Development and validation of a Symptom Illustration Scale from the Patient-Reported Outcome Common Terminology Criteria for Adverse Events for patients with breast cancer

Yoko Suzuki[1,2], Takayuki Iwamoto[1,3*], Maya Uno[1,2], Minami Hatono[1], Yukiko Kajiwara[1], Yuko Takahashi[1], Mariko Kochi[1], Tadahiko Shien[1], Yuichiro Kikawa[4], Yukari Uemura[5], Yasuhiro Hagiwara[6], Seiichiro Yamamoto[7], Naruto Taira[1,3], Hiroyoshi Doihara[1,8], Shinichi Toyooka[2]

1. Breast and Endocrine Surgery, Okayama University Hospital, Okayama, Japan
2. Department of General Thoracic Surgery and Breast and Endocrinological Surgery, Okayama University
3. Department of Breast and Thyroid Surgery, Kawasaki Medical School Hospital, Kurashiki, Japan
4. Breast Surgery, Kansai Medical University Hospital, Osaka, Japan
5. Biosciences, Department of Data Science, Center for Clinical Sciences, National Center for Global Health and Medicine, Tokyo, Japan
6. Department of Biostatistics, School of Public Health, The University of Tokyo, Tokyo, Japan
7. Shizuoka Graduate University of Public Health, Shizuoka, Japan
8. Department of Surgery, Kawasaki Medical School General Medical Center, Okayama, Japan

Corresponding author:

Takayuki Iwamoto, M.D., Ph D.,

Breast and Thyroid Surgery, Kawasaki Medical School Hospital

Address: 577 Matsushima, Kurashiki City, Okayama, Japan 701-0192

Phone: +81-86-462-1111; Fax: +81-86-462-7897

tiwamoto@med.kawasaki-m.ac.jp

Abstract

Purpose: Emojis are commonly used for daily communication and may be useful in assessing patient-reported outcomes (PROs) in breast cancer. The purpose of this study is to develop and validate a Symptom Illustration Scale (SIS) as a new PRO measurement.

Methods: Eighteen original SIS items were developed from the PRO-CTCAE. In cohort one, the SIS validity and reliability were examined in patients with breast cancer, using a semi-structured five-question survey to investigate content validity. PROs with PRO-CTCAE and SIS were examined twice to determine criteria validity and test-retest reliability. In cohort two, the responsiveness of the scales were examined in patients treated with anthracycline, docetaxel, paclitaxel, and endocrine therapy. PROs with PRO-CTCAE and SIS were investigated two or three times, depending on the therapy.

Results: Patients were enrolled from August 2019 to October 2020. In cohort one (n=70), most patients had no difficulties with the SIS, but 16 patients indicated that it was difficult to understand severities in the SIS. For criterion validity, Spearman rank correlation coefficients (r_s) between PRO-CTCAE and SIS items were ≥0.41, except for "Decreased appetite." For test-retest reliability, κ coefficients of the SIS were ≥0.41 for 16/18 items (88.9%). Response time was significantly shorter for the SIS than for PRO-CTCAE (p<0.001). In cohort two (n=106), score changes between PRO-CTCAE and SIS for relevant symptoms all had correlations with r_s≥0.41.

Conclusion An original SIS from the PRO-CTCAE for patients with breast cancer were verified the validity, reliability, and responsiveness. Further studies to improve and validate the SIS are needed.

Keywords: Quality of life, patient-reported outcome measurements, breast cancer, electronic data capture of PROs (ePRO), Symptom Illustration Scale.

Introduction

Patient-reported outcomes (PROs) have an important role in oncology research and are used as key endpoints in some oncology drug trials [1–3]. With development of electronic communication devices, electronic data capture of PROs (ePRO) has also become a promising approach. ePRO is a web-based system that enables medical staff to monitor patients' symptoms in real time [4–9]. Attempts to integrate PRO assessment into routine clinical care have increased, with comparative studies showing that ePRO improves readmission rates, emergency room visits, medication compliance, accuracy of subjective symptoms, symptom management, and possibly overall survival (OS) [8, 10–14].

Introducing ePRO into daily clinical situations requires good compliance, ease of understanding, and familiarity. Additionally, the patient burden and language barrier should be resolved. PRO measurements (PROMs) have been developed to evaluate PROs in clinical trials. PROMs have been translated into multiple languages and are widely used to evaluate PROs. One aim of assessing PROs in clinical trials is to evaluate adverse events. Patients are required to answer multiple questions to evaluate symptoms for assessing adverse events using, for example, the PRO-Common Terminology Criteria for Adverse Events (CTCAE) questionnaire [15]. However, this may be a burden to patients with a poor performance status (PS) due to progression of cancer. Moreover, these PROMs are based on text and this may cause additional difficulties for patients who cannot read.

