Usability testing of EirV3 – a computer-based tool for patient reported outcome measures in cancer

Hilde Krogstad [*1 3] and Stine Marie Sundt-Hansen [*1 4], Marianne Jensen Hjermstad[2], Liv Ågot Hågensen[7], Stein Kaasa[2 6], Jon Håvard Loge[2 5], Sunil X. Raj[1 3], Aslak Steinsbekk[8] and Kari Sand[1 3]

1. European Palliative Care Research Centre (PRC), Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, and St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway
2. European Palliative Care Research Centre (PRC), Department of Oncology, Oslo University Hospital and Institute of Clinical Medicine, University of Oslo, Oslo, Norway
3. Cancer Clinic, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway
4. Department of Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
5. Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Norway
6. Department of Oncology, Oslo University Hospital, Oslo, Norway
7. Coordination Unit Orkdal Region, Orkdal, Norway
8. Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

*shared first authorship

Number of tables: 4
Number of figures: 2
Number of references: 42
Number of words:

Corresponding author:

Hilde Krogstad

European Palliative Care Research Centre (PRC), Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, and Cancer Clinic, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway
Postbox 8905, NO-7491 Trondheim, Norway
E-mail: hilde.krogstad@ntnu.no,
Phone: +47 06800

ORCID: 0000-0002-2514-8154

Abstract

Purpose: Eir Version 3 (V3) is an electronic tool for administration of patient reported outcome measures (*Eir-Patient)* that immediately presents patient scores on the physician´s computer (*Eir-Doctor*). Perceived usability is an important determinant for successful implementation. The aim of this study was to answer the following research question evaluated at the cancer outpatient clinics, in the patients´ home, and at general practitioners´ (GP) offices: What are the number, type, and severity of usability issues evaluated by the patient (*Eir-Patient* module) and by the physician (*Eir-Doctor* module)?

Methods: A usability evaluation using observations, think-aloud sessions, individual interviews and focus group interviews in cancer patients and their physicians was conducted. Identified usability issues were graded on a severity scale from 1 (irritant) to 4 (unusable).

Results: Overall, 73 Eir registrations were performed by 37 patients, and used by 17 physicians in clinical consultations. All patients were able to complete the *Eir-Patient* symptom registration. Seventy-two usability issues were identified. None of them were graded as unusable. For the *Eir-Patient* module, 62% of the identified usability issues were graded as irritant (grade 1), 18% as moderate (grade 2) and 20% as severe (grade 3). For the *Eir-Doctor* module, 46% of the identified usability issues were graded as irritant, 36% as moderate and 18% as severe.

Conclusions: In the subsequent Eir-version, issues in the severe and moderate categories have been changed, to optimize the usability of using real time PROMs in clinical practice.

Keywords: Patient reported outcome measures, PROMs, usability, feasibility, electronic patient-reported outcomes

Introduction

Patient-centered care is essential to the mission of health care in terms of content, quality and appropriate use of services [1,2], and involves focusing on those elements of care, support and treatment that matters most to the patient, their family and carers [3]. This in turn supports patients` knowledge, skills and confidence in managing and making informed decisions about their treatment and care [3]. The increased focus on patient-centered care in health care in general and even more so in oncology, makes patients' report of symptoms and preferences for care paramount [4]. Routine use of patient-reported outcome measures (PROMs) benefits communication, better prognostic awareness and patient satisfaction [4-6], and may improve quality of life (QoL) [4]. It is even suggestive of prolonged survival [7].

PROMs cover all subjective outcome measures related to health, level of functioning and well-being as reported by the patients [8]. As the patient is the primary source of information, PROMs supplement clinical observations and objective findings, and play a central role throughout the whole cancer disease trajectory. Systematic symptom assessment in clinical practice is important to improve the quality of diagnostics and treatment [9,10]. However, systematic symptom assessment by using PROMs in direct patient care still remains uncommon in oncology [11,12].

Electronic collection of PROMS provides a platform for consistent documentation, and gives health care providers a real-time presentation of the patients´ situation when outside of hospital. Furthermore, the development of symptoms can be easily reviewed [13]. Studies have shown that electronic collection of PROMs is feasible in patients with advanced cancer [14,13]. Comparative studies examining equivalence between electronic and paper based PROMs have shown excellent agreement between the two methods [15,16]. Overall, 41% of the patients preferred assessment on tablets, 19 % preferred paper, while 40 % had no preference [15].

Uptake of electronic administration of PROMs relies on perceived usefulness in the sense that it is regarded as beneficial to use in the actual context as well as a high degree of perceived usability by the end-users, e.g. patients and health care providers. Usability is commonly defined as the extent to which a product can be used to achieve specific goals with effectiveness, efficiency and satisfaction in a defined context [17]. Usability testing is a critical step in the development of an electronic tool, as it can inform about the ease of use, and guide further development and implementation [18]. Repeated (iterative) testing on usability of any electronic medical system is the preferred method for identifying and solving usability issues [19,18].

