patients were able to complete the task. This ePRO survey method seems feasible, at least for short periods.
Depending on the situation, ePRO measurements can contribute to the tracing of a patient’s HRQOL responses in greater detail, which could be expected to help provide evidence for the usefulness of chemotherapy regimens that do not decrease the patients’ HRQOL scores. However, from our results, the effect size of the QALD by EQ-5D-5L was about 5%, which might not be very large. One idea is to use continual ePRO measurements at home when intensive chemotherapy is administered, considering the burden of the ePRO task.

There were some limitations of this study. First, most of the patients included in this study had been diagnosed as suffering from lung and breast cancer, although they received a variety of chemotherapy treatments. Patients with other cancers may experience different transition patterns of the HRQOL scores, although a similar pattern to that in this study could also have been observed. In addition, our sample was relatively small, and there are some problems with the generalizability of the results. Finally, in this study, patients who could not consume tablets were excluded, although it is also difficult for such patients to measure their health status by ePRO.

In conclusion, equivalent scores on the EQ-5D-5L and EORTC QLQ C-30 questionnaires can be obtained through paper-based estimation of the PRO at a hospital and measurement of the ePRO at home, even if the ePRO measurement is done more frequently. Second, continual measurements of the HRQOL at home by ePRO may yield more detailed profiles of the HRQOL. Depending on the characteristics of the chemotherapy or the patients, it might be useful to detect different patterns of the HRQOL of the patients between 2 hospital visits. This might also influence the calculation of the QALY and the cost-effectiveness analysis. The response rate for the ePRO measurement was high, even though patients were asked to respond repeatedly at home. Thus, continual ePRO measurement could be deemed as being feasible. We hope that our results would contribute to developing data collection strategies for better assessment of the treatment benefits of chemotherapy in cancer patients.

Acknowledgments
This study was supported by the Comprehensive Support Project for Oncology Research (CSPOR) of Public Health Research Fund (PHRF).

ORCID iDs
Takeru Shiroiwa: https://orcid.org/0000-0003-3055-9932
Yasuhiro Hagiwara: https://orcid.org/0000-0003-3792-5368

References
1. Ministry of Internal Affairs and Communications. The White Paper on Information and Communications in Japan. Tokyo: Ministry of Internal Affairs and Communications; 2018.
2. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727–36.
3. Aaronson NK, Ahmedzai S, Bergman B, et al. 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. 1993;85(5):365–76.
4. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. Value Health. 2014;17(4):445–53.
5. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Go´iri JM. Overview, update, and lessons learned from the international EQ-5D-5L valuation work: version 2 of the EQ-5D-5L valuation protocol. Value Health. 2019;22(1):23–30.
6. Takashima T, Mukai H, Hara F, et al. Taxanes versus S-1 as the first-line chemotherapy for metastatic breast cancer (SELECT BC): an open-label, non-inferiority, randomised phase 3 trial. Lancet Oncol. 2016;17(1):90–8.
7. Shiroiwa T, Fukuda T, Shimozuma K, et al. Long-term health status as measured by EQ-5D among patients with metastatic breast cancer: comparison of first-line oral S-1 and taxane therapies in the randomized phase III SELECT BC trial. Qual Life Res. 2017;26(2):445–53.
8. Luo N, Chew LH, Fong KY, et al. Do English and Chinese EQ-5D versions demonstrate measurement equivalence? An exploratory study. Health Qual Life Outcomes. 2003;1:7.
9. Julious SA. Sample sizes for clinical trials with normal data. Stat Med. 2004;23(12):1921–86.
10. Shiroiwa T, Ikeda S, Noto S, et al. Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. Value Health. 2016;19(5):648–54.
11. Fayers P, Aaronson N, Bjordal K, Groenvold M, Curran D, Bottomley A. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer; 2001.
12. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. Value Health. 2017;20(4):644–50.
13. Shiroiwa T, Fukuda T, Ikeda S, et al. Japanese population norms for preference-based measures: EQ-5D-3L, EQ-5D-5L, and SF-6D. Qual Life Res. 2016;25(3):707–19.
14. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol. 1998;16(1):139–44.
15. Cocks K, King MT, Velikova G, Martyn St-James M, Fayers PM, Brown JM. Evidence-based guidelines for determination of sample size and interpretation of the
European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol*. 2011;29(1):89–96.

16. Musoro JZ, Coens C, Fiteni F, et al. Minimally important differences for interpreting EORTC QLQ-C30 scores in patients with advanced breast cancer. *JNCI Cancer Spectr*. 2019;3(3):pkz037.

17. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*. 2010;152(11):726–32.

18. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*. 2012;308(24):2594–604.