Use of illustrations or emojis may be an option for assessment of PROs in clinical practice [16]. The Wong-Baker FACES® Pain Rating Scale, one of several face scales for pain assessment, shows excellent correlations with a visual analog scale, particularly in children who have difficulty expressing pain severity with language [16, 17]. In another study, smiley pictures appeared to fulfill the requirements for early and prompt screening of depression, particularly in older patients who had a stroke and may not have the concentration required to complete lengthy questionnaires [17].

Emojis, which comprise actual pictures and require specific software support, were created in Japan in 1997 and encoded in Unicode standard in 2010. Emojis are now commonly used for daily communication [18] and are familiar and universal, and thus understandable, regardless of language. Use of emojis in medical situations

has been suggested previously [19–21]. A correlation between Linear Analog Scale Assessment and emojis, and possible use of emojis for PRO assessment, were described at the 2018 Annual Meeting of the American Society of Clinical Oncology. Stickers can be comprehensive representations of various elements, such as environmental descriptions and facial/body language, and these are also commonly used in daily mobile messages [18].

Use of emojis or stickers has rarely been studied to assess PROs in patients with cancer. Therefore, the objective of this study was to develop the Symptom Illustration Scale (SIS), which uses emojis and stickers to assess PROs. We hypothesize that the new SIS is equivalent to PRO-CTCAE in terms of validity, reliability, and responsiveness.

Material and Methods

Development of the SIS was shown in Supplementary method.

Validation of SIS v.1.0

The study was approved by the Medical Ethics Committee of Okayama University Hospital on November 8, 2019 (application numbers 1907-038 (Cohort one) and 1907-042 (Cohort two)). Written consent was obtained from each patient. Each patient received a 500-yen gift card. To validate SIS v.1.0, two single-center, prospective, cohort studies were conducted to test reliability and validity (cohort one) and responsiveness (cohort two).

Participants

The study had a prospective design and used two cohorts for two purposes: to examine if the SIS could identify a stable symptom and a change in a symptom. Cohort one included patients who were clinically stable. The inclusion criteria were 1) age ≥20 years, 2) diagnosed with breast cancer, 3) stage 0–III breast cancer, 4) able to comprehend Japanese, 5) underwent initial treatment (e.g., surgery, chemotherapy, and radiotherapy) that ended more than one year ago, and 6) underwent prescribed endocrine therapy that started more than one year ago. Patients who were receiving treatment with a molecular-targeted drug (e.g., trastuzumab, pertuzumab) were also enrolled. Patients with active disease or with cognitive disorders identified from comorbidities and medication in medical records were excluded.

For cohort two, the inclusion criteria were: 1) age ≥20 years, 2) diagnosed with breast cancer, 3) Stage 0–III breast cancer, 4) able to comprehend Japanese, and 5) started any of the following adjuvant therapies: i) Anthracycline: patients scheduled for neoadjuvant or adjuvant chemotherapy with intravenous anthracycline and cyclophosphamide at intervals of three weeks for four cycles; ii) Docetaxel: patients scheduled for neoadjuvant or adjuvant chemotherapy with intravenous docetaxel and cyclophosphamide or docetaxel alone at intervals of three weeks for four cycles. iii) Paclitaxel: patients scheduled for neoadjuvant or adjuvant chemotherapy with intravenous paclitaxel weekly for 12 weeks. iv) Endocrine therapy: patients scheduled for adjuvant endocrine therapy with tamoxifen or an aromatase inhibitor once daily. Patients with cognitive disorders were excluded.

Data collection and survey

In both cohorts, demographic data were collected after the survey, including marital status, academic background, employment status, frequently used electronic communication equipment, frequency of computer usage, and frequency of use of emojis in daily life. All clinical data were extracted from electronic medical records at our hospital.

In cohort one, we investigated PROs at the time of enrollment (T1) and 7 days later (T2), using paper-based PRO-CTCAE and SIS questionnaires. Using the envelope method, patients were randomized into PRO-CTCAE and SIS groups. At T1, a face-to-face survey was conducted in the hospital. At T1 and T2, patients in the SIS group completed the SIS followed by the PRO-CTCAE, whereas those in the PRO-CTCAE group completed the PRO-CTCAE followed by the SIS. After collecting each questionnaire at T1, a semi-structured survey was conducted by the medical staff. The survey comprised the questions was shown in Supplementary method. At T1, questionnaires for T2 were given to patients. These were completed at home and mailed to our research office. The deadline for submission was 14 days after the registration date. If questionnaires did not arrive at the research office by the 10th day, the staff telephoned each patient and asked them to submit their questionnaires.