Eir version 3 (EirV3) is an electronic symptom assessment tool developed by the European Palliative Care Research Centre (PRC) [20] at the Norwegian University of Science and Technology (NTNU) and Trondheim University Hospital. EirV3 has been developed for and by end-users through iterative development processes including regular user testing and continuous amendments [15]. EirV3 has two modules: *Eir-Patient* and *Eir-Doctor*. *Eir-Patient* is for patient self-report on tablets or computers. When the physician logs on to *Eir-Doctor* on the computer, the patient's PROMs registrations have already been wirelessly transferred and transformed to a special format designed for immediate use in clinical consultations. As EirV3 is intended for use in outpatients, in-patients and in primary care as well as different groups of cancer patients with respect to cancer diagnosis, age and sociodemographics, usability testing in a variety of settings is warranted.

The aim of this study was to answer the following research question evaluated at the cancer outpatient clinics, in the patients´ home, and at general practitioners´ (GP) offices: What are the number, type, and severity of usability issues evaluated by the patient (*Eir-Patient* module) and by the physician (*Eir-Doctor* module)? More specifically, the focus was to gain new information about content, functionality and barriers experienced by the end-users. Intentions were to make changes based on the results to improve the tool´s usability.

Methods

This was a usability study using observation, think-aloud techniques and interviews with patients and physicians. The study was performed from September 2015 to September 2017. Patients were included at a local and at a university hospital. Use of EirV3 took place in three different settings: 1) cancer outpatient clinics at a local and at a university hospital, 2) in the cancer patients' home, and 3) at GP offices in two municipalities. The focus for these data collections was to gain new information about usability issues regarding content and functionality. In other words, what are the barriers experienced by the end users by using the Eir system.

The Eir modules

The *Eir-Patient* module includes questions assessing 19 of the most common cancer-related symptoms, and five supplemental questions assessing level of functioning and nutritional status [15]. *Eir-Patient* has a hierarchical structure starting with dichotomous assessments (yes/no) regarding the presence of initial symptoms experienced

by cancer patients. If endorsed, patients are asked to score symptom intensities on a 0-10 scale (Figure 1). If symptom intensity was above 1, in-depth questions were presented for the following symptoms: pain, breathlessness, depression, anxiety, insomnia, constipation, vomiting and diarrhea. Finally, five questions on QoL, physical functioning, nutritional intake, height and body weight are to be completed by all patients, irrespective of their reported symptom burden [15]. Details on the EirV3 development process have been presented previously [15]. The content in *Eir-patient* is based on literature reviews, clinical experience, evidence-based guidelines and well-validated forms, i.e. EAPC basic dataset [21], Patient-Generated Subjective Global Assessment (PG-SGA) [22], Patient Health Questionnaire-9 (PHQ-9) [23], General Anxiety Disorder-2 (GAD-2) [24], Insomnia Severity Index [25] and Common Toxicity Criteria for Adverse Events (CTCAE 4.0) [26].

When the physician opens *Eir-Doctor* on the computer and identifies and selects a particular patient in the list of patients, the patient's present responses are displayed, together with prior scores if they exist (Figure 2).

Subjects and data collection
Patients
Recruitment of patients was done by purposive sampling to ensure variation in age, gender, diagnosis and anticipated symptom burden. Eligible patients were diagnosed with any type of cancer in all stages of the disease trajectory. It was important to include a variety of patients to enhance the generalizability of results to the cancer populations seen at the sites of inclusion. Thus, including patients with a very high symptom burden or those having advanced stage cancer was important. Patients with obvious cognitive impairment, as judged by the attending physician according to established criteria (e.g. orientation, memory, attention span) were not included. All participants were above 18 years with no upper age limit. EirV3 was completed either (1) while waiting for a scheduled consultation at the cancer outpatient clinic, (2) at home between consultations, or (3) prior to a scheduled consultation at the GP's office. The patients who had a GP in the two municipalities, were encouraged to visit their own GP. Before a scheduled consultation at the GP´s office, the patients were observed while completing EirV3 at home.

Observation using the think-aloud method and patient interviews were used to collect data on the patients' practical use of EirV3, and to provide insight in their immediate reactions and experiences when using the tablet [27]. Patients who had used EirV3 on multiple occasions were interviewed after each completion of Eir to gain insight in usability issues that persisted over time.

Initially, all patients received a brief introduction to *Eir-Patient*. Patients were encouraged to think aloud, i.e. to constantly verbalise their thoughts, as they responded to the Eir-questions [28]. If they were unsure on how to proceed, they were encouraged to do what they found most intuitive, before being assisted by the researcher (H.K., S.SH.). Data were collected by observing the patients as they used EirV3. Field-notes were made based on a predefined observation template covering navigation errors, ease of use, apparent misunderstandings or technical difficulties. After completion, each patient took part in a structured interview, following an interview guide designed specifically for this study. The content of the interview guide was based on previous usability studies of electronic symptom assessment tools [18,29], with standardized, open questions about potential difficulties regarding understandability, practical use, design, layout and time expenditure. Specific usability issues that had been observed when the patient used EirV3, were also addressed in the interview. The whole session was audio-recorded. The same guide was followed for repeated use of EirV3.

Physicians
All hospital physicians were oncologists, and they were recruited from cancer outpatient units at a university hospital and at a collaborating local hospital. General practitioners (GPs) were recruited from two municipalities.

Prior to study start, physicians received a quick introduction to *Eir-Doctor*. A thorough instruction was not given, since the intention was to evaluate EirV3 in their clinical work. The physicians' were observed by the researchers while using EirV3. Field notes were made based on a predefined observation template covering the use of EirV3 before and during the consultation. By the end of the study, physicians were invited to attend focus group interviews or individual interviews to discuss and summarize their experience with the tool. The physicians were asked questions about how they used EirV3 before and during consultations, difficulties regarding use of EirV3, potential benefits or disadvantages and suggested changes. Inclusion and exclusion criteria for patients and physicians are provided in Table 1.

Analysis

All audio-recorded material (think aloud-sessions and interviews) was transcribed verbatim by one of the first authors and was analysed together with the field notes.

Usability issues were identified and categorised by use of simple content analysis, and rated by the authors (SSH, HK , KS.), guided by the approach by Rubin and Chisnell [27]. Identified usability issues were categorised as follows: *Understandability, visibility, workflow, content, navigation and bugs*. The number of participants experiencing each issue was registered. Each issue was graded on a scale from 1 to 4 (*1=irritant, 2=moderate, 3=severe, and 4=unusable*) [27], based on the severity of the problem, frequency and potential for affecting treatment. The grading was done by the authors (SSH, HK, KS) independently and subsequently discussed until consensus was reached on each issue of divergence.

Ethical considerations

Confidentiality issues and adherence to all regulations regarding the registration, transfer, handling and storage of data were major issues during the development process. No data were stored on the tablets. The patients logged on to EirV3 using a randomly generated study ID securing their anonymity. Accordingly, the list of patients presented to the physicians only included the patients' study ID. Patients received oral and written information about the study. Informed consent was obtained from all individual participants included in the study. Approval was obtained from the Regional Committee for Medical and Health Research Ethics (REK-2014/212 and REK2015/185).

Results

Seventy-three Eir registrations were performed by 37 patients, and 17 physicians used the results in *Eir-Doctor* (Table 2). Twenty-two (60%) of the patients received treatment with palliative intention, while 40% received curative treatment. Gastrointestinal cancers were most common (32%). Mean age was 64 years (SD 11.3), and 40% of the patients were male. The median number of reported symptoms were 7 (1-15). Median Karnofsky score was 80 (50-100). Fifty-one percent of the patients used tablets every day, while 35% had never used a tablet prior to the study.

Seventeen physicians were enrolled. Five worked as GPs and 12 as oncologists. Mean age for physicians was 48 years (SD 11.7)

Use of EirV3

Of the 73 Eir registrations completed by patients, 53 were completed at the outpatient clinics, 10 at home and 10 at the GP´s office. The 17 physicians used EirV3 in a total of 59 consultations (range 1–5). No technological difficulties were encountered in any of the three settings. All participants managed to use EirV3, and the median time used to complete the *Eir-Patient* the first time was 10 minutes and 20 seconds.

Usability issues

In total, 72 usability issues were identified in *Eir-Patient* and *Eir-Doctor*. 35% of the issues were categorized as "content" (whether the displayed information is appropriate, consistent and accurate), 21% of the issues were categorized into the "understandability" category (to which extent the text or image or task is comprehensible), 13% as navigation, 16% as workflow, 11% as visibility while 4% of the issues were categorized as "bugs" (error or defect in the software that causes malfunction).

Usability issues in *Eir-Patient*

All patients were able to complete the symptom registration on EirV3. On average, the patients reported seven symptoms per completion. Four patients reported that completion was challenging due to a high symptom burden, which led to more in-depth and follow-up questions and problems with registration of their responses because the screen was not sufficiently sensitive when being touched, or due to poor vision. Thirty-three patients found the program easy to use and the navigation by touching the arrow icons on the screen intuitive.

A total of 44 usability issues were identified for *Eir-Patient*. Of these, 27/44 (62%) were assessed as grade 1 (irritant) and 8/44 (18%) issues as grade 2 (moderate). Examples of irritant usability issues were spelling errors and unclear wording. Examples of usability issues graded as moderate (grade 2) were missing questions on urinary problems, and that the response alternatives regarding e.g. pain descriptors did not fit with how the patient would describe this. No issues were ranked as 4 (unusable), while 9/44 (20%) issues were graded as 3 (severe). Usability issues with grade 3 within each category are described in Table 3, as these issues were the most urgent to resolve.

Usability issues in *Eir-Doctor*

Observation of physicians using *Eir-Doctor* in clinical consultations showed that the module was intuitively easy to use, that it gave a good overview of the patient´s symptoms, and clearly depicted which symptoms were the most troublesome (Figure 1). The automatic ranking of symptom intensity from highest to lowest was actively used in all consultations, as was the graph that showed development of symptom intensity over time. The physicians also found the pain body map informative. The patient's answers to follow-up questions for symptoms other than pain were rarely used or noticed by the physician. A total of 28 usability issues were identified. Of these, 46% were rated as grade 1 (irritant) and 36% as grade 2 (moderate). 18% of issues were rated as grade 3 (severe) (Table 4). None were found unusable.

Based on the identified usability issues, EirV3 has been improved. In summary, the most important changes made in EirV3´s content and functionality were a) making the questions in the symptom screening section in *Eir-Patient* mandatory, b) removing a question in *Eir-Patient* regarding self-care and c) adding text to describe anchors, e.g. wellbeing where 0=great and 10=worst imaginable, and to clarify the difference between the terms drowsiness and tiredness, d) disable double-click in *Eir-Patient* in order to avoid patients unintentionally skipping questions if they double tapped the screen, e) improving the accuracy of the pain body map in *Eir-Doctor*, to prevent overlap of marked pain areas.