In cohort two to examine the change in scores, we assessed responsiveness using the PRO-CTCAE and SIS questionnaires, as in cohort one, but the patients were not randomized into SIS and PRO-CTCAE groups, because it did not matter which order the surveys were administered. Cohort two included patients with breast cancer under treatment with chemotherapeutic regimens or endocrine therapy. As such, we assumed that the type

of adverse event and time of onset would differ depending on the regimen.

The survey schedules are defined in Supplementary method. Depending on the reason (e.g., adverse events), dose reduction or postponement of treatment was allowed. Concomitant anti-human epidermal growth factor receptor type 2 (HER2) therapy (trastuzumab and pertuzumab) was also allowed. In cases of treatment discontinuation, the last survey was conducted within two weeks after discontinuation.

Outcomes

In cohort one, the primary endpoint was the content validity of the SIS in order to verify the accuracy of measuring what is truly intended as the subject, and the secondary endpoints were criteria validity based on PRO-CTCAE in order to verify how accurately a test measures the outcome it is designed to measure, test-retest reliability in order to verify the correlation between results obtained by repeating the same test on the same subjects at a consistent interval, assuming that the subjects do not undergo significant changes over a short period of time, and response time. In cohort two, correlations between changes in PRO-CTCAE and SIS were assessed in order to verify the high responsiveness of the scale, which reflects changes in symptoms such as improvement or deterioration.

Statistical Analysis

In cohort one, to verify the content validity of the SIS, responses were categorically classified based on data collected in the semi-structured survey. The distribution and ratio of participants per item were calculated. To verify criteria validity and test-retest reliability, data for symptom severity were ranked 0–4 (rash was ranked 0 or 1) from mild to severe. To verify criteria validity, the Spearman rank correlation coefficient (r_s), κ coefficient, 95% confidence interval (CI), and SIS concordance rate were calculated for each PRO-CTCAE item. Moderate correlation (r_s >0.41) and strong correlation (r_s >0.70) were defined as “correlation” and a κ coefficient ≥0.41 indicated “agreement”. The rate of concordance was calculated as the ratio of patients who responded to both questionnaires to those whose severity on both questionnaires matched. To verify the test-retest reliability of PRO-CTCAE and SIS, the degree of agreement between T1 and T2 was calculated as a κ coefficient, 95% CI, and agreement rate. The times required for completion of PRO-CTCAE and SIS were compared by t-test.

In cohort two, to verify the responsiveness of SIS, datsa for SIS and PRO-CTCAE were interpreted as

continuous variables ranging from 0 to 4, from mild to severe. Changes in each score on PRO-CTCAE and SIS from T1 to T2 and T1 to T3 were calculated. Correlation analysis was performed between PRO-CTCAE and SIS using r_s and 95% CI at each survey time point.

In cohort one, the sample size was defined as follows: to verify content validity, we assumed that theoretical saturation would be reached by conducting a semi-structured survey in 100 patients. As such, we set the target number of patients to 100. In cohort two, we assumed a correlation coefficient of 0.8 for PRO-CTCAE and SIS score changes, which required a sample size of 14 to produce results with a probability of 80%. Considering dropout cases, the target number of cases for cohort two was set to 30.

Spearman rank order correlation analysis was performed using R v.3.2.1, and κ coefficient analysis was performed using JMP v.11.2 (SAS Institute, Inc., Cary, NC, USA).

Results

Characteristics of cohort one

From August 2019 to March 2020, 100 cases were enrolled in cohort one. Of these, 70 cases were analyzed and 30 were excluded due to a printing error found in the questionnaire. There were no missing surveys at all time points. The patient background factors are shown in Table 1. All participants were female. The median age was 56 years (range: 37–76 years) and the median interval from diagnosis to enrollment was 3.7 years. Among all participants, 87% were smartphone users and 94% had used an emoji. Only three (4.3%) had an education level below junior high school, 19 (27.1%) were homemakers, and nine (12.9%) were neither a smartphone nor tablet computer user.

Content validity

To test whether the scale is measuring the appropriate elements, we assessed content validity through a semi-structured survey comprising the following questions:

Q1. Was there any SIS **item that you found difficult to understand? (comprehensibility)**

For this question, 54 patients (77%) answered “no” and approximately 80% of patients easily understood the purpose of all SIS items. Additionally, 16 (23%) answered that one or more items were difficult to understand;

three (4%) answered that “Decreased appetite” and “Swelling” were difficult to understand; and two (3%) answered “Nausea” and “Hot flushes.” The 16 patients with difficulties were asked which aspects were difficult to understand, and 9 indicated differences in severity.

Q2. Was there any SIS **item that you felt needed an additional explanation? (comprehensibility)**

For this question, 65 patients (93%) answered “no.” Of the five patients who answered “yes,” four stated “an explanation is needed regarding the differences in severity”.

Q3. Was there any other symptom that you wanted to report? (comprehensiveness)

For this question, 44 (63%) answered “no.” Among those who answered “yes” the main symptoms were “pain in the surgical wound” in (n=4, 6%), “itchiness” (n=4, 6%), and “headache” (n=3, 4%). In the free comments section, one patient stated “I had no way to report pain caused by a positional change.”

Q4. Which was easier to answer, the SIS **or PRO-CTCAE questionnaire? (response burden)**

For this question, 61 patients (87%) stated that the SIS was easier to answer than the PRO-CTCAE. In the free comments section, seven patients mentioned that the SIS were more fun to answer and had no response burden.

Q5. Did you feel discomfort or shame while using the SIS**? (impression)**

For this question, 68 patients (97%) answered “no.” In the free comments section, 30 patients stated “The SIS was cute” and five stated “I felt cheerful after answering the SIS.”

The results for Q1 to Q5 are shown in Table 3.

Criterion validity

Criterion validity was assessed to investigate whether the test values obtained were highly correlated with the external criteria. The result at T1 is shown in Figure 1. The r_s values for PRO-CTCAE and SIS items were ≥0.41, except for “Decreased appetite.” This indicated that 17/18 items (94.4%) were correlated between the PRO-CTCAE and SIS. Additionally, the κ coefficients showed agreement for 13/18 test items (72.2%) and the concordance rates were within 38.6–98.6%.

The survey results at T2 were similar to those at T1; thus, 17/18 items (94.4%) were correlated and 12/18 (66.7%) items matched (Figure S1). Likewise, the results of the stratified analysis of the PRO-CTCAE and SIS

groups were similar. The correlation coefficient for “Decreased appetite” tended to be low; however, a correlation was found between other PRO-CTCAE and SIS items (Figures S2, S3).

Test-retest reliability

Test-retest reliability between T1 and T2 was assessed to investigate whether symptoms changed in a short period of time. The κ coefficients for the SIS were ≥0.41 for 16/18 items (88.9%), excluding “vomiting” and “diarrhea” (Figure 2); the κ coefficients of these PRO-CTCAE items was low. The test-retest reliability between PRO-CTCAE and SIS was similar. In addition, at T1 and T2, the concordance rates of all SIS items were ≥50%. The test-retest reliability per group is shown in Figures S4 and S5.

Comparison of response time

The average time required to answer the SIS was 1 min 48 s, with a range of 1 min 37 s to 1 min 59 s, whereas that for the PRO-CTCAE was 3 min 20 s, with a range of 2 min 59 s to 3 min 32 s. Thus, the response time for SIS was about 1.5 min shorter than that for the PRO-CTCAE (t-test: p <0.001).

Characteristics of cohort two

From December 2019 to October 2020, 29 patients receiving anthracycline, 30 receiving docetaxel, 17 receiving paclitaxel, and 30 receiving endocrine therapy were registered in cohort two. The demographic and clinical characteristics are shown in Table 2. Compared to other groups, the anthracycline group had more patients aged >60 years. Most patients had a performance status of 0 or 1; >70% were married; all had more than high school education; >80% used a smartphone daily, >50% rarely used a personal computer; and >90% used emojis in emails or social media.

Correlation of score changes between the PRO-CTCAE and SIS

In the anthracycline group, we focused on changes in appetite, nausea, vomiting and hair loss. The results showed that 17/18 items (94.4%) were correlated in terms of score changes from T1 (before the first cycle) to T2 (day 8 of the first cycle) ($r_s \geq 0.41$) (Figure 3). Decreased appetite (S: Severity), nausea and vomiting all had $r_s \geq 0.41$. Only “Decreased appetite (I: Interference)” had $r_s = 0.35$. Correlations of score changes from T1 to T3 (day 15 of the first cycle) are shown in Figure S6. “Vomiting” had a weak correlation ($r_s < 0.41$), while “Hair loss,” “Decreased appetite,” and “Nausea” were correlated ($r_s \geq 0.41$)

In the docetaxel group, we particularly examined change of swelling. For “Swelling,” all three items from T1 (before the first cycle) to T2 (day 1 of the second cycle) and two of three items, except “I: Interference,” from T1 to T3 (day 1 of the fourth cycle) were correlated, indicating a generally good correlation (Figures S7, S8).

In the paclitaxel group, we focused on the change of numbness & tingling. Score changes from T1 (before the first cycle) to T2 (day 1 of the sixth cycle) between PRO-CTCAE and SIS for “Numbness & tingling”, as well as from T1 to T3 (day 1 of the 12^{th} cycle) were correlated, indicating a generally good correlation (Figures S9, S10).