Discussion

This study presents results from a usability test of EirV3, an electronic symptom assessment tool, in different care settings and in patients at different stages of cancer. EirV3 was found easy to use by most patients and physicians, was usable in multiple settings, and perceived relevant regardless of diagnosis and treatment. The majority of identified usability issues were classified as irritant, while none were ranked as unusable.

All participants managed to complete EirV3. A total of 72 identified usability issues is apparently a high number given that Eir had been subject to extensive and iterative test rounds before. However, almost the same number, 65 issues, were identified in a similar study [30]. Potential reasons for the high number of issues might be that audio recordings were not used in the previous test rounds. Thus, it is highly likely that important details regarding usability were not identified. Moreover, previous testing did not include repeated assessments, and as such, usability issues regarding e.g. symptom development over time were not revealed. Based on our experiences, we recommend a combination of observations with think-aloud method and individual interviews in studies like this, as we think it makes researchers identify more end-user barriers [31], and ensure a comprehensive understanding of the use and experience of the tool.

None of the identified usability issues made it difficult to complete EirV3. However, some of them caused frustration among patients. For example, in the follow-up section for depression, some patients had to answer questions about sleep and food intake which were perceived as overlapping with questions they had already answered. Removing overlapping questions would probably have reduced this frustration. Dynamic tools make it possible to reduce overlapping items. Thus, reducing the number of overlapping questions can weaken the validity of the tool. However, this could be worthwhile as it is likely to improve the user experience of the tool. When developing an assessment tool, the desired amount of information must be balanced against the perceived burden of completing the assessment tool, in order to reduce the amount of missing items [32]. On the other hand, an instrument without sufficient comprehensiveness to disclose the patients' problems has less clinical value [33].

Not all identified usability issues were resolved during the test period due to hardware, software, funding or resource constraints as reported in other studies [34,30]. Even if the intention was to change the elements with the highest severity rating, some proved difficult to change. One example is that the patients' trouble using the touchscreen due to dry fingertips is related to the hardware functioning, not the software itself. This has also been identified for other tools [35,36]. Other identified usability issues did not lead to immediate changes, but gave important knowledge before implementation. For example, some physicians asked the patients about weight or last bowel movement, even though this information could be found in EirV3. It was also observed that while the pain body map in *Eir-Doctor* was frequently used, the answers to follow-up questions for other symptoms were rarely noticed and consequently not addressed by physicians. This observation may be related to the lack of systematic attention to symptoms in general during an oncological consultation. Pain has been in focus as a key symptom for decades during undergraduate and graduate training which may explain why this symptom is given more attention. More thorough instructions for the users were suggested instead of changing the tool. However, that may not be enough to change physicians´ attention and behavior. Systematic training may be needed as well.

Patient-centered care in oncology involves focusing on the impact of the disease and treatment on each patient´s life, along with anticancer treatment and at any stage of the disease trajectory [37]. Systematic use of *Eir-Patient* gives the patient the opportunity to report information directly to the physician´s computer. Repeated testing enables use of the graphical display of symptom development over time. This may direct the focus of the subsequent consultation to issues that are bothersome to the patient, and may lead to changes in treatment based on the patient´s needs.

A recent review of electronic PROMs distinguished between patient-centered systems and treatment-centered systems, with the first being more likely to provide user-friendly features and good scoring options, whereas the other had a more administrative focus with flexibility with respect to location, automated reminders and better clinical integration [38]. As of now, we regard EirV3 as a patient-centered system but with the aim being to incorporate features from treatment-centered systems in further development, e.g. *Eir-at-home*. Automatic incorporation of clinical decision support systems (CDSs) and evidence-based guidelines for symptom management will probably be incorporated in the future, although only marginal positive effects of CDSs on symptoms and PROMs have been documented so far [39].

EirV3 is still in development, and is currently implemented into the patient care pathways and clinical practice in a Norwegian cluster randomized trial on early integration of palliative care in oncology [40]. Further work should address how to successfully integrate EirV3 into daily clinical practice.

Limitations and strengths
Generally, about ten participants are considered adequate for robust usability testing [41]. In this study, 37 patients completed EirV3, of whom 18 did it more than once. The relatively high number of participants was necessary to provide a rich material from different settings, including a diverse sample of cancer patients, representing both genders and a wide range of age, diagnoses, functional levels, symptom burden, treatments and treatment intentions. None of the patients were in the lower age group, i.e. less than 35 years. However, it is highly unlikely that younger patients who are more used to all kinds of electronic devices, would find EirV3 more difficult to use than the older patients. The study was performed over a long period of time since it was carried out in different settings. Furthermore, quantitative measures were not used to evaluate usability. Initially, we intended to use the system usability scale (SUS), which is a simple, ten-item scale giving a global view of subjective assessments of usability [42]. Previous pre-tests however, revealed that the sickest patients did not manage to complete the SUS on paper, and thus the results were not representative for the whole sample. Nevertheless, all patients managed to complete EirV3, which indicates that electronic tools might be more user-friendly than questionnaires on paper. To fully ascertain this, specific testing of paper versus electronic tools should be performed in the frailest patients, which is a challenging task. However, the patient sample in our study consisted of many patients whose condition deteriorated over time and between follow-ups, with no increase in problems with completion of Eir,

Conclusion

Patients and physicians found EirV3 easy to use. Results indicate that EirV3 is usable in a heterogeneous population of cancer patients. In the subsequent Eir version, most issues in the severe and moderate categories have been changed to optimize the usability and feasibility of using real time PROMs in clinical practice.