In the endocrine group, we assessed changes of joint pain and hot flushes. Score changes from T1 to T2 (2 months ±2 weeks after the first treatment) between PRO-CTCAE and SIS for “Joint pain” and “Hot flushes” showed correlations between PRO-CTCAE and SIS (Figure S11).

Discussion

In this study, we developed an 18-item original SIS from the PRO-CTCAE and verified the validity and reliability of the SIS in cohort one and the responsiveness of the SIS in cohort two.

The validity and reliability of the SIS was shown in cohort one. The findings were generally favorable and within expectations. Emojis are widely used in social media, which can increase the end users' fondness for these items as a fun and practical tool. Our study included many patients who used smartphones and emojis, reflecting today's society wherein these have become a part of life. In this regard, future use of emojis between healthcare providers and patients is easy to imagine, particularly if communicating by a conventional method is difficult. Based on our findings, most patients indicated that the SIS was easier to answer than the PRO-CTCAE. Further, they did not feel embarrassment or discomfort while answering the SIS, and the response time was shorter than that for the PRO-CTCAE. Lastly, some patients indicated that they enjoyed answering the questions on the SIS. It is important to note that these results may simply capture different results from different indicators.

Similarly to previous reports [17, 22], we received a few comments regarding comprehensibility, particularly

difficulty understanding differences in severity. Lee et al. developed a 5-step (1: very happy, 2: happy, 3: fair/average, 4: sad, and 5: very sad) evaluation tool with a smiley face scale for patients with ischemic stroke and noticed that older patients had difficulty telling the difference between very happy and happy, as well as very sad and sad, faces [17]. Thus, they reduced the number of expressions to 3. Since we anticipated this problem during the development phase of SIS, we added some text to the emojis. In the next version, we may need to revise the expressions to make them more easily understandable.

We also received a few comments regarding comprehensiveness. In question 3, which assesses content validity, most patients answered that they had no other symptoms to report. However, a small number of patients stated that they experienced additional symptoms, including “pain in the surgical wound,” “itchiness,” and “headache.”. In the next version, some informative items will be added to the SIS. "Pain in the surgical wound," a possible complication of total mastectomy, is referred to as post-mastectomy pain syndrome (PMPS). In a previous report, 50% of survey respondents stated that they were still experiencing PMPS at a mean of 9 years after surgery [23]. “Headache” is also a potentially significant symptom that may be suggestive of distant metastasis. Covering all symptoms, including rare symptoms, is not possible, but we will consider adding new items based on their frequency of occurrence and clinical importance.

For criterion validity, only “Decreased appetite”, one of 18 items (5.6%), was below the standard correlation value of 0.41. We noticed a discrepancy in this item between the PRO-CTCAE and SIS questionnaires. The PRO-CTCAE examines the degree of appetite loss by asking “In the last 7 days, what was the SEVERITY of your DECREASED APPETITE at its WORST?”, whereas the SIS asks “How was your appetite in the past week?” This may have led to the poor agreement between the two scales. In the next version of the SIS, we plan to update the question for “Decreased appetite”.

We confirmed that the test-retest reliability of SIS was favorable. Only two items, "Vomiting" and "Diarrhea," showed weak agreement. For “Vomiting,” there was a significant difference in κ coefficient between “F: Frequency” and “S: Severity” in the PRO-CTCAE. “Vomiting” is examined by frequency and not intensity. As such, compared with “Severity,” “Frequency” may have better agreement with the true outcomes. Likewise, in the SIS, “Vomiting” was expressed as “Severity” and not “Frequency.” Compared the PRO-CTCAE, the SIS

had a much worse κ coefficient for “Frequency” but a similar κ coefficient for “Severity.” Further, “Vomiting” in the SIS had a higher κ coefficient than “Severity” in the PRO-CTCAE. Lastly, “Diarrhea” had poor agreement between the PRO-CTCAE and SIS, which may also be due to the “Frequency” and “Severity” difference seen for “Vomiting.” These findings that emojis were not good indicators of frequency suggest a need for revisions such as using more suitable text and emoji stickers.

In cohort two, we interpreted categorical values with continuous values and assessed associations between the PRO-CTCAE and SIS. We found a generally favorable responsiveness for the SIS between baseline and post-administration. Since the expected side effects vary widely depending on the treatment regimen, each side effect was evaluated on a regimen-by-regimen basis. In this analysis, we assessed the correlation of changes between two time points. To validate responsiveness, we did not use effect size (mean of the second time − mean of the first time/SD of the first time) or standardized response means (mean of the second time − mean of the first time/SD of the change) because we assumed that the symptoms in this cohort were relatively stable with little variability and a small standard deviation. However, a question remains on the appropriateness of replacing categorical variables with continuous variables; i.e., whether each has a guaranteed equal priority [15]. Although some authors perceive responsiveness as the most important characteristic of an evaluative tool, the proper way to assess responsiveness is not apparent [24]. Terwee et al. concluded that a distinct measure of responsiveness leads to a distinct conclusion because of a distinct objective [24]. Further discussion is needed to determine the most appropriate method to validate responsiveness.

Without considering dropouts, an estimated sample size of 14 patients was needed for each treatment group in cohort two. However, the relatively small sample size of the paclitaxel group (n=17) requires a comment. In cohort two, patients were registered from December 2019 to October 2020. At that time, COVID-19 began to spread in Japan, and a reduced number of hospital visits was necessary to control the outbreak. Therefore, many physicians chose to administer docetaxel every 3 weeks for 4 doses rather than paclitaxel every week for 12 doses. Due to COVID-19, extending the enrollment period would not have increased the number of enrolled cases. Nevertheless, the study was completed as planned.

This study had several limitations. First, it was limited to female, Japanese patients with breast cancer. Many

factors, such as gender, age, type of cancer and disease, religious, cultural background, and social media platform, affect individual use of emojis. As such, each situation may need a different SIS. Second, cohort one had a small sample size. We estimated that 100 cases were needed in this cohort, but we were able to analyze only 70 cases. This small sample size may lead to a bias. Third, as mentioned above, the appropriate approach to validation of responsiveness is uncertain. Forth, as it has been validated with subjects who have a good PS, it has not been validated with subjects who have a poor PS. It is unknown whether similar performance to SIS can be obtained even with subjects who have poor PS.

In conclusion, we developed an original SIS from the PRO-CTCAE for patients with breast cancer and verified the validity, reliability, and responsiveness of the SIS. It should be noted that some of the items were correlated between PRO-CTCAE and SIS, while others were low, such as “Decreased appetite”. Different situations may be captured by different scales. Further improvement and validation are needed to clarify the differences and commons from the original PRO-CTCAE and to create a scale that more truly captures patient symptoms.

Acknowledgements

This study was funded by Pfizer Inc.

Disclosure of potential conflicts of interest

Tadahiko Shien, Yuichiro Kikawa, and Naruto Taira received an honorarium from Pfizer Inc.

Compliance with ethical standards

Research involving Human Participants and/or Animals

This study involving human participants was in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Figure legends

Figure 1 Correlation coefficients, κ coefficients, and concordance rates between the PRO-CTCAE and SIS **at T1**

The forest plots show correlation coefficients between the PRO-CTCAE and SIS at T1. κ coefficients and concordance rates are also shown. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; ROC, rate of concordance; S, severity.

Figure 2 Test-retest reliability between T1 and T2

The forest plots show κ coefficients and 95% confidence intervals for all patients between T1 and T2. Concordance rates are also shown. A, amount; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/Absence; PRO, patient-reported outcome; ROC, rate of concordance; S, severity.

Figure 3 Correlation coefficient of score changes from T1 to T2 between the PRO-CTCAE and SIS in the anthracyline group

The forest plots show correlation coefficients of score changes from T1 to T2 between the PRO-CTCAE and SIS. A, amount; CC: corelation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/Absence; PRO, patient-reported outcome; S, severity.

Figure 4 Questionnaire for Symptom Illustration Scale **v.1.0**

English translations were added to the original Japanese version for publication: 1) Decreased appetite, 2) Nausea, 3) Vomiting, and 4) Constipation, continued in Sup. Figure 1.

Supplementary Figure Legends

Figure S1 Correlation coefficients, κ coefficients, and concordance rates between the PRO-CTCAE and SIS at T2

The forest plots show correlation coefficients between the PRO-CTCAE and SIS at T1. κ coefficients and concordance rates are also shown. A, amount; CC: correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; ROC, rate of concordance; S, severity.

Figure S2 Correlation coefficients, κ coefficients, and concordance rates between the PRO-CTCAE and SIS at T1 in the PRO-CTCAE group

The forest plots show correlation coefficients between the PRO-CTCAE and SIS at T1 in the PRO-CTCAE group. κ coefficients and concordance rates are also shown. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient- reported outcome; ROC, rate of concordance; S, Severity.

Figure S3 Correlation coefficients, κ coefficients, and concordance rates between the PRO-CTCAE and SIS questionnaires at T1 in the SIS group

The forest plots show correlation coefficients between the PRO-CTCAE and SIS at T1 in the SIS group. κ coefficient and concordance are also shown. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; ROC, rate of concordance; S, severity.

Figure S4 Test-retest reliability between T1 and T2 in the PRO-CTCAE group

The forest plots show the κ coefficient and 95% confidence interval between T1 and T2 in the PRO-CTCAE group. Concordance rates are also shown. A, amount; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; ROC, rate of concordance; S,

severity.