References

1. Stacey D, Legare F, Col NF, Bennett CL, Barry MJ, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt A, Thomson R, Trevena L, Wu JH (2014) Decision aids for people facing health treatment or screening decisions. The Cochrane database of systematic reviews (1):Cd001431. doi:10.1002/14651858.CD001431.pub4

2. Richards T, Coulter A, Wicks P (2015) Time to deliver patient centred care. BMJ (Clinical research ed) 350:h530. doi:10.1136/bmj.h530

3. Person-centred care made simple. (2014). www.health.org.uk/publications/person-centred-care-made-simple.

4. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, Rogak L, Bennett AV, Dueck AC, Atkinson TM, Chou JF, Dulko D, Sit L, Barz A, Novotny P, Fruscione M, Sloan JA, Schrag D (2016) Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 34 (6):557-565. doi:10.1200/jco.2015.63.0830

5. Atkinson TM, Li Y, Coffey CW, Sit L, Shaw M, Lavene D, Bennett AV, Fruscione M, Rogak L, Hay J, Gonen M, Schrag D, Basch E (2012) Reliability of adverse symptom event reporting by clinicians. Qual Life Res 21 (7):1159-1164. doi:10.1007/s11136-011-0031-4

6. Strasser F, Blum D, von Moos R, Cathomas R, Ribi K, Aebi S, Betticher D, Hayoz S, Klingbiel D, Brauchli P, Haefner M, Mauri S, Kaasa S, Koeberle D (2016) The effect of real-time electronic monitoring of patient-reported symptoms and clinical syndromes in outpatient workflow of medical oncologists: E-MOSAIC, a multicenter cluster-randomized phase III study (SAKK 95/06). Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 27 (2):324-332. doi:10.1093/annonc/mdv576

7. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, Schrag D (2017) Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. Jama 318 (2):197-198. doi:10.1001/jama.2017.7156

8. Food and Drug Administration USDoHaHS (2009) Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical product Development to Support Labeling Claims. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm193282.pdf.

9. Zikos E, Ghislain I, Coens C, Ediebah DE, Sloan E, Quinten C, Koller M, van Meerbeeck JP, Flechtner HH, Stupp R, Pallis A, Czimbalmos A, Sprangers MA, Bottomley A (2014) Health-related quality of life in small-cell lung cancer: a systematic review on reporting of methods and clinical issues in randomised controlled trials. The Lancet Oncology 15 (2):e78-89. doi:10.1016/S1470-2045(13)70493-5

10. Gravis G, Marino P, Joly F, Oudard S, Priou F, Esterni B, Latorzeff I, Delva R, Krakowski I, Laguerre B, Rolland F, Theodore C, Deplanque G, Ferrero JM, Pouessel D, Mourey L, Beuzeboc P, Zanetta S, Habibian M, Berdah JF, Dauba J, Baciuchka M, Platini C, Linassier C, Labourey JL, Machiels JP, El Kouri C, Ravaud A, Suc E, Eymard JC, Hasbini A, Bousquet G, Soulie M, Fizazi K (2014) Patients' self-assessment versus investigators' evaluation in a phase III trial in non-castrate metastatic prostate cancer (GETUG-AFU 15). European journal of cancer 50 (5):953-962. doi:10.1016/j.ejca.2013.11.034

11. Kroenke K, Monahan PO, Kean J (2015) Pragmatic characteristics of patient-reported outcome measures are important for use in clinical practice. Journal of clinical epidemiology 68 (9):1085-1092. doi:10.1016/j.jclinepi.2015.03.023

12. McGrail K, Bryan S, Davis J (2011) Let's all go to the PROM: the case for routine patient-reported outcome measurement in Canadian healthcare. HealthcarePapers 11 (4):8-18; discussion 55-18

13. Benze G, Nauck F, Alt-Epping B, Gianni G, Bauknecht T, Ettl J, Munte A, Kretzschmar L, Gaertner J (2017) PROutine: a feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. Annals of palliative medicine. doi:10.21037/apm.2017.07.05

14. Hjermstad MJ, Lie HC, Caraceni A, Currow DC, Fainsinger RL, Gundersen OE, Haugen DF, Heitzer E, Radbruch L, Stone PC, Strasser F, Kaasa S, Loge JH (2012) Computer-based symptom assessment is

feasible in patients with advanced cancer: results from an international multicenter study, the EPCRC-CSA. J Pain Symptom Manage 44 (5):639-654. doi:10.1016/j.jpainsymman.2011.10.025

15. Krogstad H, Brunelli C, Sand K, Andersen E, Garresori H, Halvorsen T, Haukland EC, Jordal F, Kaasa S, Loge JH, Løhre ET, Raj SX, Hjermstad MJ (2017) Development of EirV3: A Computer-Based Tool for Patient-Reported Outcome Measures in Cancer. JCO Clinical Cancer Informatics. doi:10.1200/CCI.17.0051

16. Wintner LM, Giesinger JM, Zabernigg A, Rumpold G, Sztankay M, Oberguggenberger AS, Gamper EM, Holzner B (2015) Evaluation of electronic patient-reported outcome assessment with cancer patients in the hospital and at home. BMC medical informatics and decision making 15:110. doi:10.1186/s12911-015-0230-y

17. International Organization for Standardization/International Electrotehcnical Commission (ISO/IEC) 9241-11 (1998) Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11- Guidance on usability.