Figure S5 Test-retest reliability between T1 and T2 in the SIS group

The forest plots show κ coefficients and 95% confidence intervals between T1 and T2 in the SIS group. Concordance rates are also shown. A, amount;. CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; ROC, rate of concordance; S, severity.

Figure S6 Correlation coefficients for score changes from T1 to T3 between the PRO-CTCAE and SIS questionnaires in the anthracyline group

The forest plots show correlation coefficients of score changes from T1 to T3 between the PRO-CTCAE and SIS. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; S, severity.

Figure S7 Correlation coefficients for score changes from T1 to T2 between the PRO-CTCAE and SIS questionnaires in the docetaxel group

The forest plots show correlation coefficients of score changes from T1 to T2 between the PRO-CTCAE and SIS. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, Interference; P, presence/absence; PRO, patient-reported outcome; S, severity.

Figure S8 Correlation coefficients for score changes from T1 to T3 between the PRO-CTCAE and SIS in the docetaxel group

The forest plots show correlation coefficient of score changes from T1 to T3 between the PRO-CTCAE and SIS. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO: patient-reported outcome; S, severity.

Figure S9 Correlation coefficients for score changes from T1 to T2 between the PRO-CTCAE and SIS in the paclitaxel group

The forest plots show correlation coefficients of score changes from T1 to T2 between the PRO-CTCAE and SIS. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, Frequency; I, Interference; P, presence/absence; PRO, patient-reported outcome; S, severity.

Figure S10 Correlation coefficients for score changes from T1 to T3 between the PRO-CTCAE and SIS in the paclitaxel group

The forest plots show the correlation coefficient of score changes from T1 to T3 between the PRO-CTCAE and SIS. A, amount; CC, corelation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; S, severity.

Figure S11 Correlation coefficients for score changes from T1 to T2 between the PRO-CTCAE and SIS in the endocrine cohort

The forest plots show correlation coefficients of score changes from T1 to T2 between the PRO-CTCAE and SIS. A, amount; CC, correlation coefficient; CI, confidence interval; SIS, Symptom Illustration Scale; F, frequency; I, interference; P, presence/absence; PRO, patient-reported outcome; S, severity.

Figure S12 Questionnaire for the Symptom Illustration Scale v.1.0

English translations were added to the original Japanese version for publication: 5) Diarrhea, 6) Shortness of breath, 7) Swelling, 8) Rash, 9) Hair loss, 10) Numbness & tingling, 11) General pain, 12) Joint pain, 13) Insomnia, 14) Fatigue, 15) Anxious, 16) Vaginal discharge, 17) Vaginal dryness, and 18) Hot flashes, continued from Figure 1.

References

1. Harbeck N, Iyer S, Turner N, et al (2016) Quality of life with palbociclib plus fulvestrant in previously treated hormone receptor-positive, HER2-negative metastatic breast cancer: patient-reported outcomes from the PALOMA-3 trial. Ann Oncol Off J Eur Soc Med Oncol 27:1047–1054. https://doi.org/10.1093/annonc/mdw139

2. LeBlanc TW, Abernethy AP (2017) Patient-reported outcomes in cancer care - hearing the patient voice at greater volume. Nat Rev Clin Oncol 14:763–772. https://doi.org/10.1038/nrclinonc.2017.153

3. Takashima T, Mukai H, Hara F, et al (2016) 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 17:90–98. https://doi.org/10.1016/S1470-2045(15)00411-8

4. Basch E, Iasonos A, Barz A, et al (2007) Long-term toxicity monitoring via electronic patient-reported outcomes in patients receiving chemotherapy. J Clin Oncol Off J Am Soc Clin Oncol 25:5374–5380. https://doi.org/10.1200/JCO.2007.11.2243

5. Judson TJ, Bennett AV, Rogak LJ, et al (2013) Feasibility of long-term patient self-reporting of toxicities from home via the Internet during routine chemotherapy. J Clin Oncol Off J Am Soc Clin Oncol 31:2580–2585. https://doi.org/10.1200/JCO.2012.47.6804

6. Wallwiener M, Matthies L, Simoes E, et al (2017) Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial. J Med Internet Res 19:e322. https://doi.org/10.2196/jmir.8210

7. Kikawa Y, Hatachi Y, Rumpold G, et al (2019) Evaluation of health-related quality of life via the Computer-Based Health Evaluation System (CHES) for Japanese metastatic breast cancer patients: a single-center pilot study. Breast Cancer Tokyo Jpn 26:255–259. https://doi.org/10.1007/s12282-018-0905-1

8. Yang J, Weng L, Chen Z, et al (2019) Development and Testing of a Mobile App for Pain Management Among Cancer Patients Discharged From Hospital Treatment: Randomized Controlled Trial. JMIR MHealth UHealth 7:e12542. https://doi.org/10.2196/12542