18. Stinson J, Gupta A, Dupuis F, Dick B, Laverdiere C, LeMay S, Sung L, Dettmer E, Gomer S, Lober J, Chan CY (2015) Usability testing of an online self-management program for adolescents with cancer. Journal of pediatric oncology nursing : official journal of the Association of Pediatric Oncology Nurses 32 (2):70-82. doi:10.1177/1043454214543021

19. Corrao NJ, Robinson AG, Swiernik MA, Naeim A (2010) Importance of Testing for Usability When Selecting and Implementing an Electronic Health or Medical Record System. Journal of Oncology Practice 6 (3):120-124. doi:10.1200/JOP.200017

20. European Palliative Care Research Centre (PRC), . www.ntnu.edu/PRC.

21. Sigurdardottir KR, Kaasa S, Rosland JH, Bausewein C, Radbruch L, Haugen DF (2014) The European Association for Palliative Care basic dataset to describe a palliative care cancer population: Results from an international Delphi process. Palliative medicine 28 (6):463-473. doi:10.1177/0269216314521264

22. Bauer J, Capra S, Ferguson M (2002) Use of the scored Patient-Generated Subjective Global Assessment (PG-SGA) as a nutrition assessment tool in patients with cancer. European journal of clinical nutrition 56 (8):779-785. doi:10.1038/sj.ejcn.1601412

23. Kroenke K, Spitzer RL, Williams JB (2001) The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine 16 (9):606-613

24. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Lowe B (2007) Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 146 (5):317-325

25. Morin CM (1996) Insomnia: Psychological Assessment and Management.

26. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 (2009).

27. Rubin J, Chisnell D (2008) Handbook of Usability Testing : How to Plan, Design, and Conduct Effective Tests. Wiley,

28. Lewis C (1982) Using the "Think Aloud" Method In Cognitive Interface Design. Yorktown Heights, NY. IBM : Technical report RC-9265

29. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, Tucker L, Hetherington R, Tse S, Spiegel L, Campillo S, Gill N, White M (2010) Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. Journal of medical Internet research 12 (3):e30. doi:10.2196/jmir.1349

30. Wolpin SE, Halpenny B, Whitman G, McReynolds J, Stewart M, Lober WB, Berry DL (2015) Development and usability testing of a web-based cancer symptom and quality-of-life support intervention. Health informatics journal 21 (1):10-23. doi:10.1177/1460458213495744

31. Bradbury K, Watts S, Arden-Close E, Yardley L, Lewith G (2014) Developing Digital Interventions: A Methodological Guide. Evidence-based Complementary and Alternative Medicine : eCAM 2014:561320. doi:10.1155/2014/561320

32. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ (2010) The routine use of patient reported outcome measures in healthcare settings. BMJ (Clinical research ed) 340:c186. doi:10.1136/bmj.c186

33. Walker J, Bohnke JR, Cerny T, Strasser F (2010) Development of symptom assessments utilising item response theory and computer-adaptive testing--a practical method based on a systematic review. Critical reviews in oncology/hematology 73 (1):47-67. doi:10.1016/j.critrevonc.2009.03.007

34. Breakey VR, Warias AV, Ignas DM, White M, Blanchette VS, Stinson JN (2013) The value of usability testing for Internet-based adolescent self-management interventions: "Managing Hemophilia Online". BMC medical informatics and decision making 13:113. doi:10.1186/1472-6947-13-113

35. Korzeniowski M, Kalyvas M, Mahmud A, Shenfield C, Tong C, Zaza K, Howell D, Brundage M (2016) Piloting prostate cancer patient-reported outcomesin clinical practice. Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer 24 (5):1983-1990. doi:10.1007/s00520-015-2949-5

36. Mirkovic J, Kaufman DR, Ruland CM (2014) Supporting cancer patients in illness management: usability evaluation of a mobile app. JMIR mHealth and uHealth 2 (3):e33. doi:10.2196/mhealth.3359

37. Jordan K, Aapro M, Kaasa S, Ripamonti CI, Scotté F, Strasser F, Young A, Bruera E, Herrstedt J, Keefe D, Laird B, Walsh D, Douillard JY, Cervantes A (2018) European Society for Medical Oncology (ESMO) position paper on supportive and palliative care. Annals of Oncology 29 (1):36-43. doi:10.1093/annonc/mdx757

38. Jensen RE, Snyder CF, Abernethy AP, Basch E, Potosky AL, Roberts AC, Loeffler DR, Reeve BB (2014) Review of electronic patient-reported outcomes systems used in cancer clinical care. Journal of oncology practice / American Society of Clinical Oncology 10 (4):e215-222. doi:10.1200/jop.2013.001067