9. Nipp RD, El-Jawahri A, Ruddy M, et al (2019) Pilot randomized trial of an electronic symptom monitoring intervention for hospitalized patients with cancer. Ann Oncol Off J Eur Soc Med Oncol 30:274–280. https://doi.org/10.1093/annonc/mdy488

10. Velikova G, Booth L, Smith AB, et al (2004) Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. J Clin Oncol Off J Am Soc Clin Oncol 22:714–724. https://doi.org/10.1200/JCO.2004.06.078

11. Berry DL, Hong F, Halpenny B, et al (2014) Electronic self-report assessment for cancer and self-care support: results of a multicenter randomized trial. J Clin Oncol Off J Am Soc Clin Oncol 32:199–205. https://doi.org/10.1200/JCO.2013.48.6662

12. Basch E, Deal AM, Kris MG, et al (2016) Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. J Clin Oncol Off J Am Soc Clin Oncol 34:557–565. https://doi.org/10.1200/JCO.2015.63.0830

13. Denis F, Lethrosne C, Pourel N, et al (2017) Randomized Trial Comparing a Web-Mediated Follow-up With Routine Surveillance in Lung Cancer Patients. JNCI J Natl Cancer Inst 109:. https://doi.org/10.1093/jnci/djx029

14. Bond DM, Hammond J, Shand AW, Nassar N (2020) Comparing a Mobile Phone Automated System With a Paper and Email Data Collection System: Substudy Within a Randomized Controlled Trial. JMIR MHealth UHealth 8:e15284. https://doi.org/10.2196/15284

15. US National Cancer Institute. Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™). http:// outcomes.cancer.gov/tools/pro-ctcae.html. Accessed 10 Apr 2019. 50:

16. Garra G, Singer AJ, Taira BR, et al (2010) Validation of the Wong-Baker FACES Pain Rating Scale in pediatric emergency department patients. Acad Emerg Med Off J Soc Acad Emerg Med 17:50–54. https://doi.org/10.1111/j.1553-2712.2009.00620.x

17. Lee ACK, Tang SW, Yu GKK, Cheung RTF (2008) The smiley as a simple screening tool for depression after stroke: a preliminary study. Int J Nurs Stud 45:1081–1089. https://doi.org/10.1016/j.ijnurstu.2007.05.008

18. Tang Y (2019) Emoticon, Emoji, and Sticker Use in Computer-Mediated Communication: A Review of Theories and Research Findings. 27

19. Lai D, Lee J 8, He S (2021) Emoji for the Medical Community-Challenges and Opportunities. JAMA 326:795–796. https://doi.org/10.1001/jama.2021.8409

20. Van Dam L, Rietstra S, Van der Drift E, et al (2019) Can an Emoji a Day Keep the Doctor Away? An Explorative Mixed-Methods Feasibility Study to Develop a Self-Help App for Youth With Mental Health Problems. Front Psychiatry 10:593. https://doi.org/10.3389/fpsyt.2019.00593

21. Bai Q, Dan Q, Mu Z, Yang M (2019) A Systematic Review of Emoji: Current Research and Future Perspectives. Front Psychol 10:2221. https://doi.org/10.3389/fpsyg.2019.02221

22. Bieri D, Reeve RA, Champion DG, et al (1990) The Faces Pain Scale for the self-assessment of the severity

of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. Pain 41:139–150. https://doi.org/10.1016/0304-3959(90)90018-9

23. Macdonald L, Bruce J, Scott NW, et al (2005) Long-term follow-up of breast cancer survivors with post-mastectomy pain syndrome. Br J Cancer 92:225–230. https://doi.org/10.1038/sj.bjc.6602304

24. Terwee CB, Dekker FW, Wiersinga WM, et al (2003) On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 12:349–362. https://doi.org/10.1023/a:1023499322593

Statement & Declarations

Funding: This study was funded by Pfizer Inc.

Competing Interests: Tadahiko Shien, Yuichiro Kikawa, and Naruto Taira received an honorarium from Pfizer Inc.

Author Contributions: All authors contributed to the study and design. Material preparation, data collection, and analysis were performed by Yoko Suzuki, Naruto Taira, Takayuki Iwamoto, Yuichiro Kikawa, Yukari Uemura, Yasuhiro Hagiwara, and Seiichiro Yamamoto. The first draft of the manuscript was written by Yoko Suzuki and all authors commented on versions of the manuscript. All authors read and approved the final manuscript.

Ethics approval: This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Okayama University Hospital (Cohort one: Date 2019.5.28/No. 1907-038, Cohort two: Date 2019.11.29/No. 1907-042).

Consent to participate: Written informed consent was obtained from all participants in the study.

Consent to publish: The authors affirm that all patients provided informed consent for publication of all figures and tables.