39. Blum D, Raj SX, Oberholzer R, Riphagen, II, Strasser F, Kaasa S (2015) Computer-Based Clinical Decision Support Systems and Patient-Reported Outcomes: A Systematic Review. Patient 8 (5):397-409. doi:10.1007/s40271-014-0100-1

40. Palliative Care Integrated in Oncology (PALLiON). (2017). https://pallion.no/.

41. Virzi RA (1992) Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? Human Factors 34 (4):457-468. doi:10.1177/001872089203400407

42. Brooke J (1996) SUS: a "quick and dirty" usability scale, in Usability Evaluation in Industry. Jordan P.W., Thomas B., Weerdmeester B.A., McClelland I.L. (editor). Taylor & Francis: London:189-194

Table 1: Inclusion and exclusion criteria

	Physicians	**Patients**
Inclusion criteria	- At least three years of work experience as a physician - Provision of written informed consent	- A verified cancer diagnosis - Age 18 years or older - Provision of written informed consent - Having a GP in one of the nearby municipalities (only for patients recruited from the local hospital)
Exclusion criteria	- Member of Eir development group	Obvious cognitive impairment that may impact on the ability to use and understand EirV3, according to established clinical criteria (orientation, memory, abstract thinking) as judged by the patient´s physician

Table 2: Characteristics of included patients

Patient characteristics	Total n= 37
Age, mean (SD)	64 (11.3)
Gender, male n (%)	15 (40)
Treatment intention, palliative, n (%)	22 (59)
Ongoing chemotherapy, n (%)	15(41)
Ongoing radiotherapy, n (%)	6 (16)
Karnofsky score, median (range)	80 (50–100)
Number of symptoms, median (range)	7 (1-15)
Diagnosis:	
Breast cancer, n (%)	8 (22)
Gastrointestinal cancer, n (%)	12 (32)
Lymphomas, n (%)	4 (11)
Prostate cancer, n (%)	4 (11)
Gynecological cancer, n (%)	3 (8)
Lung cancer, n (%)	3 (8)
Malignant melanoma, n (%)	2 (5)
Testicular cancer, n (%)	1 (3)
Education:	
Elementary school, n (%)	8 (22)
High School, n (%)	15 (41)
Higher education, n (%)	13 (35)
Use of smartphones:	
Daily, n (%)	23 (67)
Weekly/sometimes, n (%)	3 (8)
Never, n (%)	8 (22)
Use of tablets:	
Daily, n (%)	19 (51)
Sometimes, n (%)	3 (8)
Never, n (%)	13 (35)

Table 3: Usability issues graded at level 3 (severe) in *Eir-Patient*, and possible resolutions

Usability issues	Quote/observation	Resolution/suggestion[1]
Understandability		
Physical function: - time-consuming to read all response options - having trouble finding suitable options - wanted to choose more than one option	*"There should have been more than one option to choose. Because if I go outside, I don´t run, but I walk fast"*	Resolution: Existing questions were removed due to difficulties finding the suitable option, and replaced with simpler alternatives
Self-care: - having trouble finding suitable options - the options did not necessarily reflect the need for home help	*"Washing clothes is not the most demanding task, taking a shower is much more demanding"*	Resolution: Existing questions were removed due to difficulties finding the suitable option.
Wellbeing ("how you feel overall"): - confusing that 0 equals best wellbeing, while 10 equals worst	*"In my head it just gets a bit confusing to read [this question]. It should be the opposite"*	Resolution: Added text to numbers: 0=great, 10=worst imaginable wellbeing
Content		
Follow up-questions regarding depression (PHQ-9) - confusing or too difficult due to contradictive questions	*"Bad appetite and eating too much: why are there two conflicting questions? I eat too much, but I do not have bad appetite. That was a stupid question"*	Suggestion: replace PHQ-9 with PHQ-2 (only two follow up questions)
Overlapping questions - overlapping follow-up questions for several symptoms, e.g. depression, sleep, nutrition	*"Didn't I already answer that question?"*	Suggestion: consider removing follow-up questions that are overlapping
Visibility		
Symptom screening: - a risk that patients unintentionally skip one of the three symptom screening pages if they double-clicked the screen.	One patient, who had pain, had not marked pain: *"were there any questions about pain then?"*	Resolution: Patient must either mark one or more symptoms, or tick of "neither of these"
Workflow		
Pain assessment:		

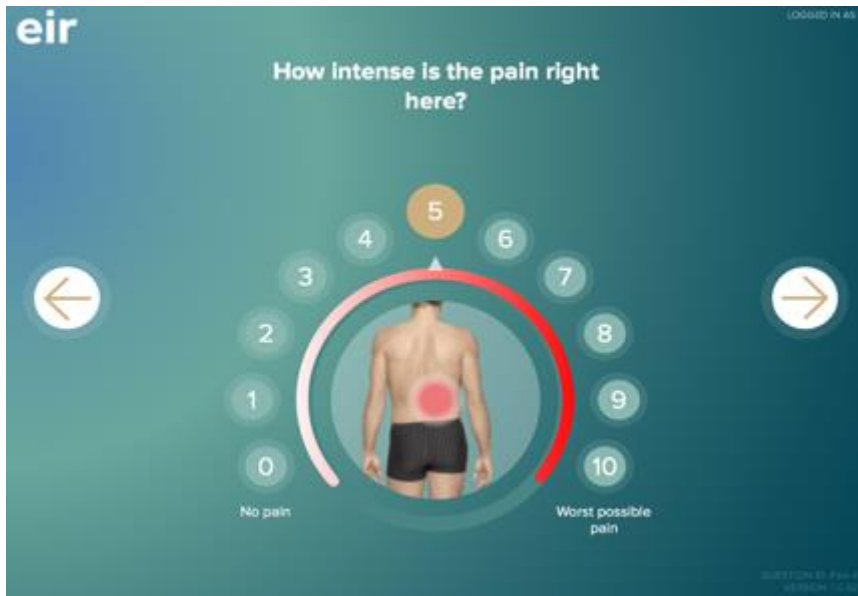|		
– difficulty understanding the term "breakthrough pain"	*"How intense is the breakthrough pain? But didn`t I have this question already?"*	Suggestion: Make the pain assessment less complicated. Simplify questions regarding breakthrough pain. Reformulate to make distinction between worst pain intensity and breakthrough pain intensity clear.
Pain body map – patients wanted to mark multiple pain locations on the same body map	*"If I marked two locations, the first one disappeared"*	Suggestion: make it possible to mark several locations on the same body map
Navigation		
Problems with the touch screen not responding to the patients' taps: – due to wrong technique	*"I have to use my fingers, not the stylus. There you see. How could I have used it at home, it's impossible"*	Resolution: Use a stylus of high quality, and spend more time on instructions and training for unexperienced tablet-users

[1]Resolution means that there has been a change in the program to solve the usability issue. Suggestion means that the issue is not resolved.

Table 4: Usability issues graded at level 3 (severe) in *Eir- Doctor*

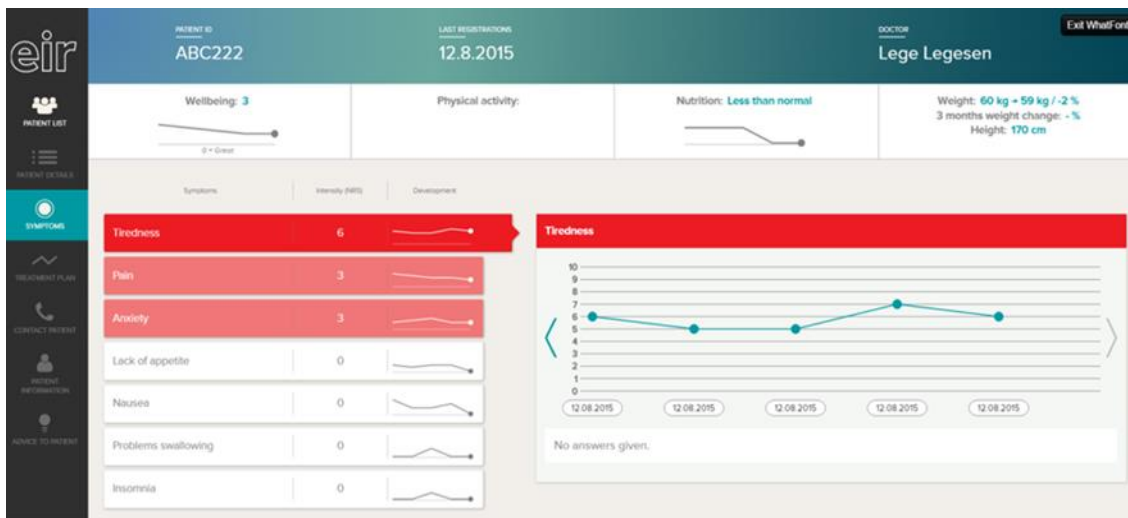| Usability issues	Observation	Resolution/suggestion[1]
Workflow		
In case of incorrect patient login: All symptom scores are set to 0, and the last registration is not visible	Patients made multiple attempts to login. The last registration is set to 0, and the time curve is incorrect	Suggestion: Possibility to click on the last registration to get the correct symptom presentation, or that an incomplete completion is not registered
Previous pain body map not available	Physicians wanted to have access to the patients´ previous pain registration	Suggestion: Permit access to all information in previous registrations by clicking on the date of interest
Understandability		
Wellbeing NRS 0-10	Observed that physician misunderstood the scale (same as for patients)	Resolution: 0 is anchored by text (0= great)
Bugs		
Numbers on pain body map disappear	Pain score not visible in pain body map	Resolution: Made visible
Content		
Relevance of PHQ-9	Physicians questioned the relevance of several questions on depression	Suggestion: replace PHQ-9 with PHQ-2 (only two follow up questions)
Relevance of follow-up questions	Physicians rarely use follow-up questions for other symptoms than pain	Suggestion: make the follow-up questions visible on the front page, e.g. by marking the symptoms where additional information is available

[1]Resolution means that there has been a change in the program to solve the usability issue. Suggestion means that the issue is not resolved.

Figure 1: Eir-Patient*

*Symptom intensity score on a 0-10 scale

Figure 2: Eir-Doctor opening screen*

*The patient`s present symptom intensity to the left, and a graphical presentation of symptom intensity over time to the right